





LEARNING OBJECTIVES

The learning objectives of the CAAT module are:

- Understand how to use office automation software for performing various tasks as relevant to services provided by chartered accountant in areas of accounting, assurance and compliance.
- How to use CAAT/SQL queries for data analysis as required.
- How to review controls implemented at various levels/layers such as: Parameters, user creation, granting of access rights, input, processing and output controls in enterprise applications.

1.1 INTRODUCTION

Auditors deal with information in myriad ways encompassing the areas of accounting, assurance, consulting and compliance and most of this information is now available in electronic form. This is true not only in case of large and medium enterprises but even in small enterprises. In case there are enterprises who have still not adapted the digital way, then it is an opportunity for Auditors to help such enterprises to ride the digital wave. Hence, it has become critical for Auditors to understand and use information technology as relevant for the services we provide. It is rightly said: "one cannot audit data which is flying in bits and bytes by using the ancient method of riding on a horse back". We are living in a knowledge era where the skill sets are keys to harnessing the power of technology to be effective as knowledge workers. Computer Assisted Audit Techniques (CAATs) refers to using technology for increasing the effectiveness and efficiency of auditing. CAATs enable auditors to do more with less and add value through the assurance process which is more robust and comprehensive. This chapter provides an overview of the process, approach and techniques which could be used across various technology platforms and in diverse enterprises.

1.2 THE ALL-ENCOMPASSING ELECTRONIC DATA

A great blessing in ancient times was: "May you live in exciting times". Indeed, we are living in exciting times without even being aware of it. We are experiencing how technology innovations are making our life and living simpler by bridging global boundaries and bringing global information on our finger tips. For enterprises as well as professionals, the question is no longer what technology can do for us but what we can do with technology. The question "do I need to use technology" is no longer relevant instead the relevant decision is about "how do I use technology to remain relevant".



Information technology is all pervasive and more so as the government and regulatory agencies also are using technology platform to provide services to citizens and compelling information to be filed in electronic form. The government at all levels has drawn up ambitious plans to implement e-Governance initiatives to improve speed, access and transparency of services. The Information Technology (IT) Act 2000 with IT Amendment Act 2008 and IT rules 2011 provide the regulatory framework and mechanism for recognizing electronic records and electronic transactions thereby facilitating ecommerce and also identifying cybercrimes and providing penalties and compensation for them. Hence, we can expect IT usage to only keep growing in the near future impacting all areas of life more so in our work as professionals.

1.3 AUDITORS AND CAATs

As auditors, we come across computers and communication technology as the most common denominator among our clients, both large and small. Further, we use computers and communication technology for providing services to our clients. In today's complex and rapidly changing technology environment, it is important to master the right techniques which could be used across enterprises and across various technology platforms. Typical of a IT environment are the speed of processing, large capacity of storage, lack of the paper based trails, the radically different way of information processing, the ease of information access, internal controls being imbedded and the ever-present risk of failure of IT and loss of data. All these factors make it imperative for auditors to harness power of technology to audit technology environment by taking into consideration the risks, benefits and advantages. CAATs empower Auditors with the key survival techniques which effective used in any IT environment. CAATs are not specialist tools designed for use by specialist IT auditors but these are common techniques which can be easily mastered to audit in a computerized environment for statutory audit, tax audit and internal audit as also for providing consulting services.

1.4 AUDITORS AND CAATs

CAATs are tools for drawing inferences and gathering relevant and reliable evidence as per requirements of the assignment. CAATs provide direct access to electronic information and empower auditors not only to perform their existing audits more efficiently and effectively but also facilitate them in knowing how to create and execute new type of IT related audit assignments. CAATs provide a mechanism to gain access and to analyze data as per audit objective and report the audit findings with greater emphasis on the reliability of electronic information maintained in the computer system. There is higher reliability on the audit process as the source of the information used provides and greater assurance on audit findings and opinion. CAATs are available in specific general audit software designed for this purpose but the techniques of CAATs can be applied even by using commonly used software such as MS Excel and by using query/reporting features of commonly used application software. CAATs can be used to perform routine functions or activities which can be done using computers, allowing the auditors to spend more time on analysis and reporting. A good understanding of CAATs and know where and when to apply them is the key to success. ICAI has published a guidance note on CAAT and publication titled: "Data Analysis for Auditors" which may be referred for more details.

1.5 NEED FOR CAATs

In a diverse digital world of clients' enterprises, the greatest challenges for an Auditor is to use technology to access, analyze and audit this maze of electronic data. CAATs enable auditors to move from the era of ticks of using pencil or pen to the era of clicks by using a mouse. CAATs will help





auditors to change focus from time-consuming manual audit procedures to intelligent analysis of data so as to provide better assurance to clients and also mange audit risks. Some of the key reasons for using CAATs are:

- 1. Absence of input documents or lack of a visible paper trail may require the use of CAATs in the application of compliance and substantive procedures.
- 2. Need for obtaining sufficient, relevant and useful evidence from the IT applications or database as per audit objectives.
- 3. Ensuring audit findings and conclusions are supported by appropriate analysis and interpretation of the evidence
- 4. Need to access information from systems having different hardware and software environments, different data structure, record formats, processing functions in a commonly usable format.
- 5. Need to increased audit quality and comply with auditing standards.
- 6. Need to identify materiality, risk and significance in an IT environment.
- 7. Improving the efficiency and effectiveness of the audit process.
- 8. Ensuring better audit planning and management of audit resources.

1.6 OBTAINING AUDIT DATA

In most cases where CAATs are used, it becomes necessary to obtain copy of data in their original format for independent analysis. The data has to be obtained in commonly accepted format. It is important to understand the format in which the data is stored in the application which is being audited. If the data is a native format which is not readable by audit software, then it is necessary to use the reporting feature of application software and export this data to commonly recognizable format of audit software. For example, auditor may not be aware of the data structure/tables of a software developed through a vendor by the client. In such case, auditor may have to study the reporting features and use the export feature to get the data in the required format. It is very important to educate the client about the need to obtain copy of the data as required for audit. Based on the audit scope and relevant audit environment, auditor may have to finalize the required approach for getting the data for audit. This may include installing audit software on client system or using the application software for audit as feasible.

1.7 KEY STEPS FOR OBTAINING DATA

- 1. Discuss with client about the requirement of raw data for audit and issue a request letter for getting the requested data in specified form as per the audit objectives.
- 2. Discuss with the IT personnel responsible for maintain data/application software and obtain copies of record layout and definitions of all fields and ensure that you have an overall understanding of the data. The record layout should describe each field and provide information about the starting and ending positions and the data type (numeric, alphanumeric, character, etc.).
- 3. Print sample list of the first 100 records in the data file and compare this to a printout of the obtained data to confirm they are correct.



- 4. Verify data for completeness and accuracy by checking the field types and formats, such as identifying all records with an invalid date in a date field.
- 5. Obtain control totals of all the key data and compare with totals from the raw data to ensure all records have been properly obtained. This can be performed by importing the data in audit software and reviewing the statistics of all the key fields.

1.8 KEY CAPABILITIES OF CAATs

CAATs refer to using computer for auditing data as per audit objectives. This requires understanding of the IT environment and most critically the core applications and the relevant database and database structure. CAATs could be used by using the relevant functionalities available in general audit software, spreadsheet software or the business application software. However, broadly the key capabilities of CAATs could be categorized as follows:

- 1. **File access:** This refers to the capability of reading of different record formats and file structures. These include common formats of data such as database, text formats, excel files. This is generally done using the import/ODBC function.
- 2. **File reorganization:** This refers to the features of indexing, sorting, merging, linking with other identified files. These functions provide auditor with an instant view of the data from different perspectives.
- 3. **Data selection:** This involves using of global filter conditions to select required data based on specified criteria.
- 4. **Statistical functions:** This refers to the features of sampling, stratification and frequency analysis. These functions enable intelligent analysis of data.
- 5. **Arithmetical functions:** This refers to the functions involving use of arithmetic operators. These functions enable performing re-computations and re-performance of results.

Precautions in using CAATs

CAATs have distinct advantages for Auditors and enable them to perform various types of tests. However, it is important to ensure that adequate precautions are in taken in using them. Some of the important precautions to be taken by Auditors are:

- 1. Identify correctly data to be audited
- 2. Collect the relevant and correct data files
- 3. Identify all the important fields that need to be accessed from the system
- 4. State in advance the format the data can be downloaded and define the fields correctly
- 5. Ensure the data represent the audit universe correctly and completely.
- 6. Ensure the data analysis is relevant and complete.
- 7. Perform substantive testing as required.
- 8. Information provided by CAATs could be only indicators of problems as relevant and perform detailed testing as required.





1.9 STEP BY STEP METHODOLOGY FOR USING CAATs

CAATs are very critical tools for Auditors. Hence, it is important to formulate appropriate strategies to ensure their effective use. Some of the key strategies for using CAATs are:

- 1. Identify the scope and objectives of the audit. Based on this, auditor can decided about the need and the extent to which CAAT could be used.
- 2. Identify the critical data which is being audited as per audit scope and objectives.
- 3. Identify the sources of data from the enterprise information system/application software. These could be relating to general ledger, inventory, payroll, sundry debtors, sundry creditors.
- 4. Identify the relevant personnel responsible for the data and information system. These personnel could be from the IT department, vendors, managers, etc.
- 5. Obtain and review documents relating to data/information systems. This should provide information about data types/data structures and data flow of the system.
- 6. Understand the software by having a walk-through right from user creation, grant of user access, configuration settings, data entry, query and reporting features.
- 7. Decide what techniques of CAATs could be used as relevant to the environment by using relevant CAAT software as required.
- 8. Prepare a detailed plan for analyzing the data. This includes all the above steps.
- 9. Perform relevant tests on audit data as required and prepare audit findings which will be used for forming audit report/opinion as required.

1.10 EXAMPLES OF TESTS PERFORMED USING CAATs

CAATs can be used for compliance or substantive tests. As per the audit plan, compliance tests are performed first as per risk assessment and based on the results of the compliance tests; detailed compliance tests could be performed. Some examples of tests which can be performed using CAATs are given below:

- **1. Identify exceptions:** Identify exceptional transactions based on set criteria. For example, cash transactions above Rs. 20,000.
- **2. Analysis of Controls:** Identify whether controls as set have been working as prescribed. For example, transactions are entered as per authorised limits for specified users.
- **3. Identify errors:** Identify data, which is inconsistent or erroneous. For e.g.: account number which is not numeric.
- **4. Statistical sampling:** Perform various types of statistical analysis to identify samples as required.
- 5. **Detect frauds:** Identify potential areas of fraud. For example, transactions entered on weekdays or purchases from vendors who are not approved.
- **6. Verify calculations:** Re-perform various computations in audit software to confirm the results from application software confirm with the audit software. For e.g.: TDS rate applied as per criteria.
- 7. **Existence of records:** Identify fields, which have null values. For example: invoices which do not have vendor name.



- 8. Data completeness: Identify whether all fields have valid data. For example: null values in any key field such as date, invoice number or value or name.
- **9. Data consistency:** Identify data, which are not consistent with the regular format. For example: invoices which are not in the required sequence.
- **10. Duplicate payments:** Establish relationship between two or more tables as required. For example duplicate payment for same invoice.
- **11. Inventory obsolescence:** Sort inventory based on data of purchase or categories as per specified aging criteria or period and identify inventory which has become obsolete.
- **12.** Accounts exceeding authorized limit: Identify data beyond specified limit. For example, transactions entered by user beyond their authorized limit or payment to vendor beyond amount due or overdraft allowed beyond limit.

1.11 ANALYTICAL REVIEW PROCEDURES

The various standards on auditing highlight need for acquiring the required skill-sets to audit in an IT environment and using relevant techniques. Many of the requirements of the auditing standards can be complied by adapting them for use in an IT environment as required. For example: Standard on Auditing (SA) 520Analytical Procedures states:

A1. Analytical procedures include the consideration of comparisons of the entity's financial information with, for example:

- Comparable information for prior periods.
- Anticipated results of the entity, such as budgets or forecasts, or expectations of the auditor, such as an estimation of depreciation.
- Similar industry information, such as a comparison of the entity's ratio of sales to accounts receivable with industry averages or with other entities of comparable size in the same industry.

A2. Analytical procedures also include consideration of relationships, for example: Among elements of financial information that would be expected to conform to a predictable pattern based on the entity's experience, such as gross margin percentages.

Most of the analytical procedures can be performed in an IT environment using CAATs which makes the audit process much more effective and efficient.

SUMMARY

CAATs enable auditors to use computers as a tool to audit electronic data. CAATs provide auditors access to data in the medium in which it is stored, eliminating the boundaries of how the data can be audited. As auditors start using CAATs, they will be in a better position to have a considerable impact on their audit and auditee as more time is spent on analysis and less time on routine verification. It is important to understand the client IT environment and chart out which techniques of CAAT could be used. Initially, time needs to be invested in this Endeavour but once the audit plan is prepared based on the IT environment as per audit scope, re-use becomes easier. However, the audit plan and tests need to be updated based on changes in the IT environment as relevant. Using CAATs provides greater assurance of audit process to the auditor and also to the auditee. The key to using CAAT is recognizing the need, learning how to use CAATs and using them in practical situations.



DATA ANALYSIS AND AUDIT TECHNIQUES

ICAI has issued guidelines on CAAT and the various assurance standards highlight the importance of using CAATs as relevant for audit. ISACA has also issued standards and guidelines on auditing and CAATs. Some of the key aspects from these standards and guidelines are given below.

2.1 NEED FOR USING CAATs

A P

T E R

As entities increase the use of information systems to record, transact and process data, the need for the auditors to utilize tools to adequately assess risk becomes an integral part of audit coverage. The use of computer-assisted audit techniques (CAATs) serves as an important tool for the auditor to evaluate the control environment in an efficient and effective manner. The use of CAATs can lead to increased audit coverage, more thorough and consistent analysis of data, and reduction in risk.

CAATs include many types of tools and techniques, such as generalized audit software, customized queries or scripts, utility software, software tracing and mapping, and audit expert systems.

CAATs may be used in performing various audit procedures including:

- Tests of details of transactions and balances
- Analytical review procedures
- Compliance tests of general controls
- Compliance tests of application controls

CAATs may produce a large proportion of the audit evidence developed on audits and, as a result, the auditor should carefully plan for and exhibit due professional care in the use of CAATs.

2.2 KEY FACTORS TO BE CONSIDERED IN USING CAATs

When planning the audit, the IS auditor should consider an appropriate combination of manual techniques and CAATs. In determining whether to use CAATs, the factors to be considered include:

- Computer knowledge, expertise, and experience of the IS auditor
- Availability of suitable CAATs and IS facilities
- Efficiency and effectiveness of using CAATs over manual techniques
- Time constraints
- Integrity of the information system and IT environment
- Level of audit risk

2.3 CAATs PLANNING STEPS

The major steps to be undertaken by the auditor in preparing for the application of the selected CAATs include the following:

- Set the audit objectives of the CAATs, which may be included in the terms of reference for the exercise.
- Determine the accessibility and availability of the organization's IS facilities, programs/systems and data.
- Clearly understand composition of data to be processed including quantity, type, format and layout.
- Define the procedures to be undertaken (e.g., statistical sampling, recalculation, confirmation).
- Define output requirements.
- Determine resource requirements, i.e., personnel, CAATs, processing environment (the organization's IS facilities or audit IS facilities).
- Obtain access to the organization's IS facilities, programs/systems and data, including file definitions.
- Document CAATs to be used, including objectives, high-level flowcharts and run instructions.

2.4 AUDIT EVIDENCE AND CAATs

Audit is primarily said to be the process of collecting and evaluating audit evidence as per audit objectives. Based on the scope and objectives of audit, auditor can obtain the audit evidence by:

- Inspection
- Observation
- Inquiry and confirmation
- Re-performance
- Recalculation
- Computation
- Analytical procedures
- Other generally accepted methods

2.5 CAATs DOCUMENTATION

Work papers

The step-by-step CAATs process should be sufficiently documented to provide adequate audit evidence. Specifically, the audit work papers should contain sufficient documentation to describe the CAATs application, including the details set out in the following sections.

Planning

Documentation should include:

CAATs objectives





- CAATs to be used
- Controls to be exercised
- Staffing and timing

Execution

- Documentation should include:
- CAATs preparation and testing procedures and controls
- Details of the tests performed by the CAATs
- Details of inputs (e.g., data used, file layouts), testing periods, processing (e.g., CAATs highlevel flowcharts, logic) and outputs (e.g., log files, reports)
- Listing of relevant parameters or source code

Audit Evidence

Documentation should include:

- Output produced
- Description of the audit analysis work performed on the output
- Audit findings
- Audit conclusions
- Audit recommendations

In audits where CAAT is used, it is advisable that the audit report includes a clear description of the CAATs used in the objectives, scope and methodology section. The description of CAATs used should also be included in the body of the report, where the specific finding relating to the use of CAATs is discussed. This description should not be overly detailed, but it should provide a good overview for the reader.

2.6 AUDIT TEST USING CAATs

If the data to be audited is available in electronic form, then CAATs could be used for:

- Inquiry and confirmation identifying accounts for which external confirmation is to be obtained. Request letters for confirmation of balances can be printed using CAAT software.
- Re-performance: The processing of transactions done by the application software can be reperformed and the resultant data can be compared to verify correctness and completeness. For example: Postings of transactions to personal ledger can be re-performed using the original transaction data base and compared with classified transactions as per ledgers.
- Re-calculation: All the computations which were done electronically by the application software used in the enterprise can be independently validated by re-performing the computations. For example, Tax deducted at source or VAT charged on sales, interest computation, etc. can be re-computed in CAAT software and validated with the computed totals from the original application software to confirm correctness of processing of transactions.



- Computation: using CAAT software, it is possible to compute totals to confirm correctness. For example, the VAT payments made for the year can be total in CAAT software to compare with the total payments as per VAT returns. The interest debited can be computed and compared with actual debit to interest a/c for the year.
- Analytical procedures: Based on the data available in electronic format, various analytical procedures can be performed by comparing and relating various aspects of financial and on-financial information.

2.7 AUDIT SAMPLING

Auditor has to design and select an audit sample and evaluate sample results. Appropriate sampling and evaluation will meet the requirements of 'sufficient, reliable, relevant and useful evidence' and 'supported by appropriate analysis. Auditor should consider selection techniques that result in a statistically based representative sample for performing compliance or substantive testing.

When using either statistical or non-statistical sampling methods, auditor should design and select an audit sample, perform audit procedures, and evaluate sample results to obtain sufficient, reliable, relevant and useful audit evidence. Audit sampling is defined as the application of audit procedures to less than 100 percent of the population to enable the IS auditor to evaluate audit evidence about some characteristic of the items selected to form or assist in forming a conclusion concerning the population.

Statistical sampling involves the use of techniques from which mathematically constructed conclusions regarding the population can be drawn. Non-statistical sampling is not statistically based, and results should not be extrapolated over the population as the sample is unlikely to be representative of the population.

Design of the Sample

When designing the size and structure of an audit sample, IS auditors should consider the specific audit objectives, the nature of the population, and the sampling and selection methods.

Auditor should consider the need to involve appropriate specialists in the design and analysis of samples.

Selection of the Sample

There are four commonly used sampling methods. Statistical sampling methods are:

- **Random sampling** Ensures that all combinations of sampling units in the population have an equal chance of selection
- **Systematic sampling** Involves selecting sampling units using a fixed interval between selections, the first interval having a random start. Examples include Monetary Unit Sampling or Value Weighted selection where each individual monetary value (e.g., Rs. 1) in the population is given an equal chance of selection. As the individual monetary unit cannot ordinarily be examined separately, the item which includes that monetary unit is selected for examination. This method systematically weights the selection in favour of the larger amounts but still gives every monetary value an equal opportunity for selection. Another example includes selecting every 'nth sampling unit





Non-statistical sampling methods are:

- **Haphazard sampling** The IS auditor selects the sample without following a structured technique, while avoiding any conscious bias or predictability. However, analysis of a haphazard sample should not be relied upon to form a conclusion on the population
- **Judgmental sampling** The IS auditor places a bias on the sample (e.g., all sampling units over a certain value, all for a specific type of exception, all negatives, all new users). It should be noted that a judgmental sample is not statistically based and results should not be extrapolated over the population as the sample is unlikely to be representative of the population.

Auditor should select sample items in such a way that the sample is expected to be representative of the population regarding the characteristics being tested, i.e., using statistical sampling methods. To maintain audit independence, the IS auditor should ensure that the population is complete and control the selection of the sample. For a sample to be representative of the population, all sampling units in the population should have an equal or known probability of being selected, i.e., statistical sampling methods.

DATA ANALYSIS USING IDEA

LEARNING OBJECTIVES

- To gain understanding on Importing different file formats into IDEA.
- To understand on how to generate field statistics for the database.
- To understand formatting Data.

There are many audit software available in the market. However, for the purpose of learning CAATs, we will be using IDEA software in this chapter and also for performing exercises in the lab. Students may refer to the ICAI publication titled: "Practical application of CAAT-case studies" for more examples and details of CAATs using IDEA Software.

3.1 IMPORTING DATA

Function Description

Important Assistant brings the selected file or files into IDEA database management system. Userfriendly Import assistant guides user through a series of steps and instructions for importing the file into the Software. All the functionalities of IDEA can be performed only when the file is available within IDEA. Hence, the first step in data analysis is ensuring that the files to be audited are in selected format acceptable in IDEA and are imported into IDEA.

Let us assume you have an Excel or Access file and you want to perform certain idea functionalities on it, then it is important to import these files into IDEA.IDEA facilitates user to import external files in different formats like Access, Excel, dbase or other ODBC/DSN formats into IDEA database.

We are explaining below step by step how to import a file into IDEA. If you have access to IDEA Server, you can also import files to IDEA Server. For now, you will have to import a file to your Working Folder. A Working Folder is simply a folder that contains the IDEA databases that you wish to analyze. It is recommended that all the files related to each audit or investigation be stored in a separate folder or directory to simplify file management and housekeeping.

Note: The sample files which are used for explaining how to perform various functionalities of idea are available in the C:\Program Files\IDEA\UserFiles\Tutorial\ when IDEA is installed with default settings. You can perform exercises using the relevant sample data files. Please check with your lab instructor about the location of these files. We will be using tutorial/sample files for explaining different functionalities of IDEA in this booklet.



You can import files of different format into IDEA. However, for this exercise, we are giving below an example for importing an Access file into IDEA.

Step by Step Procedure for Importing Data into IDEA

Location

- File>Import Assistant >Import to IDEA
- Alternatively, on the Operations toolbar, click the Import to IDEA button. Import Assistant dialog box appears as shown in Fig 3.1.1.



Fig. 3.1.1 Import Assistant Location

Step 1: Select the Format

elect the format:	
Advanced Record I AS400 dBASE Lotus Microsoft Access Microsoft Excel ODBC Print Report and Ad SAP/AIS Text XML	befinition Editor
Computer location:	My Computer
Computer location: File name:	My Computer
Computer location: File name: Optional file:	My Computer Step 10 EA/UserFiles/Tutorist/Customen/MDB
Computer jocation: File name: Optional file:	My Computer My Computer My Computer GutProgram Files/IDEA/UserFiles/Tutorist/Gustomer/MDB GutProgram GutProgr

Fig. 3.1.2 Import Assistant Step 2

- (a) In the Import Assistant Dialog Box, select Microsoft Access from the list.
- (b) Click the **Browse** button next to the File namebox to select the Microsoft Access database you want to Import.
- (c) Navigate to and select C:\Program Files\IDEA\UserFiles\Tutorial\Customer.MDB.
- (d) Click **Open**.
- (e) The **Select File** dialog box closes and the selected file name and path appear in the File namebox in the **Import Assistant** dialog box.
- (f) Click Next.

The Microsoft Access dialog box appears.

Step 2: Select Tables as shown in Fig 3.1.3

File Explored		ACCEUL-Outstan	3.40 @ Customer	Catations 340 @ Langled -Weblog 340 @ Lales Tra	martian SalesTrans.IMP @ Summarized Transaction.IMD # ×	Properties
		MyOCIN	0 NO_OF_RECS	ALTS_SUM		I + Detabase
	_	1.	8	167500	3	100
a real of the second seco	- Print	2 100	1	20500		a con
	Pactores	3 1902	1	20060		a recory
ACCESS-Database	1,000	4 1000	1	19500		 Freid Statistics
Cultomer-Catabase	1,000	8 1004	1	29000		Control Total
Sales Tenenantine, Sales Tenen		# 1905	1	18500		E Cifela
Committee Transmittee		1006	1	18000		1 + Reads
B Sample Authorization	15	8 1014	1	1,000		
B Semale Weblas	200	8 1415	1	1,000		Industry
Target History	201	10 1010	1	11000		
		40 1000		1000		🖌 No index
		12 1018	1	10000		
		14 1020	1	3400		• Comments
		18 1025	1	1100		il Ald comment
	- 3	16 1027	1	3000		
		17 1023	1	3000		
		18 1024	1	1300		
		10 1025	1	1300		
		30 1826	1	1200		
		25 1627	1	1200		
	-	22 1428	1	1200		
	-	23 1629	1	1200		
	-	24 1000	1	1000		
	- 1	28 101	1	1000		
	- 1	28 1012	1	790		
	- 1	10 100	1	730		
		The Local	2	1000		
		M HART		1700		
		10 100		A FRANCE		

Fig. 3.1.3 Import Assistant Step 3

(g) In the **Select tables** box, select **Database1**.

Note: If you import a Microsoft Access file that contains more than one table, you may simultaneously import multiple tables by selecting the associated check boxes. However, any options you select will apply to all imported tables.

- (h) All Character fields will be imported with a length of 255 characters unless changed. This is not likely to be the underlying Character field length. Therefore, leave the Scan records for field length check box selected. Also, accept the default value in the Scan only box in order to scan 10,000 records to determine the maximum field length.
- (g) Accept the default output file name (**Customer**), and then click **OK**. When the file is imported, the database name becomes filename-tablename. In this case, the file you imported becomes an IDEA database called **Customer-Database**.





Step 3: Result - Customer-Database

On clicking Ok, the Customer Database is imported into IDEA. The imported database is opened in the Database Window. In the File Explorer Window the imported database is highlighted as shown in Fig 3.1.4

Bie Edit Liew Data ;	Analysis Sampling	i Io	ok <u>H</u> indow	Beb						
343-3-2	6 2 3 3		100	2 () 000 (iii			182	J J D R D I O		
r lightere	*	1	Customer	Database1.IND						
122185	115		CUST_NO	FRST_NWE	LAST_NAME	COUNTRY	STATUS	CREDIT_LIM		
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		1	10000	DOWNE	BURROWS	SOUTH AFRICA	М	4000		
		2	1003	(EVIN	NICHOLSON-KNOWLES	SOUTH AFRICA	M	29006		
ine .	Records	3	10004	JONER	DE FRETAS	SOUTH AFRICA	M	1,3000		
ACCESS-Database	1,000	4	1005	CHABRAI	SAWIER	SOUTH AFRICA	F	19000		
Customer-Databes	341	5	10006	GATHARINE	BURROWS	SOUTH AFRICA	F	13000		
Sales Transaction	999	4	1007	DONGSAN	BIS	NIGERIA	M	8000		
Sales-Sheet1	292	7	10101	MALINDA	JOHNSTON	NECENCIA.	M	3000		
Sample-Authorizet.	15	8	1002	TANIA	PLOUFFE	NEGEREA	M	2000		
		1	10301	NATESHA	MCDONALD	NOCERCIA	£	290300		
		10	10205	KENNETH	HIGGS	NICERUA	M	1900		
		11	10204	RORN	600538N5	BELGUM	M	7000		
		12	1892	BENOT LUMMERANT BELGUM	м	5000				
	- 1		13	10400	ANNEX	VANDERVUST	BLGUM	8	1800	
		14	10500	PAUL	RUMAND	BEIGUM	M	3000		
		15	10801	CREITIAN	SIN	BELGUM	M	23000		
		18	10900	REGNE	RULGEANU	BRIGUM	F	27000		
		17	111100	NACIA	KALAMBER	BRIGHM	8	2000		
		18	11207	ALAN	BUICK	BELGUM	M	5000		
		15	11300	OHISTIANE	DUDUEDIE	BECKIM	5	2500		
		20	11301	MEALANGA	VLADESCU	BELGUM	M	5000		
		21	11400	DENSE	IONN	BARBADOS	4	2000		
		22	11600	SAMLE	GONGALVES	84784005	1	1500		
		20	11702	STERN	ARCHER	BARBADOS	1	2000		
		24	1174	CECL	CLARKE	84884005	1	2600		
		25	11805	DEAN	SPENCER	BARBADOS	1	18000		
		28	11305	LINCA	GREDNE SMALL	BARBADOS		34000		
		27	11809	MARNELA	HRISTON	BULGARIA	6	2000		
		2	11810	LUCME.	TOMOV	BULGARIA	M	6000		
		20	12368	DEMOLEN	ROINOV	BLL GARM	1	14000		
		15	12305	YERCEAL	STANTOPPIA	RAGARA		24000		
		31	12306	CLUMBER	KANTCHEV	RIE GARDE	M	1300		

Fig. 3.1.4. Customer Database is imported into IDEA

Exercises:

- 1. Please perform the following exercises in your lab.
- 2. Please import excel file following the above steps.
- 3. Please import text delimiter file and notice the different steps.
- 4. Please import excel file using ODBC option.

3.2 IMPORTING DATA

Click on field statistics after importing the files and understand the nature of data which is imported.

Function Description

You can use the Export Database task to create an external file from an IDEA database so that you can use the data in other applications, such as a spreadsheet package. IDEA exports data in a number of text, database, spreadsheet, and mail merge formats.

You can use the Export Database task to create an external file from an IDEA database so that you can use the data in other applications, such as a spreadsheet package. You can also use Copy and Paste to incorporate portions of text or sections of database into other Windows applications. IDEA also supports drag and drop into any other OLE2 container application, such as Microsoft Excel. IDEA exports data in a number of text, database, spreadsheet, and mail merge formats.

Step by step process for exporting files from IDEA.

