(Practical Case Studies on Using CAAT's)

(Revised and updated Edition)

(Software IDEA CD with 1000 records restriction enclosed with this book for practice)



The Institute of Chartered Accountants of India (Set up by an Act of Parliament) New Delhi © The Institute of Chartered Accountants of India

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form, or by any means, electronic, mechanical photocopying, recording, or otherwise, without prior permission, in writing, from the publisher.

Revised Edition	:	January, 2016
Committee/Department	:	Committee on Information Technology
Email	:	cit@icai.in
Website	:	www.icai.org/http://cit.icai.org
Price	:	₹ 400/- (including CD containing Demo copy of CAAT Software)
ISBN	:	978-81-8441-172-0
Published by	:	The Publication Department on behalf of Committee on Information Technology, The Institute of Chartered Accountants of India, Post Box No. 7100, Indraprastha Marg, New Delhi-110 002.
Printed by	:	Sahitya Bhawan Publications, Hospital Road, Agra 282 003. January/2016/P1856(Revised)

FOREWORD

It is a matter of immense pleasure that the Committee on Information Technology of the Institute has come out with this comprehensively revised publication on "*Data Analysis for Auditors – Practical Case Studies on Using CAAT's*". The publication aims to provide practical insight into use of auditing tools and techniques by Internal Auditors, Statutory Auditors, Business Control Analysts, Fraud Investigators for their data analysis requirements using Computer Assisted Audit Techniques (CAATs)/ General Audit Software (GAS).

Earnest endeavor has been made to share practical experience in using CAAT tools for professional assignments and research in wide range of industries and experiences gained from interactions with fellow professionals. This compilation of practical case studies is expected to provide the requisite exposure to enable Chartered Accountants to see how CAAT Tools can be used to analyze the operations of an enterprise and recommend steps to ensure better control and efficient functioning of the business, apart from detecting fraud/mistakes.

I am sure that this compilation would be immensely useful to the Chartered Accountants both in practice and Industry. They would be able to further enhance the effectiveness and efficiency of their services to have better insight through better data analysis and reporting using CAAT/ GAS Tools. Using the book the members will be able to easily accomplish global standards in terms of skills, knowledge and competencies as is now being demanded/ expected from the profession.

I compliment CA. Rajkumar S. Adukia, Chairman and other members of the Committee on Information Technology for

bringing out this useful publication. I am sure that this initiative would empower the members to enhance their professional offerings.

CA. Manoj Fadnis President, ICAI

Place: New Delhi Date: January 13, 2016

PREFACE

The Committee on Information Technology has great pleasure in releasing the revised publication on "Data Analysis for Auditors- Practical Case Studies on Using CAAT's as a capacity building initiative for members to enhance the bouquet of services that they can offer, apart from enhancing the efficiency and effectiveness of their operations.

Using of Computer Assisted Audit Techniques (CAAT) is gaining popularity within audit departments in enterprises and servicing of clients by Chartered Accountants as they enable us to conveniently undertake better analysis of data and report thereof in a structured manner, apart from keeping an audit trail of the analysis and intermediate results.

The CAATs/ GAS enable Professional Accountants to better analyze the data, detect fraud/ mistakes, and recommend steps to ensure better control and efficient functioning of the business, apart from offering compliance services. This initiative enables professional to address the practical audit requirements and statutory compliances for businesses using CAAT Tools.

We are grateful to CA. Deepjee Singhal and CA. Manish Pipalia for contributing the basic content of this publication.

I extend my heartiest congratulations to the IT Committee and wish the publication a grand success.

CA. Rajkumar S. Adukia

Chairman Committee on Information Technology

Place: New Delhi Date: January13, 2016

CONTENTS

S. No	Contents	Page No.
1	Introduction	1-3
2	Pre-requisites for the Application of CAATs	4-10
2.0	Tool Selection	4
2.1	Tool Application	8
2.2	Good Processes for Implementation of CAATs	10
3	Regulatory Drivers for the Use of CAATs Worldwide	11-13
3.0	Introduction	11
3.1	The Institute of Chartered Accountants Of India	11
3.2	The American Institute of Certified Public Accountants (AICPA)	12
3.3	International Audit Practice Committee (IAPC) of IFAC	12
3.4	Information Systems Audit and Control Association	12
3.5	Hong Kong Institute of Certified Public Accountants	13
4	Stages in use of CAATs	14-22
4.0	Identifying the audit objectives and areas where the CAATs can be used	14
4.1	Determining the data required and arranging for the download	15
4.2	Import of Data	18
4.3	Backward Reconciliation of Data to the Host System	20
4.4	Use of Functions within the CAAT	21
4.5	Review of Results and Housekeeping	21

S. No	Contents	Page No.
5	Examining Data and Isolating Specific Items	23-38
5.0	Introduction	23
5.1	Indexing	24
5.2	Sorting	28
5.3	Group Records	29
5.4	Go To	31
5.5	Find	31
5.6	Extraction of Data	32
5.7	Display of Data	37
6	Identifying Exceptions and Irregularities	39-45
6.0	Introduction	39
6.1	Duplicate Detection Test	39
6.2	Gap Detection	42
7	Profiling Data	46-58
7.0	Introduction	46
7.1	Summarization	47
7.2	Stratification	50
7.3	Pivot Table	53
8	Working with Multiple Files	59-67
8.0	Introduction	59
8.1	Visual Connector	59
8.2	Join Databases	62
8.3	Append Databases	64
8.4	Database Compare	65
9	Sampling	68-80
9.0	Introduction	68
9.1	Systematic Record Sampling	68
9.2	Random Record Sampling	70
9.3	Stratified Random Record Sampling	71

S. No	Contents	Page No.
9.4	Attribute Sampling	73
9.5	Monetary Unit Sampling	76
9.6	Classical Variables Sampling	79
10	Field Manipulation	81-88
10.0	Introduction	81
10.1	Editable Fields	81
10.2	Virtual Fields	84
11	Importing Various Data File Formats	89-12 1
11.0	Introduction and Stages in the Import of Data into IDEA	89
11.1	Data File Types Compatible with IDEA Import	90
11.2	Data File Types and Common Application Systems	9′
11.3	Importing a Microsoft Excel File	92
11.4	Importing ASCII, EBCDIC or Delimited ASCII Files (Comma Separated Value) ASCII Files	96
11.5	Importing a dBASE File	107
11.6	Importing a Print Report File	110
12	Working with Application Softwares	122-129
12.1	Introduction	122
12.2	Core Banking System – Finacle from the Banking Industry	122
12.3	Web Based Application Systems – Sabre from the Airlines Industry	128
12.4	SAP R/3 and SAP Reports from the Manufacturing Industry	126
13	Application of CAATs to General Insurance	130-256
13.1	Introduction	130
13.2	Application of CAATs to Banks	236
13.3	Application of CAATs to Retail Sector	243
13.4	Application of CAATs to Utility Services	253

S. No	Contents	Page No.
14	Accounts Payable and IDEA	257-284
14.1	IDEA and the Accounts Payables Process	257
14.2	Accounts Receivable and IDEA	264
14.3	Inventory and IDEA	271
14.4	Payroll and IDEA	277
14.5	Computer IT Security and IDEA	281
15	Tally Accounting System and IDEA	285-298
15.0	Introduction	285
15.1	Import of a Tally 9 ERP Data File	285
15.2	Analysis of Tally 9 ERP Data in IDEA	291
16	Fraud Investigation using IDEA	299-313
16.0	Introduction	299
16.1	Benford's Law of Digital Analysis	301
16.2	Other Case Studies on Fraud Investigations using IDEA	312
17	Documentation of CAATs	314-320
17.0	Introduction	314
17.1	Regulatory Admissions	314
17.2	Planning a CAAT	315
17.3	Objectives and Assurances	315
17.4	What Data is available?	316
17.5	The Flow of Data in the System	316
17.6	Identifying the file for interrogation	317
17.7	Specifying the information required for the CAAT	317
17.8	File Download and Conversion	317
17.9	Who or What can help in Data Identification?	318
17.10	Specimen Format of a Data Indent	318
18	Continuous Auditing Using IDEA	321-330
18.0	Introduction	321

S. No	Contents	Page No.
18.1	Benefits of Continuous Auditing	321
18.2	Implementing Continuous Auditing	322
18.3	Sample Case Study on Continuous Auditing using IDEA	323
19	Advanced Statistical Methods in IDEA	331-339
19.0	Introduction	331
19.1	Correlation	331
19.2	Trend Analysis	336
20	Do's and Dont's for IDEA Software	340-345
21	New and Emerging Features in Data Analysis	346-349
21.1	Visualization	347
21.2	Discover	348
21.3	Advanced Fuzzy Duplicates	348
	Appendices	351-405

Chapter 1 Introduction

Chartered Accountancy (CA) firms operate within an environment of high expectations from both its clients, the external public and regulators. Regardless of size, CA firms whether Statutory Auditors or Internal Auditors are always looking for new and innovative ways to deliver value to their existing customer base while attracting new clientele. Despite changing regulatory environments and fluctuating economies, one constant to a more profitable practise is lowering costs and time. CA firms require technologies that can add as much value as possible to its clients, in the most cost-effective method available. With audit enabling technologies CA firms can harness the power of data analytics to ensure a structured framework for performing organized, efficient and reliable audits that meet or exceed professional standards.

Audit, financial control and compliance professionals are faced with the challenge of reliably achieving objectives while addressing uncertainty and acting with integrity. Whether the objective is to provide business assurance, improve controls, or ensure compliance, it will be difficult to draw any credible conclusions to help the business without analyzing the underlying data in detail.

However, data analysis specifically for audit, financial control and compliance is easier said than done. The key goal behind the analysis is to identify data discrepancies that need more investigation – this requires more sophisticated capabilities than just summarization and aggregation. Add to this the importance of easily sharing, presenting and distributing analysis results to stakeholders and the search for an effective analysis solution becomes even more challenging.

This book grew out of our desire to provide to internal auditors and other professionals like statutory auditors, business control analysts, fraud investigators, students and educators of this subject the right kind of help and guidance to understand and learn the application of CAATs. We have been using CAATs in our professional practice for the last fifteen years. This book is based on the material researched and experiences gained by us by. We have also incorporated new insights gained by us by talking to fellow professionals on the subject.

Throughout this book our use of the word 'internal auditor' should be read to include other professionals like external auditors, fraud investigators, students, educators, and professionals who would like to work with or are already working with CAATs.

This book uses narrative descriptions, illustrations and case studies to explain CAATs. It includes information on pre-requisites for the application of CAATs, introduction to a data analysis tool, stages in the use of the tool, techniques to examine data and isolate specific items, identifying exceptions and irregularities, profiling data, and various case studies like travel claims review and others, use of the tool in specific industries like Insurance, banking, and in forensic review and interrogation (fraud investigation).

Audit tools may be classified thus:

- Electronic Working Papers
- Information Extraction and Analysis
- Fraud Detection
- Network Security
- Electronic Commerce and Internet Security
- Continuous Monitoring
- Audit Reporting
- Other specialized software tools like Risk Management etc.

Among all the available tools, the ones for data extraction and analysis including fraud detection and continuous monitoring are the most widely used. These have also higher payback in terms of analyzing vast volumes of data for identifying deviations. Because of this the book focuses on only Data Extraction and Analysis tool.

Though there are numerous tools available for extracting and analyzing data, we have used "IDEA" CAAT tool for case studies. The audit objectives and the analysis process covered in these illustrations and case studies are generic in nature. These can therefore be used as a learning base for its application in a specific work environment. We would like to thank Caseware Idea Inc., Canada for giving us permission to use their deeply researched material, including the training material on IDEA software tool in this book, especially screenshots which help us to demonstrate the use of the tool in various live situations. As audit tools grow more efficient and sophisticated, they also become easier to learn and use. At the same time, they must also fit into a complex and ever changing environment. The features of audit software can easily conflict with the features of other software on the computer or network, and must be carefully managed.

Pre-requisites for the Application of CAATs

2.0 Tool Selection

The Global Technology Audit Guide (GTAG) 16: Data Analysis Technologies released by the Institute of Internal Auditors (IIA) lays down detailed information which will help potential users pick the right Data Analytical Tool.

Attributes of Data Analytical Solutions for Audit

Data analysis technology for Audit use needs to have the features and functionality that auditors require to do their job effectively. Not only should it deal with the data access challenges, but it also needs to support the way in which auditors work and the types of analytics that are appropriate to the audit task on hand

Mentioned below are some of the key attributes of Data Analytic Tools -

- Able to analyse entire data populations covering the scope of the audit engagement.
- Makes data imports easy to accomplish and preserves data integrity.
- Allows for accessing, joining, relating and comparing data from multiple sources.
- Provides commands and functions that support the scope and type of analysis needed in audit procedures.
- Generates an audit trail of analysis conducted that is maintained to facilitate peer review and the context of the audit findings.
- Supports centralized access, processing and management of data analysis.
- Requires minimum IT support for data access or analysis to ensure auditor independence.
- Provides the ability to automate audit tasks to increase audit efficiency, repeatability and support for continuous auditing.

Ranking Card to choose the right Data Analytical Tool for Audit

Auditors may employ the comprehensive ranking card (below) while evaluating suitable Data Analytical Tools for his/her engagements:

Need: 4=Mano	0=Needless: 1=Nice to Have: 2=Desirable: datory	Need			
Interna	Internal Audit Strategic Objectives				
1	Software is easy to learn and use				
2	Competitive advantage				
3	Minimize reliance on IT professionals				
4	Improve work accountability, responsibility and supervision				
5	Enforces production program change controls				
6	Reliability: bug free, speed, work like a professional				
7	Portability: runs on a laptop				
8	Scalable: grow from desktop to server without learning new software				
9	Data integrity and security: client data is protected from auditor change				
10	Collaborative features				
11	Supports development of automated and continuous programs				
12	Compatible with electronic work papers				
13	Improves documentation of audit work completed.				
Provider and Implementer Support					
14	Global presence				
15	Years in business				
16	Multiple languages				
17	Help desk available				
18	Ease of doing business: knowledgeable in auditing needs				

Need: 4=Mane	0=Needless: 1=Nice to Have: 2=Desirable: datory	Need
19	Regular software upgrades	
20	Training readily available	
21	User group program for networking with other users of the Tool	
22	Knowledgeable consultants independent of the provider available	
23	Getting started programs available	
Technic	cal Features and Functionality	
24	Import all file types used by the organization	
25	Handles large file record sizes	
26	Handles large data volumes	
27	Ease in validating and reconciling data import	
28	Modify imported data field properties	
29	Support search for text, numbers and time.	
30	Project visual chart or mapping of data actions performed	
31	File join/merge/compare	
32	File append	
33	Visual Connector	
34	Sorts, indexing, filtering and fuzzy logic	
35	Summarization	
36	Extraction	
37	Pivot Table	
38	Stratification	
39	Gap Detection	
40	Aging	
41	Compare data to predicted data – Benford's Law	
42	Advanced statistical analysis: correlation, trend analysis, time series	

Need: 4=Man	0=Needless: 1=Nice to Have: 2=Desirable: datory	Need
43	Sampling	
44	Statistical analysis	
45	Export to typical office applications	
46	Create custom reports and graphics	
47	Create simple and complex calculated fields	
48	Data cleansing solutions – character @functions available	
Cost		
49	Software purchase	
50	Implementation costs – scripting and components	
51	Upgrade fees	
52	Annual help desk support	

Pre-requisites for the Application of CAATs

The IIA's analysis made several key recommendations for internal auditors to consider before using of CAATs:

- Determine the organization's audit mission, objectives and priorities: Auditors must consult with the management of organizations regarding their audit functions of the highest priority and where exactly CAATs could be used to meet those priorities.
- Determine the types and scope of audits: To determine the specific objectives of each audit and its terms of reference. Information on this is crucial for making an appropriate software decision.
- Consider the organization's technology environment: To make sure that selected CATTs will have to operate with the software, hardware and network systems of an organization.
- Be aware of the risks Applying software to any mission-critical function carries some risks, and CAATs is no exception to this. As the IIA article suggests, CAATs can prompt auditors to jump to premature, faulty and hasty assumptions and conclusions Data should not be taken for granted and every finding revealed by the CAAT should be thoroughly validated for its accuracy and relevance.

In short good CAATs should possess:

- Simple and intuitive interface
- Menu driven functions rather than command driven functions
- Ease of data extraction
- The ability to access a wide variety of data files from different platforms
- Ease of data analysis
- Simplified query building
- Comprehensive logging features with 'read' rights only unalterable
- Easy template creation
- Convenient template recall and use
- Ease in automation of tasks
- Simplicity in scheduling of automated tasks for continuous monitoring"

Appendix I gives an illustrative list of CAATs.

2.1 Tool Application

2.1.1 Management support for the optimum use of CAATs

- The CAATs should be made a vital part of the Internal Audit Charter and should have ready support of the Audit Committee as a key audit resource.
- The management of an organization should perceive the real value of CAATs and should encourage the Internal Auditor to apply it optimally, and the Information Systems (IS) department should to support the data flow/ availability for the use of a CAATs.

At one location, where the experiences with CAATs were highly successful, the senior management of the organization encouraged the Internal Auditor to make optimum use of the CAAT and MS-Excel in day-to-day data analysis. Moreover, based on the successful experiences of the CAAT within Internal Audit, the management decided to roll-out the CAAT to process divisions like Sales and Purchase. Eventually the CAAT was made an important Management Tool.

2.1.2 Data availability for the optimum use of CAATs

- Access to open data is a must for the use of CAATs. Encrypted data, like the data in 'Tally' Accounting Software, cannot be accessed directly. (Tally is a widely used accounting software in the South Asian region). The 'Tally Vault' prohibits direct link to the data tables in 'Tally'.
- Information Systems (IS) Administrator should appreciate the functionality and application of the tool. Considering the confidentiality and integrity of the data desired, an elaborate plan of data access, data storage and data use should be identified. Support is necessary in the form of organizing data availability for specific audit objectives in formats desired by the Auditor. A working procedure should be established in which the roles of IS officials and Auditors are clearly defined for ensuring the easy availability of data.

2.1.3 Internal Auditor skill-sets for the optimum use of CAATs

- The user of CAATs should be innovative and imaginative and capable of thinking out of the box.
- The user should be computer literate and possess a working knowledge of Windows and any Office Suite.
- He/She should have adequate audit knowledge and experience to define
 - ✓ Audit Objective/s
 - ✓ Data required for achieving audit objective
 - ✓ Steps to be employed within the CAAT to achieve the intended results in minimum time.
 - ✓ Likely benchmarks for validating results adequate industry domain knowledge

Our work has led us to the conclusion that a CAATs Auditor does not need to be a programmer with years of intensive programming experience.

2.2 Good processes for the implementation of CAATs

- Selecting a user champion/s for the CAATs, preferably within the existing audit group who has spent relatively long years in the function/organization/industry.
- Organising a structured training programme covering product functionality with case studies and including organization specific cases in the training programme. Involving all users in the training programme at all levels. If the audit group is large, the training programme could be a "Train the Trainer" type.
- Identifying a plan for introducing the tool by linking it to the existing audit programmes. Follow the KISS--'Keep It Short and Simple--'concept in its introduction and follow it up by using advanced functionality. The period of evolution would depend on the competence of the team members, and could range from six months to two years.
- Reviewing utilization every six months to ensure its continuing use.
- Availing continuous support of the vendor. This way the tool is updated at all times and is free from known/unknown bugs.

We have noticed in countless instances that a nominated CAATs Auditor is overburdened with day-to-day audit responsibilities, which prevent him from devoting his physical and mental energies to effective implementation of the CAATs. A nominated CAAT Auditor should be given free space and time to explore the CAATs fully. At the outset it may seem a bit costly for the organization, but it will start paying back within three to six months.

Chapter 3

Regulatory Drivers for the use of CAATs Worldwide

Pronouncements by Academic and Audit Regulators on the use of $\ensuremath{\mathsf{CAATs}}$

3.0 Introduction

Numerous regulatory bodies from all parts of the world have backed the use of CAATs: from Accounting and Audit Regulators and Watchdogs.

Some of their pronouncements are reproduced below, -

3.1 The Institute of Chartered Accountants of India (ICAI)

Recognising developments in the field of technology and their impact on the accounting profession in India, Auditing and Assurance Standards Board issued Auditing and Assurance Standard (AAS) 29, Audit in Computer Information Systems Environment in September 2003. This Guidance Note on Computer Assisted Audit Techniques is a sequel to that AAS. The Guidance Note deals extensively with the concept of CAATs, with focus on the areas where it may be used, considerations for this use, controlling the application of CAATs, documentation required for its use, and its use in small entities. The Guidance Note also contains a comprehensive appendix containing examples of CAATs, their description and comparable advantages and disadvantages of each of the CAATs.

From the Handbook of Auditing Pronouncements, Vol. II (Edn. 2005)'

An excerpt from the Guidance Note on Computer Assisted Audit Tools (CAATs)

"The application of auditing procedures may, however, require the auditor to consider techniques known as Computer Assisted Audit Techniques (CAATs) that use the computer as an audit tool for enhancing the effectiveness and efficiency of audit procedures."

3.2 The American Institute of Certified Public Accountants (AICPA)

An Excerpt from the Statement on Auditing Standards No. 94, The Effect of Information Technology on the Auditor's Consideration of Internal Control in a Financial Statement Audit, AICPA, New York, USA, May 2001. (Amends Statement on Auditing Standards No. 55, Consideration of Internal Control in A Financial Statement Audit, AICPA, New York, USA, April 1988.)

"The auditing standard states that computer-assisted auditing techniques (CAATs) are needed to test automated controls in certain types of IT environments. This paper revisits auditing-through-the-computer techniques, which should become more widely used with the issuance of SAS No. 94, and focuses on the test data technique, which can be applied in almost any audit to test automated programmed controls."

3.3 International Audit Practice Committee (IAPC) of IFAC

In March 2000, the International Audit Practice Committee (IAPC) of IFAC released an (exposure draft?) on four topics, which form a supplement to ISA (International Standard on Auditing) 401 "Auditing in a Computer Information Systems Environment (CIS)."

"According to the exposure draft, the purpose of the statement on CAATs "*is to provide guidance in the use of Computer Assisted Audit Techniques (CAATs), which are techniques that use the computer as an audit tool.*" The exposure draft "*applies to all uses of CAATs involving a computer of any type or size.*"

3.4 Information Systems Audit and Control Association (ISACA)

Excerpts from the Document G3

- 1.2 <u>Need for Guideline</u>
 - 1.2.1 Computer Assisted Audit Techniques (CAATs) are important tools for the IS auditor for performing audit.
 - 1.2.2 CAATs include many types of tools and techniques, such as generalised audit software, utility software, test data, application software tracing and mapping, and audit expert systems.

3.5 Hong Kong Institute of Certified Public Accountants

Practice Note 1009 provides guidelines on the use of CAATs.

Excerpts from the Practice Note:

"The overall objectives and scope of an audit do not change when an audit is conducted in a computer information technology (IT) environment. The application of auditing procedures may, however, require the auditors to consider techniques known as Computer-Assisted Audit Techniques (CAATs) that use the computer as an audit tool."

"CAATs may improve the effectiveness and efficiency of auditing procedures. They may also provide effective tests of control and substantive procedures where there are no input documents or a visible audit trail, or where population and sample sizes are very large."

The full Practice Note 1009 is reproduced in Appendix II.

Chapter 4

Stages in the use of CAATs

CAATs are used in the following stages:

- 1. Identifying the audit objectives and areas for the use of CAATs
- 2. Determining the data required and arranging the download
- 3. Importing data
- 4. Reconciling data
- 5. Working with the data imported: Use of Functions and Scripts within the CAATs
- 6. Reviewing of results and housekeeping

4.0 Identifying the audit objectives and areas where the CAATs can be used

A list of audit objectives to be achieved should be identified by the internal auditor at the outset. The audit objectives may include financial and operational dimensions for various business processes. They may also include special investigations and/or reviews into alleged fraudulent activities or business analysis at the request of the top management.

The objectives should be clearly documented to cover all the desired aspects of audit. The Internal Auditor needs to make a check on the feasibility and cost/benefit aspect of the application of CAATs for each of the objectives. This is a one-time exercise which will yield good results in the longrun.

By feasibility is meant the easy availability of data in an electronic format for each objective. For example, a *review of payments to contract labour at a manufacturing site cannot be checked through CAATs for validity and accuracy of payments if the attendance records are logged manually.*

By cost/benefit analysis is meant the cost and effort to be invested in obtaining the data vis-à-vis the benefits in money terms/in kind which can be reaped. For example, *instances where a borrower has overdrawn the lower of his sanctioned limit or drawing power in a bank are normally available in the 'End of Day Reports' of the Core Banking System. It will be uneconomical to verify this objective by using CAATs.*

4.1 Determining the data required and arranging for the download

The audit objectives should be mapped against the data that should be available in very clear and accurate terms. The internal auditor should seek the assistance of the data owner and IS department in this regard. This can be done as follows:

- 4.1.1 Request IS for Data
- 4.2.2 Request the Data Owner for Data
- 4.2.3 Access to the Database Server

4.1.1 Request IS for Data

The Internal Auditor should send an email to the IS Manager. It should contain

- A clear and full description of the audit objective
- The period for which the data is required
- The data fields essential to the test.
- The format in which the data is required
- The medium through which the data is to be conveyed to the internal auditor: email, CD, network LAN access, etc.

This approach is explained with the help of an example from the banking industry

Electronic 'Know Your Customer (KYC)' Audit

The team selected an objective to test KYC norms on current account customer master data.

So the team issued a Data Request to IS in the following format:

- Data required: Current account customer master information.
- Period: (As of the date of Audit: reword this because date and period are different things).
- Fields of reference: Branch ID, Customer ID, Account ID, First Holder & Joint Holder/s Name, Address, PAN No., Mobile No., Residence No., Office No., Mode of Operation and Clear Balance.

• Format of Data: Text form.

In response to this, the IS ran a SQL query on the production database and generated a text file dump which was saved in a secure folder with special access to the audit team only.

4.1.2 Request the Data Owner for Data

The internal auditor should send an email to the data owner, which should contain

- A clear and full description of the audit objective
- The period for which the data is required
- The data fields essential to the test
- The format in which the data is required
- The medium through which the data is to be conveyed to the Internal Auditor: email, CD, network LAN access, etc.

Here the data owner/process owner/department head provides data to the internal auditor either through the standard report suite of the application system or through some specialized Business Warehouse solution like Crystal Reports, BIW, COGNOS, and SAP-AIS.

This approach is explained with the help of an example from the banking industry

Audit of Loans and Advances MIS from Business Warehouse (BIW) for approval verification

The team decided to audit the MIS report on Loans and Advances generated from the Business Warehouse (BW) for approval verification. The MIS report is a comprehensive listing of Account ID, Type of Loan, Type of Security, Type of Industry, Sanction Limit, Drawing Power, Rate of Interest, Due Date of Loan, and Approval Officer.

To complete this test, the team sought assistance from the Merchant Banking Wing of the Bank. The Loans officer for Branch "X" generated the MIS report for the period--1st April 2014 to 31st March 2015. This report was saved as a Microsoft Excel file and provided to the team on a CD.

The team imported the Microsoft Excel file using the MS-Excel option for import within the Generalised Audit Software (GAS).

4.1.3 Access to the Database Server directly

There is a common belief in professional Information Technology (IT) circles that while using CAATs, connectivity to data servers should not be allowed. Normally, as the data server contains data in a raw tabular format by using CAATs on the server directly, it could compromise the integrity of the data and slow down performance. While performance may be a factor that can be considered in some circumstances, integrity is not a concern. In order to avoid controversy, production data servers need to be avoided. However, there are organizations where the IT officials are willing to provide access to a Flash Server, Disaster Recovery Site Server or a Training Server, which has a historical image of the production data server.

There is also a myth that because servers contain a million data tables, identifying the right data table for each objective is virtually impossible. This needs to be cleared at the outset. Take the SAP server, for instance. The SAP server has not more than a few thousand tables. Identification of the right table is difficult but not impossible. .

Access to any server requires the following to be in place:

- The workstation on which the CAATs is loaded, (and which attempts to access a specific server should contain the client end data server drivers this part of the sentence is not clear) For instance, Microsoft ODBC for Oracle, Oracle for Ora 92 drivers for Oracle, Microsoft SQL drivers for MS-SQL, etc.
- The CAATs should contain the ODBC facility. ODBC is a middleware which allows third party querying/reporting tools to connect to database servers and access server based table information. ODBC is a secure acceptable data conduit.
- Server reference and schema (logical modules (?) like MM, FI, SD in SAP) reference.
- User access 'READ' rights to the server.

Once the connectivity is in place, good CAATs allow the internal auditor to firstly write SQL queries on the raw data table at source itself to limit the data to be imported. Popular CAATs also allow the user to include or exclude fields of their choice at the data import stage itself.

This approach is explained with the help of an example from the banking industry.

With the help of the IS wing, the team set up connectivity between the CAATs and the FINACLE ORACLE DR site server. (FINACLE is a Core Banking Solution used in many banks in India and abroad).

The test was undertaken to identify Term Deposit accounts, where the deposit was accepted for more than 120 months.

As per the bank-rules for acceptance of Term Deposits and guidelines of the regulator, the bank cannot solicit Term Deposits in excess of 120 months. This test was undertaken to identify non-compliances with the rule.

The team connected the DR site server using the CAATs and identified the data table containing the Term Deposit account masters TAM. At the import stage, the team selected three fields – Account ID, Account Name and Deposit Period in Months from the table containing 15 different fields. The team also entered an SQL query in the GAS, filtering the data for Branch "X". The import of data of 0.1 million rows/lines was initiated and duly completed within 1 minute. This was faster in comparison to writing an SQL query which would take 3-4 hours to run in the past.

4.2 Import of data

In order to interrogate and analyse the data by using CAATs, the data must first be 'imported'. Good CAATs contain a rich, robust, easy-to-use, screen based import facility. Many a times the file to be imported is very complex (like a multi-line text report file). In such cases, CAATs should offer the facility of saving the import process in the form of a reusable template. The template can be used as long as the report format (fields and field information) does not change.

The import facility should read the following file types (extensions provided):

- Spreadsheet files MS-Excel (xls), MS-Access (mdb), Dbase (dbf)
- Lotus Files (.wk1, .wk3, .wk4)
- Text files Data Files (dat), Print Report Files (prn)
- Portable Document Format (PDF) Files. PDF files are gaining prominence on account of the security which comes with such files.
- Delimited Files Comma Separated Value Files (csv)
- Flat Files ASCII Files (asc), EBCDIC Files
- Extendible Markup Language (XML) Files commonly found in Internet

Based Web Applications like Online Air Ticket Reservation Systems – SABRE.

• Audit Information System (AIS) from SAP.

4.2.1 Easy Practical Tips on Technical Issues for Data Import

- Request File Layouts from the IS Department, which are normally called record definitions, data dictionaries, schemas or a host of allied names.
- Take IS help in identifying salient fields in file layouts.
- Review the data with some knowledgeable users on screen or through hard-copy review to identify important fields for selection and inclusion. For example, date of birth field may take a whole new meaning from a validation on child fares or child facilities. Otherwise this information may be normally excluded.
- Pay attention to type, ledger and product coding systems.
- If a report that has all the required data consistent with the audit objective can be located, then go for that file.
- Many IS support staff will automatically format a report to an 80 or 132 column width to fit standard paper sizes. As the report is to be printed to a file and not to paper it can be as wide as it needs to be. IDEA software can approximately accommodate 256 odd columns.
- Before transferring a large file over a network, carry out a test on a small file to determine transfer time and find out if resource will be inadequate for other system users. If yes, maybe the large file data import can be scheduled outside working hours or on holidays.
- Many files include "deleted" or old records that may not be included in control totals. For example, payroll in which 'left employees' (terminations) are often kept on the system for tax or social security reasons. Identify such items and exclude them.
- To import COBOL files, take assistance from the COBOL file definition Section. COBOL files can be imported by creating a record definition for the file and then using the very same Record Definition Editor (RDE) to import the file. Alternatively, it can be imported through the ASCII option.

- Printouts are generally formatted with page headers and footers as well as column headers. These should be excluded when specifying the data to import. However, reports often contain "header" information such as an account number, region or area code, etc., which needs to be included for data completeness. Mark off such headers and identify a consistent pattern.
- Data residing in columns can be identified as separate fields based on delimiters within the fields.
- Totals and sub-totals should be excluded from the selection process. Reconciliation is very critical to ensure that 'include' and 'exclude' features have been correctly dealt with while importing.
- In some cases the amount of data is too lengthy to be printed in a single line. In such cases, reports are designed to wrap onto two or more lines. Other reports lines may wrap unintentionally due to the printer set up. Such reports must be 'unstacked' to produce single-line records prior to specifying the columns of data to import.

While generating a report in print file form from an application, it is prudent for the auditor to be aware of report generation facilities and not depend fully on IS. This serves as a foolproof cross check on data integrity of files imported.

4.3 Backward Reconciliation of Data to the Host System

Once the data has been defined and read into the CAATs, it is imperative that before commencing audit testing, it is reconciled back to the "host" system. Given below is the list of errors that could occur during the process:

- Requesting wrong data;
- Being supplied with wrong data (and/or for the wrong period);
- Errors on the extraction of the data from the host system;
- Errors when importing the data like invalid or bad data in certain fields due to the inconsistent nature of transaction entry in that field.

The data can be reconciled back to the host system by reviewing:

- Numeric field totals;
- The record count;
- A sample of records compared to a print from the host system.

It is important that reconciliation reports and control totals are requested along with the data and filed with the documentation from the download/import process.

Data can be reconciled in different ways. CAATs contain Field Statistics and Control Total options to reconcile totals and to ensure that the dates match the data requested and do not include transactions from another period. This is the essence of cut-off checks.

4.4 Use of Functions within the CAATs

On the data analytical front CAATs can do the following -

- Isolate Specific Items
- Identify Irregularities
- Profile
- Categorize
- Field Manipulation
- Sampling
- File Look Ups
- Merging of Data Files
- Repetitive Analytics

4.5 Review of Results and Housekeeping

4.5.1 Review of results

Findings arising out of an analytical task are potential cases for reporting. They turn into actionable findings post validation only. The internal auditor should use CAATs as a means to an end and not as an end in itself. CAATs is simply a tool to examine and review data. The results need to be discussed with the process owner in the light of the objective being tested, data provided and circumstantial facts in the process. This also needs to be tested vis a vis the business trend, company performance trend, practical possibility of such results in the area of audit and other non- financial and operational areas.

CAATs can display instances where the net accounts receivable exceed the credit limit of a customer. This need not necessarily indicate deliberate

attempts by 'Sales' to bypass the safe credit exposure limits of a customer. The need to 'force-sales' may be justified by the need to meet year-end sales targets. This can be supported by due pre-facto approvals and post balance sheet date verification of collections.

4.5.2 Housekeeping

Data extraction and analysis should be undertaken in distinct file-folders within the CAATs. Each assignment should bear separate reference of a file-folder which is properly identified and stored. The internal auditor should follow appropriate record/ folder backup and retention practices and ensure consistent implementation by his team. The file imported, its template (if applicable), the resultant output files, queries/equations, views, graphs and more should all be stored together for easy recall and use.

For example, in a cyclical Payroll Audit, the pay file, attendance file, and allied files for historical cycles should be regularly retained on a separate drive of the same workstation or on an off-site location, appropriately named and indexed to make way for the current files. The back-up should be easily identified by any member from the internal audit Team whenever required. This is essential for electronic evidence management and audit automation – macro (audit routine) creation, record, playback and scheduling.

Chapter 5

Examining Data and Isolating Specific Items

5.0 Introduction

Data once imported can be examined and reviewed in IDEA through

- Indexing
- Sorting
- Grouping of Records
- Go To
- Find, Find Next and Find Field

Isolating specific items can be accomplished in IDEA through

- Display of Data
- Extraction of Data

Examining data and Isolating Specific items allows the user

- The conveniences of speedy processing of tasks within IDEA, since databases are largely indexed prior to performance of any major function in IDEA.
- The benefit of efficient auditing, as more can be accomplished in a lesser time.
- The advantages of effective auditing, as the user can directly move-in on exceptions, irregularities, and anomalies thereby meeting his/her objectives fully.

Examination of Data and Isolating Specific items briefly includes-

Indexing and Sorting: Displaying the data in a specified format for easy visual interrogation; for instance, visual examination of top 10 non-performing assets for a bank by product category.

Grouping of Records: Allowing the user to view multi-layer information in a tree-like structure; for instance, welfare spends by a state undertaking Zone City Taluka Village wise.

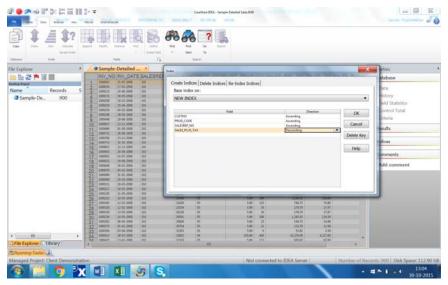
Go To and Find: Providing the interrogator with quick means of zeroing in on a specific record, value, string, or date. Find can be used for to identify all treasury transactions made on Sundays or public holidays for a foreign exchange center.

Display of Data and Extraction of Data: Facilitating identification of all records that meet specific criteria within the database. For instance, all claims admitted in a General Insurance industry wherein the Claim Loss Ratio is over 100%

5.1 Indexing

The Indices task is used to specify the sequential order of the records – high to low, earlier to latest. The records are not physically sorted into a separate database; only an index is created to describe the logical sequence in which the records are to be read, displayed, printed, and processed.

The field(s) you select to determine the order of the database is known as the key. You may select up to eight fields for the index key.



In addition to determining the sequence in which records should be displayed or printed, you can use indices to insert control breaks (or totals) in reports on change of key.

Examining Data and Isolating Specific Items

Con Index	Set Depter	Append	Hudfy										
le Explorer		*	Sample	e-Detailed	* *								operties
1 2 P 0	#			NV_DAT		Report Assistant - Del	ne Broaks	OCAN PROFILE	277 Jan. 8	1,007,340 84	ATR STORE		Database
ne nun ame ² ∎ Sample-De_	Records 900	S 33 4 5 6 7 8 9 9 100 1112 13 1 15 6 7 8 9 9 100 1112 13 1 15 6 7 7 8 9 9 100 1112 13 1 15 6 16 7 7 8 9 9 100 112 12 13 1 15 6 16 7 7 8 9 9 100 112 12 13 1 15 6 16 7 7 8 9 9 100 12 12 12 13 1 15 6 16 7 7 8 9 9 100 12 12 12 13 1 15 6 16 7 7 8 9 9 100 12 12 12 13 1 15 6 16 7 7 8 9 9 100 12 12 12 13 15 6 16 7 7 8 9 9 100 12 12 12 13 15 6 16 7 7 8 9 9 100 12 12 12 13 15 6 16 7 7 8 9 9 100 12 12 13 15 6 16 7 7 8 9 9 100 12 12 13 15 6 16 7 7 8 9 9 100 12 12 13 15 6 16 7 7 8 9 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 6 16 7 7 8 9 100 12 12 13 15 7 8 9 100 12 12 12 13 15 7 8 9 100 12 12 13 15 7 8 9 100 12 12 11 11	3000647 2000154 2000153 200053 200053 200053 200055 2000548 200048 2000648 2000652 2000648 2000652 2000665 2000655 2000665 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000655 2000555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 2005555 20055555 20055555 200555555 20055555555	3: 61:300 3: 61:300	200 300 300 300 300 300 300 300 300 300	Select what fill Base break on No index Costino Prody_cost subsiziery no subsiziery no subsiziery no		Re the report to bree Beaching According According According Enconding		• ete Key	Нир	12 12 12	Data History Field Statistics Control Total Criteria Results Indices No index Comments Add comment
III File Explorer,	Ibrary	2012223 245 226 2278 29 30 31 233 34* *	2000.330 2000.252 2000.253 2000.320 2000.320 2000.324 2000.324 2000.325 2000.255 2000.254 2000.254 2000.254 2000.254	21 49-3908 33-05-2008 13-05-2008 13-05-2008 13-05-2008 13-05-2008 13-05-2008 01-04-2008 01-06-2008 13-05-2008 13-06-2008 13-06-2008 13-05-2008 13-05-2008	202 202 202 202 202 202 202 202 202 202	21650 20154 20154 2016 2008 2008 2008 2008 2008 2008 2008 200	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	530 530 530 530 530 530 530 530 530 530	10 10 10 10 10 10 10 10 10 10 10 10 10 1	740,75 139,39 139,39 148,25 148,25 148,25 139,56 318 42,259,00 439,67	74.38 13.737 12.455 14.455 14.38 12.56 5.59 4.227.65 4.227.65 4.27.65		

The report preview would appear like what is given below -

and a set of the set o	a Data and a second sec	-							
e Explorer *	Sample-Detailed *							- *	Properties
	Print Next Page Prev Page 1	2 Pages Z	com In Zoom	Out Close					Database
Anne Records S Sample-De. 900	Res FINN NO INV. DATE SALESKEP, NO	CUSTNO	PROD_CODE	UNIT, PROX	QTV	SALIS_BIF_TAX	SALIS_TAX	-	Data History Field Statistics
	CUSTNO 1 1000011 20-05-2006 104	10000		35.13	3	175.75	17.38	1	 Control Total Criteria
	CUSINO (3 record)				5	125.25	17.54		· • Results
	CUSTNO 2 0000000 26-06-0000 200	10003	15	5.09	1094	6,07.42	60.74		· • Indices
	CUSTNO (3 recent)				1058	6.337.42	613.74	-	Mo index
	CUSTNO 1 100006 11-05-2008 104 CUSTNO (1 recent)	1004	18	5.8	790	4,151.00	419.30		CUSTNO/A + PROD_CODE/A + SALESREP_NO/A +
					700	4,331.00	413.30		SALES_PLUS_TAX/D
	CUSTNO 4 100000 23-04-200 104 CUSTNO G recent	13005		5.00	1000	5,990.00	389.00		Comments Add comment
	COSTNO (3 record)				1000	5,990.00	599.00		
	CUSTNO 5 1000155 14-05-3008 104 CUSTNO (2 recent)	10006	8	5.00	2050	12,278:50	1,227,95		
					2050	12,279.90	1,227,85		
File Explorer Ubrary			- III -						

Many IDEA tasks will automatically sequence the database(s).

- Duplicate Key Detection
- Duplicate Key Exclusion
- Gap Detection
- Join Databases

- Summarization
- Compare Databases
- Stratification
- Aging

However, it is not necessary to sequence the database(s) before selecting these tasks, as the creation of the index is carried out within the task, if required. If the index or sort order has already been generated, then it is used; otherwise it is created once the test is initiated. The index order generated is then added to the list of available indices for that database.

When selecting an index order, the most significant field is selected first: this is followed by the next most significant field, and so on down to the least significant field. The length of time to index a database is proportional to the length or number of "characters" in the index key.

Once you have created an index, IDEA displays the records in the database in the index order. All generated indices are available in the Indices area of the Properties window. The current index is highlighted with a check mark.

	1			wort* are to and lang time! Expert	Frank San	y				
Explorer	1		e-Detailed	* *						Properties
E 2 7 0 0	1			EJALESRE	P_NC. DUSTN	- ROD_COD	. UNIT_PRICE	QTY SALE	S_BEI +	• • Database
Hop Project	1	33000011	25-05-2908	204	10000	8	15.15	1058	1	✓ Data
ame Records	5 3	1000066	11-03-2008	394	22004	15	5.09	700		
Sample-De_ 900	4	3000000	23-04-2008	314	10005	15	5.00	1096		 History
	5	1000153	14-03-2008	334	20006	15 05	5.09	2090		Field Statistics
	6	3000563	15-05-3908 29-07-3908	394	33907	15	5.09	285 110		Control Total
	1	2000607	03-11-2006	114	20102		5.00	177		
	9	3300042	22-04-2968	304	38952	85	5.99	357		Criteria
	10	2000645	24-10-2008	294	20199	45	539	5		and with the second
	11	1000679 1000608	23-10-2008	204	35390	25	5.00	75 2000		• Results
	12	1000000	29-13-2908	294	38,206	15	325.69	138	-	
	14	2000034	10 07 2008	314	3525	-15	5.09	855		* Indices
	15	33007,20	19-12-2008	234	38200	-85	5.99	136		
	16	3300838	11-07-3908	204	29,301	6	5.89	39		No index
	17	2000.778	34-10-2908 10-10-2908	254 354	35201 36201	25 65	5.00	58		CUSTNO/A + PROD CODE/A + SALESREP NO/A +
	18		28-04-2008	104	25203	8	5.00	40		SALES PLUS TAX/D
		2000625	11-11-2000	204	36203	15	5.00	100		25942 1102 11920
	21	2000705	30-25-2008	203	33204	15	5.09	305		* Comments
	22	3000717	11-08-2008	383	38302	- 15	5.09	158		
	23	3300738	07-11-2008	353	10400	25 85	5.00	117		Add comment
	20 21 22 23 24 25	3000743 3000743	15-12-2008 29-09-3008	303 203	20500 20800	05	5.99	75 67		
	26	3000758	25-05-200	383	19900	85	5.00	\$7		
	27	3000775	13-08-2908	203	111000	00	5.00			
	28	300080M	11-07-2008	103	11307	15	5.09	630		
	29	1000877	19-09-2008	333	11.900 11.908	25 11	5.00	3184		
	30	1000004	28-01-2008	125	11,901	-	5.09	40		
	31	33000113	07-04-2908	125	11600		8.0	30		
	33	23004M	28-07-2908	125	11000	25	5.09	852		
m	> 34	2000.292	08-07-2908	125	11600	05	5.09	342		
File Explorer Ubrary	100	1000474	12.06.7978	175	11600	.15	5.00	(9)		
	1.0									

IDEA also displays a directional arrow on the indexed field column header to allow you to easily identify the indexed field and its order (Ascending or Descending). The primary key index is indicated with a red arrow while secondary keys are indicated with a blue arrow.

Examining Data and Isolating Specific Items

er den	Adbe Its	-	*	100	and tang 2 mail		Ŋ				
e Explorer		14	Sample	e-Detailed						• *	Properties
B: 2 * 3 III						O M. DUSTN	ROD CODE	UNIT PRICE	OTVISALES		
		1.1	Taxan III	20-05-2008	234	20000	100_0001	5.13	a l'indicatera	B	Database
Atta Prant		1	12000110	26-06-2008	214	20003	10	5.00	1058	121	✓ Data
lame Re	cords	5 3	3200066	11-01-2008	234	20004	15	5.09	790		
Sample-De_ 900	4	3700080	23-04-208	334.	23905	15	5.09	0000		 History 	
	5	1000153	14-03-2008	204	30006	05 05	5.00 2010		Field Statistics		
		6	2000043	29-07-2008	324	2000		5.09	110		Control Total
			2005607	03-11-2008	24	20032	-	5.00	177		
		9	2000142	22-04-2968	304	389072	45	5.99	357		Criteria
		10	3000645	24-10-2008	334	10.99	45	539	3		
		11	3000679	21-10-2008	334	34390	25	5.89	15		• Results
		12	3300608	17-15-2908 28-11-2008	204	38308	54	325.49	2500	- 2	
		13	2000034	10-13-2008	204	34255	15	5.09	855		Indices
		15	3300720	19-12-2008	204	38,001	-85	5.99	136		
		16	1000820	11-07-2008	204	36291		5.09	39		No index
		17	2000778	34-10-2908	204	36203	15	5.00	58		CUSTNO/A + PROD CODE/A + SALESREP NO/A +
		18	2000.707	10-10-2008 28-04-2008	334	38,308	8	5,99	45		SALES PLUS TAX/D
		19	1000250 2000625	11-11-200	204	36203	10	5.00	100		54425_PL05_140/0
		20	2000705	10-11-2008	203	25204	11	5.00	325		· Comments
		22	3000717	11-08-2008	393	38707	65	5.00	158		1 S COMPREND
		23	3000738	07-11-2008	383	10400	05	5.00	187		Add comment
		24	3000743	15-12-2008	333	38500	85	5,99	75		
		25	2000743	29-09-2008 25-09-2008	203	20405	01 05	5,00	67 37		
		26	1000/38	13-08-2008	203	11100	10	5.09	1		
		201223345567 22222222222	100080M	13-67-2908	303	11307	15	5.00	400		
		29	3300877	19-09-2008	213	11300	25	5.00	3184		
		30	3300004	14-04-2008	314	1190		15.15	25		
		31	3300633	28-01-2008	125	11400	-	5.09	40		
		32	33000113 2300488	07-04-2908	125	11800 11800	60 15	25.25	50 852		
	,	33	2000292	08-07-2908	10	11800	6	5.09	342		
		32 33 34	1000474	12.06.2008	175	11000	15	5.00	190		
File Explorer Ubra	ary	1			11 / 1 / 1 / 1 / 1 / 1 / 1 / 1 / 1 / 1						
Running Tasks											

The Indices dialog box provides three choices:

- Create Indices
- Delete Indices

You can delete unused indices through the Delete Indices tab of the Index dialog box or through the right-click menu of the Indices area of the Properties window.

Delete unwanted indices if you require additional disk space. However, as many IDEA tasks require an index and indices are often used when printing reports (to provide control breaks and totals), it is advisable to leave the indices until the end of the audit, unless additional disk space is required.

If you create an index with the wrong key or by mistake, delete the index so that you do not use it inadvertently at a later date. If you are connected to IDEA Server and you delete an index in a Shared project, you are deleting the index for all users associated with that project.

• Re-generate Indices

Occasionally, the index on a database may become invalid. This may happen if the index is based on an Editable field or a Virtual field where the equation has been changed. Use the Re-Index Indices task to regenerate the index for any predefined index, thereby eliminating the need to delete and re-create the index. Old indices can also be re-indexed through the right-click menu in the indices area of the Properties window.

Once you have re-indexed a database, IDEA displays the records in the database in the index order and updates the list in the Indices area of the Properties window. The current index is highlighted with a check mark. IDEA also displays a directional arrow on the indexed field column header to allow you to easily identify the indexed fields and their order (Ascending or Descending). A red arrow indicates a primary key field, while a blue arrow indicates a secondary key field.

If you require only a single field for a key, perform a quick index.

When using the quick index method, you can only index the database by a single field.

To index a field in ascending order:

Double-click the field name.

To index a field in descending order:

Double-click the field name a second time.

To return to the default order:

Click the No Index link in the Indices area of the Properties window.

5.2 Sorting

The Sort task is used to create a new database that is physically sorted in a specified order. If you require a sequenced database and do not wish to create a new one, use the Indices task, which creates an index describing the logical sequence in which the records are to be read, displayed, printed, and processed.

Many IDEA tasks require the database to be sequenced. These tests run faster using a sorted database rather than an indexed database.

- Duplicate Key Detection
- Gap Detection
- Join Databases
- Summarization

- Compare Databases
- Stratification
- Aging

However, sorted databases require additional disk space. Once a database is sorted, you need to consider if the original unsorted database on which the sorting was performed should be deleted.

The field(s) used to determine the sort order of the database is known as the key. You may select up to eight fields for the key.

When selecting a sort order, the most significant field is selected first, followed by the next most significant field, and so on down to the least significant field. The length of time to index a database is proportional to the length or number of "characters" in the sort key.

Note

Sorting characters is based on the alphabet as well as the case. Characters are sorted by the alphabet first, and then the lowercase alphabet is followed by its uppercase. For instance, the characters C, c, D, and d are sorted in the following order: c, C, d, D.

With the Sort task, a new database is created with the records in the sequence of the key. This new database is a child database and appears in the File Explorer below the main database. You can view the child database by clicking on the expand button next to the main database.

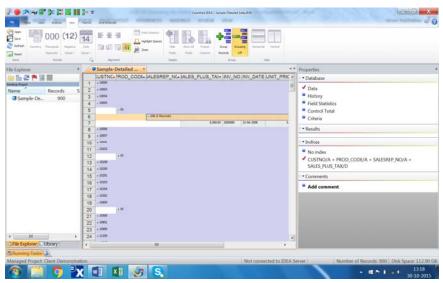
Once you have sorted a database, IDEA displays the records in the database in the sort order and updates the list in the Indices area of the Properties window. When an index is selected, a red directional arrow appears in the column header of the primary key field, while blue directional arrows indicate secondary key fields.

5.3 Group Records

Use the Group Records task to organize matching records in indexed fields (keys) into expandable/collapsible groups. This task will run on local databases and server databases (if you have an IDEA Server).

teret Conner Transfer Conner						Net A		
Explorer		e-Detailed						Properties
E 2 7 2 1	INV_N		ESREP_NC. CUSTN	C. MOD_CO		TY SALES B	EI -	• Database
me Records Sample-De_ 900	S 2 10000000 3 10000000 4 10000000	26-06-2008 204 11-05-2008 204 23-04-2008 204	Group Records	T:	. 5.0	-	×	■ Data ■ History
	5 100041	34-45-2000 304 13-45-2000 304	Group keys:			OK		Field Statistics
	7 35000M B. 3000607 9 3500042	29-47-200 254 10-13-250 254 23-04-200 254	CUSTNO PROD_CODE	und .	Direction Accending Accending	Cancel		 Control Total Criteria
	10 200040 11 2000679	24-15-2908 254 21-16-2908 264	Party of the		Ascending	Delete Ke	Yii	* Results
	12 3000000 13 9000790 14 30000044 15 3000730	13 45-398 26 38-15-398 26 16-97 288 26 18-11-398 26				Help		* Indices
	16 2000800 17 2000178	11-67-3908 294 34-15-3908 294						* Comments
	18 100000 19 100000 20 100000 21 100000 22 100000 23 100000 23 100000 25 100000 26 1000000 26 1000000 26 100000 26 1000000 26 1000000 26 1000000 26 10000000 26 100000000000 26 1000000000000000000000000000000000000	10-25-250 10 31-04-250 10 13-12-250 10 15-12-250 10 10						Add comment
	24 3000745	15-11-2008	100		15.00	8	-	e
	26 1000340	25-05-200 100 125-05-200 100	3000	11	5.09	40 31		
	27 200775	15-08-2000 10 11-05-0000 100		81. 81	5.00	1		
	29 200611	19-09-200 200	1130	13	5.00	22.84		
		38-05-2008 L25	23400	8	15.25 5.00	25		
	22 1000013 33 1000488	107-04-2008 125 28-87-2008 125		8	25.15	30		
		08-03-2508 125	11800	6	5.09	142	1	
III .		13.36.2998 629	STARL'	15	5.99	198	1	

To view the records that comprise a group, click the nodes to the left of the records.



The Group Records task will run faster if the selected database has previously been indexed on the field(s) selected as the key(s). If not, IDEA will automatically index the database by the key selected before performing the Group Records task.

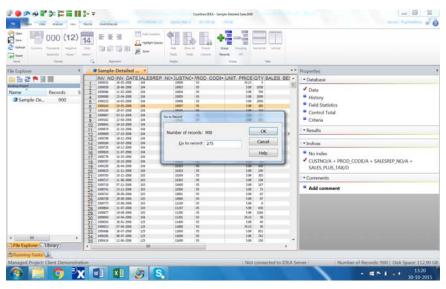
Upon completion of the Group Records task, IDEA moves the fields selected as keys to the left-most position of the database grid.

The results of the Group Records task are display-only. The Print, Print Preview, and History features are not available for this task.

Group Records mode can be turned off by clicking on Grouping Off under View in the IDEA menu tool-bar.

5.4 Go To

The Go To option is used to go to a specific record number within the active database, which is often a quicker method than paginating through the entire database.



5.5 Find

Use Find to locate records within the active database that meet the specified criteria/parameter or condition.

For example, to find the first sale entry to a customer which is a round sum sale in multiples of 1000:

Select Data – Search category – Find – Find a Record

		I = = Naria Duell Dent Rady	with I with		8 7	angle-Depiled Sales 345				
File Explorer		Samp	le-Detailed *	(w):	1000			• *	Properties	
- E 2 P 0			INV DATE ALESP	EP NC+ CUSTN	- ROD COD	DE- UNIT PRICE	QTY SALES		• Database	
Name Sample-De	Records S 900	3 32000 4 20000 5 20000 6 32000 7 20000 8 320000 9 32000 9 32000 10 20000 12 20000 13 200000 14 200000 15 320000 16 200000	Search oriteria:	35,20 25,20 35,20 35,20 35,20	888	OK Cancel Help Sau Sau Sau	136 30	,	Data Hotory Hotory Field Statistics Connol Total Coloria Coloria No.index Mo.index Konnex Konnex	IP_NO/A +
III Ide Explorer D Ruwing Tasks Managed Project O	SAINAU	** > < <*	>+ +>*/ AND		1		All Character Nameric Date / Time Matching Conditional Financial Custom Functi	ions	CFunctions Grantions are available for performing more such as date arithmetic, statistical calculations and test functions bypical They are very tind as in other	Disk Space: 112.90 Gf

Click the Equation Editor Button.

Enter the expression SALES_PLUS_TAX % 1000 = 0.

Click OK. You may find further records that meet the same criteria by pressing F3.

Find differs from Criteria and Direct Extraction by searching for the first occurrence of the specified criteria from the first record in the database, or optionally, the current cursor position. It does not find and list all occurrences.

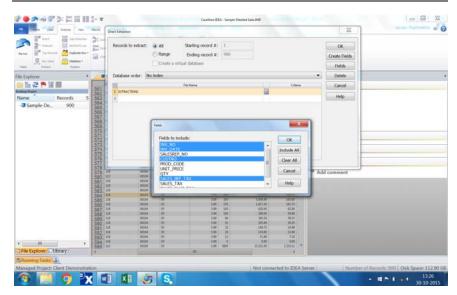
Find also differs from Search because with Search you do not need to enter criteria. With Search, you can perform keyword searches for data within fields and across records in a single database or multiple databases located in the same Working Folder or project (if you have IDEA Server).

You can also use Find Next to scan through the database searching for further records satisfying the criteria.

5.6 Extraction of Data

Direct Extraction allows the user to write and run criteria (Equations and Functions) directly on the IDEA Database. This task generates an output database which can be further interrogated within IDEA or taken up for manual validation.

Examining Data and Isolating Specific Items



Nature of the Extraction

Direct Extraction can perform a single extraction on a database, or up to 50 separate extractions with a single pass through the database. This reduces the processing time of tasks within IDEA. For example - the user can look for high value, round sum, cash payments made to all Vendors containing a specific string like '& Assoc", through a single pass of the database. For each extraction separate databases will be created within IDEA.

Basis for the Extraction

Extraction can be performed on the entire database or on its sample through data range specifications. For example, extractions can be run on the sales of all retail stores in India or for a specific range of stores, if the store references are available in the database and uniformly formatted.

Display of Results in the Output Database

An option is available to select the order of the records appearing in the output database, such as oldest inventory items appearing first on the output list.

Extractions at times run on raw data imported into IDEA. This raw data (say from an Oracle Billing Table of a Telecom Utility) contains all the relevant and irrelevant data field schemas. IDEA offers the user the flexibility of choosing the fields relevant to the extraction test which need to feature in the output database.

Writing Equations and Functions within the Extraction

The Equation Editor can be used to enter the extraction criteria. The Equation Editor allows the user to enter arithmetical, logical, and conditional equations and functions in the equation box.

A normal equation consists of

- Parameter / field within IDEA.
- Arithmetical operator like +, -, *, and more.
- Logical operator like .AND., .OR., .NOT.
- Constant like a number or character.

Sample Equation – (DebitCredit = "CR" .AND. PAN = "" .AND. Amount > 50000)

Here the Parameter / field is DebitCredit, PAN and Amount. The arithmetical operator is = and >. The Logical operator is .AND. The constant/value is 50000

A normal function consists of -

- ➤ @ function
- Parameter / field within IDEA.
- Constant like a number or character.

Sample Function - @len(PAN)<>10

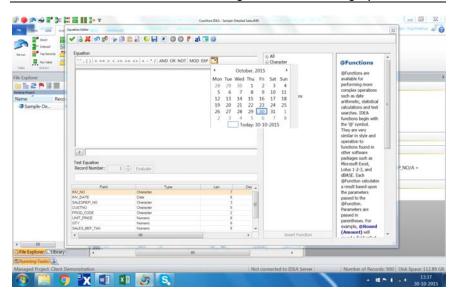
Here the @ function is @len which displays the length of a character field. The Parameter / field is PAN. The constant is 10. The @function will allow the user to test the validity of the PAN for cases where length of PAN is not 10

The Equation Editor has a Validator which checks the equation or function for its structure and syntax. Equations which have an inherent error cannot be processed until the error is identified and duly de-bugged.

The Equation Editor in IDEA provides integrated online help for building equations and functions and provides easy-to-use practical tips and good practices while devising equations.

Character and Date fields in IDEA should always be in Parenthesis. While selecting Date constants within IDEA always use the Universal Calendar in the Equation Editor.

Examining Data and Isolating Specific Items

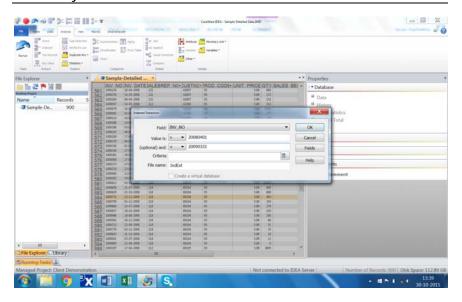


While using Character fields always sanitize the fields by using @ functions like @ strip to get rid off spaces and irrelevant punctuation marks (if the test so desires).

5.6.0 Types of Data Extractions within IDEA

5.6.1 Indexed Extraction

Indexed Extraction allows you to limit the scope of data for which IDEA searches in the database. Indexed Extraction saves time when reviewing large databases. You can select an index for the search rather than have IDEA search through the entire database.



5.6.2 Key Value Extraction

Key Value Extraction allows you to generate a series of secondary databases quickly from common values found in the main database. Key Value Extraction does not require you to create equations to perform the extraction. A key is an index created for information in a database. A key value is a piece of information indexed by the key. You have the option of creating one output database or multiple output databases.

The sector of th	Di tentan II a	Sey Value Subschart	*	
Renat 🖉 Tan Rossen 🎦 Dapkane Key *	Des Stationen auf res	Select Record Groups	OK	
Tatle Setat	Category	Existing keys: CUSTNO/A		
		Management and a second s		-
le Explorer *	Sample-Det	Orteria:	Fields	
			Help	P
Name Records S	563 2000284 08-08	Groups Found		
Sample-De 900	564 100007 34-09 565 1000079 30-04	Please select the values you wish to extract.	Clear All	
	566 199000 02-94 567 200014 15-94	There are 303 unique key values in the database.	Include All	Indistics Total
	568 XXXXX ID-10	¥ 10000	Evenentiation	TOLE
	570 1000347 III-04 571 1000389 13-08	V 10003	(m)	1
	572 3300179 94-09	V 10004 V 10005	1.000	-
	572 100179 94-09 573 200110 01-40 574 2000340 27-05	7 10006		
	575 1000.81 29-04 576 300000 27-05	V 10007 10101		
	577 1000111 27-0	¥ 10102		nts
	578 300000 13-00 579 300000 13-00	V 10199 10200		amment
	580 200081 15-01 581 2000812 05-02	V 10201 10203		
	582 100018 22-01 583 10003 23-01	2 10204	-	
	5/64 2000/71 22-12	10903		
	585 200000 HI-LI 585 2000000 34-0	Output		
	587 sources 25-10	Prefor for each database:		
	589 100011 16-12	KeyVal		
	501 2000/79 83-23	Create a separate database for each unique key		
	592 200847 B-19 593 20081 03-03	Create a vetual database		
(III)	594 200940 22.00 595 1000.00 27.44			
File Explorer Ubrary	+	H	1 mil	
Running Tasks				

Key Value Extraction can be used to generate the bank branch wise report of high value cash transactions in newly opened current accounts. Here separate database will be generated for each branch to enable branch segmental reporting and oversight on potential suspicious money laundering indicators.

5.7 Display of Data

The Display of Data function assumes that the user is aware of the numeric value or character string to be interrogated/ isolated within IDEA. Here the user identifies the first occurrence of the value or string through the Find Function or through Visual Verification.

The user then goes on to display the data through the Display All Records Containing dialog box. (Seen Below)

Illustration: As an auditor you have received a tip-off from your internal source at the client's workplace that a given Employee (Say Emp. No. 287654) is indulging in excessive overtime claims. With IDEA you can display all overtime claims submitted by Emp No. 287654 and follow up the display with actual observation of his late-sittings and the productivity thereof.

The Display Data function can also be accomplished by writing an equation or function in the Criteria, seen on the right hand side of the screen under the Database Toolbar. Writing criteria is similar to writing criteria in the Direct Extraction dialog box seen in the section on Extraction of Data.

	and the second second	24 -				Canalifying EEA - Sample-Data	olied Sales JAC					
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Analysis (res	Naria Shartina	a the	ADDRESS OF	April 1994	1. 10.004 10.004					Server.	0
a Part	Gap Detector	DE Sameranten	1 APR	Te an		Manufacty Line +						
	Sectorizan	trathater	Prost Table	-a Append	Tantos .	Tarabia *						
	Dekate fay*	Dwt.		Wave Corvector	Citer *							
Rey Take	Entering *	-		Corpers								
	(1000	(mm)		111		Pacifier				2		
Explorer			e-Detailed			Less seat have				Properties		
						- ROD CODE- UNI				Database		
	ecords S	568 10000M	10-06-2008	112 113	42111 90002	65 65	5.09	142	_	J Data		
Sample-De_	900	569 100004 570 1000047 571 1000047	08-04-2008 17-06-2008	112 112	50003	65 05	5,00	972 117	- 1	 History 		
11111111111		572 3000279	06-05-2008	112	50007	6	5.99	127	-0	Field Statistics		
		571 2000,800 572 2000,279 573 2000,388 574 2000,388 575 2000,381	01-02-2908	111	30006	6	5.00	138 298 460	_	Control Total		
		575 2000.001 576 2000.000	25-04-2008 17-03-2008	112	50013	65	5.88	400	_	Criteria		
		577 1000170 578 1000174	37-85-2908 23-04-2008	111 111	0.005	Append Field	1.05	- 19	-	· Results		
		579 200060	12-02-3908	118	90000	Modify Field				Page 1		
		580 2000383 581 2000612	19-05-2908	312 118	60304	Remove Field(s)				* Indices		
		582 20001M	23-87-2908	118	0000-	Field Manipulation				No index		
		584 1000771	33-13-3908 01-13-3908	118	60324	Create Report					ROD_CODE/A + SALESREP_N	D/A.+
		586 100000	34-07-2008	118	6010-	Views				SALES_PLUS_TA	X/D	
		587 1000857 588 1000858	20-10-2008 19-08-3008	118 118	6033-	Display All Records O	Containing	"60005"		CUSTNO/A		
		589 2000346	16-12-2908 23-89-2008	118 118	6030- 6020-	Show Field Stats	Ceremon and			* Comments		
		591 3300779	03-11-2908	118	60334	Highlight Spaces				Add comment	6	
		592 2000647 593 200062	31-15-3908 83-67-2908	118	6000-	Zoom						
		593 3300043 594 3300845 595 3300307	22-09-3908 13-04-2908	218 11.7	6033-	Define Action Field						
		596 1000M1	09-07-2008	118	6000F	Save as Selected						
		598 1000067	23-05-2008	111	60.20	Close Database			Ctrl+D			
	,	5199 2000532 600 2000445	10-04-2008	113	6520	Close All Databases						
III le Explorer CDbr		601 3300436	05-02-2008	113	60.292	Delete Database						
aming Tasks									_	the second s		
	Analyse Des	2+ = 102 - 2210		🧭 S.		Contract EEA - Sample-Data					- #5130	1343 30-10-201
an a	Andres Ves Sadres Ves Gap Detector Sadres tar	le T Nation Statute Distances		1	in tanta							30-10-201
The set	Andres Ves Sadres Ves Gap Detector Sadres tar	Z= ∓ Narit statule Dit termenatur			Anna A	Manalary Line *						30-10-201
Den	C C C C C C C C C C C C C C C C C C C	In T New particular In Summation In Content Content Content		The series of th	in tanka Landar Church	Manalary Line *			-			30-10-201
Description Product Product Tablearch Strytals Stratt	Andreas Andreas Gen Detection Mandrea Cape Andreas A	In T Nacional and the Second and Const Con	e-Detaile	Capity Af Faced Co	tantur Tantur Coner -	Manalary Line *				- Contraction of the second		30-10-201
Store Total	C C C C C C C C C C C C C C C C C C C	Contention Conten	e Detaile 0 RV/ DA 12 Post Table	The series of th	tantur Tantur Coner -	Manalary Line *			OK	base		30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	24 - Norm Secondary 25 Secondary 26 Secondary 27 Secondary 28 Secondary 29 Secondary 20018 Secondary 573 200018 574 200018 574 200018	e-Detaile 0 NV DA 2-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	Capity Af Faced Co	tareng _	Manalary Line *						30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Superior Superior	e-Detaile 0 NV DA 5-0-201 2-0-200 2-0-200 2-0-200 2-0-200 2-0-200 2-0-200 2-0-200	Constant Constant Constant Constant Constant Constant Display All Records Car Display all reco Field:	tantas tantantas tan	Manalary Line *			OK Cancel	ta itory		30-10-201
Topological States of the second states of the seco	C C C C C C C C C C C C C C C C C C C	Image: Superior Superior	Portaile O INV. D4 Total Source 124 Total Source 124 Total Tota	Corpus Corpus Corpus Display All records Corp	tantas tantantas tan	Manalary Line *			Cancel	ta itory id Statistics		30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: State State Image:		Display all reco Field:	translave translave	Manalary Line *				ta itory		30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control of the second	Company C	Paran Anard Constant Constant Display all reco Field:	translave translave	Manalary Line *			Cancel	ta tory Id Statistics ntrol Total		30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control Image: Control Image: Contro	e-Detaile 0 Roy Detaile 0 Roy Detaile 0 Roy Detaile 10 40,000 10 40,0000 10 40,0000 10 40,000 10 40,0000 10 40,0000 10 40,	Coperator:	translave translave	Andread (1997)			Cancel	ta tory Id Statistics ntrol Total		30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control of the second secon	Post	See an Auerol Static Corrector Static Corrector Static Corplex All Records Corr Field: Operators: Value:	Anthone anthone cds where: CUSTNO anti- cds (c) anthone CUSTNO anti- c) c) c) c) c) c) c) c) c) c)	Andread (1997)			Cancel	ta tory Id Statistics ntrol Total		30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control of the second secon	e-Detaile 0 Roy Detaile 0 Roy Detaile 0 Roy Detaile 10 40,000 10 40,0000 10 40,0000 10 40,000 10 40,0000 10 40,0000 10 40,	Coperator:	Add to ou	Andread (1997)			Cancel	ta tory id Statistics ntrol Total teria ults cos		30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Control	August	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	ta tory doty dotatistics ntrol Total beria ults cos i Index		30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C		Partiel Par	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	ta Itory id Statistics otrol Total beria ults cos lindex STNO/A + PI	ROD_CODE/A + SALESREP N	30-10-20
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C		Appr	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	ta tory d Statistics ntrol Total teria dts ces index STNO/A + PI ES_PLUS_TA	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	2	Aprop.	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	a tory d Statistics netrol Total teria alts ces index SINO/A + P ES_PLUS_TA SINO/A	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Comparison of the second	Approx	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Construction Construction Image: Construction Image: Construction <	niii	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	a tory d Statistics netrol Total teria alts ces index SINO/A + P ES_PLUS_TA SINO/A	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control Image: Control Image: Contro	Contaile	See Ser Search Sear	Add to ou	Andread (1997)			Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control Image: Control Image: Contro		Depley Al Record Co Operators: Value: Controls Operators: Users: Value: Controls Operators:	Add to car the formula	rend official	10		Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control Image: Control Image: Contro	Description Image: Control of the second	Dipetay of Control Con	Add to cu • '60104"	rend collecto			Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
Rever and a constraint of the second se	C C C C C C C C C C C C C C C C C C C	Image: Control Image: Control Image: Contro	control	Depley All Reads	Ambodie Ambodie Commentation of the second	rend offers	19		Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
series of the se	C C C C C C C C C C C C C C C C C C C		Image Image Image <td>Dipley All nool Operators Under All Neoreman Dipley All neoreman D</td> <td>Arthur and a second a second</td> <td>rent official</td> <td>190 190 190 190 190 190 190 190 190 190</td> <td></td> <td>Cancel</td> <td>ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments</td> <td>ROD_CODE/A + SALESREP N</td> <td>30-10-201</td>	Dipley All nool Operators Under All Neoreman Dipley All neoreman D	Arthur and a second	rent official	190 190 190 190 190 190 190 190 190 190		Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
splore references	A Augur Barran Augur Barran Aug		Image: Control of the second	ar	Imaka I	rend collecto	1		Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
splorer set of the set of the se	ecords S 900		Image Image Image <td>Dipley All nool Operators Under All Neoreman Dipley All neoreman D</td> <td>Arthur and a second a second</td> <td>rent official</td> <td>190 190 190 190 190 190 190 190 190 190</td> <td></td> <td>Cancel</td> <td>ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments</td> <td>ROD_CODE/A + SALESREP N</td> <td>30-10-201</td>	Dipley All nool Operators Under All Neoreman Dipley All neoreman D	Arthur and a second	rent official	190 190 190 190 190 190 190 190 190 190		Cancel	ta tery d Statistics etrol Total teria alts ces index STNO/A + PJ ES.PLUS,TA STNO/A ments	ROD_CODE/A + SALESREP N	30-10-201
Support	A angre en		Image Image Image <td>Dipley All nool Operators Under All Neoreman Dipley All neoreman D</td> <td>Arthur and a second a second</td> <td>rent official</td> <td>19 30 30 30 30 30 30 30 30 30 30</td> <td></td> <td>Cancel Help</td> <td>h hay hay shares hool Total and hay index index index is NIXOA - PIN STROA - P</td> <td>ROD_CODE/A + SALESREP N</td> <td>30-10-201</td>	Dipley All nool Operators Under All Neoreman Dipley All neoreman D	Arthur and a second	rent official	19 30 30 30 30 30 30 30 30 30 30		Cancel Help	h hay hay shares hool Total and hay index index index is NIXOA - PIN STROA - P	ROD_CODE/A + SALESREP N	30-10-201

The Display Data function creates a temporary view of the database. This view appears as a result on the right hand side of the screen. The view can be saved by invoking the Save As option in File in the Menu Tool Bar.

Illustration: The Criteria can be used to identify all blank address fields in a customer master file to review the authenticity of the customer.

Identifying Exceptions and Irregularities

6.0 Introduction

IDEA includes two key functions to identify exceptions, irregularities, anomalies and errors.

These are:

- Duplicate Detection
- Gap Detection

These functions assist the user to sift through large volumes of data and help pin-point specific duplicate entries or specific missing entries. These also help the user obtain an assurance on all the data reviewed by it. The duplicate or missing items identified can be taken up for tests of transactions and balances (substantive tests) after running the respective duplicate and gap tests within IDEA.

Duplicate Detection and Gap Detection tests are standard tests which are applied to every database to study the completeness and accuracy of its contents. The tests do not require much querying experience and resemble plug-n-play tests through standard Graphical User Interface (GUI). Both tests run largely on formatted sequential data fields like Invoice Number, Purchase Order Number and Indent Number as examples.

The user needs to evaluate the need for such tests, depending on the functional control being tested within the process. For example, if the user is reviewing financial controls within ERP-SAP R/3 for Procure to Pay, the chances of the system generating duplicate Purchase Order Numbers or missing numbers is quite improbable. But if the user would like to test for duplicate Vendor Bills submitted (an external document), which is a non-system generated field, the duplicate test could be employed.

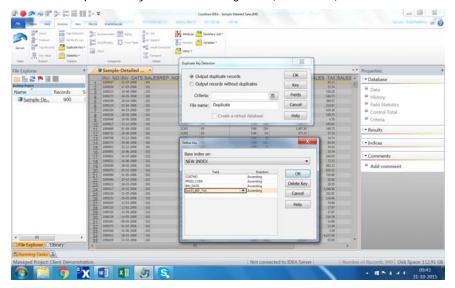
6.1 Duplicate Detection Test

Duplicate Tests within IDEA include Detection tests and Exclusion tests.

Duplicate Key Detection includes identifying Duplicates and identifying Non-Duplicates items within the active database.

6.1.1 Duplicate Key Detection – Output Duplicate Records can run on any normal database which is active within IDEA. The test allows IDEA to look for duplicates in and up to 8 fields. The fields may be numerical, date or character based.

The fields on which duplicate tests have to be run are entered as 'Key' fields within the Duplicate Key Detection dialog box. (See Below)



IDEA will look for duplicate cell-items in a combination of the Key fields selected.

Illustration: If a user is looking for duplicate sales invoices raised on customers, and if he enters the Invoice No., Customer No., and Invoice Amount as the Key fields, IDEA will look for duplicate sales made to the same customer for the same amount against the same invoice number.

The Duplicate Key Detection dialog box also offers the flexibility of Criteria.

The Criteria allow the user to look for duplicates on a filtered part of the active database and not on the whole active database.

Illustration: If a user is looking for duplicate sales invoices raised on customers, and if she enters the Invoice No., Customer No., and Invoice Amount as the Key fields, IDEA will look for duplicate sales made to the same customer for the same amount against the same invoice number. In

this case, the user can enter criteria wherein IDEA will look for duplicates only for a particular product code or a particular business division of the group. Here the search is more specific and meaningful.

The user can also choose IDEA to display certain fields in the exception output database through the option Fields. This is relevant when the database being investigated has a large group of fields like master tables. The user may like to specifically report on duplicate items in a few relevant fields like Bank Account Number, PAN Number. Here she can use the Fields option.

6.1.2 Duplicate Key Detection – Output Records without Duplicates follows the same mechanism as the Duplicate Key Detection – Output Duplicate Records test in terms of Key fields, Criteria and Fields. The only difference in the two tests is that in the Duplicate Key Detection – Output Records without Duplicates looks for unique items in the database i.e. non-duplicates. This test will exclude all the duplicate items.

The Duplicate Test within IDEA has been improvised significantly over the years in terms of speed and performance. The test generates exceptions on large databases in record time. The test is also user-friendly as the user does not have to separately sort the database or write a vertical cell-difference function (like in MS-Excel) prior to test-run. One click of a button in the Duplicate Key Detection dialog box takes care of the sort and vertical cell-difference function automatically.

6.1.3 Duplicate Key Exclusion – Output Duplicate Records where one Field is Different

The duplicate key exclusion function is similar to the duplicate key detection function in the sense that the test allows IDEA to look for duplicates in and up to 8 fields. The fields may be numerical, date or character based. In addition, Duplicate Key Exclusion offers the framework where one additional field other than the 8 matching fields may be different.

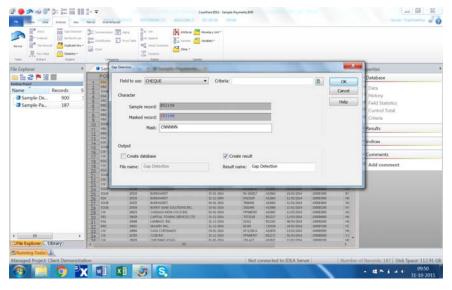
Illustration: In concurrent audits of Registrar and Transfer Agents (R&T), the AMC (Mutual Fund) pays the R&T, brokerage as entry and exit loads where the units have been purchased or sold through a SEBI registered broker. Here specific brokerage rates are fixed for specific mutual fund schemes. Moreover, brokerage is not paid where the transaction is done through a direct dealing popularly called 'ARN-DIRECT" in the Broker Code or where transactions are over Rs. 5 crores. Here, through Duplicate Key Exclusion, the auditor can identify incorrect / varying brokerage rates paid for the same mutual fund scheme for the same broker code.

Duplicate Key Exclusion is popularly called De-Dup and is used extensively for effective Know Your Customer (KYC) checks in the Banking and Financial (BFSI) sector and Commodity and Instrument Exchanges on large customer master files.

6.2 Gap Detection Test

Gap Detection within IDEA can be deployed on numeric, character and date fields.

Gap Detection can run effectively on numeric fields provided the fields are uniquely / sequentially formatted and represent the same logical field. For example, if gap detection is to be run on a Cheque Number field, cheque Number being numerically formatted, the test can be run for the entire cheque series or a cheque range. In most cases, multiple cheque series of varying ranges are used by organizations. In such instances the user can look for gaps in a specific cheque series range.



The user also has the express option of specifying a Gap Increment. This is true of a scenario where the numeric field number does not increment itself (through system or user design) by one. The increment is some number other than one like a Product Code which is in multiples of 1000, say 1000-1050, 2000-2050, etc. Here the Increment Range can be specified as 1000, so that IDEA looks for gaps within the actual number range (1000-1050) and not between various number range series (1050-.2000).

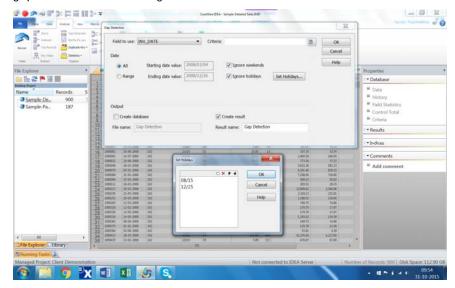
The Gap Detection dialog box also offers the flexibility of Criteria.

The Criteria allow the user to look for duplicates on a filtered section of the active database and not on the whole active database.

Illustration: If the user is looking for missing collection receipts issued to borrowers of a bank, and if she enters the Receipt No., as the Fields to Use, IDEA will look for missing receipts against instalment collections. In this case, the user can enter criteria wherein IDEA will look for gaps only for a particular loan product code or for a particular zone or collection agent. Here the search is more specific and meaningful.

Gap Detection can run efficiently on date fields also.

Gap Detection on date fields follows the same mechanism as numeric field gaps, in terms of date ranges and Fields to use.



IDEA offers the user the flexibility of excluding Holidays and Weekends from the gap testing for dates to identify genuine date gaps and not gaps attributed to holidays and weekends. Holidays can be set by enabling the option Set Holidays. Public holidays in a financial year may be entered herein. Weekends by default follow the universal weekend route of Saturday and Sunday.

The default universal weekends can be changed by altering the Weekend Settings under Date Options under File – Options – System Tab. This is relevant for organizations operating in geographical locations like the Middle

Eastern and Gulf States where weekends are normally Friday and Saturday.

THE Experimental Sample-Den. 500 Sample-Den. 500 Sample-Den. 187	System Database Grid Reports Caldons Rebotin Calcharge Robotin Qarck Access Toolbar	User interface options Automatically generate output distates ni Automatically generate each arms Automatically generate each arms Database proper help Do not record field changes made through Do not allow the distellon drino-withal fiel Database the Benan task for exponse Date options Century output D0 Veelensts Files Ausyns use OEA Library folders when option O on generates a "Bad Output Jobates and Data" distates ent O on generates a "Bad Output Jobates and DataBase ent O to generates a "Bad Output Jobates and Output Jobates and DataBase ent O to generates a "Bad Output Jobates and Output	Editable Media in History 25	tatistics It Total It It It It It It It It It It
Anaged Project: Clerit Demon	atration	(m) (S)		r of Records: 900 Disk Space: 112.90 Gr

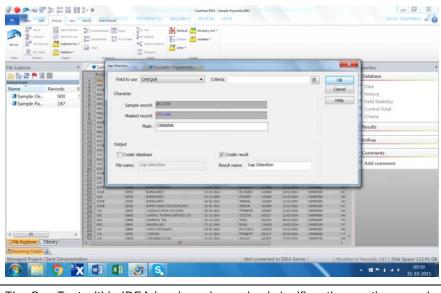
Illustration: the production server backup for an online e-tailer has to be taken every week-day on to a backup server as part of a Business Continuity Plan for the e-tailer. The backup history log can be imported into IDEA and Date Gap Detections can be run on the log to identify missing back-up dates (excluding Holidays and Weekends). This exception report will help to support the IS Auditors' assertion that controls are not being implemented regularly with regard to off-site data backups.

Finally Gap Detections run well on Character fields too.

The character fields in such cases need to be of the same length and must follow the same syntax. Character Gap Detection cannot run on Names or Addresses where the length of the field and Syntax does not follow any fixed pattern.

The perfect fields for for such tests are fixed length alpha-numeric fields like Employee ID's in a Payroll Data File. If the Employee ID is always formatted as A123456 for instance, where the syntax rule CNNNNNN works well, the Gap Detection detects missing employees within this combination of digits. This combination of digits is often called a MASK in IDEA.

Identifying Exceptions and Irregularities



The Gap Test within IDEA has been improvised significantly over the years in terms of speed and performance. The test generates exceptions on large databases in record time. The test is also user-friendly as the user does not have to separately sort the database or write a vertical cell-difference function (like in MS-Excel) prior to test-run. One click of a button in the Gap Detection dialog box takes care of the sort and vertical cell-difference function automatically.

Chapter 7 Profiling Data

7.0 Introduction

CAAT - IDEA can be used effectively as a Management Tool. It can play a vital role in assisting Management Teams to make key decisions by providing it with valuable, insightful and reliable analytical information on a near real time basis.

The Tool offers three key functionalities for Business Analysts:

- Summarization
- Stratification
- Pivot Table

Summarization - Allows the user to prepare and view a consolidated overview, i.e., a summarized position for specific fields as defined by the user. For instance, a Stock Broking firm dealing in the Options market can generate a Day-wise summary for Total Options Bought and Sold along with the Total Net Exposure for the day. The Firm can then go on to detect Net Exposures in excess of say Rs. 25 Lacs per day.

Stratification: Facilitates a range of band wise or interval wise based consolidation of data. For instance – sales made by an entity to its Customers through various Channel Agents can be stratified into value bands with respect to Channel Agents. Stratification will reveal sales made below the minimum floor price and above the ceiling price for different Channel Agents.

Pivot Table: it enables a cross-tab analysis of data. Here two data fields or more can be analysed against each other with respect to specific numeric fields. For instance, cross tab analysis can be performed on travel claims and travel policies of a company. Air travel rationalization can be performed to identify the most cost effective airline on fare basis to specific destinations, for a specific travel class and travel date too.

7.1 Summarization

Summarization accumulates the values of Numeric fields for each unique key. For example, summarizing an Accounts Payable database by supplier account number (the key) and totaling invoice amounts produces a database or Results output of outstanding liabilities by a supplier. The Summarization task provides:

- A list of unique items (keys) in the database
- The number of records/count for each key
- Totals of one or more numeric fields for each key

The user may select up to eight fields to summarize or if using one field can use the Quick Summarization option. By using multiple fields, the Summarization task works faster and is more efficient on a database that has been sorted.

Select Quick Summarization only if you are summarizing by a single field and if there are fewer than 32,000 unique values in the field.

The user may select any number of Numeric fields to total. The resultant database provides a list of unique key items and the number of records (NO_OF_RECS) for each key from where one can drill down to display the individual records for the key.

The user can select additional field data from fields from the first occurrence or the last occurrence.

The summarization result grid is limited to 4,000 rows and does not display results beyond that. If the result expected is more than 4,000 rows, then there is a need to create a database.

Summarization Examples

Example 1

The following database of cheque payments is to be summarized by Supplier with the payment amount being selected as the Numeric Fields to total and the Cheque Date selected as an additional field to include. The Statistics to include automatically appears as Sum.

Data Analysis for Auditors

SUPPLIER	CHQNO	AMOUNT	DATE
A100	348603	100	1995/12/31
A100	412591	500	1996/02/01
A100	311192	200	1995/09/04
B200	326319	300	1995/11/03
B200	521361	100	1996/01/04

Cheque Payments database to be summarized

The output database of summarized cheque payments shows three records for SUPPLIER A100 with a total amount of 800 (100, 500, 200). Although the date column displays 1995/12/31, it does not mean that there were three payments on 31/12/1995, but that this was the first entry in the cheque date field for the first record for SUPPLIER A100.

SUPPLIER	NO_OF_RECS	AMOUNT	DATE
A100	3	800	1995/12/31
B200	2	400	1995/11/03

Example 2

Stock items to be summarized

WH	STOCK_NO	QTY	PRICE	VALUE	DATE
NW	X-345-886	25	10.00	250.00	1996/06/30
NW	X-345-886	2	10.00	20.00	1996/09/15
NW	X-345-886	10	11.50	115.00	1996/11/03
NP	P-666-188	2	4.90	9.80	1996/12/02
NP	P-666-188	1	5.65	5.65	1996/10/08

To summarize the above database of stock items by stock number (STOCK_NO) within each warehouse (WH), we would select:

Fields to summarize: WH then STOCK_NO

Numeric fields to total: QTY, VALUE

Other fields to include: PRICE, DATE

The two tables below show the different resultant databases created,

depending on whether we have selected the (other) fields data from the first or last occurrence for each key.

First Occurrence

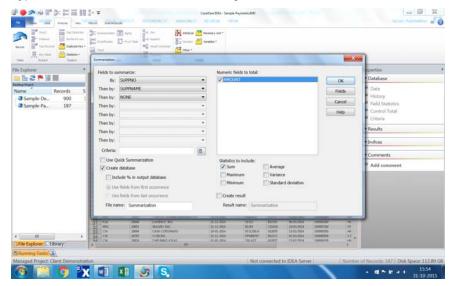
WH	STK_NO	NO_OF_RE CS	QTY	PRICE	VALUE	DATE
NP	P-666-188	2	3	4.90	15.45	1996/12/02
NW	X-345-886	3	37	10.00	385.00	1996/06/03

Last Occurrence

WH	STK_NO	NO_OF_RECS	QTY	PRICE	VALUE	DATE
NP	P-666-188	2	3	5.65	15.45	1996/10/08
NW	X-345-886	3	37	11.50	385.00	1996/11/03

The choice of first or last occurrence of (other) fields makes a difference to the PRICE and DATE field values. In this example, this information could be misleading. So one has to take care while selecting (other) fields. The (other) fields selected, in this case PRICE and DATE, do not affect the calculations, they are simply either the first or last occurrence of data for that field associated with the key. Therefore, in the examples given above, QTY + PRICE do not decide the value and are only meant for display and reference.

Typical view of the Summarization dialog box in IDEA is -



Typical view of the Summarization result in IDEA is -

Toda Parts	Constantia a const	est Table	Attitute and Handary Linit * Earder and Tanataka * Ofter * Tangte				
ile Explorer *	Sample-Pay	yments *				- 1	Properties
1 2 P G II	1 6 STET						· • Database
water Preset	Numeric fields to	ANCENT		Statistics Sum			Data
Name Records S	SUPPNO	SUPPNAME	# Records		AMOUNT - Sum	(%) AMOUNT - Sum	History
Sample-De 900	10500	AASMAN DESIGN L	# PROCORDS	(%) # Records 0.53	545.09	0.01	Field Statistics
Sample-Pa_ 187	10900	ACKLANDS - GRAI	1	0.53	485.19	0.01 =	Control Total
	11400	AG JOYERIA	1	0.53	227.62	0.01	Criteria
	11810	DKNY INC	1	0.53	4.839.92	0.13	
	12203	ANKE KRUSE ORG	4	0.53	4,639.92	0.02	• Results
	20028	BETONSTEIN GMBH	4	0.53	52.845.00	1.39	Stratification
	20041	BUOUTERIE CAR	1	0.53	290.68	0.02	 Summarization
	20129	BILEVICH BOEDO	1	0.53	1.000.00	0.03	
	20133	BISHOP PUBLISH	2	1.07	21.395.57	0.56	• Indices
	20204	BOREAL LABORAT	1	0.53	54 721 88	1.44	No index
	20352	GUESS INC	1	0.53	75.373.68	1.99	AMOUNT/A
	20403	BRIDDIG	1	0.53	6.844.63	0.18	SUPPNO/A + SUPPNAME/A
	20414	CITIZEN INC.	1	0.53	4 522 28	0.12	
	20508	FOSSIL INC	2	1.07	47 374 53	125	· Comments
	20532	WERNER KG	3	1.60	95 875 02	2.53	Add comment
	20535	BURKHARDT	6	3.21	1.09.430.24	2.88	
	20535	JOHN PETERSON	1	0.53	96,166,49	2.53	
	20536	JOHN PETERSON	1	0.53	78,262.98	2.06	
	20554	BURNT SAND SOL	i.	0.53	38,149,16	1.01	
	20625	CADEAUX MON CO	1	0.53	57,170.89	1.51	
File Explorer Dibrary	1	CADITAL TOURISM	-	0.63	1.677.50	0.04	
Running Tasks							

7.2 Stratification

The process of stratification involves creating bands based on ranges of values (normally from the minimum to the maximum values of one or more fields) and grouping the records from the database into the appropriate bands. By totaling the number of records and value of each band, one can gain insights into the profile of the values in the database. The stratification analysis is also useful for determining high and low cut-off values for testing exceptional items - outliers.

With the Stratification task, one has the option to create both a database and a result with the option to include details of the band range within which each record falls. One can then use this database for further analysis.

Stratification produces a numeric stratification analysis for each key in the database. For example, one could produce a profile of sales for each salesperson. This task can potentially create an extremely large volume of output. Therefore, there is an option to specify low and high cut-off values to restrict output to only the keys required, the keys whose total value of transactions is between the specified range are output.

One needs to remember that IDEA is case sensitive when processing data, therefore, one has to consider the case when specifying stratification bands. (If data has been input inconsistently?), such as in both upper and lowercase, we have to create a Virtual field with @upper or @lower to convert all text to the same case, and then stratify the Virtual field.

To include all items in the stratification analysis, the bands should start less than the minimum value and the upper band greater than the maximum value of all fields.

When stratifying a Numeric field, before entering the stratification bands we have to use the Field Statistics task to know the minimum and maximum values of the field(s) being stratified as well as the skewness of the values. These values help to determine the interval start and end values as well as the increment size.

Date and Character stratification are different than Numeric Stratification in the sense that different fields are totaled to the one used for banding.

Stratification Examples

Stratifying a Numeric field

When auditing a payments system, we can identify invoice splitting just below authorization by stratifying the payment amounts. We have to make sure to specify a narrow band below each limit.

A higher count and value of total payments just below the authorization limits would indicate that invoice splitting has potentially taken place.

Stratifying a Character field

Sales data can be analyzed by Business code area. We have to specify the Business Code field as the Character field to stratify and enter the band using as many characters of the code as are required to identify the area.

Stratifying a Date field

We can analyze employees by their date of birth to get an age increment profile, or we can analyze fixed assets by date of acquisition.

Stratification is the first stage process in identifying the classification of values in any population from High (A Class) to Low (C Class). This functionality within IDEA when used in conjunction with "Field Statistics" gives the user a true sense of the dispersion of values in a normal database. The user is able to easily appreciate the famed "80-20 Rule of Business" i.e. 80% value accounted for by 20% records and 20% value accounted for by 80% records". Once the stratum are created, the user can then employ "Stratified Random Sampling" where complete assurance can be sought on the confidence interval within the population.

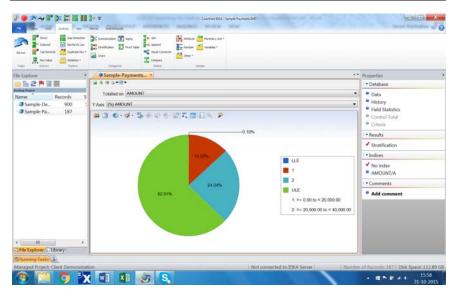
Typical view of the Stratification dialog box in IDEA is -

Array	Stratication	Increment: 20,000.00	
The property of the second sec	Field to stratify: POSTED_BY		Cut att
III I	The These Delivery these the	11.01.004 (19.217 (25.05) (19.0104) (19.	NUMBER OF STREET

Typical view of the Stratification result in IDEA is -

Rarrar Polenal III II Rarrar Polenati III II Rarrar Rarrar III III III Rarrar III III III III III III IIII IIII	e Detection Infordicion gécale Ray * alleba * science	Die Szeneroster 🚺 Crastituster 🚺 Orat Categorie		Andure Monday UNI* Forder Monday UNI* Forder Monday * Other*					
le Explorer			yments *						Properties
		10 6 4 5 · H ·	•						• • Database
lame Record	5	Totalled on	AMOUNT						Data
Sample-De 90 Sample-Pa 18	0 5	Stratum #	++ L Limit	< U Limit	# Records	(%) # Records	AMOUNT	(%) AMOUNT	 History Field Statistics
Sample-Pa_ 18	-	1	0.00	20,000 00	117	62.57	4,96,057.44	13.07	Control Total
		2	20,000.00	40,000.00	31	16.58	9,14,151.75	24.09	P Criteria
				Lower limit exceptions:	3	1.60	-3,888.17	-0.10	• Results
				Upper limit exceptions:	30	19.25	23,88,187.70	62.94	 Stratification
				Totals	187	100.00	37.94 508 72	100.00	* Indices
		1							 No index AMOUNT/A
									· • Comments
									Add comment
III JFile ExplorerUbbrary	,								

Profiling Data



7.3 Pivot Table

The pivot table is an interface where the user has the freedom to define how the data is to be displayed and organized.

When creating a pivot table in IDEA, we use the Pivot Table Field List dialog box to choose the fields to be included in the pivot table. The position on the pivot table defines how the data is organized and summarized.

Once the fields are in position, the user can move them about easily by dragging and dropping. One may even drag a field right off the window to remove it.

The user may summarize the numeric data in many ways, by selecting from the Pivot Table Field List dialog box. One can configure the pivot table to show record counts, totals, averages, or minimum/maximum values. Color code summary fields can be assigned.

The records represented in any cell are visible by selecting the Drilldown Data button on the Pivot Table toolbar to display an Extraction Preview window from which the data may be saved to another database or printed.

The user can create, record and save an IDEAScript relevant to the pivot table by using the IDEAScript button on the toolbar in the Pivot Table result.

The pivot table is capable of holding 50,000 rows x 2,000 columns within IDEA.

After creating the Pivot table the following options are available:

- Change the summary statistic for data fields.
- Export the displayed portion of pivot table to Excel and / or an IDEA database too.
- Change the sort order.
- Show/hide footer.
- Change the background colour of cells in the Total rows and columns.
- Show Top Records.
- Create IDEA Script code based on the pivot table.
- Add a new field to a pivot table.
- Delete a field from a pivot table.
- Filter displayed field items.
- Expand or collapse fields.
- Drill down field data.
- Print the pivot table.

Tata Street	Gap Detects	" Petrature IT not		the Content	Manadary Line *						
ile Explorer		* Sample-Payn	nents ×								Properties
- E- 2 P G I		RES- 100	-								• Database
Inatio Proved	211	Drop Page Fields He									Data
Name R	lecords	S Sum: AMOU/AUTT		<i></i>					-		
Sample-De	900	SUPPNA E	BC.	CB	CW	H.M.		H.V.	HMV	HV =	History
Sample-Pa_	187	1 MOORE					4,792.00				Field Statistics
		A MEADOW							4,34		Control Total
		A RAID		-695.50							Criteria
		AASMAN D			545.09						I = Results
		ACKLANDS	485.1								
		AG JOYERIA	Prest Table P				1	1			Stratification
		ANKE KRUS	1000					1	820.63		Summarization
		BETONSTEL	Drag fie	Drag fields to the pivot table:						Pivot Table	
		BUOUTERIE	AMOU	AMOUNT (0)					790,68	5	* Indices
		BILEVICH B		AUTH (C) CHEQUE (C) INVOICE (C)							
		BISHOP PUB							2.713.16	j	No index
		BOREAL LAB		INVOICE, DATE (D) E						AMOUNT/A	
		BRIDOIG							SUPPNO/A + SUPPNAME/A		
		BULOVA IN	PAY_D/ POSTE	ATE (C) D_BY (C)							· Comments
		BURKHARD	PURCH	PURCH_ORDE (C) 61,067.50					3		
		BURNT SAN	SUPPN	AME (C)			-	38.149.16			Add comment
		CADEAUX	Contraction of the second s			1 Dames	10.000				-
		CAPITAL TO	Add 1	0		Remove	Close		1,677.20)	
		CARTER BO		-		-	-	1	1.000		
< III		CARY S MA							18,688.49		
C File Explorer Ub	rary/	() () () () () () () () () ()			100						

Typical view of the Pivot Table dialog box in IDEA is -

Pivot Table Examples

In the General Insurance industry, motor premiums and claims form a crucial part of the underwriting business of a company. Historically, motor claims have always proved to be a loss making unit of a company due to the sheer volume of claims with premiums not following pace with the escalation in claims. Recently a leading General Insurance company undertook a novel exercise within IDEA where the motor claims admitted and paid for the last 5 years were analysed within IDEA's Pivot Table. The vehicle number of the insured was captured in the Pivot Row and the Claim Admitted for the last 5 years was captured in the column. Vehicles making repetitive claims were identified through this activity.

Working with Multiple Files

8.0 Introduction

Just as the extraction of data into IDEA from a client's application system is imperative for the use of IDEA in the same manner tasks within IDEA can be optimally performed only with the help of multiple file operations and activities.

Even though data is getting centralized all over the world, it is distributed in data tables and schemas. A single buy in ERP-SAP for instance will affect data tables in FI and MM modules.

The CAATs auditor, with the assistance of IT Support, can bring data from different data sources into IDEA. The independent data sources can be merged, compared, reconciled to prepare a single base file for testing in IDEA.

IDEA offers the following techniques for working with multiple files:

- Visual Connect
- Join
- Append
- Compare

These techniques allow the user to combine multiple files conveniently for analysis and review.

8.1 Visual Connector

Many databases created by large systems, such as SAP and Oracle, are composed of several tables. When importing those tables (using Open Database Connectivity - ODBC), they appear in IDEA as separate files or databases. In order to rebuild the relationships, one can use the Visual Connector.

Visual Connector allows the user to generate a single database from several other databases that share common or "key" fields. To create a visual connection between databases, we need to select a primary database, and then connect databases that contain matching records.

The relationship that the Visual Connector creates between the databases is one to many, meaning that the primary database can have several matches in connected databases. All records in the connected databases that match records in the primary database are included in the output database. For example:

Database A

Primary Database

Output Database

Name	Name	Value	Name	Value
Mary	Mary	A	Mary	A
Bill		В		В
		С	ĺ	С
	Bill	М	Bill	M
		N		N
	Fred	G		
		L		

Database A contains several records that match records in the primary database. All matching records are included in the output database. Records in Database A that do not have matches in the primary database Fred - G, and Fred - L are not included in the output database.

Illustration of the Visual Connector

Let us select the following three databases:

- Database 1 contains product codes, product numbers, and prices.
- Database 2 contains product codes and descriptions.
- Database 3 contains product numbers and regional sales.

We select Database 1 as the primary database, because the other two databases have at least one key common to Database 1. After we define visual connections between the three databases, the output database contains all records relating to the common keys in the three databases.

Data Analysis for Auditors

Database 1

PRODUCT_CODE	PRODUCT_NUMBER	PRICE
47	8000	6.29
48	6722	7.80
49	4467	3.29

Database 2

PRODUCT_CODE	DESCRIPTION
47	Cup
48	Glass
49	Bottle

Database 3

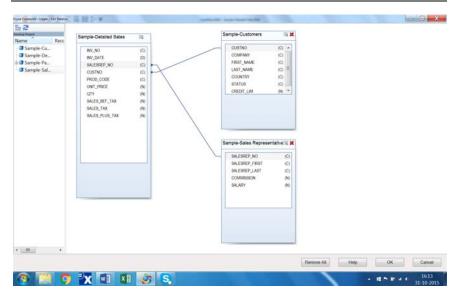
PRODUCT_NUMBER	REGIONAL_SALES
8000	Toronto: 3700
8000	London: 5800
4467	Toronto: 6800
4467	London: 4400

Output Database

Information for PRODUCT_NUMBER 6722 is absent from the output database because Database 3 contains no information for PRODUCT_NUMBER 6722.

Typical view of the Visual Connector dialog box in IDEA is -

Working with Multiple Files



8.2 Join Databases

We can use the Join Databases to:

- Combine fields from two databases into a single database for testing.
- Test for data which matches or does not match across systems looking for unreconciled items.

Тір

The Join Database runs considerably faster on sorted rather than indexed databases.

We can join or match databases only if they contain a common field (referred to as the "key"). The common field/s (up to eight in number) do not need to have the same name or length but they must be of identical field type. So while joining a bank book with the bank statement if bank book contains the cheque number as numeric and the bank statement contains the instrument number as character, the bank reconciliation join cannot take place till one of the fields are changed to unify the same.

We can join only two databases at a time.

- IDEA joins the secondary database to the primary database. We have to make sure to join the databases in the correct order.
- The databases to be joined must be in the same project folder in IDEA.

 IDEA provides five join options. We have to ensure that we choose the correct option. .

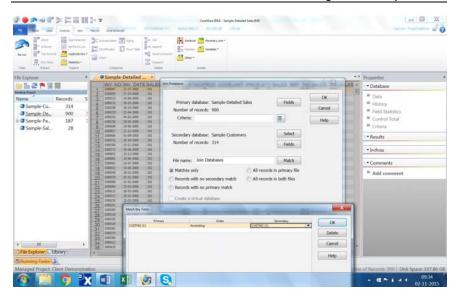
- 'Matches only' provides the first match between two files like citizens having a PAN number who have filed their income tax returns for the last financial year.
- 'Records with no secondary match' provide records which exist in the primary file but are absent/missing in the secondary file, like citizens having a PAN number but who have not filed their income tax returns for the last financial year.
- 'Records with no primary match' provide records which exist in the secondary file but are absent/missing in the primary file, like citizens who have filed their income tax returns but do not exist in the PAN database for the country.
- 'All records in primary file' provides a representation of all records from the primary file with due matches from the secondary file and even unmatched primary file records, like sales to customers with matching orders or without matching orders.
- 'All records in both files' provide a combined representation of all records from the primary file with and without matches and all records from the secondary file with or without primary matches. This step is different from Matches only in that the Matches only looks for the first match, whereas all records in both files follow the one-way match technique.

The Join Database takes each record in the primary database in turns and looks for a matching key in the secondary database. It does not support many-to-many relationships. To support many-to-many relationships, we have to use Visual Connector.

When we join two Character fields with different lengths, IDEA pads the shorter field contents with trailing spaces.

Typical view of the Join dialog box in IDEA is -

Working with Multiple Files



8.3 Append Databases

We use the Append Databases to append or concatenate IDEA databases into a single database for audit testing. We can append or concatenate up to 32,768 databases, in the order selected, into a single database.

For example, by using Append Databases, we can append 12 monthly payroll databases to produce a database of all payroll transactions for a year. We can then summarize the database by employee to ensure that every employee has a count of 12 payroll payments in the year.

It is to be noted that when using Append Databases, we cannot append or concatenate:

- Databases that contain one or more fields with the same name but different field types. So if the primary file has a field Transaction Number with character field type and one of the secondary files has the same field with numeric field type then the append would fail. Both the fields have to be of character or numeric field types.
- The same database more than once. So File A cannot be appended once again to File A.
- Databases from different locations: ensure all the files to be appended are in the same project folder.

Note

We have to retain each of the monthly archive databases and append them at the year end to produce a database of transactions for a year. Once databases have been appended or concatenated, we have to consider whether the individual databases are required, in addition to the appended or concatenated database. If they are not required, these databases can be deleted.

If we import the opening and closing balances databases into IDEA, we can reproduce the closing balances from the opening balances and transactions databases and compare them with the closing balances database. This test is the central theme of Financial Audits.

A line to the second	Las Deser Series des	File name: Append Databases Orberia:				OK Add
An Incide	Shelpitta -	Deales frant			Databases to append to active database:	Remove
	Carre	Name	Records	Size Modifi Created	KeyVal=02 KeyVal=03	Remove All
Explorer		Sample-Customers	314 900	51 31-10 31-10-201	KeyVal=04	Cancel
市ちょう		Sample-Detailed Sales KeyVal=01	900	135 02-11 30-10-201! 28 02-11 02-11-201!	KeyVal=05 KeyVal=06	Allow Property and
ne R	ecords		4			Help
Sample-Cu	314	KeyVal=02 KeyVal=03	39	18 02-11 02-11-201! 21 02-11 02-11-201!		
Sample-De_	314 900	KeyVal=04	39	19 02-11 02-11-201		
Sample-De_ KeyVal=_		KeyVal=04	834	83 02-11- 02-11-201		
KeyVal= KeyVal=	3	KeyVal=05	834	18 02-11 02-11-201		
KeyVal=	39	Bample-Payments	187	116 02-1131-10-201!		
KeyVal=_	17	Summarization	125	26 31-10 31-10-201		
KeyVal=	834	Sample-Sales Representativ	28	16 31-10 31-10-201		
KeyVal=_	3	Sample-Sales Representativ	20	16 31-10-201		
Sample-Pa_	187					
Sample-Sal_	28					
Surface Sec.						
		4				
			_			
111						
e Explorer Ubr		-				
	ary					
aming Tasks						

Typical view of the Append Databases dialog box in IDEA is -

Note

Append appears both under Analysis and Data in the IDEA menu tool-bar. To clear the use of the Functions Append under Analysis and Relate stands for Append Databases and Append under Data and Fields stands for Field Manipulation and Append.

8.4 Database Compare

We use the Database Compare to identify differences in a single Numeric field within two databases (referred to as the primary and secondary databases), for a specified common key.

The databases being compared need to be located within the same project folder.

Generally we compare the same database at two points in time, such as the Trial Balance at the beginning and end of the year to identify significant and material changes in each of the Heads of Accounts. We can also use it to compare a Numeric field from different systems (providing the key is the same), such as the quantity of inventory on the inventory master database versus the quantity of inventory on the stock take database (where the key is the item number) INVENTORY RECONCILIATION. Neither the key fields, nor the fields for comparison need to have the same names in the two databases, but the key fields must be of identical type. If necessary, we can change field types to allow the IDEA to match fields.

In statutory audits, File Compare can be used to view deviations in specific trial balance account heads between quarterly figures and annual unaudited figures. Where the deviations exceed 25%, the account heads under question can be taken up for detailed substantive testing. In fact, this procedure is at the center of the Audit Standard on Materiality.

As there may be more than one record for each key in either or both the databases, the Compare Databases task compares the totals for the Numeric field in each database (it essentially summarizes than compares the databases).

In addition to determining the differences in a Numeric field in a key, it also identifies the number of records there are in each database for each key.

If we need to compare more than one Numeric field, we can repeat the process for each field in turn.

Typical view of the Compare Databases dialog box in IDEA is -



Chapter 9

Sampling

9.0 Introduction

As a tool, IDEA has robust and stable capabilities to electronically audit 100% data in any application and database.

But CAATs do not replace traditional audit checks in terms of observation, inquiry, walk-through, substantive testing, and corroboration.

As an example – one can user IDEA to check manual adjustments to inventory balances for shortages and excesses. Armed with the list of such manual adjustments the auditor may want to check only a sample of such cases to ensure proper management authorization for the adjustments. This is where sampling comes into play.

Audit firms of international and national repute use IDEA primarily for sampling in their Financial Audits.

Sampling in IDEA is broadly statistical and probability-based.

The probability-based sampling techniques are:-

- Systematic
- Random
- Stratified Random

The statistical sampling techniques are:-

- Attribute
- Classical Variable
- Monetary Unit

9.1 Systematic Record Sampling

Systematic Record Sampling is a method to extract a number of records from a database at equal intervals to a separate database. It is often referred to as interval sampling.

There are two methods of determining the sample:

- Entering the number of records, in which case IDEA computes the interval size.
- Entering the selection interval, in which case IDEA computes the number of records.

IDEA calculates the above parameters on the number of records in the database and defaults to the first to last records. However, we can extract the sample from a range of records, if required.

Tip – Systematic sampling can be used where there is a need to cover samples evenly across the whole population, like in Payroll Audits. In Payroll activities the number of employees stays fairly uniform and even over a financial year. Here systematic sampling can be used for each quarter to cover different lots and intervals of Employees. Herein opinion can be expressed on a fairly representative lot of employees from the total employee base.

Typical view of the Systematic Sampling dialog box in IDEA is -

	tar brinning	- Det .		The state	1	lystematic Sample	10-m			×	n in the second s		
Renat B ^{ar} Tanknowl	and the second	· .	fur)		2	Number of Records	Selection Inter	val					
Talla Rotart	Explore		Callege		3	4	lection interval:	18		Fields			
ile Explorer		1	Sampl	le-Detailed			cords to select:			Compute		· Properties	
E 2 P 0			INV N	O INV DAT	E.S.	Starting record nu	mber to select:	1		Compose.	BALES TAX BALES	Database	
tation Printed	211 C		2200043 1355854	23-07-2008	12					Help	4313 37.34		
Name	Records	5 3	20000115	10-06-2008	10	Ending record nu	mber to select:	900		formers of second	724.39	H Data	
Sample-Cu_	314	4	35885.75 23800.94	38-05-3908	18	File name: Sys 5	-				349.75 260.57	History	
Sample-De_	900	. 2	1996.119	25-04-2508	12	rive native, says :	samp				2533	Field Statistics	
		1 2	3000254	94-81-2909	10						403.30	Control Total	
KeyVal=	3	6	1000448	25-05-208	1.148	Create a virtual da	tabase		OK	Cancel	140.25	Criteria	
- KeyVal=	- 4	10	1000417	23-13-200	82.		cote/cl in			and a sum provide statement	1.5	- Criteria	
KeyVal=	39	11	3200mint	11-09-2008	225	21107			534 254	(1875	140.71	Results	
KeyVal=	17	12	3350732	36-26-2008	385	23.395	-		5.00 83				
KeyVal=	834	14	2000714 2000772	15-12-208	103 121	25490	0		5.09 5.0	347.4		• Indices	
KeyVal=	3	15	3000012	33-43-3998	885.	25403	23		5.00 12				
		16	3300HEE	34.06-208	- M	2425	42		25 11			· Comments	
Sample-Pa_	187	1 17	10000003	34-07-2008	101	21426	65		15.15 N 15.15 LI			Comments	
Sample-Sal_	28	19	5000048	26-02-2908	152	25.462	25		5.00 872	1,822.2	982.23	# Add comment	
		20	2000079	45-43-2908	382	23464	45		5.00 1052				
		21	1000385	11-15-2938 25-03-2908	111	25466 . 25467	8		5.09 1200	7145			
		23	2000111	18-01-2908	342	25480	15		5.05 45				
		24	3000217	19-02-2908	382	25496			5.00 3128	12,000.0	1,196,39		
		25	1000,70	21-03-2008	142	21544	12		5.00 307	2,111.1			
		212234255728930	3000,013	13-01-2008	142	21830	10		5.89 1.0	745.7			
		28	33990.130	13-03-3308	342	25154	11		530 31	179.7	11.01		
		29	3300.520	13-05-2000	303 367	30236	11		5.00 30	176.7			
			200033M	34-05-2008	352	30505	10		5.09 294				
		22	2220173	15-82-2008	111	30704	15		5.04 21	187	12.58		
		33	3300384	13-06-2008	252	32,305	15		5.30 9	50.8			
III III		31 32 33 34 35	1000403	28-83-2008	周	32962	64		5.09 112	42,2%.0			
File Explorer	brary	4					11				+		
Running Tasks										_			

Sampling

The sector of th	Di terretario		ſ	ystematic Sample	-	-			×	1	
Renat 🔐 Tankanan 🍊 Dakate Key+ An Tankanan 🏠 Dakate Key+	(in 1997)		9	Number of Records	Selection Inte	rval					
Taka Futurt Eulere	(Dillip	64 C	1	Number of re	cords to select	: 45			Fields		
ile Explorer *	Sampl	e-Detailed		54	ection interval	20			Cimpute		Properties
1 2 P 0 1		O INV DATES	24	Starting record nur	wher to select	. 1			Compose	BALES TAX BALES -	• Database
extop Proat	2 1000040	23-03-2008							Help	4113 1774	I Data
Name Records S	3 1000115	10-06-2008	3	Ending record nur	nber to select	900				724.39	History
Sample-Cu 314	5 32000.00	18-01-2008 1	245	File name: Sys Sa	mp					280.57	 History Field Statistics
Sample-De_ 900 :	6 3990239 7 2090234	25-04-2000	2							21543	Control Total
KeyVal= 3	B 3300,04	29-05-208	1	Create a virtual dat	abase		OK		Cancel	140.75 4.75	Criteria
SkeyVal= 4	10 3000417	23-13-2508	а.		01570				Print Conde Condensity	1.5	- Criteria
KeyVal=39	11 20000000 1777 2000730	11-05-203	11	21107	11	_	539	0	217.11	148.75 31.34	/ = Results
KeyVal=17	11 2000000 12 2000722 13 2000704		10	25400	6		5.00	14	347.42	34.74	Sec. 1
KeyVal=	14 2200772		201	25.452	8		5.09	311	\$54.99	\$0.50	* Indices
KeyVal=3	15 2000et.2 16 2000et2		101. 102	240	21 42		5.00	9 11	341.43	34.34	
Sample-Pa 187 1	17 100000	34-07-2008	101	21425	61		25.25	11	2,480.50	246.25	- Comments
Sample-Sal_ 28	18 300000J 249 3000048		期	23490	10		2515	11	775.30 5.827.20	77.33	
a particular and	20 200010		142	25464	8		5.00	1052	6,311.41	63625	# Add comment
	25 3300089		111	25466	65		5.99	1298	7.148.95	714.80	
	22 3300940		国	25463 25488	10		5.00	10	308.25	30.00 20.75	
	24 3000217		112	25496			5.00	3118	ILMIAL	LINAN	
	25 1000/10		144	21544	41		5.88	307	2,318.13	.231.85	
	26 1000,00		加加	23546	8		5.88	10	1,186.62	210.00	
	28 3300100	13-03-2000	10	25154			5.00	1.00	179.79	11.07	
	29 1000.00		202	36226	13		5.89	38	178.70	17.87	
	30 30000M		352 312	30%5	25 15		5.00	200	1265.62	128.59	
	32 3300175	15-82-2008	\$82	30704	15		5.01	21	125.78	1238	
	33 2000384		252 352	11305	15		5.30	1.8	57.82 42.276.00	5.35	
< III >	18, 100010 19, 100048 20, 200048 22, 200048 23, 200048 23, 200048 24, 200048 25, 200048 25, 200048 26, 200048 27, 200048 28, 200048 28, 200048 30, 200034 31, 200048 33, 200038 34, 409401 35, 100419 35, 10		用口	32102	0		5/7	111	42,0%,00	4,227,00	
File Explorer Ubrary	1										
Running Tasks	terre a										

9.2 Random Record Sampling

Random Record Sampling is a commonly used method of sampling. With it we enter the sample size as well as the range of records from which the sample is to be extracted to a separate database. Then, using a random number seed, IDEA generates a list of random numbers and selects the appropriate records associated with these numbers.

Random sampling is normally applied in every regular audit project regardless of a preliminary look at the nature of the data being audited. It should practically be applied where the data is normally distributed, in the sense that there are no artificial limits within the data or major troughs or spikes.

Typical view of the Random Sampling dialog box in IDEA is -

Anna Anna Anna Anna Anna Anna Anna Anna	Las Detector Sector De Las Suplicate Rey * Suplicate Rey * Esplore			13 montan	3	nent nationale nationale	Monstary Unit *						
le Explorer		1	Sample	e-Detailed	-							- 1	Properties
	1		INV NO	DINY DAT	E SAL	ESREP NO CUSTN	OPROD COD	EUNIT: PR	CEIQT	ALES BEF TAX	ALES TAX SAL	ES ×	• Database
sitio Privat	2 T	2	3353654	12-01-2008	381	23296	8		3.09 6	8	31,74	用	III Data
	lecords S	3	30000115	10-06-2908 30-05-2908	直調			-	4.00 1.00	X	774.19 549.75		History
Sample-Cu	314	15	32000398	18-03-3008	315	and the second second	-		10. 10		280.57		 History Field Statistics
Sample-De_	900 ;	6	3996239	25-04-2008	两日		records to select:	-			215.80		
KeyVal=	3	8	3335(254	25-05-2004	1200	Humber of				OK	148.75		Control Total
KeyVal=	4	9	3000448 3000817	23-13-2908	11	Rando	om number seed:	16369		Fields	4.25		Criteria
KeyVal=	39	11	3200mint	111-099-2008	111				141	Canori	140.75		- Results
KeyVal=	17	12	3000732 3000764	38-29-2008 15-12-2008	胡田		number to select:	1	A. .7.	Lances	37.34 34.34		
KeyVal=	834	14	3900772	20-08-2008	371	Ending record r	number to select:	900	0	Help	\$0.50		• • Indices
KeyVal=	3	15	3000012	13-13-2008 34-06-2008	10	Allow d	uplicate records:	-			34.34		
Sample-Pa_	187 1	17	1000000	34-07-2008	352	10000					246.01		Comments
Sample-Sal_	28	18	3000002	28-02-2008	10	Create a v	(rtuis) database:				77,33		# Add comment
			22000175	45-43-2908	342		d Samo				630.15		Aug comment
		2011213455678890	3000385 3000940	11-15-2908 25-03-2008	田田				-		714.00		
		23	1000111	18-01-2908	342	21400	10	1		2011	28.75		
		24	3000217	19-62-2008 21-65-2008	載	25496 25544	e e		5.00 311 5.00 30		1284.00		
		26	1000252	15-05-2008	312	23546	8		5.85 2.9	8 1,186.02	112.86		
		27	3000,001	13-03-2008	はは	21850	11		5.89 12 5.89 3		74,88		
		29	3000.020	13-05-2008	303	36226	11		5.80 3	0 178.70	£7.87		
		30	3000334 3005.M.2	14-05-200	線加	30%S	15		5.00 20		124.59		
		31	2000323	11-12-200	312	30704	15		509 2		12.58		
		33	2000384	13-06-2908	111	31305	15		5.30	8 50.8L	5.36		
		22 33 34 35	2000413	28-83-2008	周辺	33902 32107			5.09 11		4,227,80		
File Explorer Ub	rary	30					11 III					+	
Running Tasks													
lanaged Project: Clie	1.0	0.L							-	nected to IDEA Serve	1 14	Contract of	of Records: 900 Disk Space: 107.8

9.3 Stratified Random Record Sampling

Stratified Random Record Sampling is based on the materiality of transactions. Here due emphasis is given to the 80-20 rule in business, wherein 80% of value resides in 20% quantum and 20% value in 80% quantum. The 80% value representing the Class A is taken up for detailed scrutiny. The 20% value representing the Class C is taken up for limited random selection and scrutiny based on either rule of thumb or existing techniques based on business and data conditions.

Stratified Random Record Sampling may be used to extract a random sample with a specified number of records from each of a series of bands. This task requires then the database be first stratified into a series of numeric, character or date bands. IDEA then presents us with a table displaying the number of records within each band. We then enter the number of sample records we want IDEA to extract at random from each band.

IDEA provides a Stratification Assistant to guide us through the steps for stratifying and generating the sample.

It is not necessary to stratify the file first, although if this has been done and a stratified database used, then the Assistant skips the initial stages and goes directly to the sample selection stage. To ensure that we extract at least one transaction of each type in a random sample, we have to use Stratified Random Sampling: to stratify the database by the type field and then select at least one record from each band.

Stratified Random Sampling can be used in General Ledger Scrutiny as different values and risks are associated with different ledger items.

0 • • • • * × = = =		Carolifying EEA - Sample-Datailed Sales.3MD		
the new Data Andres New	Narie Bratthauter	and but got a crist		0
Arrent Fater Fater Fater	Conserver Conser	Interface Revelop Line Interface Interface Interface Interface Interface Interface		
File Explorer	Sample-Detailed *			Properties
ine coporer	INV NO INV DATE SALESRER	Stratification Accesses	BEF TAX SALES TAX BALES +	Database
Destas Pased	1 200040 23-45-2008 1m	The current database has not been stratified.	411.38 4333 317.37 317.34	
Name Records 5	2 3300054 13-61-3008 10 3 300011 10-06-300 201 4 3300111 20-65-3008 101		7.241.01 724.39 1.467.50 549.75	Data
Sample-Cu 314	5 2200100 18-03-3000 201	In order to perform a stratified random sample the database must first be stratified.	2,405,85 280,57	History Field Statistics
Sample-De 900	6 1806213 25-04-2008 281 7 1000254 04-85-2008 281		2,198,33 225,83 4,181,30 4(38,30	Control Total
KeyVal=3	B 2000/298 25-055-2008 201 9 2000448 15-06-2008 201	Please select the type of stratification you would like to perform.	LAI7.36 148.25 17.52 4.25	Criteria
-3 KeyVal= 4 -3 KeyVal= 39	10 3303817 22-13-2008 888		1.008.27 195.43	
KeyVal=		Perform a numeric stratification	L 607.50 140.75 377.37 37.74	 Results
KeyVal=	12 2300722 36-89-2008 248 12 2500764 15-12-2008 249 14 2500772 36-66-2008 249	Perform a character stratification	347.42 34.24 604.09 80.50	• Indices
KeyVal=3	15 mm 10-10-00 mm	Perform a date stratification	341.43 34.24 317,75 113,74	
B Sample-Pa_ 187	17 100000 34-07-2000 103		2,41150 240,81	- Comments
Sample-Sal 28	18 1990101 19-96-2009 181 19 200048 26-83-2008 182 20 200019 85-63-2008 182	Once the task has been completed, the assistant will take you to the Stratified Random Sample dialog.	773,30 773,31 3,822,38 582,23 4,382,48 636,25	# Add comment
	$\begin{array}{cccccccccccccccccccccccccccccccccccc$		7.148.00 724.00 309.25 30.42	
	23 200031 38-05-2008 542 24 2000217 19-02-2008 183		2033 23,05 W.MEL 26,06	
	25 1000,00 21-05-2008 102	Cancel Continue Help	2,319.32 231.81	
	26 1000252 18-85-2008 142 27 2000,051 12-85-2008 142		748.75 74.88	
	28 300030 13-05-3000 143 749 100030 23-05-3000 333	2014 EL 5.00 EL 2018 EL 5.00 EL	179.79 (21.97 179.70 (21.97	
	30 1000004 34-05-2008 352	3050 25 5.00 200	1,245.62 124.59	
	3.1 2005MJ 08-04-2008 283 3.2 2000325 05-02-2008 883	Intos 05 5.09 25 30704 85 5.08 21	148.75 14.98 125.79 12.58	
	33 2000384 03-06-2008 282 34 2000413 28-03-2008 282	22325 US 5.70 S 32902 04 225-89 480	53.81 5.39 42.2%.00 4.227.88	
< III >	35 1300420 13-02-3006 342	32103 01 5.09 113	6%.07 67.08 ····	
C File Explorer (Dbrary	1	.III.		
BRunning Tasks		7.2017 (10.10)	The second second	
Managed Project: Client Demonstra	and the second se	Not connected b	b IDEA Server Number	of Records: 900 Disk Space: 107.83 GB
l 🚳 📋 🧿 🏅	K 🛯 🗶 💱 😪			• ■ ► i ◄ • 09.52 02-11-2015
				02-11-2013
9 • • • * * * * * * *	17	Canality and EKA - Sample-Detailed Sales IME		
And a state of the				
The Marke Data Analysis Des	Narie Breitheuter			
in Dest in Gebetter		Attribute 🚰 Marelary Unit *		
P Island Sinthetran	Staticator	E		
Renat Parlament Calculate	121 d			
Tatle Factor Patter		Increment: 15,000.00	OK	
		Value of Street Street Street	Cut off	
File Explorer	Field to stratify:	** Lower Limit * Up	per Lient	Properties
	UNIT_PRICE	- 2 15:000.00 M:500.00	Long - Annalytic -	• • Database
Name Records S	SALES_BEF_TAX	3	Insert	🖬 Data
Sample-Cu 314	SALES TAX	3	Remove	History
Sample-De_ 900	SALES PLUS_TAX	1	Cancel	Field Statistics
KeyVal=3	7.8	8	And a second second	Control Total
KorVal=4	9	10	Help	H Criteria

Typical view of the Random Sampling dialog box in IDEA is -

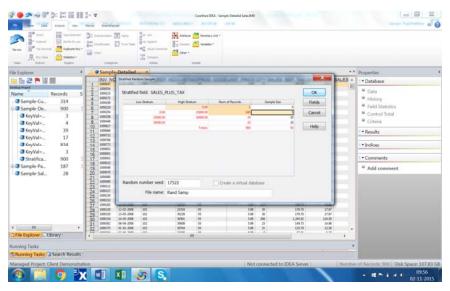
 Image: Participation
 Description: Under Standing

 Image: Participation:
 Description: Under Standing

 Image: Participation:
 Description: Under Standing

 Image: Participation:
 Description:

 Image: Participation:
 De



9.4 Attribute Sampling

Attribute Sampling refers to the examination of a subset of a population (a sample) in order to assess how frequently a particular event or attribute occurs in the population as a whole. An attribute has only two possible values: true or false. In auditing, typical attributes are whether an item is in error or not, whether a particular control has been exercised or not, or whether the entity complied with a particular law or not.

The Attribute Planning and Evaluation task is meant to plan attribute samples and evaluate results. The Planning dialog boxes are used to determine the minimum sample size required to meet specified audit objectives, as well as the critical or maximum allowable number of sample deviations in the sample.

The two Planning options, Beta Risk Control and Beta and Alpha Risk Control, allow us to (control or cover) the risk of one or two types of incorrect decisions. Auditors largely set only the risk of relying on controls when they are not effective (Beta Risk) and the tolerable deviation rate. To use this approach, we select the Planning (Beta Risk Control) tab in the Attribute Sampling dialog box.

Sampling

Port Hardener De	Athibute Langting	N	-			×==
2" Island 52 Saladium Date	Planning (Beta Risk Control) Planni	ing (Beta and Alpha Risk (Control) Sample Evaluatio	on		
E Talance Cakan fay*						
Tetter Tolar Color	Population size: 90	00	5	Expected deviation rate:	10.00	
eer #	% Tolerable deviation rate: 15	5.00	Confidence lev	vel (to control Beta Risk)	85.00	Properties
2 🖻 🖉 🖩 🖉						- Database
e monte a particular and a la	Sample stre : 70		Critical number of devia	ations in sample: 7		Data
Records S A A A A A A A A A A A A A A A A A A	Devators N.Deviators		Achieved Confidence (Bet			History
nple-De_ 900 : 5	1 1.43					Field Statistics
eyVal=3	2 2.00				98.0	a Control Total
eyVal= 3 8 eyVal= 4 9 eyVal= 39 11	3 43				99.0	Criteria
eyVal=39 11 eyVal=17 12	4 5.72				94	
eyval=1/ 12 eyVal=834 14	3 1.14				96.	Indices
leyVal=	4 4.57				92.0	
tratifica	3 1100				85.1	Continuents
аучин 17 аучин 814 аучин 814 аучин 814 аучин 900 1 5 аран 900 1 5 аран 900 1 5 аран 900 1 6 аран 900 1 аран 900 1 6 аран	8 11.4				751	- Add comment
20 20 110 110 100	8 12.86 30 14.29				41	
23	11 157					
25						
27	Conclusion: If no more to that the pop	han 7 deviations are observation deviation rate is no	rved in a sample of size 7 at more than 15,00%.	0, you can be at least 85	.00% confident	
29						
30						
22						
III , 34 35			Print	Close Comp	de Help	
	VI XI 🐼 S.	Central Ed	Not.com	nnected to IDEA Serve		Aumber of Records: 900 Disk Span
In Table () Project: Clerit Demonstration () () () () () () () () () ()	VIII VIII S	Contra Ed		nnected to IDEA Serve		- #Fi44 0
	Statistica	N	1 - Sample Devolut Sales 2011			- #>i44 o
	Institution Anticula Sanging [Planning (Beta Risk Control)] Planni	ing (Beta and Alpha Risk (l- Sample-Densited Sales, Mill Contrype) Sample Evaluatio	on		- #>i44 o
Project: Clerit Demonstration Project: Clerit Demonstration	Annual Sensing (Planning (Beta Risk Control) Planni Population size: 90	ing (Beta and Alpha Risk 6	A - Sample-Densind Sales, MAR Constrat) Sample Evaluatio Number of deviation	on		- #>i44 o
	Institution Anticula Sanging [Planning (Beta Risk Control)] Planni	ing (Beta and Alpha Risk 6	A - Sample-Densind Sales, MAR Constrat) Sample Evaluatio Number of deviation	on		- #>i44 o
Protection (Centre Demonstration) Project Centre Demonstration Project Centre Demonstration P	Annual Sensing (Planning (Beta Risk Control) Planni Population size: 90	ing (Beta and Alpha Risk 6	A - Sample-Densind Sales, MAR Constrat) Sample Evaluatio Number of deviation	on		- 85144
Professional Constructions of the second sec	Institution Anthone Senging (Planning (Deta Risk Control)) Planni Population size: 90 Sample size: 70	ing (Beta and Alpha Risk (00	A - Sample-Densind Sales, MAR Constrat) Sample Evaluatio Number of deviation	on		- E - e - o
Professional Constructions of the second sec	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 6	L. Sample-Detailed Sales, MIC Control() Sample Evaluatio Number of deviation % Desired con	on Ins is sample: 10 fidence level: 85.00		Properties Properties Postabase Postabase W Data History
Professional Constructions of the second sec	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tig	per Land	Properties · Data · Hotory · Fad Statistics
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	L. Sample-Detected Sales, MI Control) Sample Evaluatio Number of deviation % Desired con	on Ins is sample: 10 fidence level: 85.00		Properties
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tig	per Land	Properties Properties Postabase
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tig	per Land	Properties
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tig	per Land	Properties • Outabase • Data • Field • Fi
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tip	per Land	Properties • Outabase • Outabase • Outabase • Data • Heatory • Fried Statistics • Control Total • Control Total • Results • Results • Indices
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tip	per Land	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tip	per Land	Properties • Outabase • Outabase • Outabase • Data • Heatory • Fried Statistics • Control Total • Control Total • Results • Results • Indices
Project Clear Demonstration Project Clear Demonstrati	Interfactories Pleaning (Beta Risk Control) Pleani Population size: 90 Sample size: 70 Sample deviation rate: 14.29	ing (Beta and Alpha Risk 0 00 0 % (10 / 70)	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on in sample: 10 indexce level: 55.00 ded Lawer (unit 2 State tip	per Lent	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	A - Sangle-Dasket See 200 Control) Sample Evoluatio Number of deviation % Desired con	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Project Clear Demonstration Project Clear Demonstrati	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pr Los	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control
Specifie	An divide Services Peopulation state: 94 Sample deviation rate: 14,29 L 56	ing (Beta and Alpha Rok (00 0 75 (10 / 70) and logar Lond	Longie-Dinisi Sex.M Control) Serrufie Evoluatio Number of deviation % Desired con 14.0 2.56	on	pp Link A.D Cat	Properties Data History Field Statistics Control Total Control Total Results Hereins Control Total Control Total Control Total Control Total Control Total Control Total Control Total Control

Example - An internal control requires that all checks over \$5,000 be signed by a manager. Whether the check associated with a transaction has been signed by a manager is an example of an attribute of the transaction, it has a value of true or false. Attribute Sampling can be used to determine whether the control is operating properly, to some specified level of confidence or reliability, without examining every \$5,000 check in the population. The population in this example consists of all checks over \$5,000. A subset of this population, called the sample, is selected for examination. Based on the number of control deviations (unsigned checks) found in the sample,

inferences about the number of control deviations in the population as a whole can be made.

One of IDEA's two Sample Planning tasks can be used to determine the minimum sample size necessary to infer that the percentage of control deviations in the population is no more than a specified tolerable deviation rate, while controlling either the level of Beta Risk or the levels of both Beta and Alpha Risks. Roughly speaking, Beta Risk is the risk of incorrect acceptance, while Alpha Risk is the risk of incorrect rejection. The most crucial of these in audit situations is Beta Risk.

Here is an example of planning to control Beta Risk, since this is the most common situation. Suppose the auditor wishes to calculate the sample size using the following parameters:

- Population size (number of checks over \$5,000): 10,000
- Expected deviation rate (expected percentage of checks over \$5,000 that are not signed by a manager): 2%
- Confidence level (to control Beta Risk): 95%
- Tolerable deviation rate (maximum allowable percentage of checks over \$5,000 that may be not signed by a manager while still allowing reasonable reliance on this control): 5%

Using these parameters, the required sample size is 180, with a critical number of sample errors equal to 4.

Planning to control both Beta and Alpha Risks would proceed in the same way, except that a second confidence level, equal to 100% - the maximum allowable Alpha Risk, would be specified as well. For example, with confidence to control Alpha Risk = 95%, lower threshold deviation rate = 2%, and all other parameters as above, the required sample size is 382, and the critical number of deviations is 12.

Suppose only Beta Risk is controlled, so the sample size of 180 is used. A sample of 180 items is selected from a population of checks over \$5,000 (such as may be obtained by performing a Direct Extraction using Random Record Sampling, and the true number of errors or deviations in the sample is determined through audit procedures. Next, Attribute Sample Evaluation is used to determine the upper limit on the number of population deviations to a specified confidence level. Suppose the true number of deviations is equal to 4. Then, in the Sample Evaluation dialog box, enter:

- Population size: 10,000
- Sample size: 180
- Number of deviations in sample: 4
- Confidence level: 95%

The one-sided upper limit on the population deviation rate is calculated to be 4.98, which is just under the tolerable deviation rate specified when the sample was planned. This indicates that the sample size was efficient. If, on the other hand, five errors are found, the upper limit is 5.72, which exceeds the tolerable deviation rate. If only three errors are found, the upper limit is 4.22. IDEA also calculates a two-sided interval, delineated by a two-sided upper limit and a two-sided lower limit.

IDEA's Attribute Sampling programs are used to plan for and evaluate samples obtained by Random Record Sampling. If the sample was obtained by Stratified Random Sampling, then we do not use this evaluation program; instead, we seek the guidance of a statistician.

"A detailed Technical Specification on Attribute Sampling – CaseWare IDEA Research Department Document of release date June 17, 2003 has been reproduced in Appendix III for the benefit of the academically inclined readers. This document describes the mathematical procedures followed by IDEA in planning and evaluating Attribute samples."

9.5 Monetary Unit Sampling

A numeric amount field is the basis of the sample in Monetary Unit Sampling, in which the monetary unit is sampled. Having chosen the monetary units to test on an interval (or random item within each interval) basis, the items which contain these units are identified. To do this, it is necessary to calculate the cumulative value of the relevant field for each item.

Record	Ref	Amount	CUMULATIVE
1	A123	50	50
2	A124	75	125
3	A125	40	165

With an interval of 100 starting at a random start point of 35, the first selection would be at 35, and the second at 135. It is the cumulative field that is used for the selection.

Any items greater than the sampling interval are to be chosen by the Auditor. However, these key or high value items have to be extracted from the population and subjected to separate testing, leaving the items below a specified high value amount (which can, if desired, be set to a different limit to the interval) to be evaluated statistically.

IDEA offers options for determining the sample size for both compliance and substantive tests, based on confidence levels, population size, materiality and certain options on the weighting statistics used.

Whether or not the planning option is used, samples can be selected using the Monetary Unit Sampling techniques.

The output file contains an Editable field called AUDIT_AMT. If desired, this can be updated from the results of audit testing by using the Field Manipulation task and then a detailed evaluation of the results can be performed against book value.

The Monetary Unit Sampling technique is used extensively by the Comptroller and Auditor General of India.

Monetary Unit Planning - Examples

You wish to sample from the salary data contained in the Sales Representatives file. You wish to sample from both positive values and negative value. You have the following information for your planning:

- Confidence level 90 percent.
- Tolerable error \$200,000. In auditing, this is often known as the materiality amount.
- Expected error in the population \$40,000.

Open the dialog box for Monetary Unit Sampling - Plan. Enter the information as indicated below. IDEA will compute the sampling interval and sample size in the dialog box. If you are ready to select the items, you may do so directly from the dialog box using the Accept button.

Note that the sample size is only an approximate value since the effect of high value items on the sample size is not known.

- Use value from database field: SALARY
- Absolute values selected
- Confidence level : 90 percent

- Tolerable error: 200,000
- Expected error: 40,000
- No change to basic precision

Estimate

- Approximate sample size : 15
- Sampling Interval : 64,733.33
- Sum of tolerable sample taintings : 61.98 percent

You must accept the population at the 90 percent confidence level when no more than 0.6198 total taintings are observed in a sample size of 15.

"A detailed Technical Specification on Monetary Unit Sampling – CaseWare IDEA Research Department Document of release date May 7, 2003 has been reproduced in Appendix IV for the benefit of the academically inclined readers. This document describes the mathematical procedures followed by IDEA in planning and Monetary Unit samples."

Monetary Unit Sampling is most effective in the following areas -

- Accounts Receivable confirmation testing for the existence of accounts receivable.
- Accounts Receivable testing for unrecorded receivables by sampling on cash receipts subsequent to year end.
- Allowance for doubtful accounts testing for the valuation of accounts receivable when the allowance is calculated item by item.
- Inventory costing where overstatement of inventory is a concern.
- Loans or mortgages receivable confirmation of new issues.
- Loans and mortgages receivable confirmation of existing loans. In this case, sampling efficiency may be obtained by an appropriate choice of basic precision pricing and the application of the small tainting assumption.
- Accounts Payable tests of existence. Sampling from disbursements subsequent to the period's end.
- Commission income where income is based on certain types of revenues such as revenues from insurance premiums. In this case,

sampling efficiency may be obtained by choosing the basis precision pricing to reflect he maximum commission rate.

- Payroll expenses testing for the valuation and existence of recorded payroll transactions.
- Other expenses testing for the valuation and existence of other expenses.

9.6 Classical Variables Sampling

We use Classical Variables Sampling to provide estimates of quantities that are not recorded in the main database (also known as the population), by projecting results obtained for a sample drawn from the population.

A common use of Classical Variables Sampling is to estimate the total error and the total corrected, or audited, amount in a population, using an audited sample from the population

Classical Variables Sampling is also useful for making accounting estimates. For example, we may estimate how many rupees of inventory are represented by obsolete inventory, based on a sample for which the portion of each entry that is obsolete has been determined from manual records.

The sampling unit used in Classical Variables Sampling is a single record.

The sample database contains the key field from the population database, that is, the Numeric field whose entries are being checked by the audit, as well as an Editable field in which corrected amounts, determined through audit, can be entered.

Classical Variables Sampling techniques for projecting errors are most useful for sample auditing of accounts in which there may be errors or misstatements. If we wish to determine limits on the error in the population though we expect no errors or very few errors in the population, and if the principal risk is overstatement, we should consider using IDEA's Monetary Unit Sampling instead of Classical Variables Sampling. If we do not wish to project errors between the total book value and audited amounts, and require only limits on the total audited amount based on a sample of audited amounts, the mean estimator supplied by Classical Variables Sampling may be a good choice. The mean estimator can be computed for samples having no errors or very few errors.

Classical Variables Sampling is performed in three steps: sample preparation, sample audit, and sample evaluation. In sample preparation, the population database is stratified, the required sample size is determined based on our audit objectives, and the sample is extracted. In auditing the sample, we must update the field containing audited amounts with correct values determined through the audit procedure. Finally, the audited sample is evaluated to obtain projections of the most likely total audited amount and error in the population, and precision limits on this quantity. IDEA automates the sample preparation and sample evaluation phases of the Classical Variables Sampling process.

Classical Variables Sampling may use stratified estimates, in which case the population and sample databases must be stratified, or un-stratified estimates, in which case they are not. Generally, stratification reduces the sample size needed to achieve a desired result, but requires a larger number of errors in the sample. Various estimation techniques exist in Classical Variables Sampling theory. The most efficient technique generally depends upon the characteristics of the population database. IDEA supports six Classical Variables Sampling estimation techniques:

- Mean
- Difference
- Combined Ratio
- Separate Ratio
- Combined Regression
- Separate Regression.

Mean estimation is the least efficient of all and generally not used. except for making comparisons, or estimating quantities for which no recorded values exist prior to the audit.

Chapter 10 Field Manipulation

10.0 Introduction

Field Manipulation can be used within IDEA to view field definitions, add or delete fields, change field properties such as field name and type, or add tags (field templates) for the Smart Analyzer Financial component within IDEA.

Field Manipulation within IDEA can be performed through Virtual Fields and Editable Fields.

- Virtual Fields, as the name suggests, are computed fields through specific user based equations and/or functions. For example, adding a virtual numeric field using the @Function, @justnumbers will remove all numbers from an alpha-numeric narration within IDEA. This new virtual numeric field will contain only numbers like Cheque Numbers, for instance.
- Virtual Fields are generally used when we need to do a recalculation, re-computation, or field –conversion.
- Editable Fields, as the name suggests, are fields that are updated by user entry. For example, creating an editable field for recording the inventory count numbers during a physical inventory verification.
- Editable Fields are generally used when a user-based instruction or entry is to be made within the database. This is the only occasion where a user can add values into the database which cannot otherwise be altered in any manner.

10.1 Editable Fields

Editable fields are modifiable fields that can be added through:

- Field Manipulation
- Direct Extraction
- Import Assistant (with Delimited, Fixed length, EBCIDIC, and dBASE files only)

Editable fields are valuable for:

- Entering comments/data
- Amending fields, such as when evaluating monetary unit samples
- Performing new calculations and ratios from fields in the database

To help us distinguish Editable fields from protected fields, by default, IDEA displays editable data in the blue font. One can change the display colour through Grid Settings.

By default, IDEA does not record the definitions of Editable fields and any entries/changes we make to the data in the History. To record these, we need to clear the Do not record field changes made through editable fields check box in the System options.

If we make any changes to the data in an Editable field, any index that is based on the Editable field becomes invalid. To validate the index, we have to re-index the database.

Editable Fields: Examples

Example 1: Add an Editable field to a database for comments

Select Data > Extractions > Direct Extraction....

In the first row of the grid, specify a file name for the resultant database.

Click Create Fields and append an Editable field with the following definition:

Name	COMMENTS
Туре	Character
Length	100
Parameter	

Click OK on the Create Fields dialog box.

Click OK on the Extract to File(s) dialog box. This creates a new database with an additional empty field, COMMENTS.

Select the field and enter appropriate comments. The **History** records any possible (?)changes.

Example 2: Add an Editable field to a database for recording an audited value

Select Data > Extractions > Direct Extraction....

In the first row of the grid, specify a file name for the resultant database.

Click Create Fields and append an Editable field with the following definition:

Name	AUDIT_VALUE
Туре	Numeric
Dec	2
Parameter	VALUE (the field name of the items being audited)

Click OK on the Create Fields dialog box.

Click OK on the Extract to File(s) dialog box. This creates a new database with an additional field, AUDIT_VALUE. This field is the same as the field VALUE but it is editable.

When auditing the items, if a value is different to VALUE, change the entry in AUDIT_VALUE.

Extract where VALUE <> AUDIT_VALUE. If testing a sample of items, extrapolate the results for the whole population, if required. The History records any and all changes.

ð 🖡 🛔	1?					888	ľ								
Allief Die	Sarted Order			-		ne Palad Tig Second	ħ		_						
e Explorer En 🍣 🗭 🕅 🕅		-		DINV DATE	- Ĉ.	pend Field	=	States and South	-	-	and Section	and Person	i	riopense	
Atta Print	-	1	3200042	23-03-2008	1	and rest		10000		1000	-	and the second		· Datab	0.02
	cords	5 3	1000113	12-01-2008		Field name		OMMENTS				OK		🗸 Data	
Sample-Cu	314	4	3500571	30-05-2008	1									 Histo 	ry.
Sample-De_	900	. 6	2000UW 2000UW	18-03-2008 25-04-2008		Field typ	e:	Rable Character			•	Cance		Field	Statistics
KeyVal=	3	7	2000254 2000254	94-85-200	1	Lengt						Help		Cont	rol Total
KeyVal=	4	9	3000448	15-08-2008	18	Lengu	n.					Those	- 1	Criter	ia
KeyVal=_	39	10	1000H17 1000H11	23-13-3908	а.	Number of decimal	ls;							- Read	2
KeyVal=	17	12	2300732 2000754	26-26-2008	18						100				-
KeyVal=	834	14	3000772	30-08-2008	a.	Paramete	82							Strati	fication
KeyVal=	3	15	3000012 3000002	31-13-2008	а.	Description	e:						_ 1	· Indice	5
Stratifica	900	1 17	1000000	34-07-2008	18								_ 1	✓ No ir	
Sample-Pa_	187	1 18	3000012	19-16-2008 26-02-2008	117	34(0	-	5.0		W21	3,812,08	54727	_		NO/A + PROD_CODE/A
Sample-Sal_	28	022022202220222222222222222222222222222	2000079	15-43-2908	112	23464 05		5.00		1952	8,315.48 7148.00	60525 716.80			SREP_NO/A +
		22	2000040	25-03-200	111	2140 0		5.0		85	308.25	MAG			S PLUS TAX/D
		23	3000833	38-03-2908	342	23480 18		5,0		45	293.51	21.75		CUST	
		24	2000217	23-42-298	10	25496 05		5.0		367	22,008,02	1286.00			
		20	1000252	15-85-208	342	23546 19		5.0		104	1,186.02	118.86			D_CODE/A
		27	3000,045	13-01-2008	182	21850 85		5.0		1.0	748.75	74,88		SALE	S_PLUS_TAX/A
		28	1000130	13-63-2008	342	25154 01 36226 01		5.8		10	179.79	1110			
		30	1000334	34-05-2008	m	101 25		5.00		24	126.62	124.59		· • Comm	ients.
		31	2000.MJ	38-04-2008	382	300.01 01		5.0		25	148.75	34,98		Add.	comment
		22	2000375	15-42-208	111	30704 15		5.0		21	125.79	12.58			
		, 33	1000413	25-03-2008	141	32902 94		325.49		405	42,276.00	4,221,80			
		35	3300425	13-62-2006	342	3210) en		520		113	676.67	\$7.89			
File Explorer Clubra	ary	1 1 10				III.					No. of Street				
Running Tasks															

10.2 Virtual Fields

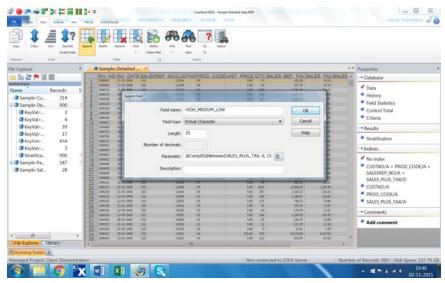
Appended (Virtual) fields are calculated fields that can be added and used for:

- Proving existing calculations in a database.
- Performing new calculations and ratios from fields within the database.

Converting data from one type into another, such as converting cheque numbers stored in a Character field into a Numeric field prior to testing for gaps in the cheque number sequence.

As virtual fields are calculated fields, IDEA computes the results "on the fly" and the results are not physically stored in the database. However, any files we create from a database with a Virtual field then contains the virtual field as a non- calculated value.

Virtual fields may also be added through Field Manipulation or the Import Assistant. If we want to add an empty field for entering comments/data, we need to add an Editable field.



Virtual Field: Examples

A virtual field can be used to

Convert a cheque number field (CHQNO) stored as a character field into a numeric field for testing for gaps by adding a Virtual Numeric field with the expression: @VAL(CHQNO)

Convert a date field (TRANS_DATE) stored as a character field with the format DD-MMM-YYYY (03-NOV-2015) into a date field by adding a date virtual field with the expression: @CTOD(TRANS_DATE, "DD-MMM-YYYY"). The date will then be displayed in the IDEA date format YY/MM/DD.

Caution:

Making changes to the database through Field Manipulation may show anything based on that field (results, drill-downs, indices, views, etc) as incorrect or invalid. Results can be made valid again by returning the settings to what they were when the result was created.

Field Manipulation	and Commonly	/ used	@Functions	in
IDEA Software with	Examples			

S No	@Function	Purpose	Example
1	@age	For creating a new virtual numeric field with difference in days.	To arrive at the difference in days between 'Token Date' and 'Cheque Date' in ANY BILL PAY FILE use @age(Cheque Date, Token Date) through Field Manipulation.
2	@dow	Identify transactions made on any day of the week.	To capture cheques issued on Sunday use @dow(Cheque Date)=1 through a Direct Extraction. (1 is Sunday in IDEA)
3	@isblank	Filter out blank character fields	To present blank scheme agency name in GRANTS-IN- AID data use @isblank(Scheme Name) through Criteria.
4	@isini	Capture specific character strings from a character or alpha-numeric field where the <u>strings are not</u> <u>case sensitive</u> .	To display cheque in favour of field containing 'Shri' for cheques issued inter-se Government use @isini("shri", cheque in favour) through Direct Extraction.

Field Manipulation

S	@Function	Purpose	Example
No		rupose	Lvampie
5	@compif	Conditional function which allows for preparing an expression with multiple conditions and consequent results. Useful for 'If then analysis'	In the Disaster Claim file, to create a new virtual numeric field titled Sanctioned Amount using @compif(Category="0400", 100000, Category = "0104", 100000, "0103", 500000, 1, 0) through field manipulation. This means if the category is 0400 the result in the new virtual numeric field will be Rs. 1000000 and so on so forth.
6	@strip	Removes all special characters, spaces and punctuation marks from a character or alpha-numeric field.	To check if the same sanction number has been used more than once in the same controller, DDO, PAO, Function Head and Scheme by inserting special characters or spaces to bypass inbuilt system checks. Append a virtual character field with criteria @strip(Sanction Number). Then run a duplicate key detection test on the stripped sanction number field.
7	@left	Removes the left most occurrences of a character field or alpha- numeric field.	In the Disaster Claim file, every case claim is administered a Unique Case ID of 12 digits. The first 4 digits represent the category of claim i.e. death (0400) etc. can be captured separately by creating a virtual character field using @left(Unique Case ID, 4) through Field

S No	@Function	Purpose	Example
			Manipulation.
8	@right	Removes the right most occurrences of a character field or alpha- numeric field.	In the Disaster Claim file, every case claim is administered a Unique Case ID of 12 digits. The last 6 digits represent the claim serial number which can be captured separately by creating a virtual character field using @right(Unique Case ID, 6) through Field Manipulation.
9	@mid	Removes the central most occurrences of a character field or alpha-numeric field.	In the Disaster Claim, every case claim is administered a Unique Case ID of 12 digits. The 5 th and 6 th digit together represent the case tribunal number which can be captured separately by creating a virtual character field using @mid(Unique Case ID, 5,2) through Field Manipulation.
10	@year	Derives the year from any date field in IDEA	In Tax collection data via APPLICATION SYSTEM containing nodal scroll date, a new virtual numeric field can be created using @year(nodal scroll date) to derive the year of collection through Field Manipulation.
11	@month	Derives the month from any date field in IDEA	In Tax collection data via APPLICATION SYSTEM containing nodal scroll date, a new virtual numeric field can be created using

Field Manipulation

S No	@Function	Purpose	Example
			@month(nodal scroll date) to derive the month of collection through Field Manipulation.
12	@day	Derives the day from any date field in IDEA	In Tax collection data via APPLICATION SYSTEM containing nodal scroll date, a new virtual numeric field can be created using @day(nodal scroll date) to derive the day of collection through Field Manipulation.
13	@DtoDays	Computes the days corresponding to any date field in IDEA	In Gradation List to compute the days for the retirement date add a virtual numeric field with criteria @DtoDays(Retirement Date) through Field Manipulation
14	@DaystoD	Reverse compute Date from days.	In Gradation List to compute the date from the days add a virtual date field with criteria @DaystoD (Retirement Date Days+30) through Field Manipulation. The +30 is to suggest adding 30 grace-days as an example.
15	@match	Looks for multiple values in a field of choice through one single equation without the need of having to write multiple equations.	In Pension payments file to filter out basic pension equal to Rs. 3500 or Rs. 40000 (i.e. outliers) @match(Basic, 3500, 40000) can be run through a Direct Extraction. This avoids the need to run two equations like (Basic=3500) .OR. (Basic=40000)

Chapter 11

Importing various Data File Formats

IDEA and Data Files

11.0 Introduction and Stages in the Import of Data into IDEA

To analyse data to arrive at meaningful insights it is imperative to identify, request and download the required data onto the computer before starting IDEA. The "Guide to Downloading Data" lists several methods that can be used for downloading data to personal computers for audit purposes. It also provides pointers and tips on how to deal with a number of situations, on the basis of practical experience of innumerable downloads. The Guide is available on the CaseWare IDEA web site for registered CaseWare users.

11.0.1 Planning

The planning stage involves discussions with users and IT staff on what data is to be kept, how it is to be is stored, and how it can be transferred. Discussions on these issues can help us choose the best available method for our use.

11.0.2 Requesting the Data

Once the transfer has been planned, a formal request should be made to the relevant parties, asking permission where necessary, and describing what is required (format, method of storage, method of transfer, and time frame for the transfer of data). IDEA and other similar software rely on the transfer of data from the host computer system to the personal computer used by the auditor. The host computer may be a mainframe or a minicomputer, a personal computer, or possibly some form of distributed computing over a network. The data may be in a single file, a database, or spread across a number of systems. There are several means of transferring data, many conventions for the file types, and a myriad of structures for storing data.

11.0.3 Performing the Transfer

The transfer may be achieved through a direct connection from the host computer to the personal computer via a network that includes dedicated leased lines, or using media such as an external hard disk. The transfer may be achievable in a single action, or there may be an intermediate step, such as via a file server or a personal computer on a network and then to a laptop.

In order to interrogate data using IDEA, the data must first be imported. The **Import Assistant** guides the user through the process of import.

11.1 Data File Types compatible with IDEA Import

IDEA integrates and connects seamlessly with the following Data file formats:

- 1. Microsoft Excel
- 2. Microsoft Access
- 3. ASCII (Text) Fixed Length
- 4. ASCII (Text) Delimited (including Comma Separated Value)
- 5. EBCDIC Fixed Length
- 6. Print Report and Adobe PDF files
- 7. XML
- 8. SAP/AIS
- 9. dBASE III and IV
- 10. Lotus 1-2-3 (.wk1, .wk3, .wk4)
- 11. AS400
- 12. Proprietary Accounting Packages
- 13. ODBC

11.2Data File Types and Common Application Systems

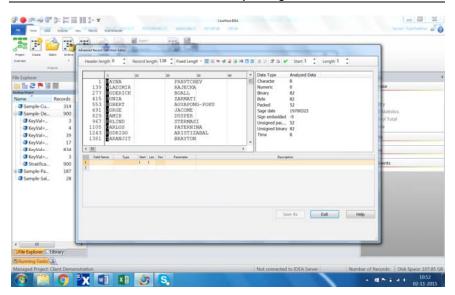
These data file formats are frequently used for varying mainstream commercial applications.

Report File Format	Illustrated Applications					
Spreadsheet	Tally, SAP ERP, Oracle Apps.					
ASCII Fixed Length	Tally					
Comma Separated Value	BAAN ERP					
Print Report	Home-grown ERP Systems					
XML	Tally, Sabre (Airline Reservation System)					
SAP / AIS	SAP ERP					
Dbase	TBA, PBA Banking Applications.					
Lotus	Lotus Workbook					
ODBC	Direct interface to SQL, Oracle Server					

For ASCII Fixed Length, EBCDIC Fixed Length, and ASCII Delimited, the data must be defined for IDEA in terms of size, type and number of fields (often called a record definition). Defining the data involves creating a record definition that describes the format of the file. The **Import Assistant** guides the user through the process of defining and importing the data.

Record definitions are file based and can be reused as required.

Importing various Data File Formats



Other types of files, such as dBASE, Lotus 1-2-3, Microsoft Excel and Microsoft Access can be read directly. However, Lotus 1-2-3, SAP, and Microsoft Excel files must confirm to their respective database formats.

IDEA connects to a wide variety of databases, such as Oracle or Sybase, with Microsoft's ODBC (Open Database Connectivity) for data import.

Where problems arise in obtaining the raw data, or an extract of the data, due to technical reasons, lack of IT assistance, or time constraints, a very successful and simple means of obtaining data is to import via Print Report files. Data formatted as a report is printed to a file and can be extracted using the Report Reader module in IDEA.

11.3 Importing a Microsoft Excel File

IDEA reads dBASE, Lotus 1-2-3, Microsoft Excel, and Microsoft Access files directly because the information about the data is detailed in the file header. These file types are the easiest to import, and are imported in generally the (same what same?) way.

- 11.3.1 Click the **Desktop** button to display the **Import Assistant** dialog box.
- 11.3.2 Select Microsoft Excel

Col Col Address Col Address Col Address Col Col	Terret Rome	~ Q
Tele Caplorer ● ■ ■ ●	Sietche format: Advace Record Definition Editor - 45400 - 4000 - 4	Properties • • Obtabase # Data # Bistory # Faid Statistics # Control Total # Criteria • Results • Consenses • Consenses • Consenses
Managed Project Clerit Demonstration	Not connected to IDEA Server	Number of Records: Disk Space: 107.85 GB

- 11.3.3 Click the **Browse** button adjacent to the **File name** box to select the Microsoft Excel file to be imported.
- 11.3.4 Navigate to and select C:\Users\PCNAME\Documents\My IDEA Documents\IDEA Projects\Samples\Source Files.ILB\Sample.xls

lardes 1	Select Fie			X
Preside		ects Client Demonstration Source Files.ILB	• 4 Search	Source File_ P
ile Explorer *	Organize * New fo	lder	11	
				-
lame Records 5	A Favorites	Name	Date modified	Type
Sample-Cu 314	Desktop	Q Medicalatsx	19-03-2015 16:43	Microsoft I
Sample-De. 900 1	Us Recent Places	Motor Premium Policy Dumpals	19-03-2015 16:43	Microsoft I distics
-3 KeyVal= 3	 OneDrive 	C Payroll Analytics Base Dataxisx	18-11-2014 19:34	Microsoft I
KeyVal=4	Downloads	Q Payroll August xlsx	20-07-2012 11:04	Microsoft I
KeyVal= 39		Payroll July alsx	09-08-2012 09:38	Microsoft I
KeyVal= 17 KeyVal= 834	Cibraries	Payroll September.xlsx	09-08-2012 09:39	Microsoft I
KeyVal= 814	- Documents	Sampleals	18-08-2015 19:33	Microsoft
Stratifica. 900	4 Music	SAP Procurement Base Dataxisx	17-09-2014 10:38	Microsoft (1994
Sample-Pa187	- Pictures	G Ser Tax Credit Availment Dataxisx	09-02-2015 17:09	Microsoft (
Sample-Sal_ 28	- Videos	Travel Analytics Base Data.dsx	14-11-2014 15:16	Microsoft (
		Vendor Master Changes Base Dataxiss	09-12-2014 17:23	Microsoft (
	-& Homegroup	WorldTotalEmissions.als	07-12-2012 19:08	Microsoft I +
	File	pame: Sample.xls	Microsoft Excel file	s (*.xlsx: *.) •
			Open 💌	Cancel
			- Phil	

- 11.3.5 Click Open.
- 11.3.6 Click Next.

The **Excel** dialog box appears. The **Excel** dialog box displays a preview of the Microsoft Excel file.

IDEA determines the field types and widths by scanning all rows in the spreadsheet. During the scan, if any unusual non-numeric characters are found in Numeric or Date fields, the affected fields will be imported into IDEA as Character fields.

In the **Preview** area of the **Excel** dialog box, the Date and Character fields are left-justified, while Numeric fields are right-justified.

11.3.7 The first row of the spreadsheet contains column field names. We check the box **First Row is Field Names** to ensure that the field names from Excel and duly reflected in IDEA. This allows for meaningful analysis of data imported based on the field names captured.

The Explore The	History Edd Statistics
Mare Naue Weblog Ør Pint row in lind names Index names Index names Index names Index names Index names Index names Index names Index names Index names Index names Index names	AA U U Data History Field Statistics Control Total
mme Records S Sample-Cu. 314 Sample-Cu. 314 Control Sample-Cu. 314	History Field Statistics Control Total
3 ample 0 314 Z Detailed Seles Concernent 3 kmyle 0 38 smyle 0 * Concernent Concernent 3 kmyle 0 3 * Socielle n * Concernent 3 kmyle 0 3 * * Meller Meller Meller 3 kmyle 1 4 * * Meller Meller Meller 3 kmyle 1 3 * * * Meller Meller </td <td>Field Statistics Control Total</td>	Field Statistics Control Total
Skmylet 3 Sknylet 3 Sknylet 3 Sknylet 4 Prefer: Prefer: Meyore 17	Control Total
KeyVal	
KeyVals 39 Preview: XexVals 17 Inv.nki	- Criteria
TANAN IT IN A TANAN IT IN A TANAN INA NA TANAN IN A TANAN INA NA	
	E + Results
1 100000 21/07/208 101 21/24 10 5.09 72	
Image: Constraint of the state of	I Indices
Keyvall	· Comments
4 100/28 25/6/200 10 2234 25 5.00 30	
Sample-Pa., 187 : 7 100054 84/01/208 20 210 05 53/9 200 Sample-Sal, 28 9 10 100054 2010200 201 2130 05 53/9 20	
9 10/0449 10/96/2008 101 22340 05 5.00 8	•
10 1009417 121.11.1009 519 21.51 (m 4.66 566 56	
Input file name: C:\Users\Jairam Rajsbekhar\Documents\My IDEA Documents\JDEA P	Terms and the second
and the Country	
Output file name: Sample1	<u> </u>

- 11.3.8 If the First row in Excel is blank, default field names (Col1, Col2) will be provided during the import.
- 11.3.9 If the spreadsheet being imported has one or more empty cells in a column, the data in that column is imported into IDEA as a Character field. If the data is numeric, date or time and we want to import it into IDEA as a Numeric, Date or Time field, we have to select the **Import empty numeric cells as 0** check box.

- 11.3.11 In the Select sheets to import box, select the worksheet Detailed Sales. If we select multiple worksheets to import, each worksheet will be imported as a separate database.
- 11.3.11 If the user does not enter a unique Output File Name, then the default name i.e. Excel file name will take over.
- 11.3.12 Click **OK**.

The file is imported into IDEA and displayed in the **Database** window.

	1. I.		. 0		- 6.	ind lang E na	Prigat Salest Adventulari Sarrar ERA Sarrar							
ile Explorer			1	Sample	-Detailed	×							- 5	Properties
				INV NO	DINV DAT	ESALES	REP NO CUSTN	OPROD C	ODE UNIT PRICE	OTY SAL	ES BEF TAX SA	LES TAX SALE	5 -	Database
Aame F Sample-Cu. Sample-Cu. KeyVal KeyVal KeyVal KeyVal KeyVal Strafica Strafica Sample-Pa	Records 314 900 3 4 39 17 834 3 900 187 28	5	12134567891011121341551677892022223455627	2000254 2000254 2000255 2000254 2000254 2000254 2000254 2000556 2000576 2000576 2000776 20000775 20000775 2000075 2000075 2000075 2000075 2000075 2000075 2000075 2000011 2000075 2000011 2000025 2000012	$\begin{array}{c} 110.300\\ 100.200\\ 100.200\\ 100.200\\ 100.200\\ 100.000\\$	444 445 345 345 345 345 345 345 345 345	2,126 2,127 2,128 2,128 2,128 2,138 2,138 2,148 2,148 2,148 2,148 2,145 2,145 2,145 2,145 2,145 2,145 2,145 2,145 2,145 2,1462,146 2,1462,146 2,146 2,146 2,146 2,1462,146 2,146 2,146 2,146 2,1462,146 2,146 2,146 2,146 2,146 2,146 2,1462,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,146 2,1462,146 2,146 2,146 2,146 2,146 2,1462,146 2,146 2,146 2,1462,146 2,146 2,146 2,1462,146 2,146 2,1462,146 2,146 2,1462,146 2,1462,146 2,1462,146	"你这样的你的你想是你你的你就是你的你的你?"	199 199 199 199 199 199 199 199 199 199	22 43 1236 437 236 437 236 437 236 43 348 236 43 348 348 348 348 348 348 348	277.07 7.34.04 1.407.55 2.209.51 2.209.51 2.209.51 2.209.55 2.209.	77.44 77.44.55 246.95 246.95 245.95 245.95 245.95 245.95 245.95 245.95 245.95 245.95 245.95 245.95 25.95 25.95 25.95 20.		Data History Field Statistics Control Total Criteria Criteria Stratification Indices Minimum Control No index CUSTNO/A + PROD_CODE/A + SALESFEV.STAV/D USTNO/A PROD_CODE/A SALESFEV.STAV/D SALESFEV.STAV/D SALESFEV.STAV/D SALESFEV.STAV/A
			28 29 30 31 32 33 34	2000320 1000320 2000334 2000362 2000362 2000364	12-02-2008 73-05-2008 14-05-2008 08-04-2008 01-02-2008 03-06-2008	382 103 382 202 382 382 382	25154 30128 30908 30704 71505	85 85 85 85 85 85	5.00 5.00 5.00 5.00 5.00 5.00 5.00	30 30 21 21 21 9	129.30 179.30 1.245.92 149.35 125.79 53.91	17.87 17.87 124.59 14.88 12.58 5.39		Comments Add comment
III III			34 35	3000413 3000435	28-83-2908	調	32902	04. 15	305.69	400	42,276,00 476,87	4,327,80		
	rary													

Note

Worksheet/Spreadsheet Import Rules

It is strongly recommended that before importing data from Microsoft Excel or lotus 1-2-3 worksheets/spreadsheets, the format of the data is checked to ensure that it conforms to the following set of rules:

- The data to be imported is in table format, such as a block of rows and columns.
- The first row of the table should be the field name or contain the data to import.
- All columns containing both text and numeric data are formatted as text.

- All numeric columns contain a fixed number of decimal places.
- Blank lines, sub-total lines and total lines are removed from the data area.
- Merged cells are un-merged.
- The range of data to be imported is selected.
- The above steps are performed for all tables of data within the worksheet/spreadsheet.
- The worksheet/spreadsheet is closed during the import process.

11.4 Importing ASCII, EBCDIC or Delimited (Comma Separated Value) ASCII Files

11.4.1 Introduction

File types such as ASCII (American Standard Code for Information Interchange) and EBCDIC (Extended Binary Coded Decimal Interchange Code) are very common. Most mid-range computers and personal computers hold data in ASCII format, while the EBCDIC format is typically found on mainframe and mid-range computers.

Unlike Microsoft Excel, Microsoft Access, Lotus 1-2-3, and dBASE, ASCII and EBCDIC files do not contain standards or headers defining the structure of the data. Since no such header exists, we have to define the file structure in IDEA.

The following print-screen displays the ASCII file format:

99788	87 81	J1		244	ı			2	11	5	:	28	. 98		8.	. 8	899	/8	876	91 J	н	2	44				58	18	h	13	.94
8.86	999/	88	/8	1J1	1	24	4				23	88		8	. 86		442	. 9:	299	//	187	01	J2		248	1			11	88	
1676.	. 16		8.0	889	97	08	178	нJ	2	2	48				481	8		8	. 86	916	76	.1	699	971	88/	84	13	- 2	237		
1188	47	۰.	88		8.	. 8 (999	/6	87	84	J3	- 2	237				21	16		8			-30	8.1	8 8 9	97	88	784	4J3		237
	436			0.	. 86	14	46		89	97	88,	7 84	4J4		238	•			11		1	64	2.0	88		۰.	00	99,	/ 88	/8	4J4
238			21	16					- 4	2.	88	99,	/ 88	78	4.14		23	8			4	82	8		۰.	88					
688.6	9899	7.	87	84.	4	2	38				2	11!	5		7.3	11		۰.	889	97	88	/ 8	4.J	4	23	8			- 5	88	8
17.88	3	8.	889	99/	88	17	94.J	4	2	38				58	18		143	1.1	Ő	8	1.8	89	970	88,	/ 84	J4	F 3	231	8		
2818		۰.	88	10	68.	35	99	/8	87	87	Jó	-	132	39			21	15	1	26	1.4	9		8.1	889	97	88	/87	7J6		
13239	,		50	18	18	807	1.6	0		۸.	88	99,	/ 88	/8	736	j	13	23	9		2	01	8								
0.001	933	1.4	999	77	38/	87	' J7		28	Ő			2	81	8 1	9	33.	49		8.	88	99	/81	871	87 J	7	2	Øő			
1818		۰.	88	193	з.	49	99	/8	8/	86	J8	- 2	286	35	8		21	15		53	1.1	3		8.1	889	97	88	/80	6J8		
28835																															
8.889	9978	18/	86.	J 8	2	81	135	8		2	30	8		۴.	88	8	12.	13	99/	788	7.	17 J	9	23	34			1	118	8	
2917.	. 89		8.0	889	9/	88	1/1	7.	9	2	34				218	10			. 86	1 2	16	.1	499	971	88/	87	J9	- 2	234		
2116		۰.	88	17	6.	75	599	/6	8/	87	J9	- 2	234				43	88			. 8	182	52	5.1	889	97	88	/87	7J1	8	895
	101	8	29	17.	89			. 6	89	97	88,	/ 87	7J1	8	895	,			11				8.0	882	291	7.	89				

Technical Notes

Explanation of the ASCII Delimited File Format

ASCII Delimited files are a common format of variable length file where each field is long enough to contain only the data stored within it. This is an advantage with text fields, such as names, addresses, and narratives.

In order that software use such data files, there is a separator (a special character) at the end of each field within the record. Additionally, text fields may be enclosed (encapsulated) with another character, typically quotation marks (" ").

The separator is often a comma (,) and such files are known as comma separated value (CSV). However, the separator could be another ASCII character, for example a ; : ^ ~ These files are referred to as ASCII Delimited (comma separated files are also often referred to as ASCII Delimited).

There must also be a record delimiter at the end of each record, typically the carriage return (<CR>) and line feed (<LF>) characters, or sometimes only the line feed character.

When this type of file is imported, a fixed length database is created within IDEA.

Example of an ASCII Delimited – CSV:

"6776",1865,1,190,"04-02-2006" <CR><LF>
"6714",1865,1,97,"04-02-2006" <CR><LF>
"11905",2366,11,9,"25-02-2006" <CR><LF>
"7555",4352,1,79,"03-03-2006" <CR><LF>
"4547",4815,5,17,"03-03-2006" <CR><LF>
"6344",4815,1,7,"03-03-2006" <CR><LF>
"151",2366,1,3,"29-03-2006" <CR><LF>
"145",2366,1,1,"05-08-2006" <CR><LF>
"206",2366,1,2,"29-03-2006" <CR><LF>
"207",2366,1,3,"29-03-2006" <CR><LF>

The following specimen data request has been reproduced.

Request for Data Files for Analysis Purposes – DATA INDENT

Please supply the following data from the Inventory system in ASCII Delimited format:

FIELD NAME	TYPE	Dec	DESCRIPTION							
PRODCODE	Ν	0	Product code							
DEPOT	С		Depot name							
QTY	Ν	0	Quantity on hand							
AV_COST	Ν	3	Average cost for item							
TOTALCOST	Ν	3	QTY * AV_COST							
OBSOLETE	С		Obsolete = "Y"							
MAX	Ν	0	Maximum inventory level on hand							
MIN	Ν	0	Minimum inventory level on hand							
DELQTY	Ν	0	Last delivery quantity							
ORDERNO	Ν	0	Last delivery order number							
DELDATED	D		Date of last delivery (YYYYMMDD)							
CURSELLPRI	Ν	2	Current selling price							
CUREFFDATE	D		Date selling price effective (YYYYMMDD)							
PREVSELPRI	Ν	2	Previous selling price							
USAGE	Ν	0	Sales quantity in current year							
PREVCOST	Ν	3	Unit cost of last purchase							

Data is required for the following period: December 31, 2014

Please also supply the following control totals for reconciliation purposes:

- 1. Total cost of inventory on hand at December 31, 2014.
- 2. Number of transactions supplied, i.e., product lines.
- **3.** Record definition to support the data file provided.

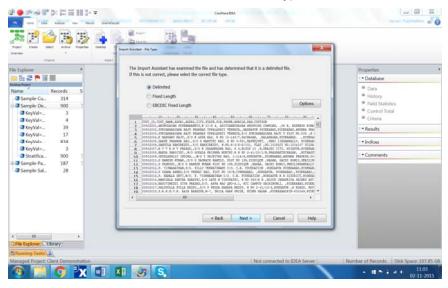
Importing an ASCII Delimited File

You have been provided with an ASCII Delimited file:

- 11.4.2 From IDEA select the **Import Assistant** by clicking the Desktop button under Home in the IDEA toolbar.
- 11.4.3 Select Text from the list. Click the Browse button and the Select File dialog box will be displayed. Navigate to and select the file from the Working Folder. Click Open, and then click Next.

Col Date Date <thd< th=""><th></th><th>• •</th></thd<>		• •
The Explore	AS400 Bits Star Hirosoft Access Hirosoft Access Hiros	Popertie Postasee Data Data Data Hatory Frad Statistics Control Total Control Total Postat Persuits Comments
CFRetExplorer, Ubray	Not connected to IDFA Server	Number of Records: Disk Space: 107.85 C8

11.4.4 The Import Assistant - File Type screen will be displayed. Ensure Delimited is selected as the correct file type.



View the data and determine the field separator and the record delimiter.

Click Next

11.4.5 The Import Assistant will try to determine the field separators and text encapsulators (if any) for the file. Accept the defaults and click Next to proceed.

Some (but not this) ASCII Delimited files have field names as the first row of the file. To use these as field names, select the First visible row is field names check box.

Popet Dealer Select Write Properties Dealtop Dearters Projects	I separat Accutent - Specify field Separation and Test Encapsision	
File Explore * <t< th=""><th>Place impact the file displayed in the proteins and make change, if required, in the options below. Field Segmentary Place impact the file displayed in the proteins and make change, if required, in the options below. Place impact the file displayed in the proteins and make change, if required, in the options below. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and the proteins and</th><th>Properties • Outase • Unta • Hotory • Fred Statistics • Control Total • Control • Reads • Indices • Comments</th></t<>	Place impact the file displayed in the proteins and make change, if required, in the options below. Field Segmentary Place impact the file displayed in the proteins and make change, if required, in the options below. Place impact the file displayed in the proteins and make change, if required, in the options below. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and make change. Place impact the file displayed in the proteins and	Properties • Outase • Unta • Hotory • Fred Statistics • Control Total • Control • Reads • Indices • Comments
II II In	Not connected to IDEA Server	Number of Records: Disk Space: 107.85 G • • • • • • • • • • • • • • • • • • •

11.4.6 The **Import Assistant - Field Details** step will be displayed. Click on each field heading in turn and using the record definition, define the field details.

FIELD NAME	TYPE	Dec	DESCRIPTION
PRODCODE	Ν	0	Product code
DEPOT	С		Depot name
QTY	Ν	0	Quantity on hand
AV_COST	Ν	3	Average cost for item
TOTALCOST	Ν	3	QTY * AV_COST
OBSOLETE	С		Obsolete = "Y"

FIELD NAME	TYPE	Dec	DESCRIPTION
MAX	Ν	0	Maximum inventory level on hand
MIN	Ν	0	Minimum inventory level on hand
DELQTY	Ν	0	Last delivery quantity
ORDERNO	Ν	0	Last delivery order number
DELDATED	D		Date of last delivery YYYYMMDD
CURSELLPRI	Ν	2	Current selling price
CUREFFDATE	D		Date selling price effective YYYYMMDD
PREVSELLPRI	Ν	2	Previous selling price
USAGE	Ν	0	Sales quantity in current year
PREVCOST	Ν	3	Unit cost of last purchase

Enter the correct field name in the **Field name** box. The **Import Assistant** suggests the file type for that field in the **Type** box. If this is incorrect then change it to the field type specified in the record definition.

Enter the description of the field in the **Description** box.

If the field is a Numeric field with decimals, it is necessary to specify whether or not the decimals are implied. However, if there are decimals in this database, the decimal place is stored in the number, therefore the Check here if decimals are implied check box must not be selected. Specify the number of decimals in the Number of decimals box.

Define the Date fields as type **Date**, i.e., change from **Character** to **Date** in the **Type** box. A mask must be defined for the Date fields. Click the **Date Mask** box and enter the appropriate date format i.e., **YYYYMMDD**.