Location

• File>Export Database



Fig. 3.2.1 Export Database Location

Database Used

• Customer-Database1





Export Dialog Box

oort Database			
Records to Select			OK
() Al	Starting record number:	1	
Bange	Enging record number:	341	Eields
	5.50		Cancel
Export Details			
Database order:	No index	•	Help
<u>C</u> riteria:			
Export Type			
Export as:	Text Delimited	•	
Field separator:	, Decim	al separator:	
	Include	field names: 🛄	
File name:	C:\Users\Saranya\Docur	ments\IDEA\Samples	

Fig. 3.2.2. Export Database

- (a) In Records to Select, on selecting **All** will select the entire records. On selecting Range allows the user to select the **Starting** and **Ending record number**.
- (b) By default the path is set to C:\Users\Saranya\Documents\IDEA\Samples\Customer-Database1.
- (c) In Export Type, user can select the format in which the current file has to be exported and also name the resultant exported file.
- (d) In **Filename** allows the user to select the path to which the file has to be exported.
- (e) On clicking **Fields**, we can select or unselect the fields that have to be exported.
- (f) During Exporting process, condition or criteria can be applied using the criteria button. On clicking this button, Equation Editor Dialog box is opened which facilitates the user to write in query or condition.
- (g) Clicking Ok, exports the active file into given desired format.

Exercise:

1. Export actual idea file to different formats such as: HTML, Excel, text delimited file and check the results.

3.3 SUMMARIZATION

Learning Objectives:

To total the sales transaction by INVOICENO to produce a list of outstanding Sales as well as to identify the number of INVOICENO and the Sales per INVOICENO.

Function Description

Summarization accumulates the values of Numeric fields for each unique key. For example,

summarizing an Accounts Payable database by account number (the key) and totalling invoice amounts produces a database or **Results** output of outstanding liabilities by supplier. The Summarization task provides:

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

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

You 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 you can drill down to display the individual records for the key. You can select additional field data to be selected 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 4,000 rows. If you expect your result will have more than 4,000 rows, you must choose to create a database.

Step by Step Procedure for summarization

Ensure that **Sales Transactions** is the active database and the **Data** property is selected in the **Properties** window.

Location

File Edit View Data	Analysis	Sampling	To	ols Wi	ndow	Help	
	₿Σ Sum	nmarization				EB C	
File Explorer	Pive	t Table tification				action-Sa	lesTrans.IMD
	Due	licate Key			- 6	DATE	INVOICENO
	= ~	D. t. t'			- 11	1-Apr-09	1001
IDEA Files	I Gap	Detection	8			1-Apr-09	1002
Name	Agir	ng				1-Apr-09	1003
ACCESS-Database	Ben Ben	ford's Law				1-Apr-09	1004
- Ustomer-Database1	Adv	anced Statis	tical N	Aethods		1-Apr-09	1005
Sales Transaction-Sale			10	KA		1-Apr-09	1006
- Sales-Sheet1		192	7	RA	0	1-Apr-09	1007
Sample-Authorization		15	8	RA	0	1-Apr-09	1008
			9	RA	0	1-Apr-09	1009
			10	RA	0	1-Apr-09	1010
			11	RA	0	1-Apr-09	1011
			12	RA	0	1-Apr-09	1012

• Analysis > Summarization

Fig. 3.3.1 Summarization Location

- (a) In the Summarization Dialog Box, select the following:
 - Fields to summarize: INVOICENO



- Numeric fields to total: SALES
- (b) Click **Fields**. The **Fields** dialog box appears. Note that no fields are selected. This stops unnecessary information from being included in the summarized database.
- (c) Click **OK** to return to the **Summarization** dialog box.

Select the **Use Quick Summarization** check box. The Use Quick Summarization check box may be selected as a faster means to summarize your database. However, Quick Summarization may only be used if the database has no more than 4,000 unique keys. In addition, Quick Summarization allows you to select only one field to summarize.

There are two types of output from the **Summarization** task:

- Summarization database
- Summarization result

Note that as with most tasks in IDEA, you may apply a criterion to the task, for example; only summarize transactions for a specified period. As with all other tasks where you can apply a criterion, if you apply the criterion to the database using the **Criteria** link in the Properties window, the criterion equation appears in the **Criteria** text box on the task dialog box. However, you may enter a new criterion or modify an existing one using the Equation Editor as shown in Fig 3.3.2.

Fields to su	immarige:		Nymeric fields to total:		
By:	INVOICEND	-	SALES		OK
Then by:	NONE	-			Fields
Then by:		*			C C
Then by:		+			Cancel
Then by:	16L	-			Help
Then by:		-			
Then by:		*			
Then by:	0	+			
<u>C</u> riteria:	(f)		-		
Use Q	uick Summarization	2002	Statistics to include:		
Create	gatabase		🕼 Sum	Average	
E Inc	lude % in output database		Magimum	Vgriance	
@ Us	e fields from first occurrence		Minimum	Standard deviation	
O Us	e fields from Jast occurrence		Create jesuit		
File on	mer Summarized Transaction		Beauti name: Sum	matization	



- (d) In the **Statistics to include** area, accept the default selections of **Sum**.
- (e) Click OK.

View the resultant database and note the following fields:

- INVOICENO- List of unique INVOICENO
- NO_OF_RECS Number of records for each INVOICENO
- Sales_SUM- Total Sales of the transactions for each INVOICENO

Also note the number of records (989) on the **Status** bar displayed in the below figure. Note that IDEA automatically creates an Action Field link to the parent database (**Sales Transactions**). It allows you to display the records from the **Sales Transactions** database by clicking on a value (in blue) in the **NO_OF_RECS** field as shown in Fig 3.3.3.

D - 2 - 2 - D - 0	S. S. 15	a 10 -011	15 (13 and 117) a		Pr 17 15 15 15 T @ I	
De To Bar Do en en i	04 04 110	P Hat Service	8 C m [11+] .	a can real full range real size ran in a lite	CONTRACT BALLEN ON CONTRACT	
the Explored		ACCESS-Outables	e.340 🕼 Customer Ox	skaue.340 🐨 SempleS-Weblog.340 🐨 Seles Trans	ction SalesTrans.MC* @ Summarized Transaction.IMD **	Properties
422PK		MVOICEN	O NO_OF_RECS SA	S_SUM		Teleboie
Ma Port	_	1		167500		J Deta
lana -	Records	2 1000	1	20500		M History
CONTRACTOR Detailers	1.000	3 1002	1	20080		IR Field Determine
- Customer-Database	1,000	4 1000		29500		I Control Land
PLOWS B2000 HTMLDR ACT.	0	18 1004 ·	A)	29000		an Coloria
Sales Transaction-Sales Trans	9990	# 1000		1000		E Conta
Summarized Transaction	989	2 10m		2,000		 Results
a Bample-Authorization	15			1 2000		
Sample-Weblog	290	10 1015	1	13000		= w Indices
Samplel-Meblog	2015	11 1017	1	113900		
		12 1018	1	10000		 No index
		12 1019	1	10000		T # Company
		14 1829	1	3800		Carnet
		18 1025	1	38.00		ill Add comment
		16 1822	1	3080		
		17 1023	1	3000		
		18 1024	1	1300		
		18 1025	1	1300		
		30 3826	1	1200		
		26 1627	1	1200		
		22 1628	1	1200		
		23 1629		1200		
		24 3830	*	1000		
		28 10.0		1000		
		20 1000		100		
	1	20 2024		130		
		Distantion.	1	1200		
		30 1007	-	27900		
· · · · · · · · · · · · · · · · · · ·		20 1 10100	1	2020		

Fig. 3.3.3 Summarization Result

Exercises:

- 1. Use the sample employees file in the sample database of idea and perform summarization on branch and total on salary. Click on the hyperlink of the number of records and check the result.
- 2. Use the sample employees file in sample database of idea and perform summarization on country and total on salary and check the result.

3.4 STATISTICS

Learning Objectives

Understanding to view the field statistics for the Numeric fields in the active database.

The statistics is used for reconciling totals, obtaining a general understanding of the ranges of values in the database, and highlighting potential errors and the area of weakness to focus subsequent tasks.

Function Description

The Field Statistics property provides statistical information about all Numeric, Date, and Time fields within the active database. The field statistics are available and displayed for all records in the database, with any applied criteria ignored.

By default, the Field Statistics window displays the statistics for Numeric fields. Ensure that **Customer-Database** is the active database and the **Data** property is selected in the Properties window.

Step by Step Procedure for Statistics

Location

• In the **Properties** window, click **Field Statistics**.





Properties +P
II 🕶 Database
🖉 Data
History
Field Statistics
Control Total
Criteria
II ▼ Results
II ▼ Indices
No index
CUST_NO/A
II 🔻 Comments
Add comment

Fig. 3.4.1 Statistics Location

- (a) On clicking Field Statistics in property window, IDEA displays below message box.
- (b) Click **Yes** to generate field statistics for all fields.

IDEA	-	-	-			×
•	Statistics are not avai for all fields without s	lable for a statistics?	ll fields. D	o you wish to creat	e statistics	
				Yes	No	

Fig. 3.4.2 Statistics

- (c) By Default By default, the Field Statistics window displays the statistics for Numeric fields. In this case, field statistics appear for all the available numeric fields. To view field statistics for the Date fields and Time fields in a database, click Date and Time in the Field Type area of the Field Statistics window. In the current database, there are no Date or Time fields.
- (d) If the database contained more than one Date, Numeric, or Time field, multiple date, numeric, or

Time field statistics would appear together on the same screen for easy comparison of values.



Statistics for Numeric Fields

Customer-Dat	dML easts						~
0 0 0 -01	8 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	A CONTRACTOR OF THE OWNER		·			
Field Type	Numeric Statistics	COUNTRYNO	UNIT_PRICE	QTY.	PROD_CODE	AMOUNT	PAYMENT_AMT
Peterseen	Fiet Value	8.07%	8.084.43	90.074	9.627.416	710.356.44	710.356.44
Type	Abeolute Value	8.075	8.084.43	90.074	9.627.416	710.356.44	710.356.44
Numeric Fields	If of Records	1.000	1.000	1,000	1.000	1.000	1.000
COUNTRYND	# of Zero Kerns	0	0	-49	0	49	49
2 UNIT_PRICE	Positive Value	8.075	0.084.43	90.074	9.627.416	710.356.44	710,356.44
PROD_CODE	Negative Value	0	0.00	0	0	0.00	0.00
2 AMOUNT	If of Positive Records	1000	1000	951	1000	951	951
Contraction of the second second	If of Negative Records	0	0	0	0	0	0
	# of Data Errors	0	0	0	0	0	0
	# of Valid Values	1,000	1,000	1,000	1,000	1,000	1,000
	Average Value	8.08	8.08	90.07	9,627.42	710.36	710.36
	Minimum Value	3	3.84	0	9,598	0.00	0.00
	Maxpreum Value	10	9.99	3,496	9.951	27,793.20	27,793.20
	Record # of Min	1	640	6	1	6	6
	Record # of Max	968	191	26	191	26	
	Sample Std Dev	2.14	0.66	227.30	96.49	1.758.47	1.758.47
	Sample Variance	4.60	0.43	51,663.19	9.310.30	3.092.214.89	3.092.214.89
	Pop Std Dev	2.14	0.66	227.18	96,44	1.757.59	1.757.59
	Pop Variance	4.60	0.43	51.611.53	9.300.99	3.089.122.68	3.089.122.68
	Pop Skewness	-1.882584	0.056460	6.793836	3.016606	6.973450	6.973450
	Pop Kurtosis	1.645513	14,736605	68.926272	7.176397	74.416881	74.416881

Fig. 3.4.3 Numeric Statistics

As shown in the above image, active database contains six numeric fields and the statistics for all six numeric fields are displayed in Field Statistics window which allows easy comparison between them.

Study the field statistics for the all Numeric fields. Note in particular:

		UNIT_	QTY	PROD_	AMOUNT	PAYMENT_
		PRICE		CODE		AMT
# of Records	1,000	1,000	1,000	1,000	1,000	1,000
# of Zero Items	0	0	49	0	49	49
Average Value	8.08	8.08	90.07	9627.42	710.36	710.36
Minimum Value	3	3.84	0	9598	0.00	0.00
Maximum Value	10	9.99	3496	9951	27793.20	27793.20

Below two Images displays statistics for Date and Time fields for current active database.

Statistics for Date Fields

Customer-Database.	MD			
103 49 109 - 407 100 Her	CT -10 -40 -020 -			
Field Type	Date Statistics	ORDER_DATE_DATE	PAYMENT_DAT_DATE	SHIPMENT_DA_DATE
héumenic .	 At of Valid Values 	1,0001	19/8	199
Torre	Af of Zero Keme	0	801	801
Date Fields	Af of Plecords	1,000	1,000	1.000
EXCLUSION INCOMENCIES	M of Data Errors	0	0	0
2 SANPARENT DAT DATE	Earland Cate	13-1404-93	11-1404-93	13.860.03
all an out highly a manufacture of	Latest Date	23-Nov-01	21-Herv-01	23-Mov-01
	Record # of Earliest	610	610	610
	Flacord # of Latest	150	150	150
	Most Common Day	Turneday	Monitary	Wednesday
	Most Common Month	2.04	August	August
	Reserve in Jakrousky	27	19	22
	Barris In Fabricary	17	13	34
	Barros on March	31		11
	Barros in April	23	6	
	Berris in May	36	10	9
	Berns in June	60	10	14
	Barris In July	199	20	210
	Reme to August	100	31	31
	Barris in September	161	38.	. 16
	Rema as October	105	22	24
	Barros an Pécreenitrer	67	10	10
	Barris In December	23	10	17
	Rema on Sunday	10	27	6
	Barrow oro Montelland	14	44	

Fig. 3.4.4 Date Statistics





Statistics for Time Fields

Customer-Database.IMI	ס			
0 8 8 - 2 B M C] �			
Field Type	Time Statistics	ORDER_DATE_TIME	PAYMENT_DAT_TIME	SHIPMENT_DA_TIME
Numeric	# of Records	1,000	1.000	1,000
Time	# of Valid Values	1.000	1.000	1.000
Time Fields	# of Data Errors	0	0	0
ORDER_DATE_TIME PAYMENT_DAT_TIME SHIPMENT_DA_TIME	# of Zero Items	0	801	801
	Latest Time	11:55 PM	11:59 PM	11:59 PM
	Earliest Time	12:01 AM	12:00 AM	12:00 AM
	Most Common Hour	11	. 0	0
	Most Common Minute	12	0	0
	Most Common Second	0	0	0
	# of Records in AM	788	801	801
	# of Records in PM	212	199	199
	# of Records Before 6 AM	373	801	801
	# of Records After 6 PM	104	95	100
	Average Time	8:39 AM	3:31 AM	3:35 AM
	# of Records Less Than	1000	1000	1000
	# of Records More Than	0	0	0

Fig. 3.4.5 Time Statistics

Exercise:

- Click on field statistics for sample employees and understand profile of data.
- Click on field statistics for sample-inventory and understand profile of data.

3.5 SAMPLING

Learning Objectives

- To draw a number of records with fixed interval for testing.
- To select a random sample of records for testing.
- To extracts a random sample with a specified number of records from each of a series of bands.

Function Description

Sampling in IDEA is broadly statistical and probability-based.

The probability-based sampling techniques are: Systematic, Random and Stratified Random

The statistical sampling techniques are: Attribute, Classical Variable and Monetary Unit sampling

We will be covering examples of systematic and random sampling only in this training. Students may try other forms of sampling in the lab exercise.

Systematic Record Sampling

Systematic Record Samplings 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.

Step by Step Procedure for performing Systematic Sampling

Ensure that **Sales Transaction-Sales Trans** is the active database and the **Data** property is selected in the **Properties** window.

Location

- Sampling >Systematic.
- Alternatively, click the **Random Record Sampling** button on the **Operations** toolbar. The **Random Record Sampling** dialog box appears as shown in Fig 3.5.1



Fig. 3.5.1 Systematic Sampling





Step 1: Number of Records

Systematic Sample			×
Number of Records	Selection Inter	rval	
Sel Number of <u>r</u> ec	ection interval: cords to select:	10	<u>Fi</u> elds Compute
Starting record nu	mber to select:	1	Help
Ending record nu	mber to select:	1000	Пор
File <u>n</u> ame: System	matic Sales		
		ОК	Cancel

Fig. 3.5.2 Number of Records

- (a) In the Number of records to select box, enter 100.
- (b) Accept the random number seed provided by IDEA. IDEA uses the random number seed to start the algorithm for calculating the random numbers. If a sample needs to be extended, then entering the same seed but with a larger sample size produces the same original selection plus the required additional records.
- (c) Accept the defaults in the **Starting record number to select** and the **Ending record number to**
 - (a) Select boxes.
 - (b) IDEA sets the defaults as the first and last records; in this case 1 and 1000.
- (d) In the **File name** box, enter **Systematic Sales**.

Step 2: Selection Interval

Systematic Sample		×
Number of Records Selection Inter	val	
Number of records to select:	100	Fields
Selection interval:	10	Compute
Starting record number to select:	1	Hala
Ending record number to select:	1000	
File name: Systematic Sales		
	ОК	Cancel

Fig. 3.5.3 Selection Interval



- (a) In **Selection Interval** tab page, enter 10 for selection interval. This means it picks every 10th records from 1 to 1000 records. Total records picked are 100.
- (b) Click Ok.

Step 3: Result of Systematic Sampling



Fig. 3.5.4 Systematic Result

(a) As displayed in the result, every 10th record is picked from the active database and total numbers of records are 100.

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.

Step by Step Procedure for performing Random Sampling

Ensure that **Sales Transaction-Sales Trans** is the active database and the **Data** property is selected in the **Properties** window.





Location

- Sampling >Random.
- Alternatively, click the **Random Record Sampling** button on the **Operations** toolbar. The **Random Record Sampling** dialog box appears as shown in Fig *3.5.5*

IDEA - Sales Transaction-SalesTrans.IM	MD	Statement of the local division of the local		
File Edit View Data Analysis	San	npling Tools Window Help Systematic Random Stratified Random Sample Monetary Unit	•	
IDEA Files		Classical	•	
Name Record		Attribute Planning and Evaluation		
ACCESS-Database 1,0		Generate Random Numbers		
	20	5 KA 01-Apr-09 10	005	
Bales Transaction 95	99	6 RA 01-Apr-09 10	006	1
Sales-Sheet1 19	92	7 RA 01-Apr-09 10	007	
Sample-Authoriz 1	15	8 RA 01-Apr-09 10	800	
		9 RA 01-Apr-09 10	009	

Fig. 3.5.5 Random Sampling Location

Number of records to select:	10		OK
Random number seed:	2120	3	<u>Fields</u>
Starting record number to select:	1	A. V	Cancel
Ending record number to select:	999	A. V	Help
Allow duplicate records:			

Fig. 3.5.6 Random Sampling

- (a) In the **Number of records to select** box, enter **10**.
- (b) Accept the random number seed provided by IDEA. IDEA uses the random number seed to start the algorithm for calculating the random numbers. If a sample needs to be extended, then entering the same seed but with a larger sample size produces the same original selection plus the required additional records.
- (c) Accept the defaults in the **Starting record number to select** and the **Ending record number to Select** boxes.

IDEA sets the defaults as the first and last records; in this case 1 and 999.

- (d) Leave the Allow duplicate records check box unselected.
- (e) In the **File name** box, enter **Random Sales**.
- (f) On clicking **Ok** total **10 records** are extracted from active database.



Random Sampling Result

and the second se		all come of the							Married Works	-
-1 50 m st -3 12 -	_	Ti V Systematic Custome	ers.3HD @ Sales Trans	action-Sales1	sans 3MP 1	W Random Sales JMD		•*	- Contraction	
「日のも見る」	R	DATE INVOICEN	IO NAME	SALIS	USCRO	SAR RECNO			Catabase	
DEA Film		1 00-Apr-00 1111	Page and	75500	1	101			J Data	
Name	Records	3 25-441-09 1627	satuka	1150	54	627			# Hiday	
ACCESS-Dutabase	1,000	4 07-554-09 1825	richard	23650	MM	525			Field Statistics	
- @ Customer-Detabase	1,900	8 01-May-09 1292	mathy	1010	54	297			Control Tatul	
Bratification	1,000	8 04-May-09 1858	Destroy -	3945	54	868			E Criteria	
Systematic Customers	100	7 04-844-09 1118	diversi	4150	54	118				
FLOWS, 520300 HTMLD		# 07-Apr-09 1187	taket	1130	SA	387			1 . fesufts	
Sales Transaction-Sales	999	# 23-Apr-09 1563	ridenta	1,3000	RA	363				
Random Sales	20	10 06-May-09 1955	kanak	35400	AA	855			1 winders	
Systematic Sales	100									
B Sample-Authorization	25								af Neinder	
Sample-Weblog	200								I + Comments	
- W Sample1-Weblog	205									
									E Add comment	
					_					
					N	umber of	Records: 10			
					N	umber of	Records: 10	Ì		
					N	umber of	Records: 10]		
					N	umber of	Records: 10	1		
					N	umber of	Records: 10	l		
					N	umber of	Records: 10	1		
					N	umber of	Records: 10	l		
					N	umber of	Records: 10			
					N	umber of	Records: 10	Ì		

Fig. 3.5.7 Random Sampling Result

Exercises:

- 1. Perform sampling on different sample files available in sample database.
- 2. Perform stratified sampling using one of the tables in the sample database.

3.6 STRATIFICATION

Learning Objectives

To stratifying the data from the file into bands and gaining the profile of the data. The data can be stratified based using the numeric, date or character field to select a random sample of records for testing.

Function Description

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 accumulating the records from the database into the appropriate bands. By totaling the number of records and value of each band, you can gain a profile of the data in the database. You can then investigate any deviations from expected trends. You may have up to 1,000 stratification bands. The stratification analysis is also useful for determining high and low cut-off values for testing exceptional items.

A numeric stratification analysis can also be created for each unique value or key in a field by selecting that field from the **Group by** drop down list. For example, you could produce a profile of sales for each salesperson. This can potentially create an extremely large volume of output however the maximum number of groups that will be displayed in the result is 80. If there are more than 80 groups, only the first 80 are displayed. Therefore, there is an option to specify low and high cut-off values to restrict output. Only groups whose total value of transactions is between the specified range are output.

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.

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





Step by Step Procedure for performing Numeric Stratification Ensure that **Sales Transaction-Sales Trans** is the active database.

Location

• Analysis >Stratification.



Fig. 3.6.1 Stratify Location

- (a) In the **Field to stratify** box, select **SALES**.
- (b) In the Fields to total on box, select SALES.
- (c) Group By **USERID**.
- (d) Specify the stratification bands:
 - Change the increment to 10,000.
 - Click in the **< Upper Limit** text box of the first row. Note the text box is filled with the value 10,500.
 - Click the second row of the spreadsheet area. This automatically fills with 10,500 to 20,500
 - Highlight the next three rows of the spreadsheet area to take the range to 50,500.

				100	Cut off
Field to stratity:		>= Lower Limit	< Upper Limit	1	Fields
DATE	1	500	10,500	-	Linuar
NAME	2	10,500	20,500		Insert
SALES	3	20,500	30,500		Bemove
USERID	4	30,500	40,500		Cancel
	5	40,500	50,500		Caricer
	6				Help
	7				
Fields to total on:	8				
SALES	9				
	10				
	11				
	12				
	13				
	14				
	15			-	
				141	
Criteria:	1				
Consta database 🖂 Jackeda	attention latencele	IZ Create would			

Fig. 3.6.2 Stratification

202



Ensure the **Create result** check box is selected. In the **Result name** box, enter **Stratification Sales**.

- (e) Click **OK**.
- (f) The **Numeric Stratification** result output for the **Sales Transactions** database becomes active and appears as a link in the **Results** area of the **Properties** window. Note that there are 6 accounts in the first band (>= 500 and < 10,500).

Systematic Custom	ners.IMD T Sales Transaction-S	alesTrans.IMD				~ 3
688-8-						
Totalled on: SALES		•	USERID= RA			
Stratum #	>+ L Limit	< U Limit	# Records	(%) # Records	SALES	(%) SALES
1	500	10,500	ē	4.88	61,200	3.36
2	10,500	20,500	117	95.12	1,760,700	96.64
3	20,500	30,500	0	0.00	0	0.00
4	30,500	40,500	0	0.00	0	0.00
5	40,500	50,500	. 0	0.00	0	0.00
		Lower limit exceptions:	0	0.00	0	0.00
		Upper limit exceptions:	0	0.00	0	0.00
		Totals:	123	100.00	1,821,900	100.00

Fig. 3.6.3 Stratification Result

Exercises

Use sample-employees file and perform numeric stratification on salary with various bands. Use sample-employees and perform character stratification based on name.

3.7 SORTING

Learning Objectives

• To create a new database in which its records are physically sorted in a specified order.

Ensure that **Sales Transaction-Sales Trans** is the active database.

Location

• Data>Sort.





File Edit View	Data	Analysis Sampling	To	ols Windo	w Help
	133	Extractions Search	۲		1 🔛 🤅
File Explorer		Indices	_	Sales Tra	insaction-Sa
	01	finances		USERID	DATE
	11.2	30IL		MM	01-Apr-09
IDEA Files		Describe Sorted Order		RA	01-Apr-09
Name	2-	Group Off		RA	01-Apr-09
- ACCESS-Datab	20	Group Records		RA	01-Apr-09
- Customer-Dat			-	RA	01-Apr-09
Sales Transact	國	Define Action Field		RA	01-Apr-09
- Sales-Sheet1	0	Field Manipulation		RA	01-Apr-09
- Sample-Autho		Chart Data		RA	01-Apr-09
	and .	1	0	RA	01-Apr-09
			10	RA	01-Apr-09
			11	RA	01-Apr-09



Function Description

The fields you select to sort the records by are known as keys. A sort order may contain up to eight keys. When creating a sort order, the most significant field is selected first (primary key), followed by the next most significant field, and so on down to the least significant field (secondary keys).

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. 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.

Step by Step Procedure for performing Sorting

Click on the **Field** list button (down arrow) on the **Field** column of the first row of the dialog grid to display a list of all fields in the database. Select the required field.

The **Direction** column displays **Ascending** order for each field selected. To change the order for a field, click the **Direction** column to activate the **Direction** list button.

To add further fields to the sort order, click the next available row in the **Field** column and select the field and its direction as above. A maximum of eight fields may be entered. **Note**: If required, you can delete a key from the sort order by selecting the required row and then clicking **Delete Key**.

Here **DATE** field is selected in descending direction.

Data Analysis Using IDEA



ase sort order on:		
NEW INDEX		
Field	Direction	ОК
DATE	Descending	Delete Key
		Cancel
		Help

Fig. 3.7.2 Sort Database

- (e) In File name box enter **Sorted database**.
- (f) Click **OK**. Date field is sorted in descending direction as shown in the below figure.

Sorted-Database

IDEA - Sorted database.IMD	The other Designation of the local distribution of the local distribut					and the second	
Eile Edit View Data Analys	sis Sampling	Io	ols <u>W</u> indo	ow <u>H</u> elp			
	w 🗞 🗞	ľ) 000 🗐 🕄 ·	• 1 🗑 🐻 🖪	
File Explorer	4	1	Sales Tra	nsaction-Sale	esTrans.IMD	Sorted database.IN	ID
			USERID	DATE	INVOICENO	NAME	SALES
IDEA Ellar		1	AA	09-May-09	1999	vivek	44900
IDEA Flies	2	2	AA	09-May-09	1998	martin	45150
Name	Records	3	AA	09-May-09	1997	nisha jain	45400
ACCESS-Database	1,000	4	AA	09-May-09	1996	ayushmaan	45650
- Ustomer-Database1	341	5	AA	09-May-09	1995	siddharth	45900
E-W Sales Transaction-Sale	999	6	MM	08-May-09	1994	ajmera	28650
Sorted database	999	7	MM	08-May-09	1993	maullik	28900
- I Sales-Sheet1	192	8	MM	08-May-09	1992	modi	29150
Sample-Authorization	15	9	MM	08-May-09	1991	lainy	29400
	1	10	MM	08-May-09	1990	lara	29650
		11	MM	08-May-09	1989	dhupia	29900
		12	AA	08-May-09	1988	vishvanath	30150
		13	AA	08-May-09	1987	maria	30400
		14	AA	08-May-09	1986	mallika	30650
		15	AA	08-May-09	1985	malayka	30900
		16	AA	08-May-09	1984	tanisha	31150
		17	AA	08-May-09	1983	dimple	31400
		18	AA	08-May-09	1982	khanna	31650
		19	AA	08-May-09	1981	tiwary	31900
		20	AA	08-May-09	1980	shaily	32150
		21	AA	08-May-09	1979	sandra	32400
		22	AA	08-May-09	1978	sapna	32650

Fig. 3.7.3 Sorted database





Exercises:

- 1. Use Sample-Employees and sort on city.
- 2. Use Sample-Employees and sort on First Name, City and Salary.
- 3. Duplicate Detection and Gap Detection

3.8 DUPLICATE DETECTION

Learning Objectives

- To test the validity of invoices.
- To test for duplicate invoice numbers.

Function Description

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

These are: Duplicate Detection and 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 testing after running the respective duplicate and gap tests within IDEA.

Duplicate Detection and Gap Detection tests are standard passing tests which are run on every database right at the (inception prior to detailed excepting and analytical testing within IDEA. The tests do not require much querying experience and resemble plug-n-play tests. Both tests run largely on formatted sequential data fields like Invoice Number, Purchase Order Number, Cheque Number, etc.

Step by Step Procedure for Duplicate Detection

Ensure that **Sales Transaction-Sales Trans** is the active database.

Location

• Analysis > Duplicate Key > Detection.

File Edit View Data	Analysis Samplin	g To	ols Win	dow	Help			
File Explorer	BΣ Summarizatio BΣ Pivot Table B↔ Stratification	n			action-Sa	000	◆ 🖥 🍘 💽 👯 🕥 Sorted database.IM	D
G C 2 P C	Duplicate Key	<u>1</u>		×	III Det	tection	NAME	SALES
IDEA Files	Gap Detection	l			Exc	lusion	tonima ch himanshu	20500 20000
Name	Aging	Aging				1003	meenal pasari	19500
ACCESS-Database	Benford's Law	4444			1-Apr-09	1004	s tibrewal	19000
Customer-Database1	Advanced Sta	tistical N	Methods		1-Apr-09	1005	dhwani	18500
Sales Transaction-Sale		10	1KA	- 0	1-Apr-09	1006	chahat	18000
- Sales-Sheet1	192	7	RA	0	1-Apr-09	1007	jayesh panchamiya	17500
Sample-Authorization	15	8	RA	0	1-Apr-09	1008	jayshree chandan	17000
		9	RA	0	1-Apr-09	1009	p m murthy	16500
		10	RA	0	1-Apr-09	1010	dinesh	16000
		11	RA	0	1-Apr-09	1011	varun gollecha	15500
		12	RA	0	1-Apr-09	1012	nekita	15000
		13	RA	0	1-Apr-09		p m mittal	15000
		14	RA	0	1-Apr-09	1014	ruhi fatima	12000
		15	RA	0	1-Apr-09	1015	shah ankana	12000

Fig. 3.8.1 Identify Duplicates Location



(a) The **Duplicate Key Detection** dialog box appears.