Importing various Data File Formats

You can now specify field details. Select a field by clicking on the column heading below, and then modify its information below.								
Field name	PRODCOL	DE		Type: Nu	imeric			*
Description	n: Product co	ode		Number of	decimals: 0			
🗌 Do not	import this fiel	ld		Check I	here if decimals	are implied		
	verted Examp	le		110,211,059				
	PRODCODE	DEPOT	QTY	AV_COST	COTALCOST)	MAX	MIN	^
1 2	11021105 11120110		32 125	4.368 3.511	139.776 438.875	100 300	25 100	1
3	11120110		123 670	0.69	462.3	900	200	
4	11180705	QUEBEC	684	0.591	404.244	750	100	5
5	11180755 11181305	-	61 000	1.877	114.497 542.52	300 3000	100 500	4 5
7	11181555		990 140	0.540	79.94	5000	100	1
8	11182160	OTTAWA	360	0.482	173.52	388	169	4
9 10	11182170 11182180		380	0.401	152.38	404 250	127 100	2 1 🌱
oort Assis	stant - Fie	ld Detail		Back	Next >	Cancel		Help
You can n		eld details. :	s		Next >		elow, and	
You can n	ow specify fie nformation be	eld details. S elow.	s	ield by clickin			elow, and	then
You can n modify its i	ow specify fie nformation be PRODCO	eld details. : elow. DE	s	ield by clickin Type: N	g on the colum	n heading b	elow, and	
You can n modify its i Field name Description	ow specify fie nformation be PRODCO	eld details. S elow. DE code	s	ield by click in Type: Number o	g on the colum lumeric	n heading b		then
You can n modify its i Field name Description	ow specify fie nformation be e: PRODCO n: Product c	eld details. S elow. DE code eld	s	ield by click in Type: Number o	g on the colum Jumeric of decimals: 0 < here if decima	n heading b		then
You can n modify its i Field name Description	ow specify fie nformation be PRODCO Product c import this fie verted Examp PRODCODE	eld details. Solow. DE code eld DEPOT	s	ield by clickin Type: Number o Check 110,211,05	g on the colum Jumeric of decimals: 0 < here if decima	n heading b		then
You can n modify its i Field name Description Do not Con	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105	eld details. : elow. DE code eld DEPOT QUEBEC	Select a f	ield by clickin Type: N Number o Check 110,211,05 AV_COS 4.368	g on the colum Jumeric of decimals: 0 k here if decima 9 T roTALCOST 139.776	n heading b	d MII 25	
You can n modify its i Field name Description Do not Con 1 2	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105 11120110	eld details. : slow. DE code eld DEPOT QUEBEC OTTAWA	s Select a fi	ield by clickin Type: N Number o Check 110,211,05 AV_C0S 4.368 3.511	g on the colum lumeric of decimals: 0 k here if decima 9 T <u>rOTALCOST</u> 139.776 438.875	n heading b als are implie	d MII 25 100	then
You can n modify its i Field name Description Do not Con	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105 11120110 11160190	eld details. Selow. DE code eld ole DEPOT QUEBEC OTTAWA QUEBEC	Select a f	ield by clickin Type: N Number c Check 110,211,05 AV_C0S 4.368 3.511 0.69	g on the colum Jumeric of decimals: 0 k here if decima 9 T roTALCOST 139.776	n heading b	d MII 25	
You can n modify its i Field name Description Do not Con	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105 11120110	eld details. : slow. DE code eld ole DEPOT QUEBEC QUEBEC QUEBEC	Select a fi	ield by clickin Type: N Number o Check 110,211,05 AV_C0S 4.368 3.511	g on the colum lumeric of decimals: 0 < here if decima 9 T TOTALCOST 139.776 438.875 462.3	n heading b als are implie	d <u>MII</u> 25 100 200	then
You can n modify its i Field name Description Do not Con 1 2 3 4 5 6	ow specify fie nformation be PRODCO r: Product c import this fie verted Examp PRODCODE 11021105 11120110 11160190 11180705 11180755 11181305	eld details. s elow. DE ode eld DEPOT QUEBEC QUEBEC QUEBEC QUEBEC TORONTO	Select a f	ield by clickin Type: N Number of Check 110,211,05 AV_COS 4.368 3.511 0.69 0.591 1.877 0.548	g on the colum Jumeric of decimals: 0 k here if decima 9 T <u>fOTALC037</u> 139.776 438.875 462.3 404.244 114.497 542.52	1 MAX 100 300 900 750 300 3000	d 25 100 200 100 100 500	then
You can n modify its i Field name Description Do not Con Con 1 2 3 4 5 6 7	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105 11120100 11160190 11180705 11180755 11181555	eld details. S slow. DE code eld ole <u>DEPOT</u> QUEBEC QUEBEC QUEBEC QUEBEC QUEBEC QUEBEC	Select a f	ield by clickin Type: N Number of Check 110,211,05 AV_COS 4.368 3.511 0.69 0.591 1.877 0.548 0.571	g on the colum lumeric of decimals: 0 k here if decima 9 T TOTALCOST 139.776 438.875 462.3 404.244 114.497 542.52 79.94	n heading b als are implie 100 300 900 750 300 3000 500	d 25 100 200 100 500 100	then
You can n modify its i Field name Description Do not Con 1 2 3 4 5 5 7 8	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp 11021105 11120110 11160190 11180705 11181305 11181355 11181355 11182160	eld details. S slow. DE code eld ole OTTAVA QUEBEC QUEBEC QUEBEC QUEBEC QUEBEC OTTAVA	S Select a fi Select a fi 32 125 670 684 61 990 140 360	ield by clickin Type: N Number of Check 110,211,05 AV_C0S 4.368 3.511 0.69 0.591 1.877 0.548 0.571 0.482	g on the colum lumeric of decimals: 0 k here if decima 3 T TOTALC037 139.776 438.875 462.3 404.244 114.497 542.52 79.94 173.52	n heading b als are implie 100 300 900 750 3000 3000 3000 3000 3000 3000 300	d 25 100 200 100 100 500 100 100	then
You can n modify its i Field name Description Do not Con Con 1 2 3 4 5 6 7	ow specify fie nformation be : PRODCO n: Product c import this fie verted Examp PRODCODE 11021105 11120100 11160190 11180705 11180755 11181555	eld details. s Jow. DE code eld DEPOT QUEBEC OTTAWA QUEBEC QUEBEC QUEBEC TORONTO QUEBEC OTTAWA QUEBEC	Select a fi Select a fi 32 125 670 684 61 990 140 360 380	ield by clickin Type: N Number of Check 110,211,05 AV_COS 4.368 3.511 0.69 0.591 1.877 0.548 0.571	g on the colum lumeric of decimals: 0 k here if decima 9 T TOTALCOST 139.776 438.875 462.3 404.244 114.497 542.52 79.94	n heading b als are implie 100 300 900 750 300 3000 500	d 25 100 200 100 500 100	then

Data Analysis for Auditors

Click Next.

11.4.7 The Import Assistant - Create Fields screen will be displayed. The Import Assistant -Create Fields screen allows you to add fields in the imported file. This can be done during the import or at any time while using IDEA. For this exercise, no fields will be added.

Click Next, and then Next again.

	Assistant - Creat					
You d and d	can add one or more lescription for the fiel	fields to the imported d(s) in the columns b	d file by below.	enterin	g the field name, type, le	ngth, parameter
						Delete
		-		-		
_	Field Name	Туре	Len	Dec	Parameter	Tag Nan
1						
]	
<						>
		C <	Back		Next > Cance	Help

- 11.4.8 The definition will automatically be saved in the Working Folder. IDEA will give the definition the same name as the source file. Click the Browse button adjacent to the Save record definition as box if you want to change the name.
- 11.4.9 Select the Generate field statistics check box and enter Inventory at Dec 31 2014 in the Database name box.

Importing various Data File Formats

Angel Data Meter Antonio Datas Dense	The second secon	
Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Import Impor	Specify a discription name for the database. Click Rhink to import the database into IDEA. Working directory : C:(Javes)Jairam Rajdhekhar(Documents)/JDEA Documents)/JDEA Projects/Client D IVen encirity rou life to use the data Re. Disport - IDEA runs failer when a file is imported. Generate field attatios Generate f	Properties - Data - Data - Hetory - Field Statistics - Control Total - Creteria - Results - Comments - Comments
Alle Explorer , Library / Reasenge Franke & Control Control Common training	Not connected to IDEA Server	Number of Records Disk Space: 107.85 1105

- 11.4.10 Click Finish to import the file into IDEA.
- 11.4.11 Maximize the *Database Window*.
- 11.4.12 Click the **Control Total** link in the *Properties Window* and select the **TOTALCOST** field and then click **OK**.

Select Control Total	×
MAX MIN ORDERNO PREVCOST PREVSELLPRI PRODCODE QTY TOTALCOST USAGE	OK Cancel Help

Results The control total of 626,963.915 will appear beside the Control Total link in the *Properties Window*. There are 767 records as can be seen on the Status bar at the bottom of the screen. You are now ready to verify that the data has been imported correctly and to commence testing. Verifying That the Database Has Been Correctly Imported 11.4.13 Click the Field Statistics link in the *Properties Window*.

Field Statistics are displayed for the Numeric fields in the database.

	The Field Statistics are available instantly as they were generated when the database was imported.
Notes	Field statistics are also available for Date fields. These will be reviewed in the following exercise.

11.4.14 Study the field statistics for all Numeric fields, but especially, the QTY and

TOTALCOST fields. Notice in particular the Net Value, Maximum Value, Minimum Value, # of Zero Items, and # of Negative Records statistics.

Results

Please check your results against the figures below.

	다 술 추 팬·									
Field Type	Numeric Statistics	PRODCODE	QTY	AV_COST	TOTALCOST	MAX	MN	DELQTY	ORDERNO	CURSELLPR
Numeric Date	Net Value	1,214,074,724,378	214,954	4,178.563	626,963.915	263,856	91,399	204,443	581,830,762	8,603.78
Time	Absolute Value	1,214,074,724,378	214,904	4,170.563	627,153.615	263,056	91,399	204,443	581,800,762	8,603.78
Numeric Fields	# of Records	767	767	767	767	767	767	767	767	767
PRODCODE	# of Zero Items	0	68	2	69	7	152	0	0	1
V QTY V AV COST	Positive Value	1,214,074,724,3	214,969	4,178.563	627,061.765	263,056	51,399	204,443	581,830,762	8,603.78
TOTALCOST	Negative Value	0	-15	0.000	-97.850	0	0	0	0	0.00
IV MAX IVI MIN	II of Positive Records	767	636	765	695	760	615	767	767	766
DELQTY	# of Negative Records	0	3	0	3	0	0	0	0	0
CIRCERNO CURSELLARI	# of Data Errors	0	0	0	0	0	0	0	0	0
PREVSELUPRI	# of Valid Values	767	767	767	767	767	767	767	767	767
V USAGE	Average Value	1,582,887,515	200	5.440	817,424	344	119	267	758,500	11.22
PREV0051	Minimum Value	110,211,059	-8	0.000	-67.000	0	0	2	266	0.00
	Maximum Value	4,527,504,110	10,800	128.436	70,555.641	12,000	3,000	3,000	960,494	249.00
	Record II of Min	1	202	519	90	144	15	87	541	140
	Record # of Max	767	576	662	384	576	577	255	281	662
	Sample Std Dev	1,297,729,561.90	653.53	10.93	3,863.12	614.4)	216.44	540.09	30,257.61	20.06
	Sample Variance	1,684,102,016,0	427,098.71	113.54	14,923,698.38	377,601.52	46,844.74	292,562.93	915,522,953.86	402.44
	Pop Std Dev	1,296,883,308.51	653.10	10.93	3,860.60	614.09	216.30	540.54	30,237.88	20.05
	Pop Variance	1.681.906.315.8	426,541.07	119.39	14.904.241.14	377,109.22	46,783.66	292,101.49	914.329.312.46	401.92
	Pop Skewness	0.915824	8.798266	5.973602	13:599907	11.116465	6.159424	8.341977	-20.496621	5.864411
	Pop Kurtosis	-0.310923	112.051908	49.738079	212.937655	181.196105	59.017565	105.643266	513.865118	43721371

Importing various Data File Formats

Field: QTY

STATISTIC	VALUE	COMMENT
Net Value	214,954	
Average Value	280	
Minimum Value	-8	Should not be any negative quantities
Maximum Value	11,800	
# of Records	767	
# of Zero Items	68	
# of Positive Records	696	
# of Negative Records	3	These should be identified

Field: TOTALCOST

STATISTIC	VALUE	COMMENT
Net Value	626,963.915	
Average Value	817.42	
Minimum Value	-67.00	
Maximum Value	70,555.64	
# of Records	767	
# of Zero Items	69	Different number or zero-quantity
# of Positive Records	695	
# of Negative Records	3	

11.4.15 If your computer is attached to a printer, print the field statistics for the QTY and TOTALCOST fields by clicking the Print button on the Field Statistics toolbar and selecting each field.

- 11.4.16 The **Field Statistics** reports should be filed with the audit documentation as proof of reconciliation.
- 11.4.17 Return to viewing the database by clicking the **Data** link in the *Properties Window.*

Conclusion

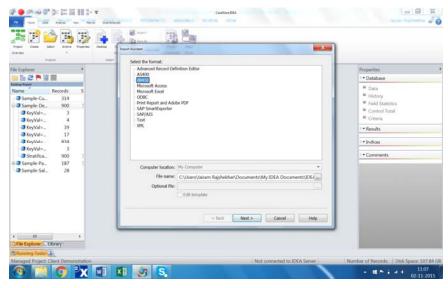
The file has been imported correctly and the client's report footer total reconcile correctly. However, there are some negative items, which we should identify and report them to the client.

11.5.0 Importing a dBASE File

11.5.1 Introduction

dBASE files are one of the most common computer formats used for exporting data from computer systems and the easiest file format to import into IDEA. The DBASE file structure consists of a file header holding information about the file, including the field definition information. dBASE files are imported directly when the file is dBASE III or newer. If we have older Dbase files, it is recommended that we read them into Microsoft Excel or Microsoft Access and then export or import them to IDEA from there.

- 11.5.2 From IDEA, select the **Import Assistant** by clicking on the Desktop button under Home in the IDEA toolbar.
- 11.5.3 Select **dBASE** from the list of available formats.



11.5.4 Select the file to import by clicking the **Browse** button then navigate to and select the following file:

Click Open, and then click Next.



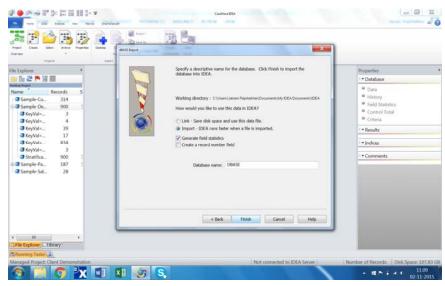
When selecting a file, the **Files of type** may need to be changed to *dBase* or *All Files*.

10.5.5 The dBASE Import – Create Fields screen dialog box will be displayed. Create Fields allows you to add Virtual or Editable fields in the imported file. This can be done during the import or at any time while using IDEA. For this exercise, no fields will be added. Click Next.

dBA	dBASE Import - Create Fields									
	You can add one or more fields to the imported file by entering the field name, type, length, parameter and description for the field(s) in the columns below.									
							Delete			
		Field Name	Туре	Len	Dec	Parameter	Tag Nan			
	1									
	2							_		
	<							>		
-			<	Back		Next > Cancel	Hel	p		

- 10.5.6 The final step of the **Import Assistant** will ask you to change or confirm details about the database to import. Select or confirm the following options:
 - Leave the default selection to Import rather than Link the file
 - Accept the option to Generate field statistics on import

- Accept the default not to Create a record number field
- Enter the database name: Accounts Receivable



Click **Finish**. The **Accounts Receivable** database will be imported, opened, and selected as the active database.

Technical Notes

Differences between Importing and Linking to a Data File

Each IDEA import or link results in an IDEA master database, *filename.imd*, being created in the **Project Folder**.

However

- In both cases, a file (*filename.imd*) is created that stores all the information about the database.
- **Importing** results in a copy of the data being included in the *filename.imd* file.
- If the option to Link to the database is selected, the original data is not included in the *filename.imd* file. The advantage of linking is that it can save disk space since the IMD file is smaller than it would be if the source data had been copied. However, both the source data file and the *filename.imd* file must be stored because IDEA must have access to the source data file to obtain the data needed to perform analytical functions.
- Testing will be faster on imported data files.
- Housekeeping is simplified when data is imported as there is only one file to backup/restore/copy/move and it is stored in the Working Folder.

It is recommended that the **Import** option be used provided we have sufficient disk space.

11.6.0 Importing a Print Report File

11.6.1 Introduction

Report files which can be imported into IDEA have a unique format unlike spreadsheet and delimited files. Report files have headers, footers, subheaders, sub-totals and transactions.

IDEA's **Import Assistant** provides importing capabilities for data files that are in a report format. **Report Reader** is part of the **Import Assistant** within IDEA and can be accessed through it.

Report Reader extracts data from plain text reports, print reports saved to file or Adobe Acrobat PDF files and translates it into the specialised file formats of IDEA. The reports might have come from an application on the computer or from a mainframe computer.

11.6.2 What Kind Of Files Can Be Imported?

With **Report Reader**, we can import data from any ASCII file (TXT, ASC, CSV, DAT, etc.) or an Adobe Acrobat Portable Document Format (PDF) file.

An ASCII file is a plain text file or a printer output file. Most DOS applications will produce text output by a process called "print to disk". This simply means sending a report to a disk file instead of printing it. Windows applications usually have an ASCII text output option or can print a report to a text file with the Generic/Text Only printer driver.

The following is a list of Text formats that Report Reader can read.

Mainframe text reports

Minicomputer text reports

DOS Print to Disk files

Generic/Text Only files (Windows)

Word Processor Text Only files

Text only accounting reports

E-mail reports

Online communications text captures

Print to Files (PRN)

Tab Separated Variables

Standard Data Format (SDF)

Unconverted Files from SAP

Adobe Acrobat PDF files can be opened in **Report Reader**. The text from the PDF file will be converted into plain text format which does not maintain the font size, color, text bounding box and rotation angle as in the PDF file. A page break is inserted at the end of each page and graphics and images are not imported.

11.6.3 How Does Report Reader Work?

Report Reader displays our entire original file in a section of the **Report Reader** window called the **Report** view. Using the scroll bars, we can scroll through the report to view the details of the data and the format. Data to be imported can be selected with a simple point-and-click operation like those in Excel and other applications. The user selects the portions of the file to import. No special knowledge of the file structure or command language is needed to use **Report Reader**.

Report Reader imports the data we select into IDEA. Our selections and instructions for importing data in a report can be saved as a "template". This template can be re-used to import data from other similar data files in the same format. Using templates saves time, particularly if we need to import the same or similar reports on a regular basis.

Report Reader does more than simply extract data from one file and put it in another. It distinguishes between numeric and non-numeric information and handles both appropriately. It also automatically formats data and time fields in their native field type without the need for any manual intervention.

To import a sample Print Report file:

11.6.4 Click the **Desktop** button under Home to access the Import Assistant.

The **Import Assistant** dialog box will be displayed.

11.6.5 Select Print Report and Adobe PDF.

Angel Cardon State State States		~ Q
Name Records S Sample Ca. 314 Sample Ca. 314 Sample Ca. 314 SeyVal= 3 SeyVal= 3 Startple Ca. 90 Startple Ca. 90 Startple Ca. 90 Startple Ca. 3 Startple Ca. 28	Select the format: Selec	Properties * * Data Data Data Data Data Data Data D
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	Not connected to IDEA Server	Number of Records: Disk Space 107.31 G8

- 11.6.6 Click the **Browse** button to navigate to and select the required file.
- 11.6.7 Select the sample file, and then click **Open**. Notice the various types of Print Report files IDEA can import.

Select File					?×
Look in:	accounts Rec	ceivable 🗸	GØ	19 🖽	
My Recent Documents	🗐 Customer.txt				
Desktop					
My Documents					
My Computer					
	File name:	Customer.txt		*	Open
My Network	Files of type:	Print Report (*.prn; *.txt; *.csv; *.as		*	Cancel
		Print Report (*.prn; *.txt; *.csv; *.ase Adobe Acrobat (*.pdf) All Files (*.*)	c; *.dat)		Help

- 11.6.8 Click Next to open the Report Reader.
- 11.6.9 Maximize the **Report Reader** window. Scroll through the report and notice the information that is included in the report:
 - Account Number
 - Name
 - Address
 - Credit Limit

** Report Reader - Customer.txt			
File Edit Layers View Traps	Fields Help		
COMOX/TNS.	X (솔솔)루(x-x) ::: 알 볼 뿌(:::		
			< Teid Details
	Bright ID Customer Mast		Details
Account Number	- Name	Address	Credit Limit
A001	Dan Ackroyd	Audenshaw 125 New Street	20000
		Montreal Quebec H28 3H2	
A123	Mike Atsil	The Vetinary House 123 Dog Row Thunder Bay Ontario	20000
A128	Ivan Aker	KSA 7G1 The Old House Ottawa	10000
2001	Tin Danianan	Ontario P1D 8D4	10000
8001	Kim Basinger	Mesh House Fish Street Rouyn Quebec	12000
Ready			In , Col Page:



Technical Notes

The **Report Reader** uses a set of layers to define the data to be imported into IDEA. The first step is to identify the **Base Layer**. The **Base Layer** is the line(s) that contains the lowest level of detail in the report and is often the transaction row in the report. It is very important to select the line from the first record in the report to ensure that all records are duly selected.

WARNING: The **Report Reader** is a top-down processor; therefore, any information that precedes the defining record in the report is ignored. This feature is helpful when very detailed header

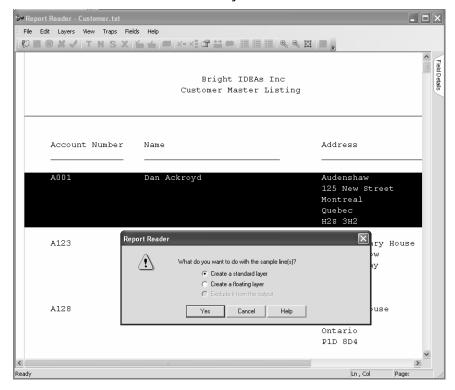
information is included in the report and we want that data excluded from the database in IDEA.

Once the **Base Layer** is defined, **Append Layers** that provide related information to the **Base Layer** can also be defined. The **Base Layer** and **Append Layers** are identified by **Traps**.

Traps can be Text, Numeric, Space, Non-Blank, Floating, or actual words that uniquely identify each occurrence of the record in the report. Keep the Traps as simple as possible. Do not overlook the possibility that spaces or spaces combined with numbers or text could be the best Traps.

Define a **Base Layer**:

- 11.6.11 Scroll through the report and identify the line(s) that contain the lowest level of detailed information and highlight the lines with the cursor (click and drag). Since the address information is a multiple line field, the **Base Layer** will contain 5 lines.
- 11.6.11 Select Create a standard layer, and then click Yes.



The **Report Reader** will display a **Field Editor** (the section between the two yellow lines) where the fields that need to be included in the database are defined. Scroll through the report and notice each section of data begins with an account number. The best **Trap** for the **Base Layer** is a combination of text and numbers, based on the account number (**TNNN**).

11.6.12 Place the cursor in the Anchor Editor area above the account number in the Field Editor. On the Report Reader toolbar, click

the Text Trap T button once and click the Numeric Trap N button three times.

÷,~	Report Reader - Customer.txt*	•				
	lle Edit Layers View Traps					
	🖩 🖲 🎗 🗸 T N S 🛛	X 🖌 🖆 💻 x- x] 🖷 🔛 🖙 🗏				
	TNNN				Field Details	ņ
Þ	A001	Dan Ackroyd	Audenshaw	^	24	
			125 New Street			
			Montreal			
			Quebec H2S 3H2			
			H25 3H2	~		
				^		
		Bright ID	EAS Inc			
		Customer Mast				
Ŀŀ				-		
11						
	Account Number	Name	Address			
L L						
Ш	A001	Dan Ackroyd	Audenshaw			
Ш			125 New Street			
Ш			Montreal Ouebec			
Ш			Uuebec H2s 3H2			
Ľ			120 382			
	A123	Mike Atsil	The Vetinary House			
			123 Dog Row			
			Thunder Bay			
			Ontario	¥		
<			¥24 701 >			
Read	ły				Ln , Col F	Page:

11.6.13 Notice each section of data that begins with an account number is highlighted. Scroll through the report to make sure all the information is highlighted.

Identify Field Anchors:

Selecting the information that will become the data fields in the database is done by identifying **Field Anchors**. Using the cursor, highlight each **Field Anchor** contained in the **Base Layer**. Scroll to the right to include the values for credit limit. Make sure the widths are large enough to include information that may be slightly off position or extremely long. Watch for items that appear as spaces in the **Base Layer** line, but contain values in other lines.

🖅 Report Reader	- Customer.txt*					- D X
File Edit Layer	s View Traps Field	s Help				
	TNSX	<u>4</u> = x-x = = ≈ ≈ = = = ≈ ≈				
TNNN					Field Details	ą.
				_	₽ 2↓	
A001		Dan Ackroyd				
			123 14600 2010000		Layer Information Layer Type	Standard
			Montreal		Layer Name	Laver-1
			Quebec		 Field Informal 	
			H2s 3H2	~	Name	AUDENSHAW
					Туре	Character
				^	Decimals	(none)
					Implied Decimals	
		Bright IDEAs Inc			Mask	(none)
		Customer Master Listing	3		Position	
					Offset	56
					Display Width Actual Width	0
					- Attributes	30
					Blank Cells	Leave blank
					Multi-line	Yes
ACCO	unt Number	Name	Address		End Field On	Specify line length
					Line Length	5
					Filters	
A001		Dan Ackroyd	Audenshaw		Include/Exclude	
			125 New Street		Range	All
			Montreal			
			Ouebec			
			H2s 3H2			
			125 312			
				Ì		
A123		Mike Atsil	The Vetinary House			
			123 Dog Row			
			Thunder Bay			
			Ontario			
				~		
15)			>			
Ready					Ln , Col	Page:
	lf a mis Remove Field Ar	al Notes stake is made while e Field Anchor hchor or right-click the F soth actions remove the ed.	tton on the toolba Field Anchor to di	ir İsp	to remo play a s	ove the shortcut

Now that each **Field Anchor** in the **Base Layer** has been highlighted, it is necessary to name and detail each field. This is done using the **Field Details** window.

11.6.14 Click the Save Layer 💙 button on the toolbar.

- 11.6.15 The Field Details window appears that contains the data values that were in the Base Layer. This information will need to be changed. The names that originally appear (for or in?) the data in the Field Details window will be the Field Anchor values of the data in the database. Notice that the fields are listed in the order that they were highlighted in the Base Layer. The order of the fields is easily changed once the data is imported into IDEA.
- 11.6.16 Use the following chart to update the **Base Layer** information. Select each field and change the name and details.

Field as shown in Field Details	Name	Data Type	Offset	Actual Width
A001	ACCOUNT_NO	Character	5	4
DAN_ACKROYD	CUST_NAME	Character	23	28
AUDENSHAW	ADDRESS	Character	57	33
20000	CREDIT_LIM	Numeric	95	7

ADDRESS is a multi-line field. Make sure that the multi-line field option is selected and that the field is 5 lines long.

11.6.17 In the Layer Name box, delete the default layer name (Layer-1) and enter Customer.

The Field Details window should look like the following example.

Fie	eld Details	. P				
	. ≜ ∔					
Ē	Layer Informal	tion				
	Layer Type	Standard				
	Layer Name	Customer				
	Field Informati	on				
	Name	ACCOUNT_NO				
	Туре	Character				
	Decimals	(none)				
	Implied Decimals	(none)				
	Mask	(none)				
Ξ	Position					
	Offset	5				
	Display Width	0				
	Actual Width	4				
	Attributes					
	Blank Cells	Leave blank				
	Multi-line	No				
	End Field On	Blank line				
	Line Length	1				
	Filters					
	Include/Exclude	All				
	Range	All				
Ļ						
Layer Name						
Ni	Name for this layer.					

Since all the required information can be defined in the **Base Layer**, **Append Layers** will not be defined for this exercise.

Importing into IDEA:

11.6.18 Before importing the file into IDEA, it is recommended that we preview the entire database. Once the **Base Layer** has been defined, the IDEA database can be previewed at any time by clicking

the **Preview Database** button on the toolbar. The preview should look like the following example.

	ACCOUNT_NO	CUST_NAME		1	DDRESS 🔥	Close
1	A001	Dan Ackroyd	Audenshaw	125 New Street	Montreal	Refres
2	A123	Mike Atsil	The Vetinary House K3A 7G1	123 Dog Row	Thunder E	nellesi
3	A128	Ivan Aker	The Old House	Ottawa	Ontario	
4	B001	Kim Basinger	Mesh House	Fish Street	Rouyn	
5	B002	Richard Burton	Eagle Castle	Leafy Lane	Sudbury	
6	B004	Jeff Bridges	Arrow Road North 2S1	Lakeside	Kenora	
7	B008	Denise Bent	The Dance Studio H5S 3H2	Covent Garden	Montreal	
8	B010	Carter Bout	Removals Close M2A 7D3	No Fixed Abode Roa	d Toront	
9	B022	Ronnie Biggs	Gotaway Cottage	Thunder Bay	Ontario	
10	C001	Tom Cruise	The Firm 6H5	Gunnersbury	Waskaganish	
11	C003	John Candy	The Sweet Shop 3A1	High Street	Trois Rivieres	
12	C004	Jamie Lee Curtis	Wanda Avenue H4S 3H2	Welham Green	Montreal	
<	101				>	

11.6.19 Before importing the file into IDEA, it is also recommended that we run the Scan for Errors task. Scan for Errors identifies any items that may not import properly such as fields where the length is too short, fields where there are excessive amounts of space, and fields that are not in alignment with the preceding fields. Click the Scan for Errors button on the toolbar.

🖗 Report Reader - Customer.txt*				- OX
File Edit Layers View Traps Field	s Heb ▲ ☑ 曰 ו ×: ☞ 註 ● 圖 圖 圖 Bright IDEAs Ir		Field Details	ų tion
	Customer Master Lis	sting	Layer Type Layer Name Field Informati Name Type Decimals	ACCOUNT_NO Character (none)
Account Number	Name	Address	Implied Decimals Mask	(none)
A001	Dan Ackt	n errors in this report. Percent and the second an	Offset Display Width Actual Width = Attributes Blank Cells Multi-line End Field On Line Length	5 0 4 Leave blank No Blank line 1
A123	Mike Atsil	The Vetinary H 123 Dog Row Thunder Bay Ontario K3A 7G1	Filters Include/Exclude Range	All
A128	Ivan Aker	The Old House Ottawa Ontario P1D 8D4	Layer Name Name for this layer.	
< Ready		>	Ln , Col	Page: //

- 11.6.20 Once the Scan for Errors validation is complete, save the Report Reader template for future use with similar reports. Select File – Save Template As. Give the template a name and save it to the Import Definitions sub-folder under the project folder. The Import Definitions is the preferred location unless the templates are shared items and need to be stored elsewhere.
- 11.6.21 Click the Import into IDEA State button on the toolbar.
- 11.6.22 Click Yes to proceed with the import.
- 11.6.23 Select the Generate field statistics check box.
- 11.6.24 Name the database Customer Master and then click Finish.

Chapter 12 Working with Application Softwares

12.1 Introduction

CAAT – IDEA is a robust, time-tested, stable, data analytical tool. IDEA has evolved steadily to integrate conveniently and seamlessly with well-known Application Softwares worldwide.

IDEA connects well with simple user-friendly Accounting Packages like Tally, ACCPAC, SAGE and Peach Tree. It also gels well with more complex Applications like ERP Applications – SAP, and Oracle. In the Banking Sector, IDEA mixes well with Core Banking Applications like Finacle, Flex Cube and FNS. In the General Insurance sector, IDEA serves well with Applications like Genisys.

The sections that follow illustrate through actual experiences the import and analysis/interrogation of data in diverse Application environments and setups.

We have depicted connectivity to

- SAP and SAP Flat Files from the Manufacturing Industry
- Core Banking System Finacle from the Banking Industry
- Web Based Application Systems Sabre from the Airlines Industry

12.2 Core Banking System – Finacle from the Banking Industry

12.2.1 Import of Data

Set-up of ODBC environment for linking IDEA to Oracle database for FINACLE Core Banking System (CBS)

• Install IDEA V 9 Electronic License on the designated client machine.

- With the help of the CBS implementation partner, load the relevant Microsoft for Oracle client drivers on the above machine. Note the drivers should match with the Oracle database version
- Update the Internet Protocol Address and Port Number of the destination server in the "TNS Ora" file on the client machine.
- Create an ODBC Data Source (DSN) to the Flash, Training or DR site server as decided. In this case, we connect to the training server 'testbanktraining'.
- To create the DSN, name the DSN connection, input the server name correctly, we select the relevant drivers and enter the user access ID and password for gaining entry to the server.
- Once the DSN has been created, it will appear in the list of 'Select an ODBC driver that is configured to connect to a specific data source:' within IDEA > Import Assistant > Import into IDEA > ODBC.
- In this case, our user access ID is 'testbankuser', password is 'testuser123' and the training server is 'testbanktraining'. So 'testbanktraining' will appear in the list of 'Select an ODBC driver that is configured to connect to a specific data source:'
- Each time the user navigates through the above-mentioned DSN and by entering the password for access, the user can view (read only access) to all the authorized data tables in the server. The tables which the user can view are the ones to which the user has view right access.
- By selecting a specific data table, the user can navigate to 'Advanced' and choose to select specific fields to download while importing the data. The user can also choose to enter an SQL statement, so as to import data for a particular period, branch or some other parameter.

12.2.2 Interrogation of Data

For example, in Deposits –Term Deposit domain, we choose to filter cases where deposit period has exceeded a tenor of more than 120 months (Maximum Tenor of a Term Deposit being 120 Months as per Reserve Bank of India).

Here IDEA connects through the ODBC process mentioned above to the tables $\mbox{-}$

- GAM Account Masters
- TAM Term Deposit Masters.

The fields of reference relevant to the objective being tested which may be chosen through the Advanced Options in IDEA ODBC Import are

- GAM ACID (Account Number), FORACID (Customer ID)
- TAM ACID, Deposit_Period_Mths (Deposit Tenor), SOL_ID (Branch Number)

The process of interrogation followed in IDEA:

- Import GAM with the fields of reference.
- GAM has 14 Lac rows. The ODBC takes 5 seconds.
- Open GAM
- Write a criteria where SOL_ID="00015" where 00015 represents large Branch X.
- Save the criteria as a database within IDEA. Name it X
- Import TAM with the fields of reference.
- Open TAM
- Write a criteria where SOL_ID="00015" where 00015 represents large Branch X.
- Save the criteria as a database within IDEA.
- Perform a direct extraction on the above with the criteria Deposit_Period_mths>120.
- Save the database as Greater than 120
- Now join Greater than 120 and X based on matching field ACID, using the Join option "Matches only".
- The resultant database Filename Objective 4 has given us a list of instances of Term Deposit accounts accepted with tenor over 120 months.

12.3 Web Based Application Systems – Sabre from the Airlines Industry

12.3.1 Introduction

Nowadays we can book air-tickets very conveniently from our home or office. This is possible with the help of Internet based Air Reservation Application System, Sabre. In fact, advanced versions of the Internet Application allow users to check-in their seat through Mobile Short Messaging Service (SMS).

As an application system Sabre generates reports in an Extendible Mark Up (XML) format. The XML data format is a markup language that defines a set of rules for encoding documents in a format which is both humanreadable and machine-readable. Nowadays many applications are generating reports in these formats for easy portability and recognition across platforms and applications.

The Airline Reservation System generates Sales Reports and other relevant Flight Capacity Utilization Reports.

These Reports contain the following fields:

- Passenger First Name
- Passenger Last Name
- Address
- Contact Number
- Email
- Destination From
- Destination To
- Basic Web Fare
- Taxes and Levies
- Other Charges
- Payment Information and any other required information

12.3.2 Import of Data

The XML Data Files once saved on the workstation containing IDEA or on the Local Area Network in a shared audit data folder can be accessed through IDEA's Import Assistant AML component. The process of import is simple and easy to apply, since XML formats are readily recognized by IDEA just like importing a spreadsheet file.

12.3.3 Interrogation of IDEA

With the recent credit card fraud in the Airlines Industry where a noted carrier took a revenue hit of 18-20 Crores due to criminal usage of stolen credit cards through web-bookings, Airlines are now closely monitoring Passenger Personal particulars along with their Credit Card particulars.

In this case, the Passenger Personal particulars:

- First Name
- Middle Name
- Last Name

These are extracted from the Sales Account for a given period of flight reservations.

This information is matched with the Credit Card particulars:

- First Name
- Last Name

The match is done through a superior @function in IDEA titled @soundslike.

@soundslike throws up exceptions where names or character strings like addresses match each other not necessarily on syntax but on mnemonics (Sound and Rhyme).

The seemingly matching cases are extracted and taken up for substantive testing to check the bona fide of the cases. Special care is taken to identify group bookings, family bookings, geographical concentrations, repeat bookings, and incomplete names.

12.4 SAP R/3 and SAP Reports from the Manufacturing Industry

12.4.1 International Practice

In certain countries like Germany and Canada, advanced statutory measures have been put in place by the Ministry of Revenue for usage of CAATs with SAP ERP. Clients are mandated to generate reports from SAP - MM, FI, SD,

FA Modules in a specific format to meet specific Excise and Customs (Revenue) Audit objectives. This mandate provides a stable platform for the growth of CAATs.

12.4.2 National Practice

Manufacturing majors – IDEA Clientele using IDEA with SAP in specific areas follow the following approach to integrate it with SAP. The practice within India is similar to the practice outside India:

- 1. The Audit User can generate standard reports through his login ID from SAP.
- 2. These reports can be generated either from the standard report menu OR

In case the Company has procured and enables AIS – Audit Information System along with SAP (The audit information system (AIS) is an auditing tool that you can use to analyze security aspects of SAP NetWeaver Application Server (SAP NetWeaver AS) for ABAP system in detail.

AIS presents its information in the audit infomation structure (similar to the Implementation Guide) so that you can easily determine which activities you need to perform and which you have accomplished.

- 3. These reports once generated can be saved either as an unconverted file (Text File) or spreadsheet. It is recommended to save the file as an unconverted file if it is larger than a standard spreadsheet row length capacity of 1 million odd rows (Excel 2013).
- 4. Once saved, these text files or spreadsheet files can be conveniently imported into IDEA using IDEA's Import Assistant.

Merits of the Report Approach are:

- Easy to apply
- Capture meaningful data in report forms
- Generally sufficient to express management assurance on audit objectives

Demerits of the Report Approach are:

• Reports may not reflect the right data

 Reports are a personalized view of data; raw data is comprehensive and cannot be manipulated

12.4.3 Theoretical Approach: Complex and Seldom Followed

IDEA can be integrated directly with the SAP server provided

• The IT department is convinced on the software and non-integrity violation of data in the backend.

In one of our recent client acquisitions, a leading Engineering and Projects Company decided to opt for IDEA since it has a robust/stable SAP Report Import Component. The General Manager – IT in the subject Company was not for IDEA connecting to the Live or DR Server on grounds of security and confidentiality of production data. Given this concern, IDEA was implemented on Spreadsheet and Unconverted Report Files from SAP.

 The Audit user is knowledgeable about the data tables in the backend, and its several fields. SAP has over 60000 odd physical tables in its back-end. It is a challenge linking the right table for the right objective.

The following methods can be used to identify the tables within SAP

- Technical Information button within a transaction
- Where Used List for Data Elements
- Where Used List for the Domain
- Search on Fieldname / Partial Fieldname within table DD03L
- Search on Fieldname / Partial Fieldname within table DD08L
- SE16 fuzzy Search on Table Descriptions
- System Trace

For example: SAP Server - MM module contains tables like EKPO or EKKO which represent Header and Transaction detail tables for Ordering.

Once the integration is approved, the user can establish a link to the SAP Server through the middleware – ODBC, using the Oracle drivers for SAP.

The desktop having IDEA is linked to the server, with due access authorizations, tables viewed and subsequently imported into IDEA. This is the raw data being imported.

Merits of the Direct Link:

• Pure unaltered Data can be mined from the server.

Demerits of the Direct Link -

• To arrive at a logical representation of data, many data tables may have to be joined to arrive at a table having more or less all the logical fields necessary for achieving audit objectives. This will necessitate that the Audit User has a thorough knowledge of each and every table and field for file joins.

Chapter 13 Application of CAATs to General Insurance

13.1.1 Introduction

IDEA Data Analysis Software can play a very vital role in the Insurance industry in delivering enhanced, effective and value added results in a significantly short span of audit execution time.

With significant de-regulation in the Insurance sector, widespread availability of information through integrated insurance application systems and heightened awareness about Audit Software, concerted efforts need to be made to spread the use of the tool amongst insurance auditors, academicians and specialists like risk underwriters and loss prevention teams.

This research publication is an attempt to create awareness about the use and application of IDEA to realize simple, common, and generally accepted audit objectives of the General Insurance industry.

Some assumptions made during this project have been duly mentioned against each research objective. The publication is an attempt to generalize the objectives with reference to common practices across the globe.

Objectives have been established for the main streams of activity covering premium and claims. Analysis of objectives and financial tests have also been explored.

In Premiums, the objectives dealt with include establishing the accuracy of premiums collected, authorizing discounts offered, validating manual policies prior to system integration, justifying policies canceled shortly after inception and accuracy of penalties levied in case of repetitive past claim history.

In Claims, objectives dealt with include establishing the validity of claims, validating close proximity claims, justifying operational delays in claim settlement, accuracy of agency commission and settlement of claims on policies underwriting risks for stolen motor vehicles.

Certain aspects of analysis include rolling comparison of incurred losses over three years to identify cases for premium enhancement in case of repetitive

claim history. Development officer business solicitation practices have also been studied in case of high loss policies over time.

Finally, certain financial tests have been undertaken to establish the accuracy and validity of the system of balances and postings between sub ledgers and general ledger control accounts.

13.1.2 Potential Audit Tests

INSURANCE PREMIUM

The main tests are concerned with ensuring that all premiums are raised and the correct rate is used. However, analysis of risk and exception tests of old or forward dated items are useful.

Typical tests include:

Calculations

Recalculate premiums

Prove posting totals

Analysis

Stratify premiums by size and product type

Summarize new business by agent or area

Summarize debts by agent

Exception Tests

Extract old premiums due on policies that have not lapsed

Identify policies that were cancelled soon after inception

Identify policies with initial discounts that have then been lapsed and credited

Extract negative debts

Identify unallocated cash

Extract policies with blank or invalid policy numbers

Identify policy premiums that have only been part paid

Direct Debits that have not been collected

INSURANCE CLAIMS

Claims are a classic payments area. Claims can be matched back to policy

details and there are usually various claim types, reason codes and risk codes.

Calculations

Cast (or total) outstanding claims

Prove posting totals

Analysis

Analyze by month and type and calculate average values

Group claims by product type

Matrix analyze date of incident to date claim reported as part of assessing claims incurred but not recorded

Exceptions Tests

High value claims

Identify claims with no movement during the last 6 months (some claims get reported but queries on the claim are not responded to leaving old expired claims on the system)

Identify negative balances (recoveries)

Gaps and Duplicates

Check for duplicate claims (same policy number and either same amount or same date of incident)

Check for multiple claims on the same policy (some policy holders are particularly unlucky; however, others make a business out of claiming more than once for an incident. This can be picked by testing claims for the same amount or the same date of incident)

REASONABLENESS TESTS ACROSS FILES

Existence of a policy in the same time period as the claim

Checking details between claims and policies

New policies experience claims shortly after inception

HEALTHCARE FRAUDS

Excessive procedure billing of same diagnosis, same procedures

Identify excessive number of procedures per day or place of service per day / per patient

Identify multiple billings for same procedures, same date of service

Analyze for mismatched services to diagnosis codes

Review diagnostic fees in excess of masters

Review procedure fees in excess of masters

Identify and report diagnosis and / or treatment that is inconsistent with a patient's age and / or gender

13.1.3 Obtain the Data

The following data files for the period April 1, 2005 – Oct 31, 2005 will be used for this exercise:

1	Premium Report for Motor Class (Tariff)	Spreadsheet Excel File
2	Premium Report for Earthquake Class (Tariff)	Spreadsheet Excel File
3	Premium Report for Motor Class (Non- Tariff)	Spreadsheet Excel File
4	Premium Report for Motor Class (Non- System)	Text File
5	Canceled Policy Report for Motor Class	Spreadsheet Excel File
6	Cover Note Report for Motor Class	Spreadsheet Excel File
7	Premium Master for Motor Class	Dbase File
8	Premium Master for Earthquake Class	Dbase File
9	Discount Master for Motor Class	Dbase File
10	Policy Number Report for Rolling Policies Motor Class	Spreadsheet Excel File
11	Agent Commission Report for Motor Class	Spreadsheet Excel File
12	Commission Master for Motor Class	Dbase File
13	Claim Report for Motor Class	Text File
14	Stolen Vehicles Report from the National Crime Records Bureau (NCRB)	Text File
15	Surveyor Activity Report for Motor Class	Spreadsheet Excel File
16	Incurred Loss Summary Report for 2005	Spreadsheet Excel File
17	Incurred Loss Summary Report for 2004	Spreadsheet Excel File
18	Incurred Loss Summary Report for 2003	Spreadsheet Excel File

Application of CAATs to Utility Se	ervices
------------------------------------	---------

19	Incurred Loss Detail Report for 2005	Spreadsheet Excel File
20	Business Performance Report for 2005	Spreadsheet Excel File
21	Incurred Loss Report Agent wise for 2005	Spreadsheet Excel File
22	Sub Ledger for Cash Deposit Account	Dbase File
23	General Ledger for Cash Deposit Account	Dbase File
24	Healthcare Report on Healthcare Claims	Spreadsheet Excel File
25	Healthcare Master for Healthcare Claims	Spreadsheet Excel File

The reports cited above have been generated from the Insurance Application Software - M.I.S. function and have been captured in different data file formats.

13.1.4 Audit Program

Test Ref	Auditing Procedures	Objective	Page Ref
	Premium		
1.0	Re-verification of premium collected for insurance against natural disasters like earthquakes.	Revenue Integrity	
2.0	Re-verification of premium collected for insurance against motor claims.	Revenue Integrity	
3.0	Re-verification of discretionary discount allowed on premiums in the non-tariff free market insurance segment		
4.0	Verification of rating fields for special additional impose / penalty incase of frequent claim history on sensitive policies.		
5.0	Input field checks on non-system generated or manual policies	Validity	
6.0	Identifying policies canceled soon after inception along with justification by endorsements.	Validity	
7.0	Input field checks on cover notes (temporary policies) for motor insurance and conversion to bona fide policies through receipts.	•	

Test Ref	Auditing Procedures	Objective	Page Ref
	Claim		
1.0	Identification of duplicate motor claims in partial loss policies	Validity	
2.0	Identification of duplicate motor claims due to irregularities and discrepancies in survey	Validity & Fraud Detection	
3.0	Isolation of policies experiencing claims shortly within risk commencement – close proximity		
4.0	Identification of claims without bona fide policy numbers within the system – orphan claims	Validity	
5.0	As part of operational studies and work flows identification of irregular time lags in claim settlement.		
6.0	Re -verification of agency commission on business solicited	Payments Integrity	
7.0	Identification of motor insurance business claims settled on policies pertaining to stolen vehicles		
	Analysis		
1.0	Analysis by client for different class of risks insured	Business Studies	
2.0	Analysis by development officer and / or agent of insured losses	Business Studies	
3.0	Identification of premium enhancement cases for loss making and high risk policies	Business Studies	
	Finance & Accounts		
1.0	Reconciliation of Sub Ledger and General Ledger for principal control accounts	Integrity of Accounts	
2.0	Identification of un adjusted credits in Cash Deposit Account	Integrity of Accounts	

Application of CAATs to Utility Services

Test Ref	Auditing Procedures	Objective	Page Ref
	Healthcare Frauds		
1.0	Excessive procedure billing of same diagnosis, same procedures		
2.0	Identify excessive number of procedures per day or place of service per day / per patient		
3.0	Identify multiple billings for same procedures, same dates of service		
4.0	Analyze for mismatched services to diagnosis codes	Validity & Fraud Detection	
5.0	Analyze for mismatched services to diagnosis codes	Delection	
6.0	Review diagnostic fees in excess of masters		
7.0	Review procedure fees in excess of masters		
8.0	Identify and report diagnosis and / or treatment that is inconsistent with patient's age and / or gender		

PROJECT SET-UP

LOAD DATA

OBJECTIVE

To be able to load data, run IDEA 9 and set a projct folder for the audit within $\ensuremath{\mathsf{IDEA}}$

EXERCISE DESCRIPTION

This exercise covers copying data into a folder, loading IDEA, setting a project folder and entering client properties that will be printed on all reports.

IDEA FUNCTIONALITY COVERED

- Setting a Project Folder
- Entering Client Properties

The following data files have been copied on the hard disk in the folder c:\Program Files\IDEA\Insurance Research

Control totals have also been provided for each of the files.	Control totals have	e also been	provided for	each of the files.
---	---------------------	-------------	--------------	--------------------

1	Premium_Source – Premium – Motor. xls	Spreadsheet Exce	I Sum Insured – 98,70,000
			Net Premium – 2,17,464.37
2	Premium_Source – Premium –	Spreadsheet Exce File	5,82,95,000
	Earthquake. xls		Net Premium – 12,96,162.5
3	Premium_Source – Premium – Motor Non	Spreadsheet Exce File	I Sum Insured – 98,70,000
	Tariff. xls		Additional Excess Impose – 1000
			Net Premium – 2,18,464.37
4	Non system policies. txt	Text File	Sum Insured – 98,70,000
			Net Premium – 2,17,464.37
5	Premium_Source – Canceled Policies. xls	Spreadsheet Exce File	I Sum Insured – 98,70,000
			Net Premium – 2,18,464.37
6	Premium_Source – Cover Note. xls	Spreadsheet Exce File	I Sum Insured – 98,70,000
			Net Premium – 2,18,464.37
7	Motor Premium Master. dbf	Dbase File	No. of Master Records – 6
8	Earthquake Premium Masters. dbf	Dbase File	No. of Master Records – 7
9	Discount Master. dbf	Dbase File	No. of Master Records – 6

10Premium_Source-SpreRolling Policies. xlsFile	eadsheet Excel No. of Records – 4
11 Premium_Source – Spre Commission. xls File	eadsheet Excel Net Premium – 2,17,464.37 Gross Commission – 9,366.03
12 Commission Master. Dba dbf	se File No. of Master Records – 2
13 Claim Transaction. txt Tex	t File Estimated Loss Amount – 3,29,19,750 Estimated Expense Amount – 4,40,000 Paid Loss Amount – 2,46,89,812.50 Paid Expense Amount – 2,42,000
14 Stolen Vehicles.txt Tex	t File Number of Records – 10
15 Claim_Source – Spre Surveyor Activity. xls File	eadsheet Excel Number of Records - 19
16 Analysis – Incurred Spro Loss 05. xls File	eadsheet Excel Premium – 2,17,464.38 Claim – 8,59,980 Loss – 7,74,080
17 Analysis – Incurred Spre Loss 04. xls File	eadsheet Excel Premium – 2,17,464.38
	Claim- 1,33,400 Loss – 61,112.5
18 Analysis – Incurred Spro Loss 03. xls File	

Application of CAATs to Utility Services

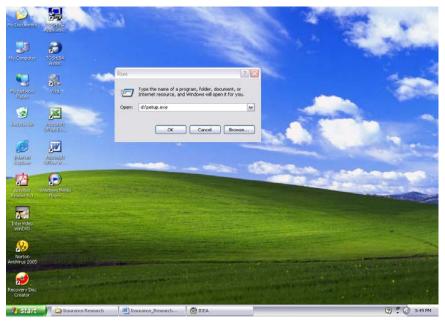
-		1	
	Loss Detail 05. xls	File	6,17,808.75 Claim – 25,04,940 Loss – 22,49,115
20	Analysis - Business. xls	Spreadsheet Excel File	Premium – 6,17,808.75 Claim – 25,04,940 Loss – 22,49,115
21	Analysis – Incurred Loss Agent. xls	Spreadsheet Excel File	Sum Insured – 98,70,000 Net Premium – 2,17,464.37
22	Sub Ledger CD. dbf	Dbase File	Debit - 2,57,500 Credit - 8,50,000 Balance - 63,42,500
23	General Ledger CD. dbf	Dbase File	Opening Balance – 23,43,500 Debit – 2,57,500 Credit – 8,50,000 Closing Balance – 29,36,000
24	Healthcare Fraud Source Files – Healthcare. xls	Spreadsheet Excel File	Diagnostic Fees – 14,550 Procedure Fees – 10,40,000
25	Healthcare Fraud Source Files – Masters. xls	Spreadsheet Excel File	Diagnostic Fees – 5,250 Procedure Fees – 3,65,000

Throughout this Chapter we refer to the files located in c:\Program Files\IDEA\Insurance Research.

If the SETUP.EXE file has been run earlier, then there is no requirement to run it again.

Place the compact disk into the CD reader.

- Click the Start button
- Select Run
- Type D:\SETUP.EXE. (D: being the letter assigned to your CD reader)



The files are installed into the proper directory.

Accessing IDEA

From the Windows START Menu select: Programs - IDEA - IDEA



Alternatively, click the **IDEA** shortcut icon on the desktop, if available. IDEA will load.

Select a Working Folder and Enter Client Information

To facilitate housekeeping, it is recommended that a separate project folder be used for each audit / investigation.

All information relating to the audit, including data files, equations, views / report definitions import definitions must be stored in the project folder.

The project folder and client information will be printed on all reports. Once the project folder is set, it remains the active folder until changed.

1. To set the project folder, click the **Create** button under Home in the IDEA menu tool-bar:

Imaged Faces

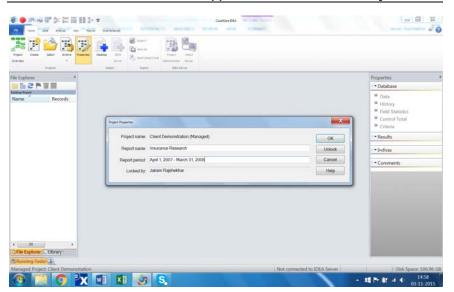
Imaged Face

Enter Client Demonstration as the Managed Project Name

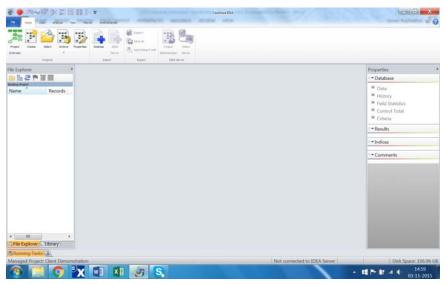
2. Click on **Properties** under Home in the IDEA menu tool-bar and enter the Report Name and Report Period:

Report Name:	Insurance Research
Report Period:	April 1, 2007 – March 31, 2008

Application of CAATs to Utility Services



Click the OK button. Client Demonstration will become the Managed Project Folder, as is seen from the working folder path at the left bottom corner of the IDEA screen.



13.1.5 Import of Data Files

OBJECTIVE:

To import data files for testing.

EXERCISE DESCRIPTION:

Three categories of data file types need to be imported: Spreadsheet Excel, Dbase and Text. Spreadsheet Excel and Dbase files can be directly imported into IDEA. Text files can be imported through the Report Reader, where a text file in report form is converted into a spreadsheet file in an array of rows and columns within IDEA.

IDEA FUNCTIONALITY COVERED

This exercise will cover:

- Import of Spreadsheet Excel files
- Import of Dbase files.
- Import of text files.

IMPORT OF SPREADSHEET EXCEL FILES

Access the Import Assistant by clicking the **Desktop** icon under Home in the IDEA menu tool-bar.

The Import Assistant loads to guide us through the process of importing the data.

Select Microsoft Excel.

Click Finish

Application of CAATs to Utility Services

Pasen (Sect	Select the format:	120000
plorer #	Advanced Record Definition Editor AS400	Properties
	dBASE	• Database
Records	Microsoft Access	Data
	ODBC	History
	Print Report and Adobe PDF SAP SmartExporter	Field Statistics
	-SAP/AIS	Control Total Criteria
	100 1391	- Criteria
		Resurs
		· • Indices
		· • Comments
	Computer location: My Computer	
	File name: C\Users\Jairam Rajshekhar\Documents\My IDEA Documents\JDE/	
	Optional definition file:	
	Edit template	
	< Book Next > Cancel Help	
6		
m		
Explorer Ubrary		

Click the Input file name box.

This takes us to the Source Files folder.

Select the MS-Excel file Analysis, and click Open.

ane Records	anize - IDEA Proj anize - New foi avorites Desktop Recent Places	ilder [Client Demonstration + Source Files.1.B +			thes tabase
1日2日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日日	avorites Desktop	î (tabase
me Records	Desktop			Anapoe by		
	Recent Places		FOULDE PRESIDER	Conserved to a first	Folder *	ata atory
	OneDrive		lame	Date m	sodified	old Statistics antrol Total
	Downloads		MIS	02-11-	2015 15:48	eteria
	L'omnour I		Analysis.ds		2005 16:26	a/ts
(m)	ibraries		Billing.xls	28-01-	2013 16:06	
	Documents		Claim_Source.ds	27-12-	2005 16:26	ces
a de la compansión de la c	Music		CountryEmissions.xls	07-12-	2012 19:08	mments
	Pictures	1	deposits.xls	27-04-	2009 12:51	
	Videos	1	Financial_Tests.als	27-12-	2005 16:27	
			Healthcare Fraud Source Files.ds		2005 14:40	
	Iomegroup	1.1	💼 itemmast.xls	20-08-2012 08:52		
	File	name:	Analysis.ds	Microsoft Excel files	(talso ta 💌	
				Open 💌	Cancel	
					_	

The preview of the first worksheet in the file Analysis appears in the **Excel** box.

Check the box **First row is field names**, to enable the transaction headers in Excel to get captured in the IDEA headers.

Click OK.

nouro p	Mensel free 8			Properties
	Select sheets to import:		OK	Database
a Property	V Tide Clients	Eist row is field names	Include All	I Data
ne Records	Incurred Loss 03	Import empty numeric cells as 0	Qear All	History
	Incurred Loss Agent		Cancel	Field Statistics
	Lincurried Loss of			Control Total
			Help	Criteria
	Preview: POUCY.NO CLAID. OF JEST CUST.JD	INSURD NA. MEMON QUAN U	010	Results
	1 REFERENCES Manage Durn D AL	A & Car \$1500.000 \$15700	-21396.3	· · Indices
	2 Address Mater Own D., 81 3 Address Mater Own D., Cl	C & Ca 23750.000 0	64	- monoes
	4 ########### Mutter Own D. 11 5 ############## Mutter Own D. 11	D & Ce 1125-000 0 E & Ce 12000-000 1500	64 19863	- Comments
	Autor Dun D., F1 Autor Dun D., F1 Autor Dun D., G1	F & Co 3000.000 13700 G & Co 29950.000 25000	-13780.0 3950.0	
	B REPORTED Manue Own D., HE	H & Ce 1300.000 0	84	
	3 PROBABBBBB Mintor Own D., R 10 REPRESENTED Montor Own D., K1	16 Co 1500.000 225000 2.6 Co 2175.000 250000	-223980-0 -3470-25-0	
	10 Annexament Matter Own D., K1 11 Annexament Matter Theat., 11	1 & Ca 2375.000 350000 1 & Ca 30000.000 8	-365555	
		ishekhar\Documents\My IDEA Documen		
		pheknar/Documents/My IDEA Documen	SUDEA PIC	
	Output file name: Analysis			

The MS-Excel file once imported appears as below.

The name of the file in IDEA is Analysis – Business. IMD, Analysis is the name of the Excel file and Business is the name of the worksheet.

The Laplow 	National Construction		and the second s	(Creat)						
Nume Records 1 Manual Manu	ile Explorer	Analysis-1	lide Clie *						• *	Properties
Nume Records 1 Manual Manu	1 - 2 = 10 III	POLICY_NO	CLASS_OF_RIS	KCUST.	DINSURED_NA	ME PREMIUM	CLAIM	LOSS		
Nume Records 2 Massimon Machenhamp 0 8.6.2 L3286 0 8.6 Avalysis-LL 57 2 Massimon Machenhamp 0 6.6.2 L3286 0 8.6 3 Avalysis-LL 57 4 Massimon Machenhamp 0 6.6.2 L3286 0 8.6 3 Avalysis-LL 57 4 Massimon Machenhamp 0 6.6.2 L3286 6 8.6 3 Massimon Machenhamp 0 6.6.2 L3286 6 8.6 4 Machenhamp 0 6.6.2 L3286 6 8.6 6 8.6 3 Machenhamp 0 6.6.2 L3286 6 8.6 6 8.6 1 Machenhamp 0 6.6.2 L3286 6 8.6 8.6 9.6 8.6 1 Machenhamp 0 6.6.2 L3286 6 8.6 <t< td=""><td></td><td>1 101001300500000</td><td>Motor Dwn Damage</td><td></td><td></td><td>12,500.000</td><td>15700</td><td>-23,200.8</td><td></td><td></td></t<>		1 101001300500000	Motor Dwn Damage			12,500.000	15700	-23,200.8		
 Analysis-TL 57 4 80000000 Matrix brance 0 0 6 6 6 0 000 000 000 000 000 000 00							0			 Data
1 1811100000 Machemberge 0 0.6.0 100000 100000 100000 1000000 1000000 1000000 1000000 10000000 10000000 10000000 10000000 10000000 100000000 100000000 100000000 1000000000000000000000000000000000000							0			Listens
1 10000000 Max 0.0<	Analysis-Ti_ 57						0.			
10 -100010000 Marchenberge 4 6.4 6.0 2000 3000 3000 6 10001000 Marchenberge 4 4.6 0 1.0000 3000 3000 10 10001000 Marchenberge 4 4.6 0 2.0000 3000 3000 10 100010000 Marchenberge 4 4.6 0 2.0000 9000 3000 11 100010000 Marchenberge 4 4.6 0 2.0000 9000 3000 11 100010000 Marchenberge 4 4.6 0 2.0000 900 3000 12 100010000 Marchenberge 4 4.6 0 2.0000 900 3000 13 100010000 Marchenberge 4 4.6 0 2.0000 900 4.000 14 100010000 Marchenberge 4 4.6 0 1.0000 900 4.000 900 15 100010000 Marchenberge 4 6.6 0 1.0000 900 4.000 900 4.000 900 4.000 900 4.000 <td< td=""><td rowspan="2"></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>Field Statistics</td></td<>										Field Statistics
0 10000000 Macheology 0 0.6 c 1.000 0 0.000 1 1000000 Macheology 0 0.6 c 1.000 0 0.000 1 1000000 Macheology 0 0.6 c 1.000 0 0.000 1 1000000 Macheology 0 0.6 c 1.000 0 0.000 1 1000000 Macheology 0 0.6 c 0.000 0 0.000 1 1000000 Macheology 0 0.6 c 0.000 0										Control Total
10 Hill Hill Hill Hill Hill Hill Hill Hill										
11 1001120000 Martinescher 4 6.5.0 N.8000 6 100 12 1001120000 Martinescher 4 6.4.5.0 4.1000 100 13 1001120000 Martinescher 6 6.5.0 100 100 14 100010000 Martinescher 6 6.5.0 100 100 15 100110000 Martinescher 6 6.5 11500 100 15 100110000 Martinescher 6 6.5 11500 100 16 100110000 Martinescher 6 6.5 11500 100 16 100110000 Martinescher 6 6.5 11500 100 17 100110000 Martinescher 6 <td< td=""><td rowspan="3"></td><td>9 1010013000009</td><td>Motor Own Damage</td><td></td><td></td><td>1,500,000</td><td></td><td></td><td></td><td> Criteria </td></td<>		9 1010013000009	Motor Own Damage			1,500,000				 Criteria
12 1182125980 Mon. You Fey 4 A 5 Co 11800 0 180 13 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 14 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 15 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 15 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 16 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 16 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 17 1182125980 Mon. You Fey 6 A 6 Co 1.000 1000 18 1182125980 Mon. You Fey 6 C Co 1.000 1000 18 1182125980 Mon. You Fey 6 C Co 1.000 1000 18 1182125980 Mon. Hour Fey 1 C Co 1.000 1000 18 1182125980 Mon. Hour Fey 1 C Co 1.000 1000 18 1182125980 Mon. Hour Fey 1 C Co 1.000 1000 18 1182125980							150000			Contraction of the Contraction o
13 11102139900 Mon Trout Pary 4 40.6.0 20.000 9000 41.000 14 11102139900 Mon Trout Pary 4 40.6.0 2000 1000 1000 15 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 15 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 15 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 40.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 6.6.0 1203 1000 1000 17 1100129900 Mon Trout Pary 6 6.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 6.6.0 1203 1000 1000 16 1100129900 Mon Trout Pary 6 6.6.0 1203 1000							6			 Results
14 100000000 Machinese 0 0.5.0 0.0000 10000 1000 10000 10										1
1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market A C A 1 INULINOUS Monormative market A C A Lation Monormative market										T Indices
10 1001120000 Munchmedragi 0 <										
10 1011210000 Number berging 0 0.4.5.2 123,82.0 0 0.8.0 10 1011210000 Number berging 0 0.4.5.2 123,800 0 0.8.0 10 1011210000 Number berging 0 0.4.0 123,800 0 0.8.0 10 1011210000 Numberson 0 0.4.0 123,800 100.0 100.0 Numberson 0 0.4.0 123,800 Numberson 0 0.4.0 12,800 Numberson 0 0.4.0 12,800 <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>No index</td>										No index
10 1001210000 Machineling 0 0.6.0 0.100 0 0.00 10 1001210000 Machineling 0 0.6.0 0.100 0.00 <td></td> <td></td> <td>Motor Third Party</td> <td></td> <td>R & Co</td> <td>12,796,875</td> <td>6</td> <td>0.00</td> <td></td> <td></td>			Motor Third Party		R & Co	12,796,875	6	0.00		
III IIIILIINUUM Muchimum AG IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII			Matur Third Party	11		787,500		8.00		Comments
Initialization Mundhemen R C-4 L508 0 6.0 22 INITIALISMON Mundhemen C C-5 2.10 6		19 1010017050000								
Mill Mill Col Col Col Dist Dist Dist 55 INTELLISTON Mill Col Col Dist Dis Dist Dist Dis		20 10100130500000					15790			Add comment
Mill Mill Col Col Col Dist Dist Dist 55 INTELLISTON Mill Col Col Dist Dis Dist Dist Dis		21 100018050000								
Mill Mill Col Col Col Dist Dist Dist 55 INTELLISTON Mill Col Col Dist Dis Dist Dist Dis		22 101001900000								
0 11002.150000 Monufacewa fi C+F 2,0000 1100 1100 0 1100.150000 Monufacewa fi C+C 2,0000 1000 22 1100.150000 Monufacewa fi C+C 2,0000 1000 22 1100.150000 Monufacewa fi C+C 2,0000 1000 22 1100.150000 Monufacewa fi C+C 2,0000 1000 23 1100.150000 Monufacewa 0 C+C 2,0000 1000 30 3100.150000 Monufacewa 0 C+C 2,0000 9 0.000 31 1000.150000 Fm M C+C 11,0000 9 0.000 31 1000.150000 Fm M C+C 11,0000 9 0.000 31 1000.1500000 Fm M C+C 11,0000 9 0.000 32 1000.1500000 Fm Fn <c< td=""> C+D 10,0</c<>		24 10.001850005					2506			
III IIII.LISING Multi-lisere G C-6 JU000 State 22 IIIII.LISING Multi-lisere 6 C-6 JU000 State 24 IIIII.LISING Multi-lisere 6 C-1 LISE LISE 25 IIIII.LISING Multi-lisere 6 C-1 LISE LISE 36 IIIII.LISING Multi-lisere 0 C-1 LISE LISE 36 IIIII.LISING Multi-lisere 0 C-1 LISE LISE 32 IIIIII.LISING Multi-lisere 0 C-1 LISE LISE 32 IIIIII.LISING Multi-lisere C-1 LISE LISE LISE 32 IIIIII.LISERE Multi-lisere C-1 LISE LISE LISE 33 IIIIII.LISERE Multi-lisere C-1 LISE LISE LISE 34 IIIIII.LISERE Multi-lisere C-1 LISE LISE LISE		26 1010018050000								
27 IIIII.ISDOB Michaemar # C.11 L.Date 20 LB 23 IIIII.ISDOB Michaemar # C.11 LD 20 LB 24 IIIII.ISDOB Michaemar # C.11 LD 20 LB 25 IIIII.ISDOB Michaemar # C.11 LD 20 LB 31 IIIII.ISDOB Michaemar # C.14 LD 90 40.05 31 IIIII.ISDOB Michaemar # C.41 LD 6 40 31 IIIII.ISDOB Michaemar # C.41 LD 6 40 31 IIIII.ISDOB Michaemar # C.41 LD 6 40 31 IIIII.ISDOB Michaemar # C.41 LD 60 40 40 31 IIIII.ISDOB Michaemar # C.44 LD 40 40 40 32 <td< td=""><td></td><td>26 10100100000</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>		26 10100100000								
28 IRELIDENDS Mundmense 8 C-1 LUBIS 2000 -123.08.0 29 IRELIDENDS Mundmense 6 C-1 2000 -123.08.0 31 IRELIDENDS Mundmense 6 C-1 40.05.0 6 1.00 32 IRELIDENDS Mundmense 6 C-1 40.05.0 6 1.00 33 IRELIDENDS Mundmense 6 C-1 40.05.0 6 1.00 33 IRELIDENDS Mundmense 6 C-1 40.05.0 6 1.00 10 33 IRELIDENDS Mundmense 6 C-1 40.05.0 6 1.00 11 34 IRELIDENDS Mundmense 6 C-1 40.05.0 1.00 1.410.0 11 34 IRELIDENDS Mundmense 6 C-1 40.05.0 1.00 1.410.0 11 35 IRELIDENDS Mundmense 6 C-1 40.05.0 1.		27 111/011050000								
23 1001120000 Feb Understand G C-1 2.558 0000 471C51 31 1001120000 Feb C C-1 2.558 0000 471C51 32 1001120000 Fe 46 C-1 42,000 60 32 1001120000 Fe 46 C-1 42,000 60 33 1001120000 Fe G C-1 42,000 600 4100120000 Fe G C-1 42,000 600 4000 51 1001120000 Fe G C-1 42,000 600 4000 51 1001120000 Fe G C-1 2000 600 4000 51 10011200000 Fe G C-2 12,000 7000 71250 51 10011400000 Fe G C-2 12,000 7000 71250		28 1010013050000								
1 100000000 Fm 40 C-1 40,0000 0 0.0 23 10000000000 Fm 46 C-54 40,0000 9000 4000 31 10001400000 Fm 46 C-54 40,000 9000 40,000 10 43 10001400000 Fm 46 C-54 40,000 9000 40,000 10 43 10001400000 Fm 41 C-0 L 2003 9000 40,000 10 43 10001400000 Fm 41 C-0 L 2003 9000 41,000 10 43 10001400000 Fm 41 C-0 L 2003 9000 41,000 10 43 10001400000 Fm 41 C-0 L 2004 9000 +12,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 1000 40,000 <td>29 10100138/500018</td> <td></td> <td></td> <td></td> <td></td> <td>150000</td> <td></td> <td></td> <td></td>		29 10100138/500018					150000			
32 INITIAL SHOREMONE Fac. 64. C, 64. ZL, 80000 9000 -41,404.00 33 INITIAL SHOREMONE Fac. C, 64. C, 50.										
13 Illinucacomos for 04 64.0 04.00 M 44 Statuset Setting Fit 6.0 4.400 Bit Diplome 14 G.0 4.400 9.00 4.400 File Explorer 1.00 4.400 9.00 4.500 4.400 File Explorer 1.00 4.400 9.00 4.500 4.000										
III + 54 Instructions (n_e n Go Lation None 433.05 File Explorer Lthorary 35 Instructions n Go Lation None 433.05 File Explorer Lthorary 35 Instructions n Go Lation None 433.05										
File Explorer LtDrany 25 36 Inscitements 0 C+P LtDsto 7/1253 File Explorer LtDstary 36 Inscitements File C+Q 12/3425 6 6.66 +										
			Fire							
Busines Tarks (2)	and the second s	36 101001340500007	Fire	14	C+Q	12,796,875	0	6.00	*	
Leaven 1 same 2	Running Tasks									

Invoke the Field Statistic Tab and view the Net Value of Premium, Claim and Loss fields.

A close comparison with the Control Totals, provided with the file, show that the totals match.

Now the file is ready for IDEA testing.

Project Charles Select Archive Arrendes Select Archive Arrendes Select	e Propries Dailto 203 IR bardians front	and and a second s					
le Explorer	Analysis-Tide Clie *						Properties
1 2 P G III	D 6 9 * 2 W # 0 9 + 8 *						Database
Also Proved	Field Type	Numeric	POLICY	PREMIUM	CLAIM	LOSS	
lame Records	Nameric	 Net Value 	5,35,30,61	6,17,808.7		-22,49,115.0	Data
Analysis-Ti_ 57		Absolute Val	5,35,30,61	6,17,808.7		23,29,815.0	 History Field Statistics
> PO > PR		# of Records	57	57	57	57	
	Numeric Fields	# of Zero Its	4	. 4	28	28	Control Total
	V POLICY_NO	Positive Value	5,35,30,61		25,04,940	40,350.0	Criteria
	CLAIM	Negative Val	0	0.000		-22,89,45	· Results
	V LOSS	# of Positive	53	53	29	0	
		# of Negative	0	0	0	23	* Indices
		# of Data Err	0	0	0	0	No index
		# of Valid Val	57	57	57	57	
		Average Value	9,39,13,35	10,838.750	43,946 32	-39,458.16	 Comments
		Minimum Val	0	0.000	0		Add comment
		Maximum Va	10,10,01,1	41,055.000	3,50,000	9,500.0	
		Record # of	54	54	2	10	
		Record # of	29	12	10	5	
		Sample Std	2,60,29,32	12,514.09	89,430.18	89,658.95	
		Sample Vari_	67,75,25,5	15,66,02,4		8,03,87,28	
		Pop Std Dev	2,57,99,98	12,403.83	88,642.23	99.688,88	
		Pop Variance	66,56,39,0	15,38,55,0	7,85,74,44	7,89,76,97	
		Pop Skewness	-3.365334	0.963654	2.548071	-2.603314	
	E	Pop Kurtosis	9.325472	-0.305850	5.429468	5.606579	
File Explorer Ubrary							

The process of import is the same for the other MS-Excel files intended to be used for various Insurance Audit Tests.

While undertaking the Insurance Audit Tests, only a passing reference will be made to the import of the MS-Excel files required for each of the tests.

IMPORT OF TEXT FILES

Access the Import Assistant by clicking the **Desktop** button under Home in the IDEA menu tool-bar

The Import Assistant guides us through the process of importing the data.

Click Print Report and Adobe PDF.

Select the text file to be imported from the Source Files folder by clicking on the select file browse box.

The **Select File** dialog opens, displaying the list of text files available in the working folder.

Select the file Claim Transaction.txt

Click Next.

The text file opens in the Report Reader software.

Data Analysis for Auditors

e Edit Layers View Traps							
SEG OF X Y TH	8 × 40	1981 # X+ XI # 1962 전	3 1 1 1 1 1 1 1	RQQ	81		1.0
Claim No	El No	NOL Dev Off	Code	Veh Regis	tration No. E	ngine No. Acci	dent
101701160502000	1	Total Loss	9	MH 01 1	E	01 1 31.03.05	
101701160502001	1	Partial Loss	99	MH 01 2	E	01 2 01.04.05	4
101701160502002	1	Total Loss	9999	MH 01 3	28	01 3 28.02.05	-
101701160502003	1	Total Loss	999	MH 01 4	E	01 4 20.02.05	
101701160502004	1	Total Loss	99	MH 01 5	85	01 5 01.04.05	
101701160502005	1	Partial Long	999	MH 01 6	E	01 6 01.05.05	
101701160502006	1	Total Loss	999	MH 01 7	Е	01 7 07,05.05	
101701160502007	1	Total Loss	999	MH 01 8	E	01 8 15.05.05	
101701160502008	1	Partial Loss	99	MH 01 9	E	01 9 01.06.05	
101701160502009	1	Total Loss	999	MH 01 10	E	01 10 17.05.05	
101701160502010	1	Partial Loss	99	MH 01 2	R.	01 2 18.05.05	
101701160502011	1	Total Loss	9	MH 01 12	E	01 12 01.06.05	
101701160502012	1	Partial Loss	99	MH 01 13	E	01 13 20.07.05	
101701160502013	1	Total Loss	99	MH 01 14	E	01 14 20.07.05	-
101701160502014	1	Partial Loss	9999	MH 01 15	E	01 15 20,07.05	
101701160502015	1	Total Loss	999	MH 01 16	12	01 16 15.07.05	
101701160502016	1	Partial Loss	9999	MH 01 15	R	01 15 20.07.05	
101701160502017	1	Total Loss	999	MH 01 10	82	01 18 23.08.05	1
101701160502018	1	Total Loss	9999	MH 01 19	E	01 19 01.06.05	1
101701160502019	1	Total Loss	999	MH 01 20	E	01 20 01.09.05	
101701160502020	1	Partial Loss	9	MH 01 21	E	01 21 15.09.05	
101701160502021	1	Partial Loss	999	MH 01 22	R	01 22 10.09.05	
101701160502022	1	Total Loss	99	MH 01 23	E	01 23 15,00.05	
101701160502023	1	Partial Loss	999	MH 01 6	80	01 6 20.09.05	
101701160502024	1	Total Loss	9	MH 01 25	E	01 25 01,07,05	- E
101701160502025	1	Total Loss	999	MH 01 26	E	01 26 15.10.05	1
101701160502026	1	Total Loss	9999	MH 01 27	10	01 27 28.06.05	÷.
101701160502027	1	Total Loss	9	MH 01 28	E	01 28 20.10.05	6

Select the first transaction in the text file as the base layer / standard layer.

Click Yes.

.

laim No B	1 No	NOL	Dev Off	Code	Veh	Re	gistration No.	312 1	ngin	• 1	No.	Acciden	e î
.01701160502000 1		Total	LOUS	9	1616	01	1	10	01	1	31.03.05		
01701160502001 1			al LOBB	99	мн				01		01.04.05		-
01701160502002 1		Total	LOBE	9999	MH			E	01	3	28.02.05		- 1
01701160502003 1		Total		999	MH			E	01	4	20.02.05		- 1
01701160502004 1		Total		99	MH	01	5	E	01	5	01.04.05	8 C	- 1
01701160502005 1		Partis	al Long	999	MH	01	6	в	01	6	01.05.05		- 41
01701160502006 1		Total	Long	999	MH	01	7	B	01	7	07,05.05	6 C	- 41
01701160502007 1		Total	Logg	999	MH	01	8	E	01	8	15.05.05	÷	1
01701160502008 1		P. Repor	t Reader					- ×	01	9	01,06,05		- 61
01701160502009 1		T						-	01	10	17.05.05		- 43
01701160502010 1		P. /	What	do you want to o	do with th	ie 240	ple line(s)?		01	2	18.05.05	1 C C C C C C C C C C C C C C C C C C C	- 41
01701160502011 1		т. 4	<u></u>	G Create a	standard	Launer			01	12	01.06.05		- 43
01701160502012 1		P		C Create a					01	13	20.07.05		- 81
01701160502013 1		т		C Dichade a	A frank the	100			01	1.4	20.07.05		- 41
01701160502014 1		P.				- 1			01	15	20.07.05	6 C	- 48
01701160502015 1		T.		Yes	Caricel		Help		01	16	15.07.05		- 41
01701160502016 1		PALLA	AL DORD	7777		U.A.	10	- 6	01	15	20.07.05	6 C	- 41
01701160502017 1		Total	Long	999	MH	01	18	R	01	1.0	23,08,05		1
01701160502018 1		Total	Long	9999	MH	01	19	B	01	19	01.06.05		- 11
01701160502019 1		Total	Long	999	MH	01	20	B	01	20	01.09.05		- 41
01701160502020 1		Partie	al Loss	9	MH	01	21	B	01	21	15.09.05		- 41
01701160502021 1		Partie	al Long	999	MH	01	22	E			10.09.05		- 61
01701160502022 1		Total	Long	99	1614	01	23	E	01	23	15.08.05	1	- 41
01701160502023 1		Partis	al Lonn	999	MH	01	6	R	01	6	20.09.05		- 63
01701160502024 1		Total	Long	9	2424	01	25	30	01	25	01.07.05		1
01701160502025 1		Total	Long	999	1424		26	31			15.10.05		:
01701160502026 1		Total	LOBB	9999	1614	01	27	R	01	27	28,06.05		1
01701160502027 1		Total	LOBB	9	1414	01	28	Б	01	28	20.10.05		- 6
												ol Pag	2

The first transaction is captured between two yellow lines called the Field Editor.

Application of CAATs to Utility Services

	Fields Help					
- II ® X / T N	S X 🖆	1 MB 🖉 🗵 🛪 🖬 🖾 🖲	2			
						Field Details
101701160502000	1	Total Loss	9	MH 01 1	E 01 1 🚖	RD 24
Claim No	81 No	NOL Dev Off	de de	Veh Registration No.	Engine	
	1 NO	Total Loss	0	MH 01 1	E 01 1	
101701160502000	1	Partial Loss	99	MH 01 2	E 01 2	
101701160502002	1	Total Loss	0000	MH 01 3	E 01 2	
101701160502002	-	Total Loss	9999	MH 01 4	E 01 3	
101701160502003	-	Total Loss Total Loss	999	MH 01 5	E 01 9	
101701160502004	-	Partial Loss	999	MH 01 6	CO (C) 7 - C	
101701160502005	1	Total Loss	999	MH 01 7	E 01 6 E 01 7	
101701160502006	1	Total Loss	999	MH 01 8		
101701160502007	-	Partial Loss	999			
	1	Total Loss	999			
101701160502009	1	Partial Loss	999			
101701160502010	1	Total Loss	99		E 01 2 E 01 1	
			99			
101701160502012	1	Partial Loss		MH 01 13	E 01 1	
101701160502013	1	Total Loss	99	MH 01 14	E 01 1	
101701160502014	1	Partial Loss		MH 01 15	E 01 1	
101701160502015	1	Total Loss	999	MH 01 16	E 01 1	
101701160502016	1	Partial Loss	9999	MH 01 15	E 01 1	
101701160502017	1	Total Loss	999	MH 01 18	E 01 1	
101701160502018	1	Total Loss	9999	MH 01 19	E 01 1	
101701160502019	1	Total Loss	999	MH 01 20	E 01 2	
101701160502020	1	Partial Loss	9	MH 01 21	E 01 2	
101701160502021	1	Partial Loss	999	MH 01 22	E 01 2:	
101701160502022	1	Total Loss	99	MH 01 23	E 01 2	
101701160502023	1	Partial Loss	999	MH 01 6	E 01 6	1
101701160502024	1	Total Loss	9	MH 01 25	E 01 2	
1012011/01/000005		m.a	0.0.0	MH 03 07		
sady						Ln 15, Col 43 Page: 1/1

Specific Traps (Numeric N, Text T, Space S or Special Character X) can be set depending upon the data which needs to be trapped for output into IDEA. The Trap is a mask that identifies the common character or space vertically amongst all the records and picks only those records that meet the criteria of the Trap.

In the file under import, we can place a Numeric N Trap over the first digit of the Claim No. field.

The Trap is set by clicking N in the menu tool-bar of the Report Reader.

The Trap is set in the space above the top yellow line called the Anchor Editor, and the Traps are anchored to a specific layer.

Upon setting the Numeric N Trap, the records which meet the Trap criteria are t highlighted in blue.

Data Analysis for Auditors

in Edit Layers View Traps	Pields Help	570					(m) (tP) [2
- III @ X - T N	S X C	(@) # (X+ XE 07 날 C					
N						Field Details	
101701160502000	1	Total Loss	9	MH 01 1	E 01 1 🚖	ED 21	
5.2					^		
Claim No 101701160502000	Sl No	NOL Dev Off Total Loss	Code	Veh Registration No. MH 01 1	Engine E 01 1		
101701160502000		Partial Loss	99	MH 01 2	E 01 2		
101701160502002		Total Long	9999	MH 01 3			
101701160502003		Total Loss	999	MH 01 4	E 01 4		
101701160502004		Total Loss	99	MH 01 5	E 01 5		
101701160502005		Partial Loss	999	MH 01 6	E 01 6		
101701160502006		Total Loss	999	MH 01 7	E 01 7		
101701160502007	1	Total Loss	999	MH 01 8	E 01 8		
101701160502008	1	Partial Loss	99	MH 01 9	E 01 9		
101701160502009	1	Total Loss	999	MH 01 10	E 01 1		
101701160502010	1	Partial Loss	99	MH 01 2	E 01 2		
101701160502011	1	Total Loss	9	MH 01 12	E 01 1:		
101701160502012	1	Partial Loss	99	MH 01 13	E 01 1		
101701160502013	1	Total Loss	99	MH 01 14	E 01 1		
101701160502014	1	Partial Loss	9999	MH 01 15	E 01 1		
101701160502015	1	Total Loss	999	MH 01 16	E 01 1		
101701160502016	1	Partial Loss	9999	MH 01 15	E 01 1		
101701160502017	1	Total Loss	999	MH 01 18	E 01 1		
101701160502018	1	Total Loss	9999	MH 01 19	E 01 1		
101701160502019	1	Total Loss	999	MH 01 20	E 01 2		
101701160502020	1	Partial Loss	9	MH 01 21	E 01 2		
101701160502021	1	Partial Loss	999	MH 01 22	E 01 2:		
101701160502022	1	Total Loss	99	MH 01 23	E 01 2		
101701160502023	1	Partial Loss	999	MH 01 6	E 01 6	-	
101701160502024	1	Total Loss	9	MH 01 25	E 01 2		
10120110000000	· ·	m	000	MI 01 02	n		
1					(*)		
ady filiation (1990) Insuran		IDEA - IDEA				Ln , Col	Page:

We then start selecting the fields that need to be present in the IDEA file.

After selection of each field, the **Field Details** appear on the right hand side of the screen.

In the Field Details – the Name and Type of the field is entered. IDEA by default sets the Type of the field based on the characteristics of the field. The field type can be changed from the default type to a user choice type.

		i iii i iii iii iii iii iii iii iii ii	and a second sec	and the second state is a second		Field Details
101701160502000	1	Total Loss	9	MH 01 1	E 01 1	1 22 24
			0.701	Part of the lot	0	T. B. P. St. Station and Street Stree
Claim No	S1 No	NOL Dev Off	Code	Veh Registration No.	Engine	Laver Name Laver-1
101701160502000	1	Total Loss	9	MH 01 1	E 01 1	The Party of Tarfer strand have
101701160502001	1	Partial Loss	99	MH 01 2	E 01 2	Name Claim Mumber
101701160502002	÷	Total Loss	9999	MH 01 3	E 01 3	Type Numeric Decimals 0
101701160502003	1	Total Logg	999	MH 01 4	E 01 4	Implied Decimais No
101701160502004	1	Total Long	99	MH 01 5	E 01 5	Maych (Anonar)
101701160502005	-	Partial Loss	999	MH 01 6	E 01 6	Offset 0
101701160502006	i.	Total Loss	999	MH 01 7	E 01 7	Offset 0 Depley Width 0
101701160502007	î	Total Loss	999	MH 01 8	E 01 8	Actual Width 15
101701160502008	1	Partial Long	99	MH 01 9	E 01 9	has Alterhautas
101701160502009	1	Total Logs	999	MH 01 10	B 01 1	Blank Cells Leave blank
101701160502010	1	Partial Loss	99	MH 01 2	E 01 2	End Field On Ellark Ine
101701160502011	1	Total Loss	9	MH 01 12	E 01 1	
101701160502012	1	Partial Loss	99	MH 01 13	E 01 1	Include/Exclude Al
101701160502013	î	Total Loss	99	MH 01 14	B 01 1	Range Al
101701160502014	î	Partial Loss	9999	MH 01 15	B 01 1	
101701160502015	-	Total Loss	909	MH 01 16	E 01 1	
101701160502016	-	Partial Long	9999	MH 01 15	E 01 1	
101701160502017	-	Total Loss	999	MH 01 18	E 01 1	
101701160502018	-	Total Loss	9999	MH 01 19	E 01 1	
101701160502019	1	Total Loss	9999	MH 01 20	E 01 2	
101701160502029	÷	Partial Long	9	MH 01 20	E 01 2	
101701160502020	-	Partial Loss	000	MH 01 22	E 01 2	Name
101701160502022	-	Total Loss	99	MH 01 23	E 01 2	Name of this field.
101701160502022	1	Partial Loss	999	MH 01 6	E 01 6	
101701160502023		Total Loss	9	MH 01 25	E 01 2	
101701100502024	-	Total Long	000	MR 01 25		
1						Ln 14, Col 67 Page: 1/1

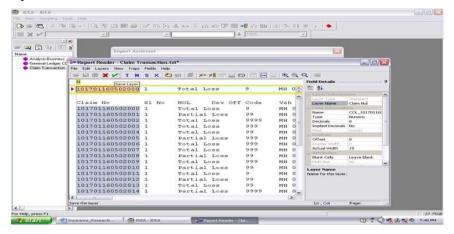
All the fields in the text file are selected and defined based on the record layout provided below –

Field Name	Field Description	Field Type
Claim No	Claim Number	Numeric
SL No	Serial Number	Numeric

Application	of (CAATs	to	Utility	Services

NOL	Nature of Loss	Character
Dev Off Code	Development Officer Code	Numeric
Veh Regn No	Vehicle Registration Number	Character
Engine No	Engine Number	Character
Accident Date	Accident Date	Date
ELA	Estimated Loss Amount	Numeric
EEA	Estimated Expense Amount	Numeric
PLA	Paid Loss Amount	Numeric
PEA	Paid Expense Amount	Numeric
V No	Voucher Number	Numeric
V Date	Voucher Date	Date
Policy No	Policy Number	Numeric
Receiver	Receiver	Character
Chq No	Cheque Number	Numeric
Chq Date	Cheque Date	Date

After selecting and defining the fields, click on the Green Check Button on the Report Reader menu bar. This saves the activities performed on the layer.

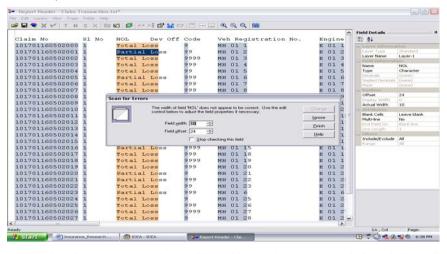


Then click the Scan for Errors Button on the Report reader menu bar. This invokes the Scan for Errors dialog, which corrects data which has been left out of the selection in the earlier step.

Final Showing Mark						
-1	a texa texa Berra		3		- Caracter and	
	Report Reader - Claim Transa					
	File Edit Layers View Traps Field					
	📽 🖬 🧐 🗶 💅 T N S	× 🕮 🗭 😂 ×= ×≣ 🖾		RAQI	101	
		Scan for Errors			d Details 9	
			Off Code	Veh III	24	
	101701160502000 1	Total Loss	9	мн о		1
	101701160502001 1	Partial Loss		мн о		
	101701160502002 1	Total Loss	9999	MH O		
	101701160502003 1	Total Loss	999	мн о		
	101701160502004 1	Total Loss	99	мн о		
	101701160502005 1	Partial Loss Total Loss		MH 0		
	101701160502006 1 101701160502007 1	Total Loss Total Loss	999	мн о		
	101701160502007 1	Partial Loss		MH 0		
	101701160502009 1	Total Loss	999	MH 0		
	101701160502010 1	Partial Loss		MH 0		
	101701160502011 1	Total Long	9	MH 0		
	101701160502012 1	Partial Loss	99	MH 0		
	101701160502013 1	Total Loss	99	MH 0		
	101701160502014 1	Partial Loss	9999	MH 0		
	101701160502015 1	Total Loss	999	MH 0		
	101701160502016 1	Partial Loss		MH 0		
10-10 Sales	101701160502017 1	Total Loss	999	мн ом		
Ca CD > CJ <	Scan for any bad traps				Ln , Col Page:	12

Increase or decrease the Field Width to adjust the right indent. Increase or decrease the Field Offset, to adjust the left indent of any field.

Notice the first case – Partial Loss (Seen Below). In this case, the letters 'ss' from Loss have been left out of the selection. Also notice that the Field Width for this specific cell is 10.



Notice that by increasing the Field Width, we have covered the whole cell and the whole word 'Partial Loss'

Application of CAATs to Utility Services

🖬 🔍 X 🖌 T N	s x	a 🖬 🚅 🗶 🕫 🖬	P 14 03			Q	Q III					
			-	and the second s					1	~	Field Details	
Claim No	S1 No	NOL De	V OFF	Code	Vel	Re	gistration No.	E	ngine		10 2.	
101701160502000	1	Total Loss		9	MH	01	1	E	01 1		- Lager Informa	
101701160502001	1	Partial Los	18	99	HH	01	2	E	01 2		Layer Type	Standard .
101701160502002	1	Total Loss		9999	MH	01	3	E	01 3		Layer Name	Layer-1
101701160502003	1	Total Loss		999	1414	01	4	E	01 4		- Field Informati	NOL
101701160502004	1	Total Loss		99	HH	01	5	E	01 5		Type	Character
101701160502005	1	Partial Los	2.52	999	MH	01	6	E	01 6		Decimals	(tatatan)
101701160502006	1	Total Loss		999	MH	01	7	E	01 7		Implied Decimals	(more) (more)
101701160502007	1	Total Loss		999	MH	01	8	E	01 8		-Pusition	
101701160502008	1	Scan for Errors		and the second second					9		Offset	24
101701160502009	1	Scan for Errors				_			1		Actual Width	10
101701160502010	1	The	width of fie	id 'NOL' does	not appe	ar to be	e correct. Use the edit	Chang	- 12		a) Attributes	10
101701160502011	1	con	rol below 6	o adjust the fi	eld propert	ies it n	ecessary.			11	Blank Cells	Leave blank
101701160502012	1		Field wid	at land	- E			gnore			Multi-line	No
101701160502013			1010000					Einish			End Field On	Black line
101701160502014	1		Field gff:	iet: 24 ;	H			Help	11		- fullers	
101701160502015	1			T Stop	checking (his field	d		- 1		Include/Exclude	
101701160502016		Partial Los		9999	MH	01	15	12	01 1		Aange	AI.
101701160502017		Total Loss		999	MH	01		E	01 1			
101701160502018		Total Loss		0000	MH		19	10	01 1			
101701160502019	1	Total Loss		999	MH		20	E	01 2			
101701160502020	1	Partial Los		9	MH		21	E	01 2			
101701160502021	1	Partial Los		000	MH		22	E	01 2			
101701160502022	1	Total Loss	100	99	MH		23	E	01 2			
101701160502022	1	Partial Los		999	MH	01	6	E	01 6			
101701160502024	1	Total Loss		9	MH		25	E	01 2			
101701160502025	1	Total Loss		000	MH		26	E	01 2			
101701160502026	1	Total Loss		9999	MH		27	E	01 2			
101701160502028	1	Total Loss		9		01			01 2			

We continue scanning the file by adjusting the width and offset as shown above, till we get the message Validation is Complete. No other errors were found.

Mah	De	ers a	tration				10 1		Accident		Entimat.	nd Long	Amt	-	Field Details	
	01	1	crackon.		R	01	1	31.03.05		100000	D	75000	0		ND 91	
MH	01	2			XC.	01	2	01.04.05	5	25000	5000	18750	2750		End Romann Statements	
MH	01	3			10	01	3	28.02.01	5	350000	0	262500	0			
MH	01	-4			B	01	-4	20.02.01	5	1500000	0	1125000	0		Layer Name	Layer-1
24.24	01	5			10	01	5	01.04.05	5	1000000	0	750000	0		Martin	PEA
HH	01	6			в	01	6	01.05.01	5	100000	0	75000	0		Type	Pératoseric:
MH	01	7			12	01	7	07.05.05	5	1250000	45000	937500	24750		Decimals	0
MH	01	0			10	01	0	15.05.05	5	350000	D	262500	0		Implied Decimals	Péro
PEFE	01	9			R	01	9	01.06.0		650000	0	487500	0		and Permitance	
1111	01	10			10	01	10	17.05.05	5	450000	45000	337500	24750		Offset	120
24.24	01	2			R	01	2	18,05,05	2	45000	5000	33750	2750		Actual Welth	0 6
2131	01	12			R	01	12	01,06,05		4500000	65000	3375000	35750		Address wears	
	01	13			10	01	13	20.07.0		37500	0	20125	0		Blank Cells	Leave blank.
MH	01	1.4			10	01	1.4	20.07.0		1150000	0	862500	0	1	Phillipping	1.Parts
	01	1.15			R	01	15	20.07.05		175000	ũ.	131250	0			Disco lese
	01	16				-	0.000					750000	19250		- Pilters	
	01	11	can for Ermit							-		750000	0		Include/Exclude	
	01	11	Common State	Report	(nad	lor.					Continue 1	1075000	0		Range	A8
	01	44	-								and the second se	2625000				
	01	21			VA	dation	1.85.004	splete. There we	tre no other error	s found.	Linke-	4875000				
	01	2:		1000							Erial -	675000	0			
		21						ÓK			Help	65625	0			
	01	21			_	-	Contract (The second state of the second			- Det	1233750				
	01	6			10	10.4	D	20.09.00		MESON		73875	0			
	01				1	01	25	01,07,05		975000	0	731250	0		10	
		26			10	01	26	15.10.0		235600	0	176700	ő			
		27			12	01	27	28,06,05		865000	65000	648750	35750			
	01				10	01	20	20, 10, 01		479650	0.0000	359737.				
MH					2	01	29	20, 10, 0		475000	0	356250		-		
PAR	O.A.	100.00			10	0.4	100.00	20120101		473000	W-	330430				

Click the button Import into IDEA

Data Analysis for Auditors

E COR		ers View Traps Pields Help											
-		t mto IDEARtion No.		gin			Accident Date		ed Loss	Amt	~	Field Details	
	01			01		31,03,05	100000	0	75000	0		TD 24	
2135	01	2	10	01	2	01.04.05	25000	5000	18750	2750		A some linterrout	
MIC	01	3	10	01	3	28,02,05	350000	0	262500	0			
2114	01	4	E	01	4	20.02.05	1500000	0	1125000	0		Layer Name	Layer-1
MH	01		E		5	01.04.05	1000000	0	750000	õ.		Norme	PEA
MH	01	6	12		6	01.05.05	100000	0	75000	0		Type	Numeric
2111	01	7	10	01	7	07.05.05	1250000	45000	937500	24750		Decimala	0
MH	01	8	10		8	15.05.05	350000	0	262500	0		Implied Decimals	No
MH	01	9	10		à.	01.06.05	650000	0	487500	0		- Presidente	
MH	01	10	12		10	17.05.05	450000	45000	337500	24750		Offset	120
MH	01	2	12		2	18.05.05	45000	5000	33750	2750		Display Width	
MH	01		10		12	01.06.05	4500000	65000	3375000			Actual Width	6
MH	01	13	10		13	20.07.05	37500	0.0000	20125	0		Blank Cells	Leave blank.
MH	01	14	10		1.4	20.07.05	1150000	0	862500	ő		Phato-krue	Fáiz
MH	01	15			15	20.07.05	175000	0	131250	0		End Field On	Blackline
MH	01	16	10		16	15.07.05	1000000	35000	750000	19250		- Fillers	
MH	01	15			15	20.07.05	1000000	0	750000	0		Include/Exclude	
MIL		18	E		18	23.08.05	2500000	0	1875000			Range	AL
MH	01		R		10	01.06.05	3500000	ő	2625000				
	01				20	01.00.05			4875000				
MH	01		E		21		6500000	100000	675000	0			
MH			10			15.09.05		0	65625				
MH		22	1C		22	10.09.05	87500	0		0			
MH	01		E		23	15.08.05	1645000	75000	1233750				
MH	01	6	в		6	20.09.05	98500	0	73875	0			
MH	01		E		25	01.07.05	975000	0	731250	0			
MH	01		IC		26	15.10.05	235600	0	176700	0	1.00		
	01		10		27	28.06.05	865000	65000	648750	35750			
	01		E		20	20,10,05	479650	0	359737.5				
MH	01	29	в	01	29	20.10.05	475000	0	356250	0	~		
	-	to IDEA								1.00		kn. Col	Pagel

Report Reader prompts us to create and save a template for the file in the working folder.

This template allows the user to import similar files in future, with the use of the template without going through the Report Reader process.

9 9999 999 999 999 999 999	МН 01 1 МН 01 2 МН 01 3 МН 01 4 МН 01 5 МИ 01 6 МИ 01 7	E E E	01 1 01 2 01 3 01 4	31.03.05 01.04.05 28.02.05	100000 25000	0 7500 5000 1875	E Laver Inference	
9999 999 999 999 999 999	МН 01 3 МН 01 4 МН 01 5 МИ 01 6	E	01 3			5000 1075		
999 99 999 999 999	MH 01 4 MH 01 5 MH 01 6	E		28.02.05				
99 999 999 999	MH 01 5 MH 01 6				350000	0 2625	Laver Name	Starubied
999 999 999	MH 01 6	EC .		20.02.05	1500000	0 1125	Layer Parne	Layer-1
999			01 5	01.04.05	1000000	0 7500	Name	PEA
999			01 6	01.05.05	100000	0 7500	Туре	Numeric
			01 7	07.05.05	1250000		Decimals Implied Decimals	0 No
	MH 01 0	Save As				2625	Mach	(moning)
99	MH 01 9	Save in	Insurance	Research		4875	- Paulinie	
999	MH 01 10	and the first of the			21 + G C C	3375	Offset	120
99	MH 01 2					3375	Actual Width	5
9	MH 01 12					1 3375	- Attributes	
99	MH 01 13					2012	Blank Cells	Leave blank
99	MH 01 14					8625	Eind Parks Ces	Charle hour
9999	MH 01 15					1312		
999	MH 01 16) 7500	Par Pillare	
9999	MH 01 15	File name				7500	Include/Exclude Bange	A3
999	MH 01 18	File name:	Claim Tran	asaction	50	1075	to a spe	
9999	MH 01 19	Save as type:	Template	files (".apm]	• Can	vcel 2625		
999	MH 01 20	-				4875		
9	MH 01 21	82	01 21	15.09.05	900000	0 6750		
999	MH 01 22	HI.	01 22	10.09.05	87500	0 6562		
9.9	MH 01 23		01 23	15.00.05		75000 1233		
999	MH 01 6		01 6	20.09.05	98500	0 7387		
9	MH 01 25	E	01 25	01.07.05	975000	0 7312		
999	MH 01 26		01 26	15.10.05	235600	0 1767		
9999	MH 01 27	E	01 27	28.06.05	865000	65000 6487		
9	MH 01 28	101 E	01 28	20.10.05	479650	0 3597		
0.0	MH 01 29	80	01 29	20.10.05	475000	0 3562 .		

We proceed to the final screen 'Specify IDEA Filename'.

We may retain the name as reflected by the source file as default name or give a user defined name.

The Net Value for the fields ELA, EEA, PLA and PEA are viewed from the Field Statistics and reconciled back to the control totals provided with the text file.

The totals match.

We can now use this file for further testing.

Application of CAATs to Utility Services

· ··· bu biline · ·	an and automatic		Concern Clarge					* 49 1 100 O
ngat Data Salat Adres	Pagente Debite and a construction of the const	9						
e Explorer *	Analysis-Tide Client Z Claim Transaction	×						Properties
1 2 P 3 1	0 8 9 * 2 W # 0 9 + 0 *							· Database
Also Proved	Field Type	Numeric	CLAIM_NO					
ame Records	Numeric	 Net Value 	3,05,10,34	30		3,29,19,750	4,40,00	
Analysis-Ti_ 57	Date	Absolute Val_	3,05,10,34	30		3,29,19,750	4,40,00	
Claim Tran 30	Time	# of Records	30	30	30	30	3	 Field Statistics
	Numeric Fields	# of Zero Its	0	0	0	0	2	Control Total
	CLAM_NO	Positive Value	3.05,10.34	30		3,29,19,750	4,40,00	Criteria
	V SL_NO V DEV OFF CODE	Negative Val	0	0	0	0		· Results
	V ESTIMATED LOSS AMT	# of Positive	30	30	30	30		
	PAID_LOSS_AMOUNT	# of Negative	0	0	0	0		• Indices
	V PAID EXPENSE AMT	# of Data Err # of Valid Val	30	0	30	30	-	No index
	VOUCHER_NO	Average Value	10.17.01.1	1.00		10.97.325	74.000.0	· Comments
	V POLICY NO	Minimum Val	10,17,01,1	1.00	2,001.00	25.000	14,000.0	
		Maximum Va_	10.17.01.1	1	9,999	65.00.000	1.00.00	Add comment
		Becord # of	10,00,00,00	1	1	2	1,000,000	
		Record # of	30	1	3	20	2	
		Sample Std	0.00	0.00	3 636 62	14.45.945	28 129 4	
		Sample Vari_	0.00		1.32.25.00	20.90.75.9	79.12.64.3	
		Pop Std Dev	0.00	0.00		14,21,642	27.656.6	
		Pop Variance	0.00		1,27,84,17	20.21.06.7	76.48.88.8	
		Pop Skiwness	0.000000	0.000000	1.732307	2.357806	1 72100	
		Pop Kurtosis	0 000000	0.000000	1.097727	5.434694	1.66991	
File Explorer (Dibrary)			III.	and the second	10022012	-		
The capacities an obtainy				-	_	_		-
Running Tasks								

The process of import is similar for the other Text files intended to be used for various Insurance Audit Tests.

13.1.6 Premium Case Studies

1.0 RE-VERIFICATION OF PREMIUM COLLECTED FOR INSURANCE AGAINST NATURAL DISASTERS LIKE EARTHQUAKES

OBJECTIVE

To verify whether the correct value of Premium on Earthquake Class of risks has been recovered from the insured as per the Premium Masters and the correct City Code.

EXERCISE DESCRIPTION

In this exercise we will join the Premium Transaction and Premium Master File based on the City and Location field respectively.

Append a computed field - virtual numeric field to re-calculate premium on sum insured by applying the master rate for the City to the Sum Insured.

Arrive at the differences between premium recovered and computed premium

Understand the reasons for differences if any in premium recovered.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

This exercise will show us how to:

- Join two files.
- Append computed fields virtual numeric fields.
- Use the Equation Editor.
- Extract data with a specified criterion.

Data Import

Import the files Premium_Source – Premium – Earthquake. xls and Earthquake Premium Masters. dbf into IDEA as per the steps demonstrated in the session on IMPORT OF DATA FILES.

The files are imported as two independent IDEA files as seen in the File Explorer Section of IDEA.

JOIN TWO FILES

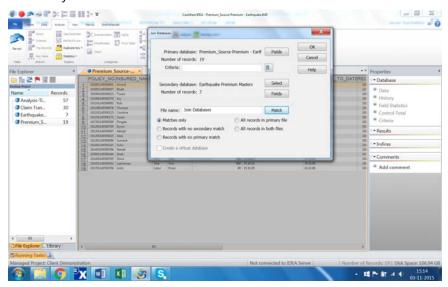
Designate Premium_Source – Premium – Earthquake. IMD as the Primary Database and Earthquake Premium Masters. IMD as the Secondary Database.

This is based on the logic that the premium masters are to be applied to the premium transaction file.

Open Premium_Source – Premium – Earthquake. IMD, and then click the Join Database option on the File Explorer Toolbar.

is relation		·				Properties
	POUCY_NOINSURED	NAME CITY LOCATION DEV_O	FFICER_AGENTRISK_EFFECT_FI	ROM_DATERISK_EFFECT_T	O_DATEREC	Database
Inne Records Andylis-Tic. 57 Claim Tran. 30 Earthquake. 7 Premium S. 19	3 1551440902 Test 4 250244000 Ref. 4 250244000 Ref. 4 250244000 Ref. 5 250244000 Ref. 7 151214000 Ref. 9 250244000 Ref. 9 151214000 Ref. 10 151214000 Ref. 10 151214000 Ref. 10 151214000 Ref. 11 151214000 Ref. 12 151214000 Ref. 13 151214000 Ref. 14 14124410000 Ref. 15 1515140000 Ref. 14 14144140000 Ref. 15 15151400000 Ref.	Model Model Marchael Model Marchael Model Marchael Model Marchael Model Marchael Mar	1 2.2446 2 2.0456 3 2.0456 4 2.0456 3 2.0456 4 1.0456 3 2.0456 1 2.04	30 (a) (a) 30 (b) (a) 50 (a) (a) 12 (a)		Data History Field Statistics Control Total Criteria Results Mo index Comments Add comment
n +		11				

The Premium_Source – Premium – Earthquake. IMD is automatically seen as the Primary Database. Select Earthquake Premium Masters. IMD as the Secondary Database.



In earthquake insurance, the location is the decisive factor for premium application.

For example, Bhuj in Gujarat and Latur in Maharashra are sensitive seismic zones and command higher premium.

The Primary database has a field 'Location' and a field 'City'.

Both the 'City' and the 'Location' fields in the Policy Creation stage are to be selected from an index list of cities within the insurance application software.

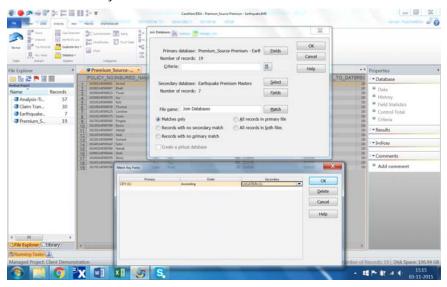
The 'City' field is selected at the entry stage, based on the documentary evidence submitted and site visits.

The 'Location' field on the other hand has no referential link to the 'City' field, and can be independently entered.

The Premium Master Rate applied to the sum insured is dependent on the 'Location' and not the 'City'.

To consider the possibility of an incorrect selection of the 'location' affecting the Master Premium Rate applied by the system, we match the two databases based on the 'City' field in the Primary Database and 'Location' field in the Secondary Database.

Click Matches Only.



The Joined database appears as a Child database of the Parent - Primary database - Premium_Source - Premium - Earthquake. IMD

Image: Second	Lite Kay *	trathuter Directate	 Ban Append Vaue Corrects Corpore Solors 	tanton Otwr *	Maradary LAIL *			
e Explorer	Z	Premium_Source-Pr						Properties
12 2 P 2 III		POLICY_NOINSURED_N	AME CITY	LOCATION Pure	DEV_OFFICER_AGENTRISK_EFFECT	FROM_DATERISK_EFFECT_TO	DATEREC	Database
mer Maxim Records Analysis-Tit. 57 Cickim Tran. 30 Earthquake. 7 Permian 5. 19 Join Dat. 15	23456789101111231415		Mambai	16ageur	Image Image Image <th>1738.00 1737.00 1747.000 1747.000 1747.000 1747.000 1747.00000000000000000000000000000000000</th> <th>86 93 93 94 94 94 94 94 94 94 94 94 94 94 94 94</th> <th>Otata History Field Statistics Control Total Criteria Results No index No index Add comment</th>	1738.00 1737.00 1747.000 1747.000 1747.000 1747.000 1747.00000000000000000000000000000000000	86 93 93 94 94 94 94 94 94 94 94 94 94 94 94 94	Otata History Field Statistics Control Total Criteria Results No index No index Add comment
m +			.10				,	

APPEND A CALCULATED FIELD – VIRTUAL NUMERIC FIELDS

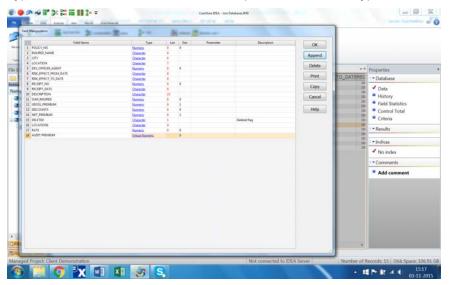
Double click on any field cell in the joined database

This invokes the Field Manipulation dialog box.

Append a field with the following specifics:

Name:Write 'Audit Premium'

Type: Select 'Virtual Numeric' from the drop down list of Field Types.



Click the Parameter cell.

USE THE EQUATION EDITOR

This opens the Equation Editor.

Logical criteria in IDEA are written through the Equation Editor.

In this specific audit test, apply the master premium rate to the sum insured.

The result is to be deducted from the premium recovered to identify errors in premium recovery.

The logical criteria / equation can be entered in two ways.

The first one is used if we don't know the exact field name, but are aware of the starting letter of the field.

Type the starting letter of the field to be entered and then select the field from the List of fields displayed on the screen that begin with that letter.

Type 's'. IDEA displays 'sum_insured'. Press enter.

The second way is to click 'Insert database field'.

This option gives a list of all the fields in the database.

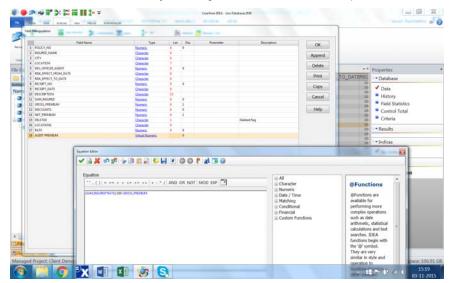
Scroll to the field of choice, and click the field.

In this way enter the equation (sum_insured*rate)/100-Gross_Premium.

Constants and logical operators can be entered by clicking on the Equation Editor calculator or using the keyboard.

After entering the equation in the equation box, click the Validate Equation button.

This is done to check the syntax, logic and structure of the equation.



The equation is checked by IDEA.

The equation is valid. Click OK in Valid Equation box.

Application of CAATs to Utility Services

Manpulation	>	1	N					
Failt Na POLICY NO	***	Туре	Lan Dec	Paramatar	Description	ОК		
POLICY NO		Mameria Character	1 1			[
UTV .		Chevadas				Append		
LOCATION DRV.OFFICER_AGENT		Cheadar Nomers	1			Delete	101	
REX EFFECT FROM DATE		Character				-		Properties
REK, EFFECT, TO, DATE		Datadat				Print	TO_DATEREC	Database
RECEPT,NO		fames.	4 4			Copy	32	✓ Data
RECEPT, DATE		Character	15			cohi	10	
DESCRIPTION SUM INSURED		Character Networks				Cancel	10	
CROSS, PREMIUM		Names.	1				100	Field Statistics
DISCOUNTS		thenesis:	8 0			Help		Control Total
NET_PREMUM		formers:	8 1	Caurillan ERA		1.	381	Criteria
5 DELETED 6 LOCATIONS		Chevelet	1.	CoorWere DIA			20	- Unteria
E LOCATIONE		Character Romens					10	· Results
AUGUT PRESIDENT		Votual Nameria	0	Valid Equation!			30	
1.1	<mark>ම් 🖊 හ</mark> ේ නේ 🦛 🗃 ශාමාන			e ok				int
	** , () = ++ > < <+		/ AND OR NOT	MOD EXP	All Charact	er.	@Funct	tions
Project: Clent Demo	an yaund yang yan Group	немени			III Numer III Date / III Matchin III Conditi III Financi III Custom	Firme 10 onial al	@Function available fr performicy complex o such as da arthmetic, calculation searches. I functions to the '@' syn They are v stermar in s	or g more perations te statistical s and text IDEA begin with mbol. ergn

Click on the Validate and Exit button to process the equation on the open Database.

Click OK in the Field Manipulation dialog box.

The resultant computed virtual numeric field appears in a GREEN FONT

ran Parkan Tar Reserve Rey Taken	Gap Deter Gap Deter Gaberbeite Dapkanne Etabliste Etabliste	ider Kary *	Constantine Di Apro Constante Di Anos Const Constante		Attribute 🧰 Have E Familier 🍏 Grade Other *						
Explorer			Premium_Source-		abases.IMD ×						Properties
12 P 0			DESCRIPTION SUI	M INSUREDGROS	S PREMIUMDISCO	DUNTSNET	PREMIUM DELI	ETED LOCATION	RATE AUDIT	PREMUM	Database
Print	111 C	1	Building	8500000	212,500.0		212,500.0	Pro Contraction	14	127908	1-
me 1	Records	3	Building	2500000	90,000.0	0	50,000.0	Latur		30000	🖌 Data
Analysis-Ti_	57	4	Farm House	4500005	112,500.0	0	132,506.0	Latur	4	67500	History
		5	Stock	7500000	187,500.0	0	187,500.0	Membai	11	37500	Field Statistics
Claim Tran	30	6	Farm House.	530000	12,500.0	8	12,500.0	Munba	3	2509	
Earthquake_	7	7	Stock Building	250000	8,250.0		6,251.0 10,005.0	Munital Negtur		L290	Control Total
		8	Farm House	750000	15,500.0		25,008.0	Region	1		Criteria
Premium_S_	19	10	Businets Center	1300000	30,000,0		30,000.0	Negrut	1		Criteria
Join Dat	15	11	Business Center	1500000	20.000.0	0	30.008.0	Nagour .		1	· Results
		12	PartingLot	430000	11,295.0	.0	31,250.0	There	1	3798	- Nesoics
		13	Factory	450000	112,500.0	0	112,500.0	There	- 1	22588	
		.14	Perking Lot Fectory	850000 6300000	21.250.0	0	21,258.0	Thate	1	4250	• Indices
		15	Felloy	6300000 [162,500.0		142,500.0	(There	1.01	7200	✓ No index
		L .									· • Comments
		L .									Add comment
m											
lle Explorer Utit	brary	1									
	0	-									1
anning Tasks 🕼											

EXTRACT DATA WITH A SPECIFIED CRITERIA

Now perform a data extraction so as to identify all the differences in the premium charged vis-à-vis audit premium from the computed field.

Click Analysis – Extract - Direct

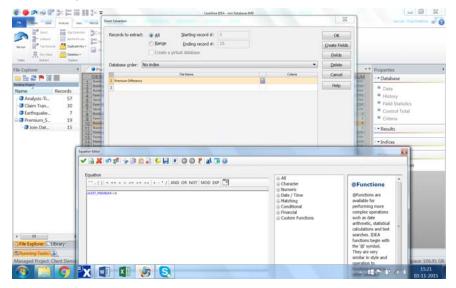
This invokes the Direct Extraction dialog box

Enter the File Name as 'Premium Difference' and click on the calculator button.

This opens the Equation Editor.

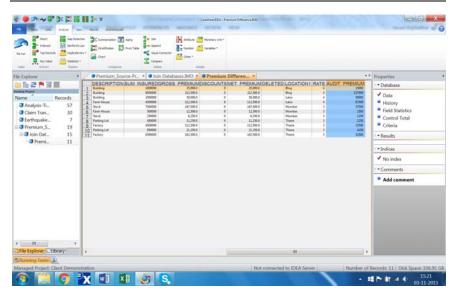
Enter the equation 'audit_premium<>0' in the Equation Box.

This equation will identify separately differences in premium charged vis-àvis the computed field.



The extraction gives 11 cases of short premium charged, as seen from the field "Audit_Premium".

Application of CAATs to Utility Services



Click the Set Control total button.

Set the Control Total for the computed field Audit_Premium

The Control Total appears as 362750.

This is indicative of the Net Premium Under-Recovered during the period of review.

CAUSE FOR THE UNDER-RECOVERY

A closer review reveals the following:

In Policy Number 103001140500007 and 103001140500444 the city being BHUJ, location was wrongly selected as PUNE while preparing the policy. Hence the system picked the premium rate for PUNE, which is 2.5, whereas 4 should have been applied for BHUJ.

In Policy Number 101701140500500 and 101701140500356 the city being LATUR, location was wrongly entered as SHOLAPUR and PUNE respectively. Hence incorrect rates were applied by the system based on incorrect location selected by the development officer.

LEARNING

The development officer while selecting the Location field in the policy generation process has inadvertently selected the wrong location from the index in the insurance application software.

As premium applied is based on the location field, and the location does not match with the city in certain cases, incorrect (short) premium has been applied.

As a system suggestion the Location field should not be a stand-alone field. The field should be linked to the city field in the application system.

2.0 RE-VERIFICATION OF PREMIUM COLLECTED FOR INSURANCE AGAINST MOTOR CLASS OF RISKS

OBJECTIVE

To verify whether the correct value of Premium on Motor Class of risks has been recovered from the insured as per the Premium Masters and the correct City Code represented by the Vehicle Registration Number.

EXERCISE DESCRIPTION

In this exercise we will join the Premium Transaction and Premium Master File based on the Zones respectively.

Append a computed field - virtual numeric field to re-calculate premium on sum insured by applying the master rate for the Zone to the Sum Insured.

Arrive at possible differences between premiums recovered and computed premium

Understand the reasons for differences if any in premium recovered.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Use @ Functions
- Join two files.
- Append computed fields virtual character fields.
- Use the Equation Editor.
- Extract data with a specified criterion.

Data Import

Import the files Premium_Source – Premium – Motor. xls and Motor Premium Master. Dbf into IDEA as per the steps demonstrated in the session on IMPORT OF DATA FILES.

USE @ FUNCTIONS

The location in the Policy Creation stage is ordinarily guided by the first 5 digits or characters or location codes of the Vehicle registration number.

This represents the city of registration and use, with the assumption that the vehicle is insured in the same city in which it has been registered and used.

Since the master file also contains the location code, the location code should be captured as a separate field in the premium file for further comparison.

The location code should be captured as a separate field from the Vehicle Registration field by writing an @ function - @ left in the Equation Editor.

The @ Left function isolates the left most occurrences of a fixed number of digits; in our case the first 5 digits of the Vehicle Registration number.

Go to Field Manipulation in Data in the Menu Tool-Bar.

Append a field with the following specifics:

Name: Write 'Location Code'

Type: Select 'Virtual Character' from the drop down list of Field Types.

Length: 5

Click the Parameter cell.

Enter the equation '@left(Veh_Reg_No, 5) in the Equation Box of the Equation Editor.

Manpulation	-92			10					
	Faid Name	Type	Lar	Dec	Faramatar	Description	ок		
POLICY_NO INSURED_NAME		Mameria Character		.0			in the second second		
VEH REG NO		Cheater	11				Append		
LOCATION		Ownetter	1				Delete		
DEX,OFFICER, AGENT		thermos.					Desete		Properties
REX, EFFECT, FROM, DATE		Operated	4				Print	FFECT_TO_D	
REX, EFFECT, TO, DATE		Character					eren	FREGT_10_D	Database
RECEPT,NO RECEPT,DATE		Batters					Copy		✓ Data
DESCRIPTION		Character Character	11						
SUM INSURED		Transis					Cancel		History
2 GROSS FREMEMAN		Names	- 1	1				_	Field Statistics
DECOUNTS		Network		3			Help		Control Total
A NET, FREMUM		Transis .		3			And and a second second		Criteria
S LOCATION CODE		Votual Character	3	0					 Criteria
									· Results
									- NESUIS
									- 6 A
			-	_					• Indices
	Institut Liditor								
			11.00	0.0	0.4 - 0.0				
	A 10 1 20	- 10 E KI	H (10)	00	r a				
									nt.
	Equation					III All		Proceedings.	and the second second
		a an an an la ca	1.440	OR NOT					
		C 44 34 43 [4.1	/ AND	OR NOT	MOD EOP	Character		@Funct	ions
	··· , () = ··· >	< <* >* <> + +	/ AND	OR NOT	MOD DP	E Charach			
		< (*)* () * · *	/ AND	OR NOT	MOD DP	Charach Numeria Date / T	ime	@Function	sare
	··· , () = ··· >	c (*)* () + · ·	/ AND	OR NOT	MOD DØ	Charado Charado Numeria Dote / T Matchiny	c Ime 0	@Function available fo	sare
	··· , () = ··· >	< <* >* <> + - *	/ AND	OR NOT	MOD DOP	Charach Charach Numeria Date / T Matchin Condition Condition	r Ime 9 mai	@Function available fo performing	s are Xr 1 more
	··· , () = ··· >	< <* >* <* >* <> * · *	/ AND	OR NOT	MOD DOP	Charach	r Ime onal A	@Function available fo	s are X 1 more perations
	··· , () = ··· >		/ AND	OR NOT	MOD DP	Charach Charach Numeria Date / T Matchin Condition Condition	r Ime onal A	@Function available fo performing complex op such as dat	s are pr proce perations te
	··· , () = ··· >	*******	/ AND	OR NOT	MOD DP	Charach	r Ime onal A	@Function available for performing complex op	s are y more perations te statistical
	··· , () = ··· >		/ AND	OR NOT	MOD DP	Charach	r Ime onal A	@Function available fo performing complex op such as dat arithmetic,	s are or perations te statistical s and hext
	··· , () = ··· >	C (*)* (*) * *	/ AND	OR NOT	MOD DP	Charach	r Ime onal A	@Function available fo performing complex op such as dat arthmetic, calculation	s are x proce perations statistical s and text DEA
	··· , () = ··· >		/ AND	OR NOT	MOD DØ	Charach	r Ime onal A	@Function available fo performing complex of such as dat arithmetic, calculation searches.1	s are x proce perations ta statistical s and text DEA egin with
	··· , () = ··· >		/ AND	OR NOT	MOD DØ	Charach	r Ime onal A	()Function available fo performing complex or such as dat arithmetic, calculation searches. 1 functions b	s are proce perations te statistical sand text DEA mpin with bool.

The virtual character field 'Location Code' appears in GREEN FONT.

Notice the contents of this field. It represents the location code from the vehicle registration number.

ile Explorer			Premium Source1 *						Properties .
12P	S 100		RECEIPT NORECEIPT D	DATE DESCRIPTION SU	JM INSUREDGROS	S PREMUMDI	SCOUNTSNET	PREMIUMLOCATION CODE	Database
Atop Project	and the second	2	103001360500007 15-54-05	Own Damages	65000	1.625.0	8.00	1625.000 58410	
ame	Records	3	101004140500023 27.04.05	Third Party	1000000	25,000.0	1,250.000	21,790.000 MH L2	 Data
Analysis-Ti_		4	101/06160506070 21.05.05	Dun Damages	45000	1,125.3	6.05	1325-000 10413	History
		.5	201201340500085 02.06.05	Third Party	650000	13,000.0	0.00	11.000.000 MH+10	Field Statistics
Claim Tran	30	6.7	10139636590078 07.06.05	Own Damages	100000	2,008.8	9.00	2,900.000 MH4 25	
Earthquake_	7	7	101301340500125 1336.05 101301340500275 18.07.05	Third Party Ount Samager	1300000	30,000.0	3,050.000	28,950.000 MH L1 1,300.000 MH 20	Control Total
		8	101701348700700 20.07.25	Own Dantages	75000	1,300.0	8.00	1.500.000 40+17	Criteria
Premium_S_		10	101.309340500.000 11.07.05	Own Damages	85000	23758	8.00	2,375,000 Milet14	Contraction of the second s
i Join Dat	15	11	101305360500007 1736.05	Third Party	1290000	31,250.0	1,250.000	30,000.000 58+14	· Results
Premi	11	12	101205380500015 10:06.05	Third Farty	1785000	465253	1,570,000	41,015.000 Aller 14	
and the second se		.13	101200360500090 2538-35	Third Party	900000	22,500.0	900,000	21,500.000 5.0+1.4	
Premium_S	19	.14	101405360500047 05.05.05 101400360500094 10.08.05	Own Damages Own Damages	25000 #1000	612.5 1.487.5	4.00 8.00	612.500 549-04 1,467.500 549-10	* Indices
		15	10100310700444 18.06.05	Out Datages	75000	1,497.0	8.00	1.575.000 584 22	✓ No index
		17	10140418050055 17.09.05	Third Party	750000	13,125.0	128.125	12.796.875 1414 12	 No index
		18	101401360500412 25.30.05	Oun Damages	43000	787.5	8.00	767.530 Alerbe	
		19	101701360500756 13.30.25	Third Party	850000	25,250-0	2,525.000	18,125,000 MH+10	 Comments
									Add comment
m									
	and an owner of the local division of the lo								
File Explorer		1.							

JOIN TWO FILES

Designate Premium_Source – Premium – Motor. IMD as the Primary Database and Motor Premium Masters. IMD as the Secondary Database.

Join the two databases on the basis of 'Location Code' from the Primary

Database and 'Code' from the Secondary Database.

Distance Data Data (spin) Data Image: Strate (spin) Image: Spin (spin)	The Summarian Constant	The second seco			ina tatati d
File Explorer #	Premium_Source1-	x in Inin Databaset I x		SS BION DE	Properties FECT_TO_DA Patabase
Term Point Point Name Records Johnshitz Statistical		C Records with no secondary match	n - Mot Fields	OK anci	FECT_TO_DA = Data Data History Field Statistics Control Total Control Total Results Field Statistics No index No index Add comment
The taplorer		Records with no primary match	<u>leensky</u> 096 63 ⊨¥	OK Delete Cancel Help	• • Namber of Records: 0). Disk Space: 106.89 • 152 • 152 • 152 • 152 • 152 • 152 • 151 •

Tip – For a file join the matching fields do not need to have the same field name. Only the field type should be the same for the match to be effective.

USE THE EQUATION EDITOR

Append a field through Field Manipulation with the following specifics:

Field Name: Audit Premium

Type:Virtual Numeric

Enter an equation in the Equation Box of the Equation Editor as

(sum_insured*rate)/100-gross_premium

The resultant computed virtual numeric field appears as -

Table	rr truthune		terdan	Marelary UNI *					
le Explorer #	Join Databa	ses11 *							Properties
Image Point Records Analysis-TL S7 Claim Tan. 30 Detropuse. 7 Motor Pre. 6 Premium SL. 19 Join Dat. 15 Premium SL. 11 Premium SL. 19 Join Dat. 11	1 2 3 4 5 6 7 7 8 9 9 10 11 12 13 14 15 16 17	MILMESCUNTSNE 413 50 1151 00 1160 00	T PREMIUM 62.300 187.500 14.25.00 11.000.000 1.875.000 11.25.000 12.200.000 12.200.000 12.200.000 12.200.000 12.200.000 12.000.000 12.000.000 12.000.000 12.000.000 12.000.000	No+104 No+104 No+10 No+10 No+10 No+10 No+11 No+12 No+12 No+12 No+12 No+12 No+12 No+12 No+14 No+14	E DELETED LOCATIO Gra Gra Fure Fure Fure Fure Name Name Monitor Monitor Monitor Name Name Name Name Name Name Name Name	N1 CODE 1 M+01 M+03 M+03 M+03 M+03 M+03 M+03 M+03 M+03	RATE AUDIT. 2 3 3 3 3 3 3 3 3 3 3 3 3 3	PREMIUM	Database Data History Field Statistics Control Total Criteris Results Indices No index
ioin Dat 19	19	.386.0	L 200.000	10-20	Nagar	1912	1	•	Comments Add comment

EXTRACT DATA WITH A SPECIFIED CRITERIA

Now extract data through Analysis – Extract - Direct with the help of the equation

audit_premium<>0

The Extraction gives 16 cases where motor premium has been under – recovered. The Control total of such cases is 64964.

Table Fred	Gap Detects		Szenarzotter 🔐 Agrag (statistister 🚺 Rost Table Overt Categorier	Corpore Corpore Corpore Laters	E fanten	Montary UNI *						
e Explorer		/ .	oin Databases111.	* EXTRACT	TONI.IMD	*						Properties
12 P 0	11 · · · · · · · · · · · · · · · · · ·		COSS PREMIUMDIS		PREMIUM	LOCATION CO	DE DELETED			RATE AUDIT		Database
Atop Propert	A A COLUMN	1	6133	0.00	612,500 767,500			Ges Ges	50104		113	1
ame	Records	3	1,625.0	0.00	1,625,000	MH130		Pune	84430	3	325	 Data
Analysis-Ti_	57		110003	0.00	17,000.000			Pune	MH10		4506	History
		5	1,487.5	0.00	1,487,500			Pune Pune	5/H10 5/H10		1063	Field Statistics
Claim Tran	30	7	21,250.0	2125.000	18,125,000			Pute	A9410	i	4250	Control Total
Earthquake_	7	8	12,900.0	8.00	12,500.008	MH 12		Muniter	89432	3	2508	
Motor Pre	6	9	25,000.0	1,250.000	23,790.008			Mumbe	AH+12	2	5008	Criteria
Premium S_	19	10	1,121.0	0.00 1.052.000	1,125,000			Muniter Muniter	MH12 MH12		.225 15000	
		12	11125.0	128.125	12,796,875			Martin	10112	- 1	8175	 Results
🗄 🗊 Join Dat	15	13	2,375.0	0.00	2,375.000	3,6134		Thate	50114	1	475	
Premi	11	.14	31,250.0	1,250,000	30.090.000			Thang	80114	- 1	6250	* Indices
Premium_S_	19	11 12 13 14 15 16	44,625.0	1,570-000	45,055.000			Thate	65+34 68+34		1025	✓ No index
Join Dat.	0		tapara (199.000						 No index
Join Dat	19											· · Comments
EXTR	16											Add comment
m												
File Explorer CD	brary	٤										
Running Tasks	9											
	ient Demonst							at connects				ecords: 16 Disk Space: 106

CAUSE FOR THE UNDER-RECOVERY

A closer review reveals the following:

In Policy No 101101160500085 and 101403160500784 the vehicle has been registered and is being used in Pune, but the officer has wrongly selected the Location as Nagpur and Goa respectively. Hence 2% and 1.75% have been applied by the system for Nagpur and Goa respectively in place of 2.5% for Pune.

CONCLUSION

The development officer while selecting the Location field in the policy generation process has inadvertently selected the wrong location from the index in the insurance application software.

As premium is applied based on the location field, and the location does not match with the Location code in certain cases, it means that incorrect (short) premium has been applied.

As a system suggestion, the Location field should not be a stand-alone field. The field should be linked to the Location Code in the application system.

3.0 RE-VERIFICATION OF DISCRETIONARY DISCOUNT ALLOWED ON PREMIUMS IN THE NON-TARIFF FREE MARKET INSURANCE SEGMENT

OBJECTIVE

To verify whether the discretionary discounts have been correctly allowed as per the Discount Masters

EXERCISE DESCRIPTION

In this exercise we will join the Premium Transaction and Discount Master File based on the Development Officer Code.

Append a computed field, -virtual numeric field to re-calculate discount on gross premium by applying the master rate for the Development Officer to the Gross Premium.

Arrive at the difference between discount allowed and computed discount

Understand the reasons for difference, if any, in the discount allowed.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

• Join two files.

- Append computed fields, virtual numeric fields.
- Use the Equation Editor.
- Extract data with a specified criterion.

Data Import

Import the file Premium Source-Premium Motor Non Tariff. xls and Discount Master. dbf into IDEA as per the steps in IMPORT OF DATA FILES.

EXTRACT DATA WITH A SPECIFIED CRITERION

Perform a Direct Extraction from Data in the Menu Tool-Bar.

Isolate only those premium receipts where discount has been allowed.

Do the same by writing an equation – discounts<>0

The resultant database will be a sub-database by the name 'Discount Allowed'.

This sub-database will be used to test the correctness of discount allowed.

JOIN TWO FILES

Designate Discount Allowed. IMD as the Primary Database and Discount Masters. IMD as the Secondary Database.

Join the two databases on the basis of 'Development Officer Code' available in both the databases.

APPEND COMPUTED FIELDS – VIRTUAL NUMERIC FIELDS

To the joined file, append a field through Field Manipulation with the following specifics:

Field Name: Audit Discount

Type: Virtual Numeric

Enter an equation in the Equation Box of the Equation Editor as

(gross_premium*discount_r)/100-Discounts

The computed numeric field appears in PINK FONT.

Extract the differences to a separate file through Data Extraction. Write an equation

Audit_Discount<>0

The resultant difference in discount computed vis-à-vis discount allowed is reflected in a separate child file.

						💩 🐎 💵 🕏	• •		
🔟 🗙 🖌	V No Index		-	No control to		-			
		INSURED_ 1 Alok		VEH_REG_NO_ IH 14 CG 555	LOCATION Thane	DEV_OFFICER_	_AGENT 11	AUDIT_DISCOUNT	RISK_EFFECT_
ê 🔍 🕃 🦦 🛍 👬	* 🐮 🛄	2 Joshi		H 14 CG 555	Inane IPune		99		15.10.05
lame — 🔷 Analysis-Business		2 00 011	14	1110111401	1 dile		22	-1003	13.10.03
Clemer Transaction General Logar DD Control Logar DD Co	arthquake lotor Non Tariff Motor								
- 21				ry 🖌 Field Stats 🖊					

CAUSE FOR THE EXCESS DISCOUNT ALLOWED

A closer review reveals the following:

For Vehicle registration number MH 14 CG 555, Insured Alok, and Agent code 11, the discount allowed is 8% whereas agent 11 has the authority to give discount only up to 4% of the Gross Premium.

The Division Office has given a special approval for Policy No. 101205160500015, Insured Alok for a higher discount rate from the norm, for which the necessary approval documents are available on record prior to policy release.

CONCLUSION

The agent while exercising his discretionary right for discount has gone beyond the limit in two cases.

The system has inbuilt controls to control the discount rate as per the masters, and any exceptions over the masters have to be electronically approved by the DO, which was actually done in the first case.

It appears that the system has been bypassed in the second case without a preventive message. This matter needs to be taken up with the in-house IT Team for their comments and action.

4.0 VERIFICATION OF RATING FIELDS FOR SPECIAL ADDITIONAL IMPOSES / PENALTY IN CASE OF FREQUENT CLAIM HISTORY ON SENSITIVE POLICIES

OBJECTIVE

To verify whether special additional impose / penalty has been correctly recovered in the rating fields, as per the instructions of the Divisional Office for policies experiencing frequent past claims.

EXERCISE DESCRIPTION

In this exercise we will join the Premium Transaction and Rolling Policy File based on the Old Policy Number.

Identify policies where special additional imposes have not been charged

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Join two files.
- Extract data with a specified criterion.

BACKGROUND

The management audit team at the Divisional Office has identified the following Policies in the earlier year having excessive claim history: 101106160400045, 103001160400333, 101004160400478, & 101701160400174.

This inference has been made from the Incurred Loss Statement for Motor Claims for 2004.

Due to repetitive claim history the Divisional Office has ordered an Additional Excess Impose of Rs. 500 to be charged as penalty for every renewal in the current year.

Based on the policy numbers forwarded by the Divisional Office to the Branch office, the management audit team at the branch has used the conversion facility in the insurance application software to arrive at the corresponding current policy numbers.

This is seen in the file 'Premium Source 2 – Rolling Policies.IMD' as imported into IDEA.

JOIN FILES

Designate Premium Source1-Premium-Motor Non Tariff. IMD as the Primary Database and Premium Source 2- Rolling Policies. IMD as the Secondary Database.

Join the two databases on the basis of 'Policy No' available in both the databases.

The Policy No is the current file to be matched with the Policy No 2005 in the Rolling file.

EXTRACT DATA WITH A SPECIFIED CRITERIA

The resultant joined file represents policies prior to 2005, which have experienced repetitive claims and have been subsequently renewed with clear instructions to levy penal charges.

As a corollary, all the policies finding place in this file should have a special additional impose / penalty of 500 added in the rating field.

To test this file perform a direct data extraction through the equation

Additional_Excess_Impose=0.

DEA (Join3) Pile Edk Yerr Data Analysis Sampling			Charles
	en an an air (+). A see ii gh air (n en ∞ des Ditract to File(s)	818 HU CO 94 84 18	
Another Strategy Control of the State Strate Str	Hecords to extract: (• All Start	ing record II 1	ОК -2004 РОЦСУ, НО. 2005 30047 10104169500023 Сламе Pieks 3400474 10105169500028 3400476 300174 3101116950346 Pieks 3400373 103001160500344 3001160500344 3001160500344 Dates Cancel Help 103001160500344 3001160500344
Discount Difference Join3 Motor Premium Master			
Discount Difference Join3	문 Equation Editor 같 An X 드 산 분 한 번 알 두 신 ?		
Oiscourt Difference Join3 Motor Premium Master Discourt Master		Functions:	
Oiscourt Difference Join3 Motor Premium Master Discourt Master	メ Ai X 二公は X 他 他 G 日 クマ	Functions	
Join3 Motor Premium Master Discount Master	✓ A1 X <td>Functions:</td> <td>C C S S OF functions List of 90 GPHRCTICKS available for performing more complex operations with a date artifuence, statistical</td>	Functions:	C C S S OF functions List of 90 GPHRCTICKS available for performing more complex operations with a date artifuence, statistical

The extraction gives a sub-file with 2 record exceptions.

	sis Sampling Tools							- 8
🕒 🏕 🖪 👘 🖄 🗠	🖪 🕲 🗁 📟 🖨	🛯 🖓 🖓 🖓 📥 🕷	ĸ 🗄 kất Nh 🌌	🔛 =6 a'e 💷 🗠) 🎊 💶 😫 🤉	• •		
🎫 🗙 🖌	No Index		No control tot					
A Construction of the second sec	arthquake fotor Non Tailf Motor ence Chargesi	POLEY R0. 1010418050023 101701180500258		VH FEG NG MH 12 CC BUS MH 10 TT 1497	LOCATION Mumbal Mumbal Pune	DEV OFFICER	AGENT 1 27 0 90 15.11	
	> T	> Database His	I State Charles I		1.1			

CAUSE FOR THE NON-RECOVERY

A closer review reveals the following:

The development officer while renewing the existing policy in the system has not referred to the list of policies experiencing repetitive claim history as informed by the Divisional Office to the branch.

CONCLUSION

The procedure of manual comparison of current policies against past policies is flawed, tedious and subject to error of oversight or advertent omission.

List of policies experiencing frequent claims in the past should be mastered within the insurance application system.

Whenever an agent renews an existing policy that comes within the frequent claim history category, the system should automatically prompt the agent to levy the impose, without which the new policy cannot be created in the system.

The discretion to waive the impose should vest with the Divisional office through electronic validation.

5.0 INPUT FIELD CHECKS ON THE(?) NON-SYSTEM GENERATED OR MANUAL POLICIES

OBJECTIVE

To verify whether there is any missing, duplicate, fabricated policy number in the list of non-system or manual insurance policies.

EXERCISE DESCRIPTION

In this exercise we will run tests for duplicate detection on the Policy Number field.

Specific tests for number format verification will be undertaken on the Policy Number field to establish conformity with the standard numbering pattern.

Gap tests will be run to identify missing numbers.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Identify Duplicates
- Perform Number format Tests
- Display Function
- Identify Gaps
- Extract data with a specified criterion.
- Use @ Functions

Data Import

Import the file Non system policies. prn into IDEA as per the steps in IMPORT OF DATA FILES.

IDENTIFY DUPLICATES

Select Duplicate Key Detection from Analysis on the Menu tool bar.

This opens the Duplicate Key Detection dialog box.

Click Key. This opens the Define Key dialog box. Let the Field be defined as POLICY_NO.

Click OK. Check Output Duplicate Records in the Duplicate Key Detection dialog box. Click OK.

File Explorer 9	Premium_Sou	ce2-Non-system				÷×	Properties 4
Control of Clerks Control of Clerks Control of Clerks Control of Clerks Control Scarcel Penniam Mod Penniam Scarcel Penniam Mod Penniam Scarcel Penniam Mod Penniam Scarce Penniam Scarcel Penniam Scarcel Penni	POLICY / NO POLICY /	INSURED_VAME As Co As Co So Co So Co So Co Co	Group Motor Mu Group Motor Mu Detection Acoln Records and Without Duplicates	nbal Nbal Ke Fiel Care He	1 01.046 1 22/040 1 22/040 1 22/040 7 (20.060 7 (20.060 9 7 13.000 9 7 13.000 9 10 10 07 1 1 10 700 1 1 10 700 9 10 10 07 1 1 10 700 9 10 10 07 9 10 10 07 10 10 10 10 10 10 10 10 10 10 10 10 10 1	*****	Otobose Osto Osto

The test gives us two instances of policies having the same number MOTOR04014.

These cases need to be taken up for further investigation with the development officer and / or agent.

The Insured would need to be contacted in secret to establish his or her identity / validity.

Application of CAATs to Utility Services

le Explorer 4	Premium Sou	rce2-Non-system policier	Uplicate			₩ ×	Properties
📮 🐮 🏖 🚩 🕵 词 👘	POLICY_NO	INSURED_NAME	VEH_REG_NO	LOCATION	DEV_OFFICER_AGENT	RISK_EFFECT	II - Database
A Nane Analysis-Tide Clents Analysis-Tide Clents Analysis-Tide Clents Moor Prenium Masters Moor Prenium Masters	1 MOTOR04014 2 MOTOR04014	C & Co D & Co	Group Motor Group Motor	Mumbai Mumbai		27.04.04 31.05.04	Data History Field Statistics Control Total Criteria
Premium_Source1-Premium - Mot							Results
 Duplicate Premium_Source Premium - Earth 							Volindex
							II - Comments
							Add comment
	214					2	

PERFORM NUMBER FORMAT TESTS

Each manual non-system generated policy is in a standard format of 10 digits and has an alpha numeric field type.

The purpose of this test is to identify policies that do not match the standard number format.

In the first stage we write an @ Function to give us the actual length of each policy.

Append a field through Field Manipulation with the following specifics:

Field Name: Policy Length

Type:Virtual Numeric

Write an equation @len(@trim(policy_no)) in the Equation Editor.

-		4 Pramin	m Source?-Non-eve	tem					÷×	Properties
6										vatabase
Г	1	Field Name	Туре	Len	Dec	Parameter	Tag Name	Description	OK	Data
	9	RECEIPT_DATE	Character	8			«No tag»			History
	10	DESCRIPTION	Character	11			<no tag=""></no>		Apper	Field Statistic
	11	SUM_INSURED	Numeric	8	0		<no tag=""></no>		Delet	e Control Total
	12	GROSS_PREMIUM	Numeric	8	1		<no tag=""></no>		Print	Criteria
	13	DISCOUNTS	Numeric	8	3		«No tag»			- neutre
	14	NET_PREMIUM	Numeric	8	1		<no tag=""></no>		Copy	
1	15	POLICY LENGTH	Virtual numeric		0		<no tag=""></no>		Cano	ndices
Ľ				-				~		
1	<							2	Help	No index
	1		666 - CONTRACTOR			arte trattigete	1.00		dy - they	POLICY_NO
		16 MOTOR04	026 P & Co		Group Mo			9 18.08.0		I - Constants
	_	16 MOTOR04 17 MOTOR04	026 P&Co 027 Q&Co		Group Ma Group Ma			9 18.08.0 999 17.09.0		II - Comments
0	Equ	16 MOTORD4 17 MOTORD4 ation Editor	026 P&Co 027 Q&Co							add
0		17 MOTORD4 ation Editor	027 Q & Co	F	Group Mo	otor Mumba				
•	1	17 MOTORD4 ation Editor	027 Q & Co	6	Group Mo	otor Mumba				add
•		17 MOTORD4 ation Editor	027 Q & Co	0	Group Mc	otor Mumba				add
	Equa	17 MOTORD4	027 Q & Co		Group Mc	otor Mumba				add
	Equa	17 MOTORD4 ation Editor	027 Q & Co		Group Mc	otor Mumba		999 17.09.0		add
	Equa	17 MOTORD4	027 Q & Co	0	Group Mc	otor Mumba		999 17.09.0		add
	Equa	17 MOTORD4	027 Q & Co	S	Group Mc	otor Mumba		999 17.09.0		add
	Equa	17 MOTORD4	027 Q & Co		Group Mc	otor Mumba		999 17.09.0		add
	Equa	17 MOTORD4	027 Q & Co		Group Mc	otor Mumba		999 17.09.0 PFunctions Functions on waloble for performing non complex operations uch as date arthmetic, tastiscid calculations		add
	Equa	itiz Motorou Nation Editor	027 Q & Co		Group Mc	otor Mumba		999 17.09.1		add
	Equa	17 MOTORD4	027 Q & Co		Group Mc	otor Mumba		999 17.09.0 PFunctions Functions on waloble for performing non complex operations uch as date arthmetic, tastiscid calculations		add
•	Charles Contraction of the second sec	(Orin (cito string)	027 Q & Co		Group Mc	otor Mumba		999 17.09.1		add
	Char Char 2 Test	IZ MOTORU4 Idlan Editor Solution (@firm(pakcy_no)) (@firm(pakcy_no)) [@firm(cabo atting) Equation	199 (199 (199 (199 (199 (199 (199 (199		Group Mc	otor Mumba		999 17.094		add
	Char Char 2 Test	IZ MOTORU4 Idlan Editor Solution (@firm(pakcy_no)) (@firm(pakcy_no)) [@firm(cabo atting) Equation	027 Q & Co		Group Mc	otor Mumba		999 17.09.1		add
	Char Char 2 Test	IZ MOTORU4 Idlan Editor Solution (@firm(pakcy_no)) (@firm(pakcy_no)) [@firm(cabo atting) Equation	927 Q&G		Group Mc	otor Mumba		999 17.09/ PFunctions #functions #functions one workbin for performing workbin for performing workbin for performing to be dea athenets, ratested cabletors and best exercises. TEA workbin begin with the 0'symbol. They are erry similar in style and peration to functions ound in differ software antopos such as		add
	Char Char 2 Test	IZ MOTORU4 Idlan Editor Solution (@firm(pakcy_no)) (@firm(pakcy_no)) [@firm(cabo atting) Equation	927 Q&G		Group Mc	otor Mumba		999 17.094 PFunctions on wakkin for performing makking opposing transmit cablestone name and bet seedness. It for inclose begin with the of yondub. They are every similar in style and peraktion to functions ound in other selfware actiops such as denoges such as den	ж 3	add
	Char Char 2 Test	IZ MOTORU4 Idlan Editor Solution (@firm(pakcy_no)) (@firm(pakcy_no)) [@firm(cabo atting) Equation	927 Q&G		Group Mc	otor Mumba		999 12.094	ж 3	Add comment

Tip – If @ Len is to be used on a character field as in the case above then always use the function in conjunction with @trim so as to trim the field of spaces before calculating the length. Blank spaces are also interpreted as digits for the purpose of length, and hence the need to trim the field prior to applying the length function.

Now perform a direct data extraction on the computed virtual numeric field.

Write a logical expression

Policy_Length<>10

This expression will extract all records where the policy number does not match the standard numbering convention of 10 digits.

Application of CAATs to Utility Services

File Edit Wew Data Analysis Sampling Tools		■ # E # # # # # #	T. 🔴 🖢
File Explorer 4	Source2-Non-system	* ×	Properties
Construction C	Percent to extract: (© All Stating record # Percents to extract: (© All Stating record # Percents to extract: () All Stating record # Percent to refer to r	KISK_EFFEC 01.06.04 01.05.04.04 00.06.04 00.06.04 00.05.06.04 00.05.06.04 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 01.07.06 11.00.04 00.07.06 00.07.06 11.00.04 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.07.06 00.00	II - Database Data History Field Statistics Control Total Conterns II - Results II - Indices
Premium_Source-Premi Given Editor			I - Comments
Equation policy_length <> 10 [2] Test Equation	Image: Second		Add comment

The extraction gives two cases where the numbering convention has not been met.

In the case of G & Co and Perl the policies do not follow the standard convention of 10 digits and actually show 12 digits.

These cases need to be taken up for further investigation with the development officer and / or agent.

The insured would need to be contacted in secret to establish his or her identity / validity.

Explorer 4	Dramam Sou	rce2-Non-system policier	EXTRACTION			• X	Properties
222 4 3	POLICY_NO	INSURED_NAME	VEH_REG_NO	LOCATION	DEV_OFFICER_AGENT	RISK_EFFECT	II - Database
A Flore Name Carlboard Forman Marker Carlboard Forman Marker Premiss, Societ Forman Marker M		G B CO	Group Motor Group Motor	Munita Munita		13.06.04	Deta Hetory Hetory Field Statistic Control Total Control Total Control Total The stats H Indices H Indices No index Il Comments Add comment
	e					×	

IDENTIFY GAPS

Let us now identify missing policies.

Tip – The Gap test has to be run on the file containing only those policies that meet the 10 digit numbering convention so that only real gaps are identified, and not gaps on account of differences between the 10 digit and 12 digit policies.

Extract a file containing only the 10 digit policy numbers.

Scroll to the virtual field 'Policy_Length' as created in the last exercise.

Locate the first record containing 10.

Place your cursor over the cell, right click, and scroll to the option "Display all records containing "10".

Click OK.

Application of CAATs to Utility Services

Explorer 9	Premium	a Cource2.	Non-system		and the second	Contract of the second designed	- ×	Properties
Analysis-Tide Clients Earthquake Premium Masters Motor Premium Master Premium Source1 Premium - Mot Premium Source2Non-system p. Divelacte	4 5 6 7 7 10 10 11 11 12 13 14	500000 65000 000000 45000 65000 5000 500000 500000 500000 500000 50000 50000 50000 50000 5000	ent display criteria	DISCOUNTS 0.00 0.00 1,220000 0.00 1,020000 0.00 0.00 0.00 0.00 0.00 0.00		POLKY_LENGTH 10 10 10 10 10 10 10 10 10 10		II - Database Data Hetory Field Stockiele Control Total Criteria II - Results II - Results II - Indices Notice Politicy (NO) II - Comments Add comment
	¢						2	

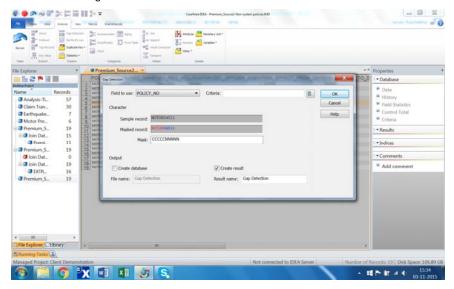
The view generated contains all the policies having 10 digits only.

🗟 🗠 🎭 🐜 📝 🖄 🛙	n 8	A LAN I A					he 1992	T1 👝
le Explorer 4				Bt 🗚				
		Premium_Source	2-Non-system				**	Properties
🎑 🔁 🥙 📕 🖋 词 👘		SUM_INSURED	GROSS_PREMIUM	DISCOUNTS	NET_PREMIUM	POLICY_LENGTH		II 🕶 Database
	1	500000	12,500.0	0.00	12,500.0	10		🥒 Data
	2	65000	1,625.0	0.00	1,625.0	10 10		History
EA Files	1 2 3 4 5 6 7	1000000 45000	25,000.0 1.125.0	1,250.000 0.00	23,750.0 1,125.0	10		Field Statistic
Name	5	650000	13.000.0	0.00	13,000.0	10		-
Canalysis-Tide Clients	6	100000	2,000.0	0.00	2,000.0	10		Control Total
Earthquake Premium Masters	7	65000	1,300.0	0.00	1,300.0	10		Criteria:
Motor Premium Master	8	75000	1,500.0	0.00	1,500.0	10		POLICY_LENG
Premium Source1-Premium - Mot	9	95000	2,375.0	0.00	2,375.0	10		= 10
Premium Source2-Non-system p.	10	1250000	31,250.0	1,250.000	30,000.0	10		II 🖛 Results
Duplicate	11	900000	22,500.0	900.000	21,600.0	10		
EXTRACTION1	12	35000	612.5	0.00	612.5	10		
Premium_Source-Premium - Earth	13	85000	1,487.5	0.00	1,487.5	10		II - Indices
	14	75000	1,875.0	0.00	1,875.0	@len(@trim(policy_no))		🥒 No index
	15	750000	13,125.0	328.125	12,796.9	10		
	12 13 14 15 16 17	45000 850000	787.5 21.250.0	0.00 2.125.000	787.5 19.125.0	10 10		POLICY_NO/
	٣	000000	21,20010	2,223,000	19,120.0	10		II - Comments
								bba 🛢
								comment
X	<						>	
🖁 Running Tasks 🔮 To-Do List [🚮 Se	arch R	tesults 🛛 🚑 Conclusions						
Help, press F1			Weeking	Colders Coldenaria	Files\IDEA V 7.3\Insura	nce Research Number of Rec	oude: 17/10	Disk Space: 11.55

Proceed to run a gap test on this view to locate the missing policy numbers.

As the Policy No is an alpha-numeric field type, navigate to Character Gap Detection under Data in the menu tool bar.

The MASK is automatically generated by IDEA in the Character Gap Detection dialog box.



Click OK.

After running the gap test, the result appears as a TAB along with Database, History and Field Statistics.

The test reveals three gaps / missing policy numbers from the list of manual policies.

The cases need to be further investigated as to whether the policies were really not created or created but not entered into the system by the agent.

Application of CAATs to Utility Services

e Explorer 4		2 🕥 000 💷 🗖 🍋 📷 🔣		
i takar 🗸 🕨 📢 🟹	Premium_Source2-Non-sys	stem		A Properces
	8 8 · 💀 ·			Data
U	Key Value MOTOR			History
EA Files	From: POLICY_NO	To: POLICY_NO	Number	Field Statistic
Name	■ MOTOR04013	MOTOR04013	1	🗧 Control Total
- Earthquake Premium Masters	HOTOR04017	MOTOR04017	1	Criteria: POLICY_LENC
- 1 Motor Premium Master	B MOTOR04022	MOTOR04022	1	= 10
Premium_Source1-Premium - Mot Premium_Source2-Non-system p.				II 🕶 Results
 Duplicate EXTRACTION1 		Total number of items detected	3	🖌 Gap Detectio
8- 🧊 Premium_Source-Premium - Earth		Total number of gaps detected	3	II 🖛 Indices
				POLICY_NO/
				Add comment
🖁 Running Tasks [🌛 To-Do List 🛛 🛃 Se Help, press F1	earch Results 🛛 🖓 Conclusions		7.3/Unsurance Research Number of Record	ds: 17/19 Disk Space: 11.55

6.0 IDENTIFYING POLICIES CANCELED SOON AFTER INCEPTION ALONG WITH JUSTIFICATION BY ENDORSEMENTS

OBJECTIVE

To verify whether any policies were canceled soon after risk inception, and if so, whether they are supported by endorsement notes.

EXERCISE DESCRIPTION

In this exercise we will write an @ Function to arrive at the time lag between policy cancellation and policy issue.

Extract policies canceled shortly after issue for close scrutiny.

Extract policies canceled, not backed by endorsement notes.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Append a Field
- Extract data with a specified criterion.
- Use @ Functions

Data Import

Import the file Premium Source 3-Canceled Policies. xls into IDEA as per the steps in IMPORT OF DATA FILES. (check if this change is appropriate?)

BACKGROUND

The insurance company management audit guidelines have a norm that policies canceled within 7 days of risk commencement should be identified and taken up for close scrutiny along with the endorsement records and documentary proof.

APPEND A FIELD

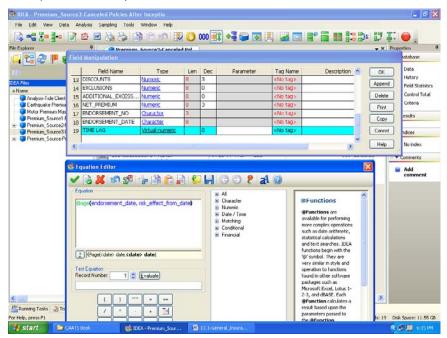
Append a field through Field Manipulation with the following specifics:

Field Name:- Time Lag

Type: Virtual Numeric

Write an @ Function in the Equation Editor.

@age gives the difference between two dates. In this specific test we arrive at the difference between the Risk Effect from Date and the Endorsement Date (Canceled Date)



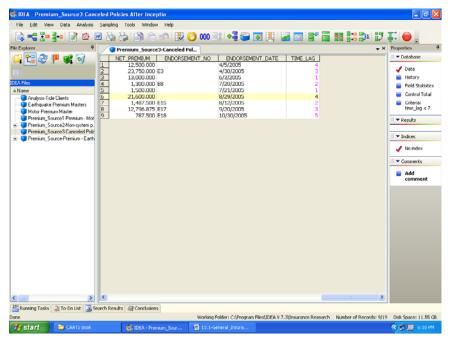
EXTRACT DATA

Based on the virtual numeric field, Time Lag isolates all the policies canceled within 7 days of risk commencement.

Perform a data extraction using the equation

Time Lag < 7

The extraction gives 9 cases of policies canceled within 7 days of risk commencement.



As per the Management Audit guidelines, all policies canceled within 7 days of risk commencement, should be supported by an Endorsement note reference.

Perform a data extraction using the equation

Endorsement_No=" "

The extraction gives 4 cancellations without Endorsement references.

IDEA - Premium_Source3-Canc						. 8
File Edit View Data Analysis 🚉 📲 🌠 📑 🛥 📝 🛃 🖻			1 📲 😂 💽 👢		18 8-0 Ghail 99 1	-
📸 📲 🎦 📑 🖬 📝 🖄 🖻			* 🚜 👝 🐻 64 8			100
	Premium_Source: NET_PREMDUM		ENDORSEMENT DATE	TIME LAG	+ ×	Properties
A Hee None Catilyake Penium Matters Catilyake Penium Matters Catilyake Penium Matters Catilyake Penium Matter Catilyake Penium Matter Catilyake Penium Matters Catilyake Penium Matters Penium Source Plenium - Earth	1 2 12,500,000 2 13,000,000 3 1,500,000 4 21,600,000	ENGORGEMENT. NO	445/2005 67/2005 7/21/2005 9/29/2005			Data Hictory Field Statistic Control Total Criteria: time.log < 7 ANO. If < Results If < Indices If < Comments Comments Comments
	c.				>	
Running Tasks 🌙 To-Do List 🛃 Se		1				
	sarch Results 🔐 Conclusions		older: C:\Program Files\IDEA V 7.			Disk Space: 11.55
ne						

CONCLUSION

It seems that the insurance application system does not compulsorily insist on electronic validation of canceled policies through endorsement note references. This is a major shortcoming of the software.

The development officer should not be able to proceed with policy cancellation unless the event is authorized by an endorsement reference with valid reasons.

Moreover. policy cancellation should not be the prerogative of just one officer. At least, two officers should handle this sensitive high exposure area.

All canceled policies should get reflected in the exception statement, which should be duly checked and authorized by the Branch Manager on a daily basis, filed and kept on record.

7.0 INPUT FIELD CHECKS ON COVER NOTES (TEMPORARY POLICIES) FOR MOTOR INSURANCE AND CONVERSION TO BONA FIDE POLICIES THROUGH RECEIPTS

OBJECTIVE

To verify if there is any missing, duplicate, fabricated cover note and whether receipts support all cover notes.

EXERCISE DESCRIPTION

In this exercise we will run tests for duplicate detection on the Cover Note field.

Specific tests for number format verification will be undertaken on the Cover Note field to establish conformity with the standard numbering pattern.

Gap tests will be run to identify missing cover note numbers.

Tests to identify cover notes not supported by receipts

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Identify Duplicates
- Perform Number format Tests
- Display Function
- Identify Gaps
- Extract data with a specified criterion.
- Use @ Functions

Data Import

Import the file Premium Source 4 – Cover Note. xls into IDEA as per the steps demonstrated in the session on IMPORT OF DATA FILES.

IDENTIFY DUPLICATES

Select Duplicate Key Detection from Analysis on the Menu tool bar.

This opens the Duplicate Key Detection dialog box.

Click Key. This opens the Define Key dialog box. Let the Field be defined as COVER_NOTE.

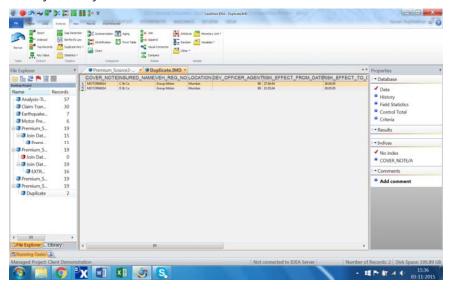
Click OK. Check Output Duplicate Records in the Duplicate Key Detection dialog box. Click OK.

Area Protect	Ben brachaster BJ Post lass	Attribute 🛄 Haveley UNI *		_	
Konstanting of the second	Constant Anno Anno Anno Anno Anno Anno Anno An	Depinet for Densition	SE OK Key Fields Cancel Help Marrie Marrie Marrie Cancel Help Cancel Can	FROM CATPOIN EFFECT TO	Properties
III Flate Epiderer Library / Rueving Tasks Rueving Tasks Rueving Caset. Client Demons 		Net co	Gancel Help	and the second se	Neords 19 Disk Space 106.69

The test gives us two instances of cover notes having the same number ${\sf MOTOR04014}.$

These cases need to be taken up for further investigation with the development officer and / or agent.

The Insured would need to be contacted in secret to establish his or her identity / validity.



PERFORM NUMBER FORMAT TESTS

Each cover note is in a standard format of 10 digits and has an alpha numeric field type.

The purpose of this test is to identify cover notes which do not match the standard number format.

In the first stage we write an @ Function to give us the actual length of each cover note.

Append a field through Field Manipulation with the following specifics:

Field Name: Cover Note Length

Type:Virtual Numeric

Write an equation @len(@trim(cover_note)) in the Equation Editor.

Tip – If @ Len is to be used on a character field as in the case above then we always use the function in conjunction with @trim so as to trim the field of spaces before calculating the length. Blank spaces are also interpreted as digits for the purpose of length, and hence the need to trim the field prior to applying the length function.

Now perform a direct data extraction on the computed virtual numeric field.

Write a logical expression

Cover_Note<>10

This expression will extract all records where the cover note does not match the standard numbering convention of 10 digits.

The extraction gives two cases where the numbering convention has not been met.

In the case of G & Co and Perl the cover notes do not follow the standard convention of 10 digits..

These cases need to be taken up for further investigation with the development officer and / or agent.

The insured would need to be contacted in secret to establish his or her identity / validity.

IDENTIFY GAPS

Let us now identify missing cover notes.

Tip – The Gap test has to be run on the file containing only those cover notes that meet the 10 digit numbering convention so that only real gaps are identified, and not gaps on account of differences between the 10 digit and 12 digit cover notes.

Extract a file containing only the 10 digit Cover Notes.

Scroll to the virtual field 'Cover_Note_length' as created in the last exercise.

Locate the first record containing 10.

Place the cursor over the cell, right click, and scroll to the option "Display all records containing "10".

Click OK.

The view generated contains all the cover notes having 10 digits only.

Proceed to run a gap test on this view to locate missing cover note numbers.

As the Cover Note is an alpha-numeric field type, navigate to Character Gap Detection under Data in the menu tool bar.

The MASK is automatically generated by IDEA in the Character Gap Detection dialog box.

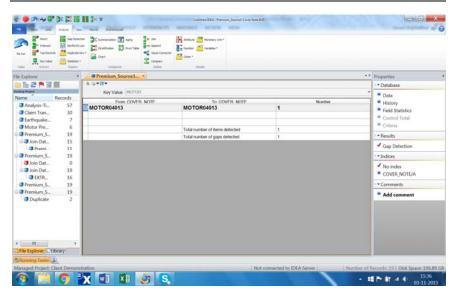
Click OK.

After running the gap test, the result appears as a TAB along with Database, History and Field Statistics.

The test reveals three gaps / missing cover note numbers.

The cases need to be further investigated to confirm if the cover notes were really created or created but not entered into the system by the agent.

Application of CAATs to Utility Services



EXTRACT DATA WITH A SPECIFIED CRITERION

Every cover note should be supported by a receipt generated by the system.

Proceed to identify cover notes not supported by receipts.

Perform a data extraction using the equation

Receipt_No=" "

The extraction reveals 5 cover notes not backed by receipt numbers.

IDFA - Premium_Source4-Cove	r Note							
		Tools Window	Help					
🗟 📲 🔚 🚰 📝 🖄 🖻	🛛 🔌 S		ଜା 🔡 🄇	5) 000 💷 📲 🎬) 💽 👢 🐷 🖸	1 2" 🖬 🖩 🕽	} Σ 17	II 🔴 🚬 👘
e Explorer 🛛 🗘	🖉 Pri	emium_Source	-Cover Note				• ×	Properties
A Flos A Plos A Control Clent:	1 2 3 4 5		RECEIPT_D	ATE DESCRIPTION Own Damages Own Damages Own Damages Third Park Own Damages	SUM_INSURED 45000 65000 178500 178500	GROSS PREMIUM 1,1250 2,0000 1,0000 44,0250 1,407.5	DISCOLANT 0. 0. 3,570. 0.	Database Data Dat
	<						2	
🚰 Running Tasks 🛛 💆 To-Do List 🔃 🚮 Se	arch Results	Conclusions						
r Help, press F1				Working Folder: C:\Progr	am Files\IDEA V 7.3\Insur	ance Research Number (of Records: 5/19	Disk Space: 11.55 G
🛃 start 📄 CAATS Book		👹 IDEA - Premi	um_Sour	13.1-General_Insura				🔇 🗊 🔜 6:22 PM

CONCLUSION

It seems that the insurance application system does not compulsorily insist on electronic validation of cover notes while converting them into bona fide policies through the receipt and cover note conversion process. This is a major shortcoming of the software.

The development officer and / or agent should not be able to proceed with policy booking unless the cover note is coupled with the receipt.

13.1.7 Claim Case Studies

1.0 IDENTIFICATION OF DUPLICATE MOTOR CLAIMS IN PARTIAL LOSS POLICIES

OBJECTIVE

To identify duplicate motor claims in partial loss policies.

EXERCISE DESCRIPTION

In this exercise we will run tests for duplicate detection on the policy number and vehicle registration number field.

Prior to running the duplicate detection, we will isolate partial loss claims from the file of total and partial loss claims.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Identify Duplicates
- Extract data with a specified criterion.

Data Import

Import the file Claim_Source - Transaction. xls into IDEA as per the steps in IMPORT OF DATA FILES.

EXTRACT DATA WITH A SPECIFIED CRITERION

A close examination of the file imported will reveal that the claim file is an assortment of total and partial loss claims.

The first step in this test is to isolate the partial loss claims, and convey these claims into a separate file.

Run a direct data extraction on the file imported through the equation

Nol = "Partial Loss"

IDENTIFY DUPLICATES

Now proceed with running a duplicate key detection test on the file Partial Loss. IMD generated through the data extraction.

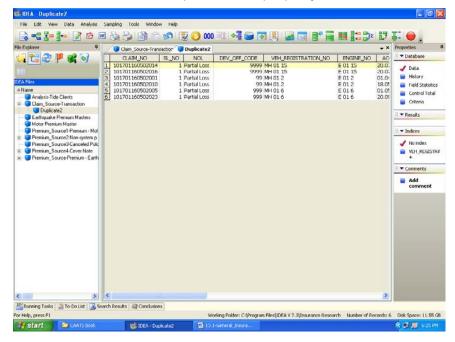
The objective of the test is to identify claims having the same policy number and vehicle registration number.

So select POLICY_NO and VEH_REGISTRATION_NO as the Key fields in the Define Key dialog box.

IDEA will run the duplicate check on these specific fields only.

Sile Explorer 9			Properties 9
Earliquade Prenum Nades Motor Prenium Nades Motor Prenium, Souto I-Inerxion. Mot Prenium, Souto I-Inerxion. Mot Prenium, Souto I-Inerxion. Mot Prenium, Souto I-Convelt folde Prenium, Souto - Convelt folde Prenium, Souto - Convelt folde Prenium, Souto - Convelt folde	CLUMP, NO SL, NO NO DEV. CFE. CODE VEH. REGISTRATION 101701160502000 Duplicate Key Detection X	N.300 ENGINE 400. PANILE 401. 31.1 E 012 0.01.2 31.1 E 012 0.01.2 24.4 E 013 24.4 24.4 E 014 20.01 24.4 E 015 0.1.1 61.0 0.1.1 E 016 0.1.1 61.0 0.1.1 E 017 0.4.2 20.0 60.0 E 018 133.2 20.0 60.0 E 0110 11.4 20.0 60.0 60.0 E 0112 100.0 E 01.15 20.0 60.0	II T Results
· · ·		2	

The test gives cases of Policy Nos 11050180000075, 11050180000235 and 11050180000666 that have duplicate claims per policy.



CONCLUSION

Since these claims are towards settlement of partial losses, the list of duplicates may not reveal inconsistencies or irregularities.

Some policyholders may have genuinely suffered accidents repeatedly on the same vehicle as seen above. However, one cannot deny the need for an inquiry into these duplicates.

The internal audit team should take up these cases for further scrutiny with the claim file, surveyor reports, and picture evidence.

These cases serve to provide empirical evidence to r the audit team of negligent policyholders or policyholders who continuously run into bad luck, and validate future premium rate enhancements on such policies.

2.0 IDENTIFICATION OF DUPLICATE MOTOR CLAIMS DUE TO IRREGULARITIES AND DISCREPANCIES IN SURVEY

OBJECTIVE

To identify repetitive motor claims due to irregularities and / or discrepancies in the surveyor's report.

EXERCISE DESCRIPTION

In this exercise we will run tests for duplicate detection on the policy number and accident date field.

Prior to running the duplicate detection, we will isolate partial loss claims from the file of total and partial loss claims.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Identify Duplicates
- Extract data with a specified criterion.

Use the child file Partial Loss. IMD generated in the earlier exercise.

IDENTIFY DUPLICATES

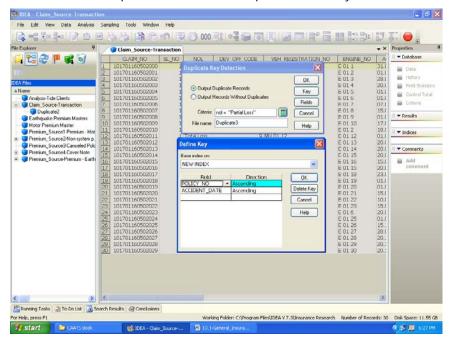
Run a duplicate key detection test on the file Partial Loss. IMD

The objective of the test is to identify claims having the same policy number and accident date.

The reason for selecting the ACCIDENT_DATE as a critical key field is to give weightage to oversight, and / or deliberate manipulation in the survey

results by the surveyor. It is not far fetched to visualize a situation in which the surveyor in connivance with the insured tries to fabricate an accident along with the required documents. To trap such cases we also select ACCIDENT_DATE as a critical field.

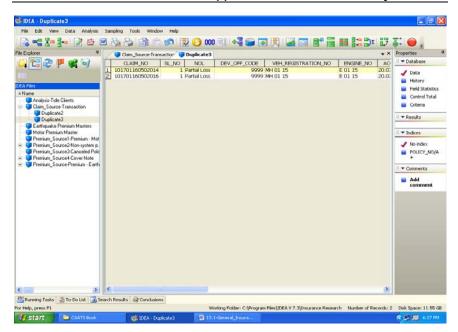
So select POLICY_NO and ACCIDENT_DATE as the Key fields in the Define Key dialog box.



IDEA will run the duplicate check on these specific fields only.

The test gives (2 cases or one case?) of Policy No 11050180000666, which has (witnessed?) claims twice.

Application of CAATs to Utility Services



CONCLUSION

The internal audit team should take up (these or this) (cases or case) for further scrutiny with the claim file, surveyor reports, and picture evidence.

If the internal audit head suspects long standing association between particular surveyors, insured and even development officers of the insurance company, then an attempt should be made to gather vital information on past dealings for such surveyors, insured and development officer to examine the length and depth of their involvement.

As a mode of building on a case for systematic well informed fraud allegation, the internal audit head should interview close associates of such surveyors, and development officers in strict confidence to attempt identifying rampant, extravagant changes in life style patterns, social habits and other references.

3.0 ISOLATION OF POLICIES EXPERIENCING CLAIMS SHORTLY WITHIN RISK COMMENCEMENT - CLOSE PROXIMITY

OBJECTIVE

To identify policies experiencing claims within a short span of risk commencement – Close Proximity Claims

EXERCISE DESCRIPTION

In this exercise, we will join the claim transaction file to the policy transaction file, so as to capture the risk commencement date and the accident date in a joined file.

We will then append a computed - virtual numeric field containing an @ Function to arrive at the difference between the risk commencement date and claim date.

Subsequently, we will extract those cases where claims have been submitted within 5 days of risk commencement, as per the difference in dates field computed above.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Join Files
- Append a Field
- Write an @ Function
- Extract data as per Logical Criterion

BACKGROUND

As per the internal guidelines, close proximity claims are claims, which appear within 5 days of risk inception against the relevant policy.

DATA IMPORT

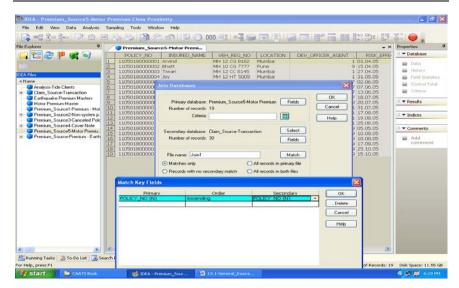
Import the file Premium Source-Motor Premium Close Proximity. xls into IDEA as in IMPORT OF DATA FILES. (check for accuracy)

JOIN FILES

Designate Premium_Source 5 – Motor Premium Close Proximity. IMD as the Primary Database and Claim Transaction. IMD as the Secondary Database.

Select the Policy_No field from both the Databases as the matching field.

Application of CAATs to Utility Services



CONCLUSION

Close proximity claims should be viewed very carefully by the internal audit team for its occurrence, nature and frequency. The auditors should try to establish whether these claims have genuinely arisen due to the bad luck of the insured or whether they have resorted to any malpractice with the intent to deceive or defraud the insurance company. The role of the development officer and / or surveyor should also be investigated in such cases through discussions and inquiry. The auditor should apply the principles of Arms Length dealing and fair practices in such cases.

4.0 IDENTIFICATION OF CLAIMS WITHOUT BONAFIDE POLICY NUMBERS WITHIN THE SYSTEM – ORPHAN CLAIMS

OBJECTIVE

To identify claims processed and settled without bona fide policy numbers within the system.

EXERCISE DESCRIPTION

In this exercise, we will run an extraction to identify claims processed without matching policy numbers within the system.

The exceptions will be independently validated with the Premium Non-System Policies file to prove the existence and validity of claims.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Extract data as per Logical Criterion
- Use the Search Function

BACKGROUND

Orphan claims are claims admitted, processed and paid, for which there are no matching policy numbers within the insurance application system.

These claims generally pertain to policies that have been booked prior to system migration from old legacy systems.

These claims need to be dealt with separately by the auditors, as there is no internal control mechanism for cross-reference within the system.

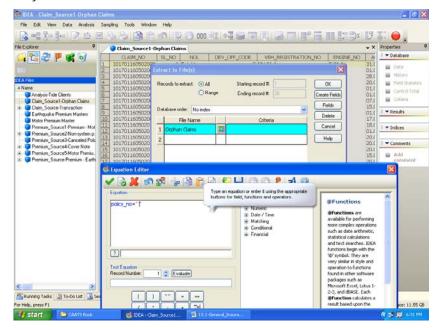
DATA IMPORT

Import the file Claim Source-Orphan Claims. xls into IDEA

EXTRACT DATA AS PER LOGICAL CRITERION

Perform a data extraction through the equation

Policy_no=" "



The extraction will give us an independent file of 8 claims paid where the policy number field is blank.

		Claim_Source1-On	phan Claime	😈 Orphan Cl	laims			÷×	Properties
Constant of the second se	8 × 10 10 10 10 10	0.40M NO 10170116550205 01070116550205 010701165502015 101701165502015 101701165502015 10170116550205 101701165502056	1 1 1 1 1	NOL Partial Loss Partial Loss Total Loss Total Loss Partal Loss Partal Loss Total Loss Total Loss	99 9 9999 9999 9999 9999 9999	VH-REGISTRATION_NO M4 01 6 M4 01 7 M4 01 10 M4 01 12 M4 01 12 M4 01 15 M4 01 15 M4 01 22 M4 01 23 M4 01 27	Profile_INC Cold Cold	A0 01.0¢ 01.0¢ 20.0; 20.0; 10.0¢ 01.0; 28.0¢	Postabase Data Data Hictory Eicld Statistics Control Total Criment Voriments Add comments

USE THE SEARCH FUNCTION

To prove the validity of the 8 claims, compare the claim details with the available list of non-system policies.

A direct join with the Insured Name as the matching field may not yield the right results as consistency in the name entry may not be maintained by the development officers and the claim officers.

We now introduce the concept of DATA SEARCH where variants of names as entered by the claim officer can be located and reviewed.

Open the file Non system policies. prn. IMD in IDEA

The first claim paid is pertaining to Arjun S.

Now navigate to Search within Data

In the box 'Look for the following text' enter *arjun*.

The logic of using the string 'arjun' is to include every possible variant and juxtaposition of the name and intial. ' denotes a wild card character that can take any form.

The search needs to be made in the field INSURED. Check the field box for INSURED.

The search should be character case independent, so we do not check the box MATCH CASE SENSITIVITY

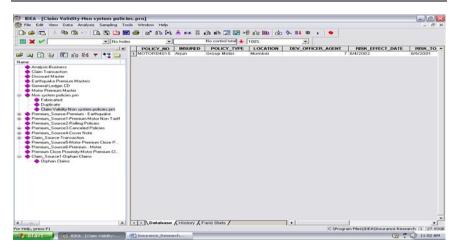
The output should be captured in an independent file, so check 'Create an extraction database'

File Explorer	P Claim, Source1-Ornhan Claims		× Properties
📮 😋 🧶 📕 💕 🗟	Search		Av Database
CEA Hes Alone Catagoria Tide Clerits Catagoria Claming Clerits Catagoria Pernium Nater Catagoria Pernium Nater Catagoria Pernium Nater Pernium Souce4 Pernium Catagoria Souce4 Non-pyt Pernium Souce4 Cover N Pernium Souce4 Clover N Pernium Souce4	Text to find.	Lock notifier defabases 2 Fields to lock in Type 2 Name ESTIMATED_D>PENSE_A. Nummo: 2 PADD_DUSS_MOUNT Nummo: 1 PADD_DUSS_MOUNT Nummo: 1 VUUCHER_NO Nummo: 1 VUUCHER_NATE Owardin 1 VUUCHER_NATE Owardin 1 VUUCHER_NATE Owardin 2 CHEQUEY_NO Nummo: 2 CHEQUE_DATE Owardin: 2 CHEQUE_DATE Owardin: 2	1.1 Osco 1.1 Pield Stotistics 2.1 Carbol Total 2.1 Carbol Total 1.1 Carbol Total 1.1 Carbol Total 1.1 In Results 1.1 Tables 1.1 Comments 1.1 Comments
	About the Search Task, Banlean Op To Perform a Search Wildcard Sr Searching Data: Fields Proximity 5 Searching Numeric Fields	Search Features	20 LL LL LL LL SL SL SL SL SL SL SL SL SL
< com	Help	OK. Cancel	2

The SEARCH gives us a single policy record with the name Arjun. This proves the claim validity.

If an attempt had been made to match the claim file and the policy file on the Insured field, no matches would have resulted, because the policy file has **Arjun** and the claim file has **Arjun S**.

Application of CAATs to Utility Services



CONCLUSION

Orphan cases need to be specially audited by the internal audit team.

Details should be sought from the claims officer and after running our own search tests as seen above, identifying the corresponding manual policy numbers.

The policies should be carefully studied and examined by the auditors to check on the insured, asset, period of coverage, sum insured and type of policy, to prove the validity of the claim.

5.0 AS PART OF OPERATIONAL STUDIES AND WORK FLOWS IDENTIFICATION OF IRREGULAR TIME LAGS IN CLAIM SETTLEMENT

OBJECTIVE

To identify irregular time lags in claim settlement due to process delays in survey.

EXERCISE DESCRIPTION

In this exercise, we will append computed – virtual numeric fields to arrive at differences between the accident date, surveyor intimation date, surveyor report receipt date and cheque paid date.

The day differences in each of the survey related activities will then be summarized surveyor-wise.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Append Fields
- Use the Equation Editor
- Apply Quick Field Summarization

DATA IMPORT

Import the file Claim Source-Surveyor Activity. xls into IDEA as in IMPORT OF DATA FILES (check?).