Fig. 3.8.2 Duplicate Key Detection

- (b) Leave the **Output Duplicate Records** option selected.
- (c) Click Fields and select **INVOICENO**, **INVDATE** and **USERID**.

ields	×
Fields to include:	ОК
DATE	
NAME	
USERID	Liear All
	Cancel
	Help

Fig. 3.8.3 Fields

- (d) Click Key. The Define Key dialog box appears
- (e) .In the **Field** column, select **INVOICENO** and leave the direction as **Ascending**.

Define Key		x
Base index on:		
NEW INDEX		-
Field	Direction	ОК
INVOICENO	 Ascending 	Delete Key
		Cancel
		Help

Fig. 3.8.4 Define Key





- (f) Click **OK** to return to the **Duplicate Key Detection** dialog box.
- (g) In the **File name** box, enter **Duplicate**.
- (h) Click **OK** to run the task.

Duplicate Result

- (a) In the resultant database of 8 transactions with DATE, INVOICENO and USERID duplicate values are investigated.
- (b) Duplicate file is opened as current active database.
- (c) The given result is sorted based on INVOICENO field in ascending direction.



Fig. 3.8.5 Duplicate Result

Exercises:

- 1. Use Sample-Employees file and find out duplicates based on name.
- 2. Use Sample-Employees and identify duplicates based on fields: First Name and Name.
- 3. Use Sample-Employees file and identify duplicates based on Address.

3.9 GAP DETECTION

Learning Objectives

- To test for completeness.
- To test for gaps in the invoice number sequence.

Ensure that **Sales Transaction** is the active database.

Location

• Analysis >Gap Detection.


IDEA - Sales Transaction-S	alesTra	ans.IMD								
File Edit View Data	Anal	ysis Sampling Tools	Wind	ow	Help					
	34	Pivot Table				₩ 🕲 000 🏦 + 🖥 🕯				
File Explorer	Ben Stratification				ales Transaction-SalesTrans.IMD					
	2+0	Dunlicate Key			SERID	DATE	INVOICENO			
IDEA Files	1	Gap Detection		м	01-Apr-09	1001				
Name	53	Aging Benford's Law Advanced Statistical Methods				01-Apr-09 01-Apr-09	1002			
Customer-Database1 Sales Transaction-Sale						01-Apr-09	1005			
Sales-Sheet1		192 7				01-Apr-09	1007			
				R/	4	01-Apr-09 01-Apr-09	1009			
			10	R/	4 4	01-Apr-09 01-Apr-09	1010 1011			
		12 RA				01-Apr-09 01-Apr-09	1012			
			14	R/	4	01-Apr-09	1014			
			15	R/	4	01-Apr-09	1015			

Fig 3.9.1 Gap Detection Location

Step by Step Procedure for Gap Detection

- (a) In the **Field to use** list box, select **INVOICENO.**
- (b) In the **Output** area, ensure the **Create result** check box is selected.
- (c) In the **Result name** box, enter **Gap Detection Sales**.
- (d) On clicking Ok, the result database contains the gaps occurring **From: INVOICENO**, **To: INVOICE** and **Number** in the active database.

Field to use: INVOID	CENO 👻	<u>C</u> riteria:	ОК
Character			Cancel
Sample record:	1001		Help
Masked record:	1001	and the second	
<u>M</u> ask:	NNNN		
Dutput			
Create gatabase		🕼 Create result	
Ele name: Gap Det	ection	Besult name: Gap Detection Sales	

Fig 3.9.2 Gap Detection





Sales Transaction-SalesTrans.IMD		•
88.1.		
Key Value		
From: INVOICENO	To: INVOICENO	Number
E 0001	1000	1,000
H 1013	1013	1
⊞ 1035	1035	1
⊞ 1158	1158	1
⊞ 1455	1455	1
⊞ 1709	1709	1
	Total number of items detected	1,005
	Total number of gaps detected	6

Fig 3.9.3 Gap Detection Result

Exercise:

1. Use Sample-Bank Transactions file and identify gaps based on date.

3.10 AGING

Learning Objectives

- To age a selected database from a particular date for up to six specified intervals. These intervals can be days, months, or years.
- To age the outstanding debts at the yearend in order to determine provisions required against bad debts.

Function Description

Aging function presents aged summaries of data. This summary may be based on the current date or a specified cutoff date. Use the Aging task to age a selected database from a particular date for up to six specified intervals. These intervals can be days, months, or years. For example, you can age the outstanding debts at the yearend in order to determine provisions required against bad debts.

The most common use of the Aging task is with Accounts Receivable or Debtor Ledgers. However, also consider using Aging on inventory databases (date of last movement) or on short-term loan databases.

The Aging task optionally produces:

- A detailed aging database
- A key summary database
- A Results output

Ensure that **Sales Transaction-Sales Trans** is the active database.



Location

• Analysis >Aging.

File Edit View Data	Analysis	Sampling	To	ols Wi	indow	Help	
🗟 🗝 🗄 - 🖄	B∑ Sum B Pivo	imarization t Table					● 000
File Explorer	Pro Stratification				action-SalesTrans.IMD		
	Dun	licate Key				DATE	INVOICENO
IDEA Files	Gap Detection				1-Apr-09	1001	
Name	Agir	ng				1-Apr-09	1003
ACCESS-Database	Benford's Law					1-Apr-09	1004
Customer-Database1	Adv	Advanced Statistical Methods			•	1-Apr-09	1005
Sales Transaction-Sale		0	O KA		1-Apr-09	1006	
Duplicate		8	7	RA	(01-Apr-09	1007
		192	8	RA	0	01-Apr-09	1008
Sample-Authorization		15		RA	0	01-Apr-09	1009
			10	RA	(1-Apr-09	1010
			11	RA	0	01-Apr-09	1011
			12	RA	(01-Apr-09	1012

Fig. 3.10.1 Aging Location

Step by Step Procedure for performing Aging:

- (a) Enter the Aging date to 2012/06/27 i.e yyyy/mm/dd format.
- (b) Set the Aging field to use to **DATE.**
- (c) Set the Amount field to total to **SALES**.
- (d) Aging interval in Days. The other options are Months and year.
- (e) The **Aging** task optionally produces:
 - A detailed aging database.
 - A key summary database.
 - A result output.
- (f) Select only Create result option.
- (g) In the Name box enter **Aging Sales**.
- (h) On clicking Ok, the below output is displayed.





Aging date:	2012/06/27		OK
<u>C</u> riteria:			Lancel
Aging field to use:	DATE	-	Help
Amount field to total:	SALES	•	
Aging interval in:	Days •]	
1: 30 🔹	3 90 🔅	5 150 0	
2 60 💠	4: 120 💠	<u>6</u> 180 ≑	
Generate detailed	aging database:		
File name: Aging		Eields	
🗌 Generate key gum	mary database:		
File name: Age Sum	mary	Key	
Create result:			
Name: Aging Sal	40		

Fig. 3.10.2 Aging Location

Aging Result

ſ	Sales Trans	saction-SalesTrans.II	ND						▼ X
ľ	Int (Days)	# Records	(%) Records	Debits	(%) Debits	Credits	(%) Credits	Net Value	(%) Net Value
ľ	0	0	0.00	0	0.00	0	0.00	0	0.00
ľ	30	0	0.00	0	0.00	0	0.00	0	0.00
ľ	60	0	0.00	0	0.00	0	0.00	0	0.00
	90	0	0.00	0	0.00	0	0.00	0	0.00
ľ	120	0	0.00	0	0.00	0	0.00	0	0.00
ľ	150	0	0.00	. 0	0.00	0	0.00	0	0.00
ľ	180	0	0.00	. 0	0.00	0	0.00	0	0.00
ľ	180+	<u>999</u>	100.00	18,366,960	100.00	0	0.00	18,366,960	100.00
ľ	ERR	0	0.00	0	0.00	0	0.00	0	0.00
ľ									
ľ	Totals:	999	100.00	18,366,960	100.00	0	0.00	18,366,960	100.00

Fig. 3.10.3 Aging Result

Exercise:

Use Sample Bank Transactions file and perform aging by entering the aging date as: 2003/11/06.

3.11 DATA EXTRACTION

Learning Objectives

• To perform an extraction to identify accounts where the new credit limit has been exceeded.

Function description

Extract selected data from a file for further investigation for creating a new file of logically selected records. For example: you can use Direct Extraction to perform a single extraction on a database, or up to 50 separate extractions with a single pass through the database.



Step by Step Procedure for Data Extraction

Ensure that **Sales Transaction** is the active database.

Location

• Data >Extractions >Direct Extraction.

File Edit View	Data	Analysis Sampling To	ols	Window Help	
		Extractions +	8	Direct Extraction.	
File Explorer		Search Indices		Indexed Extractio Key Value Extract	n ion
IDEA Files	¥4	Sort Describe Sorted Order	-	Top Records Extr	action
Name	-22	Group Off	RA	01-Apr-09	1002
ACCESS-Datab	20	Group Records	RA	01-Apr-09	1004
- Ustomer-Dat	1000	R/		01-Apr-09	1005
Sales Transact	Flot	Define Action Field	RA	01-Apr-09	1006
- Sales-Sheet1	3	Field Manipulation	RA	01-Apr-09	1007
Sample-Autho	100	Chart Data	RA	01-Apr-09	1008
l	1000	11.0	RA	01-Apr-09	1009

Fig. 3.11.1 Direct Extraction Location

Rec	ords to extrac	t @ Al	Starting record #	1		Οκ
		© Bange	Ending record #:	999		Create Field
		al de la companya de		La la construcción de la constru		Eields
a	abase groei.	File N	ime	1	Criteria	Delete
1	Sales Greate	r than 65,000			SALES >= 65000	Cancel
2						Help

Fig. 3.11.2 Direct Extraction

- (a) In the **File Name** column, replace the default file name with **Sales Greater than 20,000**.
- (b) Click the **Equation Editor** Button, and then enter the equation **SALES** >= 20000.
- (c) Click the **Validate and Exit** button to return to the **Extract to File(s)** dialog box.
- (d) On clicking Ok creates a new child database with name **Sales Greater than 20,000** which contain records where Sales is greater than 20,000.





Record Extraction Result

<u>File Edit View Data Analysis</u>	Sampling Too	ls <u>W</u> in	dow Help	>			
			n 🔛	O 000	1 +7 🞬		
File Explorer		4	Sales Ti	ransaction-Sa	alesTrans.IMD	Sales Greater th	an 65,000.IN
			USERID	DATE	INVOICENO	NAME	SALES
IDEA Files	_	1	SF	22-Apr-09	1505	karuna	65150
Name	Records	2	SF	23-Apr-09	1539	munshi	69900
	necords	3	SF	23-Apr-09	1540	merlyn	69650
ACCESS-Database 1 ACCESS-Database 1 Customer-Database1 Sales Transaction-SalesTrans Sales Greater than 65,000 Sales-Sheet1	1,000	4	SF	23-Apr-09	1541	smith	69400
	341	5	SF	23-Apr-09	1542	vienetta	69150
	999	6	SF	23-Apr-09	1543	becker	68900
	20	7	SF	23-Apr-09	1544	jacquillene	68650
	192	8	SF	23-Apr-09	1545	jack	68400
Sample-Authorization	15	9	SF	23-Apr-09	1546	yohana	68150
		10	SF	23-Apr-09	1547	shirley	67900
		11	SF	23-Apr-09	1548	tina	67650
		12	SF	23-Apr-09	1549	dheeru	67400
		13	SF	23-Apr-09	1550	ratan	67150
		14	SF	23-Apr-09	1551	jamshedji	66900
		15	SF	23-Apr-09	1552	jachson	66650
		16	SF	23-Apr-09	1553	micheal	66400
		17	SF	23-Apr-09	1554	stephen	66150
		18	SF	23-Apr-09	1555	apeksha	65900
		19	SF	23-Apr-09	1556	poojashree	65650
		20	SF	23-Apr-09	1557	nanjapa	65400

Fig. 3.11.3 Direct Extraction Result

Exercises:

- Use Sample-Suppliers file and extract supplier details where supplier no. is from 30000 to 90000.
- Use Sample-Suppliers file and extract supplier details where country is "Chile".

3.12 BENFORD'S LAW

Learning Objectives

To compare the data with the data pattern predicted by **Benford's Law** analysis.

Function Description

Benford's Law states that digits and digit sequences in a data set follow a predictable pattern. The **Benford's Law** task generates a database, and optionally a **Results** output, that you can analyze to identify possible errors, potential fraud, or other irregularities. If artificial values are present in the selected database, the distribution of the digits may have a different shape, when viewed graphically, than the shape predicted by **Benford's Law**.

A **Benford's Law** analysis is most effective on data:

- Comprised of similar sized values for similar phenomena.
- Without built-in minimum and maximum values.
- Without assigned numbers, such as bank accounts numbers and zip codes.
- With four or more digits.

Step by Step Procedure for using Benford's Law function

Ensure that **Sales Transaction** is the active database.



Location

• Analysis >Benford's Law.

File Edit View Data	Analysis	Sampling Tool	s Windo	w Help						
File Explorer	B∑ Summarization J Pivot Table Stratification			iales Tra	ales Transaction-SalesTrans.IMD					
022021	Duni	icate Key		SERID	DATE	INVOICENO	NAME	SALES		
	note ney		M	01-Apr-09	1001	tonima ch	20500			
IDEA Files	Gap I	Detection		1	01-Apr-09	1002	himanshu	20000		
Name	Agin	g		1	01-Apr-09	1003	meenal pasari	19500		
- ACCESS-Database	Benf	ord's Law		1	01-Apr-09	1004	s tibrewal	19000		
- Customer-Database1	Adva	nced Statistical Me	thods		01-Apr-09	1005	dhwani	18500		
- Sales Transaction-Sale			10	TOA	01-Apr-09	1006	chahat	18000		
- Sales-Sheet1		192		RA	01-Apr-09	1007	jayesh panchamiya	17500		
Sample-Authorization		15	8	RA	01-Apr-09	1008	jayshree chandan	17000		
			9	RA	01-Apr-09	1009	p m murthy	16500		
			10	RA	01-Apr-09	1010	dinesh	16000		
			11	RA	01-Apr-09	1011	varun gollecha	15500		
		12		RA	01-Apr-09	1012	nekita	15000		
			13	RA	01-Apr-09		p m mittal	15000		
			14	RA	01-Apr-09	1014	ruhi fatima	12000		
			15	RA	01-Apr-09	1015	shah ankana	12000		

Fig. 3.12.1 Benford's Law Location

The analysis counts digit sequences of values in the database and compares the totals to the predicted result according to **Benford's Law**. Non-zero digits are counted from left to right and values below 10 are ignored.

Fjeld to analyze: SALES			- OK		
Include Values <u>Positive</u> Negative	Option V	ns S <u>h</u> ow boundaries	Cancel		
Result					
Create result	Benford	ord			
Analysis Type					
First digit:	Benford First Digit	8			
First two digits:	Benford First Two	Digits			
First three digits:	Benford First Thre	e Digits			
Second diait	Benford Second I	Digit			

Fig. 3.12.2 Benford's Law

Steps for Benfords Law

- (a) Select the field to analyze. In the **Field to analyze** drop-down list, select the Numeric field for which you want to analyze with Benford's Law.
- (b) Select the number type. In the **Include Values** area, select the check box for the required number type (**Positive** or **Negative**).
- (c) Optionally, specify the upper and lower boundaries. The upper and lower boundaries define the acceptable range of values where the actual result can appear. Select the **Show boundaries** check box to include the boundaries in the **Results** output and resultant database.
- (d) Select the required analysis types. In the **Analysis Type** area, select the required analysis types, and then accept or change the associated database file names.
- (e) Optionally, create a Results output.





- (f) On Clicking Ok, we get results for Benford's First Digit, Second Digit, First Three Digit and First Two Digit.
- (g) In the below image, Benford First Digit result is displayed. To view Second, Third and Fourth Digit result, click on the IDEA files explorer.

DEA - Benford First Digit.IMD				-	Second second	Acres in the local division of the local div	of the local division of the local divisiono
File Edit View Data Analysis Sampling Iools		ndow He	Transaction-Sale	t + T 😂 🖪	Benford First Dig		
	Í	DIGITS	EXPECTED	LOWBOUND	HIGHBOUND	ACTUAL	DIFFERENCE
NEA Eiler	1	1	300.73	282.57	318.89	142	158.73
A A A A A A A A A A A A A A A A A A A	2	2	175.92	160.80	191.03	158	17.92
Name Records	3	3	124.81	111.67	137.96	310	-185.19
ACCESS-Database 1,000	4	4	96.81	85.03	108.60	131	-34.19
- Customer-Databasel 341	5	5	79.10	68.33	89.88	102	-22.90
🖶 🗊 Sales Transaction-Sales Trans 🛛 999	6	6	66.88	56.89	76.87	84	-17.12
- Benford First Digit 9	7	7	57.93	48.58	67.29	46	11.93
- Benford First Three Digits 900	8	8	51.10	42.27	59.93	12	39.10
- Benford First Two Digits 90	9	9	45.71	37.33	54.09	14	31.71
Benford Second Digit 10	P	-					
Sales-Sheet1 192							
The second state of the se							

Fig. 3.12.3 Benford's LawFirst Digit Result

Exercise: Use Sample-Bank Transactions file and perform Benford's Law function on amount.

3.13 CONSOLIDATION OF DATA

Learning Objectives

- To summarize data and create a report based on many calculations.
- To define how your data is displayed and organized.

Function Description

Consolidation is the process of combining values from several ranges of data. Data can be consolidated by Pivot Table.

Step by Step Procedure for reporting

Ensure that **Sales Transaction** is the active database.

Location

• Analysis>Pivot Table.





In the **Result name** box, enter a name for the **Pivot Table Results** output.

vot Table		×
Result name:	Pivot Sales Transaction	ОК
		Cancel

Fig. 3.13.2 Pivot Table

Click **OK**. The **Pivot Table Field List** dialog box appears. It contains the list of available fields in the database, with their field types in adjacent brackets.

▶ == 3= 3= 1 ≥ ⊕	= & & B	60 80 ·		ÞIPRRI,
File Explorer	-	Sales Transact	ion-SalesTrans.IMD	
	1		• & 🖽	
IDEA Files		Drop Page Fields Here		
Name	Records	Drop Co	fumn Fields Here	
ACCESS-Database	1,000	Drop Now Pi Drop Da	La merris mere	
- Customer-Database1	341			
Sales Transaction-Sales Trans	999			
			PivotTable Field List Ding fields to the pivot table: DATE (D)	
			Add To Renove Does	

Fig. 3.13.3 Pivot Table Field List

- (a) Create the table by dragging fields to summarize from the **Pivot Table Field List** dialog box to the following areas: **Drop Column Fields Here**, **Drop Row Fields Here**, **Drop Data Items Here** and **Drop Page Fields Here**:
- (b) You can drag the same field more than once into the data area creating a new statistic each time sum, amount, and so on.
- (c) You can add multiple fields to the Column and/or Row area. When more than one field is added to a column or row area the data in the rightmost header column or lowest header row is grouped by data in the preceding row or column.
- (d) You can also select a field from the **Pivot Table Field List** dialog box, select the area of the pivot table to which you want to add the selected field from the drop-down list, and then click **Add To**. Once a field has been added to the table, it is highlighted in the **Pivot Table Field List** dialog box.
- (e) For every field added to the pivot table, its sort order is added to the **Indices** area of the **Properties** window. You may delete these indices from the **Indices** area without affecting the pivot table.
- (f) From the **Pivot Table Field List** dialog box, click **Close.**





Pivot Table Result

/ 🗑 Sale	s Transaction	SalesTrans.I	MD										•	- ×
	29.428	2 53 5												
Drop Page I	Fields Here													
Sum: SALES	DATE 🗖													
USERID 💂	2009/04/01	2009/04/02	2009/04/03	2009/04/04	2009/04/05	2009/04/06	2009/04/07	2009/04/08	2009/04/09	2009/04/10	2009/04/11	2009/04/12	2009/04/13	20
AA						400,000	1 · · · · · · · · · · · · · · · · · · ·							
MM	20,500	339,500				125,000								
RA	253,500				1		1				51,750	250,000	195,300	
SA	52,100	25,850	93,100	165,750	223,550	61,000	12,460	74,750	116,490	111,360	225,420			
SF						216,000	183,000							
Total	326,100	365,350	93,100	165,750	223,550	802,000	195,460	74,750	116,490	111,360	277,170	250,000	195,300	
		Drog field DATE (D INVOICE NAME (C EXTERNO USERID	Held List is to the pivot to 0 NIXO (C) 0 NIXO (C) 0 NIXO (C) 0 NIXO (C)	able:	• Remove) Close	8							

Fig. 3.13.4 Pivot Table Result

Exercises:

- 1. Use Sample-Employees file and perform pivot table function with branch as row and salary as column name.
- 2. Use Pivot table to find out salary as per country/branch and total of salary.

3.14 EQUATION EDITOR

Learning Objectives

• To create formulae and equations in IDEA.

Functions

Use functions to perform more complex calculations and exception testing. You can use them for date arithmetic, text searches, and some statistical operations. They are very similar in style and operation to functions found in other software packages such as Microsoft Excel, Lotus 1-2-3, and dBASE. Each function calculates a result based upon the parameters passed to the function. Parameters are passed in parentheses.

Equation conot Cash Cash	8		8 al	Al Character Numeric Nateric Matching Conditional Financial Custom Functions	●Functions offunctions are available for performing more complex operations such as dette anthreets, statistical calculations and text searches. IDEA functions begin with the "g" symbol. They are very similar in style and operation to functions found in other software packages such as Microsoft Excel, Johan 1-2-3, and BURSE. Each offunction calculates	× ·
Field	Туре	Len	Dec		a result based upon the parameters passed to	
DATE	Date	8			the @Function.	
INVOICENO	Character	9	-		Parameters are passed	
NAME	Character	17			example, @Round	
SALES	Numeric	8			(Amount) will round	
USERID	Character	2			a field called "Amount"	
					to the nearest whole	

Fig 3.14.1 Equation Editor

Select the type of function you require from the following category list to view the details for the associated functions:

- All Complete list of all functions.
- Character Text manipulation or conversion Numeric Calculations and statistics
- Date and Time Date and time calculations and conversions
- Matching Matching of multiple and similar items
- Conditional IF statements to select different items
- Financial Financial calculations
- Custom User-defined functions

Entering Functions in Equation Editor

Entering a function involves entering the function name along with any necessary parameters. Functions are entered in the **Equation** area of the **Equation** Editor.

Entering the Function Name

To enter the function name in the **Equation** area, expand the appropriate function category and then double-click the required function name. Alternatively, if you know the name of the function, you can type it directly into the **Equation** area. As you type the function name, the **Equation Editor** provides a list of possible function names.

Entering the Parameters

Parameters may be complete equations with references to other functions, such as nested functions, or simple constants. Parameters are separated by a list separator. The list separator is defined in your Windows Regional Settings and is usually a comma or a semi-colon. Most functions require a specific number of parameters. A syntax error occurs if you provide the wrong type of parameter to a function.

ТҮРЕ	EXPLANATIONS	EXAMPLES
String	Represents text or a character expression. A character expression can contain characters or a Character data field.	"Smith" "123456" "Joe" + "Bloggs"
Number	Represents a numeric expression. A numeric expression can contain numeric constants or Numeric data fields.	123.45 ACCNT BAL QTY*COST
Time	Represents a time expression. It is displayed in the grid as HH:MM:SS. A time expression can contain a time constant expressed as a string, a numeric constant, a Time field or a Numeric field. Numeric values are interpreted as the number of seconds.	"23:59:59" TRANS_ TIME
Date	Represents a date expression. A date expression is a character string that is eight characters long and contains a valid date in YYYYMMDD format.	"19960131" TRANS_ DATE





3.15 REPORTING

Learning Objectives

• To generate a report for the current active database.

Function Description

You can use the reporting feature to generate various types of reports from idea.

Step by Step Procedure for reporting

Create a report using the view settings by selecting **File >Print >Create Report**. Alternatively, click the **Create Report** button on the **Operations** toolbar. The **Report Assistant** dialog box appears.

Report Assistant

Accept all defaults and select the **Allow headings to span multiple lines** check box. Click **Next**.

	What tune of report would you like to build?
	What type of report would you like to build?
	<u>H</u> orizontal
	⊘ Vertical
Nome Steel City Per	Qne record per page
	Itee existing report
	Create new report
	-
	Allow headings to span multiple lines
	Show record numbers
	V Include database name in report
< <u>B</u> ack	k Next > Finish Cancel Help

Fig. 3.15.1 Report Assistant

Step 1: Heading Step

Modify the report headings as follows, and then click **Next**:

- Date Invoice Date
- INVOICENO Invoice Number
- NAME User Name
- SALES Sales
- USERID User Id

Data Analysis Using IDEA



Report Assistant - Headi	ngs
DATE INVOICENO NAME SALES USERID	Headings font: Alignment Left O Bight O Center O Justified Text Invoice Diste
	ack Next > Finish Cancel Help

Fig. 3.15.2 Report Assistant-Heading

Step 2: Define Breaks Step

- (a) During this step, define if and where control breaks (required for totals in reports) are required.
- (b) Sequence the database records in the following order:
- INVOICENO: Ascending

ase break on:	rule report to break or	
No index		•
Field	Direction	Delete Key
NVOICENO	★ Ascending	

Fig. 3.15.3 Report Assistant - Define Breaks

(c) Click Next.

Note: IDEA displays the records in the report in the order of the index. IDEA displays the index description in the **Indices** area of the **Properties** window once you have completed the report.

INFORMATION TECHNOLOGY

221



Step 3: Report Breaks Step

Create a break and total the **Sales** field for each **INVOICENO**. Select the options as in the image below, and then click **Next**.

Break Keys INVOICENO	Count records in break Show break line Show leading break	Fields to Total
Preak on this field Show shading Text: Background: Eont	Break gpacing: 1 Line	Show shading Use currency symbol Text: Background: Fort.

Fig. 3.15.4 Report Assistant - Report Breaks

Step 4: Grand Totals Step

Create grand totals for the **Sales** field only, set the font style to bold, and then click **Next**.

Select the fields to total:	
SALES	✓ Show shading ✓ Use currency symbol
	Text:
	Background:
	Eont

Fig. 3.15.5 Report Assistant -Grand Totals

Step 5: Header/Footer Step

Enter/select the following information, and then click **Finish**.

- **Print cover page**: Select this check box.
- **Title**: Sales Transactions.
- **Comments**: Ordered by User Id and date.
- **Prepared by**: Enter your name or initials.
- **Header**: Enter the name of your organization.
- **Date/Time**: Accept the defaults unless you have particular preferences.

Note: The options you have selected effect how the report is printed. The name entered into the **Prepared by** field appears on reports accessed via the **Print Preview** of a **Results** output.

Data Analysis Using IDEA



IN FILL COVE	rpage	
<u>T</u> itle	Sales Transaction	A
<u>C</u> omments	Ordered By User ID and Date	
Prepared by:	CWI	Cover Page Font
Header:	CaseWare Idea Inc.	Date: Upper right 🗸
Eooter:		Ti <u>m</u> e: Upper right ▼
		Header/Footer Font

Fig. 3.15.6 Report Assistant – Header/Footer

On clicking finish, you can preview the report generated.

3.16 FIELD MANIPULATION

Learning Objectives

To view field definitions, add or delete fields, change field properties such as field name and type.

Function Description

You can use this to modify the fields as required. Please note that making changes to a field through **Field Manipulation** may cause any output based on that field (results, drill-downs, indices, views, etc) to appear incorrect or become invalid. Results may be made valid again by returning the settings to what they were when the result was created. To avoid this, instead of changing a field definition, append a new field to the database with the required definition. For example, instead of changing the type of a field from Character to Numeric, create a new Virtual Numeric field using @Val.

Select Data>Field Manipulation. Field Manipulation Dialog Box is opened.

	riciuivanie	Туре	Len	Dec	Parameter	Tag Name	Description	OK
1	DATE	Date	8		YYYYMMDD	<no tag=""></no>		
2	INVOICENO	Character	9			<no tag=""></no>		Append
3	NAME	Character	17			<no tag=""></no>		Delete
4	SALES	Numeric	8	0		<no tag=""></no>		Print
5	USERID	Character	2			<no tag=""></no>		C Low

Fig. 3.16.1 Field Manipulation

- Add a field
- Click **Append**.
- Enter the field definition:





- **Field Name**: A unique field name up to 40 characters in length that does not contain spaces or special characters.
- **Type**: Click the **Type** text box to display the type options, and then select the required type.
- **Length**: The total length of the field in characters. IDEA automatically determines the length for Editable Numeric fields.
- **Decimal**: If applicable, enter the number of decimal places. If the decimal is implied, enter 0. The maximum number of decimals places allowed is six.
- **Parameter**: Click the **Parameter** field to invoke the **Equation Editor**.
- **Tag Name**: If you have **Smart Analyzer** installed, click the **<No tag>** link to add a tag.
- **Description** (optional): A brief description of the field. The maximum length of the description is 256 characters.
- Click Ok
- Delete a field
- Click in the row corresponding to the field to be deleted.
- Click **Delete**.
- Click **Yes** to confirm the deletion.

Note: As a security feature, by default, native fields cannot be deleted

Change a field type

Click the **Type** text box to display the type options, and then select the required type.

Notes:

- Virtual field types can only be changed to other Virtual field types. For example, a Virtual Character field could be changed to a Virtual Date field.
- Field types cannot be changed for Editable, Boolean or Multistate fields; however the data within this field in the database can be changed.

Change a field name

Click in the **Field Name** cell of the field name to be changed and enter the new name.

Exercise:

Use Sample-employees file and modify different field and understand the impact.

Annexure - Using Audit software: Some Practical Examples

These are samples of analysis which can be performed using audit software for various types of audits/reviews.