APPEND FIELDS, USE THE EQUATION EDITOR

Append 4 computed virtual numeric fields with the following specifics:

Field Name: Gross Time Lag

Type:Virtual Numeric

Parameter: @age (Cheque_Date, Accident_Date)

oplorer	a Claim	iniera7.Sievevor (maina					+ X Pr	operties
22 2 1									atabase
	Field Name	Type	Len	Dec	Parameter	Tag Name	Description	OK	Data
Hes	2 ACCIDENT_DATE	Character	8	1000		<no tag=""></no>			History
me	3 CHEQUE DATE	Character	8			«No tag>		Append	Field Statistic:
Analysis-Tide Client	4 ESTIMATED_LOSS_A	Numeric	0	0		<no tag=""></no>		Delete	Control Total
Claim_Source1-Orp	5 SURVEYOR_CODE	Character	3			<no tag=""></no>		Print	Criteria
Claim_Source2-Sur	6 INTIMATION_DATE	Character	8			<no tag=""></no>			esuks
Claim_Source-Tran	7 REPORT_RECEIPT_DA	Character	8			<no tag=""></no>		Copy	COURS .
Motor Premium Mas	8 GROSS TIME LAG	Virtual numeric		0		<no tag=""></no>		Cancel	ndices
Premium_Source1-F		and the second second second	-				~		naces.
Premium_Source24	<						() ()	Help	No index
Premium_Source4-Co Premium_Source5-Mr	nk 🐝 Equation Editor	0902015 1507 09 🔊 👍 🗈 🖡		51 O		al 🧿	1000000 5 1		Add comment
Premium_Source3(Premium_Source4-Ce Premium_Source5-M Premium_Source7re	eto mi Leguation Editor California Equation	st 👆 🗈 f	Type on	equation or	enter it using the app	propriate			Add
Premium_Source4-Co Premium_Source5-Mr	nk 🐝 Equation Editor	st 👆 🗈 f	Type on	equation or	008	propriate	@Functions		Add
Premium_Source4-Co Premium_Source5-Mr	eto mi Leguation Editor California Equation	st 👆 🗈 f	Type on	equation or	enter it using the app ctions and operators	propriate	@Functions @Functions are		Add
Premium_Source4-Co Premium_Source5-Mr	eto mi Leguation Editor California Equation	st 👆 🗈 f	Type on	equation or	enter it using the app ctions and operators	propriate	@Functions @Functions are available for performing		Add
Premium_Source4-Co Premium_Source5-Mr	eto mi Leguation Editor California Equation	st 👆 🗈 f	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions #Functions are available for performing more complex operations such as date antifranter,		Add
Premium_Source4-Co Premium_Source5-Mr	eto mi Leguation Editor California Equation	st 👆 🗈 f	Type on	equation or	enter it using the app ctions and operators	propriate	©Functions @Functions are available for performing more complex operations such as date arithmetic, stabistical calculations		Add
Premium_Source4-Co Premium_Source5-Mo	er C Equation Editor	scident_date	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions #Functions are available for performing more complex operations such as date antifranter,		Add
Premium_Source4-Co Premium_Source5-Mr	eto mi Leguation Editor California Equation	scident_date	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions @functions are available for performing more complex operations such as date arithmetic, statistical calculations and text searches. IDEA functions begin with the (92 symbol. They are		Add
Premium_Source4-Co Premium_Source5-Mr	de Cuation Editor Causion Ca	scident_date	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	©Functions @Functions are available for performing more complex operations such as date arithmetic, statistical calculations and text searches. IDEA functions begins with the 'd' symbol. They are very smallen in otyle and		Add
Premium_Source4-Co Premium_Source5-Mr	Caultion Editor Causion Caus	accident_date)	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions @functions are available for performing more complex operations such as date arithmetic, statistical calculations and text searches. IDEA functions begin with the (92 symbol. They are		Add
Premium_Source4-Co Premium_Source5-Mr	Caultion Editor Causion Caus	scident_date	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions #functions are available for parforming more complex operations and text searches. IDEA functors begins with the Var symbol. They are very similar in tybe and for all in other software packages such as		Add
Premian_SourceFd Premian_SourceFH Premian_SourcePre	Caultion Editor Causion Caus	accident_date)	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	Functions wronking are analytic pre-mong analytic		Add
Premium_Source4-Co Premium_Source5-Mr	C C C I quation / ditor	accident_date)	Type on	equation or	enter it using the app chons and operators Date / Time Matching Conditional	propriate	@Functions #functions are available for parforming more complex operations and text searches. IDEA functors begins with the Var symbol. They are very similar in tybe and for all in other software packages such as		Add

Field Name: Pre – Internal Time Lag

Type:Virtual Numeric

Parameter: @age (Intimation_Date, Accident_Date)

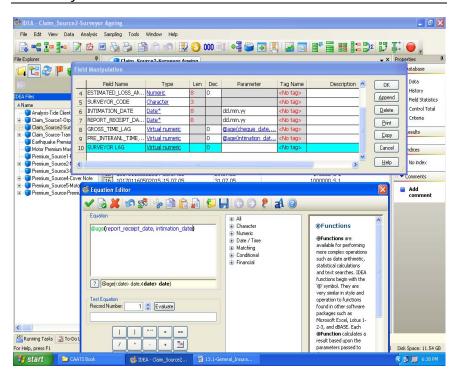
Application of CAATs to Utility Services

[mra2.Survavor A	naina					- X Pro	oerties :
22 2 1	Field	d Manipulation								atabase
		Eield Name	Type	Len	Dec	Parameter	Tag Name	Description	OK	Data
	3	CHEQUE DATE	Date*	0		dd.mm.yy	<no tag=""></no>			History
iles me	4	ESTIMATED LOSS AM	Numeric	8	0		«No tag»		Append	Field Statistics
Analysis-Tide Client	5	SURVEYOR CODE	Character	3			«No taga		Delete	Control Total
Clam Source1-Orp		INTIMATION DATE	Date*	8	-	dd.mm.yy	<no tag=""></no>		Ent	Criteria
Claim_Source2-Sur		REPORT_RECEIPT_DA	Character	8	-		<no tag=""></no>		Elex	esuts
Clam_Source-Tran	8	GROSS TIME LAG	Virtual numeric		0	@age(cheque_date	<no tag=""></no>		Copy	esuks
Earthquake Premiur Motor Premium Mac	9	PRE INTERANL TIME L.	Virtual numeric		0		<no tag=""></no>		Cancel	ndices
Premium Source1-F	-			-	-			~		Races
Premium_Source24	<	9						2	Help	No index
Premium_Source30 Premium_Source4-Co		lote 16 101701160		_					1	Comments
		🗸 🗟 🖊 🖉 🧋	ै 😓 🗋 🕯		6	. 00 ?	al 🗿			comment
		Equation deage(Intimation_date, (2))(Reget(date> date, c) TestEquation			_	Al Al Character Numeric Date / Time Matching Control into Financial		Functions available for performing more complex operations such as date arithmetic, statistical calculations and text scanders. IDEA functions begin with the 10' symbol. They are very smaller in style and operation to functions found in other software		

Field Name: Surveyor Lag

Type: Virtual Numeric

Parameter: @age (Report_Receipt_Date, Intimation_Date)



Field Name: Post – Internal Lag

Type: Virtual Numeric

Parameter - @age (Cheque Date, Report_Receipt_Date)

Application of CAATs to Utility Services

le Explorer	Fiel					🗊 🗣 🍧 🛃 🐺				
			IFCO7-SIIFVOVOF Ar					In the second	🗙 🗙 🛛 Pro	100
										atabase
	-	Field Name	Туре	Len	Dec	Parameter	Tag Name	Description	ОК	Data
	5	SURVEYOR_CODE	Character	3			<no tag=""></no>			History
Files ame	6	INTIMATION_DATE	Date*	8		dd.mm.yy	<no tag=""></no>		Append	Field Statistic
ame 🕤 Analysis-Tide Client	7	REPORT_RECEIPT_DA	Date*	8	-	dd.mm.yy	<no tag=""></no>		Delete	Control Total
Claim_Source1-Orp		GROSS_TIME_LAG	Virtual numeric		0	@age(cheque_date,	<no tag=""></no>		Print	Criteria
😈 Claim_Source2-Sur		PRE_INTERANL_TIME	Virtual numeric		0	@age(intimation dat	<no tag=""></no>			esults
😈 Claim_Source-Tran: 🔲 Earthquake Premiu		SURVEYOR_LAG	Virtual numeric		0	@age(report_receipt	<no tag=""></no>		Сору	iosuits
Eartnquake Premiur Motor Premium Mas	1000	POST INTERNAL LAG	Virtual numeric		0		<no tag=""></no>		Cancel	ndices
Premium_Source1-F				2				×		
Premium_Source2-1								>	Help	No index
Premium_Source3-0 Premium_Source4-0		Note 16 101701160				07.05		1000000 5 1		Comments
	i.	V 🗟 样 🕼 🕉 Equation	8 🗞 🗎 😭		6		A 💿		_	
		@age(cheque_date, re	port_receipt_date	D	_	 Character Numeric Date / Time Matching Conditional Financial 		@Functions @Functions are available for performing more complex operations such as date arithmetic, statistical calculations and text searches. IDEA functions begin with the		

APPLY QUICK FIELD SUMMARIZATION

The Claim Source-Surveyor Activity. IMD file appears, as can be seen below, after appending the virtual fields.

Navigate to Quick Field Summarization in the menu tool bar.

	Ø 🏹		🛛 🐺 🜔 000 💷 📲 🧯		1 8 H 2 2 2 0 D	
s Explorer 🕴 🦂 📲		laim_Source2-Surve	PRE INTERANL TIME LAG	SURVEYOR LAG	POST INTERNAL LAG	× Properties Database
Arde Name Andyro Tde Cleris Class, Source J Ophin Clains Class, Source J Ophin Clains Class, Source J Tamacotin Class, Source Tamacotin Charling Source H Penkim - Mot Premium, Source H Penkim - Mot Premium, Source H Penkim - Mot Premium, Source H Penkim - Cath Premium, Source H Penkim - Cath Premium, Source H Penkim - Cath	1 2 3 4 5 6 7 7 8 9 10 11 11 12 11 14 15 6 17 18 19 20 21 22 23 24 5 27 28 29 30	5 163 122 37 17 18 10 10 10 10 10 10 10 10 10 10 10 10 87 87 87 91 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1 144 15 8 9 6 3 5 5 5 5 5 6 13 3 6 6 13 3 9 21 4 4 4 4 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	105 50 22 23 2 2 3 3 2 2 3 3 5 7 2 2 1 0 2 2 1 0 2 2 3 3 7 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2 2 2 3 5 4 2 5 1 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	e, accident_date))
	<					

This opens the Quick (Unindexed) Field Summarization dialog box.

Select the SURVEYOR_CODE as the Summarization Field.

Select the Gross Time Lag, Pre Internal Time Lag, Surveyor Lag and Post Internal Time Lag as Fields to Summarize.

Click OK

This step will result in a ready summary of the internal stage wise time lags in the claim settlement process.

This summary will be available Surveyor wise.

Application of CAATs to Utility Services

2 2 F & 3	Summarization				Properties
	Fields to summarize.		Numeric fields to total.		Data
	By: SURVEYOR_C	ODE 🛩	CLAIM_NO	OK	History
85	Then by: NONE	*	ESTIMATED_LOSS_AMT	Fields	📑 Field S
e Analysis-Tide Clients	Then by:		PRE_INTERANL_TIME_LAG	Cancel	😸 Caritra
Claim_Source1-Orphan Claims	Then by:		SURVEYOR_LAG		Criters
Claim_Source2-Surveyor Ageing Claim_Source-Transaction	Then by:	2		Help	II - Results
Earthquake Premium Masters	Then by:				
Motor Premium Master	Then by:	3			II ▼ Indices
Premium_Source1-Premium - Mc Premium_Source2-Non-system p	Then by:				-
Premium_Source3Canceled Po	Criteria:		Statistics to include		- Commen
Premium_Source4-Cover Note Premium_Source5-Motor Premiu	Use Quick Summarizati		Count Average		Add comm
Premium_Source-Premium - Eart	Create database		Sum Vasiance		
	Include % in output (database	Minimum	u demalium	
	Dan fields from first e				
	Use fields from last o		Create result		
	File name: Summarizat		Result name: Summarization		
	autiliarea.	100	The sol many. a singliar canon		
	27	0	0 0		-
	28	0	0 0	0	
	30	0	0 0	D	
	Search Results 🛛 🖓 Conclusio				
nning Tasks 🔮 To-Do List 💩 1, press Fl Start 💫 😂 CAATS Vook		Workin	g Folder: C:(Program Files)(DEA V 7.3)(Insu eneral_Insura	rance Research Number of Record	ds: 30 Disk Space:
h press F1 Start CAATS Book A - Summarization	😻 IDEA - Ge	Workin_Source2 🖬 13.1-6		rance Research Number of Record	
s, press F1 Start CAATS Dook A - Summarization Edit Vew Data Analysis	Sampling Tools Windo	Workin sim_Source2 🖬 13.1-6 w Help	eneral_insura		، ورونه -
A press FI Start CAATS Book A Summarization Edt New Data Analysis Caats Data	Samping Tools Windo	Workin am_Source2 🗊 13.1-5 w Help	enreljegra	1 2° = 11 (2)× (• ۲۵) - : • ∓ •
h, press Fi start CAATS book A Summarization Edit Wew Data Analyss Cat Wew Data Analyss Cat Wew Data Analyss Cat Wew Data Analyss	Samping Tools Winds	Workin sim_Source2 😭 10.1-5 w Help Son 😥 So 000 rveyor Agenn 🐨 Summari	nora) Insura	18° = 11 (2×1	C D L C
A press FI Start CAATS Book A Summarization Edt New Data Analysis Caats Data	Samping Tools Windo	Workin aim_Source2 1111-0 w 1Heb % 1100 00 000 weyer Agenr 1 Summari NO_OF_RECS GROSS	enreljegra	18° = 11 (2×1	Properties WE
h, press Fi start CAATS book A Summarization Edit Wew Data Analyss Cat Wew Data Analyss Cat Wew Data Analyss Cat Wew Data Analyss	Samping Tools Windo	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10			Properties VE
s prese FL start CALLS Look A Summarization Edit View Duta Analyss weight and the start of the start kerr P C CALLS CAL	Samping Tools Windo	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	entral_insura	NNL TIME LAG AVERAGE SUR	Properties V Data History
s prese FL Start CALIS Uses. A Summarization Eds: View Data Analyse California Californi California Californi California California California	Clam Source2-Sur Clam Source2-Sur Clam Source2-Sur Surveyor, CODe 12/5 1 2/5 2 3/5 2 4/5 3 5/5 4	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	eneral Joans	NNL TIME LAG AVERAGE SUR 0 5 20 9 9	Properties VE
La prese Fi La ration La Summarization La View Data Analysis derr Data Analysis d	Clam Source2-Sur Clam Source2-Sur Clam Source2-Sur Surveyor, CODe 12/5 1 2/5 2 3/5 2 4/5 3 5/5 4	Workin aim_Source2 1111-0 w 1Heb % 1100 00 000 weyer Agenr 1 Summari NO_OF_RECS GROSS	Reveal Joanna	NIL TIME LAG AVERAGE SUR	× Properties VE VE VE VE VE VE VE VE VE VE
Lard CARS Look A Summarization Edit View Data Analysis Care Data Carly Cook Care Data Carly Cook Care Data Care Care Care Cook Care Care Care Cook Care Care Care Cook Care Care Care Cook Care	Samping Tools Windo	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties VVE VT Database J Data Hatory Field 3 Contro Criters
s prese FI IEIT I I CANIS USSA. A Summarization Eds: View Data Analyss IEIT I IIII IIIIIIIIIIIIIIIIIIIIIIIIIII	Clam Source2-Sur Clam Source2-Sur Clam Source2-Sur Surveyor, CODe 12/5 1 2/5 2 3/5 2 4/5 3 5/5 4	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	XVE Vroperties Votabes Vo
Light Construction Light View Data Analysis Calify View Data Analysis Calify View Data Analysis Calify Construction Calify Constru	Clam Source2-Sur Clam Source2-Sur Clam Source2-Sur Surveyor, CODe 12/5 1 2/5 2 3/5 2 4/5 3 5/5 4	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties VVE VT Database J Data Hatory Field 3 Contro Criters
Lard Carls Look A Summarization Lat View Data Analysis Carl View Data Analysis Carl View Data Analysis Carl Carls Concerning Carls Carls Conce Carls Carls Carls Conce Carls	Clam Source250 Clam S	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Postables Posta Pred 3 Contro Critered Venue Venue Venue Venue
Light Construction Light View Data Analysis Construction Light View Data Analysis Construction Construction Analysis Construction Cons	Clam Source250 Clam S	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Properties Posta Prover Product Product Contro
Lart Construction Later Construction Later Very Data Analyses Construction Later Very Data Analyses Construction Construction Analysis Construction	Cam Source250 Cam Source250 Ca	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Postables Posta Pred 3 Contro Critered Venue Venue Venue Venue
Land Constant Land Constant Land View Data Analysis Care Data Analysis Care Data Analysis Care Data Analysis Care Data Analysis Care Data Care Constant Care Data Constant Care D	Cam Source250 Cam Source250 Ca	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Properties Posta Pata History Posta Control Contro Control Control Control Control
Land Constant Land Constant Land View Data Analysis Care Data Analysis Care Data Analysis Care Data Analysis Care Data Analysis Care Data Care Constant Care Data Constant Care D	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Properties Protocol Pred3 Criters Tendos
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for the Source of Source retyor Agent Source of Source No_CF_RECS (GROSS 111	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Yroperties You Data Hetrory Field Control Contro
Light Constraints and a second	Contraction of the second seco	Workin am_Source2 111-1-5 w Help for 111 100 100 100 100 100 100 100 100 10	Reveal Joanna	NL TIME LAG AVERAGE SUR 10 9 5 5 10 9 5	Properties Properties Properties Posta Pata History Posta Control Contro Control Control Control Control

The screen shows that Surveyor 'S2' has the highest absolute Surveyor Lag Average of 120 days for only 2 surveys.

CONCLUSION

Monitoring the Surveyor Lag can help in streamlining and disciplining the Surveyor's attempt to achieve efficient survey turnaround and output.

LEARNING

The same exercise can be carried out on the Pre-Internal and Post-internal Lag to isolate cases where claim officers have not issued intimation letters to surveyors and not acted upon survey reports within 7 days per claim case.

The @age function can help to identify delays, bottlenecks in specific functions, handled by specific individuals.

This can help in removing bottlenecks and improving claim throughput.

6.0 RE -VERIFICATION OF AGENCY COMMISSION ON BUSINESS SOLICITED

OBJECTIVE

To re-verify commission paid to agents on policy business solicited.

EXERCISE DESCRIPTION

In this exercise, we will run a data search on the agency commission transaction file to identify business solicited from Corporate Clients.

Then we will perform a data extraction to identify incorrect 'Individual' category Codes placed for corporate business.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Use the Search Function
- Extract data as per Logical Criterion
- Join Files

BACKGROUND

Agency commission varies based on the category of the insured, viz., corporate or individual.

DATA IMPORT

Import the file Premium Source-Commission. xls, Premium Source-Agency Commission. xls and Commission Master. dbf into IDEA as per the steps in IMPORT OF DATA FILES. (check?)

USE THE SEARCH FUNCTION

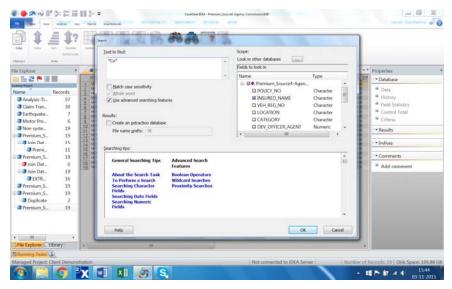
The file Premium Source-Agency Commission. IMD has a field CATEGORY. This field could either reflect Individual or Corporate, based on the status of the client.

The first test is to identify incorrect CATEGORY. That is, Individual category given to corporate clients.

Locate all transactions where business has been solicited from Corporate Clients.

Do the same by writing a search command in the 'Look for the following text box' in the Search dialog box.

The search command is * Co



The resultant file is a file containing business solicited from corporate clients, as is evident from the '& Co' notation in the Insured Name field.

anten Carrent 1	000 (12)	14		getjune 🚹 🔝 🖬 🥐				
ile Explorer			emium_Source4					Properties
	10		POLICY NO	INSURED NAME	VEH REG NO	LOCATION	CATE	· • Database
Analysis-Ti Claim Tran Earthquake Motor Pre Non syste Premium S	Records 57 30 7 6 19 19	1 2 3 4 5 6	MOTOR04011 MOTOR04012 MOTOR04014 MOTOR04016 MOTOR0401756 MOTOR04019	A & Co B & Co D & Co F & Co G & Co I & Co	Group Motor Group Motor	Mumbai Mumbai Mumbai Mumbai Mumbai Mumbai	Corpora Individu Corpora Corpora Corpora	Criteria Results
Premi Premi Join Dat	15 11 19 0	7 8 9	MOTOR04020 MOTOR04023 MOTOR04024	J & Co M & Co N & Co	Group Motor	Mumbai Mumbai Mumbai	Corpora Corpora Corpora	Indices No index Comments
 Join Dat EXTR Premium S Duplicate Premium S SE-Prem 	19 16 19 19 2 19 12	10 11 12	MOTOR04026 MOTOR04027 MOTOR04028	P & Co Q & Co R & Co	Group Motor	Mumbai Mumbai Mumbai	Corpora Corpora Corpora	Add comment
III The Explorer (Cit Search Results	brary	•						

EXTRACT DATA AS PER LOGICAL CRITERIA

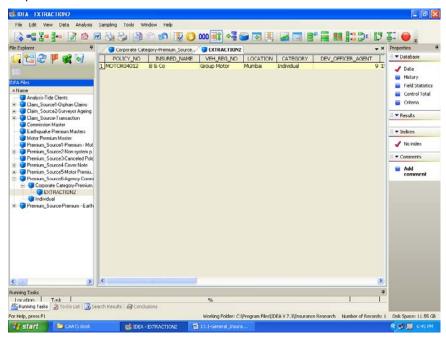
Extract data to a separate file through an equation

Category = 'Individual'

This is to identify incorrect Individual category given to Corporate Clients.

Sile Explorer 9	Premium Source6-Agency Com			Properties
Compared Control	POLICY NO Corract to File(s) MOTORNADIA MOTORNADI	Criteria	Create Fields Create Fields Create Fields 7 Dolder 7 Dolder 7 Cancel 11 11 11 11 11 11 11 11 11 1	Intervention Intervention Intervention Intervention Intervention Control Total Control Total Control Total Tensults Intervention Add comments Add comment
E Pernan, Souce Pernan - Eath Sach Reads Parning Taska (2) To to List Parning Taska (2) To to List Pernang Taska (2) To to List	Equation Category= "Individual" 2 Test Equation Record Number: 1 () (Evaluate)	Al Character Noneric Date / Time Marrie Conditional Financial	@Functions #functions are available for performing more complex operations such as did arithmski, stabilistic advalations and text searches. TERA functions begin with the W symbol. Then are very sindler in sight and operation to functions found in other software peskages such as Microsoft Excel, total 1 - 2-3, and diaxet. Each #function calculates a reack based upon the	

One case is shown where an Individual appears in the CATEGORY for corporate clients.



In this particular case, represented by Policy No MOTOR04012, 3% commission (Individual) has been paid to the agent in place of 5% (Corporate). (Source Commission Master. IMD)

JOIN FILES

Designate Premium_Source – Agency Commission. IMD as the Primary Database and Corporate Business-Premium Source. IMD as the Secondary Database.

Select the Policy_No field from both the Databases as the matching field.

Check the option Records with no secondary match to identify and isolate all business solicited other than corporate from the combined file of corporate and individual.

The resultant file should contain only INDIVIDUAL category transactions.

	No Inde		No control	total + 100%					
		POLICY NO	INSURED NAME	VEH REG NO	LOCATION	CATEGORY	DEV OFFICER	AGENT	RISI
Al C Anton Control Contro	pokcies.pm rhquake	1 MOTOR04015 2 MOTOR04013 3 MOTOR04021 4 MOTOR040221 5 MOTOR04025 6 MOTOR04029	Arjun Asif Liz Pert Zafar Ismail	Group Motor Group Motor Group Motor Group Motor Group Motor Group Motor	Mumbai Mumbai Mumbai Mumbai Mumbai	Individual Individual Corporate Corporate Individual Individual		7 11 11 999	02.06.0 18.07.0 17.00.0 19.00.0 10.08.0 15.10.0
Pennian, Soutes 2 Antel 49 Pennian, Soutes 2 Antel 49 Pennian, Soutes 4 Correct Note Dam, Soutes 1 Correct Mot Pennian, Soutes 4 Correct Note Pennian, Soutes 6 Pennian, Mot Pennian, Soutes 6 Pennian, Mot Pennian, Soutes 6 Pennian, Mot Pennian, Correct Ophan Claims Ordin, Soutes 2 Surrops Activity Pennian, Soutes 7 Correstore Activity Correstore Master	ies okcies tum Close P foto Pierrium Cl by mmission um Source								

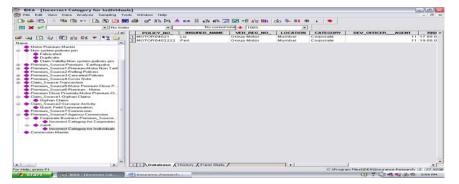
EXTRACT DATA AS PER LOGICAL CRITERIA

Extract data to a separate file through an equation

Category = 'Corporate'

This is to identify incorrect Corporate category given to individual clients.

The extraction gives 2 resultant exceptions to the logical expression.



In the specific cases represented by Policy No MOTOR04021 and MOTOR0402222, 5% commission (Corporate) has been paid to the agent in place of 3% (Individual).

CONCLUSION

The Category field determines the agency commission.

The Category field can be altered to reflect a category not in line with the insured status as can be seen above.

Hence the tests for agency commission have to be run on the Insured field to gather the true status of the insured.

7.0 IDENTIFICATION OF MOTOR INSURANCE BUSINESS CLAIMS SETTLED ON POLICIES PERTAINING TO STOLEN VEHICLES

OBJECTIVE

To identify if claims have been settled on policies, insuring stolen vehicles as per the National Crime Records Bureau (NCRB) alert file of stolen vehicles

EXERCISE DESCRIPTION

In this exercise, we will join the claim transaction file with the NCRB list of stolen vehicles on the basis of the vehicle engine number.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

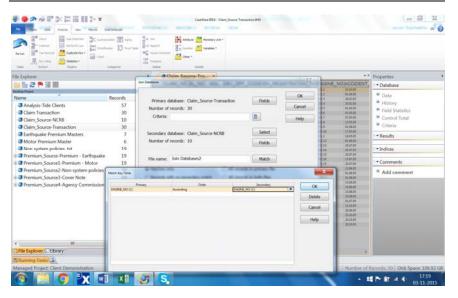
• Join Files

DATA IMPORT

Import the file Stole Vehicles. prn into IDEA as per the steps demonstrated in the session on IMPORT OF DATA FILES.

JOIN FILES

The file Claim Source-Transaction. IMD (Primary File) is joined to the file Claim Source – NCRB. IMD (Secondary File) through File > Join Databases. When a vehicle is stolen the vehicle registration number undergoes a change but the Engine_no does not change, hence the matching key is Engine_No. The process is seen below.



After joining the two files we get a list of 12 policies, pertaining to 12 stolen vehicles against which claims have been admitted, processed and settled. Details about these can be seen below. The value of such claims (Paid Loss Amount) from the control totals amounts to Rs. 87,29,812.50

	neroster 🔝 Apro shaser 🚺 Prostado s	Corporation	Andure Monetary Unit *			
e Explorer a	Claim Sour	ce-Trans_ 🗡 🖲	Join Databases2.1 *			Properties
1 2 m 3 m					NOACCIDENT_DATEESTIMATED_LOS	
Non Proved	1 500700160902014	1 Partial Love	0000 Advitt 15 0000 Advitt 15	68115	25.07.05 25.07.05	
ame	3 101701160502017	1 Total Loss	999 MH-11118	6 01.15	23.04.05	✓ Data
Analysis-Tide Clients	4 101701380902018	1 Total Loss	9999 145112 29	6 (0.19	IL.06.05	History
	5 101701340902024	1 Tetalizes	9 MH10L25	6 (0. 75	61.07.09	Field Statistics
Claim Transaction	6 301702360502525	1 Totalizes	999 144101-25	£ 01.26	15.35.05	
Claim_Source-NCRB	7 100700160502627 B 100700160502629	1 TotalLess 1 TotalLess	9 58401.28	E 01.20 E 01.30	20.00.00	Control Total
Claim Source-Transaction	g 100701160902905	1 FatialLoss	9 MPHELDU	501.6	20.2040 EL.05.05	Criteria
	10 100701100002523	1 Fatializa	990 181216	611.6	20.09.05	
Join Databases2	11 101701140502006	1 TotalLoss	999 MR+017	6 41 7	87.85.65	Results
Earthquake Premium Masters	12 30170160902007	1 TotalLoss	399 149102.8	6.01.8	15/05/05	
Motor Premium Master						I = Indices
Non system policies, txt						
Premium Source-Premium - Earthqu						No index
Premium_Source1-Premium - Motor						· Comments
Premium_Source2-Non-system polic						Add comment
Premium_Source3-Cover Note						
Premium_Source4-Agency Commissi						
RI		111			,	_
Running Tasks						441
				Not connected		Records: 12 Disk Space: 106

CONCLUSION

The development officer and agents should be provided with the current and dynamic list of stolen vehicles every fortnight from the NCRB.

Preventive checks should be run by way of Vehicle Engine No. to ensure that no stolen vehicles are being insured by the branch.

13.1.8 Business Analysis Case Studies

1.0 ANALYSIS BY CLIENT FOR DIFFERENT CLASSES OF RISKS INSURED

OBJECTIVE

To analyze the premiums collected, claims paid and losses sustained per insured for different classes of risk insured.

EXERCISE DESCRIPTION

In this exercise, we will create a Pivot Table for two variables: Class of Risk and Customer ID.

These variables will be analyzed for the main heads premium, claims and losses incurred.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

• Create a Pivot Table

DATA IMPORT

Import the file Analysis Business. xls into IDEA

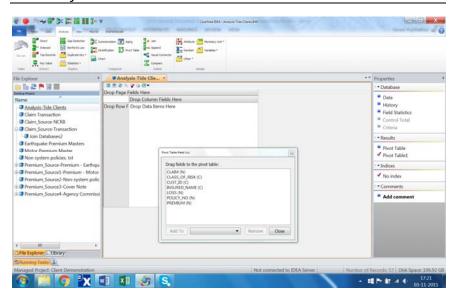
CREATE A PIVOT TABLE

In this exercise, capture two key variables Cust_ID and Class_of_Risk.

Proceed on the basis of the Insured_Name because there may not be consistency in the naming convention for the Insured_Name across Class_of_Risk.

The variables need to be analyzed across numeric fields Premium, Claim and Loss. Normally multi-key analysis is done through the Pivot Table feature in IDEA 9.

The Pivot Table is invoked through the Analysis option on the menu tool-bar. Once invoked, it appears like what is seen below



The Cust_ID from the Pivot Table Field List is picked and dropped in the Row section. The Class_of_Risk from the Pivot Table Field List is picked and dropped in the Column section. Finally the numeric fields – Claim, Premium and Loss are picked one by one and dropped in the Data Grid section of the Pivot Table. The result is seen below as a tab in the file Analysis-Tide

	reservation 🔝 A natification 🔝 m net Cologram		1	Attribute 🚰 Hanadary C Random 🍰 Vanability * Other *						
le Explorer #	Anah	sis-Tide Clie *	-					100 C	- Pr	óperties
1 2 P 3 II	IT MON	¥ 3 8*	_							Database
extup Provid	Drop Page I	Fields Here								
Name				CLASS_OF_HIS						Data
Analysis-Tide Clients	CUST_ID		•		Fire	Marine	Miscellaneous	Motor Own Damag Motor Th		History
Claim Transaction		Sum: PREMIUM		0.000						Field Statistics
Claim_Source-NCRB	-	Sum: CLAIM		0						Control Total
Claim_Source-Transaction	A1	Sum: PREMIUM			12,500.000		12,500.000	12,500.000	1	Criteria
Join Databases2		Sum: CLAIM			35,700		35,700	35,700		Results
Earthquake Premium Masters	81	Sum: PREMIUM			1,625.000		1,625.000	1,625.000	- 11	
Motor Premium Master	-	Sum: CLAIM			0		0	0		Pivot Table
Non system policies. txt	a	Sum: PREMIUM	_		23,750.000		23,750.000	23,750.000		Pivot Table1
🛙 🗊 Premium_Source-Premium - Earthqu	-	Sum: CLAIM	_		0		0	0		Indices
Premium Source1-Premium - Motor	D1	Sum: PREMIUM	_		1,125.000		1,125.000	1,125.000		
Premium_Source2-Non-system polic		Sum: CLAIM	_		0		0	0		No index
Premium Source3-Cover Note	E1.	Sum: PREMIUM Sum: CLAIM	_		13,000,000		13,000.000	13,000.000	1.	Comments
Premium Source4-Agency Commissi	FI	Sum: CLAIM Sum: PREMIUM	_		3,500		3,500	3,500		Add comment
	12	Sum: PREMIUM Sum: CLAIM	_		2,000.000		2,000.000	2,000,000		Add comment
	G1	Sum: PREMIUM			15,780	28.950.000	28.950.000	28.950.000		
	01	Sum: PREMIUM Sum: CLAIM				28,950,000	28,950.000	28,950,000		
	H1	Sum: PREMIUM	-			1.300.000	1.300.000	1.300.000		
	or .	Sum: CLAIM				1,500,000		1,500,000		
	n	Sum: PREMILIM	-			1.500.000	1,500,000	1.500.000		
< III +	÷	Sum: CLAIM				225.000		225.000		
File Explorer (DUbrary)		President and		m			110,000			
and the second sec	-						_		_199	
Running Tasks										ordic 57 Disk Space: 106

By placing the cursor of the mouse over the Row Header 'Data' and right clicking it, it invokes the drop down list of Numeric fields, which can be

checked or unchecked, as is seen below, to change the view. We uncheck Premium and Claim to just reflect the Loss position.

	eneration 🔝) adhudan 🔝) art Coloption	host Table	Arabure 🚵 Horatory (E Garden 🏄 Varable * Other * Sangle						
ile Explorer *		ysis-Tide Clie ×						Propert	les
■ L: 2* 件 図 Ⅲ		¥ 3 H+						· · Data	abase
nation Project	Drop Page	Fields Here	and the second s					B Dat	
Name			CLASS_OF_HIS		()	10 M	G 1	His	
Analysis-Tide Clients	CUST_ID.			Fire	Marine	Miscellaneous M	lotor Own Damag Motor Th		
Claim Transaction		Show All)	0.000						ld Statistics
Claim_Source-NCRB		Sum: PREMOUM	0						ntrol Total
Claim_Source-Transaction	A1	Sum: CLAIM	_	12,500.000		12,500.000	12,500.000	* Crit	teria
Join Databases2		Sum: CLAIM		35,700		35,700	35,700	· Rest	drs.
Earthquake Premium Masters	81	Sum: PREMIUM		1,625.000		1,625.000	1,625.000		
Motor Premium Master		Sum: CLAIM		0		0	0		ot Table
Non system policies. txt	a	Sum: PREMIUM		23,750.000		23,750.000	23,750.000	- Piv	ot Table1
Premium Source-Premium - Earthqu	-	Sum: CLAIM	_	0		0	0	• Inde	ces.
Premium Source1-Premium - Motor	D1	Sum: PREMIUM	_	1,125.000		1,125.000	1,125.000		
Premium Source2-Non-system polic		Sum: CLAIM		0		0	0	✓ No	index
Premium Source3-Cover Note	E1.	Sum: PREMIUM	-	13.000.000		13,000.000	13,000.000	· · Corr	oments
Premium Source4-Agency Commissi	-	Sum: CLAIM		3,500		3,500	3.500		
and the second second second	F1	Sum: PREMIUM		2,000.000		2,000.000	2,000.000	= Ad	ld comment
	-	Sum: CLAIM		15,780		15,780	15,780		
	G1	Sum: PREMIUM			28,950.000	28,950.000	28,950.000		
		Sum: CLAIM			25,000	25,000	25,000		
	H1	Sum: PREMIUM			1,300.000	1,300.000	1,300.000		
		Sum: CLAIM			0	0	0		
	n	Sum: PREMIUM	-		1,500.000	1,500.000	1,500.000		
• • • •		Sum: CLAIM			225,000	225,000	225,000		
File Explorer Ubrary	•		ш						
Running Tasks									

By placing the cursor of the mouse over the actual Data Grid and right clicking on any data cell, it invokes the drop down list of Pivot Table field list containing significant arithmetical functions like sum, min, max, count and average. A single function can be selected, as is seen below, to change the view. We select Average to reflect the average loss sustained per Cust_ID per Class_of_Risk.

	entration III Au athunan III Au at Cologetae	est Table	M Attitude 🔐 Handa E Fander 🎒 Ganda Otar *						
ile Explorer *		vsis-Tide Clie ×							Properties
■ E 2 [*] 2 [*] 2 [*]	RM C 4								• Database
unter Propert	Drop Page 8	Fields Here	and the second s					-	Data Data
Vame	_		CLASS_OF_HIS		1	100	A		History
Analysis-Tide Clients	CUST_ID			Fire	Marine	Miscellaneou	is Motor Own Damag Motor	Th	 History Field Statistics
Claim Transaction		Sum: PREMIUM	0.00						
Claim_Source-NCRB	-	Sum: CLAIM		0					Control Total
Claim_Source-Transaction	Al	Sum: PREMIUM		12,500.000		12,500.00		12	Criteria
Join Databases2		Sum: CLAIM		35,700		35,70			· Results
Earthquake Premium Masters	81	Sum: PREMIUM	-	1,625.000	<u>1</u>	1,625.00			
Motor Premium Master		Sum: CLAIM	PustTable Field			×			Pivot Table
Non system policies, txt	a	Sum: PREMIUM	-		-	0			✓ Pivot Table1
Premium Source-Premium - Earthqu		Sum: CLAIM	Source field:	PREMIUM	C	OK	0 0		• Indices
Premium Source1-Premium - Motor	D1	Sum: PREMIUM	Name: Ave	age: PREMIUM	10	0	0 1,125.000		
Premium Source2-Non-system polic		Sum: CLAIM			100	Cancel	0 0		✓ No index
Premium Source3-Cover Note	E1.	Sum: PREMIUM	Summarize by	¢		0	0 13,000.000		· • Comments
		Sum: CLAIM	Sum			0	0 3,500		Comments
Premium_Source4-Agency Commissi	F1	Sum: PREMIUM	Count			10	0 2,000.000		Add comment
		Sum: CLAIM	Min			18	0 15,780		
	G1	Sum: PREMIUM	Max			10	0 28,950.000		
		Sum: CLAIM				10	0 25,000		
	H1	Sum: PREMIUM				10	0 1,300.000		
		Sum: CLAIM	Show data as:				0 0		
	л	Sum: PREMIUM					0 1,500.000		
· · · · ·	100	Sum: CLAIM	Normal			• 6	0 225,000		
File Explorer Ubrary			C		_				
Running Tasks								-	-

In the Fire risk category, Cust_ID Q1 seems to have caused the highest damage to the branch of Rs. 73,125.

The branch has taken a significant hit on Cust_ID K1 for three categories of risks Marine, Miscellaneous and Owned Motor of Rs. 347625, 347625 and 347625 respectively. The Owned Motor risk category seems exceptionally high in comparison with the other Cust_ID's.

Terrer States St	etuare D Post	Table Append State Open Company Append App	80.	nten 🚮 i	Aperatory Lott * Landbles *				
Explorer *		s-Tide Clie *	1					- 1	Properties
	RHONN								· • Database
top Project	Drop Page Fie	ids Here							Data
ime	SUM: CLAUM			Marine	Microffenerus	Motor Own Damag	Motor Third Dot	Total	History
Analysis-Tide Clients Claim Transaction	ALCONTRACTOR OF	0	cie.	wante	moreianeous	wotor own barnag	MOTOR THE PART	0	Field Statistics
Claim Transaction	Al	, in the second s	35,700		35,700	35,700		107.100	Control Total
Claim Source-Transaction	81		0		0	0		0	Criteria
Join Databases2	C1		0		0	0		0	· Results
Earthquake Premium Masters	D1		0		0	0		0	* Results
Motor Premium Master	E1		3,500		3,500	3,500		10,500	Pivot Table
Non system policies, bd	F1		15,780		15,780	15,780		47,340	Pivot Table1
Premium Source-Premium - Earthqu	G1			25,000	25,000	25,000		75,000	• Indices
Premium Source1-Premium - Motor	H1.			0	0	0		0	
Premium Source2-Non-system polic	11			225,000	225,000	225,000		675,000	No index
Premium Source3-Cover Note	K1		-	350,000	350,000	350,000		1.050.000	· Comments
Premium Source4-Agency Commissi	L1 M1		0	0			0	0	Add comment
	N1 N1		65.000	65.000			65.000	195.000	- Add comment
	01		15.000	15.000			15.000	45.000	
	P1		50.000	50.000			50.000	150.000	
	Q1		75,000				75.000	150.000	
	R1		0				0	0	
	51		0				0	0	
	T1		0				0	0	
File Explorer Ubrary	Total	0	259,980	730,000	654.980	654.980	205.000	2,504.940	
Running Tasks									ee ee

Upon re-viewing the Pivot Table for summary of Premiums across Cust_ID and Class_of_Risk, we get the position seen below.

	untratur 🚺 Ag		H. test	Ar 🚰 Hanalary Lin •• ざ Vanakins • •				
Take Fundation Contents of Fundation of Fund	Categorie	Tide Clie *	-	Service.	_			 Properties
	THE							 Database
	Drop Page 8							
lame	Sum: PREM							Data
Analysis-Tide Clients	0.5100	Fire	Marine	Miscellaneou	Motor Own D	Motor Third Party	Total	History
Claim Transaction							0.000	Field Statistics
Claim Source-NCRB	A1	12,500.000		12,500.000	12,500.000		37,500.000	Control Total
Claim Source-Transaction	81	1,625.000		1.625.000	1,625.000		4,875.000	Criteria
Join Databases2	CI	23,750.000		23,750.000	23,750.000		71,250.000	· Results
Earthquake Premium Masters	01	1,125.000		1.125.000	1,125.000		3,375.000	* Results
Motor Premium Master	E1	13,000.000		13,000.000	13,000.000		39,000.000	Pivot Table
Non system policies, bd	F1	2,000.000		2,000.000	2,000.000		6,000.000	Pivot Table1
Premium Source-Premium - Earthqu	G1		28,950.000		28,950.000		86,850.000	 • Indices
Premium Source1-Premium - Motor	H1		1,300.000	1,300.000	1,300.000		3,900.000	
Premium Source2-Non-system polic	/1		1,500.000	1,500.000	1,500.000		4,500.000	No index
Premium_Source3-Cover Note	1(1		2,375.000	2,375.000	2,375.000		7,125.000	· • Comments
Premium Source4-Agency Commissi	11	30,000,000				30,000.000	90,000.000	
riting and a signed with the comments	M1		41,055.000			41.055.000	123,165.000	Add comment
	N1					21,600.000	64,800.000	
	01	612.500	612.500			612.500	1,837.500	
	P1	1,487.500	1,487.500			1,487.500	4,462.500	
	Q1	1,875.000				1,875.000	3,750.000	
	R1	12,796.875 787.500				12,796.875	25,593.750	
· · · ·	51 T1	787.500				787.500	1,575,000 38,250,000	
		19,125,000	-				38,250,000	
File Explorer Ubrary					1	0		
Running Tasks								

M1, N1, L1, G1, and C1 seem to be the Tide Clients drawing in cumulative premium business in excess of Rs. 15000 per annum per class_of_risk

Of the above Tide Clients a careful examination reveals that N1, seems to be causing heavy losses to the branch despite contributing handsome premiums to the Tide Client kitty.

CONCLUSION

The branch should inform this case to the Region and take a concerted decision to raise the risk rating on N1, thereby hiking up the premium rates substantially on future business accepted from N1.

A careful examination of J1 and K1 reveals that these customers are contributing insignificant premium to the branch (average of Rs. 1500 and 2375 respectively per class), but are invoking heavy back breaking losses on the branch (average of Rs. 223500, and 347625 per class).

This dis-proportionate premium to loss result is placing an enormous burden on the branch.

2.0 ANALYSIS BY DEVELOPMENT OFFICER AND / OR AGENT OF INSURED LOSSES

OBJECTIVE

To analyze the losses sustained in motor class risk for policies underwritten by different development officers and / or agents.

EXERCISE DESCRIPTION

In this exercise, we will join the agent-wise incurred loss file with the file for tide clients to capture the policy number, agent code and incurred loss together.

This file will then be analyzed through a Pivot Table for the business soliciting patterns of agents that burden the branch with heavy claims and losses.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

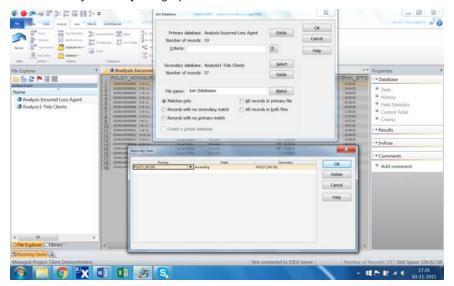
- Join Files
- Create a Pivot Table

DATA IMPORT

Import the file Analysis Incurred Loss Agent. xls into IDEA as in IMPORT OF DATA FILES.(Check)

JOIN FILES

Join the files Analysis Incurred Loss Agent. IMD (primary file) with the file Analysis Business. IMD (secondary file) on the key field POLICY NO, with 'Matches only' as the joining option.



CREATE A PIVOT TABLE

Run a Pivot Table on the joined file generated.

The Development Officer / Agent Code is captured in the rows section and the numeric fields premium, claim and loss incurred are captured in the data grid.

Application of CAATs to Utility Services

ile Explorer	* Analysis-Incurred L. 🗡 🖲 Join Datab	ases.IMD ×		• * Properties
E 2 * G III	周囲の人 19 2 日・			• Database
KATLE Project	Drop Page Fields Here	The second second second second second		P Data
Name		Drop Column Fields Her		History
Analysis-Incurred Loss Agent	DEV_OFFI Data	•		 Field Statistics
Join Databases	1 Sum: PREMIUM	37,375,000		Control Total
Analysis1-Tide Clients	Sum: LOSS 7 Sum: PREMILIM	45,250,000		- Control Total
	Sum: LOSS	45,250,000	Prost Table Field List	E3
	9 Sum: PREMIUM	1.625.000		
	Sum LOSS	0.0	Drag fields to the pivot table:	
	11 Sum: PREMIUM	2.375.000	CLAIM (N) CLASS_OF_RISK (C)	A
	Sum LOSS	-347.625.0	CUST_ID (C)	
	99 Sum: PREMIUM	1 500.000	DESCRIPTION (C) DEV_OFFICER_AGENT (N)	
	Sum: LOSS	-223,500.0	DISCOUNTS (N)	
	Total Sum: PREMIUM	88,125,000	GROSS_PREMIUM (N)	
	Total Sum: LOSS	-594.655.0	INSURED_NAME (C) INSURED_NAME1 (C)	
			LOCATION (C)	
			1055 (N)	Contraction of the local division of the loc
			Add To .	Remove Close

From the above it can be seen that Dev Officer 99 and Dev Officer 11 have brought in policies that have saddled the branch with enormous losses.

Dev Officer 99 has brought in a motor policy of Rs. 1500 premium for which the loss incurred in the year is Rs. 223500.

Dev Officer 11 has brought in a motor policy of Rs. 2375 premium for which the loss incurred in the year is Rs. 347625.

CONCLUSION

These cases should be specifically reported and taken up for close scrutiny.

The past trend for these policies should also be viewed carefully to examine history of repetitive losses. It is quite possible that the current losses were one-time abrupt losses due to bad luck.

Based on the findings, the development officer should be informed and asked to defend his decision to solicit policies experiencing repetitive claims.

In dire circumstances the development officer may have to be asked to step down from his post because his activities are blatantly hard hitting on the finances of the branch.

3.0 IDENTIFICATION OF PREMIUM ENHANCEMENT CASES FOR LOSS MAKING AND HIGH-RISK POLICIES

OBJECTIVE

To identify trends in policies experiencing repetitive losses for over 3 years.

The evaluation of such cases for premium enhancement.

EXERCISE DESCRIPTION

In this exercise, we will learn to use Visual Connect to join more than two files at the same time.

The files will be joined based on the key 'Insured Name'.

Trends in incurred losses will be identified and acted upon.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to

Use Visual Connect

DATA IMPORT

Import the file Analysis Incurred Loss 2005. xls, Analysis Incurred Loss 2004. xls and Analysis Incurred Loss 2003. xls into IDEA

USE VISUAL CONNECT

Use the Visual Connect feature in IDEA 9 to join the files on the basis of the Key 'Insured_Name'.

The files reflect renewal policies from 2003 to 2005, so it can be taken as given that the Insured_Name field will stay the same over the three years.

Invoke the Visual Connect feature from Analysis > Visual Connect.

Open the three files consecutively in the Visual Connect dialog box.

Join the three files together based on the Insured_Name field by linking the output node from the Analysis-Incurred Loss 03 file to the input node from the Analysis-Incurred Loss 04 file, and the output node from the Analysis-Incurred Loss 04 file to the input node from the Analysis-Incurred Loss 05 file.

Application of CAATs to Utility Services

12 C		Analysis2-Incurred Loss 04 🛛 🗮 🗶	
ana mont Jame ⁶ Recc ² Analysis-In ² Analysis-In ³ Analysis-I ³ Analysis-I ³ Analysis-I	Analysis2-Incurred Loss 03 (1) POLICY, MO (16) NOURED, MME (C) PREMUM (10) CLAM (10) LOSS (10)	- POLICY.NO 90 NORRED.WARE (C) PREMIM 80 CUMM 80 LUBB 80	
		Analysis2-incurred Loss 05 III X	

After Visual Connect the joined file appears as below -

1 🗙 🛩 [No Index		-	No contre	ol total 🔶 🚺	00%	-			
	21 M		INSURED_NAME	PREMIUM	LOSS	PREMIUM1	LOSS1	PREMIUM2	LOSS2	
ି 🕰 😭 ମିଆ ଛାଇ କାଳ	10 12 🛄		A & CO	12,500.000	-2,500.0	12,500.000	-13,200.0	12,500.000	-23,200.0	
me	Flecords		9 & Co C & Co	1,625.000	0.00	1,625.000	0.00	1,625.000 23,750.000	0.00	
Analysis-Tide Clients	57		0800							
Analysis1 Incurred Loss Ag	19		E & Co	1,125.000	0.00	1,125.000	0.00	1,125.000	0.00	
🗆 🔷 Join	10		E&Co F&Co	13,000.000	0.00	13,000.000	0.00	13,000.000	9,500.0	
Analysis2-Incurred Loss 03	19			2,000.000	-10,500.0	2,000.000	-10,500.0	2,000.000		
Analysis2-Incured Loss 04	19		0 & Co	28,950.000	-1,050.0	28,950.000	8,950.0	28,950.000	3,950.0	
Analysis2-Incurred Loss 05	19		H&Co	1,300.000	0.00	1,300.000	0.00	1,300.000	0.00	
VisConn	19		& Co	1,500.000	-4,000.0	1,500.000	-4,000.0	1,500.000	-223,500.0	
			1 & C 0	2,375.000	-3,625.0	2,375.000	-3,625.0	2,375.000	-347,625.0	
			L & Co	30,000.000	0.00	30,000.000	0.00	30,000.000	0.00	
			M & Co	41,055.000	0.00	41,055.000	0.00	41,055.000	0.00	
			N&Co	21,600.000	2,600.0	21,600.000	-24,100.0	21,600.000	-43,400.0	
		14	0 & Co	612.500	0.00	612.500	0.00	612.500	-14,307.5	
			P & C 0	1,487.500	-6,012.5	1,487.500	-8,612.6	1,487.500	-48,612.6	
			Q & Co	1,875.000	-6,125.0	1,875.000	-6,125.0	1,875.000	-73,125.0	
			R & Co	12,796.875	0.00	12,796.875	0.00	12,796.875	0.00	
		18	8 & Co	787.500	-1,712.5	787.500	0.00	787.500	0.00	
		19	T & Co	19,125.000	0.00	19,125.000	0.00	19,125.000	0.00	

By re-formatting the above file to position the Loss fields and the premium fields together, we get a clear linear view of the regular loss making cases, as can be seen below

Image: Control of the Contro				-		ol total 🔶 100%	No contri	*	No Index	III 🗙 🛩 [
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-23,200.0 0.00 0.00 9,500.0 -13,780.0 0.00 -23,500.0 -223,500.0 -223,500.0 -247,625.0 0.00 -347,625.0 0.00 -43,460.0 -14,387,5 +46,512,5 -46,512,5 0.00 0.00	-13,200.0 0.00 0.00 -10,500.0 0.00 -10,500.0 0.00 -3,625.0 0.00 -2,625.0 0.00 -2,6125.0 0.00 0.00 -24,100.0 0.00 0.00 0.00 0.00	-2,500.0 0.00 0.00 0.00 -10,500.0 -10,500.0 -1,050.0 0.00 -3,625.0 0.00 2,600.0 0.00 2,600.0 0.00 -6,012.5 -6,125.0 0.00 -1,712.5	12,500,000 1,622,000 23,750,000 11,125,000 13,000,000 20,950,000 1,300,000 20,950,000 1,500,000 2,375,000 21,600,000 612,500 1,875,000 1,875,000 1,875,000 1,875,000 1,875,000 1,875,000 1,875,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,000 1,750,00	12,500,000 1,625,000 23,750,000 1,125,000 13,000,000 2,000,000 1,500,000 1,500,000 2,375,000 30,000,000 1,500,0	12,500,000 1,625,000 1,125,000 1,125,000 1,125,000 2,000,000 2,0,950,000 1,500,000 2,175,000 2,175,000 2,1,605,000 1,487,500 1,487,500 1,475,500	A & CO B & CO C & CO C & CO C & CO F & A CO F & A CO H & CO J & A CO	** *	Analysis Tide Clients Analysis Tide Clients Join Analysis Tincured Loss 03 Analysis Zincured Loss 03 Analysis Zincured Loss 05

Take the first case, for example, of Insured A & Co. There has been a steady increase in the Losses Incurred from Rs. 2500 in 2003 to Rs. 13,200 in 2004 and Rs. 23,200 in 2005.

Despite this steady increase in losses due to higher and repetitive claims over the years, the premium has not been appropriately enhanced to reflect the high-risk profile of the Insured.

The same premium of Rs. 12,500 has been maintained from 2003 to 2005.

The same is true of N & Co, P & Co and Q & Co. These cases falling under the non-tariff segment of motor insurance can be treated appropriately as sensitive risk profiles and the premiums on such cases can be suitably hiked up to justify the high risk nature of the Insured.

CONCLUSION

By making a running comparison of Insured Losses for 3-5 years, the audit department can arrive at a well-based intelligence on specific policies, loss assets and claim patterns.

Given this intelligence, the audit department can strongly recommend such cases for premium enhancement in the best interest of the branch (thereby delivery its role as a value added function this part is not clear).

13.1.9 Finance & Accounts Case Studies

1.0 RECONCILIATION OF SUB LEDGER AND GENERAL LEDGER FOR PRINCIPAL CONTROL ACCOUNTS

OBJECTIVE

To reconcile the Cash Deposit Control Account in the General ledger with the Cash Deposit Sub Ledger to establish the accuracy of transactions recorded and posted.

EXERCISE DESCRIPTION

In this exercise, we will learn to add a computed virtual numeric field in the sub-ledger file to arrive at the month for each transaction undertaken.

This computed field giving monthly results would be based on the transaction date field.

Then we will summarize the sub-ledger on the basis of the month field to arrive at gross transaction details-debit and credit per month.

This summarization will finally be compared with the General Ledger monthly summary.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Tile Files
- Append a Field
- Use Equation Editor
- Write @ Functions
- Perform a Quick Field Summarization

DATA IMPORT

Import the file Analysis Sub Ledger CD. Dbf and General Ledger CD. dbf into IDEA as in IMPORT OF DATA FILES.(check?)

TILE FILES

Import the files as specified.

Navigate to window and select TILE from the windows menu

Tiling allows the files to be viewed concurrently on the screen.

le Explorer 9		Sub Ledger	0								- ×	Properties
1 🔁 🏖 🟴 📽 🗟	K	DELETED	CUST I	DIDATE	1 proc	RIPTIO	DEBIT	CREDIT	BALANCE	-		II - Database
A Has Xine Analysis1-Incoured Lose 03 Analysis1-Incoured Lose 03 Analysis1-Incoured Lose 04 Analysis1-Incoured Lose 05 Analysis1-Incoured Lose 30 Analysis1-Incoured Lose 30 Analysis1-Incoured Lose 30 Clam, Source Oxphan Claims Clam, Source Oxphan Claims Clam, Source Transaction Commission Martle Enthraalse Pennam Marter	123456789111211415167		100 100 100 100 100 100 100 100 100 100	00 01.04.0 00 07.04.0 00 03.04.0 00 30.04.0 00 31.05.0 00 31.05.0 00 31.07.0 01 30.04.0 01 30.04.0 01 30.04.0 01 23.05.0 01 23.05.0 01 25.05.0 01 25.05.0 01 31.05.0 01 31.05.0 01 31.05.0 01 31.05.0 01 31.05.0 01 30.06.0 01 30.06.0 01 30.05.0 01 30.05.0 00 30.0 00 30.05.0 00 30.05.0 00 30.05.0 00 30.05.	Balance Out Balance Out Balance Out Balance Out Balance Out Balance In Balance In Balance In Funds In Cho Voyage Mani Voyage Mani Salance Out Voyage Mani Salance Out	t BL 24 No. 234650 (No. 112709 fest 15 fest 16 fest 17	0 43000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 100000 250000 0 0 0 0 500000 0 0 0 0 0 0 0 0 0 0 0	35000 135000 90000 340000 340000 340000 340000 85000 565000 565000 565000 565000 560000 400000 400000 415000			✓ Data History Field Stability Control Total Criteria II ← Results II ← Indices ✓ No index II ← Comments
General Ledger CD Motor Premium Master	-	🍞 General L		JI 30.06.0	5 Balance Out		U	U	415000		• X	Add comment
Premium, Soucel Prenium, Mot Premium, SouceS Cancel Neroydem Premium, SouceS Canceld Pale Premium, SouceS Canceld Pale Premium, SouceS Agency Come Premium, SouceS Agency Come Premium, SouceS Agency Come Premium, SoucePremium - Earth Sob Ledger CD Vision	1204		CD 1	31.05.05	DESCRIPTIO ppl day une uly	OPENING_BA 19400 40900 87900 79150	00 450 00 1100 00 875	00 35000 00 50000	0	K3_BA 409000 879000 791500 776500		
Running Tasks 🚽 To-Do List 🐊 S	sarch (Results 🚑 Ce	inclusions		Working Fo	lder: C:lProgram P	inside v	7.3iInsuance	Research 1	umber of Reco	ords: 4	Disk Space: 11.55

We are required to reconcile the two files on the monthly totals.

APPEND A VIRTUAL FIELD, USE THE EQUATION EDITOR AND WRITE AN $\ensuremath{ @ \ }$ Fucntion

Add a virtual numeric field called 'MONTH' to the file Sub Ledger CD. IMD.

In this field write the equation @month (date) to convert all the transaction dates into month details.

This is with the objective of summarization of balances, debit and credit fields month- wise.

Application of CAATs to Utility Services

xplorer		P Suble	A second s	a constant and	and the second second	the start of the start				Properties
	Field	d Manipulation	doer (h							atabase
		and the second second		1	12				_	Data
	-	Field Name CUST_ID	Type	Len	Dec	Parameter	Tag Name	Description	OK	History
Files	-	DATE	Date*	0	0	dd.mm.yy	<no tag=""></no>		Append	Field Statistics
ame		DESCRIPTIO	Character	20	-	domininyy	<no tag=""></no>		Delete	Control Total
Analysis1-Incurred L Analysis1-Incurred L	4 5	DEBIT	Numeric	8	0		<no tag=""></no>		Denota	Criteria
Analysis1-Incured L	-	CREDIT	Numeric	8	0		<no tag=""></no>		Print	
Analysis1 Incurred I		BALANCE		0	0		and the second se		Copy	esuks
Analysis1-Tide Clier		MONTH	Numeric Virtual numeric	0	0		<no tag=""></no>			
Analysis Tide Client	0	MONTH	ALTON LYDING		U		KIND (MAS		Cancel	ndices
Claim_Source1-Orp	•	(M	Help	No index
Claim_Source Trans										
Commission Master Earthquake Premium I General Ledger CD Motor Premium Master Premium_Source1-Pre	etn s	Equation Editor				e Manifest 19	65000 aA @	0 415000		Add comments
Commission Master Earthquake Premium I General Ledger CD Motor Premium Master Premium_Source1-Pte Premium_Source2-No Premium_Source2-No	ti etn xr:-	Equation Editor				F AI				add a
Commission Master Earthquake Premium I General Ledger CD Motor Premium Master Premium_Source1-Pre Premium_Source2-Nor		Equation Editor				Al B Character B Numeric		@Functions		bbA 😈
Commission Master Canthquake Premium 1 General Ledger CD Motor Premium Master Premium_Source3-Pre Premium_Source3-Pre Premium_Source3-Co Premium_Source5-Mo Premium_Source5-Mo Premium_Source5-Mo	em an- anc ave sto ger	Equation Editor				Al Character Numeric Date / Time				bbA 😈
Commission Master Earthquake Premium I General Ledger CD Motor Premium Master Premium, Source1/the Premium, Source3/Lau Premium, Source3/Lau Premium, Source3/Lau Premium, Source6/Ag Premium, Source6/Ag Premium, Source6/Ag	em an- anc ave sto ger	Equation Editor				Al B Character B Numeric		@Functions @Functions are available for performing more complex operations		bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nor Premum, Source3-Nor Premium, So	em an- anc ave sto ger	Equation Editor				Al Character B Numeric Date / Time B Matching		Functions Grunctions Grunctions are available for performing more complex operations such as date arithmetic,		bbA 😝
Commission Master Earthquake Premium I General Ledger CD Motor Premium Master Premium, Source1/the Premium, Source3/Lau Premium, Source3/Lau Premium, Source3/Lau Premium, Source6/Ag Premium, Source6/Ag Premium, Source6/Ag	em an- anc ave sto ger	Equation Editor				Al Character Numeric Date / Time Metching Conditional		@Functions @Functions are available for performing more complex operations	X	bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	em an- anc ave sto ger	Equation Editor	2 ² 🔶 🗎 肯			Al Character Numeric Date / Time Metching Conditional		Functions		bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	em an- anc ave sto ger	Equation Editor	2 ² 🔶 🗎 肯			Al Character Numeric Date / Time Metching Conditional		Functions for performing more complex operations such as date antimetrix, stabilized calculations and text searches. IDEA functions begin with they are % symbol. They are		bbA 😈
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	em an- anc ave sto ger	Equation Editor Equation Equation Generati (date) Generati (date) Test Equation	🕵 <table-cell> i i i i i i i i i i i i i i i i i i</table-cell>			Al Character Numeric Date / Time Metching Conditional		Functions		bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	em an- anc ave sto ger	Equation Editor Equation Equation Generati (date) Generati (date) Test Equation	2 ² 🔶 🗎 肯			Al Character Numeric Date / Time Metching Conditional		Granctions Granctions are available for performing more complex, operations such as date withmetic, subsitical dividuations and text searches. IDEA functions begin with the W symbol. They are very sinder in side and operation to functions fond in other software		bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	em an- anc ave sto ger	Equation Editor Equation Equation Generati (date) Generati (date) Test Equation	🕵 <table-cell> i i i i i i i i i i i i i i i i i i</table-cell>			Al Character Numeric Date / Time Metching Conditional		Grunctions Grunctions Grunctions are available for performing such as date authmetic, add et authmetic, add et authmetic, det authmetic, det authmetic, add et al. add et a		bbA 😝
Commission Matter Commission Matter Carthquake Primain Motor Premian Matter Premian, Souce 14 Premian, Souce 340 Premian, Souce 340 Premian, Souce 340 Premian, Souce 440 Premian, Souce	em an- anc ave sto ger	Equation Editor Equation Equation Generati (date) Generati (date) Test Equation	date)			Al Character Numeric Date / Time Metching Conditional		Granctions Granctions are available for performing more complex, operations such as date withmetic, subsitical dividuations and text searches. IDEA functions begin with the W symbol. They are very sinder in side and operation to functions fond in other software		bbA 😝
Commission Mastel E anthquake Premium 1 General Ledger CD Motor Premum Source 1-Pre Premum, Source3-Nou Premium, Source3-Nou Premium, Source3-Nou Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo Premium, Source3-Mo	H Store	Equation Editor Equation Equation Generati (date) Generati (date) Test Equation	🕵 <table-cell> i i i i i i i i i i i i i i i i i i</table-cell>			Al Character Numeric Date / Time Metching Conditional		Granctions Granctions Granctions are available for performing more complex, operation such as date withmetic, subsitical deviations and text searches. IDEA functions begins with the Wory sindler in sighe and operation to functions found in other software pockages such as Merceroft Execut, Letus 1:		🖬 Add

After applying the virtual numeric field we get the following file with monthwise details in the virtual numeric field, as can be seen below.

e Explorer 🕴	Sub Lede	er CD						. x	Properties
्र 🔁 🖉 🚩 🖋 🗑	DELETED	CUST ID DATE	DESCRIPTIO	CEBIT	CREDIT	BALANCE	MONTH	~	I Tatabase
Andynii Hincured Loss 00 Andynii Hincured Loss 00 Andynii Hincured Loss 01 Andynii Hincured Loss 03 Andynii Hincured Loss 04 Andynii Hincured Loss 05 Andynii Hincured Loss 05 Andynii Hincured Loss 04 Andynii Hincured Loss 05 Andynii Hin	1 2 3 4 5 5 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1000 4/1/2005 1000 4/1/2005 1000 4/7/2005 1000 4/75/2005 1000 4/30/2005 1000 5/31/2005 1000 7/31/2005 1000 7/31/2005 1001 4/1/2005	Elaboro In Funds In Cha No. 234567 Marine Export BL 24 Funds In Cha No. 234567 Balance Out Balance Out	0 0 0 0 0 0 0 0 0 0 0 0 0 0		25000 90000 940000 340000 340000 340000 940000 940000 95000 555000 555000 555000 555000 555000 555000 555000 480000 480000 415000 40000 40000 40000 17500 17500 17500 2500 2500 2500 2500 2500 2500 2500	44444507445555550674455566077774456744		Data History Heid Subject Control Total Control Total Criteris Control Total Criteris Ventos No Index No Index Comments Add comment
	36 37	1004 5/31/2005 1004 6/30/2005		0	0	14000	5		
(A)	39	1004 7/31/2005		0	0	14000	7	~	

QUICK FIELD SUMMARIZATION

Perform a quick field summarization on the above file through Analysis > Field Summarization > Quick.

In the Quick Field Summarization dialog box select 'MONTH' as the Summarization Field and 'DEBIT', 'CREDIT' as the Fields to summarize. The process can be seen below

File Explorer				DESCRIPTIO ance In Inds In Cho No. 23456	0	EDIT 6 00000	ALANCE MONTH 35000 4 135000 4	*	Properties 7 II - Database
Altone Analysis Horound Loss (Analysis Horou	(8) Use fields C/Usin fields	mmarization		Numeric fields to total CLIST_ID CLIST_ID CLIST_ID CLIST_ID CLIST_ID CLIST_ID CLIST_ID CLIST_ID Statistics to include Clistic Clistics	Average Vationce Standard devi	stion	OK Felds Carcel Holp		Plef Soberce Plef Soberce Control Total Control Total Control Total Control Total Control Total Control Total Add Comments Add Comments
C Interning Tasks ≥ To Do Us For Help, press F1	32 33 34 35 35 35 37 36 37 38 4 2 38 5 8 27 38	1003 7 1004 4 1004 5 1004 5 1004 5 1004 6	/30/2005 Bal /31/2005 Bal /32/2005 Bal /30/2005 Bal /31/2005 Bal /30/2005 Bal /30/2005 Bal /30/2005 Bal	ance Out ance In ance Out ance Out ance Out ance Out	0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	5000 6 5000 7 14000 4 14000 5 14000 5 14000 6 14000 6 14000 7	× 38	Disk Spars: 11.55 GB

After the quick field summarization, we get the result Quick Field Summarization. IMD as can be seen below.

Simultaneously open the file General Ledger CD. IMD, and tile the two files together through Windows in the menu tool bar.

The comparison of the DEBIT and CREDIT fields from the two files shows that the values match uniformly.

Application of CAATs to Utility Services

Image: Solution and a construction	Explorer 9		4		and a second					× Properties
Image:		The second secon	and the second se							
Conserved Ledger CD Add Mode Phenam, Sourcel Present, Main Conserved Ledger CD × X Present, Sourcel Present, Sourcel Present, Main Conserved Ledger CD × X Present, Sourcel Present, Sourcel Present, Main Conserved Ledger CD × X Present, Sourcel Present, Sourcel Present, Main Conserved Ledger CD × X Present, Sourcel Present, Sourcel Present, Main Conserved Ledger CD × X Present, Sourcel	CA Flos None Analyma Hincared Loss (2) Analyma Hincared Loss (2) Analyma Hincared Loss (2) Analyma Hincared Loss (2) Analyma Hincared Loss Agent Analyma HinCared Loss Agent Analyma HinCared Loss Agent Analyma HinCared Loss Agent Analyma HinCared Loss Agent Compassion Marter Commission Marter	1 4		13 10	45000 3 110000 87500	350000 500000 0				History Field Statistic Control Total Criteris H Results H No Index No Index
■ Premain_Sourcel/Premain_Move ■ CC 30.04.05 April 104000 #Stock #Stock<	General Ledger CD	1								faammon
• Premin_Source/NorryWein p. 1 CD 30.04.05 April 194000 495000 • Premin_Source/Loweld Pdb CD 31.05.05 May 409000 10000 50000 079000 • Premin_Source/Loweld Pdb CD 31.05.05 June 079000 17500 0 791500 • Premin_Source/Area Premin_ CD 31.07.05 July 791500 0 776500 • Premin_Source/Person Edit CD 31.07.05 July 791500 150000 0 776500 • Premin_Source/Person Edit CD 31.07.05 July 791500 150000 0 776500 • Premin_Source/Person Edit CD 31.07.05 July 791500 150000 0 776500		100		Statistical states in the second	L DESCRIPTIO		DEDIT	CREDIT	CONTRACTOR OF A DESCRIPTION OF A DESCRIP	-
	Premium_Source3 Canceled Pole Premium_Source4-Cover Note Premium_Source5-Motor Premiu Premium_Source6-Agency Comm Premium_Source6-Agency Comm Premium_Source6-Premium - Eath Sub Ledger CD Sub Ledger CD	1	888	30.04.05 31.05.05 30.06.05	April May June	104000 409000 879000	45000 110000 07500	350000 500000 0	409000 679000 791500	
	Summarization1									

CONCLUSION

This is representative of the arithmetical accuracy of the sub-ledger postings into the Control accounts.

2.0 IDENTIFICATION OF UN ADJUSTED CREDITS IN CASH DEPOSIT ACCOUNT

OBJECTIVE

To identify unadjusted credits in the Cash Deposit Account from the Sub Ledger.

EXERCISE DESCRIPTION

In this exercise, we will create a pivot table by capturing the Customer ID in the row section, month in the column and balance field in the data grid.

Thereafter we will be in a position to identify unadjusted credits based on the pivot table result.

IDEA FUNCTIONALITY COVERED

This exercise will show us how to

• Create a Pivot Table

Use the file Sub Ledger CD. IMD created in the previous exercise.

BACKGROUND

Cash deposits are advance sums of money placed by corporate insured with insurance companies, wherefrom regular policy premiums are appropriated towards running policies.

CREATE A PIVOT TABLE

Create a Pivot table on this file.

Capture the CUST_ID in the row section, the MONTH in the column section and the field BALANCE in the data grid.

ile Explorer 4	😗 Sub Ledger CD	× Properties
CA Fels Ahome Analysis Horcared Loss 103 Ahome Analysis Horcared Loss 103 Analysis Horcared Loss 103 Analysis Horcared Loss 103 Analysis Horcared Loss 105 Analysis Horcared Loss Agent Analysis Horcared Loss Agent Analysis Horcared Loss Agent Clain, Source-Tophan Clains Clain, Source-Tophan Clains Clain, Source-Tophan Clains Clain, Source-Tophan Master Commission Master	Image: Control (Control (Contro) (Contro) (Control (Control (Control (Contro) (Control (Contro) (II ← Datase Date I history Field Statistics Control Tutal Control II ← Results ✓ Proct Table II ← Indices ✓ No Index MOMIN(A ← CUST_ID(A ← CUST_ID(A
Premium_Source5 Motor Premiu Premium_Source6-Agency Comm Premium_Source6-Agency Comm Sub Ledger CD	Add To Remove Close	Add comment
C	erh Results / @ Condusions	

A closer examination of CUST_ID 1003 and 1004 reveals that the balance of Rs. 5000 and Rs. 14000 respectively has been lying dormant and in credit since May 2005 (Ageing over 3 Months).

CONCLUSION

These cases reflect old outstanding advance receipts, which have not been appropriated towards any specific policy.

Reminders should be sent out to the Insured, informing them of the same and seeking their confirmation on the balances held.

13.1.10 Healthcare Frauds

1.0 EXCESSIVE PROCEDURE BILLING OF SAME DIAGNOSIS, SAME PROCEDURES

OBJECTIVE

To identify instances of excessive medical procedure billing for the same diagnosis and same medical procedure.

EXERCISE DESCRIPTION

In this exercise, we will learn to join the Healthcare Claims transaction file with the master file on the basis of the Diagnosis Code.

Append a computed virtual numeric field to arrive at instances where excessive procedural charges have been claimed by the insured, in comparison to the current master charge list.

Extract cases where the difference exceeds 15% (Acceptable variance norm across hospitals)

IDEA FUNCTIONALITY COVERED

This exercise will show us how to:

- Join Files
- Append a Virtual Field
- Use the Equation Editor
- Apply A Direct Data Extraction

DATA IMPORT

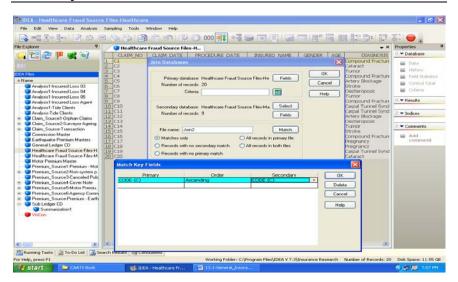
Import the file Healthcare Fraud Source Files. xls

JOIN FILES

Designate Healthcare Fraud Source Files - Healthcare. IMD as the Primary Database and Healthcare Fraud Source Files - Masters. IMD as the Secondary Database.

Join the two databases on the basis of 'Code' from the Primary Database and 'Code' from the Secondary Database. The 'Code' represents the Diagnostic Code.

Use the option 'Matches Only'



APPEND A VIRTUAL FIELD

Append a virtual numeric field by the name 'Diff in Proc Fees'.

III 🗙 🛩		· No In	sebe	-	14	o contro	i total 🔸	100%	-				
L I II	D	AGNOSTI			URE_FEES		AGNOSIS		AL_PROCEDURE			SPITAL1	DIAGNOSTIC,
× 🕰 🔂 🐙 🕅 a	1 2			500	50000 A	rtery BI	ockage	By-Pass	Heart Surgery	AB	Litewat		
ame	3		- Fie	Id Manipulatie							- EX	tetralaya	
162300UW0405	4	r		Field name	Type	Len	Dec	Parameter	Description	-	OK.	tetralava	
Agency Commission N	5		12 0	ROCEDURE P		10	0	Parameter	Concription			010	
Commission Paid	6 7			DIAGNOSIS1	Character	22				- 1	Append	one	
	7					21				- 1	Delete	bre	
Premium Receipts Healthcare Fraud Soc	9			ODE1	Character	3			-	-	Print		
A Mashboara Frand Sor	10			HOSPITAL1	Character	17					PHER	1	
 Join 	11			DIAGNOSTIC FE		8	0			- 1.	Cancel	1	
100 C 200 C	12		10 1	ROCEDURE P	Numeric	12	0				Help	1	
	13		10 0	OFF IN PROC F	Virtual numer +						(TROP	1	
	16	Ł					+						
	12 13 14 16 16 17			500	50000 9			Burr Tap		8	Breach		
	17			000	125000 9			Burr Tap		8	Breach		
	10 19 20			500 350	45000 T			Hystered		T	Breach		
	20			350	90000 T			Hystered		÷	Breach		
	100				0000011							- arroy	

USE THE EQUATION EDITOR

After adding the Virtual Numeric field, click parameter.

This invokes the Equation Editor.

Enter the equation procedure_fees-procedure_fees1 where procedure_ fees represents the fees claimed and procedure_ fees1 represents the master fees.

Validate and exit the Equation Editor.

Application of CAATs to Utility Services

-	🕴 🕜 Haatle aval	Frank Source Films Me	able .	n toin	2				operties
2 📕 🖬 🖬	d Manipulation								atabase
	Field Name	Туре	Len	Dec	Parameter	Tag Name	Description 👗	OK	Data
13	DIAGNOSIS1	Character	22		52000000000	<no tag=""></no>		Append	History
	MEDICAL_PROCEDURE1	Character	21			<no tag=""></no>			Field Statistic
	CODE1	Character	3			<no tag=""></no>		Delete	Control Total
Incurred 16	HOSPITAL1	Character	17			<no tag=""></no>		Print	Criteria
Incurred L 17	DIAGNOSTIC_FEES1	Numeric	8	0		<no tag=""></no>	100		esuits
1-Incurred I 1-Tide Clier 18	PROCEDURE_FEES1	Numeric	8	0		«No tag»		Сору	CARD .
Ide Clerit 19	DIFF IN PROCEDURE F	Virtual numeric		0		<no tag=""></no>		Cancel	ndices
rce1-0rp			-				×		No index
urce2-Sur <							0	Help	No index
guak e Premium Ma val Ledger CD scare Fraud Sourc an 2 scare Fraud Sourc Premium Master	Equation Editor		20.04		IA I	al 💿		-	Comments Add comment
Ledger CD are Fraud Source 2 are Fraud Source emium Master _Source1-Prem _Source2Non- _Source2Non- _Source5-Moto _Source5-Moto _Source6-Ager _Source6-Premi.	🗸 🗟 🖊 🔊 🕏	: 🍫 🖹 📫					OFunctions Ofunctions are available for performing more complex operations such as dote arithmetic, statistical calculations and text searches. IDEA	-	
re Fraud Source re Fraud Source smixm Matter Source1-Prem Source3-Cance Source3-Cance Source3-Cance Source3-Moto Source6-Moto Source6-Ager	🖌 😼 💥 🧐 🧟 Equation	: 🍫 🖹 📫			Al Character Numeric Date / Time Matching Conditional		Functions Grunctions are available for performing more complex operations such as date arithmetic, statistical calculations	-	Add

The List of exceptions are the non zero values seen in the table below under the field DIFF IN PROCEDURE FEES

Explorer 9	Heal	thcare Fraud Source File	s-Health			+ × Properties
	CODET	HOSPITAL1	DIAGNOSTIC FEESI	PROCEDURE_FEES1	DIFF_IN_PROCEDURE_FEES	II 🕶 Database
A Hes Name Analysis I knowed Loss (3) Analysis I knowed Loss (3) Calm, Scured Jophan Clains Claim, Scured Jophan Clains Community (3) Community (3)	1 A8 A8 3 C 5 CF 7 CF 9 CTS 1112 O 9 CTS 1112 O 1115 F 17 S 17 S 1	Llawati Llawati Shunica Netrologo Shunica Netrologo CMC Velore CMC Velore CMC Velore CMC Velore CMC Velore Hinduja Hin	500 1000 3850 3950 7950 7950 5000 5000 5000 5000 4950 4950 4950 4	50000 15000 15000 10000 10000 25000 25000 25000 25000 25000 25000 25000 25000 45000	3000 0 40000 0 0 0 0 0 0 0 0 0 0 0 0 0 0	 ✓ Date ■ History ■ Field Stotistic ■ Control Todal ■ Criteris □ < Tensils □ < Indees □ < Comments □ < Comments □ Add comments
VisCon						,

APPLYING A DIRECT DATA EXTRACTION

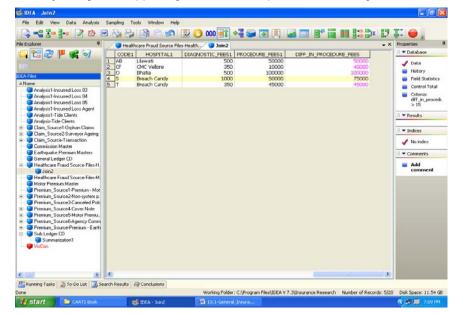
The normal acceptable variance between procedure fees across hospitals is pegged at 15%.

Hence proceed to identify exceptions beyond 15% of procedure_fees claimed.

This is done by writing an equation in Direct Data Extraction.

The extraction gives a child listing of 5 cases, as can be seen below.

These cases should be identified, flagged off and taken up for special scrutiny along with supporting documents, inquiries and investigations.



Other IDEA Tests which can be Run to Identify Healthcare Frauds

- Identify excessive number of procedures per day or place of service per day / per patient
- Identify multiple billings for same procedures, same date of service
- Analyze for mismatched services to diagnosis codes
- Review diagnostic fees in excess of masters
- Review procedure fees in excess of masters
- Identify and report diagnosis and / or treatment that is inconsistent with a patient's age and / or gender

13.2 Application of CAATs to Banks

Banks: Application of CAATs

Introduction

IDEA is useful in most areas of banking, including current accounts, personal and corporate lending, treasury functions and investment accounts. The usual guidelines on volume and depth apply. However, in some banking systems the volume of transactions may be too heavy to download to a PC. It may be appropriate to download a segment of the data or a selected number of fields. Alternatively, a data link may be setup with the Application Server to access raw data in Back-end Tables directly.

IDEA is time-tested, robust and stable. It has been proven globally as an efficient Audit Tool with all the Banking Application Systems:Core Banking Systems, Total Branch Automation Applications, and specialized Banking Domain applications like Applications for Treasury Management and Control.

The following tests cover a variety of account types:

Calculation

- Cast (or total) ledgers.
- Cast (or total) account balances.
- Cast (or total) accrued interest.
- Perform interest calculations.
- Check calculation of bank charges.
- Reprove arrear calculations and reporting (loans).
- Prove interest capitalization

Analysis

- Cast (or total) ledgers.
- Cast (or total) account balances.
- Cast (or total) accrued interest.
- Perform interest calculations.
- Check calculation of bank charges.

- Reprove arrear calculations and reporting (loans).
- Prove interest capitalization

Exception tests

- Identify accounts with missing standing data.
- Identify invalid or unusual standing data.
- Identify dormant accounts and transactions therein
- Identify accounts with statement suppression
- Check charges are being raised where appropriate
- Identify staff loans
- Identify large loans
- Test for unusual interest rates
- Identify loan balances greater than original advance
- Identify negative balances
- Identify forward dated transactions
- Provide totals of forward dated transactions
- Identify overdue maturities
- Identify customers over their overdraft limit or customers with expired limits
- Perform currency conversions
- Provide details of currency exposure

Other useful tests include stratifying interest rates against each risk rating. Generally higher risk loans should bear a higher rate of interest. The analysis followed by exception testing can identify anomalies or branches/individuals that are either risky or particularly prudent in their lending.

Matches and Compares

- Multiple loans to the same address by setting up a key based on the address and using duplicates test
- Match addresses of loans to employee addresses
- Compare balances at different periods to identify movements.

Revenue Assurance for the Bank using IDEA

CASE

Verification of accuracy of Demand Draft charges applied by the Branch Software

OBJECTIVE

To verify the correctness of Demand Draft charges levied.

EXERCISE DESCRIPTION

In this exercise, an equation is written that expresses logical criteria of DD charges applicable to DD amount as per the DD charges handbook issued by the Head Office of the Bank to the Branch for Retail Banking.

IDEA FUNCTIONALITY COVERED

This exercise includes the following steps -

- Append a computed Virtual Numeric Field
- Use the Equation Editor

METHOD

The Equation is composed as follows and can also be saved for future use

@compif((@between(dd_amount, 0,5000),25, @between (dd_amount, 5001, 25000), 40, @between(dd_amount, 25001, 50000), 75, @between(dd_amount, 50001,150000), 150, dd_amount>150001, 250+2.5*(dd_amount-150001)/100))

Multiple Approaches to Data Collation and Interrogation within IDEA for Banks

Approach 1

1. Requesting IT for Data

Case Study 1.1: Electronic KYC Audit

The team selected an objective to test KYC norms on Current account customer master data.

To test this objective the team issued a Data Request to IT in the following format:

Data required: Current account customer master information.

- Period: As of the date of Audit (period cannot be a date. Rethink).
- Fields of reference: Branch ID, Customer ID, Account ID, First Holder & Joint Holder/s Name, Address, PAN No., Mobile No., Residence No., Office No., Mode of Operation and Clear Balance.
- Format of Data: Text form.

IT in turn ran a SQL query on the production database and generated a text file dump, which was saved by IT in a secure folder with special access to the Audit team only.

The Audit team imported the text file, using the Text Report import option within the GAS.

Post import, the team used the Duplicate Key test within the GAS to identify fictitious accounts opened with similar PAN No., or Mobile No., or Address, or Office No., or Residence No., but different Customer ID.

50 cases out of 65,000 were identified where key account opening fields like PAN No., Mobile No. were same for different Customer ID's. These cases have been taken up for further checking with the account opening forms from the respective Branch to ascertain the validity of the accounts opened.

Additionally, the Team employed a SOUNDEX test on the Customer Name field. The SOUNDEX test generates a unique alpha-numeric code in a new computed field. This code depends on the mnemonic of the Name rather than the spelling. For example, Two Customer Accounts with names Prasanna Iyer and P A Iyer will have the same SOUNDEX code. After arriving at this code, duplicate tests can be run on the code rather than the Name. The advantage here is that the Duplicate on Code will generate similar sounding names (Essence of De-Dup Tests Globally), whereas a pure Duplicate will perform an exact match.

The Team also performed De-Dup Tests to identify Customer's having multiple Savings Accounts within the same Branch of the Bank, which is a potential red flag.

Case Study 1.2: Data Migration Audit

The team then decided to check the integrity of loan data migrated from the Legacy TBA application to the CBS.

To test this objective the team issued a Data Request to IT in the following format:

- Data required: Cash Credit master information for large scale branch X.
- Period: Data immediately post-migration.
- Fields of reference: Customer ID, Sanction Limit, Drawing Power, and Rate of Interest.
- Format of Data:Text form.

IT in turn ran a SQL query on the production database and generated a text file dump which was saved by IT in a secure folder with special access to the Audit team only. The corresponding data from the TBA legacy system (immediately pre-migration? Not clear) was available with the Migration Team. IT sourced the Text data from the Migration Team through a formal email request and placed the data in the secure folder.

The Audit team imported both the text files using the Text Report import option within the GAS.

Post import, the Team linked the pre-migration and post-migration data through the JOIN function in the GAS. The two data files were linked based on the Customer ID available in both the files. Post join, three new fields were created by the Team containing the differences in the Sanction Limit, Drawing Power and Rate of Interest in each field. The Team finally queried each of the fields for non-zero data.

Accounts that showed a difference in the masters migrated (non-zero data) were identified by using the procedure given above. These accounts were taken up for further testing with the legacy system data and the loan appraisal forms to ascertain their accuracy.

Approach 2

2. Using existing End of the Day (EOD) reports or existing Business Warehouse (BW) solution reports.

Case Study 2.1: Audit of movement on in-operative accounts

In the second data retrieval approach, the Team decided to audit the movement on inoperative accounts over a period of one financial year.

To test this objective, the Team identified the Report on Inoperative accounts as on 1st April 2014 and 31st March 2015 respectively. These reports were a part of the EOD suite of reports. These reports were generated as a part of the end of the day routine at Branch X and by default saved in a compressed print report format.

The Team uncompressed the reports using WIN-ZIP. Then both the print report files were imported using the Text Report import option within the GAS.

Post import, the Team linked the Report on Inoperative accounts as on 1st April 2014 and 31st March 2015 using the COMPARE function within the GAS. The two data files were linked based on the Customer ID available in both the files. Post compare, a new file was created with differences in the Clear Inoperative balances. The Team finally queried the difference field for non-zero data.

Accounts in which there was a reduction in the inoperative account balances (non-zero data) were identified through the above approach. These accounts were taken up for further testing with EOD Exception reports to ensure that the movement on the inoperative account was authorized by the Branch Manager.

Case Study 2.2: Audit of Loans and Advances MIS from Business Warehouse (BW) for approval verification

In this case, the Team decided to audit the MIS report on Loans and Advances generated from the Business Warehouse (BW) for approval verification. The MIS report is a comprehensive listing of Account ID, Type of Loan, Type of Security, Type of Industry, Sanction Limit, Drawing Power, Rate of Interest, Due Date of Loan, and Approval Officer.

To complete this test, the Team sought assistance from the Merchant Banking wing of the Bank. The Loans officer for Branch X generated the MIS report for the period 1st April 2014 to 31st March 2015. This report was saved as a Microsoft Excel file and provided to the Team on a CD.

The Team imported the Microsoft Excel file using the MS-Excel option for import within the GAS.

Post import, the Team entered a conditional criteria/query on the file, identifying non-compliances with the Sanction Limit-Approval Officer table. The CM, DGM, AGM-HO and AGM-ZO of the Bank were all vested with specific approval Sanction limits. The Team tested the file for cases where loans sanctioned by officers were not in accordance with their financial Sanction Limit powers.

Some cases were identified in which the loans approved did not fall within the Approval Officer's sanctioning limit. These cases were noted and taken up for review with the respective officer and the Branch Manager.

Approach 3

3. Access to Disaster Recovery (DR) site databases or mirror databases

The Team, with the help of the IT wing, set up connectivity between the GAS and the DR site server. This connectivity was the third and final mode of data retrieval, which gave the Team direct access to raw data tables residing on the DR site server.

Audit Tools like IDEA can work optimally on very large databases. For example, the Team connected to the DR Site Server during peak business hours imported 6 Tables from the Core Banking System Server. The Tables had a cumulative record size of 5 crore records, and the imports were completed within 45-50 minutes cumulatively (Key Performance Metrics). The Teams having specialized skill-sets in SQL Query Building, performed consecutive ODBC links on the 6 Tables. For each link, simple SQL queries (with written?), to slice and dice the data down to exact requirement:. for example, Data for Branch A and Branch B from 2050 Branches. In the Transaction Tables, the Team performed additional SQL Queries for transactions with value dates between 1st April 2014 and 31st August 2014.

The Teams also automated the import links to run as scheduled tasks during non-peak hours. This was done by converting the import tasks into MACROS and MACROS into Executable files. These executable files were programmed to run through IDEA by Windows Tasks Scheduler. The import tasks for the 6 Tables ran automatically at 7 PM in the evening, gave results similar to the peak hour imports with a throughput of 38 minutes vis-à-vis 50 minutes during business hours.

The successful Automation and Scheduling of tasks gave the Team the confidence to deploy Continuous Control Monitoring on a regular basis in the Bank. The Team graduated to running complex imports and queries on weekends, holidays and night-time slots to arrive at real time results at the start of work on the next business working day.

Case Study 3.1: Identifying Term Deposit accounts where the deposit was accepted for a period greater than 120 months.

As per the bank-rules for acceptance of Term Deposits and guidelines of the regulator, the Bank cannot solicit Term Deposits in excess of a period of 120 months. This test was undertaken to identify non-compliances with the rule.

The Team connected to the DR site server using the GAS and identified the data table containing the Term Deposit account masters. At the import stage, the Team selected three fields – Account ID, Account Name and Deposit Period in Months from the table containing 15 different fields. The Team also entered an SQL query in the GAS, filtering the data for Branch X. The import of data of 1 Lac rows/lines was initiated and duly completed within 1 minute. This was faster in comparison to writing an SQL query, which would take 3-4 hours to run in the past.

Post import, the Team queried the data to identify accounts where the Deposit Period in Months was greater than 120.

The result was exported from the GAS back to MS-Excel, printed and taken up for discussion with the Branch Manager and the Head of Term Deposits.

Case Study 3.2: Isolating loan accounts where the Drawing Power was greater than the Sanction Limit.

The Team connected to the DR site server by using the GAS, and identified the data table containing the account masters. At the import stage, the Team selected four fields – Account ID, Account Name, Sanction Limit, and Drawing Power from the table containing 35 different fields. The Team also entered an SQL query in the GAS, filtering the data for Branch X.

Post import, the Team applied criteria to the file identifying instances where the Drawing Power was greater than the Sanction Limit.

The result was exported from the GAS back to MS-Excel, printed and taken up for discussion with the Branch Manager and the Head of Loans & Advances.

13.3. Application of CAATs to Retail Sector

Retail Malls – Application of CAATs

Introduction

In the current business environment, the burgeoning growth of Retail Malls/Chains across India and their internal controls need to be controlled and monitored by Auditors for their continued adequacy and effectiveness.

Retail Malls are highly tech-centric with online point of sale billing applications, automatic inventory updation and Electronic Data Interchange re-ordering. There is widespread use of bar coding and bar scanning for billing. Given the high proliferation of technology, it is best to deploy technology in the form of CAATs to control Retail Technology.

Our case-study looks at Point of Sale Billing Systems, the import of data from such systems and the interrogation of data generated through such systems.

Point of Sale Systems from the Retail Industry – Malls

The Point of Sale Applications deployed in Malls generate comprehensive Sales Reports. These can be conveniently saved in MS-Excel Data formats. This is an alternative to File Print options.

These Sales Reports contain fields like

- Date of Sale
- Time of Sale
- Transaction Number
- Cashier Name
- Cashier ID
- Product Sold
- Quantity
- Rate
- Gross Value
- Taxes
- Scheme Discounts
- Net Value
- Scheme ID
- Scheme Details
- Collections in Cash
- Collections by Card
- Dues

Import into IDEA

These MS-Excel Data Files once saved on the workstation containing IDEA or on the Local Area Network in a shared audit data folder can be accessed

through IDEA's Import Assistant, Excel component. The process of import is simple and easy to apply, since Excel file Record Definitions are readily recognized by IDEA.

Interrogation within IDEA

Case Study 1: Reconciliation of Net Sales with Cash Takings, and Card Receipts

The fields of reference relevant to the objective being tested are

- Net Sales
- Cash Collected
- Card Receipts

The process of interrogation followed in IDEA is:

- Import the Sales Report for a given period through IDEA's Import Assistant MS Excel.
- Navigate to the Field Statistics in the Database Toolbar.
- View the numeric control totals for the Net Sales, Cash Collected and Card Receipts fields respectively.
- Normally the Net Sales should be arithmetically balanced by Cash Collections and Card receipts.
- In the case under review we notice a high percentage of Unpaid Bills, almost 25% of the period's Net Sales.
- An overview of the Unpaid Bill cases through Field Summarization reveals that the due amounts are significantly concentrated on Cashier A and Cashier D.

These cases may be specially looked into by interrogating the concerned Cashier's, to ascertain their motive.

Case Study 2: Inconsistent scheme discount rates offered by Cashier's to different customers against the same Scheme ID

The fields of reference relevant to the objective being tested are

- Cashier ID
- Scheme ID
- Scheme Discounts
- Gross Value

The process of interrogation followed in IDEA is:

- Navigating to Data in the Menu Tool Bar and selecting Field Manipulation
- In Field Manipulation, appending a computed Virtual Numeric Field Discount % with the Criteria (Scheme Discounts*100 / Gross Value), rounded off to the nearest integer.
- Navigating to Analysis in the Menu Tool Bar and selecting Duplicate Key Exclusion
- In Duplicate Key Exclusion, we identify different Discount % values for the same Scheme ID.
- We get a list of cases where varying Discount % have been applied for the same Scheme ID.
- Some cases are extremely glaring, with the Discountbeing as high as 45% though the Scheme ID warrants a Discountof 15% only.

These cases may be specially looked into by interrogating the concerned Cashier, to ascertain his motive.

Case Study 3: Identifying the most efficient Cashier across all the Malls – Operational Throughput

The fields of reference relevant to the objective being tested are

- Cashier ID
- Cashier Name
- Start Time
- End Time
- Quantity

The process of interrogation followed in IDEA is:

- Create a new computed numeric field in the imported Sales File with the difference between the Start Time and the End Time using the criteria '@agetime(End Time, Start Time)'. This new field will give us the time taken by each cashier in seconds to scan, bag and bill all the items against that transaction.
- Create another computed numeric field with the criteria 'Difference in Time/ Quantity' to arrive at the Time taken to scan each item.

- Now perform Field Summarization on the Cashier ID and Cashier Name with regard to the numeric field containing the Time taken to scan each unit.
- In the Field Summarization also include additional statistics like Count and Average along with Min and Max. These statistics will give us the number of scans by a single Cashier in a given period, the minimum time for scan, the maximum time for scan and the average time for scan.
- In the Summarization result, sort the Time to Scan (Average) on a Descending basis.
- The Cashier with the best scan rate appears right on top.

The Management of the Mall Chain can select the Top 10 Cashiers from the above exercise and place them at front desks on crucial days like weekends or public holidays. This test is a concrete measure of customer service.

Case Study 4: Detecting transactions Out of Office hours

The fields of reference relevant to the objective being tested are

- Start Time
- End Time
- Cashier ID
- Cashier Name
- Net Sales

The process of interrogation followed in IDEA f is:

- Perform a Data Extraction on the imported Sales File.
- Build a criterion using the function .NOT. @betweenagetime(Start Time, "10:00:00", "22:00:00") .OR. .NOT. @betweenagetime(End Time, "10:00:00", "22:00:00")
- This criterion will isolate all transactions out of the normal Mall working hours of 10 AM to 10 PM. Here we trap both Start Time and End Time.
- The Direct Extraction function within IDEA is workable on large databases, say upwards of 1 Crore Transactions. The function first sorts the entire database and then runs the equation through the sorted database. Hence the results arrive faster than running an ordinary command on an unsorted database.

The case reveals around 50 out of 1 Lac transactions where the Start Time and End Time is after office Hours at 10:30 PM. An explanation from the Mall Incharge reveals that all these transactions pertained to Public Holidays, when schemes were launched and working hours were extended by two hours with permission from the local administration.

Case Study 5: Demand Study of New Products introduced into Pilot Mall Outlets across India

Malls introduce new products into Pilot outlets to study the Customer Behavioral patterns, spending patterns, loyalty to existing products rather than new substitutes.

In this Business Case Study the following fields are considered

- Transaction Number
- Product Number
- Quantity
- Net Sales

The process of interrogation followed in IDEA is:

- The Transaction Number contains a combination of the Mall Outlet ID, the financial year and the transaction ticket number. The first 3 digits of the Transaction number represent the Mall Outlet ID.
- We append a new computed character field with the aim of getting the Mall Outlet ID into this field. This is performed through a criterion /equation @left(Transaction ID, 3). This function removes the first 3 digits from the Transaction Number and places it in a separate computed field.
- Field Summarization is performed on the imported Sales file on Mall Outlet ID and Product Number/Code with respect to Quantity and Net Sales.
- In the Summarization result, a direct extraction is performed on the Pilot Product Malls through the Equation @ list(Mall Outlet ID, "003", "005"...)
- A Top Records Extraction is performed with the Key field being the Mall Outlet ID, and the top 5 products in each of the 10 pilot malls are identified.

A detailed review of the final result broadly confirms the Management's

expectation. All the new products have fared well in the 10 pilot malls, save one product, which has not been preferred over its existing competitor. This exercise has armed the Management with factual historical data from a truly representative sample of Mall Outlets. Now the Management is in a position to slowly roll out the most liked products to the remaining Mall outlets over the next month.

'Retail Performance Management' enables decision-making for retailers of all sizes and segments by empowering them with comprehensive relevant 'enterprise business intelligence', across technology platforms.

With CAAT, users can jumpstart their analytic journey, and enjoy improved margins, better customer retention, inventory efficiency and promotion effectiveness.

CAATs :

- helps accelerate retail organisation's analytic maturity, taking the user one step closer to achieving excellence.
- (2) creates benefits by delivering enhanced usability by anticipating the evolving needs of decision-makers and ensuring a faster adoption rate.
- (3) through simple screen-guided analytics, empowers every decisionmaker in every role in the retail organisation. It takes the load off the IT group by being easily extendable and maintainable.
- (4) avoid latency, cost and project management challenges associated with traditional business intelligence methods and enjoy unparalleled speed to benefits.
- (5) transforms business intelligence from being a 'decision support system' to a 'decision-making system'.
- (6) makes business intelligence pervasive across the retail business by impacting both the top-line and bottom-line performance.
- (7) take on a whole new revolutionary role in 'Retail analytics' where the tool is used for continuous monitoring by 'process owners' rather than the erstwhile traditional continuous auditing by internal auditors.

However, the significance of CAAT is greatly accentuated by the understanding of the underlying business process by the 'Process Owner'. 'Retail Analytics' can thus deliver immeasurable business benefits.

Merchandising and assortment :

It's a well-known fact that shoppers prefer to visit places that offer them the maximum options, deals and not to forget a good shopping experience. As a retailer, one does everything to attract and retain customers by providing excellent customer service in terms of variety of products, price and promotions to achieve 'customer satisfaction'.

Here is where 'merchandise and assortment planning' comes in. CAAT provide merchandisers an analytic framework to plan and analyse business activities related to 'merchandise and assortment'. It :

- (1) compares historical, planned or forecasted data against actual data to define and optimise merchandise plans
- (2) analyses merchandise hierarchy across departments, categories, product lines and 'Stock Keeping Units' SKUs
- (3) analyses the performance of new products
- (4) determines customers' seasonal requirements/ preferences.

In short, CAAT provides data of when and what customer needs and all this leads to customer satisfaction which results in customer loyalty.

Loss prevention :

In the retail industry, it is well known that losses due to fraudulent transactions, theft, pilferage, excessive stocking, wastage, shoplifting, internal theft, refunds, exchanges and excessive discounting are inevitable. While one can't do away with these problems, retailers are always on the lookout for ways to minimise losses while keeping costs minimum.

Loss prevention analytics helps the retailer diagnose the root cause of the problem by identify exceptions and taking corrective measures. CAAT substantiates its analyses with historical, geographic, and demographic trends. It can identify :

- (1) incorrect refunds which could be fraudulent
- (2) spoilage, damage and write-offs
- (3) price overrides and improper discounting
- (4) supplier or warehousing issues
- (5) administrative errors

- (6) sales to customers with dubious shopping records such sales indicate fraud.
- (7) erroneous entries for product returns.

Supplier performance :

This is an area that tends to get neglected. However, focusing on this area can help you bring down operation costs because CAAT also enables to forecast optimal levels of inventory, optimise lead time, manage orders, improve fill rates, negotiate trade promotions, manage risks and improve supply chain efficiency — these are just a few challenges faced by retailers when it comes to managing supplier performance.

CAAT provides a decision-making framework that enables the retailer to identify new areas of synergy and avenues for bringing about operational excellence. It enables :

- (1) improved inventory management by tracking slow and fast-moving goods, helps avoid 'out-of-stock' situations which lead to avoid loss of sales and profit. In other words, it enables the retailer gain better visibility into stock movement from warehouse to stores.
- (2) improve vendor performance
- (3) assist in rewarding performing vendors
- (4) remove non-performing vendors
- (5) negotiate supplier involvement in product trade promotions for better deals
- (6) negotiate longer credit periods and shorter delivery cycles
- (7) manage supply-related risks and take corrective measures proactively.

Store performance :

CAAT provides retail operation managers a framework to analyse 'store performance' to increase 'store productivity'.

- reclassify stores by local demographics, competitive density, store locations, size and age
- reclassify merchandising categories based on the relationship between the customer, product and store — merchandise assortment required at the store.

- measure contribution and competence of store employees by monitoring their contribution to total sales
- benchmark, compare and rank peer groups based on metrics like yield per square area and average price per item sold.

Customers :

Customer data has long been touted as a key determinant in making better merchandising decisions; however it is an area/asset most retailers have struggled to use to its maximum potential. CAAT provides the retailer with the critical platform he needs to leverage customer loyalty data, sales transaction data, and store data to improve merchandise planning and sales tactics.

CAAT unveils hidden relationships with the customer, because it helps in identifying what the customer buys, how he pays and with what frequency he buys. This enables the manager to :

- (1) identify which store needs what and when thus reducing waste and ensuring optimum use of the shelfspace.
- (2) evaluate performance of promotion schemes.
- (3) measure campaign effectiveness what media can be used to achieve better results by having customer feedback.
- (4) identify regular high-value customers who can be suitably rewarded
- (5) identify what products customers buy together for example floor cleaners are bought alongwith utensil and crockery cleaners and vicea-versa. In retail trade this is known as 'product affinity' or 'market basket'. This enables the retailer to bunch products in promotional schemes and by placing them close to each other on the store-shelf, thereby increasing sales.
- (6) identify seasonal trends for example increase in chocolate sales during festive season – what brand and in what 'pack size'.
- (7) identify 'impulse buying' generated by a new product or new packing.
- (8) weigh loyalty schemes consider whether the customer should be given more of what he normally buys or offer him/her a new product and/or offer him/her better service — for example — having a dedicated counter.

Conclusion :

CAAT creates an environment where the process owners — the retailer — can make informed and real-time decisions on :

- Which customer segments are the most profitable ?
- Which customers should a campaign target ?
- When and how should he communicate with a customer ?
- Which customers should he spend resources on retaining ?
- Which products should be 'cross-sold' ?

We are at the dawn of mature retail analytics for the discerning retail customer. It is an 'idea whose time has come'.

The use of CAAT by the internal auditor on a sample basis on 'customer analysis' motivates the retailer client to increase greater use of CAAT. The tool introduced by the internal audit led to increased 'customer satisfaction' and increased profits.

13.4 Application of CAATs to Utility Services

Electricity Utilities – Application of CAATs

Introduction

Revenue Audits take on a pivotal role in the audit of Electricity Companies. These companies have specialized Billing applications to record the distribution of power to households and industry. These sophisticated Billing systems are able to generate Billing Sheets and Customer Personal Ledgers under normal and exceptional circumstances like temporary meter disconnection, new consumers not yet metered and the like

These application systems are designed to generate exceptional reports like Zero Value Billing, Customer Overdue, Security Deposit Recoverable, and Ageing of Faulty Meters.

Majority of these applications run on Oracle as the back-end database due to its stability, performance and security. All the billing related information is stored in a single billing data table.

The user can integrate IDEA with the Oracle database through an ODBC link, and replicate the billing data onto his system. Alternatively, the user can request the IT Cell of the Electricity Company to provide the Billing Data Table in a flat file data format like (Delimited for instance). These file formats can also be conveniently imported into IDEA.

Bird's Eye View of Functions which can be tested using IDEA for Electricity Companies

1. Recalculation of Revenue Heads

Electricity Companies levy a host of charges to both household and industry customers. These vary from fixed processing cycle charges like Meter Rent to variable charges like Energy Charges. The Company also buys power from state private players to meet the power deficit. Charges levied by the private player to the electricity company are passed on to the final consumer. These Additional Supply Charges or Fuel Escalation Charges vary on an average from month to month.

The objective, description, rationale and authority to raise these charges are stated in a Power Tariff Manual, which is issued in the common interest of operators and customers for transparency. The manual contains the formulae for the calculation of each charge very clearly.

The user can get the manual and identify the charges to be recalculated along with the formulae. Nowadays, because of the Right to Information Act, the Tariff Manuals are also available on the website of the Energy Major.

The user can then create a new computed numeric field (Virtual Numeric Field) within IDEA and replicate the tariff manual formulae in IDEA's Equation Editor.

Illustration – (Units_Consumed * 0.45) - ASC

ASC – Additional Supply Charges stands for purchase of power from private players. The rate of 0.45 per unit as seen in the formulae above is announced by the State Electricity Regulatory Commission on a month-on-month basis.

In this equation we recalculate the ASC and arrive at possible differences between the recalculated ASC and the ASC levied by the billing system.

2. Verification of Subsidies provided by the State Government to defray the levy of Full Energy Charges to the Final Customer

In many States in India, the energy charges to be recovered by the Electricity Company are distributed between the final customer and the State itself. Here the State steps in and contributes to the Energy charges by providing Subsidy.

The State Subsidy is invariably a percentage of the Net Billed Amount.

The objective of the test is to ensure that the right amount is being recovered from the State as Subsidy.

The IDEA user can accomplish this task easily by creating a new computed numeric field with the formulae

(Net_Amt * 0.40) – Subsidy

Here we assume that the State supports the Energy Bill up to 40%.

Subsidies are normally provided to households and critical industry units. Hence the IDEA user has to obtain a listing of Customer Code Categories, to know who are entitled to Subsidy and who have to bear the entire Energy Bill.

With this information n the User can make use of the Display or Extraction function in IDEA to identify eligible and ineligible Customer Categories.

3. Monitoring of Faulty Meters

Next to power theft, which ranges between 10%-30% in India, faulty meters are the high-risk area for an Energy Major.

Billing Data Files invariably contain a field for Faulty Meters. If the meter is Faulty a flag 'Y' appears against the concerned Customer. If the meter is running, the flag 'N' appears in the respective field and cell. There is also an additional field available which states the date when the meter became faulty and the average units consumed. The average units consumed is updated on the basis of past usage and history of each user. This field is invariably updated by the Billing clerk manually. In the absence of automatic system generation of average units consumed, this is a ripe area for mismanagement of revenues.

With the help of IDEA, the user can link the faulty meters, the average units consumed and the last consumption prior to the meter going faulty through the Visual Connector. The user can then create a new computed numeric field where the average units being billed fall short of the last active consumption reading by say 20% (norms can be decided on a case-to-case basis depending on the Customer Class, Geographical Region and the like). These cases can be taken up for review, discussion and scrutiny with the respective Regional Chief Electricity Engineers.

4. Duplicate Tests on Billing Information

The Billing file contains the Customer Number, Meter Number, Bill Number and Receipt Number where the Bill has been duly paid.

As a part of the standard financial and revenue integrity testing, the user can employ the Duplicate Key test within IDEA to look for possible duplicates in the combination of the following fields

Customer Number

Meter Number

Bill Number and

Receipt Number

The existence of duplicates could indicate an irregularity which needs to be investigated further in the Billing Application System.

Accounts Payable and IDEA

14.1 IDEA and the Accounts Payables Process

14.1.0 Introduction

It is vital to establish that controls in Accounts Payable processes are in place and working. For example, to make sure that the client is procuring the right products, at the right price, at the right time and from the right vendor.

CAATs – IDEA helps us to identify areas where the controls are failing because of actual data errors. These factual cases serve well in building up a strong defense and back-up for the Auditor to express opinions on the processes.

IDEA can be applied to Accounts Payable processes from an Internal Audit, Statutory Audit and Forensic Review perspective. In this chapter, we deal with IDEA applied to Accounts Payable from a Sarbanes Oxley perspective.

Most of the tests within IDEA can be automated and scheduled, which is the heart of Continuous Control Monitoring and Pre-Audit of Accounts Payable.

14.1.1 Potential Risks

The following table identifies the key risks, explains the business and audit implications of each risk and the audit objectives that can be realized by audit tests.

S No	Risk	Implications	Audit Objectives
1	Payments are made to unauthorized suppliers.	Unauthorized suppliers could represent former suppliers that supplied goods or services of unacceptable quality and should have been removed from the list of suppliers; or they could be fictitious suppliers set up by dishonest personnel to	Existence, Validity

Payroll and IDEA

S No	Risk	Implications	Audit Objectives
		receive automated payments. Payments made to unauthorized suppliers could therefore represent either error or fraud.	
2	Payments are made to individuals or employees.	Payments made to individuals or employees could represent a diversion of company payments, indicating fraud.	Existence, Validity
3	Unauthorized premiums are given to suppliers.	Unauthorized premiums may represent overpayments to suppliers in return for kickbacks.	Existence, Validity
4	Invoices are paid late.	Delays in processing Accounts Payable approvals can result in a loss of available discounts for timely remittances and understatement of liabilities in a particular period.	Cut-off, Completeness
5	Invoices are paid as per irregular or inconsistent credit period schedules	Irregular payments may reflect processing errors or fraud.	Existence, Validity
6	Invoices are processed twice.	Duplicate payments can result from the failure to cancel documents to prevent re-use or processing errors in Accounts Payable such as restoring a backup file twice.	Existence, Validity
7	Payments are made in a way	Perpetrators of fraud may arrange payments to avoid	Existence, Validity

S No	Risk	Implications	Audit Objectives
	that these are not detected by audits.	detection. For example, large amounts may be split into several smaller payments to coincide with a perpetrator's transaction approval limits or to avoid limit checks on large payments.	
8	Items (e.g., P.O.s, checks) are missing.	Since Accounts Payable are often not authorized for payment until there is a three-way match of purchase order, receiving document and supplier invoice, missing documents could result in Accounts Payable being understated.	Completeness

14.1.2 Potential Tests

The following audit tests are suggested for auditing an Accounts Payable system. However, the exact tests to be carried out for a particular client will depend upon the system used and the data available. Common tests include:

Mechanical Accuracy and Valuation

- 1. Total the file. It often pays to separate debits and credits.
- 2. Revalue foreign payables, if applicable.
- 3. Check transaction totals to the balance on each account.

Analysis

- 1. Stratify the size of payments and extract any exceptionally high payments.
- 2. Analyze payment days and identify suppliers with favorable payment terms.
- 3. If the computer system captures the approving authority for a transaction, examine the value distribution for each manager.

Exception Tests: Existence and Validity

- 1. Identify payments made to unauthorized suppliers by matching the payments and authorized suppliers list.
- 2. Search payments file for payees without "Inc", "Ltd", or "Co" in their name to identify payments to individuals.
- 3. Test for large discounts.
- 4. Test for duplicated invoices using value and supplier code as the key fields for one test and purchase order number for another. The second processing of invoices can be used to establish a value on the P/L to make a fraudulent payment. This will also pick up accidental duplication.
- 5. Identify payments made on Sundays or other days/dates that are not valid.
- 6. To find if amounts are being approved at or just below break points in authority level by a value distribution across the whole ledger. If approval authority is not directly available, perform subsidiary analysis by types of supplier or approving department (i.e. marketing).
- 7. Look for split invoices to enable approval to be kept by an individual. Extract all invoices within 90% of an approved limit (preferably for a suspected manager or department) and search for all invoices from that supplier. Sort by approving manager, department and date to identify possible split invoices or summarize payments by invoice number to determine how many partial payments have been made for each invoice.
- 8. Test for sudden increase or decrease in supplier payments between two time periods like two years.
- 9. Test for large one-off payments to suppliers.
- 10. Using the first five or six characters of the name, match supplier names against a list of employee surnames from a payroll or personnel file.
- 11. Test for suspected duplicate supplier names.
- 12. Test for incomplete or unusual supplier details.

Gaps and Duplicates

Test for missing items or gaps in the check number sequence.

14.1.3 Accounts Payable Fraud Tests

Employees as Vendors

An employee sets up his or her own company and then funnels purchases to that company. Variations include a "ghost" approach where invoices are sent from the employee's company but no actual goods or services are provided. In other instances, actual goods may be shipped.

To detect this type of fraud, ask operational managers to review new vendors as they come aboard. A quick phone call to the vendor may reveal suspicious activity. Make sure, however, that the person doing the review is not the perpetrator. With IDEA, we can use sampling techniques to generate a list of vendors to verify.

Favorable Treatment of Vendors

Look closely at vendors who have a relationship with an employee – a spouse, friend, social partner/ buddy, etc.

Pivot tables in IDEA are effective in detecting this type of fraud.

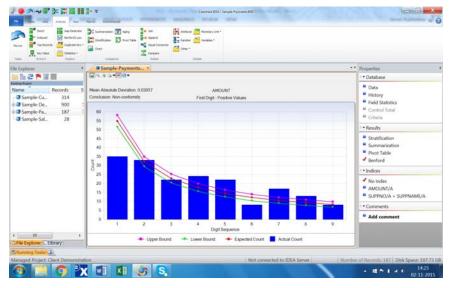
Once created, the table shows the percentage of purchases by each vendor. Always work with an operational expert to interpret the results. It may be normal, that 60% of purchases are from a single vendor, if there are only two or three vendors who can provide the item.

Transactions at or Near Spending Authorities

Spending limits are often referred to as the "trigger price". As soon as a preset limit is reached, the system automatically triggers some form of action or examination.

The perpetrator's goal is simply to avoid the threshold approval price. For example, it may be known to almost everybody that the spending authority limit is Rs. 50,000. The fraud is designed to stay under the threshold approval price.

This can be detected by employing Benford's Law of Digital Analysis, which detects instances where prices are just below the threshold approval price. The Benford Curve will reveal sharp, striking indications of actual frequency



occurrences in excess of expected and upper tolerance frequency occurrences.

Another common scenario involves paying for large purchases over several invoices. To close a sale, the vendor may accommodate this request. Look for transactions with the same vendor that are close to the threshold approval price. Comparing inventory records is also an effective way to detect this. If three invoices are used to count for the full amount, an inventory check will reveal only one set of goods.

Sarbanes-Oxley Section	Requirement	IDEA Test
Section 302: Corporate Responsibility for Financial Reports	Review of Financial Transactions of Significant Accounts	Use IDEA to summarize all detailed financial transactions for Accounts Payable and agree to General Ledger
	Financial Transaction Analysis	Use IDEA to re-age Accounts Payable, use the "Field Statistics" feature to identify unusual financial

14.1.4 Accounts Payable - Sarbanes Oxley Testing

Sarbanes-Oxley Section	Requirement	IDEA Test
		transactions i.e. outliers.
Section 404: Management Assessment of Internal Controls	Matching Controls	Use IDEA in Accounts Payable to verify the matching of Purchase Orders, Invoices and Receivers.
	Cut - off Controls	Use IDEA to test Expense booking by extracting all Creditors for Expenses transactions before and after cut-off dates to verify inclusions and exclusion in General Ledger for the accounting period.
	Limit Controls	Use IDEA to extract all checks (vendor payments, paychecks, commission checks, bonuses, employee reimbursements) above threshold limits for dual approval authorizations.
Section 806: Whistleblower Protection	Protection of Employees who provide evidence of fraud.	Use IDEA to select a sample of phone calls and incidents from Hotline database and test the effectiveness of the Anti-Fraud Hotline program.

14.2 Accounts Receivable and IDEA

IDEA and the Accounts Receivable Process

14.2.0 Introduction

It is vital to establish that controls in Accounts Receivable processes are in place and working as intended. For example, to find out –

- Whether the client is monitoring the overdue Receivables over 180 Days?
- Whether sales are being affected in excess of the Customer's Credit Limit?
- Whether excessive credit notes are being offered to specific customers towards returns, rebates, incentives.

CAATs –IDEA provide us instances to identify where controls are failing because of actual data errors. These factual cases serve well in building up a strong defense and back-up for the Auditor to expresso pinions on the processes.

IDEA can be applied to Accounts Receivable processes from an Internal Audit, Statutory Audit and Forensic Review perspective. This chapter shows how IDEA can be applied to Accounts Receivable from a Sarbanes Oxley perspective.

Most of the tests within IDEA can be automated and scheduled, which is the heart of Continuous Control Monitoring and Pre-Audit of Accounts Receivables and Collections Monitoring.

14.2.1 Potential Risks

The following table identifies the key risks, explains the business and audit implication of each risk and the audit objectives that could be realized by audit tests.

S No	Risk		Implications	Audit Objectives
1	The file	is	Items could be omitted, or the	Completeness,
	incorrectly		listed items may not be	Accuracy
	consolidated	or	included in the totals reported	

S	Risk	Implications	Audit
No	summed.	in the financial statements. The Accounts Receivable could be overstated or understated depending on the direction of error.	Objectives
2	Foreign currency transactions are not translated correctly.	Management may not be aware of the impact of transactions in foreign currency and may fail to take steps to manage currency risks. The Accounts Receivable could be overstated or understated depending on the direction of error.	Accuracy
3	Credit is granted to customers that are likely to default.	The business will sell goods to parties from which they will not be able to collect cash. This has potential implications for liquidity and bad debt expenses.	Valuation
4	Customers are double billed.	Double billing can negatively affect customer satisfaction. Also, revenues and receivables would be overstated.	Existence, Validity, Valuation
5	Accounts Receivable are invalid or incorrectly stated.	Accounts could be entirely or partly invalid. Partially invalid accounts may be the result of delays in processing transactions or errors in applying credits and payments to accounts. Fictitious accounts could be due to fraud.	Existence, Validity, Accuracy
6	Improper allocation of credits and payments	As mentioned earlier, improper allocation of payments to accounts could affect the aging of the Accounts Receivable and this would affect management's ability to determine an effective	Existence, Validity, Accuracy

Payroll and IDEA

S No	Risk	Implications	Audit Objectives
		course of action for handling the customer's account. For example, sales to customers could be blocked or the customer may be sent to a collection agency; (even though the customer is current on their accounts Does not make sense. Improper allocation of payments to accounts could also be an indicator of fraud.	
7	Accounts Receivable is not properly aged.	If the aging of the Accounts Receivable is not correct, then management may fail to take action on overdue accounts in a timely manner and expose sales to poor credit risks. Also, the calculation of the allowance for doubtful accounts and bad debts expense would be affected.	Valuation
8	A significant percentage of the receiva-bles is concentrated in a few customers.	The business could be exposed to a combination of credit and liquidity risks if these large customers do not pay their debts in time. Also, the company may be deemed to be economically dependent on the identified customers, and this may need to be noted in the financial statements.	Presentation
9	Improper classification of amounts.	If a credit balance is classified as an AR instead of an AP, then it could distort the current ratio which could be part of a debt covenant.	Presentation

14.2.2 Potential Tests

The following audit tests are suggested for auditing an Accounts Receivable system. However, the exact tests to be carried out for a particular client will depend upon the system used and the data available. Common tests include:

Mechanical Accuracy and Valuation

- 1. Total the file. It often pays to separate debits and credits.
- 2. Revalue foreign debts, if applicable.
- 3. Check transaction totals to the balance on each account.

Analysis

- 1. Profile debtors using **Stratification** to see the number of large debts and what proportion of value is in the larger items.
- 2. Produce an aged debt analysis. Consider how to deal with unallocated cash and credit notes. IDEA, by default, ages these on their date rather than allocating against the oldest item or any other treatment. It is often worthwhile to split the file into invoices, unallocated cash, and so on using multiple extractions, and then to age the individual files.

Exception Tests:- Existence and Valuation

- 1. Identify old items (i.e., more than three months old).
- 2. Identify large balances either in their own right or compared to turnover.
- 3. Select accounts for which no movements have been recorded in a set time.
- 4. Report credit balances.
- 5. Identify unmatched cash or credits.
- 6. Compare balances with credit limits and report exceptions (i.e., accounts with balances in excess of their credit limits or accounts with no credit limits etc.).
- 7. Test for items with invoice dates or numbers outside the expected range.
- 8. Identify partial payments of debts.
- 9. Identify invalid transaction types.

10. Identify customer addresses that are "care of" or flagged not to be sent out.

Gaps and Duplicates

- 1. Test for duplicate invoices (both invoice number and customer/value).
- Use duplicate exception testing for less obvious input errors, such as the same customer ID assigned to two different customer names, the same customer name assigned to two different customer IDs, receipt against the same invoice number and amount to two different customer, and so on.

Matching and Comparison Tests

- 1. Compare the balance on an account with its turnover.
- 2. Match the sales transactions to the customer master information to identify sales to new or unauthorized customers and those with exceeded credit limits.
- 3. Compare Accounts Receivable to Accounts Payable for possible contra accounts.

Sampling

Select samples (random or specific) for functional testing and confirmation (and produce confirmation letters).

14.2.3 Accounts Receivable Fraud Tests

Teeming and Lading

In Teeming and Lading, (especially in Cash Collection centric Business or Dunning Centres), the first lot of collections are defalcated by the perpetrator. To avoid clear obvious notice of overdue, cash collections from the subsequent customer are applied to the first customer. In this systematic scheme of events, old invoices are always shown/made up as clear/settled/paid and new invoices for latest customers are always shown as overdue, even though they have been paid.

In such cases, IDEA can be used to extract current bills for current sales and current customers based on Stratified Random Sampling (Materiality of the Sale). These customers can be called by the Auditor (telephonically) to ascertain the balance in his books for Debtor Confirmation. Any foul-play can

be identified at once through such a scheme of events unless there is a systematic collusion between the Vendor and Customer.

Forced Fictitious Sales to meet Sales Targets

We all have encountered enormous flurry of activity in the Sales Cell of an Entity towards each month end, quarter end, and year end. It is a known fact that Sales are pushed over phone to secure orders. These book orders are eventually accompanied by stock movements within a few days, hence the order is accrued at the period end and the targets are met.

However, we need to take care to identify unusual sales patterns at period ends which are counter-balanced by reversals in ensuing period beginnings. These unusual trends can be identified through the Field Statistics – Date Statistics in India. High number of sales transactions / records which almost resemble one another between say March 2015 and April 2015 need to be investigated, especially when the transactions in April 2015 pertain to reversals of March 2015 sales.

Transactions at or Near Selling Limits

Selling limits are often referred to as the "trigger price". As soon as a pre-set limit is reached, the system automatically triggers some form of action or examination.

The perpetrator's goal is simply to avoid the threshold approval price – trigger price. For example, it may be known to almost everybody that the spending authority limit is Rs. 50,000. The fraud is designed to stay under the threshold approval price.

This can be detected by employing Benford's Law of Digital Analysis, which detects instances where the sale price is just below the trigger price. The Benford Curve will reveal sharp, striking indications of actual frequency occurrences in excess of expected and upper tolerance frequency occurrences.

Another common scenario involves receipts for large sales over several invoices. To close a sale, the customer may accommodate this request. Look for transactions with the same customer that are close to the trigger price. Comparing inventory records is also an effective way to detect this. If three invoices are used to account for the full amount, an inventory check will reveal only one set of goods.

14.2.4 Accounts Receivable: Sarbanes Oxley Testing

Sarbanes-Oxley	Requirement	IDEA Test
Section	Requirement	וטבא ופטנ
Section 302: Corporate Responsibility for Financial Reports	Review of Financial Transactions of Significant Accounts	Use IDEA to summarize all detailed financial transactions for Accounts Receivable and agree to General Ledger
	Financial Transaction Analysis	Use IDEA to re-age Accounts Receivable, use the "Field Statistics" feature to identify unusual financial transactions.
Section 404: Management Assessment of Internal Controls	Matching Controls	Use IDEA in Accounts Receivable to verify the matching of Sales Invoices / Orders and Stock Movements.
	Cutoff Controls	Use IDEA to test Sales Booking by extracting all Debtors created before and after cutoff dates to verify inclusions and exclusion in General Ledger for the accounting period.
	Limit Controls	Use IDEA to extract all receipts (sales, customer payments, incentives, Selling Agent commission checks, Annual Target Bonuses) above threshold limits for dual approval

Sarbanes-Oxley Section	Requirement	IDEA Test
		authorizations.
Section 806: Whistleblower Protection	Protection of Employees who provide evidence of fraud.	Use IDEA to select a sample of phone calls and incidents from Hotline database and test the effectiveness of the Anti-Fraud Hotline program.

14.3 Inventory and IDEA

IDEA and the Inventory Process

14.3.0 Introduction

It is vital to establish that controls in Inventory processes are in place and working. To see, for example, whether the client is stocking material in excess of the maximum stocking limit, whether obsolete inventory is being monitored and provisioned, and whether inventory is being valued at the correct price.

CAATs IDEA identify instances where controls are failing because of actual data errors. . These factual cases serve well to build a strong defense and back-up for the auditor to express opinions on the processes.

IDEA can be applied to Inventory processes from an Internal Audit, Statutory Audit and Forensic Review perspective. In this chapter we show how IDEA can be applied to Accounts Payable from a Sarbanes Oxley perspective.

Most of the tests within IDEA can be automated and scheduled, which is the heart of Continuous Control Monitoring and Pre-Audit of Inventory.

14.3.1 Potential Risks

The following table identifies the key risks, explains the business and audit implication of each risk and the audit objectives that could be addressed by audit tests.

Payroll and IDEA

S No	Risk	Implications	Audit Objectives
1	Inventory is not correctly recorded.	Management will not have accurate information to manage inventory (e.g., ordering) effectively. Inventory could be materially misstated on the financial statements	Accuracy, Existence, Validity
2	Inventory management reports have inadequate supporting information.	Management may identify issues, but may not be able to "drill down" and identify the root cause of issues. This, in turn, hampers management's ability to make effective decisions.	Existence, Validity, Valuation
3	Quantity of inventory is not maintained within the specified range.	If too much inventory is maintained, then there is an increased risk of obsolescence and additional storage costs (i.e., for keeping the inventory). If too little inventory is maintained, then the company will not be able to meet the customer demand.	Valuation
4	Obsolete inventory items are not identified.	Management will not be able to make proper decisions on getting rid of obsolete items. From an audit point of view, obsolete items may need to be re-valued if the market price is less than their cost.	Valuation
5	Differences between the physical inventory and the inventory on the system are	Management will not have reliable information to make inventory management decisions. Also, it may have a problem identifying shrinkages. From an audit perspective, the	Existence, Validity

S No	Risk	Implications	Audit Objectives
	not identified.	inventory could be materially misstated on the financial statements.	
6	Items are not recorded in the correct period.		Cut-off
7	Gaps in sequentially numbered documents are not accounted for.		Completeness

14.3.2 Potential Tests

The following tests are to be undertaken for analyzing an inventory system. However, the exact tests carried out for a particular client will depend upon the system used and the data available.

Common tests include:

Mechanical Accuracy and Valuation

- 1. Total the file, providing sub-totals of the categories of inventory.
- 2. Re-perform any calculations involved in arriving at the final stock quantities and values.
- 3. Re-perform material and labor cost calculations on assembled items.

Analysis

- 1. Age inventory by date of receipt.
- 2. Compute the number of months each inventory item is held, based on either sales or purchases. Produce a summary of this information.
- 3. Stratify balances by value bands.
- 4. Analyze gross profit.
- 5. Analyze price adjustment transactions.

Exception Tests: Existence and Valuation

- 1. Identify and total inventory held in excess of maximum and minimum inventory levels.
- 2. Identify and total obsolete or damaged inventory (identified as such in the database).
- 3. Identify balances in excess of a reasonable usage period that are probably obsolete.
- 4. Identify items past their shelf life (if a sell by date or bought date is present on the system).
- 5. Identify any items with excessive or negligible selling or cost prices.
- 6. Identify differences arising from physical stock counts.
- 7. Test for movements with dates or reference numbers not in the correct period (cut-off).
- 8. Identify balances that include unusual items (i.e., manual adjustments).
- 9. Identify work in progress that has been open for an unreasonable period.
- 10. Identify inventory acquired from group companies.

Gaps and Duplicates

- 1. Test for missing inventory ticket numbers.
- 2. Test for missing transaction numbers.
- 3. Identify duplicate inventory items.

Matching and Comparison Tests

- 1. Compare files at two dates to identify new or deleted inventory lines or significant fluctuations in cost or selling price.
- 2. Compare cost and selling price and identify items where cost exceeds net realizable value.
- 3. Compare holdings and inventory turnover per product between stores.

14.3.3 Inventory Fraud Tests

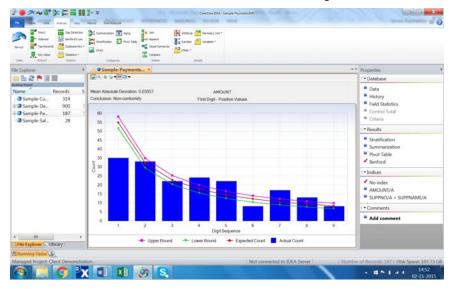
Fraudulent inventory valuations

In certain industries, Finished Goods inventory can be divided into two main classes: For Sale and For Samples / Gifts / Testing.

The inventory held for sale is correctly valued at net realizable value. Fraudsters often inflate inventory valuations for inventory held for Samples / Gifts / Testing.

As a common practice inventory held for Samples / Gifts / Testing, are valued at minimum system recognizable value like INR 1.00 to extend a notional value to the inventory in the accounting system.

The antithesis of the Benford Law may be applied here to test unexpected frequency counts at Digits other than 1 for inventory held for Samples / Gifts / Testing. If the entity is deliberately overstating, such inventory will show up on the Benford Curve for further examination and interrogation.



14.3.4 Inventory: Sarbanes Oxley Testing

Sarbanes-Oxley Section		Requirement		IDEA Test		
Section	302:	Review of	Financial	Use	IDEA to s	ummarize
Corporate		Transactions	of	all	detailed	financial
Responsibility	for	Significant Accounts		transactions for Inventory		

Payroll and IDEA

Financial Donarta		and agree to Conoral
Financial Reports		and agree to General
	Financial Transaction Analysis	Ledger Use IDEA to re-age Inventory, use the "Field Statistics" feature to identify and drill-down to unusual financial transactions.
Section 404: Management Assessment of Internal Controls	Matching Controls	Use IDEA in Inventory to verify the matching of Purchase Orders, Invoices and Receivers.
	Cut off Controls	Use IDEA to test inventory booking by extracting all receipt transactions before and after cut off dates to verify inclusions and exclusion in General Ledger for the accounting period.
	Limit Controls	Use IDEA to extract all inventory holdings above inventory control limits for dual approval authorizations.
Section 806: Whistleblower Protection	Protection of Employees who provide evidence of Fraud.	Use IDEA to select a sample of phone calls and incidents from Hotline database to ascertain the effectiveness of the Anti- Fraud Hotline program.

14.4. Payroll and IDEA

IDEA and the Payroll Process

14.4.0 Introduction

It is vital to establish that controls in Payroll processes are in place and working. For example, to find whether an employee is entitled to a specific location allowance, whether the head-count analysis for a month-on-month basis matches, and whether different employees have the same bank account number for salary credit.

CAATs –IDEA help us identify instances where controls are failing because of data errors. These factual cases serve well in building a strong defense and back-up for the auditor to express opinions on the processes.

IDEA can be applied to Payroll processes from an Internal Audit, and Forensic Review perspective. In this chapter we show how IDEA can be applied to Payroll from a Sarbanes Oxley perspective.

Most of the tests within IDEA can be automated and scheduled, which is the heart of Continuous Control Monitoring and Pre-Audit of Payroll. Cyclical audits for the same division payroll or for the same company across multiple locations can be automated.

Payroll tests essentially focus on ELIGIBILITY, ACCURACY, ADMISSIBILITY and CONSISTENCY of payment of allowances and credit of deductions.

14.4.1 Payroll – Control Risk Perspective

Payroll auditing is an excellent application of IDEA. The main objective is validity and accuracy by testing the existence of employees and their correct pay. There are many regulations and taxes associated with payroll and compliance with these can be checked. Privacy concerns may limit our testing though.

Analysis

- Summarize and stratify salaries by department/grade, etc
- Profile employee ages/years of service to assist in forward planning
- Analyze costs for special pay, overtime, premiums, etc

- Summarize payroll distribution for reconciliation to general ledger
- Summarize and compare costs for special pay, overtime, premium, etc

Calculations

- Total gross pay, net pay, deductions and any other value fields
- Check calculation of gross pay
- Check calculation of net pay

Exception Tests

• Extract all payroll checks where the gross amount exceeds a stipulated amount

Reasonableness of:

- Tax rates
- Pay/grade comparison
- Hours worked
- Overtime claimed
- Leave consumed
- Date of birth (under 18, over 60 years of age)
- Identify bonuses and other allowances
- Report activity on records for new or terminated employees
- Find changes in key payroll data, such as gross pay, hourly rates, salary amounts, exemptions, etc
- Identify records with missing information (National Insurance number/Social Security number, tax code, employee number etc)

Gaps and Duplicates

- Duplicate employees (Social Insurance, National Insurance, Social Security numbers, Employee numbers, addresses) on payroll file
- Duplicate bank account details
- Duplicate names and dates of birth

Matching and Comparing

- Comparison of payroll file at two dates to determine recorded starters and leavers, (hires and terminations) and changes in pay, etc, are as expected
- Join attendance file, to payroll master to determine if there are "ghost" employees on the payroll
- Compare time-card entries and pay to payroll and indicate variances
- Compare vendor addresses/phone numbers and employee addresses/phone numbers to identify conflict of interests (e.g. postcodes, phone numbers)

Sampling

• Most sampling options apply

14.4.2 Payroll Fraud Tests

A representation of typical fraud tests with regard to Payroll is as follows:

- Employees having the same First Name and Last Name
- Employees having the same First Name, Last Name and Bank Account Number.
- Employees having a different First Name, Last Name but having the same Bank Account for Salary Credit
- Employees having similar sounding Names (De-Dup Tests using Soundex Functions)
- Payments to Employees after they have left the Entity
- Payments to Employees who are not on the Employees Master List
- Overtime payments to Employees when normal hours have been worked.
- Payments of Location Allowances to Employees who are not entitled to them.
- Payment of Grade Allowances to Employees whose Grade does not permit it .

• Payment of both Asset Maintenance Expenses and Asset Maintenance Fixed Allowance like Vehicle Allowances when the Entity Policy allows only one.

Sarbanes-Oxley Section	Requirement	IDEA Test
Section 302: Corporate Responsibility for Financial Reports	Review of Financial Transactions of Significant Accounts	Use IDEA to summarize all detailed financial transactions for Payroll Heads of Allowances and Deductions from the Pay Ledgers and agree to General Ledger
	Financial Transaction Analysis	Use IDEA to analyse Pay Ledgers by Age, Salary, and Demographics for cases of attrition. Use the "Field Statistics" feature to identify and drill-down to unusual payroll indicators.
Section 404: Management Assessment of Internal Controls	Matching Controls	Use IDEA in Payroll to match Pay Ledger with Payroll masters to test for unauthorised payments.
	Cutoff Controls	Use IDEA to test for unentitled arrear payments or leave without pay returns to employees out of due periods for payment.
	Limit Controls	Use IDEA to test for Net Pay out of Salary Ranges in line with Employee Grades.

14.4.3 Payroll : Sarbanes Oxley Testing

Sarbanes-Oxle Section	у	Requirement	IDEA Test
Section Whistleblower Protection	806:	Protection o Employees who provide evidence o Fraud.	sample of phone calls

14.5 Computer IT Security and IDEA

IDEA and the Computer IT Security Process

14.5.0 Introduction

It is vital to establish that controls in the IT/IS processes are in place and working. For example, to find out whether out of office hour internet access is possible, prohibited web-site internet access is possible, data back-ups are taken as per the scheduled frequency, virus updates are received regularly, and scans are conducted periodically.

CAATs –IDEA identifies instances where controls are failing because of data errors. . These factual cases serve well in building a strong defense and back-up for the auditor to express opinions on the processes.

IDEA can be applied to the Computer IT Security process from a Systems Audit perspective of testing IT General and Environmental Controls at any location of the entity.

Most of the tests within IDEA can be automated and scheduled, which is the heart of Continuous Control Monitoring and Pre-Audit of IT Security.

The main tenet of tests within IDEA for IT Security is availability, confidentiality, compliance and integrity of IT processes. IDEA can be used as a vital facilitator for detecting non-compliance with the Information Technology Act in India.

14.5.1 Potential Tests

System Logs

When auditing system logs, we should check:

- List:
 - Accesses outside standard office hours or during holiday/sick leave
 - All users with their normal computers
 - All computers with their normal users
 - Users on unusual computers
- Identify users, particularly those with supervisory rights, who are logged in for a long period of time.
- Analyze by user identify those with higher use than might reasonably be expected.
- Summarize by network address to identify.
- Summarize charges by user to determine resource utilization.
- Analyze utilization by period, such as daily, weekly, and monthly, to show historical trends.

File Lists

When performing auditing tests in regards to computer security, we may :

- List duplicate names (both software for multiple copies and data where there is a risk of accidental deletion).
- Identify old files.
- Analyze by directory.
- Analyze file size by owner.
- Identify last access dates for old files.
- Analyze file type (by file name extension).
- Identify all files without an owner, such as where user accounts have been removed from the system.
- Test for .com, .exe or .bat files in areas where there should not be programs. DOS/Windows systems.

Access Rights

In regards to access rights, we must check:

• Lists of:

- Accounts with passwords not set or not required for access
- Group memberships
- Accounts with:
 - Short access passwords (less than the recommended six characters)
 - No activity in the last six months
 - o Access to key directories
 - o Supervisor status
 - Equivalence to users with high level access, such as supervisory equivalence
- Aging of password changes

E-mail Logs

E-mail logs generally contain information such as the sender and recipient address, subject title, date and time of transmission, size of file, service provider etc. Ensure the organization has a published policy related to employee use of e-mail before undertaking any of these tests.

Common tests include:

- Total length of time spent on e-mails (receiving and responding) by organization as a whole, by individuals, by month
- Analysis of internal and external e-mails
- Summarizing by service providers
- Summarizing number of e-mails by an employee and sorting to capture top users.
- Isolating, summarizing and examining personal e-mail usage
- Stratifying time and examining any unusual activity e.g. lunchtime, weekends, bank holidays
- Stratifying by size of files
- Analyzing file attachments, by size, by type
- Analyzing incoming e-mails, identify common domain addresses

- Calculating length of time individuals spent on e-mail in a given time period, sort in order
- Matching with the list of employees and extract any e-mails that are sent by invalid employees or terminated employees
- Analyzing any dormant accounts
- Identifying non-work related e-mails by searching for specific words in the subject title e.g. weekend Auditing E-mail Logs

Firewalls

Many organizations implement a range of controls, including installing sophisticated firewalls to eliminate the risk of unauthorized access to their networks, especially via the Internet.

Firewall logs record all incoming and outgoing transmissions on a network and it is not unusual to record hundreds of thousands of activities in a single day. IDEA can be used to analyze the logs, identifying trends and exceptional items to follow up.

Firewalls generally contain information like the source and destination IP address, date and time of admission, action by the firewall on receipt of transmission, the service type and the service port accessed.

Common tests include:

- Summarizing the type of service being requested or being used
- Identifying the most common IP addresses attempting access to the network
- Summarizing actions upon connection, i.e. control, accept or drop
- Analyzing trends to determine the most common access times and identifying requests at unusual times
- Extracting all dropped transmissions
- Identifying potential attacks by looking for a pre-defined sequence of port scans e.g. SATAN, ISS attacks or searches for ports which can be used.

Tally Accounting System and IDEA

IDEA and the Accounts Payables Process

15.0 Introduction

IDEA, like every other CAAT, can easily connect to an open data source. In fact, the precondition for the use of CAATs like IDEA is that there is an open unencrypted data source available for interaction.

In Tally Accounting System the data source is housed in the Tally Vault and hence protected. Direct interaction between IDEA and Tally Data Tables has not been established yet. In the given situation, IDEA interacts with reports from Tally. This is an established and well-recognized approach.

Here Account Ledger views (Display) are generated through the Balance Sheet and / or Profit and Loss Account option in Tally. These views are converted into a columnar representation, similar to an analytical Day Book view. Instead of printing out these columnar views, the view is exported to a File Format through the F12 command (for instance). The view can be saved as an ASCII, TXT, CSV, PRN, XML data format based on the need of the user. These formats can then in turn be imported through the relevant option in IDEA's import assistant.

In the case study to follow, we have imported an ASCII data file format.

15.1 Import of a Tally 9 ERP Data File

"The auditor entered the Expense Group of the Profit and Loss Account within Tally 9 for his client. He set the account period using the F2 option within Tally and generated a display of direct and indirect expenses. He then entered into direct expenses and generated a display of payments made towards manufacturing for the period 01.04.14 to 31.03.15. The expenses were displayed in a columnar ledger view with specific flags enabled for narrations and daily balances. Subsequently, the "Print to File" command was employed to capture the ledger display as an ASCII/SDF file. Steps used to import the Tally Data are

- 15.1.1 Click the Desktop icon under Home in the IDEA menu tool-bar
- 15.1.2 Select the Text option from the Import Assistant.

	A Const. III Const.	
e Explorer • • Final Control of the second	Select the format: A Advances Recear Dehnition Editor A Strong A Marces Record Dehnition Editor A Marces Record Dehnition Editor Perform Record and Adde PRF Sub Start Record Bandwares Sub Start Record Bandwares	Properties Catabase Catabase Catabase Catabase Catabase
III		

15.1.3 Click the browse button by the side of File name and select the file from the Source Files folder location on the computer.

	a least lie		X	
	The APro	jects + Client Demonstration + Source Files.ILB +	• 4 Search Source File_ P	-
ixplorer *	Organize • New fo	older	11 ·] 0	
ne Records S Sample-Cu 314	 Favorites Desktop 	Documents library Source Files ILI	Arrange by: Folder •	
Sample-De 900 : Sample-Pa 187 : Sample-Sal 28	Recent Places OneDrive Downloads	Name Mill Payroll Analytics base Latabiss QII: Payroll Augustubs	Date modified 18-11-2014 19:34 20-07-2012 11:04	distics Total
	C Libraries	Payroll July.slsx Payroll September.slsx	09-08-2012 09:38 09-08-2012 09:39	
	3 Documents	Gill SAP Procurement Base Data.xlsx	17-09-2014 10:38	-
	💰 Music	Ser Tax Credit Availment Dataxisx	09-02-2015 17:09	yts
	- Pictures	Travel Analytics Base Dataxisx	14-11-2014 15:16	
	Videos	Vendor Master Changes Base Dataxisx	09-12-2014 17:23	
		Webaccess.pm	19-08-2015 06:11	
	•& Homegroup	🖻 sdffixedwidth.sdf	19-09-2005 09:50 -	
		• •		
	File	name: sdffixedwidth.sdf	 All Files (".") 	
			Open 👻 Cancel	

15.1.4 Click Open and then Next.

15.1.5 IDEA scans the tally file and recognizes it to be a Fixed Length ASCII file.

(SDF files are "Fixed length ASCII" files)

Angel Costo Sant Anton Paperton Costo Page Costo Sant Anton Paperton Costo Page Costo Sant Anton Paperton Costo	Paper Randorf - The Tage	
File Explorer *	The Import Assistant has examined the file and has determined that it is a fixed length file. If this is not correct, please select the correct file type.	Properties • • Database
Norme Records S 3 Semple Ca. 314 3 Semple Ca. 900 t 3 Semple Pa. 187 t 3 Semple Sal. 28	Control of the second sec	Data Hotoy Hotoy Field Statistics Control Total Control Total Control tets t Comments Comments
• III • •	< Box Not> Carot Help	

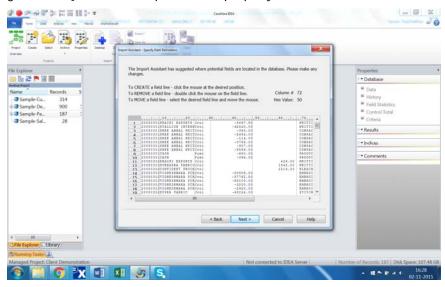
15.1.6 IDEA scans the record length of the source file and, by default, determines the file record length. In the case under demonstration it is 106.

Popel Create Select Addres Properties Cealing Concess Property Select Addres Properties Cealing Property Select Se	Depert Acrises - Specify Record Lergen	
File Explorer *	Please verify that the Import Assistant has determined the correct record length for the file. If the length is incorrect, you can adjust it by modifying the value contained in the record length edit box.	Properties •
Name Records 5 Sample-Ca. 314 Sample-Ca. 900 1 Sample-Pa. 187 Sample-Sal. 28	Record length - length of each record in the database	Das Hotoy Fried Stantso Control Control Control * Facults * Facults * Sodice: * Comments
· III · ·	< Back Next > Cancel Help	

15.1.7 Click Next

15.1.8 IDEA automatically places Field Separator Lines at the points where each logical field ends and the next begins. This at times

requires to be changed to ensure that all logical units of data (Fields) get correctly selected/ imported and properly formatted.



15.1.9 Insert field separators at the end of each field by clicking the cursor in the data content screen once at the required point.

To remove an unwanted field, click the mouse point twice over the unwanted line.

To move a field line place the mouse point on the select line, keeping the left button of the mouse pressed, and drag the line to the point of choice.

15.1.10 After giving the field line breaks, click Next.

15.1.11 Now enter the Field Names.

15.1.12 We have an option of not importing any field. Click Do not import this field in the check box.

						Properties
						II 🕶 Database
Import As	sistant - Field D	etails				Dista .
1.11			_			History.
You can	now specify field d	etails. Select a fir	eld by clici	king on the column heading below.	and then	Pield Statistics
modify it:	s information below.				Sector 1	Control Total
Field rise	THE CHAPLE			Chalacter		🔛 Griteria
Description						II 🖛 Results
						II - Indices
	sriverted Example					-
						Comments
	DATE	PAYEE	TYPE	DEBIT	CREDI	
1				-3697.50		
- 2				-42820,00		
4				-2454.00		
7	20050301 SPE	EE AMBAL KNT	Jinl	-907.00		
10			Pync	-366.00	×	
<					×	
		-	. Bash		-	
			CDACK	Lancer	Lueb	
100					28	
	modity # Feld nor Descript Polon Ce 1 2 3 4 4 5 6 6 7 7 8 9 9 10	moddy is information below. Field name CPUHBE Operations (1), p(2), p(3) Image: Convented Example Do not import this field Image: Convented Example 0004000 Image: Convented Example 00040000 Image: Convented Example 00040000 <td>Dock of a reformation televic. Personal and the second /td> <td>moddly is reformation below. Figure 10005 Figure 10005 Toper- Demonstrate 10005 Toper- Demonstrate 10005 Convented Example Do not impact this field Toper- Demonstrate 10005 Toper- Demonstrate 10005 Toper- Demonstrate 10005 Do not impact this field Do not impact this field Toper- Demonstrate 10005 Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impac</td> <td>become Description Type Description Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His His His His His His His His His His</td> <td>No. can now specify field detail. States a field by clocking on the cahann heading below, and then: Image: State in the state in</td>	Dock of a reformation televic. Personal and the second	moddly is reformation below. Figure 10005 Figure 10005 Toper- Demonstrate 10005 Toper- Demonstrate 10005 Convented Example Do not impact this field Toper- Demonstrate 10005 Toper- Demonstrate 10005 Toper- Demonstrate 10005 Do not impact this field Do not impact this field Toper- Demonstrate 10005 Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impact this field Toper- Demonstrate 10005 Do not impact this field Do not impact this field Do not impac	become Description Type Description Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His Hold Construct His	No. can now specify field detail. States a field by clocking on the cahann heading below, and then: Image: State in the state in

15.1.13 Then click each field heading, enter the field name, type and description.

15.1.14 For example, for the first field the name is Date, type is Date and Date Mask is YYYYMMDD, which matches with the actual view of the date data in the data content section. For example 20050301 will necessitate as mask as YYYYMMDD

Field Name	Field Type	Remarks
Date	Date	Mask YYYYMMDD
Vendor	Character	
TXN TYPE	Character	
DEBIT	Numeric	
CREDIT	Numeric	
AC HEAD	Character	
BALANCE	Character	

15.1.15 Define all the fields as per the following File Layout

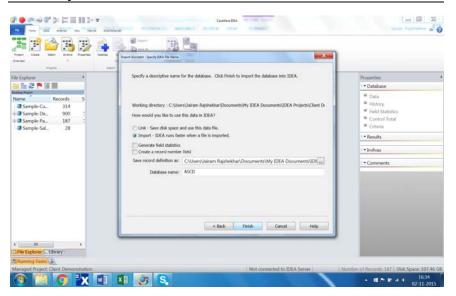
15.1.16 We may create additional fields over and above the existing fields at this stage.

But in this presentation we will not create any additional fields. Just Click Next.

Tally Accounting System and IDEA

Image: Da. 1000	nove Data Analysis new Marcia Dravid	Const which and the second states are		
Image: set of the set	s (2) 😂 (3) (3) (4)	Part 12 Ca		
Image: Construction	et Deale Select Ardrive Properties Dealtop	Singuri Assistent - Conde Falds		
Image: Construction	Page 19	You can add one or more fields to the immediat file to extended the field same to a least	E000575	_
Imple C. 314 Imple C. 314 <td></td> <td>parameter and description for the field(s) in the columns below.</td> <td></td> <td></td>		parameter and description for the field(s) in the columns below.		
Note	Manufacture and the second			
ht be 900 _ 1 ht be 900			History	
Image: market in the mark intermed into the intermed	ple-De_ 900 1	Delete	Field Statistics	
Image: Section of the section of th	de-Pa 187 :	Fuld Name Type Len Dec Parameter Description	Criteria	
	-30			
Image: Comments			1 * hading	
Contraction Not concreted to DEA Server Note on Receive Life To Dea Server 16-11 Contraction Contraction Contraction 16-11 16-11 Contraction Contraction Contraction 16-11			- Comments	
Contraction Not concerted to DEA Server And a				
Recerching the series of the s		< Back Next > Cancel Help		
Image: Control demonstration Not concreted to 200.45 server Image: Control demonstration				
de la contration par la contration par la contration la contrat				
And concerted to DDA Server Nucleon of Records - 127 Obt Space 2004 Image: Concerted to DDA Server Image: Conce				_
Image: Data Data Descention Descention Image: Data Data Data Data				
Image: Data in the image: Data in the loss of t	ed Project: Client Demonstration			1631
Image: Solution of the second seco	📋 🧿 🗽 💷	M 🛷 S.	• #►₽.4.4 ₀₂ .	16:31 11-2015
Image: Control of the section of th	📋 🧿 🗽 💷	Carther Bld	• #►₽.4.4 ₀₂ .	16:31 11-2015
Precords To can enter an equation to limit the data that is imported. Clok the Equation Editor botton to enter PRecords The equation pie Dca. The equation the equation The		Carther Bld	• #►₽.4.4 ₀₂ .	1631 11-2015
Be equified. Image: Constraint of the equified. Proceedings State Spin Pa_1187 State Spin Pa_1187 State Image: Sale Image: Sale Image: Sale Image: Sa			• #►₽.4.4 ₀₂ .	1631 11-2015
Be equation Image: Data State St			• #►₽.4.4 ₀₂ .	1631 11-2015
Records S Sple-Cu. 314 Vel-De. 900 1 sple-Pa. 187 2 sple-Pa. 187 2 sple-Pa. 28 sple-Pa. 28 sple-Pa. 187 2 sple-Pa. 28 sple-Pa. 28 sple-Pa. 187 2 sple-Pa. 187 2 sple-Pa. 28 sple-Pa. 28 sple-Pa. 187 2 sple-Pa. 28		Cariford Ed.	• #►₽.4.4 ₀₂ .	1631 11-2015
Records 5 0.04 3.04 F(Au. 33.4) 16000 16000 16000 F(Pa. 1327) 163 16000 16000		Currier DG	Properties	1631 11-2015
Du 900 1 # feld Statistics Pa_ 187 2 0 # control total Sal_ 28 0 # control total # control total * Control ** ** ** ** ** * Control ** <t< td=""><td></td><td>Curritor Ed.</td><td>Properties</td><td>1631 11-2015</td></t<>		Curritor Ed.	Properties	1631 11-2015
Pa_ 197 : Sal_ 28 Sal_ 28 (*Control Total * Chevits (*Results (*Indices (*Control Total * Results (*Indices (*Control Total * Control Total * Results (*Indices (*Control Total * Control Total		Curritor Ed.	Properties • Database • Data	1631 11-2015
book Net> Carel Hop		Curritor Ed.	Properties * Database * Database * Database * Database * Database	1631 11-2015
< Box Next > Cancol Help		Curritor Ed.	Properties • Database • Database • Cata • History • Field Statistics • Control Total	1631 11-2015
<pre></pre>	No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. Portor No. No. No. No. No. No. Portor No. No. No. No. No. No. Portor No. No. No. No. No. <	Curritor Ed.	Properties • Database • Database • Cata • History • Field Statistics • Control Total	1631 11-2015
< Reck Next > Cancel Help some Tubrary	Image: Constraint of the second sec	Curritor Ed.	Properties	1631 11-2015
m ,		Curritor Ed.	Properties • Database • Database • Data • Control Control • Contr	1631 11-2015
m ,		Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
n , , , , , , , , , , , , , , , , , , ,		Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
m ,		Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
m ,	Non- Non- <th< td=""><td>Curritor Ed.</td><td>Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total</td><td>1631 11-2015</td></th<>	Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
m	No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. No. Portor No. No. No. No. No. No. Portor No. No. No. No. No. No. Portor No. No. No. No. No. <	Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
m ,	Non- Non- <th< td=""><td>Curritor Ed.</td><td>Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total</td><td>1631 11-2015</td></th<>	Curritor Ed.	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
er (tbray)		Currier Eld	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
er (Dibrary	Image: Control of the contro	Currier Eld	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
	Records S Description Description <thdescription< th=""> <thdescription< th=""> <thdescriptio< td=""><td>Currier Eld</td><td>Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total</td><td>1631 11-2015</td></thdescriptio<></thdescription<></thdescription<>	Currier Eld	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015
Tester Not connected to IDFA Server Number of Records: 187 Disk Space: 107.4	Control C	Currier Eld	Properties • Database • Database • Database • Data • Database • Database • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total • Control Total	1631 11-2015

15.1.17 In the final 'Specify IDEA File Name' screen enter the file name ASCII and click Finish



15.1.18 Once imported, the file looks like what is shown below:

		and a second sec	
Presente Explorer #	Part Newt Citize	*	• * Properties
1 2 P 1 1	DATE PAYEE TYPE DEBIT C	REDIT EXPENSE	* Database
a based	1 GE-ES-2005 PRACHE EXPORTS Int -36IR	B INTTINU CHARGE	
me Records S	2 01-03-2015 STALLION INTERN And -42620	II KNETTING CHARGE	Data
	3 02-05-2005 SRIE AMRAI, KRET Avei -363 4 02-05-2005 SRIE AMRAI, KRET Ivei -3454	8 COMPACTING & DK 8 COMPACTING & DK	History
Sample-Cu 314	4 02-03-2005 SHEE AMBAL ENET June -104	0 COMPACTING & DR	
Sample-De_ 900 1	6 01-03-2005 SPEE AMBAL KINET June -1786	8 COMPACTING & DR	Field Statistics
Sample-Pa187 1	7 02-03-2005 SPEE AMBAL KHET And -007	0 COMPACTING & DR	Control Total
	8 05-03-2005 SPEE AMBAL KNET Just 3589	8 COMPACTING & DR	Criteria
Sample-Sal_ 28	9 01-03-2005 CASH Pyret 480 50 01-03-2005 CASH Pyret -386	8 PRODUCTION EXPE	Chierd
sdffixedwi567	10 02-03-005 CASH Pyert -306 11 02-03-005 PRACHEDPORTS Intil 0	43 FRODUCTION EXPE	I • Results
	12 02-05-2005 SUBHADRA KABROC 3ml 0	1842 IONITING CHARGE	TRANS.
	13 62-85-3089 CONFERENT PROCE and B	2214 BLEACHING & DHE	
	14 02-03 2005 VIGNESHWARA SCR And 80505	9 EMERCIEERY & PR	• • Indices
	15 02-05-3015 VIGNESHWARA SCR 3ml -57763	0 EMBROEKRY & PR	✓ No index
	16 02-03-205 V30HESHWARA SCR Anti -45153 17 02-03-205 V30HESHWARA SCR Anti -1220	EMBRODERY & PR EMBRODERY & PR	* No index
	17 02-03-2005 V05A6344MARA SCR 3ml -3200 18 02-03-2005 V05A6344MARA SCR 3ml -2005	E EMERCICIEN & PR	• • Comments
	10 02-03-2005 SUPER FABRIC JHI 43034	E STEDBIG OMAS	Comments
		8 COMPACTING & DR	Add comment
	25 03-03-2025 BANAREKA GARMEN Juli -20	8 COMPACTING & DR	Add comment
	22 03-03-3985 SRELAKSHMETEK And -3890 23 03-03-3985 PEACHEEXPORTS and -362	8 XMETTING CHARGE 8 XMETTING CHARGE	
	23 03-03-305 PRACHE DIPORTS Int -902 74 03-03-305 KADAARAR GARME Int -3043	8 COMPACTING & DR	
	20 61-61-205 SUAMERA CAMMEN Her HE 21 61-61-205 SUAMERA CAMMEN He HE 22 61-61-205 SHE LASSMER CAMMEN He HE 23 61-61-205 SHE LASSMER TEX He 289 23 61-61-205 SHE LASSMER TEX He 289 24 61-61-205 KADAMENA GAMENA He 5601 25 10-61-205 KADAMENA GAMENA SHE 3601	0 COMPACTING & DR	
	26 03-03-2005 KADAAWAR GAMME Int 0001	8 COMPACTING & DR	
	27 05-05-2005 KAGAMPAREGARME Invit -404	0 COMPACTING & DR	
	28 43-43-3005 KADAMPAREGARME Inst -368	B COMPACTING IN DR	
	29 CF-CF-3305 KADAAMAD GARME Int UB 30 CF-CF-3955 KADAAMAD GARME Int -2245	8 COMPACTING & DR	
	30 01-02-2005 KADAARAAT GARAKE Juni - 2203	E COMPACTING & DE	
	32 01-03-305 CASH Pyret 585	PRODUCTION EXPE	
5.02.1	3.3 04-03-3005 COLLARS N FAMEL Intel -15014	8 KDATTENG CHARGE	
	34 04-03-2005 COLLARS N PANEL And -1980.7	6 KNETTING CHARGE	
le Explorer Ubrary	34 64-65-3005 COLLARS N PANEL and -18823 35 64-65-3005 SAIG FABRIDIS and -58/26 36 64-65-3005 SAIG FABRIDIS and -58/26 36 64-65-3005 SAIG FABRIDIS and -58/26	6 KNETTING CHARGE	
	36 04-03-2005 SANT FABRICS and -6267	0 KHETTENS CHURGE	
Lunning Tasks			

15.2 Analysis of Tally 9 ERP Data in IDEA

15.2.1 Extract Cash Payments

Go to Analysis - Extract - Direct

Tally Accounting System and IDEA

A Constant C	Crathuser D PostTate	Append 1	Ambula 🦉 Hunday Lat * Kanton 🏄 Kanaton * Offer *		
Exple B" sweet #	• sdffixedwidth.IMD			- 1	Properties
En containing the records that	DATE PAYEE	TYPE DEBIT	CREDIT EXPENSE		· • Database
And Annual Par others From the	1 RE-RE-2005 PRACHE ENFORTS	Jani -3638	B I INTTING CHARGE	(m)	
me Records S	2 01-03-3015 STALLION INTERN 3 01-03-3005 SPEE AAMEAL KNET	Jest -42520 Avai -363	KNETTING CHARGE COMMACTING & DR		 Data
Sample-Cu 314	4 02-02-2005 SHEE AMERAL KINET	3ml -3454	B COMPACTING & DR		History
	5 OL-00-2005 SPEE AAMEAL KNET	Jed -114	0 COMPACTING & DR		Field Statistics
Sample-De_ 900 1	E IT 43-2005 SHE AMERI KNET	Just -3796	8 COMPACTING & DR		
Sample-Pa_ 187 1	7 05-03-3005 SHEE AMBAL KHET B 05-03-3005 SHEE AMBAL KHET	Jed -907 Jed -5509	0 COMPACTING & DR		Control Total
Sample-Sal_ 28	0 01-05-2015 CASH	Port 480	PRODUCTION EXPE		Criteria
	10 (1-45-2015 CA5H	Pyers -366	8 PRODUCTION EXPL		
sdffixedwi 567	11 02-03-2005 PRACHE EXPORTS	Just 0	438 KNATTING CHAMIE		Results
	12 62-85-305 SUBHADKA FABRIC	Jed 0	1842 KNETTING CHARGE		
	13 62-65-3019 CONFERENT PROCE	Just 0	2214 BLEACHING & DVE 9 EMBROBIERY & PE		- Indices
	14 02-03-2005 VIONESHWARA SCR 15 02-03-2005 VIONESHWARA SCR	And 90509 Jest -57763	EMERCIDERY & PR EMERCIDERY & PR		
	16 02-05-2005 VIONESHWARA SCR	And -41223	0 EMERCIDERY & PR		No index
	17 02-03-2005 VIGNESHWARA SCR	Jeel -3200	B EMERCIDERY & PR		
	18 02-03-2005 V90AESHWARA SCR	3eei -2900	8 EMERCIDERY & PR		· Comments
	19 02-03-305 SUPER FABRIC 20 02-03-305 BARL/REXA GARMEN	Jest -63034 Line -81	E STITCHENG CHARG E COMMACTING & DR		
	20 05-05-2005 BANUREKA GARMEN 21 05-05-2005 BANUREKA GARMEN		8 COMPACTING & DR		Add comment
	22 01-03-3915 SPELAKSHMETER	Ard -0800	8 KNATTING CHAME		
	23 03-03-3985 PRACHED/PORTS	Jed -902	8 KNETTING CHARGE		
	24 03-03-3005 KADAMPARI GARME		8 COMPACTING & DR		
	25 03-03-2005 KADAMPART GARS 6		0 COMPACTING & DR		
	26 03-03-2005 KADAMPAR GARME 27 03-03-2005 KADAMPAR GARME		COMPACTING & DR COMPACTING & DR		
	27 05-05-2005 KADAMPAR GARME 28 05-05-2005 KADAMPAR GARME		B COMPACTING & DR		
	29 03-03-3015 KADAMPAR GARM		8 COMPACTING IS DR		
	30 05-03-2005 KADAARPARI GARAR	Just -2245	II COMPACTING & DR		
	31 00-00-2005 KADAAMARI GAAM		8 COMPACTING IS DR		
	32 00-00-2005 CASH	Pyred -585	#ROBUCTION EXPE		
	33 04-03-305 COLLARS N FAME 34 04-03-305 COLLARS N FAME	Jost -15014 Jani -13023	0 KDETTENG CHARGE 0 KOETTENG CHARGE		
	34 04-03-3005 COLLARS N PANEL 35 04-03-3005 SANI KABRICS	Jeni -59050	8 KNETTING CHARGE		
le Explorer Ubrary	36 04-03-2005 SANT FABRICS	3ml -6365	0 IDETTING CHARGE		
	and the second se				14

This will open the Direct Extraction dialog box.

Enter the File Name as Cash Payments

Click the CRITERIA button to the right of the File Name

This will open the Equation Editor.

In the Equation Editor, write the Equation

PAYEE="CASH". AND. TYPE="Pymt"

Enter the field names PAYEE and TYPE by entering the first letters of the field names in the Equation Box. The Field names appear automatically by clicking the first letter. Then select the same and press enter. Enter CASH and Pymt in " " since they are Character fields. Enter these words by hand. Enter .AND. by clicking the button .AND. in the Equation keypad

Part Gardenaria Parter Gardenaria Parter Gardenaria	Di tameram III aire Di trattate II fortat	a Res Append	Anthule 🚰 Havelary U					
Total State	Curl	The Conserver	00w *					
fotal fotal	Direct Longiture	Labora .	Sample	_			Properties	
· · · · · · · · · · · · · · · · · · ·		and the second se	-				· Databa	C#
and the state of the	1. Records to extra	act: 💿 All	Sarting record #:	1		OK	Data	~
Records 5	3	C Bange	Ending record #:	567		Greate Field	Histor	
ple-Cu 314 ple-De 900	4	Create a st	rtual database			Fields	Field S	y itatistics
ple-Pa	7 Database grder	No index				• Delete	· Contro	of Total
ple-Sal. 28	g g	NO BRAN	File Nome		Criteria	Cancel	P Criteria	
xedwi 567	10. 11. 1 Cash Payments				PAREL CASH AND THEY P	ent'	· Results	
	1 Records to extra 1 Records to extra 1 Castopics grader 2					Help	• Indices	
	14						Indices	
	16.						I = Comme	ints
	18.						= Add o	omment
	20							
	222							
	24							
	26							
	27 28			and a second		_		
	29 01-01-2015 KADALIPAR 30 05-05-3015 KADALIPAR	A CARACE INTE 12	US B COMPACTS DIG B COMPACTS	d a DE				
	21 0-05-2005 6ADAMPAR 22 03-05-2005 6ADH	N GARAE Intel -2	EUII 8 COMPACTS -MII 8 PRODUCTIO	IG B. DE				
	33 04-03-3325 COLLARS N 34 06-03-2325 COLLARS N	17. heat 123.	NDA B KNETTRAS C RDS B KNETTRAS C	4ARCS				
Aorer CUbrary	36 04-03-2005 SANSFABRE 36 04-03-2005 SANSFABRE	CS Intel -53	RISE E KNETTING O KNET E KNETTING O	GLRCE			10	
g Tasks	1_39_1 wanter i wanter	14 (14 (14 (14 (14 (14 (14 (14 (14 (14 (ALL ALL MAILER AND	when it				
Project Clerit Demonstra		🧭 S.	Contine) T	kot connected to IDEA Server		lumber of Records	16 02-11
Project: Clernt Demonstra Constraints of the second secon		-			kot connected to IDEA Server		17.75	• • • • 16 02-11
Project: Client Demonstra Control Client Demonstra Client Demonstra Control Client Demonstra	(W) X 	а с ы и о	0 P al 1 0		- () Al		- 117	16 02-11
Project: Clerit Demonstra Clerit Demonstra Clerit Clerit Demonstra Clerit D	C W X X	а с ы и о			- all		17.75	16 02-11
Project: Clerit Demonstration	(W) X 	а с ы и о	0 P al 1 0		All Character Numeric Date / Time	QI QI	Functions	16 02-11
Project: Cleart Demonstri Clear Cleart Demonstri Project: Cleart Demonstr	C W X X	а с ы и о	0 P al 1 0		Ali Ocurater Warnic Date / Time Micring		Functions	16 02-11
Popet Clert Demonstra Clerk Clerk Company Clerk Company C	C W X X	а с ы и о	0 P al 1 0		All Duracter Duracter Duracter Macring Conditional Phantal		Functions Functions are abable for forming more majes operations	16 02-11
Project: Cleart Demonstri Clear Cleart Demonstri Project: Cleart Demonstr	C W X X	а с ы и о	0 P al 1 0		Al Ourader Winnerk Ourader Winnerk Oude/Time Mudzing Conditional		Functions Functions are alable for forming more mplex operations th as date	16 02-11
Poper Cleve Demonstra	C W X X	а с ы и о	0 P al 1 0		All Duracter Duracter Duracter Macring Conditional Phantal	(B) (B) (B) (B) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C	Functions Functions are alable for dorming more mplex operations th as date timetic, statistical culations and test	16 02-11
Project Clerk Demonstra	C W X X	а с ы и о	0 P al 1 0		All Duracter Duracter Duracter Macring Conditional Phantal	G G S S S S S S S S S S S S S S S S S S	Functions Functions are allable for forming more majes operations ch as date timetic, statistical	16 02-11
Project Clerk Demonstra Project Demonstra Proje	C W X X	а с ы и о	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	G) G) any per co sum art ca sum the	Functions Functions are alable for forming more majers operations ch as date timetic, statistical culations and test rations begin with rid's symbol.	16 02-11
Project Clerk Demonstra	C W X X	а с ы и о	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	di avi pp co ss fur th Th Th sh	Functions Functions are alable for forming more index copy allows index. IDEA utilizes begin with utilizes begin with utilizes begin with utilizes begin with inter in style and	16 02-11
Project Clerk Demonstra Project Demonstra Pr	C W X X	а с ы и о	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	e d d av ar c se f se f se f se f se f se se f se se f se se f se se se se se se se se se se se se se	Functions Functions are adable for forming more majers operations ch as date timetic, statistical culations and test functions. IDEA timetic, statistical culations and test traffic operations of the statistical regression to water very mar newsy	16 02-11
Project Clerk Demonstra Project Demonstra Pr	C W X X	а с ы и о	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	e d d s s s s s s s s s s s s s s s s s	Functions Functions are aidele for draming marker spectra aide for any spectra draming and test archives. IDCA arditors begin with "9" spectra articles. IDCA articles. The spectra articles. The spec	16 02-11
Project Clerk Demonstra Clerk Demonstra	C W K K K K K K K K K K K K K K K K K K	а с ы и о	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	di ani pen son son son son son son son son son so	Functions Functions are aliable for forming methods methods index operations index operations index operations index operations index operations without by an every index operations index opera	16 02-11
Project Clerk Demonstra Clerk Demonstra		à ♥₩ ♥ 0 //AND 00	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	ari ari pre so so so so so so so so so so so so so	Functions Functions are alable for forming more fractions are alable for fractions and test testings how and test testings how and test without substantial without substantial testings fractional in testings fractional in	16 02-11
Project Clerk Demonstra Clerk Demonstra		à ♥₩ ♥ 0 //AND 00	0 P al 10 0		All Duracter Duracter Duracter Macring Conditional Phantal	e an an an an an an an an an an an an an	Functions addition are addition are addition for the second addition and the second additions additions and additions additions and additions additions and additions additions additions additions additions additions additions additions additions addition additions addition additions addition additions additions additions addition additions addition additions addition additions additi	16 02-11
Project Clark Demonstra Comparison Comparison Project Clark Demonstra Project Demonstra Projec	C W X X	Collecte	0 P al 10 0	CG	All Character Date / Time Maching Costional Financial Custom Functions	di ati po po po se fat the the the the the the the the the th	Functions are institutes are aduated for functions are aduated for functions are aduated for the state of the state the state of the state the state t	16 02-11
Project Clerk Demonstra Control Clerk Demonst	C V V X V X V V V V V V V V V V V V V V	Image: Section 1 Image: Section 2 Image: Section 2<	NOT MOD BP	CG	All Character Date / Time Maching Costional Financial Custom Functions	Contraction of the second seco	Functions Functions are alable for draining more inplex operations in drain and the contents and the resolutions of the second resolutions of the second resolutions of the resolutions	16 02-11
Project Clerk Demonstra Composition Project Clerk Demonstra Project	Image: Section 1 Image: Section 2 Image: Section 2 Image: Section 2 Imag	Svaluate Svaluate Svaluate Svaluate	NOT MOD BP	EL- 107-14-14-14-14 	All Character Date / Time Maching Costional Financial Custom Functions	Contraction of the second seco	Functions and additional addite additional additional additional additional additional a	16 02-11
Project Clerk Demonstra Composition Project Clerk Demonstra Project Demonstra	Curr JAC Tork Your Curr JAC Tork Your Curr JAC Tork Your	Created and Constant	NOT MOD BP	10 - 10 ⁻	All Character Date / Time Maching Costional Financial Custom Functions	Construction of the second sec	Functions Functions are alable for forming more regions and back and back taking back and back and back taking back and back and back and back and back taking back and back and back and back and back taking back and back and back and back and back and back taking back and back and back and back and back and back taking back and back and back and back and back and back taking back and back and back and back and back and back taking back and back taking back and	16 02-11
Project Clerk Demonstra	Curr JAC Tork Your Curr JAC Tork Your Curr JAC Tork Your	Contractor	NOT MOD BP	EL- 107-14-14-14-14 	All Character Date / Time Maching Costional Financial Custom Functions	G G G G G G G G G G G G G G G G G G G	Functions Functions are atable for forming more reprice particular to a date forming, more reprice particular to taking the standard that within the shared taking the shared	16 02-11
Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project D	Curr JAC Tork Your Curr JAC Tork Your Curr JAC Tork Your	Contractor Contractor	NOT MOD BP	10 - 10 ⁻	All Ourader Menetic Boods Coddgal Coddgal Coddgal Coddgal Coddgal Coddgal	Construction of the second sec	Functions are adapted for the second	16 02-11
Project Clerch Demonstration	Curr JAC Tork Your Curr JAC Tork Your Curr JAC Tork Your	Created and Constant	NOT MOD BP	10 - 10 ⁻	All Character Date / Time Maching Costional Financial Custom Functions	Construction of the second sec	Functions Functions are atable for forming more reprice particular to a date forming, more reprice particular to taking the standard that within the shared taking the shared	16 02-11
Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project Clerch Demonstra Project D	Curr and Tork Your Curr and Tork Your Curr and Tork Your	Contractor Contractor	NOT MOD BP	CD- 107	All Ourader Menetic Boods Coddgal Coddgal Coddgal Coddgal Coddgal Coddgal	Construction of the second sec	Functions are adapted for the second	16 02-11

Click the Green Check button to Validate and Exit.

Then Click OK in the Extract to file (s) dialog box.

We will get a list of 28 Cash Payments, as can be seen below, which will serve as the basis for further tests to follow.

Tally Accounting System and IDEA

Total	Crathuter D Post Tale	20 20 20 Australy (M**) 20 20 20 20 20 20 20 20 20 20 20 20	
e Explorer *	sdffixedwidth.IMD	Cash Payments.IMD *	- * Properties
1 2 P 3 III	DATE PAYEE	TYPE DEBIT CREDIT EXPENSE	• Database
Also Project / Control of Control	1 05-40-3005 CASH 7 05-40-3005 CASH	Pyret 480 8 PRODUCTION DOR	
ame Records S	2 0140-305 CA94	Pyet -34 0 PRODUCTION D/H Pyet -585 0 PRODUCTION D/H	✓ Data
Sample-Cu 314	4 09-40-3005 CASH	Pyret 400 8 PRODUCTION EXPE	History
Sample-De_ 900	5 11-03-2005 CADH	Pyret 4870 0 PRODUCTION DIPL	Field Statistics
	6 19-49-3905 CASH 7 12-49-3905 CASH	Pyret -3271 0 8LEACHING 8.015 Pyret -3240 0 PRODUCTION DOM	Control Total
Sample-Pa_ 187 1	8 15-03-3085 CASH	Pyret -250 8 CONE WINDING &	
Sample-Sal. 28	9 16-03-2005 CASH	Pyret -2013 8 PRODUCTION DVPE	Criteria
sdffixedwi567	10 29-83-2005 CASH 11 22-00-3005 CASH	Pyret -277 2 CONE WENDING & Pyret -396 0 PRODUCTION D/PE	- Results
Cash Pa. 28	11 22-00-3905 CASH 12 26-00-3905 CASH	Pyret 134 0 PRODUCTION DOPL	- Nesurs
20	13 11-40-2005 CASH	Pyret -2000 0 PRODUCTION DVPE	
	14 31-00-2005 C494	Pyret 5200 8 PRODUCTION DPE Puret 3008 8 PRODUCTION DPE	* Indices
	15 11-03-2005 CASH 16 31-03-3005 CASH	Pyme -3826 0 PRODUCTION DVR Pyme -3825 0 PRODUCTION DVR	No index
	17 11-40-3005 CASH	Pyret -725 8 PRODUCTION EXPE	- Ho more
	18 11-410-31619 CASH	Pyrei -338 8 PRODUCTION EXPE	· · Comments
	19 11-40-305 CASH 20 11-40-305 CASH	Pyret -1271 0 PRODUCTION D/R Pyret -171800 0 WAGES	
	20 18-40-3005 CASH 21 18-40-3005 CASH	Pyret -173800 8 WAGES Pyret -3364 8 CONE WEVERIG IS	Add comment
	22 11-40-2005 CASH	Pyret -EBER 8 APPROVINCE-ST	
	23 11-41-2015 CASH	Pyme -328700 0 WAGES	
	24 II-01-2005 CASH 255 II-01-2005 CASH	Pyret 3828 0 PRODUCTION D/PE Pyret 480 0 83,14CH046 & D18	
	25 10-40-2005 CASH 26 30-40-2005 CASH	Pyret 430 0 BJACHNG & Drift Puret 1590 8 SCHEIN PRIVIDING	
		Pyret -720 0 BLEACHING & DIE	
	27 11-43-3915 CASH 28 11-43-3915 CASH	Pyret -200 8 PRODUCTION EXPE	
III			
Running Tasks			
Contraction of the second second	tion		

15.2.2 Identify Cash Payments made on Sundays which is the Company Weekly Off

Go to Analysis - Extract - Direct

Enter the File Name as Sunday Payments in the Direct Extraction dialog box.

Click the Criteria button to open the Equation Editor.

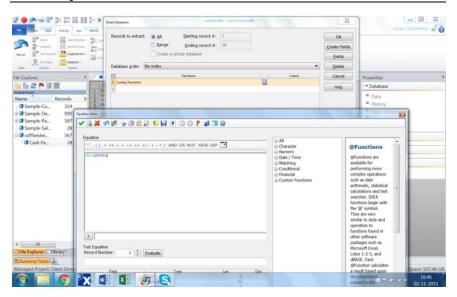
In the Equation Editor write the function

@dow(date)=1

Enter the @ function - @dow by pressing shift @ in the Equation box. This will bring up a list of all @ functions. Now enter dow, and we will get to view the function @ dow.

Press enter, open bracket enter d, select date as the field on the basis of which we have to run the @ dow function, close bracket and enter = 1.

This function will extract all cash payments made on Sundays, since Sunday is 1 within IDEA.



Click he Green Check sign in the Equation Editor to Validate and Exit the Equation Editor.

The function gives one instance of cash payment made on a Sunday.

This case needs to be taken up for investigation with the Accounts section.

Total Forder	New Selections December 20 Annual Selections December 20 An	0
e Explorer *		Properties
	DATE PAYEE TYPE DEBIT CREDIT EXPENSE 1 2545-365 CASH Aver - 277 8 CON WIDENS &	I - Database
atte Pouet Jame Records 5	Tananana ayak an a consistences	✓ Data
Sample-Cu 314		History
Sample-De_ 900 1		Field Statistics
Sample-Pa_ 187 1		 Control Total
Sample-Sal. 28		Criteria
sdffixedwi567		· Results
Cash Pa 28		
Sunda 1		• • Indices
		✓ No index
		• • Comments
		Add comment
<u> </u>		
File Explorer Dibrary		
Running Tasks		
naged Project: Client Demonstrat	on Not connected to IDEA Server Num	ber of Records: 1 Disk Space: 107.4

Close the file Sunday Payments and open Cash Payments from the File Explorer.

15.2.3 Identify the most expensive Cash Payment Expense Head

Go to Analysis – Field Summarization – Quick

Image: Source Control For The DEBT CREDIT OF DEPT OF DET OF DEPT	Arrise Street Sectors	P tratheter 17 nor		erector	Manafay Like *		
Image: Proceedings Image: Procedings <thimage: proceedings<="" th=""> <thimage: proce<="" th=""><th>le Explorer</th><th></th><th></th><th></th><th></th><th>- Properties</th><th></th></thimage:></thimage:>	le Explorer					- Properties	
Markadow Records S T<	- L 2 P 3 III	and create totals for each				I = Database	
Hime Records S Sumple Cu. No No<	catop Project					1.	
Simple-Da. 314 1 1 16000 1 160000 16000 16000<	Vame Records 5						
Simple-Da. 900 1 <	Sample-Cu 314	4 09-03-2005 CASH	Pyret	-870 4	PRODUCTION EXPE		
Sample Pot. 100 1 1 1 100 <t< td=""><td></td><td></td><td></td><td></td><td></td><td>Field Statistics</td><td></td></t<>						Field Statistics	
a m prior tal. 28 0 0.00 m prior tal. 200 0.00 m prior tal.							
→ Bright Sol. 10 240 States 10 240 States 10 240 States 10 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>							
# officered	Sample-Sal. 28					- Criteria	
■ ① Cach Pa. 28 12 240.00 Ppr 1000 Ppr <t< td=""><td>sdffixedwi_ 567</td><td></td><td></td><td></td><td></td><td>1 Block</td><td></td></t<>	sdffixedwi_ 567					1 Block	
Conda. 1		11 2240-305 CA94				· · · resurs	
10 10 <td< td=""><td></td><td>13 11-40-2005 CASH</td><td></td><td>-2380</td><td>PRODUCTION EXPE</td><td></td><td></td></td<>		13 11-40-2005 CASH		-2380	PRODUCTION EXPE		
16 18 18 18 18 18 18 19 1000CTENERS 1	Sunda 1					* Indices	
17 10 20 0.04 Pyre 10 <td< td=""><td></td><td>15 IL-03-2905 CASH</td><td></td><td></td><td></td><td>A Marinday</td><td></td></td<>		15 IL-03-2905 CASH				A Marinday	
18 10 18 18 10 18 10 10 18 18 10 10 18 10 10 18 10 10 18 10 10 18 10 10 18 10 10 18 10 10 18 10 10 18 10<						- NO INDEX	
10 10<		18 11-01-3805 CADH		-110	PRODUCTION EXPE	1. Comments	
10 >> 21 10 200 Color Fyret 300 1000		10 11-00-3005 CASH					
10 > 10 > 10 >		20 11-03-305 CASH				Add comment	
10 > 10 > 10 >		21 11 43 305 CASH					_
10 > 10 > 10 >		23 II 41-2015 CASH					
10 > 11 > 16 > 17 > 18 > 18 >		74 8-40-3005 CA94					
10 > 10 > 10 >		25 10-00-3005 CASH					
28 1 49 300 Cole Pyee 200 8 HOOCTON 194 He captorer _ Library /		20 10-00-000 CASH					
III		28 11-40-3005 CASH					
Russing Tasks	III						
	Running Tasks	11.2.					