Category	Type of Analysis	CAAT Function
Creditors	Payment terms	• Extraction, Date difference
	Credit Control Policy	• Extraction, Date difference,
	• Aging of creditors	Aging, date difference
	• Discount policy	• Stratification, Summarization,
	• Authorised suppliers	• Join, Relation, Sort,
	• Authorization	• Filter, Extract, Summarization
	Accounting	• Sort, Total, Count, Summarization
	• Supplier classification on purchases	• Sort, Relation, Join, Summarization,
	• Analysis of purchases by period	Sort
		• Aging, MIS, date difference
Tax Audit	 Identify cash loan\deposits>20000 	• Sort, Filter, Extract, Export, Index,
	• Identify cash payment > 10000	Export, analytical tool
	• Review of TDS compliance	 Join, Relation, Sort, analytical tool
	Analysis of Inventory	• Aging, MIS, Summarization, Count,
		Total
Financial	Review of Authorization	• Filter, Extract, Summarization
auun	Review of discount policy	• Stratification, Summarization,
	• Compliance with tax rates – sales tax, excise duty, etc	• Join, relation, filter, extract, sort, count, Total, Compare files
	• Verification of financial accuracy	• Summarization, Sort, Statistics,
	• Aging of debtors	Total, Summarization
		Aging, MIS, Periodicity check
Internal	Overall statistical analysis	• Statistics, Benford's Law,
Audit	• Identification of exception items	
	• Duplicate payment for invoices	Benford's Law, Filter, Gap detection
	• Debtors outstanding > credit period	• Identity duplicates, Sequence,
	• Age-wise analysis of debtors	Join, Relation, Aging, Summarization
	• Age-wise analysis of inventory	• Aging, MIS, Summarization, Statistics
		• Aging, MIS, date difference. total, count





Purchases	Duplicate payments	Duplicates, Sequence, Relation
	Invalid vendors	• Join, Relation, Sort, Summarization,
	Duplicate invoices	Filter
	Invalid purchases	Duplicates, Sequence, Relation
	• Payments without receipt of goods	• Join, Relation, Sort, Summarization, Filter
	Inflated prices	• Join Relation Compare files Filter
	• Excess quantities purchased	Exceptions
		• Filter, Sort, Stratify, Join, Relation
		• Filter, Sort, Stratify, Join, Relation
Payroll	Ghost employees	• Join, Relation, Summarization, Sort,
	• TDS	Statistics
	• Duplicate direct credits	Re-computation using analytical tool
	• Duplicate home addresses	• Identify duplicates, Sequence,
	• PO Box addresses	Identify duplicates Sequence
	• Work Phone Nos.	Relation
	Work Location	• Identify duplicates, Sequence,
	• Deductions	Relation
	• Vacation and sick leave	• Identify duplicates, Sequence, Relation
	Wage level	Identify duplicates Sequence
	Terminated employees	Relation
	Overpayment, overtime	• Re-computation using analytical tool
		• Re-computation using analytical tool
		• Join, Relation, Summarization, Statistics
		Join, Relation, Filter, Statistics
		• Summarization, Statistics, MIS, Join, Relation

Data Analysis Using IDEA



Aging	• Overdue A.R. and A.P.	• Aging, MIS, date difference
	• Favorable credit terms	• Aging, MIS, date difference,
	• Inventory turnover rates	• Stratify, Filter, Summarization,
	Dormant accounts	Identify Gaps, date difference
	 Records with future, blank or otherwise invalid dates 	• Sort, Filter, Sequence, Identify Gaps, Date difference, Statistics, Count
	• Items past a cutoff date	Date difference, Aging, Filter,
	• Contracts awarded before contract date	• Date difference, Aging, Filter, Summarization, Statistics
	 Transactions outside of billing period 	• Aging, Filter, Summarization, Statistics
	• Length in days of various activities	• Date difference, Aging, Filter, Summarization,

QUESTIONS

- 1. Computer Assisted Audit Techniques (CAATs) refers to using ______for increasing the effectiveness and efficiency of auditing.
 - A. Technology
 - B. Standards
 - C. Documentation
 - D. Systematic Process
- 2. Which of the following statements pertaining to CAAT is true?
 - A. CAATs are specialist tools designed for use by specialist IT auditors
 - B. CAATs refer to common techniques which can be easily mastered to audit in a computerized environment.
 - C. CAATs cannot be used for compliance audit.
 - D. CAATs can be used for Information systems audit only.
- 3. Which of the following statements pertaining to CAAT is incorrect?
 - A. CAATs are tools for drawing inferences and gathering relevant and reliable evidence.
 - B. CAATs provide direct access to electronic information
 - C. CAATs provide a mechanism to gain access and to analyze data
 - D. CAATs techniques are available in general audit software only.
- 4. Which of the following is a key reason for using CAATs?
 - A. Availability of input documents
 - B. Availability of visible paper trail.





- C. Need to access information from systems having different IT environments.
- D. Need to perform audit in lesser time.
- 5. Which of the following is a not a benefit of using CAATs?
 - A. Increased audit quality and comply with auditing standards.
 - B. Identify materiality, risk and significance in an IT environment.
 - C. Identifying fraud in audit area.
 - D. Ensuring better audit planning and management of audit resources.
- 6. For effective use of CAATs it is necessary to:
 - A. Understand programming language of the software being audited.
 - B. Obtain source code of the software being audited.
 - C. Obtain copy of data in their original format for independent analysis.
 - D. Audit data is retained in the original database format.
- 7. Which of the following step is important for validating the correctness of data obtained for audit?
 - A. Issue a letter for getting the requested data in specified form as per the audit objectives.
 - B. Obtain copies of record layout and definitions of all fields.
 - C. Ensure that auditor has an overall understanding of the data.
 - D. Compare the imported in audit software with control total of key data.
- 8. File access in CAATs refers to the capability of:
 - A. Reading of different record formats and file structures.
 - B. Conversion of data to common formats of data
 - C. Using the import/ODBC function.
 - D. Reviewing access controls of application software.
- 9. File reorganization feature of CAATs refers to:
 - A. Functions provide auditor with an instant view of the data from different perspectives.
 - B. Indexing, sorting, merging, linking with other identified files.
 - C. Using of global filter conditions to select required data based on specified criteria.
 - D. Useof sampling, stratification and frequency analysis.
- 10. Which of the following statement pertaining precautions in using CAATs is incorrect?
 - A. Collect the relevant and correct data files.
 - B. Identify all the important fields that need to be accessed.
 - C. Ensure the data represents audit universe correctly and completely.
 - D. Information provided by CAATs is clear indicator of problems and no further testing is required.

- 11. Which of the following is the first step in using CAATs?
 - A. Identify the sources of data from the enterprise information system/application software.
 - B. Based on scope and objectives of audit, decide about the need and extent to which CAAT could be used.
 - C. Identify the critical data which is being audited as per audit scope and objectives.
 - D. Identify the relevant personnel responsible for the data and information system.
- 12. Which of the following is the final step in using CAATs?
 - A. Obtain and review documents relating to data/information systems.
 - B. Decide what techniques of CAATs could be used as relevant to the environment
 - C. Prepare a detailed plan for analyzing the data.
 - D. Perform relevant tests on audit data as required and prepare audit findings and include in audit report/opinion.

ANSWERS

1.	А	2. B	3. D	4. C
5.	С	6. C	7. D	8. A
9.	В	10. D	11. B	12. D

REFERENCES

Below are sample list of references:

www.icai.org

www.isaca.org

www.auditnet.org

www.caseware-idea.org

www.acl.com

www.theiia.org



ADVANCED ANALYSIS USING IDEA

LEARNING OBJECTIVES

- Using advance extraction options for performing complex data analytics
- Understanding use of Benford's Law in audit
- Understanding advance @ functions
- Learning Macro and IDEA Script

4.1 EXTRACTING SELECTED RECORDS

IDEA offers various options under "Extract Group" to select and extract records from databases and perform further analysis and/or to save the extracted records as a separate database.

Following options are provided in IDEA to perform extraction under analysis menu:

1. Direct

A P

- 2. Key Value
- 3. Top Records
- 4. Indexed
- 5. Duplicate Key Detection
- 6. Gap Detection

4.1.1 DIRECT

Direct function is used to extract records by writing one more criteria's and allows to save result in new database. With direct function we can provide one or more criteria to extract records.

Step by step procedure for using direct function

Open database 'sales register'. Total records in the file are 999 and extract records having INV_AMOUNT more than 10000.



Step 1: Analysis > Direct

Home Data Ar	ds 🗐 Ga	iew Mac	ros SmartAni	lyzer ∄Σ Summarization	🗑 Aging	물= Join ····································	ector Attribu	ute 🛅 Monetary Unit 👻		
n 晋+ Indexed 原 Key Value	56 Ber	ford's Law	Statistics *	Stratification	Pivot Table	- Append I Compare	≣⊒ Rando	m 💾 Variables 👻		
Entract		Evenie		Salare Catao	anima.	Relate	Current Courser	Campio .		
CAUGUU	la mort	Capit		Catego	Juize	Nelade		Sample		1
xp = Direct	4	SAL	ES95 × US	ales Register ×			1		* ^	Properties
Create a new database		INV	NO INVOICE	_DATE REGION	PRODUCT	CUSTOMER_TYPE	EXEC_NAME	INV_AMOUNT	^	II 🕶 Database
meet the criteria from the	at	1	1 01-04-201-	4 South	Music System	s Foreign	Shahrukh	43,430.00		
op F current database.		2	2 01-04-201-	4 West	Refrigrator	Domestic	Aamir	7,553.00		J Data
e L	Bs	3	3 01-04-201-	4 West	Music System	s Government	Salman	5,724.00		History
Account Details	1,000	4	4 02-04-201	4 South	Air Condition	er Domestic	Shahrukh	72,712.00		Field Statistics
Customer Details	1,000	5	5 02-04-201	4 East	Mobile Phone	s Foreign	Shahrukh	58,801.00		Control Total
Loan master	563	0	6 03-04-201-	4 West	Refrigrator	Foreign	Salman	58,549.00		Criteria
Master Data-Emp	133	1	7 03-04-201-	4 East	Air Condition	er Foreign	Shahrukh	42,443.00		
Master Data-Sup	363	8	8 03-04-201-	4 East	Refrigrator	Government	Rithik	25,252.00		Results
Payment-Payme	185	9	9 03-04-201	4 West	Mobile Phone	s Domestic	Shahrukh	49,469.00		
Payroll_March 20	189	10	10 04-04-201-	4 South	Music System	s Domestic	Aamir	32,838.00		II 🖛 Indices
Sales Register	999	11	11 04-04-201	4 North	Music System	s Domestic	Shahrukh	41,647.00		
SALES95	101	12	12 06-04-201	+ North	Music System	s Domestic	Saiman	32,200.00		Vo index
User_web_log-Da	200	13	13 06-04-201	4 West	Air Condition	er Foreign	Rithik	72,524.00		Comments
Vehical Claims-Cl	39	14	14 07-04-201	e Cast	Kerngrator	Domestic	Aamir	24,420.00		- comments
Vehical Policy Ma	166	15	15 00-04-201	4 North	Music System	s Foreign	Aamir	49,000.00		Add comment
		10	17 00 04 201	+ East	Mobile Phone	S Government	Chabadda	62,178.00		
		17	10 00 04 201	4 South	Alla Canaditian	s Foreign	Shahrukh	09,479.00		
		10	10 00 04 201	south	Air Condition	er Domestic	Shanrukn	5,904.00		
		15	19 09-04-201	+ East	Talaulalas Cat	s Foreign	Diabili	5,479.00		
		21	21 00-04-201	1 Morth	Mobile Phone	Eoreign	Annair	59 724 00		
		22	22 00-04-201	1 North	Air Condition	Foreign	Rithile	91 100 00		
		23	23 10-04-201	4 West	Air Condition	er Foreign	Aamir	12 532 00		
		24	24 10-04-201	4 Fast	Mobile Phone	s Government	Rithik	1.831.00		
		25	25 11-04-201	A North	Music System	s Foreign	Aamir	60,207,00		
	>	26	26 11-04-201	4 North	Music System	s Government	Rithik	75 995 00		
e Explorer Library		27	27 11-04-201	4 West	Air Condition	Government	Rithik	7 739 00		
		1			1	1	1	1		
nning Tasks										
ged Project: Adv. ITT							N	lot connected to IDEA Server	Number of	Records: 999 Disk Space: 25

Fig. 4. 1.1.1 – Direct option. Number of records 999

Step 1: Analysis > Direct Extract

Step 2: Records to extract:

- All select this option to apply criteria on entire databases
- Range select this option to apply criteria on range of records and not on entire database

Step 3: Database order: Criteria can be defined on indexed or sorted database by using this option. Step 4: File Name: Give new file name

● ♪ ⊕17123		jo v					Ca	seWare IDEA -	Sales Regist	ter			- 8 - 2
Hone Data	Analys	is Vie	541	Macros	SnertAnaly	161							۵
nun 2ª indexed R.Ke	p Records ry Value	ini Gap I Si Darfe	Detectio profis Lar	n 🖾 Dupk n 🚔 State	cate Key * atics *	Stratificati Chart	tion 🔃 Aging on 🔝 Pivot Table	∄e Join ≪ V 8-s Append II C	tsuel Connector Iompere	Random	Monetary Unit *		
sks Extract			E	Explore		Ca	rtegorize	Rela	te		Sample		,
e Explorer		÷.,	/0	SALES95	× 🗑 Sak	es Register 🗦	<u> </u>					- ×	Properties
四月公司	1		1	NV_NO 1 (INVOICE_0 01-04-2014	DATE REG South	ION PRODUCT Music System	CUSTOMER s Foreign	TYPE EX	EC_NAME	INV_AMOLINT 43,430.00	^	I * Detabase
ktop Project	Bac	ande -	2	2 0	01-04-2014	West	Refrigrator Music System	Domestic Government	Aam	ir an	7,553.00		Data Dista
Account Details	HEC 1	,000	4				Dire	t Extraction			×		Field Statistics
Customer Details Loan master	1	563	6	Records t	to extracti (€.al	Starting record #1	1			OK		Control Total
Master Data-Errip Master Data-Sup		133 363	7		0	Bange	Ending record #1	999			Greate Pields		I * Results
Payment-Payme		185	9			Create a virt	ual database				Delds		Las hadras
Sales Register		999		Detabase	grder: No.i	ndex				,	Delete		
User_web_log-Da		200		1. Cala	- Ab 100	File Nar	THE	22	Criteria		Canoel		I V Comments
Vehical Claims-CL.		30	14	1 5816	s Above 100	00					Help		Add comment
			16 17 18 19	-									
			21 22 23 24										
			25	25 1	11-04-2014	North	Music System	n Foreign	Aam	ir b	60,207.00		
File Explorer	/		27	20 1	11-04-2014	West	Air Condition	er Government	Rithi	k	7,739.00	~	
Running Testes													
Aaged Project: Adv. ITT	-		-	-	-	_	-			Nat	connected to IDEA Server	Number of	Records: 999 Disk Space: 250.25
02 0			(目)	e	w		25 1	0.					 No til at te av na 1949

Fig. 4.1.1.2 – Give file name

Step 5: Click on equation editor icon to define the criteria

Step 6: Define the criteria on equation editor screen INV_AMOUNT > 10000 and click on validate and exit option or press Ctrl+U





Hone Data Analysis Fine Pore Transformer Data Fine Pore	View Macros SmartA Gap Detection 🔐 Duplicate Key	Nalyzer - BC Summarization (1) Aging - Stratification (1) Pivot Tal	CaseWare IDEA - Sales Reg	ster ₩ Attribute @ Monetary Unit * #= Random @ Variables *			- 8 🔺
Tasks Extract			Equation Editor				
File Explorer	 ✓ I I I I 	* © 6 0 0 0 0	A SODE			-0	Properties
Retros Pariet Accord Data Catomer Data Catomer Data Meter Data Fryn, Meter Data Syn, Popret Retro 2. Popret Retro 2. Utan meter Popret Retro 2. Popret Retro 2. Visical Catomer CL. Visical Catomer March 2. Visical	***. (1) * ** > <	* >* c> * - * / AND OR	NOT MOD EUF	⊕ Al Character ⊕ Data/Time ⊕ Data/Time ⊕ Data/Time ⊕ Data/Time ⊕ Data/Time ⊕ Data/Time ⊕ Data/Time ⊕ Conditional ⊕ Condition	GFunctions gFunctions are available for performing note complex operations and so date arithmetic, statistical calutations and the cardines. IREA functions large with the variant in state and coesition to functions function on the and performed the cardinal performant functions and and the arithmetic performant functions and and the arithmetic and and the arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arithmetic arith	^	P Databane Databane Databane Databane Databane Hatany Hatany Hatany Hatany Hatany Control Control V facults T ▼ datases Add comment
	Field INV_IND INVOIDE_DATE PEGEN PRODUCT CUSTONER_TYPE	Type Nurvesio Date Chasoter Chasoter Chasoter Chasoter Chasoter Chasoter	Len Dec A	Insert Function	permeters passed to fire @Function. Provincitors ere passed in percentineses. For essample, @Recard (Amount) will round a field called "Amount" to the meaners whole number. Additional balance	•	
<	25 25 11-04-25	14 North Music Sys 14 West Air Condi	tioner Government Ri	thik 7,739.00			
Managed Project: Adv. IIT	×1 e «			Not connected to IDE	A, Server	Number o	(Records: Disk Space: 250.25 G





Fig. 4.1.1.4 – *Save file*

Step 7: Click on OK

21 W	シャビが設備部	80 7					CaseWa	re IDEA - Sales Above	10000			- C ×
Tile	Hone Data Anal	sis Viev		facros	SnartAnalyzer							△ (0
Rerun	B ^{er} Direct B ^{er} Top Records	Eap De	etection of a Law	Duplos	te Key - 🕅	Stratification 🛐	Aging 2. Pivot Table 8-	Join 🦷 Visual Connecto Append 🖫 Company	Random	Monetary Unit *		
Tasks	Extract		Ð	plore	-	Categorize		Relate	5	anple		
File Ext	plorer	+	1	🗿 Sales Re	gister 🕺 🕡	Sales Above 100	00 ×				+ X	Properties #
and a		_		INV_NO	IN/OICE_	DATE REGION	PRODUCT	CUSTOWER_TYPE	EXEC_NAME	INV_AMOUNT	^	T Database
			1	1	01-04-2014	South	Music Systeme	; Foreign	Shahrukh	43,430.00		
Desktop	Project		2	- 4	02-04-2014	South	Air Conditione	s Domestic	Shahrukh	72,712.00		🖌 Data
Name		Records	3	5	02-04-2014	East	Mobile Phone	s Foreign	Shahrukh	58,801.00		History
-80	Account Details	1,000	4	6	03-04-2014	West	Refrigrator	Foreign	Salman	58,549.00		Field Statistics
-0	Customer Details	1,000	5	7	03-04-2014	East	Air Conditione	r Foreign	Shahrukh	42,443.00		Control Total
-0	Loan master	563	6	8	03-04-2014	East	Refrigrator	Government	Rithik	25,252.00		iii Criteria
-0	Master Data-Employ	133	7	9	03-04-2014	West	Mobile Phone	s Domestic	Shahrukh	49,469.00		
-0	Master Data-Supplier	363	8	10	04-04-2014	South	Music Systeme	Domestic	Aamir	32,838.00		I ▼ Results
-0	Payment-Payments	185	9	11	04-04-2014	North	Music Systeme	Domestic	Shahrukh	41,647.00		
-0	Payroll, March 2015	189	10	12	06-04-2014	North	Music Systeme	Domestic	Salman	32,288.00		I ≠ Indices
ė-🗊 (Sales Register	999	11	13	06-04-2014	West	Air Conditione	sr Foreign	Rithik	72,524.00		
14	Sales Above 10000	396	12	14	07-04-2014	East	Refrigrator	Domestic	Azmir	24,420.00		No index
-0	SALESUS	101	13	15	08-04-2014	North	Music Systeme	5 Foreign	Aamir	49,808.00		1-0
-0	User_web_log-Datab	200	14	16	09-04-2014	East	Mobile Phone	s Government	Azmir	62,178.00		1 * Comments
-0	Vehical Claims-Claim	39	15	17	09-04-2014	South	Music Systeme	5 Foreign	Shahrukh	69,479.00		Add comment
-0	Vehical Policy Manter	166	15	18	09-04-2014	South	Air Conditione	s Domestic	Shahrukh	83,904.00		
			17	20	09-04-2014	North	Television Sets	Foreign	Rithik	68,807.00		
			18	21	09-04-2014	West	Mobile Phone	s Foreign	Aamir	58,734.00		
			19	22	09-04-2014	North	Air Conditione	sr Foreign	Rithik	81,190.00		
			20	23	10-04-2014	West	Air Conditione	s Foreign	Aamir	12,532.00		
			21	25	11-04-2014	North	Music Systeme	; Foreign	Aamir	60,207.00		
			22	25	11-04-2014	North	Music Systems	Government	Rithik	75,996.00		
			23	28	11-04-2014	North	Air Conditione	r Government	Shahrukh	23,513.00		
			24	29	12-04-2014	South	Music System	i Foreign	Shahrukh	21,114.00		
c			25	30	12-04-2014	West	Television Sets	Domestic	Aamir	77,260.00		
111.614	Embran (1712)	-	26	31	15-04-2014	North	Music System	Upvernment	Aamer	43,733.00		
7.3404	rehour E round		27	32	13-04-2014	West	Air Conditione	r Domestic	Rithik	60,985.00	¥	
45 Run	ning Tasks											
Manage	el Project: Adv. ITT								Not c	onnected to IDEA Servi	r Number of	Records: 896 Disk Space: 250.25 GB
	02 9 8	a x	1	e	wB	3	Q.					- No til ant 4a 13.03 - No til ant 4a 04-05-2017





Step 8: A new file meeting the criteria is created with 896 records out of 999 having invoice amount more than 10000

Note: Any Criteria defined in the equation editor can be saved by using Ctrt+S and it will save in Equations.ILB folder of the active project. Same can be reused as it is or edited for further references.

Exercises: 1. Extract records of sales of more than or equal to 15000 in East region

Exercises: 2. Extract records of mobile phones sold to foreign customer

4.1.2 KEY VALUE

Helps to extract records using values in a key. This function is used to extract records for a key in the entire database in a separate file. Key can be a region, product, zone, customer type, etc.

Step by step procedure for using Key Value function

Open database 'sales register'. Total records in the file are 999.

Extract records based on key being 'Product'

Step 1: Analysis > key Value

BION OF BUILDE	CaseWare IDEA - Sales Register	- 0 .
File Hone Data Analysis V	en Marros SmartAnalyzer	≏ 6
Refun 2-Indexed Ray Value In Der	Derivation Bostination Barged 3 + 3 m Nonal Convertory Mathematical Astronomy forfs Law Bostination Theat Table Bostination Bostination Bostination forfs Law Bostination Bostination Bostination Bostination Bostination	
Tasks Extract	Explore Categorize Relate Sample	
File Explorer	a 🕼 Sales Register X 🔹 🗸	Properties 9
自ちたちは間	NV_NO INVOICE_DATE REGION PRODUCT CUSTOMER_TYPE EXEC_NAME INV_AMOUNT	▼ Detabase
Retrief State Recent Recent Recent Conternet Detain Recent Conternet Detain Software Soft	Key Value Extraction Name 2 Selant Roard Grage 72,72,00 4 Delay layer Coccel 93,810,00 6 Others 94,810,00 72,72,00 7 Others 94,810,00 72,82,00 8 Grage Found 94,810,00 94,810,00 9 Grage Found 72,82,00 74,840,00 9 Grage Found 72,810,00 72,810,00 10 Image 24,850,00 94,840,00 10 Image 72,810,00 94,840,00 10 Image 74,840,00 94,840,00 <	Extra Hatagy Ental Standing: Cranol Total Cranol Total Cranol Mades Comments Add comment
٤	26. Defix for each database: 75.006.00	
En File Explorer Explorer	27 Key/al 7,738.00 v	
Managed Project: Adv. ITT	Create a sparate database for each unique key Oraste a gatad database Number of Re Number of Re	cords: 999 Disk Spece: 250.25 GB
🖬 🚥 🧿 🚔	xi 🤗 🖬 🔝 🛷 🔜 🔹 .	No. R. and No. 04-06-2017

Fig. 4.1.2.1 Using Key Value

Step 2: Click on ... besides 'Existing keys:' option





310 A 8 2 2 2 3 3 3 3 4 4	CaseWare IDEA - Sale	s Register	- 8 ×
Pore Data Analysis Wen M Pore I Top Records III Cap Detactor Rerun III - Statewood III, Vary Value Detact	teors Shertifaliyer Shertifaliyer Statester Statester Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifaliyer Shertifa	annectar MAttribute Annetary Unit *	∆ (J
File Explorer #	Sales Register X	Sampe:	• × Properties •
Rest Process P	RE_IND INVOCE_DATE REACON PRODUCT CLATENER_I Key Value Extraction States from Grave Define Key Define Key Grave Grave Define Key Field Direction Field Fi	PRE DEC_1UNE PP/_ADD.047 6.8,600 1,513.00 OK 1,513.00 OK 1,513.00 Covel 3,714.00 Define 2,814.00 Pete 2,814.00 Occord 1,714.00 1,714.00 1,111.00	I ▼ Outbase II ▼ Outbase II → Outbase II → Outbase II → Outbase II ▼ Outbase
< 24 25 26 27 26 27 26 27 26 27 26 27 26 27 26 27 26 27 26 27 26 27 27 27 27 27 27 27 27 27 27	Output Byefin for each database: Keyfid	1,831.00 60,207.00 75,996.00 7,739.00	v
Managad Project Adz. ITT	Create a grand database	Not connected to IDEA Server	Number of Records: 999 Disk Space: 250.25 GB * Pa R at 4 00 04-05-2917

Fig. 4.1.2.2 Existing Keys

Step 3: Select product from the drop down option and click on OK

Step 4: Idea will extract all the key values and will pop up under group found section

910	A ≤ F ≥ 2 = 1	3					CaseV	Vare IDEA - Sa	les Register				- 8 ×
File	Hone Data Anal	ysis View	y	iacros SmartAnal	1261								_ Ø
Re-run	Porect Top Records	Gap Det	fa Lava	Duplicate Key *	Strati	fication ()	Aging Pivot Table	Join Comp	Connector	ittribute landori lither *	Monetary Unit *		
10945	EXUBL	-	0	pore	-	Calegorite		REDUC			sarpe		1
File Exp	olorer	+	4	Sales Register >	1							* X	Properties *
1 m F	2000		-	INV_NO INVOIC	E_DATE	REGION	PRODUCT	CUSTOMER	TYPE EXEC		E INV_AMOLINT	^	I ♥ Detabase
Desiton	Project	_				Key Va	alue Extractio	nc	×		43,430.00		E Data
Name		Records		Select Record Gr	0.471						5.724.00		E History
	Account Datails	1000	4	Existing keys:					OK		72,712.00		Eeld Statistics
1.0	Customer Details	1.000	5	PRODUCT	x.			¥	Cancel		58,801.00		Control Total
-ŭ	pan master	563	6	Oriteria					Fields		58,549.00		Criteria
-01	Master Data-Employ	133	7								42,443.00		
-01	Manter Data-Supplier	363	8						Help		25,252,00		▼ Results
-9	Payment-Payments	185		Groups Found							42,405.00		
1.22	Payroll_March 2015	189		Disass calact if	a color c	no seith in anti-	and a		Clear &I		41.647.00		I ▼ Indices
	SALESOS	101		There are 5 un	que key ve	elues in the clet	tebene.		Tookuda All		32,288.00		
1.0	Jaer web Inn-Datab	200							frome w		72,524.00		I ▼ Comments
-0	Vehical Claims-Claim	39	14	Air Conditio	THE .						24,420.00		Add comment
1-01	Vehical Policy Master	166	15	Munic Sector	NEZ						49,808.00		
			16	Refigrator							62,178.00		
			17	Television 5	ietz						05,475.00		
			10								5.479.00		
											68,807.00		
											58,734.00		
											81,190.00		
											12,532.00		
			24	0 mm							1,831.00		
۲.		>		Prefix for each	clatabase:						60,207.00		
(C) File I	Explorer Library			Product w	sel						7.739.00		
(MKa				of Create	cacacata :	databasa fir a	arth union ta kaur					*	
(C) Russ	ang sasa ()			Create	vetual dat	uevauese ror e tabara	on and a set						
Manage	d Project: Adv. ITT	-					-	_	_	Na	t connected to IDEA Server	Number of	Records: 999 Disk Space: 250,25 GB
	02 0 1	X		8 1	-	3	_			1			^ Na II-al 4a 04.05-2017

Fig. 4.1.2.3 Key values displayed

Step 5: Under output section, provide file name under 'prefix for each database'

Step 6: Click on create a separate database for each unique key, will extract all the records for each unique key in a separate database

Step 7: Use criteria tab to define condition for extraction of records key wise

Advanced Analysis Using IDEA



110	シャッドを設備にな	- 7		(aseWare I	DEA - Product	wise=Air Conditi	ioner			- 0
tie	Hone Data Analysis	View Macros	SnartAn	al y zer							
3	B ^e Drect B ^e Top Records	Gap Detection 👜 Du	picate Key	Scientification	Aging Runt Table	an Join 🗠 Ve	uel Connector	ttribute 💼 Monetary U	nit *		
run	콜* Indexed 릤 Key Value	🗟 Benford's Law 👜 St	atatics *	Chart	PINOL 10090	통e Append II Co	mpare 420	ther .			
sis	Extract	Explore		Categorize		Relate		Sample			
e Exp	plorer		/ 🗑 54	les Register 🕺 🗊 Pro	duct.wise=A	ir Conditioner X				* ×	Properties
			INV	NO INVOICE_DATE	REGION	PRODUCT	CUSTOMER_TYP	PE EXEC_NAME	IN/_AMOUNT	^	I w Database
18	日本での目		1	4 02-04-2014	South	Air Conditioner	Domestic	Shahrukh	72,712.00		
itop	Project		2	7 03-04-2014	East	Air Conditioner	Foreign	Shahrukh	42,443.00		🖌 Data
me	~	Records	3	13 06-04-2014	West	Air Conditioner	Foreign	Rithik	72,524.00		History
ai.	Account Details	1.000	4	18 09-04-2014	South	Air Conditioner	Domestic	Shahrukh	83,904.00		Field Statistics
ă.	Customer Details	1.000	5	22 09-04-2014	North	Air Conditioner	Foreign	Rithik	81,190.00		E Control Total
ň	nan master	563	6	23 10-04-2014	West	Air Conditioner	Foreign	Aamir	12,532.00		E Citada
ň.	Master Data, Frenkrune Master	133	7	27 11-04-2014	West	Air Conditioner	Government	Fithik:	7,739.00		u citeta
ñ.	Manter Data-Supplier Manter	363	8	28 11-04-2014	North	Air Conditioner	Government	Shahrukh	23,513.00		1 * Results
ň.	Payment-Payments	185	9	32 13-04-2014	West	Air Conditioner	Domestic	Fithik	60,985.00		
ň.	Percell March 2015-may	180	10	33 13-04-2014	West	Air Conditioner	Foreign	Salman	3,008.00		L m badaas
ñ.	Salan Barrinter	000	11	46 16-04-2014	North	Air Conditioner	Domestic	Salman	24,001.00		· · · marces
Ti	Product wise: Air Conditione	185	12	53 18-04-2014	West	Air Conditioner	Domestic	Shahrukh	35,751.00		No index
13	Product wise - Mohile Phone	107	13	58 21-04-2014	West	Air Conditioner	Domestic	Fäthik	33,217.00		
12	Product wine-Munic System	200	14	59 21-04-2014	East	Air Conditioner	Government	Aamir	Z3,494.00		I V Comments
Li	Product wise: Refrigrator	199	15	74 29-04-2014	South	Air Conditioner	Government	Aamir	65,513.00		E All connect
Li	Product wise: Television Set	218	16	81 30-04-2014	West	Air Conditioner	Foreign	Shahrukh	60,892.00		a saucomment
Li	Salas Abrus 10000	816	17	87 05-05-2014	North	Air Conditioner	Foreign	Salman	49,661.00		
n)	SALESUS	101	18	89 05-05-2014	East	Air Conditioner	Government	Rithik	66,581.00		
ň.	User web log-Database	200	19	90 05-05-2014	North	Air Conditioner	Government	Shahrukh	6,006.00		
ň.	Vehical Claims-Claim	39	20	93 08-05-2014	North	Air Conditioner	Government	Rithik	33,933.00		
ň.	Vehical Policy Master	166	21	103 11-05-2014	South	Air Conditioner	Domestic	Fithik	34,050.00		
-	,		22	113 17-05-2014	North	Air Conditioner	Domestic	Aamir	74,343.00		
			23	114 17-05-2014	South	Air Conditioner	Domestic	Salman	52,277.00		
			24	115 18-05-2014	North	Air Conditioner	Government	Shahrukh	73,687.00		
			25	118 19-05-2014	West	Air Conditioner	Government	Shahrukh	52,703.00		
		>	26	122 20-05-2014	West	Air Conditioner	Government	Salman	39,347.00		
File	Explorer		27	123 21-05-2014	West	Air Conditioner	Government	Aamir	75,296.00	v	
Run	ning Taska										
nage	d Project: Adv. ITT		_					Not connected to IC	CEA Server	Number of	Records: 185 Disk Space: 250
	02 🕥 🚞	× e	w								• No. 11 at 4a

Fig. 4.1.2.4 Key wise databases created

Key value wise databases (product wise in our example) created and saved under master file sales register

Exercises: 1. Extract records based on key value region for south and east region

Exercises: 2. Extract records based on EXEC_NAME for INV_AMOUNT exceeding 25000

4.1.3 TOP RECORDS:

Helps to extract specified number of top or bottom records for each value of key in a separate database.