This will open the Summarization dialog box.

Select the Summarization Field as EXPENSE from the drop down list of active fields

Select the Fields to Summarize as Debit from the drop down list of active fields.

in France Statistics *	Lower Lowe		×		
Explorer *	Fields to summarize: By: EXPENSE •	Numeric fields to total:	OK Operties		
me Records S Sample-Cu 314	Then by: NONE • Then by: •	CREDIT	Beids	Data History	
Sample-De 900 1 Sample-Pa 187 1 Sample-Sal 28	Then by: • Then by: •		Help	Field Statistics Control Total Criteria	
Galifiedada 567 Galifiedada 28 Galifiedada 1	Then by: • Then by: • Then by: •			Results	
	Çriteria:			Comments	
	Use Guidk Summarization Create glatabase Include % in output datagase Use fields from first occurrence Use fields from first occurrence	Statistics to include: Sum Average Agerage Magimum Variance Blaimum Standard deviation		Add comment	
	File game: Expense Summary	Result name: Summarization			

Click OK

The summarization gives a tab result in the file Cash Payments listing the Account Head, number of records for each payment and the sum total value.

Explorer		provide the second s	sh Payments.IMD 🗡 🛡 Expense Sum		Properties
12 P 0		EXPENSE	NO_OF_RECS	DEBIT_SUM	1 * Database
	ecords S	1 APPRENTICE - ST	1	-33930	🔮 Data
Sample-Cu	314	2 BLEACHING & DYE	3	-3731	 History Field Statistics
Sample-De_	900 1	3 CONE WINDING &	3	-4451	 Field Statistics Control Total
Sample-Pa_ Sample-Sal_	187 1	4 PRODUCTION EXPE	18	-30336	Criteria
sdffixedwi_	567	5 SCREEN PRINTING	1	-1596	/ = Results
Cash Pa_	28	6 WAGES	2	-500500	1 T NESURS
Expen	6		7		i * Indices
Sunda_	1				✓ No index
					· · Comments
					Add comment
m					

Double click the field-heading DEBIT. The field gets sorted on ascending basis and reveals the most expensive cash payment head: Wages Rs. 500500.

Tally Accounting System and IDEA

Ame Aust Jame Records 5 1 WAGES Jame Records 5 1 WAGES Jame Records 5 1 WAGES APPREN Jame Records 5 1 WAGES APPREN APPR	NTICE - ST ICTION EXPE WINDING &	2 1 18	_SUM - -500500 -33930 -30336	Database Data History
Imme Records 1 WAGES Sample-Cu. 314 2 APPREN Sample-Pa. 900 3 PRODU Sample-Pa. 187 4 CONE V Sample-Pa. 28 5 BLEACH Sample-Garmen-Sample-Pa. 28 6 SCREEN Staffickell. 567 6 SCREEN	NTICE - ST ICTION EXPE WINDING &	1 18	-33930	History
		3 3 1	- 4451 - 3731 - 1596	Field Statistics Control Total Criteria Criteria Results No index DEBT_SUNVA Comments Add comment

If we click the number of records field cell in NO_OF_RECS for Wages i.e. 2, it will give us a detailed listing of cash payments to workers.

A closer scrutiny reveals that these cash payments exceed Rs. 20000 under the Tax Audit rules. However, these entries are single combined entries.

Ask the Accountant for detailed listing of wage payments in cash to identify inadmissible payments under the Tax Audit Code for single transaction cash payments.

	gis Lettadion Preview		marine Till I	9 ET		-	
	DATE PAYEE 1 8-01-385 CASH 2 8-40-385 CASH	TYPE DEBIT CREDIT E	APENSE AGIS				
korer							Properties
27							• Database
ample-Cu ample-De ample-Pa ample-Sa							 Data History Field Statistics Control Total Criteria
Iffixedwi. Cash Pa							· • Results
Expen							* Indices
Sunda							■ No index ✓ DEBIT_SUM/A
							· · Comments
							Add comment
III	CONTRACT				Save Print Dor		

Chapter 16 Fraud Investigation using IDEA

16.0 Introduction

The degree of automation in the typical business environment has led to an increase in the complexity of internal control systems. Technological advances in distributed processing, worldwide networking, and remote access to corporate systems, for example, increase organizations' vulnerability to control breaches and present new challenges for the internal auditor. The impact of such technology on forensic audit practices is especially significant.

A recent study by the Association of Certified Fraud Examiners indicates that financial losses due to fraud in the United States amounted to a staggering \$400 billion in recent years. As more business operations use computerized information, more fraud is committed via computer assisted means. The technology that enables this type of fraud, however, also provides auditors with more sophisticated weapons to fight it. In fact, fraud detection is an ideal application for computer assisted audit techniques (CAATs). In recent years, analytical techniques have become not only more powerful but also more widely used by auditors for forensic investigation.

CAATs enable investigators to obtain a quick overview of business operations, develop an understanding of the relationships among various data elements, and drill down into the details of specific areas of interest. A systematic approach to fraud investigation that involves the identification of unusual activity with the aid of CAATs, including the use of digital analysis techniques such as Benford's Law, can help to ensure that misdoings get detected as close to the event as possible.

IDENTIFYING FRAUD WITH CAATS

While it does not "take one to know one" with regard to fraud investigation, auditors seeking to detect fraud would certainly want to know what it looks like. This requires auditors to recognize the tell-tale evidence of fraud and to understand how the data they obtain can be used to verify whether or not a fraudulent act has been committed. Data analysis software can assist by highlighting transactions that contain the characteristics often associated with fraudulent activity. Internal auditors can review millions of transactions,

including data from previous years or multiple locations, and probe anomalies by analyzing information across databases. Auditors should begin their analysis by pinpointing hot spots and then searching for patterns among the data.

LOOKING FOR SYMPTOMS Data analysis programs feature many commands that review records for fraud symptoms, such as the existence of duplicate transactions, missing transactions, and other anomalies. Examples of applications for these types of functions include:

- Comparing employee addresses with vendor addresses to identify employees who are also vendors.
- Searching for duplicate check numbers to find photocopies of company checks.
- Scanning the list of vendors to identify those with post office boxes for addresses. These records can be easily extracted from the vendor file for further follow-up to ensure they represent legitimate vendors.
- Analyzing the sequence of all transactions to identify missing checks or invoices.
- Identifying all vendor companies that have more than one vendor code or more than one mailing address. Such listings may represent "phantom" vendors, which exist for the sole purpose of stealing payments from companies.
- Finding several vendors with the same mailing address. These records may also signal phantom vendors.

Investigators can narrow their search for symptoms by forming hypotheses about the data and then testing them with CAATs. For example, an auditor could reasonably assume that each employee in the organization receives only one paycheck per pay period. Using CAATs to search for duplicate records in a single pay period would quickly test the validity of the assumption by highlighting all instances where employees received more than one paycheck. If the assumption is correct, the existence of employees receiving two or more paychecks may indicate fraud. If the assumption is invalid, some employees may receive a separate check for overtime in addition to their regular salary. For instance, the auditor can then form a revised hypothesis that each employee should receive only one paycheck for regular hours per pay period. The database search would then be revised on the basis of this new assumption.

CAATs can also help to reveal fraud symptoms that may be elusive because the evidence is spread across separate databases. For example, reviewing data from the accounts payable file may identify a trend in the expenditures to a particular vendor that seems unusual, but not necessarily indicative of fraud. Combining the accounts payable data with information from the contracting database, however, may reveal that all contracts with the vendor in question were raised by one contracting officer. This type of relationship among the data would indicate the possibility of a fraudulent activity, such as kickbacks received by the officer from the vendor. Conducting a search of this type would be impractical, even impossible with either sampling techniques or other manual methods of investigation.

PATTERNS OF FRAUD Experience and training in electronic fraud investigation enables auditors to recognize and identify data patterns that may be associated with fraud. Patterns such as negative entries in an inventory-received field, voided transactions followed by a "no sale," or a high percentage of returned items may indicate of fraudulent activity.

As investigators become more familiar with the characteristics of suspicious patterns, they can develop sets of predefined criteria, and transactions meeting these criteria can trigger automatic investigations. In other words, auditors can use CAATs to create a "fraud profile," which functions as a template for future audits. The template may also be used to monitor transactions on an ongoing basis, thereby functioning as a continuous tracking device to flag any suspicious activity, as soon as it appears in the database. This technique can serve as an instrumental means of reducing the number of losses attributed to fraudulent acts. For instance, a timely review of credit card transactions can identify unusual spending patterns that might represent the symptom of a stolen credit card. Continuous monitoring for suspicious patterns allows auditors to catch fraud early, before serious damage can occur.

16.1 Benford's Law of Digital Analysis

16.1.1 Evolution of the Law & its Concept

Frank Benford was a physicist at GE Research Laboratories in the 1920s. He noted that the first parts of the log table books were more worn than the back parts. The first pages contain logs of numbers with low first digits. The first digit is the left-most digit in a number.

Benford collected data from 20 lists of numbers totaling 20,229 observations. He found that the first digit of 1 occurred 31 percent of the

time. Using integral calculus, he calculated the expected digit frequencies that are now known as "Benford's Law". It took him six years to perform his analysis and develop his law of expected digit frequencies.

The Benford's Law task in IDEA can provide a valuable reasonableness test for large data sets. IDEA only tests items with numbers over 10.00. Number sets with less than four digits tend to have more skewed distributions and do not conform to Benford's Law. Positive and negative numbers are analyzed separately. This is because abnormal behavior patterns for positive numbers are very different from those for negative numbers.

16.1.2 Application to Data Mining

The application of Digital Analysis and the Benford Module is also permissible in the framework of Data Mining where certain distinctive facts in a data supply are measured against the personal expectations of the user and interpreted according to them. In this case it is not necessary that the data to be analyzed create a Benford Set, . In fact, it is permissible under these circumstances to analyze the numerical distribution of the leading digits of each data quantity and to interpret it independent of Benford's Law.

16.1.3 Assumptions of Benford's Law

Geometrical Series

The mathematical pre-condition for the examination of a data supply based on Benford's Law is that the data supply is based on a geometrical series (thus, it is presented as a Benford Set). In reality this condition is rarely met. Experience, however, confirms that data must only partially meet this condition, i.e., the constant increase, percentage-wise of an element compared to the predecessor must only be met partially. Otherwise, this would mean that no number may occur twice, which is quite improbable in the case of business data supplies. However, the pre-condition is that there is at least a "geometrical tendency".

Description of the same object

The data must describe the same phenomenon. Examples are:

- The population of cities
- The surface of lakes
- The height of mountains

- The market value of companies quoted on the NYSE
- The daily sales volume of companies quoted on the Stock Exchange
- The sales figures of companies

Unlimited data space (non-existence of minima and maxima)

The data must not be limited by artificial minima or maxima. A limitation to exclusively positive numbers (excluding 0) is permissible as long as the figures to be analyzed do not move within a certain limited range. This applies, for example, to price data (e.g., the price of a case of beer will generally always range between 15 and 20 dollars) or fluctuations in temperature between night and day.

No systematic data structure

The data must not consist of numbers following a pre-defined system, such as account numbers, telephone numbers, and social security numbers. Such numbers show numerical patterns that refer to the intentions of the producer of the number system rather than to the actual object size, represented by the number (e.g., a telephone number starting with 9 does not mean that this person possesses a bigger telephone).

16.1.4 Statistical Explanations for the Law

Basically, data complies best with Benford's Law if it meets the rules mentioned above, namely that the data consists of large numbers with up to 4 digits and the analysis is based on a sufficiently large data supply. A large data supply is necessary to come as close to the expected numerical frequencies as possible. For example, the expected frequency of the digit 9 in any data supply is 0.0457. If the data supply consists of only 100 numbers, the numbers which have a 9 as their first digit may be 5% of the data supply. Thus in the case of a small data supply, there may be an over-proportional deviation from Benford's Law. In large data supplies, the numerical distribution is increasingly closer to the expected frequencies.

If the data supply has, or just roughly has, the characteristics mentioned above, it can be analyzed based on Benford's Law. However, the results of the Benford analyses are not interpretable on the basis of Benford's Law. As stated before, the expected frequencies according to Benford's Law often represent, in the practical use, nothing more than a type of benchmark for the observed frequencies. Since the observed frequencies will only be compared with the legality discovered by Benford, not interpreted accordingly, it is not necessary that all conditions mentioned above be met. In fact, the analysis results will help the auditor to interpret the personal expectation of the user, without including the reference value according to Benford in the investigation. If, for example, the personal expectation of the user is that the starting digit 4 must occur twice as often in the analyzed data than the starting digit 2, the results of the analyzed values must not be compared with the expected frequencies according to Benford, but with the individual expectation of the user.

16.1.5 Case Study on Benford's Law within IDEA

To run a Benford's Law analysis on the Accounts Payable database

Factor Factors	St testa	na Nay *		athuse 1		in Martin Corrector							
e Explorer	57 Beature			Sample-Pa	yments								Properties
1 2 P G	Compare I Rectamber	Na summit	data to pected	OSTED_E	Y SUPPNO	SUPPNAME	INVOICE_DATE						• Database
Non Property	read.			SA	20500	AASAAAA DESIGN BKC	81-12-2014	81344	852338	87,85/2004	200082100	CP	
	ecords	S	2	Then .	20900	ACKLANZS - GRAINGER INC	11-12-2014	8222008	95223t	17/65/2014	200085400	BC	✓ Data
		5	3	SOUR	E1400	AG JOYEFIA	11-12-2014	34835	C31006	25/95/2254	330095800	K	History
Sample-Cu_	314		4	KSA	11810	DRAM BAC	31-12-2014	256	\$52394	06/85/2054	330091200	10.7	
Sample-De_	900		6	DDC ERC	12203	ANKE KELSE ORGANICS INC BETOMOTEIN GARDIN	31-12-2014 31-13-2014	A8 3283 M 21969	C50023 852204	25/03/2014	101007400	10	Field Statistics
			6	154	20045	EDOUTIRE CARMEN	11-12-2014	21560	852207	35/03/2014	200083000	10.	Control Total
Sample-Pa	187		1	MA	20129	BLEVICH BOEDO	21-01-2014	634-567	ASJAN	04/01/2004	100083100	WH .	
Sample-Sal_	28		8	MSA	201.10	BEHOP PUBLISHING INC.	01-01-2014	82641	A52638	13/11/2014	100080000	14.0	Criteria
			10	SOUR	20133	BEHOP PUBLISHING INC.	15-01-2014	10795ABC-123	A52030	35/01/2004	2000825.00	18.	
sdffixedwi	567		11	MSA.	30,004	BOREAL LABORATORIES LTD	09-01-2014	AZ178	A5264G	63/62/2004	330060300	WE	Results
Cash Pa_	28		12	MIA	20292	GUESS INC.	25-01-2014	81340	A52842	05/02/2054	200082800	80	
Expen	6		13	KSA	39403	BRECKG	57-01-2014	371004A	A52843	04/62/2004	330080800	CB	Stratification
			14	CW	20414	CITAZEN INC.	10 10 2014	10000 4	A52544	94/02/2004	100062500	R.	Summarization
Sunda	1		15	SOU#	20508	FOSSE INC.	18-01-2014	200139	A53545 A53546	64/62/2014	190082100	+8.	
			16	CW	20908	FOSIA INC. INFRATERIA	08-01-2014 05-01-2014	130729 000496C/W	A52646 A52647	04/62/2014	230082900	145	Pivot Table
			17	CW KSA	20532	MERGER KD	10-01-2014	2024040,000	45,847	06/02/2014	130083000	W	Benford
			10	CW	205.32	INCOMER NO.	11-01-2014	147905CTR	852126	13/62/2004	130086300	10.	- Demora
			19 20 21	#5A	20535	XOHIN PETERSON	08-01-2014	34676/50	852117	13/02/2004	130084100	8.8	• Indices
			21	1008	20515	BURGHARD1	52-01-2014	CS-562-07	A52858	34/02/2004	200084500	64.8	- Brances
			22	CW	20535	EURO-GADT	10-01-2014	FR-963.32	A52650	08/82/2054	330064300	18.	No index
			23	CW	20535	BURIO LARCIT	01-01-2014	117-2287	A52660	35/62/2654	10063600	18	
			24	SOUR	20536	JOHN PETERSON	02-01-2014	TN 6408 97	A52662	35/82/2004	300084600	CF.	AMOUNT/A
			25	1048	20535	BURXHUND1	15-01-2014	94-392817	A52963	33/92/2004	200083500	BC .	SUPPNO/A + SUPPNAME/A
			26	KSA SOUR	29515	BURGHARDT	31-12-2004	104256/P	A52664	11/02/2004	180083400	78.	JULTING A SOFFICIER
			2234556789	SOUR SOUR	20535 20554	BURIOWARD1 BURINT SAND SOLUTIONS INC.	04-01-2014 10-01-2014	7030445 7503040	A52665 A52666	11/02/2014 11/02/2014	230083800	WF HD	• Comments
			-8	CW.	20594	CADEMIN MON COCO INC	52-01-2014	PP148E755	A02060	11/10/2014	130083000	WE .	Southernords
			30	DES	29625	CARTAL TOWING SERVICES LTD	31-12-2014	1573218	852217	12/05/2004	100094000	+8.	Add comment
			31	K5A	20040	LUMINON INC.	11-12-2014	81522	852380	06/55/2514	130081400	+8.	rive comment
			32	000	20658	SKAGEN INC.	21-12-2014	81345	CNIDE	34/03/2014	100097100	15	
			32 33 34 34	CW	20694	CASA CORONADO	19-01-2014	\$7/5238/A	A52670	13/82/2014	330080500	+8.	
III			34	CW	20765	COMPAC	31-12-2814	201206787	812172	05/05/2004	230089900	X8 -	
File Explorer CUbr	-	-	94	CW	398.96	CHEENING KOVAS	81-01-2014	718.627	\$59C9	13/07/0814	100980100	10.	
	my.					······································					and the second s		
Huming Tasks									nected to				r of Records: 187 Disk Space: 10

1. From the Analysis menu, select Benford's Law.

2. In the **Benford's Law** dialog box, select **AMOUNT** as the field to be analyzed. Accept all other default options as displayed in the image below.

a 2 - Johnel Si Stelledroom	E teat		tarest Bi		le 🚰 Hanalary UHI * - 🏄 Kanakin *			_	_	×	
Refer to This		enford's Law						_			
plorer	-	Field to ana	lyze: AMOUNT						OK		Properties
		Include Values		Optio					Advanced		• Database
e Records	1	Positive			Show boundaries				Cancel		I Data
e Records : Sample-Cu 314	5 1	Negative			Mean absolute deviation						History
Sample-De_ 900		Result							Help		Field Statistics
Sample-Pa., 187	7	Create result		Be	nford1						Control Total
Sample-Sal. 28	9										E Criteria
udffixedwi567	10. 11	Analysis									· • Results
Cash Pa., 28	12	Test:			ate database:						
Sunda_ 1	14	Pirst digit:	-	2	Benford First Digit1						· * Indices
Janan 1	16	First two digits:	-		Benford First Two Digit	51			Suspicious		· Comments
	18	First three digits:	2		Benford First Three Dig	p81					# Add comment
	20	Second digit:	2		Benford Second Digit1						in the comment
	21	E Last two digits:	-		Benford Last Two Digit	1			Suspicious		
	23	Second order:			Benford Second Order				September		
	25	Summation:			Benford Summation 1				Sespicious		
	27	Summation:	-						Sebours		
	29				Cryote a virtual database						
		E 2003	EASA CORDINADO	-	20-12-2014 12-12-2014	80.1206/A	ASSED	1140.00	4 133387500		
	34 0	4 2030	CORDIC DHE MAD KIYAS		31-12-3014 11-301-3014	PP566787 755.637	852572 &40821	11.00.004		12 -	
Explorer Ubrary	+ 10000		The second second	111		CONTRACTOR OF THE OWNER	and the second		22 (Salar	÷.,	

Click **OK** to perform the analyses.

The **Benford First Digit** database becomes the active database. Other databases should be opened from the *File Explorer*.

Chart						×
imes field:	DIGITS	~	Туре:	Bar	*	OK
Y field(s):	DIGITS EXPECTED LOWBOUND HIGHBOUND	ACTU	-	Do not show SD chart Show grid lines Make a snapshot	✓	Cancel Help
X axis title:			Chart title:			
Y axis title:			Criteria:		[
Result name:	Chart Data					

3. To graph the data, select **Chart Data** from the **Data** menu.

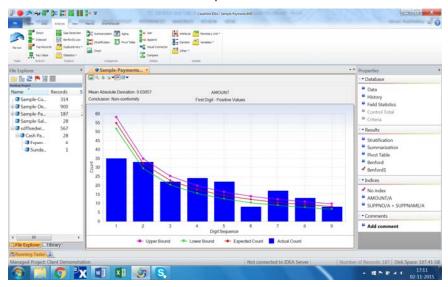
In the Y field(s) box, select ACTUAL.

In the X axis title box, enter Digit Sequence.

In the Y axis title box, enter Count.

In the Chart title box, enter AMOUNT – First Digit – Positive Value.

Click OK.



The Chart Data Results output becomes active.

The first digit graph shows a spike in the digit 7 results.



The **First Digit** test is the test of first digit proportions. The first digit of a number is the leftmost digit in the number. Zero can never be the first digit. This is a high level test. Analysts will not usually spot anything unusual unless it is blatant. This is a test of goodness-of-fit to see if the first digit actual proportions conform to Benford's Law. The **First Digit** test is an overall test of reasonableness. The upper and lower bounds are merely guidelines for the auditor. The **First Digit** graph could show a high level of conformity but the data set could still contain errors or biases.

4. Click the **Data** link in the *Properties Window* to return to the **Benford First Digit** database that was created as part of this analysis. The **DIFFERENCE** field shows the difference between the expected and actual occurrences of the digits. When the **DIFFERENCE** field is indexed in ascending order, the digit 7 results show the largest negative difference (positive spike).

Data Analysis for Auditors

	DIGITS	EXPECTED	LOWBOUND	HIGHBOUND	ACTUAL	DIFFERENCE
	DIGITS	EVLECTED	LOUBOOND	HIGHBOOND	ACTORE	DIFFERENCE
1	7	57.82	53.26	62.38	113	-55.18
2	5	78.94	73.61	84.27	106	-27.06
3	6	66.75	61.84	71.65	90	-23.25
4	9	45.62	41.57	49.67	61	-15.38
5	8	51.00	46.71	55.28	63	-12.00
6	4	96.62	90.72	102.52	108	-11.38
7	3	124.56	117.87	131.26	114	10.56
8	2	175.56	167.61	183.51	154	21.56
9	1	300.13	289.73	310.52	188	112.13

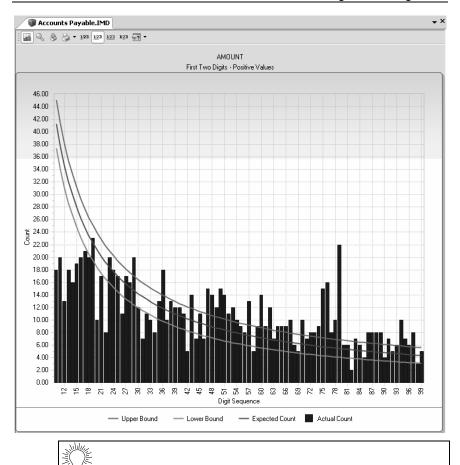
This result warrants further investigation, as the CFO has indicated that any items in excess of Rs.80,000.00 require additional approval. This spike could be indicative of an abnormal level of items being processed just below the additional approval level.

Close the **Benford First Digit** database.

5. To view the **Benford's Law** analysis result again, ensure that **Accounts Payable** is the active database and click the **Benford** link in the **Results** area of the *Properties Window*. To view the **First**

Two Digit graph, click on the **123** button on the **Results** toolbar. The 79, 76, and 75 two-digit combination spikes are clearly visible in this graph.

Fraud Investigation using IDEA



Technical Notes

The **First Two Digit** test is a more focused test. The first two digit numbers are the leftmost two digits. There are 90 possible two-digit combinations ranging from 10 to 99. This test is performed to find anomalies in the data that are not readily apparent from either the **First Digit** test or the **Second Digit** test when viewed on their own. A spike occurs where the actual proportion exceeds the expected proportion as predicted by Benford's Law. Positive spikes (above the Benford's curve) represent excessive duplication. One of the objectives of this test is to look for spikes that conform to internal thresholds like authorization limits.

6. We look at the transactions that comprise the 79 two-digit combination by clicking on the graph and selecting **Display Records**

to drill down to the transactions. We notice the number of transactions just under the Rs. 80,000 approval limit.

	SUPPNO	PAYEE	INVOICE	INV_DATE	AMOUNT	CHECK	PAY_DATE	AUTH	
1	C202	Cary S Matic	CS - 589 -97	22/03/2006	79,217.47	701241	16/04/2006	H.M.V.	
2	M025	Luke Hair	51505	24/01/2006	79,500.00	701073	17/02/2006	H.M.V.	
3	F128	Ern Payed	82 95 1	15/02/2006	79,826.53	701148	13/03/2006	WJT	
4	P007	Nellie Dunn	000511CJW	26/02/2006	7,988.72	701179	26/03/2006	WJT	
5	T004	Richmond	IN/1151/97	07/03/2006	79,839.54	701207	01/04/2006	WJT	
6	N001	Mike Atsil	AZ362	22/04/2006	79,910.95	701346	17/05/2006	HV	
7	R007	Philip Upp Garage	1421/21	01/07/2006	79,365.72	701558	28/07/2006	HMV	
8	M020	Linda Hand	GR223 97	04/07/2006	79,002.68	701572	30/07/2006	HMV	
9	R008	Philip Upp Garage	1464/21	16/10/2006	7,973.99	701910	12/11/2006	BC	
10	H025	Ivan Aker	WNZ28C	02/01/2006	79,571.88	701002	31/01/2006	BC	
11	T006	Round Table	BUR1355-G	27/03/2006	79,839.65	701260	25/04/2006	H.M.V.	
12	K001	Jackie Tupp	917356	22/03/2006	79,990.86	701262	19/04/2006	H.M.V.	
13	N005	Miles Long	FR-1026 82	30/06/2006	79,306.51	701547	28/07/2006	HMV	
14	M014	Kurt N Upp	123513	16/07/2006	79,902.90	701628	13/08/2006	HMV	
15	M014	Kurt N Upp	123522	19/08/2006	79,765.18	701732	16/09/2006	WJT	
16	8008	A Raid	3501101	07/08/2006	7,948.57	701586	05/08/2006	WJT	
17	R020	Polly Gunn	L-1347/73	02/09/2006	79,458.30	701765	30/09/2006	WJT	
18	P009	P Green	566698T	04/01/2006	79,237.49	701012	03/02/2006	BC	
19	Z001	Miles Long	FR-969 39	17/01/2006	795.59	701064	15/02/2006	HMV	
20	T005	Ri Pent	PPN98831	07/04/2006	79,086.25	701300	06/05/2006	BC	
21	R007	Peter Rabbitt	KG0385	16/09/2006	79,901.01	701820	15/10/2006	BC	
22	N005	Miles Lona	FR-1073140	25/10/2006	7,962.64	701943	25/11/2006	HMV	1

Of the 22 transactions that make up the 79 two-digit combination, 17 are between 79,000 and 80,000. Each of the payables clerks authorized some of the 17 transactions, but HMV was responsible for the bulk of these transactions, having authorized 9 out of 17.

	SUPPNO	PAYEE	INVOICE	INV_DATE	AMOUNT -	CHECK	PAY_DATE	AUTH	^
1	K001	Jackie Tupp	917356	22/03/2006	79,990.86	701262	19/04/2006	H.M.V.	
2	N001	Mike Atsil	AZ362	22/04/2006	79,910.95	701346	17/05/2006	HV	
3	M014	Kurt N Upp	123513	16/07/2006	79,902.90	701628	13/08/2006	HMV	
4	R007	Peter Rabbitt	KG0385	16/09/2006	79,901.01	701820	15/10/2006	BC	
5	T006	Round Table	BUR1355-G	27/03/2006	79,839.65	701260	25/04/2006	H.M.V.	
	T004	Richmond	IN/1151/97	07/03/2006	79,839.54	701207	01/04/2006	WJT	
7	F128	Ern Payed	82 95 1	15/02/2006	79,826.53	701148	13/03/2006	WJT	
8	M014	Kurt N Upp	123522	19/08/2006	79,765.18	701732	16/09/2006	WJT	
9	H025	lvan Aker	WNZ28C	02/01/2006	79,571.88	701002	31/01/2006	BC	
10	M025	Luke Hair	51505	24/01/2006	79,500.00	701073	17/02/2006	H.M.V.	
11	R020	Polly Gunn	L-1347/73	02/09/2006	79,458.30	701765	30/09/2006	WJT	
12	R007	Philip Upp Garage	1421/21	01/07/2006	79,365.72	701558	28/07/2006	HMV	
13	N005	Miles Long	FR-1026 82	30/06/2006	79,306.51	701547	28/07/2006	HMV	
14	P009	P Green	566698T	04/01/2006	79,237.49	701012	03/02/2006	BC	
15	C202	Cary S Matic	CS - 589 -97	22/03/2006	79,217.47	701241	16/04/2006	H.M.V.	
16	T005	Ri Pent	PPN98831	07/04/2006	79,086.25	701300	06/05/2006	BC	
17	M020	Linda Hand	GR223 97	04/07/2006	79,002.68	701572	30/07/2006	HMV	
18	P007	Nellie Dunn	000511CJW	26/02/2006	7,988.72	701179	26/03/2006	WJT	
19	R008	Philip Upp Garage	1464/21	16/10/2006	7,973.99	701910	12/11/2006	BC	
	N005	Miles Long	FR-1073140	25/10/2006	7,962.64	701943	25/11/2006	HMV	
21	B008	A Raid	3501101	07/08/2006	7,948.57	701586	05/08/2006	WJT	
22	Z001	Miles Long	FR-969 39	17/01/2006	795.59	701064	15/02/2006	HMV	×

Of the 16 transactions that make up the 76 two-digit combination, 10 are between 76,000 and 77,000. Of these 10 transactions, HMV authorized 7 transactions.

Of the 15 transactions in the 75 two-digit combination, 13 are between 75,000 and 76,000. Like before HMV authorized 6 of the 13 transactions. In addition, there are 5 transactions for exactly 75,000. Out of these 5 transactions, 4 were payments to companies with "Cash" in the payee name. Each of these 4 transactions was paid within a few days of their invoice date, clearly a violation of company policy.

	SUPPNO	PAYEE	INVOICE	INV_DATE	AMOUNT 🔺	CHECK	PAY_DATE	AUTH
1	M130	Microcomputers	5863MCC	7/28/2006	7,545.77	701672	8/26/2006	VMH
2	M123	Luke Hair	51723	10/29/2006	7,582.46	701947	11/26/2006	H.M.V.
3	M100	M Cash Inc	UP-76409	10/1/2006	75,000.00	701774	10/6/2006	HMV
4	M100	Cash Inc	CS-717-97	9/13/2006	75,000.00	701728	9/15/2006	VST
5	M100	Co Cash Inc	T5352	10/17/2006	75,000.00	701849	10/20/2006	V.S.T
6	W007	Mr Cash Co	22 31 97	10/10/2006	75,000.00	701803	10/11/2006	VST
7	W008	Willis T. J.	JK-43-JW/97	9/5/2006	75,000.00	701770	10/1/2006	BC
8	M128	Maurice Mynah	IN-392920	8/16/2006	75,031.08	701730	9/16/2006	VMH
9	P009	O Kay Yahs	BC 469897 W	10/31/2006	75,192.58	701964	11/29/2006	HMV
10	T006	Ronnie Biggs	7899499	8/23/2006	75,243.00	701742	9/20/2006	HMV
11	D025	Denise Bent	81340	1/3/2006	75,373.66	701003	1/31/2006	BC
12	M020	Linda Hand	GR278 97	10/31/2006	75,537.32	701962	11/29/2006	HMV
13	F128	Ern Payed	82 98 1	3/4/2006	75,553.42	701200	4/1/2006	HMV
14	A128	A Meadow	117-2393	9/5/2006	75,777.62	701773	10/1/2006	BC
15	C202	Cary S Matic	CS - 599 -97	4/8/2006	75,819.56	701287	5/5/2006	H.M.V.
					Save		Print	Done

7. Open the **Benford First Two Digits** database that was created as part of the analysis. Index the **DIFFERENCE** field in ascending order. Notice the large negative differences (positive spikes) in the first two digit combinations of 79, 76, and 75. These were the underlying cause of the spike in digit 7 results in the **First Digit** graph.

Data Analysis for Auditors

	DIGITS	EXPECTED	LOWBOUND	HIGHBOUND	ACTUAL	
1	79	5.45	4.05	6.85	22	-16.55
2	79	5.66	4.03	7.09	16	-10.55
2	75	5.74	4.23	7.09	15	-10.34 -9.26
<u>د</u>	60	7.16	4.30 5.55	8.76	10	-9.20 -6.84
4 5	50	8.57	6.82	10.33	14	-0.84 -6.43
8	36	11.86	9.80	13.93	13	-6.14
6 7	47	9.12	7.30	10.93	15	-5.88
8	51	8.41	6.67	10.35	14	-5.59
9	57	7.53	5.88	9.18	14	-5.47
10	94	4.58	3.30	5.87	10	-5.42
11	29	14.68	12.38	16.98	20	-5.32
12	48	8.93	7.14	10.72	14	-5.07
13	62	6.93	5.35	8.51	12	-5.07
14	78	5.52	4.11	6.92	10	-4.48
15	43	9.95	8.06	11.85	14	-4.05
16	53	8.09	6.39	9.80	12	-3.91
17	70	6.14	4.66	7.63	10	-3.86
18	67	6.41	4.90	7.93	10	-3.59
19	97	4.44	3.18	5.71	8	-3.56
20	49	8.75	6.97	10.52	12	-3.25
21	74	5.81	4.37	7.26	9	-3.19
22 23	89	4.84	3.52	6.16	8	-3.16
23	88	4.89	3.57	6.22	8	-3.11
24	87	4.95	3.61	6.28	8	-3.05
25	86	5.01	3.66	6.35	8	-2.99
26	52	8.25	6.52	9.97	11	-2.75
27	66	6.51	4.98	8.04	9	-2.49
28	95	4.53	3.26	5.81	7	-2.47
29	77	5.59	4.17	7.01	8	-2.41

8. Because of the size of this data set, a number of transactions were identified for further investigation by using the **First Digit** test and the **First Two Digits** tests. In larger data sets, a finer filter of transactions is necessary.



Technical Notes

Where there are 90 possible two-digit combinations, there are 900 possible three-digit combinations from 100 to 999. The **First Three Digits** test is a highly focused test that gives the analyst relatively smaller sections due to abnormal duplication and allows for a more narrowly focused analysis. This test is also valuable to look for spikes just below internal and psychological thresholds like authorization limits. To make the most effective use of this test, the source data set should normally exceed 10,000 records.

2. Open the **Benford First Three Digits** database that was created as part of the analysis. Index the **DIFFERENCE** field in ascending order.

				HIGHBOUND	0.071101	
_	DIGITS	EXPECTED	LOWBOUND		ACTUAL	DIFFERENCE
1	750	0.58	0.12	1.03	6	-5.42
2	807	0.54	0.10	0.98	5	-4.46
3	769	0.56	0.11	1.01	5	-4.44
4	511	0.85	0.29	1.40	5	-4.15
5	187	2.31	1.40	3.22	6	-3.69
6	962	0.45	0.05	0.85	4	-3.55
7	799	0.54	0.10	0.98	4	-3.46
8	798	0.54	0.10	0.98	4	-3.46
9	533	0.81	0.27	1.35	4	-3.19
10	492	0.88	0.32	1.44	4	-3.12
11	469	0.92	0.35	1.50	4	-3.08
12	377	1.15	0.50	1.79	4	-2.85
13	186	2.32	1.41	3.24	5	-2.68
14	299	1.45	0.72	2.17	4	-2.55
15	296	1.46	0.74	2.19	4	-2.54
16	940	0.46	0.05	0.87	3	-2.54
17	295	1.47	0.74	2.19	4	-2.53
18	922	0.47	0.06	0.88	3	-2.53
19	899	0.48	0.07	0.90	3	-2.52
20	841	0.51	0.08	0.94	3	-2.49
21	795	0.54	0.10	0.99	3	-2.46
22	278	1.55	0.81	2.30	4	-2.45
23	762	0.57	0.12	1.02	3	-2.43
24	275	1.57	0.82	2.32	4	-2.43
25	749	0.58	0.12	1.03	3	-2.42
26	739	0.59	0.13	1.04	3	-2.41
27	660	0.66	0.17	1.14	3	-2.34
28	624	0.69	0.19	1.19	3	-2.31
29	615	0.70	0.20	1.21	3	-2.30

Fraud Investigation using IDEA

The 750 digit combination has the largest negative difference (positive spike).

Select Window > Close All.

16.2 Other Case Studies on Fraud Investigations using IDEA

Other case studies on Fraud Investigations using IDEA have been summarized below with reference links to specific chapters containing detailed write-ups.

S No	Function	Investigation Area	Reference
1	Accounts	 Employees as Vendors Favorable treatment of	Point 14.1.3 in
	Payable	Vendors	Chapter 14.1

		• Transactions at or near Spending Authorities	
2	Accounts Receivable	 Teeming and Lading Forced fictitious Sales to meet Sales Targets 	Point 14.2.3 in Chapter 14.2
3	Inventory	• Fraudulent Inventory Valuations	Point 14.3.3 in Chapter 14.3
4	Payroll	Payments to Ghost Employees	Point 14.4.2 in Chapter 14.4
		• Unusual Overtime Payments	

Chapter 17

Documentation of CAATs

Documentation of CAATs: In the Planning for use of CAATs

17.0 Introduction

Data identification and its acquisition play a crucial role in the use of CAATs; these facilitate the efficient and effective use of CAATs and help realize guaranteed results and returns from its use.

The stages in Data Identification and Acquisition are

- Planning a CAAT
- Objective and assurances
- What data is available?
- The flow of data in a system
- Identifying the file for interrogation
- Specifying the information required for CAAT
- Specifying the format of the data file
- File downloading and conversion
- The Standard requirements for data
- Modes of data storage and transfer
- Who or what can help with data identification / acquisiton?

17.1 Regulatory Admissions

The Information Systems Audit and Control Association, U.S.A., has issued an IS Auditing Guideline on 'Use of Computer Assisted Audit Techniques (CAATs) Document G 3'

According to the Guideline

The major steps to be undertaken by the IS Auditor in preparing for the application of the selected CAATs are:

- Set the audit objectives of the CAATs
- Determine the accessibility and availability of the organisation's IS facilities, programs/ system and data.
- Define the procedures to be undertaken (e.g., statistical sampling, recalculation, confirmation etc.)
- Define output requirements
- Determine resource requirements, i.e., personnel, CAATs, processing environment (organisation's IS facilities or audit IS facilities)
- Obtain access to the organisation's IS facilities, programs/ system, and data, including file definitions.
- Document CAATs to be used, including objectives, high-level flowcharts and run instruction.

The documentation to be built during the stages of Data Identification and Acquisition are explained below.

17.2 Planning a CAAT

- The user needs to select the most appropriate file/ files for CAATs interrogations that best suit the objective to be tested.
- The user should make use of flow charts of the client department's system to ascertain which files are being used. It is important to select data that is as close to the original data entry process as possible.
- The flow charts and file/files selected should be documented and kept on record.
- Preferably the objectives along with the file/files selected should be documented in the manner shown below

S No	Objective	File	Form	Fields	Source	Owner	Import Method

17.3 Objectives and Assurances

• For any type of engagement it is necessary for the user to get to know the function and practices/ policies of the entity to be audited. This knowledge will help the user to plan the engagement effectively.

- For example, in a Financial Review reconciliation, High-Value item detection, exception reporting, and sampling plays a pivotal role.
- To achieve the stated objectives, a detailed knowledge of the system procedures needs to be made.
- These documents, policy notes ,and flow charts need to be documented by the user as a part of his engagement evidence.
- In forensic reviews, the engagement evidence has to be irrefutable, exhaustive and unambiguous.

17.4 What Data is available?

- The user needs to have a thorough knowledge of the operating system used by the entity being reviewed.
- The user needs to have a detailed knowledge of the application system being reviewed to identify the best possible data source and data format for each objective being reviewed.
- The user needs to engage in a meaningful discussion with the IT personnel of the entity. At times, an intermediary may be needed to accompany the user. The intermediary should have a good grasp of both audit and IT fundamentals.
- This meeting with the IT personnel needs to be undertaken formally with a clear written agenda and with space for the IT personnel response. Any format can be followed which is convenient to the user team so long as consistency of the Agenda document is always maintained.

17.5 The Flow of Data in the System

- The user needs to obtain from the entity Accounting Hand Books / Process Manuals, Operational Process Flow Charts to have a clear understanding of the details concerning transaction files, master files affecting (given?) lot of transactions.
- These flow charts need to be documented with the user and clear marks should be placed on the nature of data within the system.
- This documentation will help the user in the current engagement as well as in future cyclical engagements.

17.6 Identifying the file for interrogation

- The user must aim at selecting files which are closer to the raw data or initial input file. The study done by the user in the preceding stage will assist and facilitate in identifying such files.
- The user may choose to liaise with the IT personnel to identify appropriate data files. This liaison may be done through email, or in person, but never orally.
- For example, the user should know when to use a detailed transaction file, or a summarized version of a file or a table directly from the database or a customer master file.

17.7 Specifying the information required for the CAAT

- Once the right file has been identified, it may contain information not required for the audit. At this stage, it is prudent for the user to seek specific fields for his review.
- This has performance benefits for the user as the file import will get hastened with specific fields rather than with all fields from the file.
- For example, a Data Table with 40 Million rows and 25 fields will take around 10 minutes for import within IDEA. The same table with 5-10 relevant fields will get done faster.
- If the information to be sourced is not in the ideal format or key information is missing, it may be necessary to liaise with the IT personnel to arrange some form of file manipulation such as an embedded audit module or a file manipulation program to produce the data in a more useful format. This request to IT has to be made in writing.

17.8 File Download and Conversion

- Once the data is readied, the data may be placed on the User's PC, or it may be placed on a shared audit folder on the Local Area Network.
- If the data is not too large, it can be sent through group office mail or even through secondary storage devices like CDs and Pen Drives.
- Data Transfer should always be accompanied by a virus scan at all times.

- Upon receiving the data, the user must sanitize the data i.e., clean the data to make it suitable to an IDEA Import compatible format.
- For example, MS-Excel files should be cleaned; the first row should reflect the header row.,There should be no blanks in the data grid, and there should be no repeating header or footer rows in the data grid.
- Once the data is finally readied, the user can deploy the best technique within the Tool for Data Import.
- For example: MS-Excel for Excel Files, SAP/AIS for AIS SAP Flat Files and more.
- The user may initiate further technical liaison in writing with the IT to improve the file layout – record layout for future. This usually occurs when the entity changes its accounting system or migrates to a contemporary version of the application.

17.9. Who or What can help in Data Identification?

- The user needs to identify the right resource from the IT Section and Audit Section for the CAAT project.
- This identification can be done by seeking assistance from the HR Section and the Heads of IT and Audit respectively.
- The criteria for selection, such as the knowledge of data, knowledge of file formats, cross sectional knowledge of functions, and the like e should be documented by the user.
- These criteria need to be provided to the HR Head, IT Head and Audit Head while making the selection.

17.10 Specimen Format of a Data Indent

To –

From –

Date -

Subject – Data Request for review of (insert area)

Pursuant to our recent meeting with respect to our planned audit of Payroll, using CAAT Data Analysis Software, we would like to arrange for download of the attached data files and fields to the FTP Server set up for the purpose.

This review involves the audit of (insert audit area) in the following entities: (insert entities). We understand your office can provide information for the following areas (insert areas). Since this is an enterprise-wide review, we will be requesting similar data from (insert organizations).

We request that each of the above entities be forwarded as separate files and that three files be created for each entity, to reflect all transactions captured by the systems for (insert audit period).

To meet our planned review objectives and timing, the data must reflect all transactions captured by the system up to and including (insert audit period).

To facilitate our CAAT analysis, the data should be given to us in the form of flat files and placed in "(provide location)" on Drive F.

Together with the data, please document the file record layouts, including the skip length, if applicable and record lengths. For each fields, document the field name, field start position, field length, data type, and formats of numeric and date items. We also require documentation of key control information, including the transaction cutoff date and time, the number of records in the file and control totals of numeric fields where applicable.

This represents our initial data request associated with this audit. We anticipate that it may be necessary to make more data requests, but we will endeavor to keep these requests to a minimum.

We appreciate your assistance with data access, and look forward to working with you during the course of our audit.

If you have any questions please do not hesitate to contact me (*********), or at.

17.11 Specimen Format of a CAATs Working Paper

1.	Client and Location		-
2.	Review Period		-
3.	Audit Area	-	
4.	Work Program		-
5.	Name of Auditor and Team Leader	-	
6.	CAAT performed on		-

Documentation of CAATs

7.	Documentation of Source Data	-
	Source Reports	-
	Data File Names and Type	-
	Data Fields relevant for CAAT	-
	• Data Field Name, Type and Length	-
8.	Input Reconciliation	-
9.	Details of CAAT analysis performed	-
10.	CAAT Results	-
11.	Conclusion and Recommendations	-

Continuous Auditing Using IDEA

Continuous Auditing (CA) with IDEA

18.0 Introduction

There is a thin conceptual line between Continuous Monitoring and Continuous Auditing.

"Continuous monitoring refers to the management's responsibility to assess the adequacy and effectiveness of controls in an organization. Continuous Assurance is when auditors perform continuous control and risk assessment (i.e. continuous auditing) and evaluate the adequacy of management's continuous monitoring activities."

Continuous auditing changes the nature of evidence, timing, procedures and effort involved in modern day audits.

Transactional verification, compliance verification, estimate verification and judgment verification are the four levels of continuous audit.

18.1 Benefits of Continuous Auditing

Continuous Control Assessment provides financial executives, business process managers, and risk and compliance officers with independent and timely assurance over real time controls. Intelligent technology –enabled analytics like transaction validations and analytical reviews accompanied by drill down capabilities help discover the cause of the control failures of organizations.

Some of the key benefits of CA are:-

- Less Travel as CAATs can be deployed centrally on central data reservoirs.
- Low Staff and allied audit resource requirement.
- Low IT Hardware resource utilization like network and bandwidth as tasks can be scheduled after office hours, on holidays and weekends.
- High user satisfaction as automated tasks fired in non-peak hours give quicker results than queries fired in peak business hours.

- User prestige and recognition as CCM is the final mature stage in the use of CAATs, which is preached rather than practiced worldwide. A true CCM user gains unimaginable benefits and recognition in professional peer circles.
- Analytic monitors like Embedded Audit Modules and Statistical Tools observe material events, trigger alarms, drill down to grass root levels, facilitate data integration across multiple platforms and allow repetitive tests at much lower costs.

18.2 Implementing Continuous Auditing

Planning

- Defining the objectives for Continuous Auditing
- Obtaining Senior Management Support
- Ascertaining the degree to which Management is performing its Role
- Identifying and prioritizing areas to be addressed
- Performing a type of continuous auditing Identifying key information systems
- Understanding the underlying business processes and application systems

Executing

- Selecting and purchasing analysis tools
- Developing and maintaining auditor analysis skills and techniques
- Cleansing and preparing the Data
- Assessing the data integrity and reliability
- Indentifying crucial control points, and control rules
- Designing a technology-assisted approach to test controls
- Identifying deficiencies

Reporting

- Identifying control deficiencies
- Identifying increased levels of risk

- Prioritizing results
- Initiating suitable audit response
- Making results known to Management
- Managing results by tracking
- Reporting
- Monitoring
- Providing follow-up
- Evaluating results of actions taken.

18.3 Sample Case Study on Continuous Auditing using IDEA

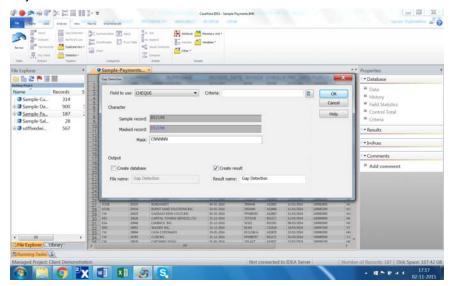
18.3.1 Start IDEA



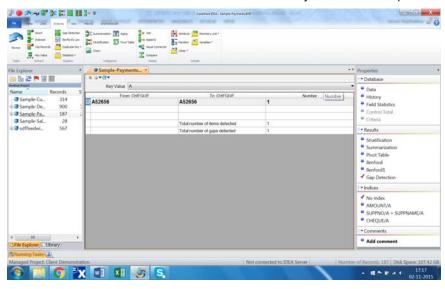
18.3.2 Record Macro (Control R)

				1	Tanti Lanta Lanta								
le Explorer			1	Sample-Pay	ments *	1							Properties
B 2 P 6				POSTED BY	SUPPNO	SUPPNAME	INVOICE DATE	INVOICE	CHEQUE	PAY DATE	PURCH ORDE	AL +	· Database
		-16		454	20500	AASAALAI DESILIN INC	11-12-2014	81344	852538	87.85.0004	200082500	67	- Database
Atop Project		_	2	MSA	20000	ACKLANDS - GRAINGER INC	11-52-3054	15212018	95223t	17/65/2004	230005400	BC I	✓ Data
ame F	Records	5	3	SOUR	E1400	AG X2YEFBA	11-12-2014	34835	C31000	25/93/2014	100095800	K.	
Sample-Cu_	314		4	KSA	33830	DAW INC.	31-12-2014	256	8523M	05/85/2054		18.	History
		10	5	0000	\$2203	ANKE KELDE ORGANECS INC	31-12-2014	A8 3285 M	C\$102	25/03/2004		10	Field Statistics
Sample-De_	900	- 1	6	ERC.	20128	BETOMITTEN GMEH	21-13-2014	21569	852204	99/63/2064		WF .	
Sample-Pa	187	1	7	MSA	20045 20029	RECUTERE CARMEN	31-12-2014	21569 634-567	852207 A52637	35/03/2014		HR. Vite	Control Total
Sample-Sal_	28	- 1	89	MSA	20129	BLENCH BOEDO BENICH PUBLENENG INC.	21-01-2014 01-01-2014	634-567	A52637 A52638	64/01/2014 13/01/2014		995	Criteria
		- 16	10	SOUR	20133	EDHOP PUBLISHING INC.	15-01-2014	N79646C-121		16/02/2014		18.	Contra Co
sdffixedwi	567	- 1	11	MSA	30.04	BOREAL LABORATORIES LTD	09-01-2014	AZ278	A53840	60/60/2004		10	I = Results
		- 1	12	MA	20252	QUESS INC.	25-01-2014	81340	A52842	05/02/2054		RC I	THEOREM
		- 1	12 13 14	KSA	39403	BADDON	17-01-2014	#71004A	A52843	04/52/2004		CB	Stratification
			14	CW	20434	CITEZEN INC.	10-10-2014	10000 A	A52544	84/62/2004		86	
			15	SOU8	20508	FOSSE INC.	18-01-2014	200139	A52645	64/62/2014		+8.	Summarization
		- 6	16	CW	20508	FOSSE INC.	06-01-2014	130729	A32846	04/02/2004	230062900	18.	Pivot Table
			17	CW	20532	INTERNATION AND	05-05-2014	000496C/W	A52647	04/02/2014		44.0	
			18	KSA .	20532	MERNER KD	10-01-2014	2525 8144	A52849	06/02/2054		W	Benford
		- 14	19	CW	20532 20535	HERER KG	13-01-2014	147905CTR	852526	13/62/2004		HR. 162	Benford1
		18	20	1008	20515	JOHN PETERSON BURKHWOT	08-01-2014 12-01-2014	34876/50 C5 - 563 -97	852117 A52656	33/82/2014		42	
		- H	21	CW	20515	EURO-CALOT	20-05-2014	FR-ME132	A52059	08/10/2014		18.	* Indices
		- 10	66	CN	20535	BURKHARDT	01-01-2014	117-2287	A52000	35/52/2014		18 L	District in
		- 11	24	SOLIE	20536	IONN PRTIRSON	02-01-2014	IN 6425 97	A52662	35/82/2004		CV.	 No index
			25	10UB	20535	BURKHARDT	15-01-2014	84 393817	A52963	33/02/2004		80	
			26	KSA	29535	BUROHARD1	31-12-2004	1042569	A52664	11/02/2004	180083400	18.	AMOUNT/A
			27	SOUR	20535	BURIONARD1	04-01-2014	7039446	A52665	11/02/2014	100063800	WE	SUPPNO/A + SUPPNAME/A
			19011223455672890012334#	5018	20154	BLRINT SAND SOLUTIONS INC.	10-01-2014	2503340	A52666	11/62/2654		HI	Sourcester & Sourcester
			29	CW	20625	CADEMIN MON COCO INC	52-01-2014	PP1486755	A5265	11/62/2014		WE -	· Comments
			30	DES.	29628	CARITAL TOWING SERVICES LTD	31-12-2014	1573298	852217	12/80/2004		+8.	
		16	31	KSA DDC	2068	IUMINOX INC. SKAGIN INC.	11-12-2014 31-12-2014	91322 81345	#52580 C3068	06/55/2014 34/65/2014		18.	Add comment
		- H	22	CN	20655	CASA CORONADO	19-01-2014	\$1,945 \$7,5238/2	A52670	13/02/2014	230080500	10.	
			23	CW		COLENC	31-12-2014	201008787	852572	01/05/2014	230089900	x8 _	
		1	29	CW	200.00	OHENNE KINS	21.01.2014	258.627	ACRON	13/07/0814	130090100	10. **	
File Explorer Ub	srary/		<	2000		m			CODE:			÷ .	
Running Tasks		stratio	-	X	S	S		Not cor	nected to	IDEA Server	N	umber	of Records: 187 Disk Space: 107.

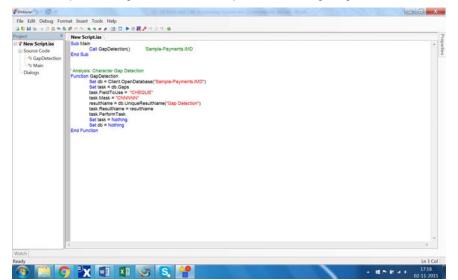
18.3.3 Run sample IDEA Function - Gap Detection on Cheques as per the Objective



18.3.4 Arrive at Result and Review in terms of the Objective. Follow Up



18.3.5 Stop recording Macro. IDEA Script for Function gets generated



18.3.6 Save IDEA Script as EXE File

New Ctrl+	N / 10 0	
Open Ctrl+	0	
Save Ctrl+ Save As	S apDetection() Sample-Payments.IMD	
Package for CasaWarn Monitor		
Send	acter Gap Detection	
Close	= Clent.OpenDatabase("Sample-Payments.IMD")	
Build New Script.exe	ix = db.Gaps	
Print Otrl+ Print Setup	ame = db.UniqueResultName("Gap Detection") esultName = resultName	
Page Setup	erformTask , jk = Nothing	
Recent IDEAScripts Exit	= Nothing	
100		
_		
		In1

18.3.7 After Review of Result, delete or archive

ran an a	Gap Certecture Barrhan Exclam Duplication Airy * Stationities * Exclama	Det Serverseter De Aproposet Serverseter De Aproposet Ap	Image: Image	
e Explorer		Sample-Payment	8	* Properties
1 2 P G III		& Export to 2 3- 2;	Filter	• Database
the Privat	-	Database	Date User	Data Data
ime Record		angle Fayment.MD		✓ History
	14	Import from Excel	31-10-2015 09:50:00 Jairam Rajshekhar	 Field Statistics
	00 1 87 2	a Index Database	31-10-2015 15:57:00 Jairam Rajshekhar	Control Total
	8/ · 28	a Stratification	31-10-2015 15:57:00 Jairam Rajshekhar	· Criteria
	67	III Index Database	31-10-2015 16:00:00 Jairam Rajshekhar	· Results
Paning and		a Summarization	31-10-2015 16:00:00 Jairam Rajshekhar	
		Create Pivot Table	31-10-2015 16:01:00 Jairam Rajshekhar	Stratification
		∃ Benford's Law	02-11-2015 14:23:00 Jairam Rajshekhar	Summarization
		a Benford's Law	02-11-2015 17:11:00 Jairam Rajshekhar	Pivot Table
		∃Index Database	02-11-2015 17:17:00 Jairam Rajshekhar	Benford Benford1
		# Gap Detection	02-11-2015 17:17:00 Jairam Rajshekhar	Gap D
		and the second second		herate_
				* Indices Deste Tault
				No inc passe at heads
				AMOUNT/A
				SUPPNO/A + SUPPNAME/A
				CHEQUE/A
				· • Comments
III File Explorer ()brary				Add comment
de Explorer L Ubrary		-		-

18.3.8 Close IDEA

18.3.9 Go To Task Scheduler in Control Panel

Deganize • 🔳 Op	en					SI -	11	0
Favorites	Name	Date modified	Type	Size				
E Desktop	Component Services	14-07-2009 10:27	Shortcut		2 KB			
15 Recent Places	Computer Management	14-07-2009 10:24	Shortcut		2 68			
OneDrive	Data Sources (ODBC)	14-07-2009 10:23	Shortcut		2.65			
a Downloads	desktop.ini	21-10-2014 13:09	Configuration settin		2.108			
	Event Viewer	14-07-2009 10:24	Shortcut		2.68			
Libraries	S. iSCSI Initiator	14-07-2009 10:24	Shortcut		2.103			
3 Documents	Local Security Policy	21-10-2014 13:09	Shortout		2.68			
2 Music	Performance Monitor	14-07-2009 10-23	Shortcut		2 KB			
5. Pictures	- Print Management	21-10-2014 13:09	Shortcut		2 KB			
Videos	Services	14-07-2009 10:24	Shortcut		2.168			
	System Configuration	14-07-2009 10:23	Shortcut		2.88			
Homegroup	P Task Scheduler	14-07-2009 10:24	Shoricut		2.08			
	Windows Firewall with Advance Schedule	computer tasks to run automa	tically rout		2 KB			
Computer	R Windows Memory Diagnostic	14-07-2009 10:23	Shortcut		2 KB			
	55 Windows PowerShell Modules	14-07-2009 11:02	Shortcut		3.68			
Network								
Task Sched	aler Date modified: 14-07-2009 10:24 Dat	e created: 14-07-2009 10:24						
+ Shortout	Size: 1.23 KB							

18.3.10 Add Scheduled Task

sk Scheduler (Loc	Task Scheduler Surrerio	Create Talk			×		Actions
Task Scheduler Li	Overview of Task Sch	10.100				11.14	Task Scheduler (Local)
			gers Actions Conditions Settings		-	20	Connect to Another Com
	You can use Tas specify. To begin	Name:	IDEA				Create Basic Task
		Location:	`				 Greate Task
	Tasks are stored Task Scheduler L	Author:	JairamRajshekha\Jairam Rajshekhar				Import Task
	Task Scheduler	Description:					💷 Display All Running Task
						E 11	II Enable All Tasks History
	Task Status					1 E	AT Service Account Confi
	Status of tasks that I						View
							ni Refresh
	Summary: 0 total - 0						Help
		Security op	tions				
	Task Name	When runn	ing the task, use the following user accourt	t		E	
	THESE THE THE	JairamRajsh	ekha\Jairam Rajshekhar	-Change User or	Group	11	
		Bun only	when user is logged on				
		💿 Run whe	ther user is logged on or not				
		Do no	ot store password. The task will only have a	access to local computer resources.			
		🔲 Run with	highest privileges				
		E Hidden	Configure for: Windows/8	0 7. Windows Server™ 2008 R2	•		
	Active Tasks						
				OK	Cancel	Π.,	
		_					
	Last refreshed at 02-11-	2015 17:19:46			Refres	h	

k Scheduler (Loc	Task Scheduler Summa	Create Talk			1	Actions
ask Scheduler Li			See After		- 14	Task Scheduler (Local)
	Overniere of Tark Sche Processor of Tark Sche specify: 70 keps Tasks are stored Task Schedder L Task Status Status of tasks that In Summary: 0 total - 0 Task Name	General 1 When y	You must specify what action this task will perform. Action: Start a program			Connect to Another Comp Create Issic Task Import Task Display All Running Tasks Fanile All Tasks History AT Service Account Config View Refersh Help
	Active Tasks Last refreshed at 02-11-	New 2015 17:19:4		aricel .	+ esh	

18.3.11 Select the program we want Windows to Run – IDEA

18.3.12 Decide on the frequency of execution and the name Task

Task Scheduler (Loc	Task Scheduler Sume	Edit Tripper			Actions
Task Scheduler Li	Overview of Task Sc You can use Taspecify. To be Tasks are store Task Schedule	Begin the task: On a schedule Settings One time One time Start: Daily Weekly Monthly Monthly	□ * 172031	•	Test Meader Savell Connect to Another Comp Create Basic Task Create Task Import Task Display All Running Tasks El Enable All Tasks History
	Task Status Status of tasks the Summary: 0 total Task Name	Stop all running tasks at end o		for a shareboy of [1 day	AT Service Account Config View Refresh Help
		Stop task if it runs longer than: Expire: 02-11-2016		Synchronike azuss bits zones	
	Active Tasks	11-2015 17:19:46		OK Cancel Refresh	-

18.3.13 Set Date and Time for execution of the Task	
---	--

k Scheduler (Loc ask Scheduler Li	Task Scheduler Sume	ter have	Actions
	Overview of Task Se You can use To specify. To be Tasks are store Task Schedule Task Status	Begin the task: On a schedule Settings One time Start: 02-11-2015 Start: 02-11-2015 Synchronize across Daily Weekly Monthly	The Underson Stand) Connect to Another Comp Connect to Another Comp Create Task Transt Task Transt Task. Display All Running Tasks Si Enable All Tasks History AT Service Account Config
	Status of tasks the Summary: 0 total Task Name	Advanced settings Codely task for up to (random delay): 1 hou Repet task every: 1 hou for a duration of: 1 day for a duration of: 1 day for a duration of: 1 day	View S Refresh Holp
		Epper: 02-11-2016 02-11-2016 02-11-2016 02-11-2016 Image: Specific control of the control	
	Active Tasks	OK Cancel	-

18.3.16 Click Apply and OK

18.3.18 The Task will run automatically at the scheduled date and time

18.3.18 Review the new result, follow up and review.

	9 Sample-Payments			 × Properties
김 🔁 🔊 🟴 🤐 🗑	8			II - Database
	Key Value			Data
A Files	and a state of the			History
Name	From CHEQUE	To: CHEQUE	Number	Pield Statistic
Sample Authorization	E C51016	C51016	1	📾 Control Total
😇 Sample-Bank Transactions		-100.00010		Criteria
Sample Customers				I - Results
Sample Cretaied Sales		Total number of items detected	1	
Sample Inventory		Total number of gaps detected	1	Gap Detection1
Sample Payments Sample Sales Represent	a second s			I - Indices
				 PAV_DATE/A + AUTHUA INVOICE/A + AMOUNT/A
				CHEQUE/A
				CHEQUE/A

Documentation of CAATs

e Explorer 4	Sample-Payments			* ×	Properties
📮 😋 🥭 🕊 🖋 🗑	8 1 2 2* 2: Fiter				II - Database
	Database	Date	User	-	Data
		Sample Payments IMD			J History
CA Tiles	C:\Program Files\IDEA V 7.3\Sa		0		Field Statistics
Name	(+) Index Database	11/04/2008 - 17:58	Javram		Control Total
Sample Authorization	Index Database	11/04/2008 - 18:03	Jayram		Criteria
Sample-Bank Transactions Sample Customers	Benford's Law	01/09/2008 - 16:29	Javram		
Sample Octailed Sales	Tag History	07/11/2008 - 12:17	Javram		II - Results
Sample Employees	i Index Database	07/11/2008 - 12:17	Jayram		Gap
Sample-Inventory	Add/Replace Tags	07/11/2008 - 12:17	Jayram		Detection1
Sarcle-Paymerits Sarcle-Sales Ptepresent Sarcle-Supples Sarcle-Weblog	(+) Tag History	07/11/2008 - 12:17	Jayram		S2X322349W
	Tag History	07/11/2008 - 12:19	Jayram		II - Indices
	Add/Replace Tags	07/11/2008 - 12:19	Javran		/ No index
	+ Tag History	07/11/2008 - 12:19	Jayram		AUTH/A
	Benford's Law	11/11/2008 - 17:12	Jayram		PAY DATE/A
					+ AUTHIA
	Index Database	07/01/2009 - 19:48	Jayram		INVOICE/A +
	Gap Detection	07/01/2009 - 19:48	Jayram		AMOUNT/A
	Gap Detection	07/01/2009 - 19:53	Jayram		CHEQUE/A
	Gip Detection	07/01/2009 - 20:04	Jayran		II - Comments
	Control Field:	No Control Total			
	Control Total:	No Control Total			Add comment
	Field used:	CHEQUE			Comment
	Mask:	CHNNNN			
	Result name:	Gap Detection1			
	IDEAScript Code:	"Gap Detection Set task = db.Gaps task.PeldToUse = "CHEQUE task.Mask = "CNMNN" resultName = db.UniqueRes task.ResultName = resultNa task.PerformTask	ultName("Gap Detection1")		

Chapter 19 Advanced Statistical Methods in IDEA

19.0 Introduction

IDEA's Advanced Statistical Methods can be used by analysts as a component of a continuous monitoring set of routines.

This chapter covers the use of two key statistical methods in IDEA and provides examples of how these methods can be used in a wide variety of settings:

- Correlation
- Trend Analysis

While calculations for these statistical methods can be done by other software programs that havestatistical routines including linear regression, or by using the formula for the Pearson product momentcorrelation, or the formulas found in statistics books for the slope and intercept, the unique feature inIDEA is that these calculations can be done on a group-by-group basis.

In IDEA's Advanced Statistical Methods module, a group is referred to as an audit unit. While the termaudit unit may be used to mean location, division, franchisee, or some other reporting entity, the testscould be run for a task that is not an audit as such, but perhaps an analysis to screen for investmentprospects, or to screen for significant data errors, or to report results for a professional or academicresearch study.

19.1 Correlation

Correlation is a statistical technique being used by auditors and other data analysts to compare thepatterns in numerical values in two data sets. A high level of correlation (that is, a correlation scoreclose to 1) means that high values in one data set are matched with high values in the second dataset, and conversely, low values in one data set are matched with low values in the second data set. The correlation score is a single number that tells us how good this match is. Correlation scores can range from -1 to +1.

Correlation was first used in auditing as a component of a continuous monitoring set of audit routines. For example, the average level of sales for each month of the year were calculated for therestaurants owned and franchised by a company, and then the correlation between each individual restaurant's sales and the average sales was calculated. A low correlation score for any restaurantshowed that the seasonal pattern for that restaurant deviated from the norm. Another example wasthe analysis of election data where it was believed that each candidate should receive the samepercentage of votes in each location. In a simplified setting assume that candidate #1 received 60percent of the votes, candidate #2 received 30 percent of the votes and candidate #3 received 10percent of the votes in an election. Correlation was used to measure how well this 60:30:10 ratioapplied to each voting location. Low correlations for any location suggested that (a) either the candidates were more or less well favored in those locations, or (b) that there was an error in thevote counts.

The applications of correlation are not limited to these types of data sets but can also be used tocompare both monthly and annual financial reports to those of prior periods. In fact, the techniquecan be used whenever a set of numbers must be compared to some norm. For example, theelectricity consumption of all schools in a school district could be compared to the normal (average)pattern, or the bad debt write-offs on a monthly basis could be compared for each branch in a bankto the average pattern for the year.

A database output or Result output could be generated from the Correlation analysis. The Resultoutput can be viewed in a grid or as a graph. The graph would display the correlation scores rankedfrom smallest to largest allowing one to see the profile of the correlation scores (that is, theproportions of low and high scoring audit units). The graphs can be further customized to suit thetask at hand.

15.1.1 Correlation Case Study

Background

An internal auditor engaged by a Manufacturing Company to review the overtime payments to staff and workers has been provided with two excel files – one – plant shop centre, week wise overtime and the other – plant shop centre, week wise production.

By linking the two files together in IDEA, the auditor would like to employ Correlation to check for specific plant shop centres where production is on a decline whereas overtime is on a rise. Correlation is an effective visualization technique to identify high risk audit units in a population. So in the given example we would like to identify high risk plant shop centres from an overtime payment control point of view amongst all plant shop centres in the Company.

Step 1 – Import of Data Files

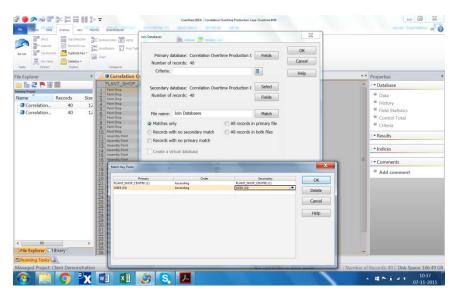
Import the two excel worksheets into IDEA using Home – Desktop – Import Assistant

PD Andrew Point Andrew New Month Search Andrew Tract Date Date Andrew Date	Depert *		×	a ()
File Explorer * Comparison of the second se	Select sheets to import:	515	OK Indude All Geer All Help	Properties
	3 Part Shop 3 4 Part Shop 4 5 Part Shop 5 6 Part Shop 6 7 Part Shop 7 8 Part Shop 7 8 Part Shop 8 9 Part Shop 8 9 Part Shop 9 10 Part Shop 9 11 Assembly Port 1	800 601 1975 1975 1975 1975 1975 1975 1975 197		= rindices ■ * Comments
STEPSore(_tbray) Running task	1 🐼 S. 🗡	Not co	onnected to IDEA Server	Disk Space: 106.49 GF

Step 2 – Join the Data Files

Open the Overtime file as the Primary database in IDEA. Click on Analysis – Relate and Join. Select the Production file as the Secondary database in the Join dialog box. Enter the matching key fields as Plant Shop Centre and Week between the Primary and Secondary files.

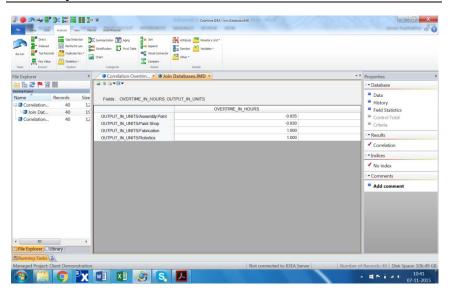
Advanced Statistical Methods in IDEA



Step 3 – Apply the Correlation

Click on Analysis – Explore – Statistics and Correlation. Select the Overtime in Hours and Output in Units as Fields to Correlate. Select Plant Shop Centre as the Audit Units. This will give correlation scores for each Plant Shop Centre.

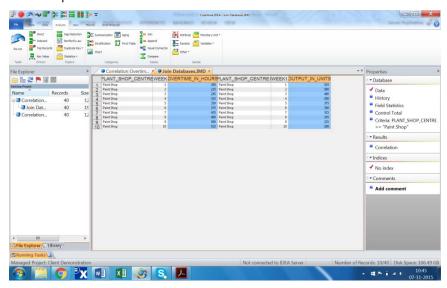
An Propagation in the second s	Fields to Correlate	Audit Units	ОК	
Rey Value Statistics -	Include all fields	Select audit unit field Audit unit field:	Cancel	
ska Extract Explore	VEEK	PLANT_SHOP_CENTRE	Help	
Explorer •	WEEK1	There are 4 units for the audit field in the database.	Trop	• × Properties
	V OUTPUT_IN_UNITS	Select the units you wish to use to calculate the correlation.		▲ Database
me Records Size	12	correlation.		Data Data
Correlation 40 12		Include all audit units		History
Join Dat 40 15				Field Statistics
Correlation 40 12		Audit Units Number of Records Assembly Point 10		Control Total
	Fabrication 10		Criteria	
		V Paint Shop 10 V Robolics 10	Results Indices	Results
				- Indian
				- Indices
				Comments
				Add comment
		Output Create result		
		Result name:		
		Correlation		
		Create database		
		File name:		
		Correlation		
le Explorer / Ubrary				
unning Tasks 3		and the second s	_	



Conclusion

The Correlation Score for Paint Shop is -0.930 indicating an inverse relationship between the Overtime in Hours and Output in Units.

A closer look at the weekly overtime and output figures for Paint Shop will reveal a falling output from week 1 to week 10 with a rising overtime for the same period. This would allow the internal auditor to focus his/her attention on reviewing, verifying, investigating the detailed overtime payments for the Paint Shop from week 1 to 10.



19.2 Trend Analysis

Trend Analysis is a statistical technique that is being used by auditors and other data analysts tocalculate the trend of data over time and to forecast values into the future on the assumption thatthe trending pattern will continue. The Trend Analysis technique used in IDEA's Advanced StatisticalMethods module is based on linear regression using the method of least squares. This technique willfit the best fitting straight line to the data and will forecast ahead using this straight line. If the data has a seasonal component (higher values associated perhaps with certain months, or days of theweek, or times in the day) then the Time Series statistical method should be used instead of Trend Analysis.

Trend Analysis can be used in auditing as a component of a continuous monitoring set of audit routines. For example, it could be used to test for unintentional or intentional errors in the reportedsales numbers of all the restaurants owned and franchised by a company. One possible indicator oferrors could be a decreasing sales trend for a restaurant when the average, over all locations, showsan increase in sales. Trend Analysis (or linear regression) could be used to quantify the trend in thesales for any location. A negative trend would be an indicator of possible problems with the salesnumbers.

Another use for Trend Analysis could be to simply provide a forecast for various accounting entities a continuous monitoring environment. The forecasts could then be compared to the actual numbers and where the forecasts differed from the actual numbers, it would be an indication that these numbers have deviated from their past trend suggesting that some change has taken place.

Trend Analysis can provide the answers to many types of monitoring questions such as:

- For a franchisor, which locations have decreases in sales compared to the amount extrapolated from past trends?
- For an investment analyst, which investments show a decreasing trend in market values?
- For a school district, which schools show a supplies usage that exceeds the amount extrapolated from past trends?
- For a cruise ship or hospital, which ships or departments show a high level or an increasing trend in employee sick days?
- For a courier service, which locations show the largest increases in fuel expenses?

 For a college, which departments show decreasing numbers of majors?

15.2.1 Trend Analysis Case Study

Background

An internal auditor engaged by a Manufacturing Company to review the overtime payments to staff and workers has been provided with two excel files – one – plant shop centre, week wise overtime and the other – plant shop centre, week wise production. The historical data in both the files is for 10 weeks of the year.

By linking the two files together in IDEA, the auditor would like to employ Trend Analysis to check for the 11th week forecast for each plant shop centre. This will allow the auditor to have a ready comparison between the 11th week forecast and 11th week actuals during the course of the audit.

Step 1 – Import of Data Files

Import the two excel worksheets into IDEA using Home – Desktop – Import Assistant

te Esplorer Seconds Size Aarme Records Size Production Produc	Ploperties	First row is field names Indude All Indude All Clear All	eets to import: rtime	Inport N	9	
Side diverts to import: Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many numeric cells a 0 Import many nu	• Database • Data • Data • History • Field Statistics • Control Total	First row is field names Indude All Indude All Clear All	rtime		* Vely total	
Concerne Induct All Import mere Production Import mere/ numeric cells as 0 Import mere/ numeric cells as 0 Production Import mere/ numeric cells as 0 Impo	Data History Field Statistics Control Total	First row is field names Indude All Indude All Clear All	rtime			
Records Size Records Size Records Size Production Profere rofere Pro	History Field Statistics Control Total	Import empty numeric cells as 0 Include All Clear All			100.000	
Cannot Holp 1	Field Statistics Control Total				Records Size	
Prefere: 1000000000000000000000000000000000000	Control Total	Cancel				
Preview: 1 Preview: 1017PU.M 1 Previdage 2 300 2 Previdage 2 300 3 Previdage 3 600 3 Previdage 3 600 5 Previdage 3 600 5 Previdage 4 305 6 Previdage 7 215 6 Previdage 9 215 7 Previdage 9 215 8 Previdage 9 215 9 Previdage 9 Prev						
Preview: 1 Preview: 1017PU.M 1 Previdage 2 300 2 Previdage 2 300 3 Previdage 3 600 3 Previdage 3 600 5 Previdage 3 600 5 Previdage 4 305 6 Previdage 7 215 6 Previdage 9 215 7 Previdage 9 215 8 Previdage 9 215 9 Previdage 9 Prev	" Criteria					
PARD 5002, MIK DOTPTIN_L 1 Fare Stap 533 2 Fare Stap 543 4 Fare Stap 548 4 Fare Stap 548 5 Fare Stap 533 2 Fare Stap 538 3 Fare Stap 538 2 Fare Stap 538 3 Fare Stap 538 1 Fare Stap 538 2 Fare Stap 538 3 Fare Stap 538 4 Fare Stap 538		Help				
1 Pare Dage 1 101 2 Pare Dage 2 500 3 Pare Dage 3 600 3 Pare Dage 5 600 4 Pare Dage 6 600 5 Pare Dage 6 500 6 Pare Dage 6 500 7 Pare Dage 6 300 8 Pare Dage 6 300 7 Pare Dage 6 300 8 Pare Dage 6 300 7 Pare Dage 6 300 7 Pare Dage 9 300 8 Pare Dage 9 200	Results					
Term Term <th< td=""><td></td><td>Î.</td><td>Paint Shop 1</td><td></td><td></td><td></td></th<>		Î.	Paint Shop 1			
4 There Using 4 100 B 3 Aver Using 5 325 6 Aver Using 6 328 7 Aver Using 6 328 8 Aver Using 6 328 9 Aver Using 6 328 9 Aver Using 6 238 30 Aver Using 6 235 30 Aver Using 6 235	I Indices					
6 Free Dags 6 356 7 Free Dags 7 315 8 Free Dags 7 315 9 Free Dags 9 325 10 Free Dags 9 325	- Comments		Paint Shop 4			
B Part Stop B 245 9 Part Stop 9 225 20 Part Stop 10 200			Paint Shop 6			
9 Paint Shop 9 225 20 Paint Shop 20 200						
			Paint Shop 9			
11 Assembly Point 1 275	•	-				
	_					
Input file name: C:\Users\Jairam Rajshekhar\Documents\My IDEA Documents\IDEA Prc	el la					
Output file name: Correlation Overtime Production Case		duction Case	file name: Correlation Overtime P			
				L		
					P.	111
le Explorer (III tibrary /						
unning Tasks 🗟					Library	e Explorer

Step 2 – Join the Data Files

Open the Overtime file as the Primary database in IDEA. Click on Analysis – Relate and Join. Select the Production file as the Secondary database in the Join dialog box. Enter the matching key fields as Plant Shop Centre and

Week between the Primary and Secondary files.

Pice Data Notes Ven Notes Sectorial State Image: Sectorial State Image: Sectorial Sta	Primary database: Correlation Overtime Production C Fields Number of records: 40	OK Cancel	a (
File Explorer • Correlation	Criteria:	Help	• × Properties
Correlation. 40 C	Secondary database: Correlation Overtime Production (Select Number of records: 40 File name: Join Databases Match Matches only All records in primary file Records with no secondary match Records with no primary with no prima		• Database Ota Ota Ota Hotoy Feld Statistics Control Tota Criteria results • Indices
17 18 19 20	America Anarding CENTE(C) Acceding ULLS, No Acceding ULLS, No Acceding	OK Delete Cancel Help	Add comments

Step 3 – Apply the Trend Analysis

Click on Analysis – Explore – Statistics and Trend Analysis. Select the Overtime in Hours as the Trend Field. Select Plant Shop Centre as the Audit Units. Click on Generate Forecasts for 11th week. This will give the trend forecast for the 11th week for each Plant Shop Centre.

Texture Select them field: Audit Units Forecasts VEEK ING: IN: INCLOS Audit Units Forecasts Forecasts VEEK ING: IN: INCLOS Audit Units field Forecasts Forecasts Inform Mark and fund: Inform Numbers of forecasts: Inform Mark and tunits for the selected audit units Numbers of forecasts: Inform Mark and tunits Inform Audit Units Numbers of forecasts: Inform Viet Hube Audit Units Inform Audit Units Number of forecasts: Inform Viet Hube Number of forecasts: Inform Viet Hube Number of forecasts Inform Viet Hube Number of forec	10 Ca	
Reference Fields IN Reference Fields IN Reference Fields Overthele State		ics al
Create result Result name: Trend Analysis Create Trend Analysis Create Forecast datato Create FMPC database FFIe name:		ent

The same the balance has	Macros SmartAnalyzer	ALTERNATION MANAGEM	ALC: NO OTHER					Salara Salaraha	0 6
	Summarization 📷 Aging	Tu Jon 2 Attrib	de 🌁 Honetary Unit *						
Indexed Bill Benford's Law	Stratification	- Append	en 🚰 Variables *						
e run 📑 Top Records 📑 Duplicate Key +	Ghert	Voual Connector							
asla Estract Explore	Categorize	Relate	Sample						
e explorer	Correlation O	Verti ×					• ×	Properties	
i 🚰 🧩 🎮 🔯 🏢	■ 8 ☆ • 3 •							• • Database	
Iame Records Siz Ocrelation 40 1 Join Dat 40 1	e Trend mo c Start fr z Foreca						Æ	 Data History Field Statistics Control Total Criteria 	
	🔉 🗈 🖾 · 💅 ·	🎭 🖶 🕈 🔂 🗱 🗱						· · Results	
	600							 Trend Analysis 	
	-							I ▼ Indices	
	S00							✓ No index	
	SUICH 400							PLANT_SHOP_CENTRE/ WEEK/A	+
	300					•••		· Comments	
	300 300 0.05						- 1	Add comment	
	8 100							Fidd Comment	
	0			• •					
	0 1	2 3	4 5	6 7 Index	8 9	10 11			
m +									
File Explorer	-	-	Actual data 🔸	Trend fits Fore	casts				
naged Project: Client Demonstrati	Marce Snettingtyse		CaseWare IDEA - Trend	Not con	nected to IDEA Serv	rer Nun		Records: 40 Disk Space: 10 10: 07-11: 07-11:	8 2015 ×
Anged Project: Client Demonstrat Client Core Angel Core C	W X Marce Smethodyper Marce Smethodyper Marce Smethodyper	Ped States		Analysis FORECAST JUD	nected to IDEA Serv	rer Nun		10: 07-11:	8 2015 ×
Anged Project: Client Demonstrat Client Core Angel Core C	Nore Smethodyse	Ped Scher		Analysis FORECAST,JAID	nected to IDEA Serv	rer Num		10: 07-11:	8 2015 ×
Anade Project: Clern Demonstration	W X Image: State Strategy	Ped States	Press Columns Re	Analysis FORECAST,JAID	y Vertos	rer Nun	-	10: 07-11:	8 2015 ×
And Project Clern Demonstration	W X Image: Struct Windows Struct Windows Image: Struct Windows Struct Windows <tr< td=""><td>Pred Statistics Pred Pred Statistics Pred Statistics Pred Pred Pred Statistics Pred Pred Pred Statistics Pred Pred Pred Pred Pred Pred Pred Pred</td><td>Data Al Prezer Frido Columno Ri ysis FO ×</td><td>Analysis FORECASTEND The Groups Group Hexandre</td><td>Verkal I Verkal</td><td></td><td>-</td><td>10: 07-11: 07-11:</td><td>8 2015 ×</td></tr<>	Pred Statistics Pred Pred Statistics Pred Statistics Pred Pred Pred Statistics Pred Pred Pred Statistics Pred Pred Pred Pred Pred Pred Pred Pred	Data Al Prezer Frido Columno Ri ysis FO ×	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal		-	10: 07-11: 07-11:	8 2015 ×
Ander Project. Clerk Demonstration Clerk Clerk Demonstration Clerk Clerk Demonstration Clerk Demonstration Cle	Ant State	Profi Standar Profi	Visis FO × RECAST_F	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal		-	Properties • Database	8 2015 ×
Andread Project Clean Demonstration	W X Torrestore Arrow	Prof Structure Prof Structure Prof Structure Prof Trend Anal HOP_CENTRPF Proint 1	Precision of the second	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC	-	Properties → Database → Database	8 2015 ×
Anore Project Clern Demonstration Clerk Clerk Clerk Demonstration Clerk Clerk Cler	Kernel Kernel Kerne	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties Detabase Data Bistory Control Total	8 2015 ×
Arrow Construction Arrow Constructio	W X Image: Second sec	Ped Stands Ped Stands Town Ped Stands Ped Stands	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254	-	Properties → Database → Database	8 2015 ×
Arrow Construction Arrow Constructio	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties Detabase Data Bistory Control Total	8 2015 ×
Arrow Construction Arrow Constructio	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties Data Properties Protection Protection Properties Protection	8 2015 ×
Arrow Construction Arrow Constructio	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties Data Hittory Hittory Red Statistics Control Total Control Total Results	8 2015 ×
Angeod Project: Clean Demonstration Clean Clean	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	10: 07-11 07-11 Properties • Data • History • Field Statistics • Criteria • Criteria • Criteria • Criteria • Criteria	8 2015 ×
Proceed Project Clean Demonstrate Provide Project Clean Demonstrate Project Clean Dem	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 ×
Proceed Project Clean Demonstrate Provide Project Clean Demonstrate Project Clean Dem	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Database • Pied Statabase • Pied Statabase • Pied Statabase • No index	8 201
Arrow Construction Arrow Constructio	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 ×
Angeod Project: Clean Demonstration Clean Clean	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 ×
Angeod Porget: Clean Demonstration Clean Clean	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 ×
Angeod Porget: Clean Demonstration Clean Clean	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 X
Anged Poject Clerk Demonstration Constraints Constraint	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015
Orm Orm Orm Orm Orm Ormelation A O Orm Ormelation A O O	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 X
Anged Poject Clerk Demonstration Constraints Constraint	Kill	Put Status Put Status Dow Dow Put Trend Anal HOP_CENTRE Point 1 1 1 1	Press Down All Press Press Colorers Press Colorers RECAST_F 1	Analysis FORECASTEND The Groups Group Hexandre	Verkal I Verkal	RS_FOREC 285 254 569	-	Properties • Detabase • Data • Hield Statistics • Criteria • Reides • Criteria • Reides • Mo index • Comments	8 2015 X

Chapter 20

Do's and Don'ts while using IDEA

This Chapter contains our collective practical suggestions for 'Do's and Don'ts' with our decade plus experience and implementation learnings with \mbox{IDEA} –

S No	Function in IDEA	Usage Tips
1	Set Working Folder	Once you create a new folder, select the new folder and click on open. This will ensure you work inside the folder.
2	Import Assistant – Microsoft Excel	 Ensure that the first row in the MS- Excel file contains the field headers. Ensure that the file is closed prior to import. Check both the boxes 'First Row Is Field Names' and 'Import Empty Numeric Cells as Zero'.
3	Import Assistant – Print Report and Adobe PDF	When selecting a date field in the field editor (Report Reader) change field type from character to date and set mask as DD- MMM-YY if the date is like 02-MAR-15. Even though the date mask has been set as DD-MMM-YY, the date field will get displayed in the IDEA database based on the regional date settings on the laptop/PC which has IDEA. So if the laptop/PC has the regional date settings as MM-DD- YYYY, even though we set the mask in the Report Reader as DD-MMM-YY, the date will get displayed in the IDEA database as MM-DD-YYYY. While writing a date equation on this field you will always use DATE="YYYYMMDD" in the Equation Editor.
4	Index	Double-click on the field of choice once to sort on ascending basis i.e. Low to High.

S No	Function in IDEA	Usage Tips
		Double-click once again to sort on descending basis i.e. High to Low. This will only create a different view of the original database.
5	Direct Extraction	 While entering an equation in the Equation Editor please ensure that every condition has a field, arithmetical operator and value/field. Eg. PASS_AMT>BILL_AMT Multiple conditions to be connected with a logical operator like (.ANDORNOT.) Eg. (PASS_AMT>BILL_AMT).OR. (TOKEN_DATE>CHEQUE_DATE) Character condition strings to be placed in text encapsulates. Eg. NAME== "ABCD" Date condition values to be place in text encapsulates. Eg. CHEQUE_DATE ="20150126". Enter the values/strings correctly depending on how the data is appearing in the database. Do not leave any leading space in the text encapsulates. Eg. SCHEME="NREGA" is wrong. This will look for NREGA with a leading space in the Equation Editor for re-use later. If you get an error message while writing an Equation – "Incorrect Syntax". If you get an error message while writing an Equation – "Mismatch Field

Do's and Dont's for IDEA Software

S No	Function in IDEA	Usage Tips
		Type" – please check the compatibility of the field and value in a given condition. Eg. CITY="DELHI" is right. CITY=50000 is wrong and will give "Mismatch Field Type"
5	Criteria	After you enter a criteria under Properties (right-side of IDEA Screen), IDEA will give you a view of the original database. Click on File-Save As to save the view as an independent file. After File-Save As, visit the original database again, right click on Criteria and clear.
6	Duplicate Key – Detection	To be used when you are looking for duplicates in one or more field/s up to 8 fields. Eg. Same Controller, Same DDO, Same PAO, Same Function Head, Same Scheme, Same Agency Name, Same Sanction Date
7	Duplicate Key – Exclusion	To be used when you are looking for duplicates in one or more field/s up to 8 fields where an independent field is different. Eg. Same PPO Number Different Pensioner Name in Central Pension Accounting.
8	Gap Detection	Can be used to look for missing items only on fixed length/formatted numeric, date or character fields. Eg. If you are looking for missing cheque numbers in a cheque series where multiple cheque books are being used. Say 123456, 123457, 123459 and 256787, 256788 and 256789. Select cheque number as Fields to Use in the Gap Detection screen and set the mask as CCCNNN. Cheque number should be set as a character field using Field Manipulation. This will give missing cheques per cheque book series only and

S No	Function in IDEA	Usage Tips
		not within the cheque books which is invalid.
9	Field Manipulation	While appending a new field to any database in IDEA, append a virtual field as it allows for the equation to be changed at any time.
10	Summarization	 Used for arriving at a count/frequency/number of times. By selecting a field in 'numeric fields to total' on the right hand side of the summarization screen, you will get sum totals as well along with counts. When you have a database where the relationship between values like say GOODS RECEIVED AMOUNT and BILL AMOUNT are not one to one. For example it may be many goods received to many bills. In such a case summarize the data by Purchase Order Number or Token No and choose numeric fields to total as GOODS RECEIVED AMOUNT and BILL AMOUNT. Once you get the summary you can apply a direct extraction equation to check for BILL AMT SUM>GOODS RECEIVED SUM. This is not required where there is a one-to-one relation though.
11	Stratification and Stratified Random Sampling	Before you perform a Stratification, visit Field Statistics and observe the MIN, MAX and AVG statistics for the field on which stratification is going to be performed. This will help you in interval creation in the Stratification dialog box. Eg. MIN is Rs. 1000, AVG is Rs. 50000 and MAX is Rs. 500000 for a field TAX_AMOUNT. While performing Stratification create intervals as 0-50000, 50000-100000. Anything above

Do's and Dont's for IDEA Software

S No	Function in IDEA	Usage Tips
		Rs. 100000 will be an Upper Limit Exception (ULE). The ULE cases are A Class cases i.e. High Value. Click on the number of records in ULE in the Stratification result and extract the records to a file titled 'Class A'. Now you can perform a random sample on the 'Class A' file.
12	Top Records Extraction	 To be used to identify 'top records' or 'bottom records' in any database. Eg. Identify top 25 assesses in the tax collection report. Enter 25 as number of records to extract and perform top records for field 'TAX_AMOUNT'. If your objective now required you to capture top 25 assesses per ward, then follow the same method as above. In Addition choose 'WARD' as the field to Group.
13	JOIN	 While joining a primary and secondary database using File – Join in IDEA, ensure that both the databases have atleast one common key field. The key field need not have the same name, but it must have the same field type like 'TOKEN_NO' (Character) and 'TOKEN_NUMBER' (Character). The Matches only option in the Join, matches the first match from the secondary database always. If you have a situation where both the primary and secondary database have multiple records per key field. Eg. TOKEN_NO in BILL FILE has multiple records per TOKEN NUMBER and TOKEN_NUMBER in CHEQUE FILE has multiple records per TOKEN NUMBER, then none of the options in

S No	Function in IDEA	Usage Tips
		 Join will help you compare the BILL with CHEQUE details since it will only pick up the first matching cheque from the CHEQUE FILE. In such a case use COMPARE under File. COMPARE will summarize the multiple records per TOKEN NO in both the primary and secondary databases and you can choose to compare BILL AMOUNT with CHEQUE AMOUNT. If you have one transaction file like BILLS and another master file like Delegation of Financial Powers. Make the BILLS file the Primary database. Always choose the transaction file as the Primary database and apply the master details to the transaction files like Pension File April 2014 and Pension File March 2015, any can be chosen as Primary. But we take Pension April 2014 as Primary given it is earlier in time period than March 2015.
14	Edit – Re-Run Task	 Always use Edit-Re-Run Task to correct a mistake made or to re-run the task last performed on any database active on your screen. While performing functions for new objectives make sure you go back to the Original Database to avoid performing a new objective on the result of the earlier objective.

Chapter 21 New and Emerging Features in Data Analysis

"More than half of chief audit executives and directors believe that continuously assessing and identifying emerging risks will be the greatest challenge for internal audit throughout 2015 - but only one-third are highly confident in their ability to identify these risks."

 The Institute of Internal Auditors Audit Executive Center. March 2015. The Pulse of Internal Audit: Navigating an Increasingly Volatile Risk Environment. Retrieved from <u>https://na.theiia.org/news/Pages/2015-North-American-Pulse-of-Internal-Audit-Released.aspx</u>

Insights: The next step in data analysis



One-click analysis that immediately identifies patterns, trends and outliers in your data.



Visually displays your data graphically in a custom dashboard that you build.



Focus on the right things and improve audit efficiency.

21.1 Visualization – Know What's in Your Data – Next Evolution in Data Analysis

The Visualize Task IDEA 10 introduces the new Visualize task, a user-

New and Emerging Features in Data Analysis

friendly visualization tool available through the Analytics tab. Working from one or even multiple databases, Visualize allows you to build customized visualization dashboards so you can see the outliers, distributions and trends in your data. Your Dashboards can then be saved and shared through the Library. Visualize also provides the ability to drill down with a grid view of your data and extract insights from any particular slice, bar, column or area of a chart. Additionally, the new Auto-Stratification feature stratifies on numeric and date fields by automatically setting appropriate ranges within your graphs and charts.

Visualization unlocks IDEA's Analytic Intelligence so you can forget the days of painstakingly scanning through a series of tables. Visualization's tasks automatically create charts and field statistics, profiling your data so you can quickly pinpoint patterns, trends, outliers and correlations that may have gone unnoticed in a table. Raising the bar even higher, Visualization displays data graphically in up to 10 panels within a dashboard that you can tailor to show exactly the results you want. All the information you need to prepare a more focused audit.

The screen – shot below depicts Visualization in IDEA.

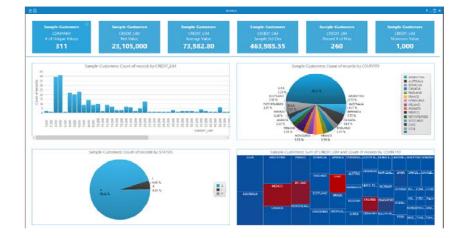


CASEWARE

21.2 Discover

Discover unlocks the power of IDEA's Analytic Intelligence. This proprietary intelligence, made up of pre-written algorithms, identifies trends, patterns or anomalies in the data that deserve investigation. Discover automatically presents charts or field statistics on a dashboard visually displaying areas of interest and anomalies that may have been hidden in the data. Any dashboard created through Discover can be modified, saved and shared with others through the Library. In addition, Discover allows you to identify key fields within a database and extract the key field statistical information you need. Based on the sound statistical method called Inter-Quartile Range (IQR), you can also identify outlier distribution and flag areas of interests in a database on numeric and date fields.

The screen – shot below depicts Discover in IDEA.

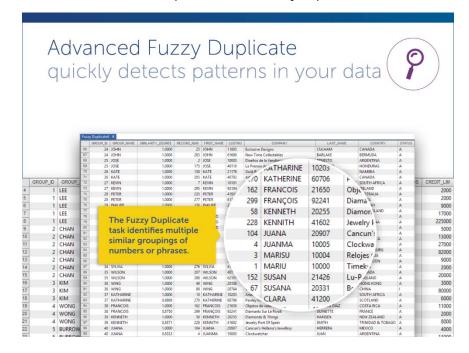


21.3 Advanced Fuzzy Duplicates

The Advanced Fuzzy Duplicate tool is more powerful than ever, identifying multiple similar records using an unheard of three Character fields. Highlighting similar records and then grouping them based on the degree of similarity, this tool does your work for you, detecting data entry errors, multiple data conventions for recording information, and fraud. You can even fine-tune your data groups based on case sensitivity.

New and Emerging Features in Data Analysis

The screen – shot below depicts Advanced Fuzzy Duplicates in IDEA.



APPENDICES

Illustrative List of Computer Assisted Audit Tools (CAATs)

- 1. IDEA Data Analysis Software (*www.caseware-idea.com*)
- 2. Methodware Risk Management and Internal Auditing Software (*www.methodware.com*)
- 3. TeamMate Audit Management System (*http://tax.cchgroup.com/TeamMate/default*)
- 4. Approva Controls Intelligence Software (*www.approva.net*)
- 5. Caseware Working Papers (*www.caseware-idea.com*)
- 6. Audit Leverage Audit Management Software (www.auditleverage.com)
- 7. Protiviti Governance, Risk, Compliance (GRC) Software (www.protiviti.com)
- 8. PAWS Audit & Risk Management Software (www.pentana.com)

Data Analysis for Auditors

PRACTICE NOTE

1009

COMPUTER-ASSISTED AUDIT TECHNIQUES

[Issued December 2003; revised September 2004 (name change)]

Contents

	Paragraphs
Introduction	1 - 3
Description of Computer-Assisted Audit Techniques (CAATs)	4 – 6
Considerations in the Use of CAATs	7 - 16
Using CAATs	17 – 25
Using CAATs in Small Entity IT Environments	26
Compatibility with International Auditing Practice Statements	27

PRACTICE NOTE 1009 COMPUTER-ASSISTED AUDIT TECHNIQUES

The purpose of Practice Notes issued by the Hong Kong Institute of Certified Public Accountants is to assist auditors in applying Statements of Auditing Standards (SASs) and Standards on Assurance Engagements (SAEs) of general application to particular circumstances and industries.

They are persuasive rather than prescriptive. However they are indicative of good practice and have similar status to the explanatory material in SASs and SAEs, even though they may be developed without the full process of consultation and exposure used for SASs and SAEs. Auditors should be prepared to explain departures when called upon to do so.

This Practice Note replaces Auditing Guideline 3.262 "Computer-Assisted Audit Techniques (CAATS)".

Introduction

- The overall objectives and scope of an audit do not change when an audit is conducted in a computer information technology (IT) environment. The application of auditing procedures may, however, require the auditors to consider techniques known as Computer-Assisted Audit Techniques (CAATs) that use the computer as an audit tool.
- CAATs may improve the effectiveness and efficiency of auditing procedures. They may also provide effective tests of control and substantive procedures where there are no input documents or a visible audit trail, or where population and sample sizes are very large.
- The purpose of this Practice Note (PN) is to provide guidance on the use of CAATs. It applies to all uses of CAATs involving a computer of any type or size. Special considerations relating to small entity IT environments are discussed in paragraph 26.

Description of Computer-Assisted Audit Techniques (CAATs)

4. This PN describes computer-assisted audit techniques including computer tools, collectively referred to as CAATs. CAATs may be used

in performing various auditing procedures, including the following:

- a. Tests of details of transactions and balances, for example, the use of audit software for recalculating interest or the extraction of invoices over a certain value from computer records;
- b. Analytical procedures, for example, identifying inconsistencies or significant fluctuations;
- c. Tests of general controls, for example, testing the set-up or configuration of the operating system or access procedures to the program libraries or by using code comparison software to check that the version of the program in use is the version approved by management;
- d. Sampling programs to extract data for audit testing;
- e. Tests of application controls, for example, testing the functioning of a programmed control; and
- f. Reperforming calculations performed by the entity's accounting systems.
- 5. CAATs are computer programs and data the auditors use as part of the audit procedures to process data of audit significance contained in an entity's information systems. The data may be transaction data, on which the auditors wish to perform tests of controls or substantive procedures, or they may be other types of data. For example, details of the application of some general controls may be kept in the form of text or other files by applications that are not part of the accounting system. The auditors can use CAATS to review those files to gain evidence of the existence and operation of those controls. CAATS may consist of package programs, purpose-written programs, utility programs or system management programs. Regardless of the origin of the programs, the auditors substan tiate their appropriateness and validity for audit purposes before using them.
 - a. *Package programs* are generalized computer programs designed to perform data processing functions, such as reading data, selecting and analyzing information, performing calculations, creating data files and reporting in a format specified by the auditors.
 - b. Purpose-written programs perform audit tasks in specific

circumstances. These programs may be developed by the auditors, the entity being audited or an outside programmer hired by the auditors. In some cases the auditors may use an entity's existing programs in their original or modified state because it may be more efficient than developing independent programs.

- c. *Utility programs* are used by an entity to perform common data processing functions, such as sorting, creating and printing files. These programs are generally not designed for audit purposes, and therefore may not contain features such as automatic record counts or control totals.
- d. *System Management programs* are enhanced productivity tools that are typically part of a sophisticated operating systems environment, for example, data retrieval software or code comparison software. As with utility programs, these tools are not specifically designed for auditing use and their use requires additional care.
- e. *Embedded Audit Routines* are sometimes built into an entity's computer system to provide data for later use by the auditors. These include:
 - i. Snapshots: This technique involves taking a picture of a transaction as it flows through the computer systems. Audit software routines are embedded at different points in the processing logic to capture images of the transaction as it progresses through the various stages of the processing. Such a technique permits auditors to track data and evaluate the computer processes applied to the data.
 - ii. System Control Audit Review File: This involves embedding audit software modules within an application system to provide continuous monitoring of the system's transactions.

The information is collected into a special computer file that the auditors can examine.

f. *Test data techniques* are sometimes used during an audit by entering data (for example, a sample of transactions) into an entity's computer

system, and comparing the results obtained with predetermined results. Auditors might use test data to:

- i. test specific controls in computer programs, such as on-line password and data access controls;
- test transactions selected from previously processed transactions or created by the auditors to test specific processing characteristics of an entity's information systems. Such transactions are generally processed separately from the entity's normal processing; and
- iii. test transactions used in an integrated test facility where a "dummy" unit (for example, fictitious department or employee) is established, and to which test transactions are posted during the normal processing cycle.

When test data are processed with the entity's normal processing, the auditors en sure that the test transactions are subsequently eliminated from the entity's accounting records.

- 6. The increasing power and sophistication of PCs, particularly laptops, has resulted in other tools for the auditors to use. In some cases, the laptops will be linked to the auditors' main computer systems. Examples of such techniques include:
- a. expert systems, for example in the design of audit programs and in audit planning and risk assessment;
- b. tools to evaluate a client's risk management procedures;
- c. electronic working papers, which provide for the direct extraction of data from the client's computer records, for example, by downloading the general ledger for audit testing; and
- d. corporate and financial modeling programs for use as predictive audit tests.

These techniques are more commonly referred to as "audit automation."

Considerations in the Use of CAATs

- 7. When planning an audit, the auditors may consider an appropriate combination of manual and computer assisted audit techniques. In determining whether to use CAATs, the factors to consider include:
 - a. the IT knowledge, expertise and experience of the audit team;

- b. the availability of CAATs and suitable computer facilities and data;
- c. the impracticability of manual tests;
- d. effectiveness and efficiency; and
- e. timing.

Before using CAATS the auditors consider the controls incorporated in the design of the entity's computer systems to which the CAATS would be applied in order to determine whether, and if so, how, CAATs should be employed.

IT Knowledge, Expertise, and Experience of the Audit Team

8. SAS 3 10 "Auditing in a computer information systems environment" deals with the level of skill and competence the audit team needs to conduct an audit in an IT environment. It provides guidance when auditors delegate work to assistants with IT skills or when the auditors use work performed by other auditors or experts with such skills. Specifically, the audit team would need to have sufficient knowledge to plan, execute and use the results of the particular CAAT adopted. The level of knowledge required depends on the complexity and nature of the CAAT and of the entity's information system.

Availability of CAATs and Suitable Computer Facilities

- 9. The auditors consider the availability of CAATs, suitable computer facilities (controlled as described in paragraphs 18-23) and the necessary computer-based information systems and data. The auditors may plan to use other computer facilities when the use of CAATs on an entity's computer is uneconomical or impractical, for example, because of an incompatibility between the auditors' package program and the entity 's computer. Additionally, the auditors may elect to use their own facilities, such as PCs or laptops.
- 10. The cooperation of the entity's personnel may be required to provide processing facilities at a convenient time, to assist with activities such as loading and running of the CAATs on the entity's system, and to provide copies of data files in the format required by the auditors.

Impracticability of Manual Tests

- 11. Some audit procedures may not be possible to perform manually because they rely on complex processing (for example, advanced statistical analysis) or involve amounts of data that would overwhelm any manual procedure. In addition, many computer information systems perform tasks for which no hard copy evidence is available and, therefore, it may be impracticable for the auditors to perform tests manually. The lack of hard copy evidence may occur at different stages in the business cycle.
 - a. Source information may be initiated electronically, such as by voice activation, electronic data imaging, or point of sale electronic funds transfer. In addition, some transactions, such as discounts and interest calculations, may be generated directly by computer programs with no specific authorization of individual transactions.
 - b. A system may not produce a visible audit trail providing assurance as to the completeness and accuracy of transactions processed. For example, a computer program might match delivery notes and suppliers' in voices. In addition, programmed control procedures, such as checking customer credit limits, may provide hard copy evidence only on an exception basis.
 - c. A system may not produce hard copy reports. In addition, a printed report may contain only summary totals while computer files retain the supporting details.

Effectiveness and Efficiency

- 12. The effectiveness and efficiency of auditing procedures may be improved by u sing CAATs to obtain and evaluate audit evidence. CAATs are often an efficient means of testing a large number of transactions or controls over large populations by:
 - a. analyzing and selecting samples from a large volume of transactions;
 - b. applying analytical procedures; and
 - c. performing substantive procedures.
- 13. Matters relating to efficiency that auditors might consider include:

- a. the time taken to plan, design, execute and evaluate a CAAT;
- b. technical review and assistance hours;
- c. designing and printing of forms (for example, confirmations); and
- d. availability of computer resources.
- 14. In evaluating the effectiveness and efficiency of a CAAT, the auditors consider the continuing use of the CAAT application. The initial planning, design and development of a CAAT will usually benefit audits in subsequent periods.

Timing

- 15. Certain data, such as transaction details, are often kept for only a short time, and may not be available in machine-readable form by the time the auditors want them. Thus, the auditors will need to make arrangements for the retention of data required, or may need to alter the timing of the work that requires such data.
- 16. Where the time available to perform an audit is limited, the auditors may plan to use a CAAT because its use will meet the auditors' time requirement better than other possible procedures.

Using CAATs

- 17. The major steps to be undertaken by the auditors in the application of a CAAT are to:
 - a. set the objective of the CAAT application;
 - b. determine the content and accessibility of the entity's files;
 - c. identify the specific files or databases to be examined;
 - understand the relationship between the data tab les where a database is to be examined;
 - e. define the specific tests or procedures and related transactions and balances affected;
 - f. define the output requirements;
 - g. arrange with the user and IT departments, if appropriate, for copies of the relevant files or database tables to be made at the appropriate cutoff date and time;

- h. identify the personnel who may participate in the design and application of the CAAT;
- i. refine the estimates of costs and benefits;
- j. ensure that the use of the CAAT is properly controlled and documented;
- k. arrange the administrative activities, including the necessary skills and computer facilities;
- I. reconcile data to be used for the CAAT with the accounting records;
- m. execute the CAAT application; and
- n. evaluate the results.

Controlling the CAAT Application

18. The specific procedures necessary to control the use of a CAAT depend on the particular application.

In establishing control, the auditors consider the need to:

- approve specifications and conduct a review of the work to be performed by the CAAT;
- b. review the entity's general controls that may contribute to the integrity of the CAAT, for example, controls over program changes and access to computer files. When such controls cannot be relied o n to ensure the integrity of the CAAT, the auditors may consider processing the CAAT application at another suitable computer facility; and
- c. ensure appropriate integration of the output by the auditors into the audit process.
- 19. Procedures carried out by the auditors to control CAAT applications may include:
 - a. participating in the design and testing of the CAAT;
 - b. checking, if applicable, the coding of the program to en sure that it conforms with the detailed program specifications;
 - c. asking the entity's computer staff to review the operating system instructions to ensure that the software will run in the entity's computer installation;

- d. running the audit software on small test files before running it on the main data files;
- e. checking whether the correct files were used, for example, by checking external evidence, such as control totals maintained by the user, and that those files were complete;
- f. obtaining evidence that the audit software functioned as planned, for example, by reviewing output and control information; and
- g establishing appropriate security measures to safeguard the integrity and confidentiality of the data.

When the auditors intend to perform audit procedures concurrently with online processing, the auditors review those procedures with appropriate client personnel and obtain approval before conducting the tests to help avoid the inadvertent corruption of client records.

- 20. To ensure appropriate control procedures, the presence of the auditors is not necessarily required at the computer facility during the running of a CAAT. It may, however, provide practical advantages, such as being able to control distribution of the output and ensuring the timely correction of errors, for example, if the wrong input file were to be used.
- 21. Audit procedures to control test data applications may include:
 - a. controlling the sequence of submissions of test data where it spans several processing cycles;
 - b. performing test runs containing small amounts of test data before submitting the main audit test data;
 - c. predicting the results of the test data and comparing it with the actual test data output, for the individual transactions and in total;
 - d. confirming that the current version of the programs was used to process the test data; and
 - e. testing whether the programs used to process the test data were the programs the entity used throughout the applicable audit period.
- 22. When using a CAAT, the auditors may require the cooperation of entity

Data Analysis for Auditors

staff with extensive knowledge of the computer installation. In such circumstances, the auditors consider whether the staff improperly influenced the results of the CAAT.

- 23. Audit procedures to control the use of audit-enabling software may include:
 - a. verifying the completeness, accuracy and availability of the relevant data, for example, historical data may be required to build a financial model;
 - b. reviewing the reasonableness of assumptions used in the application of the tool set, particularly when using modeling software;
 - c. verifying availability of resources skilled in the use and control of the selected tools; and
 - d. confirming the appropriateness of the tool set to the audit objective, for example, the use of industry specific systems may be necessary for the design of audit programs for unique business cycles.

Documentation

24. The standard of working paper documentation and retention procedures for a CAAT is consistent with that for the audit as a whole (see SAS 230 "Documentation").

25. The working papers need to contain sufficient documentation to describe the CAAT application, such as:

- a. Planning
 - i CAAT objectives;
 - ii. consideration of the specific CAAT to be used;
 - iii. controls to be exercised; and
 - iv. staffing, timing and cost.
- b. Execution
 - i. CAAT preparation and testing procedures and controls;
 - ii. details of the tests performed by the CAAT;
 - iii. details of input, processing and output; and

- iv. relevant technical information about the entity's accounting system, such as file layouts.
- c. Audit Evidence
 - i. output provided;
 - ii. description of the audit work performed on the output; and

iii. audit conclusions.

- d. Other
 - i. recommendations to entity management.

In addition, it may be useful to document suggestions for using the CAAT in future years.

Using CAATs in Small Entity IT Environments

- 26. Although the general principles outlined in this PN apply in small entity IT environments, the following points need special consideration:
 - a. The level of general controls may be such that the auditors will place less reliance on the system of internal control. This will result in greater emphasis on tests of details of transactions and balances and analytical review procedures, which may increase the effectiveness of certain CAATs, particularly audit software.
 - b. Where smaller volumes of data are processed, manual methods may be more cost effective.
 - c. A small entity may not be able to provide adequate technical assistance to the auditors, making the use of CAATs impracticable.
 - d. Certain audit package programs may not operate on small computers, thus restricting the auditors' choice of CAATs. The entity's data files may, however, be copied and processed on another suitable computer.

Compatibility with International Auditing Practice Statements

27. This Practice Note is, in all material respects, in accordance with International Auditing Practice Statement 1009 "Computer-Assisted Audit Techniques".

Appendix II

Attribute Sampling Technical Specification



A CaseWare IDEA Research Department document

June 17, 2003

Copyright ©2003 by CaseWare IDEA Inc. All rights reserved. This document is intended for use by CaseWare IDEA Inc., its employees and authorized distributors. It is not to be provided to anyone else under any circumstances, without the express permission of the Chief Operating Officer of CaseWare IDEA Inc. No part of this publication may be reproduced, transmitted, transcribed, stored in a retrieval system or translated into any language, in any form, by any means, without the permission of CaseWare IDEA Inc. IDEA is a registered trademark of CaseWare International Inc. Copyright 2002 CaseWare IDEA Inc. 145 King Street East, 2nd Floor Toronto, Ontario M5C 2Y8



Contents

1.	Attribute Sampling Overview	4
2.	Sample size determination controlling beta risk	4
3.	Sample size determination controlling alpha and beta risks	5
4.	Determination of 1- or 2-sided reliability intervals	6
Refer	ences:	8
Apper	ndix A: Technical Note	. 9
1. Int	roduction	. 9
2. Hy	per geometric Risk Function Stability	. 9
3. Pr	ecision Limit Definitions	10
4. Be	ta Risk Control	12
5. As	ymmetric Risks in Two-Sided Intervals	13

Data Analysis for Auditors

1. Attribute Sampling Overview

The purpose of attribute sampling is to estimate the number of times an event occurs in a population, based on the examination of a sample from the population. In auditing applications, the events of interest are typically misstatements (differences between recorded and audited balances), deviations from internal control procedures, or failure to comply with laws and regulations. This document describes the mathematical procedures followed by IDEA in planning and evaluating attribute samples.

Attribute sampling functionality for IDEA is provided by three main functions, each of which performs one of the following tasks:

- 1. Planning the sample size to control beta risk
- 2. Planning the sample size to control alpha and beta risks
- 3. Determination of one or two-sided reliability intervals for a given sample (also known as sample evaluation).

These functions are described in the sections that follow. Further discussion of certain issues, relating to numerical algorithms used, may be found in [2].

2. Sample size determination controlling beta risk

Inputs:

- *N*: population size
- UOR: upper threshold occurrence rate
- *EOR*: expected occurrence rate
- *BetaR*: beta reliability or confidence

Outputs:

- *n*: sample size
- x_{cr} : critical number of occurrences in sample

Other quantities referred to below:

- *OR*: true (unknown) population occurrence rate
- X: true (unknown) number of occurrences in population; $X = N \times OR$

 x_h : hypergeometric random variable (distribution depends on N, n, X)

 X_{tol} : upper threshold value of X; $X_{\text{tol}} = N \times UOR$

Description:

This planning function finds the minimum n and corresponding occurrence rate x_{cr} such that:

- 1. It can be inferred with a reliability of at least *BetaR* that the true occurrence rate is less than or equal to UOR, for population size n and number of sample occurrences up to x_{cr} .
- 2. x_{cr} / n is at least as big as *EOR*.

Please refer to [1] for a description of the meaning of reliability in this context.

Condition 1) is met in the following way: Set

 $X_{\text{tol}} = \text{floor} (N \times UOR) + 1$

where 'floor' denotes the rounding down operation. Then, n and x_{cr} must satisfy

prob $(x_h > x_{cr} | N, n, X_{tol}) \ge BetaR$.

where 'prob' denotes total probability. This means that, if the true number of occurrences X, were as big as X_{tol} , then the probability is at least *BetaR* and the number of occurrences in a sample of size *n* would be *greater than* x_{cr} . In other words, provided the sample of size *n* contains no more than x_{cr} occurrences, the interval [0, X_{tol} –1] may be expected to contain the true value X with reliability at least *BetaR*. Dividing by N to obtain occurrence rates, we see that the interval

[0, *UOR*] may be expected to contain the true value *OR* with reliability *BetaR*. Further discussion of the beta risk control can be found in [2].

Sample size determination controlling alpha and beta risks

Inputs:

- *N*: population size
- *UOR*: upper threshold occurrence rate
- *LOR*: lower threshold occurrence rate

Data Analysis for Auditors

BetaR: beta - reliability or confidence

AlphaR: alpha - reliability or confidence

Outputs:

- *n*: sample size
- *x*_{cr}: critical number of occurrences in sample

Other quantities referred to below:

- *OR*: true (unknown) population occurrence rate; $OR = N \times X$
- *X*: true (unknown) number of occurrences in population
- $x_{\rm h}$: hypergeometric random variable (distribution depends on N, n, X)
- X_{tol} : lower limit of X to ensure tolerable beta-risk (1 *BetaR*)
- X_{\min} : lower limit of X to ensure tolerable alpha-risk (1 AlphaR)

Description:

This planning function finds the minimum *n* and corresponding occurrence rate x_{cr} such that:

- 1. It can be inferred with a reliability of at least *BetaR* that the true occurrence rate is less than or equal to *UOR*, for population size *n* and number of sample occurrences up to x_{cr} . Equivalently, the risk of deciding that $OR \le UOR$ when in fact OR > UOR, is no greater than the beta risk (1 BetaR).
- 2. The risk of deciding that OR > UOR when in fact $OR \le LOR$, is no greater than the alpha risk (1 AlphaR).

The reader is referred to [1] for a description of the meanings of reliability and risk in this context.

Condition 1) is met in the following way: Set

 $X_{\text{tol}} = \text{floor} (N \times UOR) + 1$

where 'floor' denotes the rounding down operation. Then, n and x_{cr} satisfy

prob $(x_h > x_{cr} | N, n, X_{tol}) \ge BetaR$

where 'prob' denotes total probability. This is the same method as is used to control beta risk in planning to control only beta risk; its interpretation is discussed in Section 2 above.

Condition 2) is met in the following way: Set

 $X_{\min} = \text{ceil} (N \times LOR)$

Where 'ceil' is the rounding up operation. Then, n and x_{cr} satisfy

prob $(x_h \le x_{cr} | N, n, X_{min}) \ge AlphaR$

This means that, if the true number of occurrences is X_{min} (or less), then the probability is at least *AlphaR* that the number of occurrences in a sample of size *n* will be x_{cr} or smaller, and therefore that we will not falsely conclude a number of population occurrences greater than X_{tol} . Dividing by *N* to obtain occurrence rates, we see that the reliability *AlphaR* is achieved for any true population occurrence rate of *LOR* or less.

4. Determination of 1- or 2-sided reliability intervals

Inputs:

sides: 1 or 2; number of interval endpoints to calculate.

- *N*: population size
- *n*: sample size
- *x*: number of occurrences in sample
- *R*: reliability or confidence

Outputs:

 OR_{right} : upper precision limit of OR; $OR_{right} = X_{right}/N$

 OR_{left} : lower precision limit of OR; $OR_{\text{left}} = X_{\text{left}}/N$; not calculated if *sides* = 1.

Other quantities referred to below:

- X: true (unknown) number of occurrences in population
- X_{right} : upper precision limit of X
- X_{left} : lower precision limit of X; not calculated if *sides* = 1.
- $x_{\rm h}$: random variable with hypergeometric distribution specified by N, n,
- Х.
- *OR*: population occurrence rate; OR = X/N

Description:

If *sides* = 2, risk is divided equally between upper and lower limits, i.e.:

If sides = 1, R' = R.

If sides = 2, R' = (1+R)/2.

If *sides* = 1, then X_{left} is not calculated (returns 0), and X_{right} satisfies the following:

1. If $X > X_{right}$, then prob $(x_h > x \mid N, n, X) \ge R'$

2. If
$$X \le X_{right}$$
, then prob $(x_h > x \mid N, n, X) < R'$

This means that, if the true number of occurrences X were any more than X_{right} , then the probability is at least R' that the number of occurrences in a sample of size n would be greater than the observed number x. In other words, X is less than or equal to X_{right} with reliability R'.

If *sides* = 2, then X_{right} is as above, and X_{left} satisfies the following:

- 1. If $X < X_{\text{left}}$, then prob $(x_h < x \mid N, n, X) \ge R$
- 2. If $X \ge X_{\text{left}}$, then prob $(x_h < x \mid N, n, X) < R'$

This means that, if the true number of occurrences X were less than X_{left} , then the probability is at least R' that the number of occurrences would be less than the observed number x. In other words, X is greater than or equal to X_{left} with reliability R'.

The above definitions of χ_{ieft} and χ_{right} are similar to those given in [1] translated into mathematical notation, with one minor adjustment as described in [2].

Although it is a standard technique [1], the process of dividing the tolerable risk

1 - R between the upper and lower limits to obtain two-sided intervals, can lead to unnecessarily conservative intervals whenever $X_{\text{left}} = 0$ [2]. This situation most frequently occurs when x = 0, but may occur for other low numbers of occurrences, when *n* and *N* are large. When it does, IDEA displays a message saying that the two-sided interval is unavailable.

References:

- [1] Donald M. Roberts, *Statistical Auditing*, AICPA 1978; p. 184.
- [2] K.M. Adeney, "Appendix A", CaseWare Idea Inc. Internal Report, April 2002.

Appendix A: Technical Note

1. Introduction

Attribute sampling functionality for IDEA is provided by three main functions, each of which performs one of the following tasks:

- 1. Planning the sample size to control beta risk
- 2. Planning the sample size to control alpha and beta risks
- 3. Determination of one or two-sided reliability intervals for a given sample (also known as sample evaluation)

The above functions are described in [2]. This document provides extra detail on certain aspects of the methodology.

2. Hyper geometric Risk Function Stability

Risk functions for attribute sampling are calculated precisely, using the hypergeometric probability function as appropriate for sampling without replacement. The hypergeometric risk (HGR) function is based on the cumulative hypergeometric probability – that is, the sum of probabilities of observed sample occurrence numbers ranging from zero to the specified number of occurrences. This quantity may be calculated in any of a number of ways, with different methods exhibiting different numerical properties. Since direct computation of each hypergeometric probability is computationally expensive and numerically unstable, recursive relationships between probabilities are exploited in calculating the sum.

Outputs of IDEA's HGR function were tested against those produced by existing software programs such as ATSIZ1, ATSIZ2 and ATEVAL, as described in [1], as well as an additional program called SAMPLE, which performs the same calculations. The HGR function used by the BASIC ATSIZ1 and ATSIZ2 code appears to calculate HGR using a mathematical recursion that is initialized with the probability of observing zero occurrences. Although this procedure is more stable than calculation of HGR using direct evaluations of factorial expressions, it has a bug: it always returns zero when the probability of observing zero occurrences in the sample is zero, even if the total hypergeometric risk is not zero. Furthermore, this implementation becomes unstable when the probability of observing zero sample occurrences is very small (i.e. when both population occurrence rate and sample size is large).

Data Analysis for Auditors

Errors in sample sizes returned by SAMPLE are consistent with the hypothesis that the HGR function used by SAMPLE also has the bug described above. In sampling to control beta risk, SAMPLE looks for lowest n such that the risk function is below 1 - R (where R is the reliability). Sample sizes that are too low may result from the fact that, because of the bug in the HGR function, the risk condition is met sooner (for lower n) than it should be.

IDEA's attribute sampling routines calculate HGR using a mathematical recursion that is similar to the one used in previous HGR functions, except that it is initialized with a value that is guaranteed not to be zero. This value is the probability of observing whatever number of occurrences is most likely for the given population occurrence rate and sample size. This implementation does not exhibit the bug described above.

Although the new HGR function can better tolerate smaller individual event probabilities compared to the previous HGR functions, ANY discreet cumulative probability distribution will become more unstable, as the individual event probabilities summed to obtain the cumulative probability become very small. This occurs when both sample size and population size are large. For example, for a population of 1,000,000 and a population occurrence rate of 10%, the HGR function is valid for sample sizes up to 8300.

The range of allowable sample sizes has not yet been characterized for an extensive number of occurrence rates and population sizes – in general, it increases as population occurrence rate decreases, and decreases as population size increases. However, the required reliability values would need to be very high, and the tolerable and expected occurrence rates very close, in order for such large sample sizes to be required in attribute sampling. Therefore, instability of the HGR function is not expected to be of practical importance. Code has been included to detect this instability if it arises. If instability of the hypergeometric function does pose a practical problem in future, a continuous approximation to the cumulative hypergeometric probability may be used to circumvent this problem.

3. Precision Limit Definitions

The evaluation function finds reliability interval endpoints, or precision limits, while the two planning functions use precision limits implicitly in order to find sample sizes. It is important that precision limits used by all programs be defined consistently. This section describes some issues that have been identified with the definitions of precision limits found in Roberts [1] and in the outputs of past sampling programs.

Firstly, it is noted that numbers of population occurrences are always integers, so precision limits on population occurrence rates should correspond to integer population occurrence numbers. This is not the case with outputs from the previous DOS program SAMPLE, although it appears to hold for the outputs of the DOS program ATEVAL. Precision limits will be discussed here in terms of numbers of occurrences, rather than occurrence rates, since these discrete quantities are the ones actually calculated using the hypergeometric risk function.

We first consider the definitions of the left and right precision limits in Roberts' precision limit evaluation function, ATEVAL [1]. According to the description of that function, the right precision limit on the number of population occurrences is some number X_{right} such that

$$\operatorname{prob} (x_h > x_{obs} \mid N, n, X_{right}) = R , \qquad (1)$$

where x_{obs} is the observed number of sample deviations and other quantities are as described above. Here the difficulty arises. Because the hypergeometric distribution is discrete, it is highly unlikely that there is some value of *R* satisfying the above equation exactly. One may instead say that X_{right} is the *minimum* value of *X* satisfying

$$\operatorname{prob}\left(x_{h} > x_{obs} \mid N, n, X\right) \ge R.$$

$$(2)$$

The interpretation of Ineq. (2) is that, if the true number of occurrences were as big as X_{right} , then the probability is at least R that the number of occurrences in a sample of size n would be *greater than* the observed number x_{crit} . In other words, X is strictly less than X_{right} with reliability R, or equivalently X is less than or equal to $X_{right} - 1$ with reliability R. Values returned by SAMPLE are consistent with these definitions (except where they are in error due to instability of that program's hypergeometric risk function).

Alternately, one may define X_{right} to be the *maximum* value of X satisfying

prob
$$(x > x_{obs} | N, n, X) < R$$
. (3)

The difference between these two definitions is illustrated below, where X_2 is the value of X_{right} resulting from Ineq. (2) and X_3 is the value of X_{right} resulting from Ineq. 3.

Х	:	•	•	•	•	•
prob > or < <i>R</i> ?	:	<	<	\geq	>	>
				\uparrow	\uparrow	
				<i>X</i> ₃	X_2	

Note that $X_3 = X_2 - 1$ always.

In Attribute Sampling, Ineq. (3) is used, so that the left and right precision limits are included in the reliability R interval. For two-sided evaluation, X_{left} is the *minimum* value of X satisfying

prob
$$(x < x_{obs} | N, n, X) < R$$
. (4)

where R has been obtained by dividing the acceptable risk between upper and lower limits as described in [1].

4. Beta Risk Control

We now consider Roberts' descriptions of the sample size determination programs ATSIZ1 and ATSIZ2 [1], in particular the method of controlling beta risk. Roberts says that the returned sample size is the minimum possible such that the population deviation rate is *strictly less than* the upper threshold occurrence rate *UOR* (Roberts calls this the *tolerable deviation rate*) with reliability at least *R*, subject to constraints on either the expected deviation rate (ATSIZ1) or the alpha risk (ATSIZ2). This is the minimum sample size *n* satisfying (subject to constraints)

$$\operatorname{prob}\left(x_h > x_{crit} \mid N, n, X_{tol}\right) \ge R.$$
(5)

Here x_h is a hypergeometric random variable with population N, sample size n, and number of population occurrences X_{tol} , and x_{crit} is the critical number of sample deviations. Roberts does not specify the way in which X_{tol} is obtained from *UOR*. Either rounding up or rounding down (e.g. X_{tol} = ceil(*UOR*×*N*) or X_{tol} = floor(*UOR* ×*N*)) may be used. DOS versions of SAMPLE and ATSIZ1 and ATSIZ2 appear to have used rounding down.

In summary, in Roberts' ATSIZ1 and ATSIZ2 programs, the true occurrence rate *OR* is *strictly less than* the *UOR* with reliability R – so the *UOR* is excluded from the 1-sided reliability R interval. In Attribute Sampling, the above procedure is modified slightly so that the population occurrence rate can be as high as the *UOR* without going below the desired beta reliability R. This is achieved by setting

$$X_{tot} = \text{floor}(UOR \times N) + 1 \tag{6}$$

Thus, we have

$$UOR \times N < X_{tol} \le UOR \times N + 1 \tag{7}$$

The above value of X_{tol} is used in Ineq. (5). By the reasoning given following Ineq. (2), we know that with reliability R, the true population number of occurrences X satisfies

(8)

$$X \leq X_{tol} - 1$$

Dividing Ineq. (8) by N to translate into occurrence rates, and combining with Ineq. (7), we get

$$OR = X / N \le (X_{tol} - 1) / N \le (UOR \times N + 1 - 1) / N = UOR$$
 (9)

with reliability *R*. Thus the *UOR* is now included in the 1-sided reliability *R* interval.

The attribute sampling functions return slightly smaller sample sizes for some cases than do programs that exclude the *UOR* from the reliability interval. This difference becomes more pronounced as the sample size grows relative to the population size.

5. Asymmetric Risks in Two-Sided Intervals

Although it is a standard technique [1], the process of dividing the tolerable risk 1 - R between the upper and lower limits to obtain two-sided intervals, can lead to unnecessarily conservative intervals, whenever $X_{\text{left}} = 0$ [2]. In such cases the found value of $X_{\text{left}} = 0$ would work for lower confidences than the specified value of R = (1+R) / 2 [1], [2]. Said another way, the risk allowed for the left limit (i.e., (1 - R)/2) may not all be "used up" when the lower limit is zero. Therefore the true confidence of the supplied interval is higher than R.

A narrower interval (e.g., between $\chi_{\text{left}} = 0$ and some smaller value of χ_{right}) could be found which would still have a total risk less than or equal to 1 - R. For such an interval, the risks assigned to each of the upper and lower limits would be unequal, and would have to be determined by a software search. For the present time, IDEA instead displays a message saying that the two-sided interval is unavailable, whenever the calculated lower limit is 0.

Appendix III

Monetary Unit Sampling Technical Specification



A CaseWare IDEA Research Department document



May 7, 2003

Copyright ©2003 by CaseWare IDEA Inc. All rights reserved. This document is intended for use by CaseWare IDEA Inc., its employees and authorized distributors. It is not to be provided to anyone else under any circumstances, without the express permission of the Chief Operating Officer of CaseWare IDEA Inc. No part of this publication may be reproduced, transmitted, transcribed, stored in a retrieval system or translated into any language, in any form, by any means, without the permission of CaseWare IDEA Inc.

> IDEA is a registered trademark of CaseWare International Inc. Copyright 2002 CaseWare IDEA Inc. 145 King Street East, 2nd Floor Toronto, Ontario M5C 2Y8

Contents

1.	Introdu	ction	5
2.	Handlin	g of High Value Items	5
3.	Plannin	g	7
	3.1	Planning when Population Size is Known	7
	3.2	Planning when Population Size is Unknown	11
	3.3	Warning Messages from Planning	13
4.	Sample	Evaluation	14
	4.1	Single Sample Evaluation when Fewer than 20 Errors are	
		Found	14
	4.2	Combined Evaluation of Multiple Samples	19
	4.3	Single Sample Evaluation when More than 20 Errors are	
		Found	24
	4.5	Warning Messages from Evaluation	27
Ref	erences	28	

Introduction

The purpose of Monetary Unit Sampling (MUS) is to estimate an upper limit on the error in a population, based on examination of a MUS sample from the population. The sampling unit in an MUS sample is an individual dollar, rather than an item or database record. Thus, the number of units in the sampled population (commonly denoted by *N*) is equal to the total monetary value of this population (commonly denoted by *Y*). The theory behind MUS is complex and its explanation is beyond the scope of this paper – interested readers may see [2], [3], [4].

This document describes the mathematical procedures followed by IDEA in planning and evaluating Monetary Unit samples. Different methodologies are used to evaluate samples depending on whether they are found to contain a low number of errors (fewer than 20), or a high number of errors (20 or more). For the low-error case, IDEA uses a cell-based MUS evaluation similar to that described in [2]. An earlier version of this implementation is described in a previous IDEA Users' Guide [1]. For the high-error case, classical PPS theory [4] is applied. IDEA also supports joint evaluation of several samples to estimate an upper error limit for the combined population (that is, the union of all of the populations from which the samples were drawn); this evaluation uses an extension of the Stringer bound described in [3].

Note that the basic precision pricing (*BPP*) and confidence level (*CL*) are input to the screens as percentages, but appear in the formulae below as decimal fractions, for example 0.95 rather than 95.

Handling of High Value Items

High value (HV) items are records in the population having recorded amounts greater than or equal to a user-specified *high value amount*. IDEA supports two methods of dealing with HV items. The first is to extract all high value items to a separate file, the *high value database*, during sample extraction. The remaining items are then extracted to the main sample database using an average sampling interval equal to the planned sampling interval. Errors in the HV database are summed and added to the most likely error and upper error limit derived from the main sample database. They do not affect the precision gap between the upper error limit and most likely error.

The second option is to treat HV items no differently from other items in the population. In this case high value items may be sampled one, two, or more

times. Thus, the main sample database can contain multiple records for a single HV item in the population database. Errors in HV items in the main sample database affect the precision of the upper error limit in the same way as any other errors in the sample. This mode is useful for situations in which high value records in the database are aggregates of many smaller items.

The HV amount is the sampling interval by default. However, users may set the HV amount to any value.

Users may optionally specify a HV file in the evaluation dialogs (or one HV file for each sample database, in the case of a combined evaluation IDEA generates a warning whenever any recorded amount in the main sample database is greater than the lesser of the sampling interval or the specified HV amount, if one was provided in sample planning. If the sample was planned using IDEA's MUS - Plan Sample routine, the HV amount used for this test is the one specified during the planning stage. Otherwise, a HV amount equal to the sampling interval is used. The HV amount does not affect the values computed by the evaluation, but is only used for generating these warnings.

3. Planning

3.1 Planning when Population Size is known

Inputs:

Y	Total monetary value of population from which sample was drawn	No restrictions See Note 1.
X _{max}	Maximum tolerable monetary value of all errors in population (e.g., materiality for this test) – absolute value	$0 < X_{\max} \le Y$ See Note 1.
X _{exp}	Expected monetary value of all errors in population – absolute value	$0 \le X_{exp} <$ X_{max} $X_{exp} < 0.9 Y$ See Note 1.
BPP	Basic precision pricing (assumed worst- case tainting on errors in population that do not appear in the sample).	0 < <i>BPP</i> < 10
CL	Confidence level (a.k.a. reliability)	0 < <i>CL</i> < 1 See Note 1.

Data Analysis for Auditors

Note 1: As well as the above individual restrictions on the inputs, a restriction is placed on the group of inputs. This restriction is designed to filter out inputs that would lead to impractically large sample sizes or numbers of sample errors. It has the following form:

$$\frac{X_{\max}}{Y} (1 - CL) \sqrt{X_{\max} - X_{\exp}} \ge 0.07$$
(3.0)

Outputs:

- *n*: approximate sample size (for population with no HV items removed)
- SI: average sampling interval; NIn
- *TT*: Tolerable tainting in sample.

Other quantities referred to below:

- x^* : tolerable number of 100% tainted errors in sample.
- n^* : minimum sample size that can tolerate x^* 100% tainted errors

N: size of population from which sample was drawn.

For MUS, N = |Y|.

Description:

For planning the sample, and also for evaluating a single sample when few errors are present, IDEA's MUS functions use cell-based MUS methods similar to those presented in references [1], [2]. One slight change is that IDEA now uses hyper geometric factors when population size is known, while [1] and [2] describe methods in terms of Bernoulli and/or Poisson factors. It is well known that hyper geometric factors are the most appropriate for modeling sampling from finite populations (e.g., see [2]). However, at the time [1], [2] were written, computations of hyper geometric factors was not feasible, and the use of the Bernoulli and Poisson approximations was common. Since IDEA already has stable hyper geometric probability functions for its Attribute Sampling module, these are used in MUS as well. The practical effect of this change on planned sample sizes and computed error limits is very small.

When population size is known, we first compute the sample size n and then use this to compute the *SI*. This planning function finds the approximate minimum n and corresponding tolerable tainting *TT* such that it can be inferred with a reliability of at least *CL* that the true number of errors is less than the tolerable number of errors, X_{max} . Equivalently, the risk of deciding that $X < X_{max}$ when in fact $X \ge X_{max}$ is no greater than (1 - CL). The sampling interval, sample size, and number of tolerable errors in the sample are first computed for the case of 100% tainting, and then scaled by the BPP to obtain *SI*, *n*, and *TT*.

The hyper geometric distribution is used for the risk calculations. Nonetheless, the planning process is approximate, since the upper error limit obtained when a sample is evaluated depends not just upon the total tainting, but also on how the tainted dollars are distributed among the errors.

The cumulative hyper geometric distribution is given by

$$H(x,n,X,N) = \sum_{k=1}^{x} \left(\frac{X!}{(X-k)!k!} \times \frac{(N-X)!}{(N-X-(n-k))!(n-k)!} \times \frac{(N-n)!n!}{N!} \right)$$
(3.1)

H(x, n, X, N) Is the risk of observing x or fewer errors in a sample of size n drawn from a population of size N in which the total number of errors is X.

The sample size that would be appropriate if all taintings were 100%, denoted by n^* , is computed in the following way:

Set the population total error as $X = X_{max}$.

We first describe the case when the expected error is greater than zero. Then, we consider sample errors x = 0, 1, 2. For each *x* considered, we find the minimum sample size, denoted n(x), which will tolerate *x* errors at the desired confidence level *CL*. That is, if up to *x* errors are found in a sample of size n(x), the *CL*-upper precision limit on the population error rate is less than X_{max} . This is done as follows:

For each successive *x*, we find the lowest *n* such that H(x, n, X, N) does not exceed the acceptable risk, 1 - CL. Thus,

$$n(x) = \min \left[n \ \ni \ H(x, n, X_{\max}, N) \le 1 - CL \right] (3.2)$$

We thus obtain a succession of pairs (x, n(x)) giving sample sizes n(x) which could tolerate *x* errors.

For each pair we then compute the sample error rate x/n'(x), and compare this with the expected error rate for the population, $p_{\exp} = X_{\exp} / N$. Let x' and x'+1 be successive integers such that

$$\frac{x'}{n'(x')} \le \frac{X_{\exp}}{N} \le \frac{x'+1}{n'(x'+1)}$$
(3.4)

We find the values x^* and n^* by interpolating between these pairs to achieve the desired ratio $x^*/n^* = p_{exp} = X_{exp} / N$. This yields

$$\frac{x^* - x'}{x' + 1 - x'} = \frac{n^* - n'(x')}{n'(x' + 1) - n'(x')} .$$
(3.5)

Substituting $x^* = n^* X_{exp} / N$ and solving, gives

$$n^{*} = ceil\left(\frac{x' n'(x'+1) - x'' n'(x')}{[n'(x'') - n'(x')] p_{exp} - 1}\right)$$

$$x^{*} = p_{exp} n^{*}$$
(3.6)

Where *ceil* () denotes rounding up. When the expected error is zero, we set $x^* = 0$, and compute $n^*(0) = n(0)$ using (4.2).

Finally, we scale by the BPP to get

$$TT = x * \times BPP$$

$$n = floor(n * \times BPP)$$

$$SI = N/(n * \times BPP)$$
(3.7)

Where *floor* () denotes rounding down. Note that *n* is an approximate sample size – the *achieved sample size* may different from *n* due to the removal of high value items from the database during sampling. High value extraction can also change the *achieved sampling interval* slightly (that is, the final population size less high value items, divided by the final sample size).

A Further Note on High Value Items

The planning function computes a sampling interval, tolerable total taintings, and sample size required to meet the objectives on upper error limit and confidence, without consideration of high value items. When the sample is extracted, removal of HV items decreases both the population value and the sample size from the planned values. If the expected and maximum tolerable error *rates* in the population after removal of HV items is the same as that used in planning (i.e., X_{exp} / N and X_{max} / N respectively), removal HV

items will not affect the ability of the sample to achieve the planned precision to any noticeable degree.

3.2 Planning when Population Size is Unknown

Inputs:

X _{max}	Maximum tolerable monetary value of all errors in population (e.g., materiality for this test) – absolute value	$X_{\rm max} > 0$
X _{exp}	Expected monetary value of all errors in population – absolute value	$0 < X_{exp} < X_{max}$
BPP	Basic precision pricing (assumed worst- case tainting on errors in population that do not appear in the sample).	$0 < BPP < \infty$
CL	Confidence level (a.k.a. reliability)	0 < <i>CL</i> < 1

Outputs:

SI: average sampling interval; *NIn*

TT: Tolerable tainting in sample.

Other quantities referred to below:

x^{*}: tolerable number of errors in sample.

 $\lambda(x, C)$: Poisson factor (a.k.a. UEL factor) for observed number of errors *x* and confidence level *C*.

N: size of population from which sample was drawn. for MUS N = |Y|.

Description:

When population size is unknown, we compute the *SI* directly using the Poisson distribution to approximate the hyper geometric.

We first find successive integers x' and x'+1 such that

$$\frac{x'}{\lambda(x', CL)} \le \frac{X_{\exp}}{X_{\max}} \le \frac{x'+1}{\lambda(x'+1, CL)}$$
(3.8)

The Poisson factors $\lambda(x', CL)$ and $\lambda(x'+1, CL)$ are found using a

search procedure, which we shall now describe. The cumulative Poisson distribution is given by

$$P(\lambda, x) = \sum_{k=0}^{x} e^{-\lambda} \lambda^{k} k!$$
(3.9)

Where k! denotes the factorial of x, i.e. k! = k(k-1)(k-2)...(2).

The Poisson factor $\lambda(x, CL)$ is that value of λ satisfying

$$P(\lambda(x, CL), x) \approx 1 - CL \tag{3.10}$$

IDEA computes Poisson factors by incrementing a candidate λ until $\lambda(x, CL)$ satisfying (3.10) is found. The computation is accurate to five decimal places. One interpretation of the Poisson factor is as follows: $\lambda(x, CL)$ is the *CL* - upper limit on the *expected* number of errors in any future sample, when an examined sample is found to have *x* errors.

Having found x' and x'+1, and their corresponding Poisson factors, we next interpolate to find values x^* and $\lambda *$ such that $x * / \lambda * = X_{exp} / X_{max}$. Thus,

$$\frac{x^* - x'}{x' + 1 - x'} = \frac{\lambda^* - \lambda(x', CL)}{\lambda(k' + 1, CL) - \lambda(x', CL)} .$$
(3.11)

Substituting $x^* = \lambda^* X_{exp} / X_{max}$ and solving, gives

$$\lambda^* = \frac{x' \lambda(k'+1, CL) - x'' \lambda(k', CL)}{\left[\lambda(x'+1, CL) - \lambda(x', CL)\right] \times X_{\exp} / X_{\max} - 1}$$
(3.12)

 $x^* = \lambda^* \times X_{\exp} / X_{\max}$

Finally, we compute the *SI* and tolerable taintings, scaling by the BPP to get

$$SI = \frac{X_{\max}}{\lambda^* \times BPP}$$
(3.13)

 $TT = x * \times BPP$

A Further Note on High Value Items

As in the case for known population size, removal of HV items decreases both the population value and the sample size from the planned values. Once again, if X_{exp} and X_{max} are scaled to reflect the decrease in population size, removal HV items will not affect the ability of the sample to achieve the planned precision to any noticeable degree.

3.3 Warning Messages from Planning

IDEA generates warning messages in pop-up windows if certain conditions are encountered during MUS Planning. Warnings do not prevent the user from executing the planning operation as planned, but are merely notifications. The table below shows these conditions and the resulting warnings.

Condition	Warning
<i>n</i> > 0.5 × number of records in population database	WARNING: The computed sample size is more than half of the number of records in the population database. Results from evaluating this sample may be overly conservative. Classical Variables sampling may be more appropriate for this test. Consultation with a statistical sampling expert is recommended.
$X_{\max} > N$	Warning: the tolerable error exceeds the total population value.
$X_{\rm exp} > 0.5 X_{\rm max}$	Warning: you have specified an expected error of more than half the tolerable error. Monetary unit sampling is inefficient for such high-error populations. Classical Variables sampling may be more appropriate for this test.
High Value items extracted to separate file and High Value Amount > Sampling Interval	Warning: you have specified a high value amount greater than the sampling interval. Items having recorded amounts between the high value amount and the sampling interval may be sampled multiple times; statistical projection of errors found in such items is not recommended. Change HVA / Confirm.

Note: IDEA does not support sampling intervals of less than 1.0 or sample sizes of greater than 1,000,000. If the planning function computes values

Data Analysis for Auditors

outside of these accepted limits, it fails and the user is asked to try different parameters.

4. Sample Evaluation

4.1 Single Sample Evaluation when Fewer than 20 Errors are found

Inputs:

Y	Total monetary value of population from which sample was drawn	
n	Sample size	$n > 0$, $n \ge$ number of records in sample database, n < Y
BPP	Basic precision pricing (assumed worst-case tainting on errors in population that do not appear in the sample).	0 < <i>BPP</i> < 10
CL	Confidence level (a.k.a. reliability)	$0.0 < CL \le .995$
<i>RA</i> (<i>k</i>)	Recorded amount for k^{th} record in main sample, for $k = 1 \dots n$ (obtained from sample database)	$RA(k) \neq 0$
AA(k)	Audited amount for k^{th} record in main sample, for $k = 1 \dots$ number of records (obtained from sample database)	$\left \frac{RA(k) - AA(k)}{RA(k)}\right \le BPP$

Outputs:

Gross UEL o/s NoHV:upper precision limit of overstatements
projected from main sample,Gross UEL u/s NoHV:upper precision limit of understatements

projected from main sample, disregarding overstatements and not including errors in HV items

Gross MLE o/s NoHV: most likely error for understatements projected from main sample, disregarding understatements and not including errors in HV items

Gross MLE u/s NoHV:	most likely error for understatements projected from main sample, disregarding overstatements and not including errors in HV items
Net UEL o/s NoHV:	upper error limit for overstatements net of understatements not including errors in HV items
Net UEL u/s NoHV:	upper error limit for understatements net of overstatements not including errors in HV items
Net MLE o/s NoHV:	MLE for overstatements net of understatements not including errors in HV items
Net MLE u/s NoHV:	MLE for understatements net of overstatements not including errors in HV items
HVE o/s:	total overstatements in high value items
HVE u/s:	total understatements in high value items
Gross UEL o/s wHV:	upper error limit for overstatements, disregarding understatements but including errors in HV items
Gross UEL u/s wHV:	upper error limit for understatements, disregarding overstatements but including errors in HV items
Gross MLE o/s wHV:	most likely error for overstatements, disregarding understatements but including errors in HV items
Gross MLE u/s wHV:	most likely error for understatements, disregarding overstatements but including errors in HV items
Net UEL o/s wHV:	upper error limit for overstatements net of understatements, including errors in HV items
Net UEL u/s wHV:	upper error limit for understatements net of overstatements, including errors in HV items

Net MLE o/s wHV:	MLE for overstatements net of understatements, including errors in HV items
Net MLE u/s wHV:	MLE for understatements net of overstatements, including errors in HV items
Precision o/s:	Total precision of UEL for overstatements
Precision u/s:	Total precision of UEL for understatements

Other quantities referred to below:

SI:	average sampling interval
t _{o/s} (k):	tainting for the k^{th} overstatement in main sample, ordered from largest to smallest.
t _{u/s} (k):	tainting for the $k^{\mbox{\tiny th}}$ understatement in main sample, ordered from largest to smallest.
X _{0/s} :	total number of overstatements in main sample
X _{u/s} :	total number of understatements in main sample
N:	size of population from which sample was drawn. For MUS, $N = Y $.

Description:

IDEA uses cell evaluation [2] to compute a most likely error and an upper error limit for the population. These quantities are computed separately for overstatements and understatements. The resulting (gross) estimates are then combined to produce a net most likely error and a net upper error limit for the population. The hyper geometric probability distribution is used throughout.

Note that only the separate (Gross) upper error limit for overstatements has been extensively tested. The guidance of a statistical expert should be sought before the gross and net most likely errors, or the net upper error limit, are used in any audit decision.

Sampling Interval

The sampling interval is computed as

SI = NI n

Note that previous versions of IDEA have used the sampling interval used in the extraction, say SI^* , which is not exactly equal to M/n due to the presence of a remainder when N is divided by SI^* - that is, $n = \text{floor}(N/SI^*)$. Since the derivation of MUS calculations [2], [3], is based on a SI of exactly M/n, this quantity is now used in the evaluation. The value of SI^* used for extracting the sample can be found in the sample history, if the sample was extracted using IDEA's MUS extraction.

Taintings

For each item in the main sample database, let

$$\widetilde{t}(k) = \frac{RA(k) - AA(k)}{RA(k)} .$$
(4.1a)

We denote taintings of overstatements and understatements respectively by

$$t_{o/s}(k) = \left| \widetilde{t}(k') \right|; \quad RA(k') > AA(k') \quad (4.1b)$$
$$t_{u/s}(k) = \left| \widetilde{t}(k') \right|; \quad RA(k') < AA(k').$$

Note that recorded amounts may be either positive or negative. The following table gives possible ranges for tainting values, under the assumption that the audited amount has the same sign as the recorded amount.

	Recorded amount Sign	
	Positive: RA(k) > 0	Negative: RA(k) < 0
Overstatement: RA(k) > AA(k)	$0 < \widetilde{t}(k) < 1$	$-\infty < \widetilde{t}(k) < 0$
	0 < t(k) < 1	$0 < t(k) < \infty$
Understatement:	$-\infty < \widetilde{t}(k) < 0$	$0 < \tilde{t}(k) < 1$
RA(k) < AA(k)	$0 < t(k) < \infty$	0 < t(k) < 1

Since they lead to taintings of greater than 100%, understatements for positive recorded amounts or overstatements for negative recorded amounts can dominate the results. IDEA generates a warning suggesting that a sampling expert be consulted, if any taintings greater than 100% are encountered.

Taintings of both overstatements and understatements are ordered from largest to smallest.

If any tainting is greater than the input *BPP*, the *BPP* must be adjusted to a value greater than or equal to the maximum tainting before the evaluation can proceed.

Projection of Most Likely Errors from Sample

We here describe the procedure for projecting the most likely errors from the sample, without consideration of errors in high value amounts. The formulae used are

Gross MLE o/s NoHV =
$$SI \times \sum_{k=1}^{x_{o/s}} t_{o/s}(k)$$
 (4.2)
Gross MLE u/s NoHV = $SI \times \sum_{k=1}^{x_{u/s}} t_{u/s}(k)$

Projection of Upper Error Limits from Sample

We here describe the procedure for projecting the upper error limits from the sample, without consideration of errors in high value amounts. We shall provide details for *Gross UEL o/s NoHV*; the computation for understatements is directly analogous. For some quantities below we omit the subscript *o/s*; all errors and taintings in the equations below should be understood to be overstatements. The upper limit is given by

Gross UEL o/s NoHV =
$$F(x) \times SI$$
. (4.3)

The factor F(x) is computed using the cell evaluation method [2]. This technique takes advantage of certain constraints on the possible population errors, which are a consequence of the cell extraction method employed by IDEA's MUS Extraction routine. It is a recursive calculation employing hyper geometric factors $\lambda_H(k, n, N, CL)$, which will be defined below.

The recursive procedure is given by:

$$F(0) = \lambda_H (0, n, N, CL) \times BPP$$

$$F(k) = \max \left\{ F(k-1) + t(k), \lambda_H (k, n, N, CL) \times \frac{1}{k} \sum_{i=1}^k t(i) \right\}$$

$$(4.4a)$$

$$(4.4b)$$

for k = 1, ..., x.

We now define the hyper geometric factors. Let $X_{\max}(k)$ be the *CL*-upper precision limit on the number of errors in a population of size *N*, when *k* errors are found in a sample of size *n*. Note that for any *X*, it can be inferred with a confidence level 1 - H(k, n, X, N) that the true population error rate is *less than X*, where H(k, n, X, N) is the hyper geometric risk function defined in Eq. 3.1. This is because, if the error rate were as high as *X*, then the number of sample errors would be *greater than* the observed value of *x* with a probability of 1 - H(x, n, X, N). Therefore, the upper precision limit on the number of errors in the population at a confidence level of *CL* = 1 - H(x, n, X, N) is X - 1. By the above reasoning, $X_{\max}(k)$ is

$$X_{\max}(k) = \min \left[X \ni H(k, n, X+1, N) \le 1 - CL \right].$$
 (4.5)

The hyper geometric factor is given by

$$\lambda_H(k,n,N,CL) = X_{\max}(k)/SI \quad . \tag{4.6}$$

 $\lambda_H(k,n,N,CL)$ is the expected number of errors in a sample of size *n* drawn from a population of size *N* in which there are a total of $X_{\max}(k)$ errors. Note that the Poisson factor $\lambda(k,CL)$ as defined in Section 3 is a conservative approximation to the hyper geometric risk factor as defined above, with the accuracy of the approximation improving, as the population size grows large.

Incorporation of Errors in High Value Items

1

We now describe the incorporation of errors in high value amounts and the calculation of the final Gross and Net error projections appearing on the result tab. The equations used are:

HVE o/s =
$$\sum_{RA(i)>AA(i)} RA(i) ; i \text{ in HV database}$$
 (4.7a)

$$HVE \ u/s = \sum_{AA(i)>RA(i)} AA(i) - RA(i) \ ; \ i \text{ in HV database}$$
(4.7b)

$$Gross UEL \ o/s \ Whv = Gross \ UEL \ o/s \ noHV + HVE \ o/s$$

$$(4.7c)$$

Gross MLE u/s wHV	= Gross MLE u/s noHV + HVE u/s	(4.7f)
Net UEL o/s noHV	= Gross UEL o/s noHV – Gross MLE u/s noHV	(4.7g)
Net UEL u/s noHV	= Gross UEL u/s noHV – Gross MLE o/s noHV	(4.7h)
Net MLE o/s noHV	= Gross MLE o/s noHV – Gross MLE u/s noHV	(4.7i)
Net MLE u/s noHV	= Gross MLE u/s noHV – Gross MLE o/s noHV	(4.7j)
Net UEL o/s wHV	= Gross UEL o/s wHV – Gross MLE u/s wHV	(4.7k)
Net UEL u/s wHV	= Gross UEL u/s wHV – Gross MLE o/s wHV	(4.7I)
Net MLE o/s wHV	= Gross MLE o/s wHV – Gross MLE u/s wHV	(4.7m)
Net MLE u/s wHV	= Gross MLE u/s wHV – Gross MLE o/s wHV	(4.7n)
Precision o/s	= Gross UEL o/s noHV – Gross MLE o/s noHV	(4.70)
Precision u/s	= Gross UEL u/s noHV – Gross MLE u/s noHV	(4.7p)

4.2 Combined Evaluation of Multiple Samples

Inputs:

Y(k):	Total monetary value of population from which <i>k</i> th error in total sample was drawn.	Y > n
n(k):	Size of component sample in which <i>k</i> th error in combined sample was found (obtained from sample databases)	<i>n</i> > 0

Appendices

<i>BPP</i> (<i>k</i>):	Basic precision pricing corresponding to <i>k</i> th error in combined sample.	0 < <i>BPP</i> < 10
CL	Confidence level (a.k.a. reliability)	$0 < CL \le .995$
RA(k)	Recorded amount for k^{th} record in combined sample, for $k = 1 \dots n$ (obtained from sample databases)	$RA(k) \neq 0$
AA(k)	Audited amount for <i>k</i> th record in main sample, for <i>k</i> = 1 sum of sample sizes (obtained from sample databases)	$\left \frac{RA(k) - AA(k)}{RA(k)}\right \le BPP(k)$

Outputs:

Gross UEL o/s NoHV:	upper precision limit of overstatements projected from main sample, disregarding understatements and not including errors in HV items
Gross UEL u/s NoHV:	upper precision limit of understatements projected from main sample, disregarding overstatements and not including errors in HV items
Gross MLE o/s NoHV:	most likely error for understatements projected from main sample, disregarding understatements and not including errors in HV items
Gross MLE u/s NoHV:	most likely error for understatements projected from main sample, disregarding overstatements and not including errors in HV items
Net UEL o/s NoHV:	upper error limit for overstatements net of understatements not including errors in HV items

Net UEL u/s NoHV:	upper error limit for understatements net of overstatements not including errors in HV items
Net MLE o/s NoHV.	MLE for overstatements net of understatements not including errors in HV items
Net MLE u/s NoHV.	MLE for understatements net of overstatements not including errors in HV items
HVE o/s:	total overstatements in high value items
HVE u/s:	total understatements in high value items
Gross UEL o/s wHV:	upper error limit for overstatements, disregarding understatements but including errors in HV items
Gross UEL u/s wHV:	upper error limit for understatements, disregarding overstatements but including errors in HV items
Gross MLE o/s wHV:	most likely error for overstatements, disregarding understatements but including errors in HV items
Gross MLE u/s wHV:	most likely error for understatements, disregarding overstatements but including errors in HV items
Net UEL o/s wHV:	upper error limit for overstatements net of understatements, including errors in HV items
Net UEL u/s wHV:	upper error limit for understatements net of overstatements, including errors in HV items
Net MLE o/s wHV:	MLE for overstatements net of understatements, including errors in HV items
Net MLE u/s wHV:	MLE for understatements net of overstatements, including errors in HV items
Precision o/s:	Total precision of UEL for overstatements

Other quantities referred to below:

<i>SI</i> (<i>k</i>):	average sampling interval for component sample in which k^{th} error in main sample was found.
<i>RA(k</i>):	recorded amount for k th record in combined sample
<i>AA(k</i>):	audited amount for k^{th} record in combined sample
<i>t_{o/s}(k</i>):	tainting for the k^{th} overstatement in combined sample, ordered from largest to smallest.
<i>t_{u/s}(k</i>):	tainting for the k^{th} understatement in combined sample, ordered from largest to smallest.
X _{0/s} :	total number of overstatements in combined sample
<i>X_{u/s}</i> :	total number of understatements in combined sample
<i>N</i> (<i>k</i>):	size of population from which sample k was drawn. For MUS, $N(k) = Y(k) $.
N _t :	size of combined population: $N_t = \sum N(k)$
<i>n</i> (<i>k</i>):	size of sample <i>k</i>
n _t .	size of combined sample: $n_t = \sum n(k)$

Overview:

Suppose a monetary unit sample is taken in each of several distinct populations. IDEA uses the "global combined bound" proposed by Neter et al. [3] to determine the most likely error and upper error limit for the combined population – that is, the union of all of the sampled populations. This bound reduces to the bound introduced by K. Stringer, for the case of a single sample. The Stringer bound is most appropriate for samples containing few errors¹; it tends to over-estimate the upper error limits when many errors are observed.

The most likely and upper error limits are computed separately for overstatements and understatements. The resulting (gross) estimates are then combined to produce a net most likely error and a net upper error limit for the combined population. The hyper geometric probability distribution is used throughout.

Note that only the separate (Gross) upper error limit for overstatements has been extensively tested. The guidance of a statistical expert should be sought before the gross and net most likely errors, or the net upper error limit, are used in any audit decision.

The Stringer bound is slightly more conservative than that produced by the cell evaluation method. However, unlike the cell evaluation method, the Stringer bound is valid for monetary unit samples drawn using either cell selection or simple random sampling. IDEA's "combined evaluation" option can also be used to obtain the Stringer UEL for a single sample, if this is desired.

We note that [3] does not consider the use of BPPs other than 100%. In IDEA, we have extended the "worst-case" basic precision proposed in [3], to allow for arbitrary BPPs.

Details:

Sampling Interval:

The sampling intervals are computed as

$$SI(k) = N(k) / n(k)$$
 (4.8)

See comments in Section 4.1 for more details.

Taintings:

Taintings are calculated for each record in the combined sample as in Eq. (4.1a), and sorted into groups of overstatements and understatements, $t_{o/s}(k)$ and $t_{u/s}(k)$, as in E.g. (4.1b). The index *k* of a tainting denotes its position in a list of taintings from all samples, sorted from largest to smallest.

Projection of Most Likely Errors from Sample

The formula used for computing the most likely total overstatement in the combined population, based on the overstatements observed in all samples, is

Gross MLE o/s NoHV =
$$\sum_{k=1}^{x_{o/s}} t_{o/s}(k) \times SI(k) .$$
(4.9)

Here $t_{o/s}(k)$ is the tainting of the k^{th} tainting in the combined sample, ordered from largest to smallest. Taintings are described further in Section 4 – *taintings*. Note that (4.9) does not include errors in high value amounts stored in separate high value files – these errors are added onto *Gross MLE o/s NoHV* in computing the final error estimates (see *Calculation of Gross and Net Error Projections*).

Projection of Upper Error Limits from Sample

We here describe the procedure for projecting the upper error limits from the sample, without consideration of errors in high value amounts. We shall provide details for *Gross UEL*_{o/s} *noHV*; the computation of *Gross UEL*_{u/s} *noHV* is directly analogous. All errors and taintings in the equations below should be understood to be overstatements. The upper limit is given by

$$GrossUELo/s NoHV = BP + \sum_{k=1}^{x} PGW(k).$$
(4.10)

The first term is the basic precision for the combined sample. For conservatism, it is taken to be the largest basic precision of all of the component samples:

$$BP = \max_{k} \left[\lambda_{H} \left(0, n_{t}, N_{t}, CL \right) \times BPP(k) \times SI(k) \right]$$
(4.11)

PGW(k) is the precision gap widening for the k^{th} tainting in the combined sample. It is given by

$$PGW(k) = \left[\lambda_H(k, n_t, N_t, CL) - \lambda(k - 1, n_t, N_t, CL)\right] t(k) SI(k)$$
(4.12)

In (4.11) and (4.12), $\lambda_H(k, n, N, CL)$ is the hyper geometric factor, as defined by Eq.s (4.5), (4.6)² Also, n_t and N_t are the combined sample and population sizes, and t(k) and SI(k) are the tainting and sampling interval corresponding to the k^{th} tainting in the combined sample.

Calculation of Gross and Net Error Projections

The equations for computing the final Gross and Net error projections appearing on the result tab are exactly as Eq.s (4.7), (4.8), except that the misstatements in high value amounts are totaled over all specified high value files.

Footnotes:

 For example, [3] describes extensive simulations with combined samples having a total of between 10 and 50 overstatements; for a specified target confidence level of 95%, the achieved "empirical" confidence level (that is, the frequency with which the bound exceeded the actual population error) was always above 99%.

2. Note that the combined population and sample sizes are used in computing these hyper geometric factors. This prevents precision gap widening from becoming negative, which could occur if component sample and population sizes were used. Though the factors thus computed are slightly larger than would be those computed with component sample and population sizes, they are still more efficient than the corresponding Binomial or Poisson approximations.

4.3 Single Sample Evaluation when More than 20 Errors are Found

Inputs:

Y(k):	Total monetary value of population from which <i>k</i> th error in total sample was drawn.	Y > n
n(k):	Size of component sample in which <i>k</i> th error in combined sample was found (obtained from sample databases)	<i>n</i> > 0
Limits	Desired precision interval endpoints	<i>Upper</i> or <i>Both</i>
CL	Confidence level (a.k.a. reliability)	$0.1 \le CL \le .995$ if <i>Limits = Both</i> ,
		or
		$0.55 \le CL \le .995$ if <i>Limits = Upper</i> .
RA(k)	Recorded amount for k^{th} record in combined sample, for $k = 1 \dots n$ (obtained from sample databases)	$RA(k) \neq 0$ AND all $RA(k)$ have same sign
AA(k)	Audited amount for k^{th} record in main sample, for $k = 1 \dots$ sum of sample sizes (obtained from sample databases)	

Outputs:

UEL NoHV: upper precision limit of overstatements projected from main sample, disregarding understatements and not including errors in HV items

LEL NoHV ¹ :	lower precision limit of overstatements projected from main sample, disregarding understatements and not including errors in HV items	
MLE NoHV:	most likely error for understatements projected from main sample, disregarding understatements and not including errors in HV items	
HVE:	total overstatements in high value items	
UEL wHV:	upper error limit for overstatements, disregarding understatements but including errors in HV items	
LEL wHV ¹ :	lower precision limit of overstatements projected from main sample, disregarding understatements and not including errors in HV items	
MLE wHV:	most likely error for overstatements, disregarding understatements but including errors in HV items	
Precision:	Total precision of UEL for overstatements	
<i>s</i> _r :	Estimated standard deviation of ratios	
Other quantities referred to below:		
<i>RA</i> (<i>k</i>):	recorded amount for k^{th} record in main sample	
<i>AA(k</i>):	audited amount for k^{th} record in main sample	
<i>r</i> (<i>k</i>):	ratio of for the k^{th} error in main sample, ordered from largest to smallest.	
\overline{r} :	Average ratio of audited amount to recorded amount	
<i>X</i> :	total number of errors in main sample	
U(R):	Reliability factor. The reliability factor is that value such that the area under the probability distribution function from $-\infty$ to $U(R)$ is equal to R .	

Description:

The high-error evaluation uses Classical Proportional to Size (PPS) sampling theory [4] to estimate the most likely error in the population, the upper error limit, and optionally the lower error limit. The Student's t probability distribution is used throughout.

In contrast with the low-error rate method, separate estimates are not computed for each of over and understatements. Rather, estimates are based on both over and understatements. The following facts should hold in order for this method to be suitable [4]:

- 1. All recorded amounts must have the same sign (either positive or negative are allowed).
- If positive recorded amounts are used, the majority of the errors should be overstatements, and any understatements should be small. One or more large understatements can inappropriately dominate the projections. If negative recorded amounts are used, the majority of errors should be understatements.
- 3. It is recommended that the sample should contain at least 20 errors in order for the assumptions on the sampling distribution to hold [4]. At least two errors are required in order to perform the calculation.

Since both low and high-error samples are extracted in the same way, a sample planned with the low-error (cell-based) planning function can be evaluated with the high-error (classical PPS) evaluation function.

Ratios, Mean Ratio, and Standard Deviation

For each item in the main sample database, a ratio is computed as

$$r(k) = \frac{AA(k)}{RA(k)} . \tag{4.13}$$

Classical PPS theory has not been tested for cases in which negative ratios are present (i.e., there are audited amounts having different sign from the corresponding recorded amounts). IDEA generates a warning if any negative ratios are encountered. Similarly, ratios greater than one can dominate the results; such ratios result from understatements on positive recorded amounts, or overstatements on negative recorded amounts. IDEA generates a warning whenever a ratio of greater than 2.0 is encountered. (A ratio of 2 results from 100% understatement of a positive recorded amount or 100% overstatement of a negative recorded amount.)

The mean of ratios for all items is

$$\bar{r} = \frac{1}{n} \sum_{k=1}^{n} r(k)$$
(4.14)

Note that negative ratios are allowed and are not inconsistent with the assumed normality of the sampling distribution. However, the behavior of the estimate when negative ratios are present is not well characterized by empirical tests; hence IDEA generates a warning whenever negative ratios are encountered.

The estimated standard deviation of the ratios is

$$s_r = \sqrt{\frac{1}{n-1} \left(\sum_{k=1}^n r^2(k) - n \, \overline{r}^2 \right)}$$
(4.15)

Projection of Most Likely Errors from Sample

We here describe the procedure for projecting the most likely errors from the sample, without consideration of errors in high value amounts. The formulae used is

$$MLE \ NOHV = \ N \times \overline{r} \tag{4.16}$$

Projection of Error Limits from Sample

We here describe the procedure for estimating either the upper limit or upper and lower limits on the monetary value of the total error in the population. We first compute a reliability factor² U(R) from the Student's *t* tables. The degrees of freedom used for this computation is df = x - 1. For a one sided interval, the reliability is R = CL For a two-sided interval, the reliability is $R = 1 - \frac{1-CL}{2}$.

The precision is computed using the following formula:

$$Precision = \frac{U(R) N s_r}{\sqrt{n}}$$
(4.17)

The limits are then:

$$UEL NOHV = MLE NOHV + (Precision \times sgn(MLE NOHV))$$
(4.18a)

$$LEL NOHV^{1} = MLE NOHV - (Precision \times sgn(MLE NOHV))$$
(4.18b)

$$UEL wHV = UEL NOHV + HVE$$
(4.18c)

$$LEL NOHV^{1} = LEL NOHV + HVE$$
(4.18d)

Note that sgn(x) = -1 is x is negative, and sgn(x) = +1 if x is positive.

Footnotes:1 Computed only if *Limits* = *Both*

4.5 Warning Messages from Evaluation

IDEA displays notification or warning messages in the sample result, if certain conditions are encountered during MUS Evaluation. The table below shows these conditions and the resulting messages.

Condition	Warning
$n > 0.5 \times$ number of records in population database	Your sample size exceeds half the number of records in your population database. In these circumstances, the upper error limit projected by an MUS evaluation is conservative. Classical variables sampling may be more statistically efficient for the population you are considering.
Change to <i>BPP Low-error</i> single evaluation only	Warning: The Basic Precision Pricing (BPP) has been changed since this sample was planed. This happens whenever tainting is found in the sample that exceeds the planned BPP.
Sample database information shows that high value items were extracted to a file, but no HV file is specified for the evaluation.	High value items for this population were extracted to the file [<i>High Value Filename</i>], but no high value file was specified for the evaluation. The above results do not include errors in high value items.
Sample was not generated with IDEA's MUS Plan – Extract AND recorded amounts exceeding the sampling interval were found in the sample.	Warning: One or more recorded amounts exceeding the sampling interval have been found in the file [sample database name].
Sample was generated with IDEA's MUS Plan – Extract, AND HV items were extracted to a separate database, AND HV item found in sample.	Warning: One or more recorded amounts exceeding the high value amount have been found in the file [sample database name].
t(k) > 1 for some k. Low- error single evaluation and combined evaluation only.	Warning: one or more taintings of greater than 100% were found in the sample. Consultation with a statistical expert is recommended.

r(k) > 2 for some k. High- error single evaluation only.	Warning: one or more large understatement were found in the sample. Consultation with a statistical expert is recommended.
r(k) < 0 for some k. High- error single evaluation only.	Warning: one or more audited amounts whose signs differ from corresponding book values were found in the sample. Consultation with a statistical expert is recommended.
x < 20 . High-error single evaluation only	Warning: This sample contains fewer than 20 errors. You have selected high error rate evaluation, which is recommended for samples containing 20 or more errors. You may wish to use the low error rate evaluation instead

References

- [1] Canadian Institute of Chartered Accountants, *Users' Guide: IDEA for Windows Version 2*, CICA Press, 1997.
- [2] D.A. Leslie, A.D. Teitlebaum, R.J. Anderson, *Dollar Unit Sampling a Practical Guide for Auditors*. Copp, Clark Pitman, Toronto, 1979.
- [3] J. Neter, H.S. Kim and L.E. Graham, "On combining Stringer bounds for independent monetary unit samples from several populations." *Auditing*, v. 4, 1984. pp. 74-88.
- [4] Donald M. Roberts, *Statistical Auditing*, AICPA 1978; p. 184.

Notes

•	• •	• •	• •	• •	• •	• •	• •	• •	•	• •	• •	•	•	•	•	•	•	•	-	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	-	•	• •	• •	•	-	•	•	•	-	-	• •	• •	-	•	•	•	• •	• •	• •	-
•	• •	•	•	-	• •	• •	-	•	•	• •	• •	•	•	•	•	•	•	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	-	•	•	•	•	•	• •	• •	• •	•	•	•	•	• •	••	-	•	•
•		•	•	-	• •	• •	-	•	•	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	-	•	-	•	•	•	• •	• •	•	-	•	•	•	• •	• •	-	•	•
•		-	•	-	• •	• •	-	•	-	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	•	-	•	•	•	• •	• •	•	•	•
•	• •	•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	• •	• •	•••	-	•	•	•	•	•	• •	• •	• •	-	•	•	•		••	•	•	•
•		•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	-	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	-	•	•	•	• •	••	•	•	•
•	• •	-	•	-	• •	• •	-	•	-		• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	-	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	• •	-	•	•	•			•	•	•
•		•	•	-	• •	• •	-	•	-		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•	•	•	•	•	•	•		• •	• •	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•		•	• •			•	•	•	•	•	•	•		•	• •	• •	• •	• •		•	•	•	•	•	•	• •	• •		-	•	•	•	•		• •	-	•	•
•		•	•	-	• •		-	•	-			•	-	•	•	•	•	•	-	•		•		•		-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •			-	•	-	•	•	-	• •	• •		-	•	•	•			• •	•	•
•		-	•	-	• •		-	•	•			•	-	•	•	-	-	•	•	•		•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •			•	•	-	•	•	•	• •	• •		-	•	-	•			•	-	•
		•	•	-	• •		-	•	-			•	•	•	•	•	•	•	•							-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •	• •		-	•	•	•	•	-	• •	• •	• •	-	•	•	•			• •	-	•
		-	•	-	• •		-		•			•	-		•		•	•	•	•		•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •					•	-	•	•	•				-		•	•				-	•
				-			-		-				-		•		•	•	-							-	•	•	•	•	•	•	•	•	•	•	•	-								-	•	•	-				-		•	•				-	
				-			-	•	-				-				-	•								-	•	•	•	•	-	•	•	•	•	-	•	-								-	•	•	-				-		•	•				-	
		-		-			-		-				-			-	-	•								-	•	•	•	•		•		•	•		•	-								-	•		-				-		-	•				-	
				-			-		-									•										•	•	•		•	•	•	•		•									-			-				-							-	
_													_	_	_		_					_	_					_					_	_						_	_	-										_			_		_				
•																																																								•	•		•••		
•		•	•	-		••	•	•				•	•	•	•	•	•	•	-								•	•	•	•	•	•	•	•	•	•	•							•	•	•	•	•	-	•••	•	•	•	•	-	•		• •		•	•
•	••	•	•	•	•	••	•	•	•	•••	•••	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•		•	•	•	•	•	•	• •	• •	•	•	•	•	•	••	••	•	•	•
•	••	•	•	•	• •	• •	•	•	•	• •	••	•	•	•	•	•	•	•	•	• •	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	•	•	•	•	• •		•	•	•

Notes

•	• •	• •	• •	• •	• •	• •	• •	• •	• •	• •	• •	•	•	•	•	•	•	•	-	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	-	•	• •	• •	•	-	•	•	•	-	-	• •	• •	-	•	•	•	• •	• •	• •	-
•	• •	•	•	-	• •	• •	-	•	•	• •	• •	•	•	•	•	•	•	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	-	•	•	•	•	•	• •	• •	• •	•	•	•	•	• •	••	-	•	•
•		•	•	-	• •	• •	-	•	•	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •	• •	• •	•	-	•	-	•	•	•	• •	• •	•	-	•	•	•	• •	• •	-	•	•
•		-	•	-	• •	• •	-	•	-	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	•	-	•	•	•	• •	• •	•	•	•
•	• •	•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	• •	• •	•••	-	•	•	•	•	•	• •	• •	• •	-	•	•	•		••	•	•	•
•		•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	-	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	-	•	•	•	••	••	•	•	•
•	• •	-	•	-	• •	• •	-	•	-		• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	-	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	• •	-	•	•	•			•	•	•
•		•	•	-	• •		-	•	-		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•	•	•	•	•	•	•		• •	• •	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•		•	• •			•	•	•	•	•	•	•		•	• •	• •	• •	•••		•	•	•	•	•	•	• •	•••		-	•	•	•	•		• •	-	•	•
•		•	•	-	• •		-	•	-			•	-	•	•	•	•	•	-	•		•		•		-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •			-	•	-	•	•	-	• •	• •		-	•	•	•			• •	•	•
•		-	•	-	• •		-	•	•			•	-	•	•	-	-	•	•	•		•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •			•		-	•	•	•	• •	• •		-	•	-	•			•	-	•
		•	•	-	• •		-	•	-			•	•	•	•	•	•	•	•		• •	•		•		-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •	• •		-	•	•	•	•	-	• •	• •	• •	-	•	•	•			• •	-	•
		-	•	-			-		•			•	-		•		•	•	•	•		•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •					•	-	•	•	•				-		•	•				-	•
				-			-		-				-		•		•	•	-							-	•	•	•	•	•	•	•	•	•	•	•	-								-	•	•	-				-		•	•				-	
				-			-	•	-				-				-	•								-	•	•	•	•	•	•	•	•	•	•	•	-								-	•	•	-				-		•	•				-	
		-		-			-		-				-			-	-	•								-	•	•	•	•		•		•	•		•	-								-	•		-				-		-	•				-	
				-			-		-									•										•		•			•	•	•		•									-			-				-							-	
_													_	_	_		_					_	_					_					_	_						_	_	_										_			_		_				
•																																																								•	•		•••		
•		•	•	-		••	•	•				•	•	•	•	•	•	•	-								•	•	•	•	•	•	•	•	•	•	•							•	•	•	•	•	-	•••	•	•	•	•	-	•		• •		•	•
•	••	•	•	•	•	••	•	•	•	•••	•••	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•		•	•	•	•	•	•	• •	• •	•	•	•	•	•	••	••	•	•	•
•	••	•	•	•	• •	• •	•	•	•	• •	••	•	•	•	•	•	•	•	•	• •	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	•	•	•	•	• •		•	•	•

Notes

•	• •	• •	• •	• •	• •	• •	• •	• •	• •	• •	• •	•	•	•	•	•	•	•	-	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	-	•	• •	• •	•	-	•	•	•	-	-	• •	• •	-	•	•	•	• •	• •	• •	-
•	• •	•	•	-	• •	• •	-	•	•	• •	• •	•	•	•	•	•	•	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	-	•	•	•	•	•	• •	• •	• •	•	•	•	•	• •	••	-	•	•
•		•	•	-	• •	• •	-	•	•	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	-	•	-	•	•	•	• •	• •	•	-	•	•	•	• •	• •	-	•	•
•		-	•	-	• •	• •	-	•	-	• •	• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	•	-	•	•	•	• •	• •	•	•	•
•	• •	•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	• •	• •	•••	-	•	•	•	•	•	• •	• •	• •	-	•	•	•		••	•	•	•
•		•	•	-	• •	• •	-	•	•		• •	•	•	•	•	•	-	•	•	•	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	-	•	•	•	••	• •	•	•	•
•	• •	-	•	-	• •	• •	-	•	-		• •	•	-	•	•	•	•	•	•	•	• •	•	•	•	• •	-	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •	• •	• •		-	•	-	•	•	-	• •	• •	• •	-	•	•	•			•	•	•
•		•	•	-	• •	• •	-	•	-		• •	•	•	•	•	•	•	•	•	•	• •	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•	•	•	•	•	•	•		• •	• •	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•		•	• •			•	•	•	•	•	•	•		•	• •	• •	• •	• •		•	•	•	•	•	•	• •	• •		-	•	•	•	•		• •	-	•	•
•		•	•	-	• •		-	•	-			•	-	•	•	•	•	•	-	•		•		•		-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •			-	•	-	•	•	-	• •	• •		-	•	•	•			• •	•	•
•		-	•	-	• •		-	•	•			•	-	•	•	-	-	•	•	•		•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •			•	•	-	•	•	•	• •	• •		-	•	-	•			•	-	•
		•	•	-	• •		-	•	-			•	•	•	•	•	•	•	•							-	•	•	•	•	•	•	•	•	•	•	•	-	• •		• •	• •		-	•	•	•	•	-	• •	• •	• •	-	•	•	•			• •	-	•
		-	•	-			-		•			•	-		•		•	•	•	•		•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	-	• •	• •					•	-	•	•	•				-		•	•				-	•
				-			-		-				-		•		•	•	-							-	•	•	•	•	•	•	•	•	•	•	•	-								-	•	•	-				-		•	•				-	
				-			-	•	-				-				-	•								-	•	•	•	•	-	•	•	•	•	-	•	-								-	•	•	-				-		•	•				-	
		-		-			-		-				-			-	-	•								-	•	•	•	•		•		•	•		•	-								-	•		-				-		-	•				-	
				-			-		-									•										•	•	•		•	•	•	•		•									-			-				-							-	
_													_	_	_		_					_	_					_					_	_						_	_	-										_			_		_				
•																																																								•	•		•••		
•		•	•	-		••	•	•				•	•	•	•	•	•	•	-								•	•	•	•	•	•	•	•	•	•	•							•	•	•	•	•	-	•••	•	•	•	•	-	•		• •		•	•
•	••	•	•	•	•	••	•	•	•	•••	•••	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•		•	•	•	•	•	•	• •	• •	•	•	•	•	•	••	••	•	•	•
•	••	•	•	•	• •	• •	•	•	•	• •	••	•	•	•	•	•	•	•	•	• •	• •	•		•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	• •	• •	•	•	•	•	•	•	•	• •	• •	•	•	•	•	•	• •		•	•	•