Step by step procedure for using Top Records function

Open database 'sales register'. Total records in the file are 999.

Extract 3 top records INV_AMOUNT wise for each of the customer type

Step 1: Analysis > Top Records

910	*****				Ca	seWare IDEA - S	ales Regist	ler				- 8 ×	
Tie	Hone Data Analys	s View Macros	SnartAnal	/26/			-					۵ (Ð
Rerun	문 [#] Direct 문 [#] Top Records 같~ Indexed 곳, Key Value	Gap Detection Date	picate Key * distics *	Stratification	Aging Pivot Table	Se Join ≪ Vas Se Append II Cor	al Connector Inparie	Attribute	Monetary Unit * Variables *				
Tasks	Extract	Explore	_	Categorize		Relate		San	ple			,	_
File Exp	plorer	Ŧ	Sale	is Register X							+ ×	Properties	4
	12P0		1 IN/_	NO INVOICE_DATE	Tap R	ecords Extraction	CUSTOME	R_TYPE EXE	C_NAME IN	(_AMOLINT 43,430.00	î	I v Detabase	-
Name	Account Datain Continue Datain Loom mother Loom mother Loom mother March Data Signaphe Manter Payment Physical District States Register States Register University Physical District China States University Physical District Valued Datases China Valued China Valued China Valued China Valued China Valued	Records 1,000 563 163 163 165 165 166 906 30 166	1 Num 3 Num 6 Tys 6 Tys 7 Toc 8 Gar 9 0 111 12 13 14 16 16 17 18 20 Cor 21 Num 22 22 23 Cor	nhar of records to softwart ke precords for sap (pertonal): Field SATOMER_TYPE arrise: Typ 3 trivials as what a what a debaase	Top record	rds unnt v Ascending	v	OK Felds Delate Peld Canol Help	n b b b b b	3, 724,000 72, 772, 772, 772, 000 83, 849,000 42, 443,000 82, 888,000 82, 888,000 82, 888,000 82, 888,000 84, 869,000 84, 789,000 86, 779,000 88, 394,000 88, 394,000 81, 190,000 81, 190,0000 81, 190,0000 81, 190,0000 81, 190,0000 81, 190,000 81,0		History History History History Control Total Control Total Totatis Totatis Histoc T Histoc T Mate comment Add comment	
٤		>	26	26 11-04-2014	North	Music Systems	Government	Rithil		75,996.00			
File	Explorer Library		27	27 11-04-2014	West	Air Conditioner	Government	Rithil	t i	7,739.00	~		
25 Run	ning Tasks												1
Managa	d Project: Adv. ITT	Chapter 4 [Competibility	y Mode] - Word				Nat car	nected to IDEA S	erver	Number of	Records: 999 Disk Space: 250.24 G	в
	02 9 🛢	I XI e	wE	3								- Na Ri⊶al ¶a 04-06-2017	

Fig. 4.1.3.1 – Top records Extraction





Step 2: Number of records to extract - type 3

Step 3: Type: select Top records. Selecting bottom records option here will extract bottom records for the given conditions

Step 4: Top Records for: select INV_AMOUNT

Step 5: Under Group (optional): select Customer Type

Step 6: File name - type Top 3 invoices customer wise

Step 6: Enter condition to apply the function on selected type of records

ē i i	マシットを読む	o -			CaseWare	IDEA - Top 3	invoices cust	omer v	vise			-	۵ ×
Tie	Hone Data Analys	is View Macros	SnertA	nalvær									- G
Re-run	문 Direct 문 Top Records 같~ Indexed , Key Value	Gap Detection Course Du	picate Key datice *	Summarization Stratification Chart Category	😧 Aging 🚺 Pivot Table	글= Join 예를 용= Append 물	Vauel Connector Compere	왕에 Att	nbute 📑 Monetan ndom 📑 Variables her *	y Unit * .*			
10005	EXUBLI	- Copyre	/@	der Banister V	ton 3 incoloor	ourtemer wire	v.		sanpic		- ×		
	Accurate Datas Counterne Datas Counterne Datas Learn mater Mater Data-Scapitier Matter Mater Data-Scapitier Matter Mater Data-Scapitier Matter Mater Data-Scapitier Matter Mater Datas Mater Datas Mater Datas Mater Datas Unice Scapitier Matter State Repaire Trap 3 motes customer on SALESS Using Datas Unice Clarene-Claim Uniced Policy Mester	Records 1,000 563 133 185 185 185 185 185 185 185 195 195 195 195 195 195 195 195 195 19	11/1/2 3 4 5 6 7 8 8 9	 IPV/CRE_DAT 2707.2014 2707.2014 2708.10-2.2014 2708.10-2.2014 270.2015 280.00-2014 281.10-2014 383.14.02-2013 39.15-00-2014 484.35-11-2014 445.05-09-2014 	REGION West South West West West West West West	PRODUCT Music Systems Music Systems Refrigator Televisian Set Mobile Phones Music Systems Air Conditioner Mobile Phones Mobile Phones	CLISTOMER, Dometric Dometric Poreign Foreign Foreign Foreign Government Government Government	TYPE	EXEC JAME Salman Shahrudh Shahrudh Shahrudh Aanrir Salman Salman Salman Shahrudh	1072,4002,087 94,53500 94,45500 94,45500 94,45500 94,45500 94,45500 94,45500 94,45500 94,45800 94,45800			
< ট্রেনা	e Explorer 🛄 Library	3											
an a	inning Tasks												
Manag	ged Project: Adv. ITT			-					Not connected	to IDEA Server	Number	of Records: 9 Disk Spec	:#: 250.24 GB
	02 0 🗎	I XI e	w		3							• Ne Rivel 4a j	17.07 H-06-2017

Fig. 4.1.3.2 – Top records Extraction customer type wise

Exercise 1: extract bottom 5 invoices for each executive

Exercise 2: extract top 3 invoices above Rs15,500/- for each region

4.1.4 INDEXED

Uses an existing index to narrow down the scope of extraction. This function can be used along with criteria to get better results. It combines the capabilities of criteria and range extraction.

Step by step procedure for using indexed function

Open database 'sales register'. Total records in the file are 999.

Extract records of INV_AMOUNT >= 15000 having invoice date between 3^{nd} April to 10^{th} April, 2014

Step 1: Analysis > Indexed

Step 2: Field- select INVOICE_DATE

Step 3: Value is - >= 20140403 (YYYYMMDD)

Step 4: and - <= 20140410

Step 5: Criteria – INV_AMOUNT > 15000

Advanced Analysis Using IDEA

Step 6: File Name - Invoices above 15K between 3 to 10 April. 16 records extracted as an output



Fig: 4.1.4.1 –	Index	function	options

ହାଙ	マルマを設備部に	jo -			CaseWa	are IDEA -	Invoices abov	e 15K between 3	to 10 April			- 0 ×
tie	Hone Data Analys	is View Macros	Sin	artánalyze	r							≏ 0
ß	Porect Prop Records	Gap Detection	picate	5ay • 8	E Summarization 👿	Aging Pivot Table	ge Join 🖉 V	isual Connector	Attribute 📑 Honetary Randon 📑 Variables '	Unit *		
Renun	🖀 Indexed 💻 Key Value	🔠 Benford's Law 👜 Sh	statics	. 6	Chert		용e Append 물 C	orpers 12	Other =			
Tasks	Extract.	Explore			Categorize		Rela	ie –	Sample			
File Exp	plarer	+	1	🕽 Sales R	egister 🕺 🗑 lav	oices above	15Kbetwee	×			÷×	Properties #
and a				NV_NO	INVOICE_DATE	REGION	PRODUCT	CUSTOMER_TYP	E EXEC_NAME	INV_AMOUNT		T Detabase
= :			1	6	03-04-2014	West	Refrigrator	Foreign	Salman	58,549.00		
Desktop	Project		2	7	03-04-2014	East	Air Conditioner	Foreign	Shahrukh	42,443.00		🚽 Data
Name		Records	3	8	03-04-2014	East	Refrigrator	Government	Rithik	25,252.00		History
-80	Account Details	1,000	4	9	03-04-2014	West	Mobile Phones	Domestic	Shahrukh	49,409.00		Field Statistics
-0	Customer Details	1,000	5	10	04-04-2014	South	Music Systems	Domestic	Aamir	32,838.00		Control Total
-0	Loan master	563	6	11	04-04-2014	North	Music Systems	Domestic	Shahrukh	41,647.00		Criteria
-0	Master Data-Employee Master	133	7	12	06-04-2014	North	Music Systems	Domestic	Salman	32,288.00		
-0	Master Data-Supplier Master	363	8	13	06-04-2014	West	Air Conditioner	Foreign	Rithik	72,524.00		I * Results
-0	Payment-Payments	185	9	14	07-04-2014	East	Refrigrator	Domestic	Aamir	24,420.00		
-0	Payroll, March 2015-pay	189	10	15	08-04-2014	North	Music Systems	Foreign	Aamir	49,808.00		↓ ▶ Indices
0.00	Sales Register	999	11	16	09-04-2014	East	Mobile Phones	Government	Aamir	62,178.00		
14	Invoices above 15K between	n 3 16	12	17	09-04-2014	South	Music Systems	Foreign	Shahrukh	09,479.00		V Comments
-0	SALES05	101	13	18	09-04-2014	South	Air Conditioner	Domestic	Shahrukh	83,904.00		E Add comment
-0	User_web_log-Database	200	14	20	09-04-2014	North	Television Sets	Foreign	Rithik	68,807.00		
-0	Vehical Claims-Claim	39	15	21	09-04-2014	West	Mobile Phones	Foreign	Aamir	58,734.00		
-0	Vehical Policy Master	166	16	22	09-04-2014	North	Air Conditioner	Foreign	Rithik	81,190.00		
< L]File	Explorer (Library /											
25 Run	ning Tasks											
Managa	d Project: Adv. ITT								Not connected to	IDEA Server	Number of	Records: 16 Disk Space: 250.18 GB
	os 🕥 🚞	I XI e		wB								• Na R -41 4a - 17126

Fig: 4.1.4.2 – Output of indexed function

Exercise 1: Extract invoices for west region having invoice value between 8000 to 48000

Exercise 2: Extract invoices for product refrigerator having invoice date between 21st March to 28th March 2014

4.1.5 DUPLICATE KEY

IDEA includes two tasks for finding duplicate records:

Duplicate Key Detection: Identifies all records where the fields in the key are identical.

Duplicate Key Exclusion: Similar to Duplicate Key Detection, except that an additional field must be different.

The Duplicate Key Detection task is used to identify duplicate items in a database. Duplicate Key Detection can be used to identify duplicates within a single field, such as duplicate invoice numbers, or duplicates in a combination of up to eight fields.

Duplicate Key Detection can also be used to identify non-duplicate keys (unique) by selecting the **Output Records without duplicates** option instead of the default Output Duplicate Records option.





The Duplicate Key Exclusion task is used to identify duplicate items in a database where a further specified field is different, such as duplicate purchase order numbers for different suppliers. Duplicate Key Exclusion can be used to identify duplicates within a single field, such as duplicate invoice numbers, or duplicates on a combination of up to eight fields.

Step by step procedure for Duplicate Key Detection

The director of accounting is worried that several invoices may have been issued on multiple occasions. Verify if there have been any multiple invoices issued.

Step 1: Ensure that **Sales Transactions-Database** is the active database and that the **Data** property is selected in the **Properties** window.

Step 2: On the **Analysis** tab, in the **Explore** group, click **Duplicate Key** and then click **Detection**. Alternatively, from the Quick Access Toolbar, click the **Duplicate Key Detection** button.

010		l ∰2 ≩#2	≣₩₽₽	CaseWare IDEA - Sales transactions-Database
File	Home	Data	Detection	SmartAnalyzer

Fig: 4.1.5.1 Duplicate Key Detection dialog box.

Step 3: Click Key. The Define Key dialog box appears.

Step 4: Select INV_DATE in Ascending order and then select CUSTOMER_NO in Ascending order.

	•
Direction	OK
Ascending	Delete Kev
 Ascending 	Delete Key
	Cancel
	Help
	Thep
	Direction Ascending Ascending

Fig 4.1.5.2 Define Key dialog box.

Step 5: Click OK.

Step 6: Accept the default selection of the **Output duplicate records** option.

Step 7: In the File name field, enter Duplicate Customer Sales on Same Day as in Fig 4.1.5.3.

Advanced Analysis Using IDEA



	I.P.I. I	
Output of the	luplicate records	
Output r	ecords without duplicates	Key
Criteria:		Fields
File name:	Duplicate Customer Sales on Same	Cancel
		C



Step 8: Click OK.

The new database, containing all transactions with sales to the same customer more than once on the same day, is created and opened in the Database window.

Step 9: Ensure that **Duplicate Customer Sales on Same Day** is the active database and that the **Data** property is selected in the **Properties** window.

Step 10: On the **Analysis** tab, in the **Tasks** groups, click **Re-run**.

Step 11: In the Duplicate Key Detection dialog box, click **Key**.

Step 12: Add **QTY** as the third key.

Step 13: Click OK.

Step 14: In the Duplicate Key Detection dialog box, enter **Refined Duplicate List** as the file name.

Step 15: Click OK.

The new database of potential duplicates is displayed.

Step by step procedure for searching Non-Duplicate records

While Duplicate Key Detection allows you to search for more than one occurrence of a record, IDEA allows you to extract records that appear only once in the database.

Step 1: Ensure **Sales Transactions-Database** is the active database and that the **Data** property is selected in the **Properties** window.

Step 2: On the **Analysis** tab, in the **Explore** group, click **Duplicate Key** and then click **Detection**. Alternatively, from the Quick Access Toolbar, click the **Detection** button.

Step 3: Select the **Output records without duplicates** option.

Step 4: In the File name field, enter Non-Duplicate Invoices





Duplicate Key	Detection	×
Output o	duplicate records	ОК
Output r	ecords without duplicates	Кеу
Criteria:		Fields
File name:	Non-Duplicate Invoices	Cancel
	Create a virtual database	Help

Fig 4.1.5.4 Select Non-Duplicate records

Step 5: Click Key and ensure INV_NO is selected as your first field in Ascending order.

Step 6: Click OK.

Step 7: In the Duplicate Key Detection dialog box, click OK.

The new Non-Duplicate Invoices database will appear.

4.1.6 Gap Detection

The Gap Detection task identifies "missing" information in a database. The results of the test are displayed in the Results area of the Properties window. The Gap Detection task can be used on a Numeric field to identify gaps in a numerical sequence, such as missing check numbers in a payment database.

The Gap Detection task can be used on a Character field to identify gaps in a numerical sequence that is a sub-set of a Character field. For example, it may be used to identify gaps in an invoice number sequence for each store where the first character of the invoice number identifies the store (for example, A123).

The Gap Detection task can be used on a Date field to identify gaps in a range of dates. It provides the option for excluding weekends and/or holidays, if required. For example, you can identify sales transactions that occurred on non-working days.

Step by step procedure for Gap Detection on Numeric field

Step 1: Ensure that **Sales Transactions-Database** is the active database and that the **Data** property is selected in the **Properties** window.

Step 2: On the **Analysis** tab, in the **Explore** group, click **Gap Detection**. Alternatively, on the Quick Access Toolbar, click the **Gap Detection** button.



Fig 4.1.6.1 *Gap Detection dialog box*

Step 3: From the **Field to use** drop-down list, select **INV_NO**.

Select a range of values to test. By default, all values in the INV_NO field are tested and the minimum and maximum values are listed as starting key value and Ending key value. You can also change



Step 4: Ensure the **Create result** check box is selected.

Step 5: In the **Result name** field, enter **Missing Invoice Sequences**.

ap Detection				l
Field to use:	INV_NO	Criteria:		ОК
Numeric				 Cancel
() All	Starting key value:	1,000,001		Help
🔘 Range	Ending key value:	1,000,977		
	Gap increment:	1		
Output				
🕅 Create da	tabase	📝 Create re	sult	
File name:	Gap Detection	Result name:	Missing Invoice Sequences	
File name:	Sap Detection	Result name:	Missing Invoice Sequences	

Fig 4.1.6.2 Result name.

Step 6: Click OK.

The result will display the ranges of invoice numbers that are missing in the database. Note that each row shows a gap. Certain rows may contain more than one missing invoice number.

\$ \$ + ፼ +		
From: INV_NO	To: INV_NO	Number
⊞ 1,000,098	1,000,098	1
⊞ 1,000,351	1,000,352	2
	Total number of items detected	3
	Total number of gaps detected	2

Fig 4.1.6.3 Result Missing invoice

Step by step procedure for Gap Detection on Date field

Step 1: Ensure that **Sales Transactions-Database** is the active database and that the **Data** property is selected in the **Properties** window.

Step 2: On the **Analysis** tab, in the **Explore** group, click **Gap Detection**.

The Gap Detection dialog box appears.

Step 3: In the Field to use drop-down list, select INV_DATE.





You can select a range of dates to test. By default, all values in the INV_DATE field are tested and the minimum and maximum values are listed as Starting date value and Ending date value. You can also choose to ignore weekends and holidays that are set by the user.

Step 4: Accept the default option to test all values.

Step 5: Select the **Ignore weekends** check box.

Step 6: Select the **Ignore holidays** check box and click **Set Holidays...**.

The Set Holidays dialog box appears.

Step 7: From the Set Holidays toolbar, click the New button.

Step 8: Click the Browse button to select December 25 and December 26 (Fig 4.1.6.4).



Fig 4.1.6.4

Step 9: Click OK.

Step 10: In the Gap Detection dialog box, ensure the **Create result** check box is selected. Step 11: In the **Result name** field, enter **Missing Dates**.

Field to use:	INV_DATE	▼ Criteria	: [ОК
Date					Cance
() All	Starting date value:	2012/01/01	📝 Ignore weekends		Help
Range	Ending date value:	2012/12/20	Ignore holidays	Set Holidays	
~~~~					
Output					
Create dat	tabase	V	Create result		
		0.0	with parson Missing Date	2	

Fig 4.1.6.5 Missing Dates

242



#### Step 12: Click OK.

The result will display the ranges of dates that are missing in the database. Note that each row shows a gap. Certain rows may contain more than one missing date.

/	Sales Transactions	i-Database ×		¥ X
6	\$ \$ • @ •			
	From: INV_DATE	To: INV_DATE	Number	*
Ð	10/24/2012	10/26/2012	3	
Ð	11/1/2012	11/2/2012	2	
	11/8/2012	11/9/2012	2	
	11/8/2012			
	11/9/2012			
Ð	11/15/2012	11/23/2012	7	
Ð	11/27/2012	11/27/2012	1	
	11/29/2012	11/30/2012	2	
	11/29/2012			
	11/30/2012			
Ð	12/6/2012	12/7/2012	2	
Ð	12/13/2012	12/19/2012	5	h
				11
		Total number of items detected	164	
		Total number of gaps detected	58	-

## Fig 4.1.6.6 Resulted Data base

# **4.2 DEFINE ACTION FIELD**

Defines a custom action to be executed when the field is selected. It allows looking up for the corresponding values in the linked file. For example, you have two files. Account details file has details such as account number, branch code, status of the account and balance amount in the account. customer details file has details such as account number, customer name, address, etc. this function will allow you to perform a defined task on common fields present in both the files in our case account number and will fetch the data from another file corresponding to the selected field.

# Step by step procedure for using indexed function

Open databases 'Account Details' and 'customer details'

Define action on filed ACC_NO

Step 1: Data > Define Action Field

Step 2: Click on column ACCNO heading to select entire column





Pore     Data     Analysis     Vev     Vev     Data     Analysis     Vev     Decembe     Sortel Orde     Order	Macros SmartAnalyzer	Red Find Ge Search Next To Search	A (
ile Explorer	a 🖉 Account Detail	K X	• × Properties
Desktop Project	ACCNO STA 1 43908696 2 53207396	ACINE STATUS BR_CODE BALTODAY BALYESTDAY 4390656 S320736 Define Action Field	
Account Octable         Account Octable           Countorner Octable         Countorner Octable           Loan matter         Master Data-Engligver Master           Master Data-Engligver Master         Pymmerh-Payments           Payrell March 2015-pay         Sales Register           SklESS         Ublig-Database           Verbical Clains-Claim         Verbical Policy Master	3         4300399         9           0,00         4         7897590           1,000         5         27199191           1,000         5         27199191           133         7         9212005           136         8         84249982           137         9         920005           138         9         442917180           199         14         3252000           101         12         7707296           136         15         7222206           136         15         7222206           136         15         7222206           136         15         7222206           137         9220003         17           13         3230407         17           13         3230407         17           13         32304087         17           10         32304087         12           113         3204087         12           114         32304087         12           115         12         3204086           117         9270464         12           118         3204087         12 <t< td=""><td>Renove Action Link     OK       Orester Action Link to display ents factor preview     Crand       This Action will display after records from the distabase selected balaxes take.     Help       This Action will display after records from the distabase selected balaxes take.     Help       This Action will display after records from the distabase selected balaxes take.     Help       To find the Action bars on additional balaxes selected balaxes take.     Help       To find the Action bars on additional balaxes selected balaxes take.     Help       To find the Action bars on additional balaxes.     Interview       Account Details.     Interview of take.       Account Details.     Interview of take.       To select a different field on rulipite fields as the look.     Look.       Plansmeter     Interview       Plansmeter     Interview</td><td>Indiana State State</td></t<>	Renove Action Link     OK       Orester Action Link to display ents factor preview     Crand       This Action will display after records from the distabase selected balaxes take.     Help       This Action will display after records from the distabase selected balaxes take.     Help       This Action will display after records from the distabase selected balaxes take.     Help       To find the Action bars on additional balaxes selected balaxes take.     Help       To find the Action bars on additional balaxes selected balaxes take.     Help       To find the Action bars on additional balaxes.     Interview       Account Details.     Interview of take.       Account Details.     Interview of take.       To select a different field on rulipite fields as the look.     Look.       Plansmeter     Interview       Plansmeter     Interview	Indiana State
Hile Explorer Library	27 23206299		v
Sunning Tasks		Not connected to IDEA Server	Number of Records 1,000    Disk Space: 250.18 0

Fig: 4.2.1 – Define Action Field Options

Step 3: Click on ... besides file name 'account details.imd' and select 'customer details.imd' from the list

Capy Index Sort Describe Stored Order	end Modify Renove Fields	ind Define Action Mail	e sero		
ile Explorer	Ŧ	Account Details X		- ×	Properties
■四44mm		ACCNO STATUS BR_CO	ODE BALTODAY BALYESTDAY	î	1 * Detebane
	Records 1,000 1,000 543 133 145 145 145 101 200 201 201 201 201 201 201	2 5120736 3 4306590 5 273301 5 252325 1 25255 1 25555 1 255555 1 255	Literine Action metal     Literine Action metal     pais work to perform on the field "ACCING"     on UR     Link budgely estimation privilee     nel display estimation     nel display estimatinter     netable     nel display estimation     nel display estima	OK Gandi Heb	Cola     Histoy     Field Solities     Genetal Total     Jacograp.ad     diAVESTBAY     Control     Vensits     Add comment     Add comment
GRunning Tasta				2	

Fig. 4.2.2 – Create Action link

Step 4: Select look up field option and click select field ACCNO and click on OK
## Advanced Analysis Using IDEA



Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration     Image: Control Acceleration       Image: Control Acceleration     Image: Control Acce	riew Macros d Hodify Renove Fields	CaseWare IDEA - Account Details	- 0 - 6
File Explorer		Account Details ×	+ × Properties #
Antice Preset Norm Concount Details Concount Details Concount Details Concount Details Concount of the State Master Obst-Englisher Master Prove Liver State Prove Liver Antice Prove Liver Antice Prove Liver Antice StatESS Unc., etc. Igg-Database Vehical Policy Master	Records 1,000 1,000 563 133 165 188 199 101 200 39 166	ACCM0         GTATUS         BR_CODE         BALTOON*         ANXESTAW           4006666         4200666         Exclose         Exclose         Exclose           1202016         Lock-up: Fields:         Exclose         Color         Color         Color           12020176         Access Tests all lask for matheal based on the values to the lask-up field sheeth by debut is the fold that the Action is byor color sheet all different for matheal based on the values to the lask-up field.         OK           12020176         Access Tests and that the table is the lask-up field.         OK           12020176         Access to that the debut the Action is byor color sheet value on exect all different is the lask-up field.         OK           12020176         Access to that the debut the Action is byor color sheet.         The Action is byor color sheet.           1202016         Access to the color sheet on color sheet on color sheet on color sheet on color sheet.         The Action is byor color sheet.           1202017         Access to the color sheet on color sheet on color sheet on color sheet on color sheet.         The Action is the lask-up field.           1202018         Pield hames         Pield hames         Pield hames           1202018         Pield hames         Pield hames         Pield hames           1202018         Pield hames         Pield hames         Pield hames	Partalana     Partalanaa     Partalanaaa     Partalanaa     Partalanaaa     Partalanaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
< 12 file Explore (	3	28.202000         Parameter           2.202000         Parameter           3.2020000         Parameter           3.2020000         Parameter           3.2020000         Parameter           3.2020000         Parameter	
Manuaged Project: Adv. (TT		Not connected to IDEA Server	Number of Records: 1.000 Disk Spece: 250.18 G
1 or (2 62	va G		17.59

Fig. 4.2.3 – Select Lookup field

Step 5: Click ok on the Define action field dialogue box

Step 6: Observe ACCNO field. Filed converted into a hyperlink type filed.

Step 7: Click on any record in accno field to see the corresponding information available for the selected ACCNO in the customer detail file. Refer Fig. 4.2.4 below for results

Index Sort Describe Apper Sorted Order	M Modify Renove Fields	Find	Define Action Phel	Fred	Find Go Next To Search	Scandh								
Explorer		10	Account	Details ×	Employee								+ X	Properties
日本中国国		1	ADCN0 4386666	STATUS	BR_CODE 65	BALTO	0.4Y	BALYESTDAY 473.53					^	1 * Detebane
op Project		2	\$3367396		48	11,8	14.52	12,236.48						J Data
w -	Records	3	43808599	N	87	3	66.91	432.94						History
Account Details	1,000	4	72907590		45		in .		Extract	inn Dracine			×	Field Statistics
Customer Details	1,000	5	2719391		- 79		101		EXILATI	TOO PREVIES	····			Control Total:
Loan master	363	B	12131608	D	£1	1	AD	CNO CUST_NAME	CUS	T_ADD	LOCALITY	STATE		13,020,974,84
Master Data-Employee Master	133	7	92120005	1	-40		1 438	06399 PAUL WOODS	71-426 D	DMINIC ST.	Malad	Maharashtra		(BALYESTDAY)
Master Data-Supplier Master	363	B	12429892		81		-							Criteria
Payment-Payments	185	. 9	12521235	1	11	9								I w Results
Payroll_March 2013-pay	189	10	4291/193		90	- 1								
Sales Register	999	11	12522090		62	- 1								1 a holized
SALES95	101	14	122101290		14									· · march
User_web_log-Database	200	13	1013096		30	- 4								i 👻 Comments
Vehical Claims-Claim	39	14	2212289			- 14								an Add comment
Vehical Policy Master	106	16	#3736406	0		_								Add comment
		17	\$1305001		51									
		18	32304807	1	- 12	_								
	1	10	2430105		51	_								
		20	32006898		36									
		21	92702604	N	67					Save	Print	Dor	ne	
	1	22	\$2120005		37	-				. International sectors				
		23	2608199		-91	- 1	90.94	290.94						
		24	22009098	1	87	1,4	00.00	1,517,121.70						
		25	\$2013894		97	74,3	126.10	93,176.12						
	,	28	13205094	N	32	3	43.75	335.87						
te rebrares ( Transit)		-27	23206200				53.91	53.91					~	

Fig. 4.2.4 – Result of define action field

Step 8: To remove the link created by using define action field, go the option and select remove action link option





# 4.3 SEARCH

IDEA offers a variety of search options such as character, date, numeric field search. In character based search feature a particular word is searched in multiple databases and within databases in a particular field. This option is mostly used for character intensive databases having multiple character fields. IDEA offers an option of saving the search result in a separate database for future reference.

IDEA offers advanced search features as under:

Boolean Search: *Boolean Search* is a type of search allowing users to combine keywords with operators such as AND, NOT and OR to produce more relevant results.

Wildcard Search: *Wildcard searches* are not simply exact string matches, but are based on character pattern matching between the characters specified in a query and words in documents that contain those character patterns.

Proximity Search: Proximity search looks for documents where two or more separately matching term occurrences are within a specified distance, where distance is the number of intermediate words or characters.

#### Step by step procedure for using search function

#### Step 1: Data > Search



#### Fig. 4.3.1 Search

Step 2: In 'text to find' - type the word to search

Step 3: Tick case sensitive for more accurate search result

Step 4: Tick 'Create an extraction database' and provide file name to save search result

Step 5: Under scope option select field in which search word will be searched

Step 6: Cick on ... besides 'look for other databases' to look for search word in other databases

## Advanced Analysis Using IDEA



1 1 三 1? 国際服	1 5 A A 7 1	
py Didex Suit Describe Append Modify Removes Sorted Order	Search	
cord Order Fields e Explorer #	Event to find: Scope: Byles Byles B	Properties
Account Details 1,000 Custome Details 1,000 Loss master Master Dota-Tisuppler Master 363 Master Dota-Tisuppler Master 363 Paymerk-Paymerk 315 Paymerk-Paymerk 315	Image: State of the state o	Duta History Field Solitica Control Tetal Control Tetal Control Tetal
SalLSSS 101 SalLSSS 101 Use, web Jog-Database 200 Vehical Factors Claim 39 Vehical Policy Master 100	Searching Isos  General Search Ing Tips Advanced Search Peatures  About the Search Task Boolean Operators To Perform a Search Wildowd Searches Searching Character Fields Searching Tunneric Fields Vields Vi	I - Comments  Add comment
jFile Explaner/titerary/	Help         OK         Cancel           26:         Datk Genbel         3551 Dasp Rock Rd         RICHMOND         23226         8045356409           27:         Datkmash Hetek         42255 Sommand Hills Terr         RICHMOND         23226         804545673	

#### Fig: 4.3.2 Search options



Fig: 4.3.3 Search result



3.●ウルビが詳細目か。	CaseWare IDEA - Bruce-Master Data-Employee Master	- 8 ×
The Here Data Arolysis Ven Marvis Capy Index Sort Decrific Sortiad Database Capy Order Product Huddly Review Sortiad Database	ShartWeeksor Bio Prod Define * Action India Social Social	≏ €
File Explorer +	Dota-Supplier Master 🗴 🗑 Master Data-Employee Master 🛪 🗑 Bruce-Master Data-Employee. 🗴 🗑 Bruce-Master Dota-Supplier M., 🗴 🖛	Properties 9
11111111111111111111111111111111111111	EMP_NAME ADDRESS CITY ZIP MOB_NO 1 Device Remote E 363 Vertage Park Dr. 8K/HA/DKD 23226 M/0563560	1 🕶 Detabana
Dinitita Naget Name Racondi Name Racondi Account Details 1,000 A Customer Details 1,000 Customer Details 363 Moster Dota-Employee Master 313	2] Brane Cody   150 W. Cerpenter Ave.   RICHMCIND   23226   8046562783	Data     History     Field Statistics     Control Total     Criteria
Bruce-Master Data-Employee Z		i ▼ Results
- Master Data-Supplier Master 363		1 > Indess
Payment-Payments 185		I + Comments
Sales Register 999     Sales Segister 999     Sales Segister 200     Vehical Claims-Claim 39     Vehical Claims-Claim 39     C     This Explanary     This Explanary		Add comment
Search Results	9	
DATABASE         RECORD_JULINER         PELD.           1         Meter Data-Employee Master         14         EMP_N           1         Meter Data-Employee Master         15         EMP_N           3         Mester Data-Supplier Mester         37         SUPPUE	juoze troot MAE BouceBerghal MAE BouceCody RNAAME BouceBerghal	
Running Tasks G Search Results		and the second second second
Managed Project: Adv. ITT	Nut connected to IDEA Server Number	of Records: 2   Disk Specie: 250.18 GB
📲 📭 🍘 🎆 📲 🤗		- Nr R *: 4+ 2052 04-06-2017

Fig: 4.3.4 Search result

By clicking on the respective record in search window Fig 4.3.4. IDEA opens the database and points the cursor to that record in the database

## 4.4.BENDFORD'S LAW

Benford's Law, also called the first-digit law, was made famous in 1938 by Physicist Frank Benford, who after observing sets of naturally occurring numbers, discovered a surprising pattern in the occurrence frequency of the digits one through nine as the first number in a list. In essence, the law states that in numbered lists providing real-life data (e.g., a journal of cash disbursements and receipts, contract payments, or credit card charges), the leading digit is one almost 33 percent (i.e., one third) of the time. On the other hand, larger numbers occur as the leading digit with less frequency as they grow in magnitude to the point that nine is the first digit less than 5 percent of the time.

In the 1970s, Hal Varian, a professor at the University of California's Berkeley School of Information, suggested that the law could be used to detect possible fraud in lists providing socioeconomic information. Since then, Benford's law has been applied to large numbers of data to detect unusual patterns that are often the result of errors or, worse, fraud. As part of their work, internal auditors often employ tools and scientific methods that enable them to detect instances of fraud.

Benford's law states that if there is a set of non-manipulated, naturally occurring numbers, the occurrence frequency of digits one through nine as the first digit should be expected. As we can see from the numbers in table 1, naturally 30 percent of numbers have one as a leading digit, and nine occurs as a leading digit only one time in twenty. Because most financial and accounting data conform to naturally occurring numbers, by comparing the occurrence frequency of these first digits to Benford's pattern, auditors should be able to determine irregularities and possible manipulations.

In 1938, the research and calculations were performed manually, which was painstaking. Today, with computing power and the ease of accessing big data sets, one can see that Benford's Law of expected numbers is valid. It tests data such as Twitter users by followers' count, most common



iPhone passcodes, population of Indian cities, government spending, and even includes the first 652,066 Fibonacci numbers.

The expected values for any data set of the first leading digit and also for the first two leading digits are outlined in Table 6.1 below. For the first digit test, the first leading digit output is depicted in the graph in Fig 4.4.1. For example, the leading digit 1 appears 30 percent of the time, whereas the leading digit 9 appears 4.6 percent of the time. The bars are the actual data counts and the lines are the lower and upper boundaries along with the expected count. This data set conforms to Benford's Law.

	First Digit Frequency	Second Digit Frequency
0		0.11968
1	0.30103	0.11389
2	0.17609	0.10882
3	0.12494	0.10433
4	0.09691	0.10031
5	0.07918	0.09668
6	0.06695	0.09337
7	0.05799	0.09035
8	0.05115	0.08757
9	0.04576	0.08500

 Table 4.4.1: Benford's Law First Digit Frequency and First Two Digits Frequency

Following are the assumptions of Benford's Law: The numbers in the data set should describe the same object

There should be no built-in maximum or minimum to the numbers

The numbers should not be assigned, such as telephone numbers, bank account numbers, social insurance, or social security numbers

Does not apply to uniform distributions such as lottery balls where the uniform balls are selected and not the actual numbers

Primary Benford's Law tests are the first digit, first two digits, first three digits, and second digit tests. Advanced Benford's Law tests are summation and second order. Associated tests are last two digits, number duplication, and distortion factor model. All but the last two tests can be automatically executed from within the IDEA software.

The number duplication test identifies specific numbers causing spikes or anomalies in primary and summation tests. Spikes in the primary tests are caused by some specific numbers occurring abnormally too often. Abnormally large numbers in value cause spikes in the summation test.

The distortion factor model shows whether the data has an excess of lower digits or higher digits. It assumes that the true number is changed to a false number in the same range or percentage as the true number.

Most presentations and articles discuss using Benford's Law to detect numbers near their authorization limits. For example, if someone's authorization limit is Rs. 10,000, then many first two





digits in the 99, 98, and 97 areas will be detected using Benford's Law if they are trying to maximize authorizing expenditures.

Another example Bank has the limit up to Rs 50,000/- for cash deposit without PAN. For money laundering transactions one will deposit just below the limit i.e Rs 49,999/- or Rs 49,500/-. Then last two digits in the 99 and 00 will be detected using Benford's law last two digits in IDEA.

Some other practical applications include:

- Accounts payable (expenses) data
- Estimations (accruals) in the general ledger
- Sales
- Purchases
- Non-arm's-length transactions
- Customer refunds
- Bad debts
- Anti-money laundering

### **BENFORDS'S LAW IN IDEA**

The Benford's Law feature in IDEA can provide a valuable reasonableness test for large data sets. IDEA only tests positive numbers 10 and over in the data file. For negative numbers, values greater than minus 10 are excluded (exclude -9, -8, ... -1). These steps eliminate immaterial items from the analysis. Positive and negative numbers are analyzed separately.

The positive and negative numbers are evaluated on their own due to the fact that positive numbers behave very differently from negative numbers. For example, where positive earnings are manipulated for management bonuses, there is motivation to increase the earnings, moving away from zero toward larger numbers. Where there are losses and management wishes to improve stock prices, there is incentive to move the larger negative number to a smaller one toward zero.

IDEA can apply most of the Benford's Law tests and can also display suspicious results in graphical format. Tests provided in IDEA are the first digit, first two digit, first three digits, second digit, last two digits, second order, and summation tests as shown in Fig 4. 4.2

Analysis > Benford's Law



## Advanced Analysis Using IDEA

Provet Propagate	Gep Detectors 📶 Dad Derfords Leer 🖄 Stat Explore	itata Kay + 30	Elumentation (1) Age Statification (1) Pro Overt Categorize	nt Table	불= John NG Valid Connector 봄= Append 또 Company Relate	Hattrbute III Hones III Random III Variabi III Other * Sample	ryUnit * 65 *			
le Explorer 自己之内回日 Initiae Mayer		Master I ACCNO 1 40008590	Dirta Employee Master ) STATUS BR_CC 6	X Q	Mester Data-Supplier Mester × BALTODAY BALYESTDAY 473.53 473.58 Benford's Law	🕲 Customer Details	x' 🖗 Ac	count Details ×	• ×	Properties 1 * Detabure W Data
Account Details Customer Details Loan master Master Data-Employee Master Bruce-Master Data-Employee. Bruce-Master Data-Supplie -	1,000 1,000 363 133 2 1	4 5 5 7 8 0	Field to analyze: B tude Values (Positive Negative	ALTODA	Y Dptores In Show boundaries In Meen absolute deviation	v	A.	car twenced Cancel		Field Statistics Control Total Criteria
Koter Doz-Supple Moster     Paynet, Promote - Paynet, March 2017-piny     Sales Register     Sales Register     User, web, Jop Outbase     User, web, Jop Outbase     Vehicol Policy Moster	363 185 189 999 101 200 39 186	10 Ram 21	uit "Oreste result thyle: The status: Prest two digits: Second digit: Second order: Second order:	-	Bentard Create database: Bentard Frest Type Bentard Frest Type Bentard Frest Type Digts Bentard Frest Type Digts Bentard Executed Digt Pentard Executed Digt Bentard Executed Digts Bentard Bentard Executed Digts Bentard Executed Digts		✓ Suspects ✓ Suspects ✓ Suspects ✓ Suspects ✓ Suspects	a. 6. 8.		Hidesi     Comments     Add comment
File Explorer ( through	Ł	25 28 27	Duringoon	-	Coreste a virtual datateae				÷	





Fig: 4.4.3 Benford's first digit test







Fig: 4.4.4 Benford's second digit test

This first two-digit primary test output from IDEA indicates that it does not conform in Fig 4.4.4 The graph highlights the three most highly suspicious numbers and the three most suspicious items. By placing the cursor over highly suspicious bar numbers 21, 70, 99, options for extracting or displaying the records are offered. Field statistics may also be displayed.

#### Conclusion

Benford's analysis, when used correctly, is a powerful tool for identifying suspect accounts or amounts for further analysis. Benford's analysis is a tool to complement additional tests/tools. Benford's Law is a wonderful tool for initial risk assessment of the contents of a data set. It provides the auditor or investigator with a good starting point. The user must understand the business and the industry to effectively use this tool. Knowledge of the business can quickly eliminate false positives.

# 4.5 ADVANCE @ FUNCTIONS

Step by Step procedure for accessing @Functions:

- Open any database
- In the Properties window, click Criteria.
- The Equation Editor is displayed. The right-side of the window displays the list of available @Functions and displays the help page for selected @Functions.

## Advanced Analysis Using IDEA



uation Editor		CR HOT MOD EX		al		Al     Al     Content     Al     Al     Content     Numeric     Nutching     Coditional     Francial     Custom Functions	Character OFunctions The following OFunctions are used for text manpulation and conversion. Date fields can also be used as a string in these offunctions. @AllTrim -Removes all leading and trailing spaces. @Chir - Provides the character equivalent of a specified ASCII code.
Record Number: 1	Evaluate						Ignores uppercase letters when comparing expressions.
Field	type	Len	Dec		•		OCurform - Determines the format of Numeric fields
LOCUD	Numeric	8		0	91		ODulate - Deletes specified
TEAH	Numeric	8		0			characters from a string.
PAT_TIPE	Character	41		0	-		OF ind0xeOf - Finds
NO_OF_MONTHS	Pouneic .			0	-11		records with at least one
NO_OF_THANSACTIONS	Numeric	8		0			specified character.
					1011	Instant Distortions	OC et At - Behave the

- Expand a category node to display a list of associated @Functions.
- Select a @Function from the list to display its information.

IDEA provides more various @Functions for performing operations such as date arithmetic, financial and statistical calculations as well as text searches.

The @Functions are accessed through the Equation Editor. Quick help including the syntax, a description, and an example of use for each of the @Functions is available when the @Function is highlighted in the list in the Equation Editor window.

Read the definitions below to become familiar with each @Function:

- Numeric expression: Refers to a number or an equation that evaluates to a number.
- String: A series of characters, such as a name or address.

@Function	Description
@Abs	Returns the absolute value of a numeric expression.
@Afternoon	Returns 0 if time is in the AM and 1 if time falls in the PM and -1 for an invalid time.
@Age	Calculates the number of days between two dates.
@AgeDateTime	Returns the number of seconds between two dates and times.
@AgeTime	Returns the number of seconds between two times.
@AllTrim	Removes all leading and trailing spaces.
@Ascii (not in IDEA Unicode)	Provides the ASCII value of a character.
@Between	Determines if a numeric expression falls within a specific range.
@BetweenDate	Returns a number indicating whether a date value falls within a specified range (1) or not (0).





@Function	Description
@BetweenTime	Returns a number indicating whether a time value falls within a specified range (1) or not (0).
@Bit	Identifies a bit value.
@BitAnd	To mask out unwanted bits.
@BitOr	To set required bit.
@Chr (not in IDEA Unicode)	Provides the character equivalent of a specified ASCII code.
@CompareNoCase	Ignores uppercase letters when comparing expressions.
@CompIf	Determines if a record satisfies multiple criteria.
@Ctod	Converts character dates to IDEA Date Format.
@Ctot	Converts character time values to IDEA Time Format.
@CurForm	Converts numeric value into a formatted text.
@CurVal	Converts formatted character fields to numeric fields.
@Date	Returns the present date.
@Day	Returns the day in a date expression.
@DaysToD	Converts a number of days since Jan. 1, 1900 to date format.
@Db	Calculates the fixed declining-balance depreciation for a specified period.
@Ddb	Calculates double declining-balance depreciation.
@Delete	Deletes a specified number of characters from a string.
@Dow	Returns the day of the week.
@Dtoc	Converts date expressions to character.
@DToDays	Reveals the number of days between Jan. 1, 1900 and a specified date.
@Dtoj	Converts dates to Julian format.
@Exp	Calculates the exponent of a numeric expression.
@FieldStatistics	Returns the numeric value for a specified field statistic.
@FindOneOf	Finds the position of the first matching character in 2 strings.
@FinYear	Returns the financial year for a given date based on the year end.
@Format12hourClock	Returns a string representing time formatted as HH:MM:SS TT.
@Fv	Calculates the future value of an investment.
@GetAt	Returns the character that appears in a specified numeric position.
@GetNextValue	Returns the next value in the selected field.
@GetPreviousValue	Returns the previous value in the selected field.
@Hours	Returns the hours portion of a given time.
@If	Allows a choice of two results based on the evaluation of a condition.
@Insert	Inserts a string into an existing string.
@Int	Returns the integer portion of a numeric value.

# Advanced Analysis Using IDEA



@Function	Description
@Ipmt	Calculates the interest payment for a given period.
@Irr	Calculates internal rate of return.
@IsBlank	Tests if a character field is blank.
@IsFieldDataValid	Returns a 1 if the data in the field is valid or a 0 if the data is invalid.
@Isin	Returns the starting position of a string within another string (case-sensitive).
@Isini	Returns the starting position of a string within another string (NOT case-sensitive).
@Jtod	Converts Julian dates to IDEA Date Format.
@JustLetters	Returns a string with all the numeric characters removed.
@JustNumbers	Returns all the numbers (leading and trailing).
@JustNumbersLeading	Returns the leading numbers.
@JustNumbersTrailing	Returns the trailing numbers.
@LastDayofMonth	Returns the last day for any given month and year combination.
@Left	Returns the specified leftmost characters in a string.
@Len	Returns the length of a string, including any trailing spaces.
@List	Determines which criteria in a list of values is met by an expression.
@Log	Calculates natural logarithms.
@Log10	Calculates logarithm 10x.
@Lower	Converts all characters in a string to lower-case.
@Ltrim	Removes leading spaces from a string.
@Match	Determines which criteria in a list of values are met by an expression.
@Max	Returns the greater value of two numeric expressions.
@Mid	Extracts a portion of text from within a string.
@Min	Returns the smallest value of two numeric expressions.
@Minutes	Returns the minutes portion of a given time.
@Mirr	Calculates modified internal rate of return.
@Month	Returns the month in a date expression.
@NoMatch	Determines if an expression meets none of the criteria in a list of values.
@Npv	Calculates the net present value of an investment.
@Ntod	Converts a numeric expression into an IDEA date format.
@Ntot	Converts a numeric expression into an IDEA time format.
@Pmt	Calculates a loan payment.
@Ppmt	Returns the Principal amount of a loan payment.
@Precno	Returns the physical record number.
@Proper	Capitalizes the first letter of each word in a string.
@Pv	Returns the present value of an investment.





@Function	Description
@Qtr	Returns 1-4 representing the quarter a given date falls in based on the specified year end.
@Random	Generates a random number.
@Rate	Calculates the interest rate of an investment/loan.
@Recno	Returns the logical record number (index-sensitive).
@RegExp	Matches character expressions using a complex set of rules.
@Remove	Eliminates all instances of a specified character.
@Repeat	Repeats the first character of a string a specified number of times.
@Replace	Replaces a string or substring with another.
@Reverse	Reverses the order of characters in a string.
@Right	Isolates the specified rightmost characters in a string.
@Round	Rounds to the nearest integer.
@Seconds	Returns the seconds portion of a given time.
@Seed	Sets the random-number seed.
@SimilarPhrase	Measures the similarity between two specified phrases or Character fields.
@SimilarWord	Measures the similarity between two strings (either single words or character expressions) or Character fields.
@SimpleSplit	Extracts a specified occurrence of a segment of a character string that resides between certain characters like hyphens or backslashes.
@Sln	Returns the straight-line depreciation of an asset.
<pre>@SpacesToOne</pre>	Strips spaces leaving only one space between words in a string.
@SpanExcluding	Returns the characters in a string that appear before any characters in a specified string.
@SpanIncluding	Returns the characters at the beginning of a string that match any character of a specified string.
@Split	Breaks a character string into segments separated by characters, such as spaces or commas, and returns a specified segment.
@Sqrt	Calculates a square root.
@Str	Converts numeric expressions to strings.
@Stratum	Groups records by interval.
@Strip	Removes all spaces, punctuation and control characters.
@StripAccent	Removes an accent from an accented character.
@Syd	Returns the sum-of-years digit depreciation for an asset.
@Time	Returns the present time.
@Trim	Removes trailing spaces.
@Ttoc	Converts a time or number into a string with the HH:MM:SS format.
@Upper	Converts all characters in a string to upper-case.



@Function	Description
@Val	Converts a character expression to numeric.
@Workday	Returns 1 if a given date falls between Monday-Friday and 0 if the date falls on a Saturday or Sunday.
@Year	Returns the year in a date expression.

## Examples of @functions:

@Function	Purpose	Example
@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.
@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)
@isblank	Filter out blank character fields	To present blank scheme agency name in GRANTS-IN-AID data use @isblank(Scheme Name) through Criteria.
@isini	Capture specific character strings from a character or alpha- numeric field where the <b>strings</b> <b>are not case sensitive</b> .	To display cheque in favour of field containing 'Shri' for cheques issued inter-se Government use @isini("shri", cheque in favour) through Direct Extraction.
@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", 1000000, 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.
@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.





@Function	Purpose	Example
@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 Manipulation.
@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.
@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.
@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.
@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 @ month(nodal scroll date) to derive the month of collection through Field Manipulation.
@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.
@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



@Function	Purpose	Example
@DaystoD	Reverse compute Date from days.	In Gradation List to compute the date from the days add a <b>virtual date field</b> with criteria @ DaystoD(Retirement Date Days+30) through Field Manipulation. The +30 is to suggest adding 30 grace-days as an example.
@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)

## **4.6 CUSTOM FUNCTIONS**

Custom functions are written using IDEAScript. Custom functions are used just like the included @ Functions within IDEA. The custom functions are accessible in the equation editor in the same area as the @ functions. Custom functions are prefixed with the # symbol instead of the @ symbol as in regular functions and have the file extension .ideafunc. Custom functions should be located under the active IDEA project in the Custom Functions.ILB folder.

## Step by Step guide to write custom functions:

Step 1: Click on Criteria under properties section

### Step 2: Click on # sign in equation editor

Nertun 클rändered 륬Key Value III 1 Taska Extract	ap Detection 🖉 Duplicate Key + enford's Law 🚔 Statistics + Explore	Stratication Trivial Table	문제 Join 비행 Visual Connect 봄-In Append 또 Company Relate	er 副 Attribute 回 Monetary Unit * 目Randon 回 Variables * 回 Other * Serple		
ile Explorer *			Equation Edit	x	10	ies
	12405	ាំងផង ដោយ 🛛	1 000 2 ;	A TR @		date
Autom And Court Details 1,000 Customer Details 1,000 Losa matter 563 Moster Distor-Eng	Image: second	ce se ce   + . + / AND OF	Opens the Custon Fur	fahrt dalog erkt Hann / Trae Hann / Trae	© Functions A suble Brunctions are available for particineng main complex operations such as data entherates, statistical citization tracticate legis with the "Brundan. They are very selven in the and appearance in the software packages and ne Motional Experiments brucks and a software packages and ne Motional Experiments are software and and the software are and appearance and and the software are and appearance and and the software are and appearance and and appearance and and appearance and and appearance and and appearance and and appearance and appeara	eary del Strinitis write Tatal ania its mi ania irides merits M comment rabase fus beef truncated, e machinum number of
	Field	Туря	Len Dec		parameters passed to the @Punction. Parameters	cords has been reached.
	ADCND	Nurwic	8	0	are passed in	
	CUST_NAME	Diaracter	22	D	example, ØRound	
	LUST ADD	Character	22	0	(Amount) wilround a	
	LOCHLING	Thomas a			ADD FORCE MERCOLLE OF	
	LOCALITY	Character		D.	the neurest whole	
	LOCALITY STATE	Diaracter Diaracter	n	0	nunber. V	

## Fig: 4.6.1 Open equation editor

Step 3: Select new to create new custom function and define a short but very effective custom function as shown in Fig. 4.6.2 below.





Hore Data Analysis	View Mecros SnartAre	lva <del>o</del> r						- G
P Devect         P Top Records         R R           Re FUN         P Indexed         R Ney Value         R R           Tasks         Extract         Extract         R	ap Detection 🔮 Duplicate Key + enford's Leve 🚔 Statistics + Explore	Esumarization Apro Stratification Privot Table Overt Categorize	문의 Son 서울 Visual Connector Brit Append 또 Company Relate	H Attribute A Honetary Unit Randon A Variables * Other * Sample				
ile Explorer *			Fruation Friday		INPOL .		13	ies
目的のもの目	133		Edit Function					date
<ul> <li>Name</li> <li>Records</li> <li>Records</li> <li>Caccord Details</li> <li>Caccord Details</li> <li>Constructed</li> <li>Casterner Detail</li> <li>Moster Details</li> <li>Moster Details</li> <li>Moster Details</li> <li>Moster Details</li> <li>Regimeric Pergimen.</li> <li>Regimeric Pergimen.</li> <li>Sales Register</li> <li>Sales Register</li> <li>Vehical Details</li> <li>Vehical Policy M</li> <li>Moster Details</li> <li>Vehical Policy M</li> </ul>	Epakton Salari (2) Text (2) Text Record fiam Record fiam CuST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST_NAVE CUST	an gare Date Convenien angene Date Convenien gelations Date Convenience Benuit Date Convenience Benuit priorit Date Convenience Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date Date	d Ao Shing] Ao Date	Number of parameters; 0 Dawn et to a date field.		actions / Dors are available formating mana		te any di Statistico entre Tata di Antistico entre Tata di Antistico entre Tata di Antis a di Antis a di Antis
File Explorer Library	<				144	Ronal halo on		
Running Tauta								
Aanaged Project: Ade, ITT				Not connected to IDEA 5	erver	Number of	Haco	edic 1,000 Disk Specie: 250.00 GB

Fig. 4.6.2 Custom Function to Convert a Character Field into a Date Field

The custom function converts a character field containing date information into a date field in IDEA format.

CDate is a Visual Basic function that converts a character date field into an actual date field that is usable by IDEA. For example, the function can convert April 26, 2017, Apr 26/17, or 4/26/17 characters to an IDEA date field in the default format of DD/MM/YYYY.

Step4: Click on OK to save the custom function

The saved custom function is now available for use within equation editor dialogue box under custom function section as shown in Fig. 4.6.3

101 置·Indexed 亮 Key Value 图 De	ep Detection 🛗 Duplcate Key + enford's Leve 🛗 Statistics +	E Sumarization TAping Stratification Frivat Table	금n Join 서울 Visual Conne His Append I Company	dar 計Attrib 記Rand 否Othe	Note Provedany Unit *			-
ska Extract			Equation Edit	tor				
e Explorer *	V 3% 08	今日自日日	1002	al 🗔 🖗	)			les
Here Reserve The Accourt Details 1,000 Loom match 1,000 Loom match 1,000 Matter Doke Trays. 100 Matter Doke Trays. 100 M	Tgaaton       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11       *** , 11	en se es (e ) + - + / AND on	NDT   MOD EXP   0		Davachet Davachet Davachet Davachet Hathong Jondhanud Hunstal Junten Punstern L. Dete_Convention	CFunctions divident are available for jest/waysinate corpsis-operations and obta attrivers, statistical askatations and test soundes. IBBA test soundes. Default a software test soundes. IBBA test soundes. Default a software test soundes. Default as test soundes. Default as		Is day different and a second
	Field         Type         Len         Dec           ADDN0         Namelo         8         0           CUST_NAME         Divacolor         22         0           CUST_ADD         Divacolor         22         0           UOC/UTY         Divacolor         9         0           STATE         Divacolor         11         0	Brunction, Parameters are passed in poverhises, For example, @loand (Amount) will round a field called "Amount" to the nearest whole		cords has been reached.				
			III STAL	-	Lighert Function	Rahltional liabs on	~	
	<							

Fig 4.6.3 – Custom function available within equation editor

260

# 4.7 ADVANCE SAMPLING

Statistics involves the study of research designs, the collection of the data, describing the data, analyzing the data, and then forming a conclusion. We are interested mainly in the analysis of data that has already been collected by us from various business systems. We hope to be able to arrive at various conclusions after analyzing the data. Understanding some basic statistics allows you to understand the makeup or distribution of your data files. This is especially useful when your data file is large and contains millions of records. There are various types of statistical analysis, but the two major categories are descriptive statistics and inferential statistics.

## SAMPLING

The act, process, or technique of selecting a representative part of a population for the purpose of determining parameters or characteristics of the whole population. Simply put, sampling is a process of selecting a subset of the population or a number of records from the data set for the purpose of making inferences or conclusions to the entire population or data set.

Audit sampling is the audit procedure of examining of a portion of items within a class of transactions in order to evaluate one or more characteristics of that entire class. Either statistical or non-statistical sampling methods may be used against a part of the entire data set to make a conclusion regarding the entire data set.

Sampling is effective when the audit procedure or step does not require a 100 percent review of the population of the class, but a decision or conclusion is required and it is not cost effective to audit 100 percent of the transactions.

## STATISTICAL SAMPLING

Statistical sampling uses statistical mathematical calculations for selecting and then evaluating a sample from the data set. Statistical sampling outlines in numeric terms the parameters and precision levels associated with the sample conclusion.

One such use of statistical sampling that we are most familiar with are polls used to determine candidate's current standings in upcoming elections or in popularity surveys.

An auditor may be verifying an account balance through statistical sampling and conclude that it is Rs. 100,000 plus or minus 5 percent or Rs. 5,000 each way (Rs. 95,000 to Rs. 105,000), 19 times out of 20. The conclusion would be that given the precision of the sample at 5 percent (plus or minus Rs. 5,000), there is assurance that the balance is correct with a confidence level of 95 percent. In addition, if the materiality was predetermined to be at 7 percent, then it can be concluded that there would no material error based on the precision level.

Confidence level is the remaining factor when the acceptable sampling risk is eliminated. In the example of selecting a 95 percent confidence level, you allow only a 5 percent chance of getting the wrong sample that does not adequately represent the entire population.

Using statistical samples to obtain familiarity with the data set might be useful but if it is not used to reach a conclusion, it cannot be considered as part of the audit procedure. In addition to formulating a conclusion, statistical sampling must use statistical calculations, and the sample must be random.

Proper use of statistical sampling is beneficial because it:

• Requires a scientifically accepted and defined approach.





- Allows the auditor to maintain professional judgment in regard to audit risks and materiality.
- Displays the sample results and conclusion in relation to the selected data set population along with defined judgment selections.

## NON-STATISTICAL SAMPLING

Non-statistical sampling does not involve the use of statistical calculations. It relies on the subjective sampling selections by the auditor and has less of a standardized approach. Non-statistical sampling is beneficial where:

- The auditor needs to employ professional subjective judgment.
- There is a unique issue where there is a need for a less rigid standardized approach.
- The auditor should not be restricted to explicit numbers as to materiality or risk.

In order to effectively perform non-statistical sampling, the auditor must have a good knowledge of the data set. Knowing the contents of the data or population allows for a supportable sample selection choice and also supports the conclusion of the results. Sample selection may be based on random sampling or other nonmathematical techniques such as judgmental, haphazard, or block selection.

Judgmental selection is frequently used when the auditor is very experienced and selects samples based on sound judgment. Typically, the auditor will make the selections based on a combination of representativeness of the population, value of the items, and relative risk of the items.

Haphazard selection is where the auditor picks items without basis of any mathematical formula. The auditor believes that the items selected are representative of the population and no intentional bias was applied to any of the included or excluded items.

Block selection is where a contiguous sequence of items is selected as samples. These blocks may be invoice numbers from 1000 to 1100 or a specific type of transactions for the month of March. Block selection effectiveness can be much improved by sampling several blocks.

#### SAMPLING RISK

The sample selected either through statistical sampling or non-statistical sampling methods might not truly reflect the population even if done with the utmost care. This is the cause for sampling risk, where the auditor's conclusion based on the selected sample may differ from the reality of the conditions of the entire population of the data set. Sampling risk occurs due to limited time and resources that prevent an audit of the entire population.

Alpha or Type I risk is the risk of incorrect rejection. That is, the auditor incorrectly concludes from the sample that the population errors are worse than they actually are.

Beta or Type II risk is the risk of incorrect acceptance. That is, the auditor incorrectly concludes from the sample that the errors in the population are better than they really are.

Auditors are usually more interested in beta (Type II) risks, as they are concerned about the failure to detect material misstatements.

#### NON-SAMPLING RISK

Auditors may draw incorrect conclusions from using sampling techniques where the wrong inferences are not because of the selected samples themselves, but for other reasons not directly connected with the sample contents.

Non-sampling risk is where the auditor may have selected an appropriate sample but arrived at a wrong conclusion. Examples include employing inappropriate audit procedures, failure to recognize errors present, or misinterpreting the results or evidence.

## NON-STATISTICAL SAMPLING METHODS IN IDEA

The non-statistical sampling features built into IDEA are shown in Fig 4.7.1. IDEA has easy-to-use random, systematic, and stratified random sampling built in.

Porect Prop Records ()      Porect Ray Value ()      So ()      Datact	ap Detectori 🛄 Daplcate I enford's Levi 🏥 Statistica Explore	Key + BE Sumerication Stratification * B Overt Catego	The Aprop	ME Vauel Con rel I Company Ridate	neter H	Attribute 📇 Monetary Unit * Random 🔄 Variables * Other *		
Eurolessee #	Customer Deta	a x			122	Stratifed Random	* X	Properties
Contraction of the state of the state	ACCNO	CUST_NAME	CUST_ADD	LOCALITY	STA 1/4	Generate Random Numbers	^	Topchus
日本で回国	1 2004190 FRAN	KISI TAYLOR	149 1 AVE.	Malad	Maharasitte			IN CONSIDERS
ktop Project	2 2006393 WILLI	AM WECKOWSKI	48-191 67 AVE	Mahalaurre	Maharasiste	*		J Data
me Records	3 2008092 JOHN	I MAXWELL	869 KENSIGNTON AVE.	Andheni	Maharasitte			History
Account Details 1,000	4 2010999 DICK	SMITH	151 MAGNUS AVE.	Rejkut	Gujeret			Field Statistics
Customer Details 1,000	5 2014595 CATH	HE AUER	957 QUEEN ST.	Sunat	Gujarat			E Control Total
Loan master 563 Marter Onto Free 111	8 2018299 RJC P	05T	509 23 E AVE.	Vadodara	Gujarat			E Criteria
Master Data-Emp 133	7 2021392 RANE	94 SCOTT	533 DOMINIC ST.	Malat	Maharasitte			
Master Data-Sup 363	II ALASSA CARL	A INNELL	400 SPEAKE	Mahalaorre	Maharashtr	*		✓ Results
Payment-Payme 185	1 202669 RDGE	2CSNIVIPE	147.20.31.	Andhen	Numarauru	•		
Payroll_March 20 189	10 203390 CHR:	A LOWER	140.11 AVE	Fajece	Guarat			ii # Indices
Sales Register 999	12 211226 TOM	A TEMPERATE	10.05E MAIN ST	Valodau	Guintat			d Martineton
SALESPS 101	13 2117407 KATH	V MANCINELLI	635 58 AVE	Malari	Maharasivir			of repirides
User_web_log-Da., 200	14 2126092 MRS	TOM KAY	67.62 S AVE.	Manalastre	Maharashte			11 + Comments
Venical Clams-Car 39	15 2132496 DAVE	D WITCHER	556 24 ST.	Andheri	Maharashtr			14 10 10 10 10 10
weind Porcy Ma	16 2136091 LOUIS	SE BAUMAN	731 83 AVE.	Rejkot	Gujeret			Add comment
	17 2207090 RAY 0	CASEY	336 KING ST.	Samat	Gujarat			Database has been truncated
	18 2209195 MRS.	RON HART	45-333 12 5 AVE.	Vedeidara	Gigeret			records has been reached.
	19 2218295 HENR	IV HEITZMAN	332 RIVERSIDE DRIVE	Malad	Maharasitte			
	20 2221299 JEFF H	HIRSH	26-216 09 AVE.	Mehalsorre	Moharashta			
	21 2223196 JOHN	ILAGANA	499 BASELINE RD.	Andheri	Maharasitte	*		
	22 2225507 ARTH	IUR C. LINN	770 87 E AVE.	Rejkot	Gujarat			
	23 2226892 GLOR	IA GRANT	605 FOUNTAIN ST.	Sunat	Gigerat			
	24 2230191 LEWE	SLEUTTS	57-53 GARDEN RD.	Vadodara	Gigeret			
	25 2231397 GARY	FEREN	b-745 MAIN ST.	Maled	Maharashtr			
File Feelsmer / Thears	25 2233/99 RDCP	TELLE MURE/GOMERY	13-786 RIVERSIDE DRIVE	Manalaurre	Manarashtr	•		
(manual)	2254/97 CHUC	.s. DRUUDERSKE	110 80 5 MIL	Andrien	Numérasitte		~	

Fig: 4.7.1 Non-statistical Sampling Feature in IDEA

### STATISTICAL SAMPLING METHODS

There are three basic types of sampling methods that auditors may use. The choice of methods depends on the main purpose of the sample and substantive test.

## MONETARY UNIT SAMPLING (MUS)

Monetary Unit Sampling method is used to estimate the total monetary amount of potential misstatement in a population. While other methods are based on occurrences or number of records, this method is based on rupee values where the higher monetary value transactions have a higher likelihood of being chosen in a sample. MUS is similar to systematic sampling, but where systematic sampling may sample every thousandth record, MUS will sample every thousandth rupees. It is typically used to determine the accuracy of financial accounts, where size is the most important factor, and where errors are expected to be few and far between. MUS provides a substantive assessment of error or misstatement in rupee figures and is specifically designed to predict overall error.

MUS should be used when:

- The process audited is well established and known to be reliable.
- The likelihood of errors (misstatements) is low.
- Obtaining a small sample size is important.
- You want to target larger rupee transactions, and expect to see some spikes in the data.





#### Step by step procedure to perform MUS sampling:

#### 1. Planning:

- Determine the objectives of the exercise.
- Define the population.
- Define what a misstatement means.
- Determine sample size, using the following:
- **Confidence level** A percentage value comfort level that the sample will be representative and that you have the capabilities to interpret the results correctly.
- **Tolerable error** -The point of no return past which you would no longer have faith in the process audited, nor the validity of the sample.
- **Expected error -** The amount of errors or misstatements that are reasonably expected in a population.

#### 2. Performing MUS Sampling Procedures

- Select the samples.
- Perform the audit procedures.
- Record and analyze any errors observed.

#### 3. Evaluation

- Create a projected misstatement by summarizing errors and extrapolating these across population.
- Compare ranges of the projected misstatement against the tolerable error limit.
- Draw final conclusions.

As seen in Fig 4.7.2, IDEA will simplify all these steps for you.

POrect         PTop Records         III Gas           Ner(U)         III-Indexed , May Value         IIII De           Tasks         Extract	p Defection (20) Duplicate Kay + SE Summirization effort/s Leve (20) Statistica + Se Overt Explore Over	n III Aging Sin Join III Prvot Table Hri Appen gorite	ME Visual Connector di I Company Relate	計 Attribute 石 Random 凹 Other •	Planetary Unit -		
ile Explorer *	Customer Details ×	CUST ADD	LOCALITY STA	TE	HtFortReport	• × Propertie	s (
Comparison of the second		149 - 1471, 562 569 (2016) 2017 MARC 151 MARCHE ANE 569 (2016) 2017 MARC 151 MARCHE ANE 569 (2016) 2017 MARC 153 (2017) 2017 MARC 153 (2017) 2017 MARC 153 (2017) 2017 MARC 154 (2017) 2017 MARC 155 (2017) 2017 M	Nosara Mahara Makalaman Mahara Anshuri Mahara Anshuri Mahara Vastofara Gujarat Vastofara Gujarat Vastofara Gujarat Vastofara Gujarat Vastofara Gujarat Vastofara Gujarat Vastofara Gujarat Nakalama Mahara Anshur Mahara Nakalama Mahara Anshur Mahara Nakalama Mahara	antra antra antra antra antra antra antra antra antra antra antra antra antra antra antra antra antra antra antra		✓ Defail     ✓ Defail     ✓ Defail     ✓ Defail     ✓ Defail     ✓ Control     ✓ Tetail     ✓ Defail     ✓ Defail	ny Statistics na Trad is i i i i i i i i i i i i i i i i i i
Chietophony Chinny	27 22347N7 CHUCK BRODERICK	170 60 S AVE.	Andheri Mahara	nintre .		~	

Fig: 4.7.2 Monetary Unit Sampling Feature



In the example of testing sales, seen in Fig 4.7.3, we select absolute values as there are a few credit or refund values that are negative amounts in the Total field. We select a confidence level of 95 percent. Confidence levels below 90 percent are not generally recommended for MUS sampling. It should be recognized that the higher the confidence level, the larger the sample size needed.

The tolerable error amount or percentage must be entered. This is the absolute error limit that can be tolerated. The higher the tolerable error, the more errors you can accept and the lower the sample size needs to be. In this example, with sales in excess of Rs.1.7 million, 1 percent or Rs. 17,000 was decided as the maximum tolerable error. A 1 percent loss would be material enough to take actions such as to initiate an investigation or redesign the sales system. Since MUS is only to be used for processes where there is a high degree of confidence based on actual experience, significant errors may indicate fraud, embezzlement, or untrained or incompetent staff.

The expected error is the anticipated misstatement that is a realistic estimate of likely errors expected to found in the process. It is historical experience that dictates the estimate amount. Both the expected error and tolerable error can be entered either as an amount or percentage. In this example, management estimates that approximately 10 percent or Rs. 1,700 is the expected error.

When the Estimate button is selected, IDEA's MUS calculation determines that in order to be 95 percent confident, it is necessary to set the sample size at 363 records and that there should be no more than 36.25 percent tainting or total percentage of errors found.

Once the Accept button is selected, the Monetary Unit Sampling – Extract screen appears as displayed in Fig 4.7.4

	BUS	DATE_DATE	BUSDATE_TIME	TXNID	ORDERDATE_DA	TE C	RDERDAT	E_TIM
1	09/07/	2007	00:00:00	1	09/07/2007			01:0
2	09/07/	2007	00:00:00	2	09/07/2007	_	-	14:5
3	09/0	Monetary Unit S	ampling - Plan				×	5:0
4	09/0						-	5:0
	09/0	Total Value of	Sampled Population			1	Estimate	51
	09/0	Use value	s from database field:	TOTAL				51
7	09/0	Positive v	alues 🕐 Negative v	alues (	Absolute values	- !	Accept	
8	09/0	Union of the	a campled non dation		1 202 422 0	- I	Print	122
9	09/0	value of ut	e sampleu populauuri:		1,702,972.87	1 1	Cancel	
1	09/0	Settinge						20
0	00/0	Seturgs	onfidence level (%).	-	95.00		Help	100
2	09/0		without the teres ( rep.		33.0			1
14	09/0	Amoun!	O Pen	centage				51
15	09/0	Concernance of	Total to an and		17.000.00			5.4
6	09/0		Tolerable error:		17,000.00			5:4
7	09/0		Expected error:		1,700.00	)		6.5
8	09/0							5:5
9	09/0	Change b	asic precision pricing (Bl	PP) from 1	00% : 100	5		6.5
0	09/0							1:2
1	09/0	Approximate	sample size:		36	3		15
2	09/0	Countration			4 (00.0			2.2
3	09/0	Samping inte	rvai:		4,690.0	•		2:4
4	19/0	Sum of tolera	ble sample taintings:		36.2	5 %		13
5	19/0							5:3
6	19/0	You may accept	the population at the 9	5.00% cor	fidence level when no	more than		5:4
17	19/0	0.362473 total t	aintings are observed in	n a sample	of size 363.			5:5
8	19/0	This is the state	in temple size that it	-	-			5.5
9	19/0	This is the minim	um sample size that will	allow you	to draw the above con	clusion.		S:S
0	19/0							0.0
1	19/0	-		_		_		0.0

Fig: 4.7.3 Monetary Unit Sampling planning





Monetary Unit Sampling	- Plan		X	5:01:25	16:27:5
Total Value of Sampled B	ing lation			51213	15.77.5
Total veloc of semples P			Estimate	51342	16-28-0
[♀] Use values from dat	abase field: TOTAL	*	Accept	5:18:32	16:30:2
O Positive values	Negative values @ /	Absolute values	Drint	5:22:01	16:28:0
Value of the sampled	population:	1.702.477.87	Proc	5.45.70	16-27-4
a second second	Monetary Unit Sampling	- Extract			×
Settings				Station and and	
Confiden	Extraction type	High value handling	1		OK
	Pixed interval	High values in s	sample as aggregate		Fields
M Amount	Cell selection	High values in c	database		
To		High value file	name: High Values		Cancel
Exp					Help
Change basic pres	Numeric field to sample	e:	Sample interval:		
	TOTAL	•	The second s	4,690.01	
Approximate sample :	Random starting point	t	Change high va	alue amount:	
Sampling interval:		3,209.28	4,690,01		
Sum of tolerable same					
Sum of tole able same	There are 3451 items with being selected.	h a value of 0. Items	with a value of 0 will	have a 0% chance of	
ou may accept the pop			Total	Records	
1.3624/3 total taintings	Positive values	1,702,4	70.45	86,526	
This is the minimum samp	O Negative values		(2.42)	2	
	Absolute values	1,702,4	72.87	86,528	
007					
007	File name: Monetary	Sample			
007					
and the second se	Create	a virtual database			

#### Fig: 4.7.4 Extracting the sample records for Monetary Unit Sampling

Three hundred sixty-three records are then extracted and an AUDIT_AMT field is created as in Fig 4.7.5 The value of this field is equal to the field being audited but can be changed to reflect the proper amount determined during the execution of the audit procedures.

1	Header Sales IMD ×	Mo	netary Sample	× GML						+ ×
2.3	HD_NETAMOUNT	TOTAL	AUDIT_AMT	MUS_RECNO	MUS_TOTAL	MUS_HIT	MUS_REC_HIT	MUS_EXCESS	REFERENCE	
348	21.00	23.7300	23.7300	85151	1,621,270.8600	1,621,262.7300	15.6000	8.1300		
347	10.75	12.1500	12,1500	85441	1,625,956.9500	1,625,952.7400	7.9400	4.2100		
348	31.35	35,4300	35,4300	85681	1,630,655,4800	1,630,642.7500	22,7000	12.7300		
349	16.90	19.1000	19.1000	86008	1,635,334.7700	1,635,332.7600	17.0900	2.0100		
350	7.95	8.9900	8.9900	86298	1,640,023.6600	1,640,022.7700	8.1000	0.8900		
351	13.25	14.9700	14,9700	86555	1,644,718.9200	1,644,712.7800	8.8300	6.1400		
352	10.25	11.5800	11.5800	86837	1,649,406.4700	1,649,402.7900	7,9000	3.6800	-	
353	36.95	41.7600	41.7600	87122	1,654,119.8300	1,654,092.8000	14.7300	27.0300		
354	22.20	25.0900	25.0900	87429	1,658,795.5700	1,658,782.8100	12.3300	12.7600		
355	11.20	12.6600	12.6600	87711	1,663,475.5100	1,663,472.8200	9.9700	2.6900		
356	28.00	31,6400	31.6400	88028	1,668,191.7500	1,668,162.8300	2.7200	28.9200		
357	7.95	8.9900	8.9900	88289	1,672,856.8400	1,672,852,8400	4.9900	4.0000		
358	7.95	8.9900	8.9900	88576	1,677,545.0200	1,677,542.8500	6.8200	2.1700	1	
359	8.25	9.3200	9.3200	88818	1,682,236.7800	1,682,232.8600	5.4000	3.9200		
360	7.20	8.1400	8.1400	89117	1,686,925.8800	1,686,922.8700	5.1300	3.0100	and the second s	
361	6.75	7,6300	7.6300	89362	1,691,619.7100	1,691,612.8800	0.8000	6.8300		
362	57.45	64.9200	64.9200	89629	1,696,311.9100	1,696,302.8900	55,9000	9.0200	10000	
363	48,45	54.7500	54.7500	89895	1,701,008.0500	1,700,992.9000	39.6000	15.1500		
+									Long	m



Suppose three discrepancies were found, including the one in record number 250 where the actual amount was Rs. 60.00 instead of the Rs. 55.32 recorded in the sales system as shown in Fig 4.7.6. To evaluate the MUS sample, select the Single Sample option.

## Advanced Analysis Using IDEA



			œx	Ware1DEA - Mone	any some sites	-	- Company		-	-	= 0 x
25	SmartAn	0/25									○ <b>0</b>
in.	Icate Key 1	DE Summar	zation T Aging	To bin	M ^{III} Visual Connects	Attribute	Monetary (	Unit -			
Sta	istics *	Sea Stratific	ston 🗊 Pivot	Table	Compare	ERandom	tandom 🐺 Plan Other • 🔝 Extract		[		
e			Categorize		Relate	-11	Evaluat	e †	++ Single Samp	ie:	-
1	Header	Sales.IMD ×	Monetary	Sample.IMD ×			Ht Point	tReport	40 Multiple San	ples	ş 4
	TOTAL	AUDIT_AMT	MUS_RECNO	MUS_TOTAL	MUS_HIT	MUS_REC_HIT	MUS_EXCESS	REFERENCE		- Da	tabase
249	51.6500	51.6500	61230	1,166,364.4900	1,166,331.7600	18.9200	32.7300			-	
250	55.3200	60.0000	61483	1,171,024.9600	1,171,021.7700	52.1300	3.1900			10	lata
251	26.2700	26.2700	61733	1,175,718.5000	1,175,711.7800	19.5500	6.7200			- 88.2	listory
252	18.6500	18.6500	61978	1,180,402.7400	1,180,401.7900	17.7000	0.9500			H F	ield Statistics
253	70.2900	70.2900	62232	1,185,135.5600	1,185,091.8000	26.5300	43.7600				ontrol Total: 11,637.8400
254	8.9900	8.9900	62491	1,189,790.3400	1,189,781.8100	0.4600	8.5300			(	TOTAL)
255	24.2400	24.2400	62743	1,194,493.0500	1,194,471.8200	3.0100	21.2300				riteria
256	17,9700	17,9700	62996	1,199,162.6600	1,199,161.8300	17,1400	0.8300				
257	59.3800	59.3800	63233	1,203,881.6600	1,203,851.8400	29.5600	29.8200			₩ Re	sults
258	61.3600	61.3600	63489	1,208,554.9700	1,208,541.8500	48.2400	13.1200				Annature Unit Complian
259	7.3500	7.3500	63747	1,213,233.7900	1,213,231.8600	5.4200	1.9300				valuate Single Sample
260	40.5100	40.5100	63977	1,217,950.3000	1,217,921.8700	12,0800	28.4300				annue surfacisatione
261	47.8600	47.8600	64237	1,222,652.6800	1,222,611.8800	7.0600	40.8000			I = Inc	lices
262	37,2400	37,2400	64486	1,227,337.9000	1,227,301.8900	1.2300	36.0100				
263	49.2100	49.2100	64737	1,232,025.9800	1,231,991.9000	15.1300	34.0800			1	lo index
264	41.3000	41.3000	65003	1,236,713,4100	1,236,681.9100	9.8000	31.5000			I T Co	mments
265	31.5300	31.5300	65236	1,241,385.9000	1,241,371.9200	17.5500	13.9800				
266	54.1900	54,1900	65476	1,246,088.4100	1,246,061.9300	27.7100	26.4800				dd comment
267	5,5100	5.5100	65718	1.250.755.2000	1 250 751 9400	2,2500	3.2600				

Fig: 4.7.6 Audit Results of the Sampled Records

## The screen in Fig 4.7.7 appears

Method		OK
O Low error rate	High error rate (dassical)	Cancel
Fields		
Book value field:	TOTAL	- Help
Audited value field:	AUDIT_AMT	•
Reference (optional):	REFERENCE	-
Sample settings		
Confidence level (	%):	95.00
Value of the sampled populati	on: 1,70	02,472.87
Sample si	ze:	363
Change basic precision pricing	from 100%:	100.00
High value handling		
High value items are in a file	Select high values file	
High values file: C:\Users\G\Doc	uments\My IDEA Documents\I	DE/
Book value field:	TOTAL	•
Audited value field:	AUDIT_AMT	-
Reference (optional):	REFERENCE	•
Precision limits		
(i) Upper	Upper and lower	

*Fig: 4.7.7 Evaluating Single Sample Results in Monetary Unit Sampling* 

Once OK is selected, the summary shown in Fig 4.7.8 appears displaying:

- Zero overstatements were discovered.
- Three understatements were discovered.





- The net most likely error is Rs. 1,512.47.
- Both the gross and net upper error limits are less than materiality (tolerable error) of Rs. 17,000.

The conclusion is that with 95 percent certainty the projected total errors based on errors found in the sample is within the accepted range and will not exceed the tolerable error of Rs. 17,000. That is, you are 95 percent assured that the sample is representative of the population and that no actions to investigate or redesign the sales system are needed.

Header Sales.IMD × Monetary Sample.IMD ×		*
◎ @ ff ≌ ◎ • छ •		
MUS-	Low Error Rate	
5	Summary	
Confidence Level         95,00           High Value         4,690,01           Value of the sampled population excluding high values         1,70           Total value of high value items         0,00           Value of the sampled population including high values         1,70	Sampling Interval 4,690.01 Basic Precision Pricing 100.00 12,472.87 12,472.87	
Results Excluding High Value Items	Overstatements	Understatements
Sample Size	363	363
Number of errors	0	
Gross most likely error	0.00	1,512.4
Net most likely error	-1.512.47	1,512.4
Total Precision	13,990.00	13,990.0
Gross upper error limit	13,990.00	15,502.4
Net upper error limit	12,477.53	15,502.4
Results for High Value Items		
Number of high value items	0	
Number of errors	0	
Value of errors	0.00	0.0
Results Including High Value Items		
Total number of items examined	363	36
Number of errors	0	
Gross most likely error	0.00	1,512.4
Net most likely error	-1,512.47	1,512.4
Gross upper error limit	13,990,00	15,502.4
Net upper error limit	12,477.53	15,502.4
Conclusion: Based on this sample, the most likely total overstatement in the populat	ion is 0.00, and the most likely total understatement	is 1.512.47. You can infer with a
Based on this sample, the most likely total overstatement in the populat confidence of 95.00% that the total overstatement in the population do not exceed 15,502.47	ion is 0.00, and the most likely total understatement es not exceed 13,990.00, and that the total underst	is 1,512.47. You can infer with a atement in the population does

## Fig: 4.7.8 Summary and Conclusion of Audit Results of MUS

#### ATTRIBUTE SAMPLING

Attribute sampling is a statistical sampling technique often used to test internal controls; it evaluates the individual attributes of a record to be either true or false. Examples include:

- Having two required signatures for check authorizations over certain amounts.
- Whether account receivables are overdue.
- If travel expense claims are valid or not.

Attribute sampling should be used when:

- There is a need for a statistical sampling solution and judgmental sampling will not suffice.
- The objective of the review is to test compliance to internal controls.
- The compliance testing should evaluate to a true or false result.
- A random selection process will meet the objectives of your review.

## Step by step procedure to perform attribute sampling:

#### 1. Planning

- Determine the objectives of the exercise.
- Define the population.
- Define what a misstatement means.
- Determine sample size, using the following:
- **Confidence level:** A percentage-value comfort level that the sample will be representative and that you have the capabilities to interpret the results correctly.
- **Tolerable error:** The point of no return past which you would no longer have faith in the process audited, nor the validity of the sample.
- **Expected error:** The amount of errors or misstatements that are reasonably expected in a population.

## 2. Performing Attribute Sampling Procedures

- Select the sample.
- Perform the audit procedures.
- Record and analyze any errors observed.

### 3. Evaluation

- Create a projected misstatement by summarizing errors and extrapolating these across the population.
- Compare ranges of the projected misstatements against the tolerable error limit.
- Draw final conclusions.

Similar to MUS sampling, attribute sampling requires a user to set certain boundaries and tolerances for the calculations to be performed.

- **Tolerable deviation rate as a percentage:** Also known as tolerable error rate, this is the absolute maximum percentage of transactions in error (i.e., not in compliance) that is acceptable as a cost of doing business. If you have more errors than the tolerable error rate, this internal control is not working and must be redesigned. The higher the tolerable error, the more errors you can tolerate, and the lower the sample size needs to be.
- **Expected deviation rate as a percentage**: Also known as expected error rate, this is the percentage of errors (i.e., noncompliance) you would reasonably expect to see, based on experience. As a rule, the lower the expected error, the lower the sample size.
- **Confidence level as a percentage:** This is the likelihood that the sample records chosen are indeed representative of the population at large, and that you will correctly interpret the results. The more confident you need to be; the more samples you require.

In the example in Fig 4.7.9, management decided that the maximum percentage tolerable deviation rate is 10 percent. Anything above 10 percent would suggest that the control is not working and may need to be redesigned. Based on previous history, management expects a deviation rate of 3 percent.





Management is happy to accept a confidence level of 90 percent that the sample is representative of the population or data set. The population size or number of records is 89,979.

The population size, percentage tolerable deviation rate, percentage expected deviation rate, and the confidence level must be entered into the Planning tab of the Attribute Sampling box.

soute sampling			
lanning (Beta Risk Control)	Planning (Beta and Alpha Risk Control)	Sample Evaluation	
Population size	89979	% Expected deviation rate:	3.00
% Tolerable deviation rate:	10.00	Confidence level (to control Beta Risk):	90.00
Sample size :	Critical	number of deviations in sample:	
Deviations	% Deviations Ac	hieved Confidence (Beta Risk Control)	)

Fig: 4.7.9 Planning for Attribute Sampling

Once the compute button is clicked, based on the information entered, IDEA informs you that you would need a sample size of 52 and that there must not be more than two deviations or errors in the sample to achieve a 90 percent confidence level that deviation level of the population is not more than 10 percent. Refer to Fig 4. 7.10

Population s	ze: 89979	% Expected deviation rate: 3.00
% Tolerable deviation ra	ite: 10.00	Confidence level (to control Beta Risk): 90.00
Sample size :	52	Critical number of deviations in sample: 2
Deviations	% Deviations	Achieved Confidence (Beta Risk Control)
0	0.00	99
1	1.92	97
2	3.85	90
3	5.77	7
4	7.69	60
5	9.62	41
6	11.54	20
7	13.46	14
8	15.38	
Conclusion: If no that	more than 2 deviations are the population deviation rat	observed in a sample of size 52, you can be at least 90.00% confider e is not more than 10.00%.
		Print Close Compute He

Fig: 4.7.10 Conclusion to achieve attributes sampling objectives



The random record sampling feature in Fig 4.7.11 can be used to randomly select the 52 required records. A detailed audit of the 52 selected records determines whether the control was met. The number of deviations is noted.

Number of records to select:	52		OK
Random number seed:	17194		Fields
Starting record number to select:	1		Cancel
Ending record number to select:	89979	A. V	Help
Allow duplicate records:			
Create a virtual database:			

Fig: 4.7.11 Obtaining sample for attribute sampling

In the Sample Evaluation tab of the Attribute Sampling box, entries for the population size, sample size, and percentage of desired confidence level are made again. The number of deviations (five) discovered during the audit of the 52 samples must also be included, as noted in Fig 4.7.12.

Since five deviations are more than the critical number of deviations (two) in the sample calculated in the planning stage, the conclusion is that there is 90 percent certainty that the number of deviations could be significantly higher in the overall population than the 10 percent tolerable percentage. It is 90 percent certain that there could be as many as 17.11 percent of errors if the entire population was audited. As such, this sample cannot be considered as representative and it is likely the control must be redesigned.

		and the second
Population size: 89979	Number of deviations in	sample: 5
Sample size: 52	% Desired confiden	ce level: 90.00
nple deviation rate: 9.62% (5 / 52)		Sec. 1
1-Sided Upper Limit	2-Sided Lower Limit	2-Sided Upper Limit
17.11	3.87	19.16
clusion: For an observed number of 5 deviatic confident that the population deviatic can be 90.00% confident that the po 19.16%.	ins in a sample of size 52, you o in rate is no more than 17.11% pulation deviation rate is betwe	an be 90.00% . Alternatively, уоц en 3.87% and

Fig: 4.7.12 Conclusion from the audit results of attribute sampling





# 4.8 USING ADVANCED STATISTICAL METHODS

## 4.8.1 - CORRELATION

A correlation is a relationship between two things or mathematical variables that tend to vary or move together. The data is represented by the letters x and y where x is the independent variable and y is the dependent variable. The independent variable x is usually described first. There is usually some logical connection between the two variables.

Many studies have been done on income levels of high school graduates, college, and university graduates and those with post-graduate degrees. The question in those studies is whether there is a connection or correlation between educational levels and income levels. Educational level would be x, the independent variable that would influence the income level variable of y, the dependent variable.

How much does the independent variable have influence over the dependent variable can be determined by calculating the correlation coefficient and is designated as r. A perfect linear correlation would have r equaling to 1 and a perfect negative correlation would have r equaling –1. The closer r is to 1 or –1, the stronger the correlation. Where there is no correlation, r would equal to 0.

There are three basic formulas for calculating *r*. The correlation coefficient can be calculated for the population, the sample, or the product moment.

The most common one is the product-moment correlation coefficient as shown below.

 $r = \sum (xy) / sqrt[(\sum \sum x^2) (*y^2)]$ 

 $\Sigma$  denotes the summation symbol, so the formula for the top part or numerator is to multiply the *x* variables by the *y* variables and then take the sum of those numbers. For the denominator, you square the *x* variables and take the sum of them. Do the same for the *y* variables and then multiple the resulting two sums. Apply the square root to the results and then divide the final results of the numerator by the final results of the denominator to obtain the correlation coefficient.

Since we already know how to calculate the Z-scores (or have IDEA calculate it for us), a simpler method of obtaining r would be to multiply the Z-scores of the x variables by the Z-scores of the y variables and then total up all the results. Take the total of the results and divide by the number of records less one.

The formula would be  $r = \Sigma(Z_x * Z_y)/(n-1)$  where  $Z_x$  is the Z-score for the *x* variable and  $Z_y$  is the Z-score for the *y* variable. The letter *n* in the formula represents the number of records.

In general, the correlation coefficient, whether negative or positive, can be interpreted as:

.0 to .2	No correlation
.2 to .4	Weak correlation
.4 to .6	Moderate correlation
.6 to .8	Strong correlation
.8 to 1.0	Very strong correlation

The calculations of the correlation coefficients were shown for a better understanding of correlation. The auditor does not need to perform the calculations using the formulas, as IDEA has a built-in correlation feature that provides the correlation coefficient and you merely have to interpret the



results.

Using a summarized monthly sales file from a POS system, we select correlation from the statistics area of IDEA. For demonstration purposes, the fields we correlate will be the HD_NETAMOUNT_SUM field, which is the sales amount before taxes, and the PAYAMOUNT_SUM_SUM_SUM field, which are amounts paid that include taxes. We will also include payments of cash only and payments of debit cards only for the correlation calculation as shown in Fig 4.8.1.1. In addition to the results being displayed, which can be exported to various file formats, you may optionally create an IDEA database of these results.

As expected, there is a perfect positive correlation of 1.000 between sales before taxes and payments. As sales go up or down, the sales tax moves accordingly, so the total payment by customers' correlates to sales net of taxes.

There is a strong correlation between both the cash tender and the debit cards tender to the payment amounts of 0.710 and 0.792, respectively. There is no correlation between cash and debit cards payments as the correlation coefficient is 0.189. See Fig 4.8.1.2.

HTMOMTH	NO_OF_RECS TOTAL SUM TAXT_SUM	TAX2_SUM TAX3_BUM	HO_NETAKUUNT_SUM	PRIMADUNT_SUM_SUM_SUM	CASH_ONLY_BUM	DEBIT_ONLY_SUM
-07	895 18,793,8800 990,190	1.315.7500 0.00	16,487.05	18,791,8800	\$344.75	8,465.8
-08	2108 41 018 1100 2 161 830	2,879,5600 0.00	15,985.M	41,018,2100	10,293.01	19,222.9
09	Correlation			38,243,2300	8,886,00	18,041.3
10				35,219,3500	9,047.48	15,479.9
11	Fields to Correlate	Audit Units	E	OK 35,734,3800	7,813.60	16,683.7
1Z	Include all fields	Select audit unit field	-	19,9,26 0800	8,447.83	18,357.0
83.	I'' NO OF RECS	Audit unit field:	Lin Internet	Cancel 33,802,5600	8,010.57	14,986.2
12	TI TOTAL S.M.		- 6	16,967,1800	7,982.65	17,375.3
B	TAXI SIM		THE REAL PROPERTY AND	38,350,5500	8,707.41	17,529.3
4	TAX2 SUM	Please select the audit units y Contrelation task on. The audi	ou wish to perform the	12,015,4800	7,937.23	15,804.
15:	TAX3_SUM	contained in the selected aud	it unit field.	36,774-0500	6,796.04	19,480.3
6	IV HO NETAMOUNT SUM			34,762,5100	6,266.37	17,857/
π.	PAYAMOUNT SUM SUM SUM	The second second second		\$1,825.5000	4,874.14	18,5551
£ 👘	IV CASH , ONLY , SUM	Trouble as most devel		12,571.0000	5,191.31	18,392
9	DEBLT_ONLY_SUM	Audit Units	Number of Records	\$3,378,7900	6,566.45	18,242.3
0	CREDIT_ONLY_SUM	1.0.000	and the set	30,849,6100	1,277.80	20,890.
1	I BLEND_OTHER_ONLY_SUM			30,065.7200	2,838.29	17,545.
2	CASH_COUNT_SUM			30,695.7000	2,991.89	14,363
¢:	DEBIT_COUNT_SUM			26,711.0900	2,032.10	13,803.
2	CREDIT_COUNT_SOM			25,979,9200	2,452.85	14,176/
8	E ALSO SCOOL SOM			26.687.0100	2,608.17	15,750.0
6	TOTAL CONT			26.852.3700	2,225.17	16,361.2
5	E CASH DESCRIPTION COUNT			30,432,7300	2,276.90	17,372.
6	E DEBIT CREDIT COURT			25,534,5600	1,539,91	15,300.3
7	DEBIT CREDIT PERCENTAGE COUNT			27,000.3100	2,217,86	26,184.
8		Output		25,703.5000	1,213.77	18,609.1
9		Create result		25,578.3700	2,185.24	16,438.8
0		Readtname		30,154,4900	3,309.31	35,727.2
1		Churcheline Solite and Ca	of and Pathot Tandar	.29,471.3600	2,127.99	17,500.1
2		Comparison saids and Ca		13,662,9100	1,219.50	\$8,095.
1		Create database		29,196.1500	3,541.13	16,700.2
2		File name:		27,206.2400	2,539.76	15,883.5
8		Correlation		27,882.6309	3,002.86	16,523.7
4				27,720.4600	2,914.14	15,388.4
15	Contraction of the local division of the loc		COLUMN 2	30,693,1600	2,940,43	38,127.1
6	1011 11.911.7300 1.415.710	2.253.9600 0.00	78.741.45	31,011.7300	2,451.38	18.662.0

Fig: 4.8.1.1 Applying the Advanced Statistical Analysis Correlation Feature of IDEA

Monthly Sales Summary.IMD	×		
■ \$ \$ • \$ • \$		CARL TO A COMPLEX THE	and a second second second
Fields : HD_NETAMOUNT_SUM, P	AYAMOUNT_SUM_SUM_SUM.	CASH_ONLY_SUM, DEBIT_ONLY_S	ÚМ
	HD_NETAMOUNT_SUM	PAYAMOUNT_SUM_SUM_SUM	CASH_ONLY_SUM
PAYAMOUNT_SUM_SUM_SUM	HD_NETAMOUNT_SUM	PAYAMOUNT_SUM_SUM_SUM	CASH_ONLY_SUM
PAYAMOUNT_SUM_SUM_SUM CASH_ONLY_SUM	HD_NETAMOUNT_SUM 1.000 0.706	PAYAMOUNT_SUM_SUM_SUM 0.710	CASH_ONLY_SUM

Fig: 4.8.1.2 Correlation coefficient results





While IDEA does not calculate the coefficient of determination, it can be done simply by squaring the correlation coefficient that is already calculated for you. The coefficient of determination, or  $r^2$ , tells us how much of the variation in one variable can be attributed to the variation of the other variable. This calculation when multiplied by 100 will be expressed in a percentage.

In our example of the correlation between cash tender and the payment amount of 0.710, 50.41 percent ( $0.710 \times 0.710 \times 100$ ) of the variation can be attributed to the other variable.

One has to be mindful that even if the variables are calculated to have a strong correlation, there may not be a cause-and-effect relationship.

Are there direct cause-and-effect relations? That is, does *x* cause *y* or, in our previous example, does the level of education cause the income level? Maybe it is a reverse cause and effect where *y* causes *x*. That is, income levels determine your education level.

Possibly there is a third variable or a combination of several other variables that caused the relationship, such as networking relationships while in school that resulted in higher paying jobs. Maybe the whole relationship between the two variables was just a coincidence?

## 4.8.2 TREND ANALYSIS

Trend analysis is based on regression analysis. Regression analysis produces a line of best fit and predictions can be made based on the line. It is also known as the least square line, because the line passes through the distribution where the distance squared from the line is minimized.

Similar to that of correlation, the x variable can estimate or predict the y variable. That is, if x changes, then how much y changes can be estimated. For two numeric variables, you can predict y from the x variable if the correlation coefficient is strong and there is a linear pattern for the variables. Normally you would want the *r* correlation coefficient to be better than plus or minus 0.60.

Unless there is perfect correlation, the prediction for the value of y, given x, is merely a prediction. It is a guess but an educated guess with some sound scientific basis. The prediction will be subject to some amount of error. The standard error of the estimate measures how much the predicted values deviate from the actual y values. IDEA uses the mean absolute percent error (MAPE) to calculate the accuracy of predictions. MAPE is the average of the percentage errors, is expressed in percentage terms, and works best with positive amounts. The MAPE number is the predicted line on the average that is away from the actual line in percentages. Low MAPE values mean that the past data has a good fit to the regression line and you can have more confidence in the data and prediction.

To use trend analysis in IDEA, the database needs at least one numeric field, and the field where trend analysis is to be applied cannot contain any bad data. The audit unit field cannot be the same field as the trend analysis field. In addition, the database should not contain seasonal data. If it contains seasonal data, then time series analysis should be selected over trend analysis. Time series works similarly to trend analysis.

In our data file, we will perform trend analysis on the 12-month data from the debit card payments field of DEBIT_2010_2011 and generate 3 months of forecasts. It is not necessary to provide a reference field or audit units. Refer to Fig 4.8.2.1.

For debit card payments, it is trending slightly downward, and it is predicted that at the end of the three months, debit payments would drop to approximately Rs. 18,000 for that month as seen in Fig 4.8.2.2.



In contrast to the debit card payments, cash payments are trending upward when we select CASH_2010_2011 as the field to trend.

There is less reliability in the prediction as the MAPE percentage is 23.17. It is predicted that at the end of the three months, cash payments would increase to approximately Rs. 14,500, as displayed in Fig 4.8.2.3.

We will perform a trend analysis showing a reference field and audit units. There are five branches and we will include all of them as audit units. The reference field of GLOBAL_AVERAGE_SALES is the average sales for the five stores broken down by months. We have eight years of data so each store would have 96 records (8 years × 12 months). We will generate forecasts for three months as shown in Fig 4.8.2.4.

YEAR_MONTH CASH_2007_2008 DEE 2007-10 10,489,12 2007-11 9,009,12	IT_2007_2008 17,428.56 19.998.04	CREDIT_CARDS_2007_2008 6,683.98 7,677.17	NON_CASH_2007_2 24,1 26,1	008 PAYTIME_DAT 112.54 255.21	E_YEAR PA 2007 2007
2007-1 Trend Analysis					
2006 2007 2008 2008 2008 2008 2008 2008 2008	Audit Units  Audit Units  Audit Units  Audit Units  Free Analysis  are consistent  Audit Units  Audit Units	st unit field tr The audit units you with to perform the In two existences and unit listed below at the existences and unit field. I audit units Rumber of Records	Forecasts Cenerate forecasts Start from Numbers of forecasts Dates Calendar: Calendar: Calendar: Calendar: Costor Start value: Increment: Costor result Result manie: Trend Analyses of Create MAPE dashe He name: Trend Analyses	E 12 E 3 E 1 E 1 E 2 E 2 E 2 E 2 E 2 E 2 E 2 E 2	OK       Preblut       Cancel       Help

Fig. 4.8.2.1 Trend analysis in IDEA



Fig. 4.8.2.2 Trend analysis results of debit card payments with forecast of three months





Fig. 4.8.2.3 Trend analysis results of cash payments with forecast of three months

1	Sales	- Global Sales I	MD ×						 
1 2 0	MONTP 2000-01 2000-01 2000-01	A B C	SALES 1,300,123.34 609,786.34 987,243.46	GLOBAL_A	/ERAGE_SALES 672,363.64 672,363.64 672,363.64				
4 5	2000-0	Frend Analysis							
2 8 2 8 8 8 8 8 8 8 8 8 8 8 9 1 1 1 1 1 1 1 1	2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 2005-4 20	Select trend fie RATE GOBAL AVER Reference Fit SALES SALES	kk: AGE_SALES AKK KKOE_SALES		Audit Units Select audit unit Select audit unit Audit unit field: BRAHOH Three are 5 audit or field. Select the audit Trend Analysis task I include all audit Audit Units B B C C C C C E E	field  to the selected audit unit it units you wish to perform the or.  Units  Number of Records 96 96 96 96 96 96 96 96 96	Porecasts Comparison C	recasts art/from 96 brecasts: 3 t value: 1 t value: 1 ft e: byps Sales cd Analysis database cd database E database	OK Felds Cancel Help
34 35	2000-4		1	Deselect			Trend Ana	Nysis	

Fig. 4.8.2.4 Applying trend analysis of sales referencing global average sales for each branch

Branch *A* outperformed the average of the five branches. The actual data is above the reference data line and shows good promise of trending upward in Fig 4.8.2.5. The prediction is sound as the MAPE is 5.26 percent.

By selecting all five branches as the audit unit, we can display other branches by choosing from the Audit unit pull-down menu. We will look at one more by selecting



#### Fig: 4.8.2.5. Result of branch A's trend analysis with three month's forecast

Branch *B* see Fig 4.8.2.6 fits right in around the global average sales and is also trending up.



Fig: 4.8.2.6. Result of branch B's trend analysis with three month's forecast





## 4.8.3 TIME SERIES

For data with seasonal values where there are higher values in certain months (or any other time units) and lower values in other months, the time series analysis option is more appropriate. In our example in Fig 4.8.3.1, we will use gas-heating costs as the field for the time series and all five branches as the audit units. We ensure that we input the correct season length of 12 as our records are broken down by months. We will generate 12 months of forecasts.



Fig: 4.8.3.1. Applying Time Series Trend Analysis of Heating Costs for Each Branch with 12 Month Forecast



Fig. 4.8.3.2 Result of branch A's time series analysis with 12 months' forecast

From the Fig 4.8.3.2, you can see that heating costs are higher in the winter months and lower in the summer months for branch *A*. The mean absolute percentage error of 11.40% is fairly reliable in terms of the forecast. You can also view the other branches by using the pull-down menus from



Trend analysis using IDEA is simple and the auditor need not be concerned with the complex formulas to calculate the regression line and the mean average percentage error.

# 4.9 IDEA SCRIPT

IDEA Script is a programming tool in IDEA that allows for combining numerous steps into a single procedure. The programming language is similar to that of Microsoft's Visual Basic for Applications. IDEA Script files are also known as macro files. It is a file that performs a series of actions when executed. IDEA Script can record actions for editing and modification to make the macro good for general use. Any user can apply the script to any suitable data files. IDEA Script is a powerful tool that speeds up repetitive procedures.

The tools built into IDEA eliminated the need to enter your computer codes from scratch. The record macro task can be used to record a sequence of steps that will be the framework for your application or macro. The Dialog Editor will help you design customized dialog boxes that you can use to control the flow of your application and to prompt the users for required input. However, IDEAScript cannot always anticipate all of the details you may want to incorporate in your macro so you will still have to write or edit some of the code.

## IDEA Script can be used for:

- Automating repetitive tasks: Regular tasks, such as monthly data that needs to be analyzed, can use IDEA Script to run and repeat the required procedures automatically. IDEA Script can be efficiently used when certain procedures are required to be applied in multiple locations such as departments, divisions, and branches. Another frequent repetitive task is to import files. This can be automated to select the input files, specify the output names, perform the imports, and perform required data cleanup or scrubbing.
- **Creating an automated analysis system:** A set of tests or procedures can be integrated into an IDEA Script where the user may select particular tests to apply.
- **Controlling other software packages:** Using Microsoft's object linking and embedding (OLE) technology, other OLE-enabled software can be controlled from within IDEA. One example is that of sending IDEA data into an Excel spreadsheet. IDEA Script performs analysis tasks such as summarization, sends the summarized data into an Excel spreadsheet, and then the IDEA Script instructs Excel to chart the summarized data into bar graphs.
- **Creating custom tests:** IDEA Script can be used to create specific tests based on the user's or the organization's needs. Typically, these would be performing calculations that involve comparisons and equations.

In using IDEA Script, work will be performed faster and the results will contain fewer errors. The analyses are done more consistently and can be undertaken by other staff members. Standards are adhered to by all users.

## **Consideration for Automation**

The following steps should be considered when planning to create an IDEA Script.

- Identify those tasks that are frequently repeated.
- Select those procedures that are well defined and well understood.





- Plan the macros carefully to ensure the results are those expected.
- Include user-input dialog boxes so that others can use the IDEA Script.
- Make the script as flexible as possible so it is useful to other users.
- Test IDEA Script on other computers with different versions of IDEA for compatibility and to ensure that the script works as expected.

#### Visual Script versus IDEA Script

Visual Script can create and edit macros in IDEA. Visual Script is a visual representation of batch processing that allows for simple dragging (or double clicks) of desired action choices from the left side of the window to the right side for project steps of the Visual Script screen, as shown in Fig 4.9.1.

Visual Script, like IDEA Script, allows for automation of repetitive tasks. However, the automation can be performed without the complexity of knowing any programming and writing code. There are limitations, such as not being able to create message boxes, nor will Visual Script run in the IDEA Server environment. Visual Scripts have the .vscript filename extension and are saved to the default Macros .ILB folder under the active project folder.

Visual Scripts can be later converted to IDEA Scripts, through an option in the Visual Script Editor. Because of the easy conversion to IDEA Script, you can start the steps for your macro in Visual Script, convert it to IDEA Script, and then code additional functionalities within the IDEA Script editor. Visual Script has the ability to include IDEA Scripts into it. This makes it simple to group together a number of IDEA Scripts and have Visual Script run them all. Visual Script is an excellent personal tool.

IDEA Script can use dialog boxes, control other software applications, and interact with the operating system. An example of the IDEA Script editor window with a sample script is shown in Fig 4.9.2.

	AC 1854 - 20	Untitled.vscript - Visual Script Ed	litor	×	- 5
Pectra Autor Bridge Balance	Ble Edt Bun B C H M M B C X O Ffin	Project Steps	Status	-	
Name         Rates           Process         Contract           Process         Process           Process         Process           Name         Research           Name         Research           Quarter         Process           Contineer Details         1,000           Contineer Details         1,000           Contineer Details         1,000           Dammeter         940           Matter Dok-Sup.         101           Program         195           Statistic Rosel         199           Statistic Roseling         101           Under wehrlingther Con         103           Vehicat Policy Mate         104		- Exployer Master		Pn 1 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	perties 9 Colubre Data Data Fitary Fital Statis Control Columnation Results Control Co
4 > >  C file fuptony i theny / >  Stanupul Popic Adv. (1) 4 0 2 0 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2	Consen     Servic     Morror     Morror     Visual Script		Net connected to	IDEA Server	Disk Speen: 245.07 GB

Fig 4.9.1: Visual script Editor Window


A Dar Hall E. No. 201 III To	Laseware IDEA		
Home Data Analyss	New Mono Swartkreiger		
1	IDEAScript	×	Properties
Elle Edit Debug Format Ins	ert Jools Help		II 🕶 Database
Project #  Project #	High Near Ancountsis . Sou Near Call DirectEntection, "Sample-Bank Transactions MD End Stop "Data: DirectEntraction Planta DirectEntraction Solid # - Clant DirectEntraction Clant DirectEntraction End Fundbal	Tribution	<ul> <li>Files</li> <li>Heary</li> <li>Field Statistics</li> <li>Centrol Total</li> <li>Chronis</li> <li>If ▼ Results</li> <li>If ▼ Indices</li> <li>If ▼ Convertes</li> </ul>
Kenty			
eged Project: Ade, ITT		Not connected to IDEA Server	Disk Space: 340.

Fig 4.9.2: IDEA Script Editor Window

## **Creating IDEA Scripts:**

The script in Fig 4.9.3 is an example of interacting with the Windows operating system function of renaming a file in a folder. This script does not include dialog windows to allow a user to select the folder and file to be renamed. It is hard-coded into the script that the "text.txt" file must be in the C:\data folder. It can only be renamed to "text_renamed.txt." The "CreateObject" function allows IDEA Script in IDEA to access Windows file-system functions and other application functions, such as in Microsoft Excel.



Fig: 4.9.3 IDEA Script to rename a file

An example of a dialog box is shown in Fig 4.9.4. The IDEA Script extracts transactions with even thousand amounts from the active file in IDEA. The script takes all numeric fields within the file and allows the user to select the field to apply the test on. In this case, the AMOUNT field is the appropriate one to select from the pull-down selection box.





Q	-	S → 1 32		a <del>y</del>			CaseWare IDE	A - Payments.II
		Home Dat	a Analys	s Vew Macros	SmartAnalyzer			
Rec Ma Rec	cro cro	Run Open E R Macro	Sind to IDEA Ribbon	Script Visual Script Create				
B		Payments.IMI	D×					
-		POSTED_BY	SUPPNO	SUPPNO SUPPNAME INVOICE_DATE_D/				AMOUNT
E.	1	MIA	10500	AASMAN DESIGN INC 21/01/2004			G34-567	1,000.00
pio	2	MIA	10900	ACKLANDS - GRAINGE	ER INC	01/01/2004	82 64 1	18,682.41
Ter	3	SOUB	11400	AG JO			99799ABC-123	2,713.16
0	4	MIA	11810	DKNY Mod Demi	o (Even Thousands)		WNZ28C	79,571.88
1	5	MIA	12203	ANKE			AZ278	54,721.86
	6	MIA	20028	BETO Select Field	to run test on		81340	75,373.66
	7	KSA	20041	BUOU	to full test off		971004A	6,844.63
	8	CW	20129	BILEVI			10000 A	4,522.28
	9	SOUB	20133	BISHC AMOUNT		-	100139	37,754.37
	10	CW	20133	BISHC			TJ9729	9,620.16
	11	CW	20204	BORE		ncel	000496CJW	64,592.51
	12	H.M.V.	20352	GUESS		lice	G34-568	1,000.00
	13	KSA	20403	BRIDD			2828 BNA	1,185.46
	14	CW	20414	CITIZE			566698T	79,237.49
	15	MIA	20508	FOSSI			L-1221/55	6,865.13
	16	MIA	20508	FOSSIL INC.		18/01/2004	14598	50.63

Fig: 4.9.4 Dialog Box for Extracting Records with Even Thousand Amounts

回 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)			CaseWare ID	EA - Payments.IN	1D		
File Home Data Analysis View	Macros SmartAnalyzer						
Record Run Open Bird to Record Hacro Core Alboon Coreate Coreate							
Payments.IMD ×							
1         Hda         1050         AdAMAN FOR           2         Hda         1050         AdAMAN FOR           3         SCUID         File         Edit         Run           3         Hda         File         Edit         Run           4         Hda         Hda         File         Edit         Run           6         Hda         Hda         Hda         Hda         Run           6         Hda         Hda <th colspan="3">Project Steps         Status           Project Steps         Status           Status         Status           Status         Status           Divect Steps         Status           Create area         Status           Status         Status           Status         Status           Status         Status           Status         Status           Status         Status           Status         Status           Divect Straction         Status           Create a vrbail diations         Status           Divest Straction         Status           Create a vrbail diations         Status</th> <th>04/01/2004 13/01/2004 13/01/2004 03/02/2004 03/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004</th> <th>AS3637 AS3637 AS3638 AS3630 AS3640 AS3641 AS3641 AS3641 AS3643 AS3644 AS3644 AS3646 Create Fields Fields Delete Clancel Help</th> <th></th>	Project Steps         Status           Status         Status           Status         Status           Divect Steps         Status           Create area         Status           Status         Status           Status         Status           Status         Status           Status         Status           Status         Status           Status         Status           Divect Straction         Status           Create a vrbail diations         Status           Divest Straction         Status           Create a vrbail diations         Status			04/01/2004 13/01/2004 13/01/2004 03/02/2004 03/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004 04/02/2004	AS3637 AS3637 AS3638 AS3630 AS3640 AS3641 AS3641 AS3641 AS3643 AS3644 AS3644 AS3646 Create Fields Fields Delete Clancel Help		
22         CW              ← Expertine                Compositive             ←	2 It any errors.			20,107.30 39,245.38	13/02/2004	A52670	

*Fig: 4.9.5 Visual Script for Extracting Records with Even Thousand Amounts* 

In contrast, when using Visual Script, as shown in Fig 4.9.5, the procedure is:

- Under File, select Open Database. All databases in the current project are offered for selection. Choose the file to open.
- Under Extract, select Direct. The direct extraction option opens up. Enter in the File Name area MOD Thousand. Enter the equation of (AMOUNT % 1000) = 0 in the Criteria area or Equation editor.
- By double clicking on the "Payments" filename under the Project Steps, all databases in the current project are offered again.
- By double clicking on Direct Extraction under the Project Steps, the Direct Extraction option opens up where you can modify the output filename and change the equation.

#### **CREATING MACRO:**

Visual Scripts are simpler to create but the user needs some knowledge to apply, such as being able to understand the equation and modify it. IDEA Scripts are more complex for the programmer but easier for the user. They are also more accurate if error checking is built in. Creating simple IDEA Scripts can be as easy as recording the steps that you are taking to accomplish a task. Fig 4.9.6 shows how you can start and stop recording a macro.

	Open Bind to Résen	NYSE W	su Meori	s Shart	kreiyası	Cust Pro			٥
ord Macr	v (Ctrl+R)	Mod 1000	× Olm	oices above	10000 × 🖉 sales abo	ve 15000 ×		+ X	Propertie
De Derin	d or stan recording a	ATE R	GION P	RODUCT	CUSTOWER TYPE	EXEC NAME	IN/ AMOUNT	^	1 - Data
FMCTO	L	Ser	ath Ma	ic Systems	Foreign	Shahrukh	43,430,00		V Datab
-		Ser	ath Air	Conditioner	Demestic	Shahrakh	72.712.00		1 04
3	5 02-04-2014	Ee	t Ma	tale Phones	Foreign.	Shahrukh	58,801,00		III Hird
4	6 03-04-2014	W	ut Ref	ignator	Foreign.	Salman	58,540.00		10 Eat
5	7 03-04-2014	Ear	t Air	Conditioner	Foreign	Shahrukh	42,443.00		52.8
6	8 03-04-2014	Ear	t. Ref	ignitor	Government	Rithik	25,252.00		Cor
7	9 03-04-2014	W	nt Ma	bile Phones	Domastic	Shahrukh	49,460.00		Tot
8	10 04-04-2014	See	ith Ma	ic System	Domestic	Aarrit	32,838.00		E Cit
2	11 04-04-2014	No	rth Mu	sic Systems	Domestic	Shahrukh	41,647.00		
10	12 06-04-2014	No	rth Mu	ic Systema	Domestic	Salman	32,288.00		I ≠ Result
11	13 06-04-2014	We	nt Air	Conditioner	Foreign.	Rittik	72,524.00		
12	14 07-04-2014	Eat	t Ref	ignator	Domestic	Aarrie	24,420.00		1 + Inder
13	15 08-04-2014	No	rth Mu	sic Systems	Foreign	Aarrit	49,808.00		I + Come
14	16 00-04-2014	Ear	t. Ma	bils Phones	Government	Aarrit	62, 178.00		-
15	17 09-04-2014	See	ath Ma	iic Systema	Foneign	Shahrakh	69,479.00		III AN
16	18 00-04-2014	Ser	ith Air	Conditioner	Demestic	Shahrukh	83,904.00		com
17	20 09-04-2014	No	rth Tali	ninian Seta	Foreign	Rithik	68,807.00		
.18	21 00-04-2014	W	st Ma	bile Phones	Forsign	Aarrie	58,734.00		
19	22 09-04-2014	No	rth Air	Conditioner	Foreign	Rittik	81,190.00		
20	25 11-04-2014	No	rth Ma	uic Systems	Foreign	Aarrie	60,267.00		
21	26 11-04-2014	No	rth Mu	uic Systems	Government	Rithik	75,006.00		
22	28 11-04-2014	No	rth Air	Conditioner	Government	Shahrukh	23,513.00		
23	29 12-04-2014	Ser	ith Ma	sic Systema	Foneign	Shahrukh	21,114.00		
24	30 12-04-2014	W	at Tab	nision Sets	Demastic	Aarrie	77,260.00		
25	31 13-04-2014	No	rth Mu	uic Systems	Government	Aarnir	43,733.00		
26	32 13-04-2014	We	st Ar	Conditioner	Desmeatic	Ratisk	60,985.00		
27	34 13-04-2014	No	rth Ma	sic Systems	Derrestic	Salman	59,890.00		
Tunning T urged Proj	anta (B) act: Ade, ITT						Not connected to IDEA Server   IDEAScript is n	uteting   Number of Records: 832   Disk	Space: 24
THE OWNER	State of the local division in which the local division in the loc	2775 B	ALC: NO.	3 10 M	and the second second	THE OWNER WHEN THE PARTY NAMES	A ROUND WAY DODG A DECKORADINATION AND AND A DECKARADING	ACCORDENT AND ACCORDENT ACCORDENT ACCORDENT.	Statistics of the local division of the loca

Fig: 4.9.6 Recording Macro

Start recording the macro. Perform the procedures and then stop recording the macro. A window will then pop up as shown in Fig 4.9.7.

l Carlos de la composition de	10° 34 c De	詳言語3+ ギ ta Analysis View	Mecros	Shartkreiver	Case	:Ware IDEA - salı	es above 1	5000		
Pass	n P	Top Records 🔛 Gap Detec Key Value 🔛 Benford's I	an Ou	plcate Key + SE So Strice + SO Ov	warization 🗊 Aging stification 🗊 Rivot Table et	금n John 비행 View Brit Append I II Com	ul Connector quint	計Attribute @Monetary Unit * 記Random @Vortables * 通 Othe *		
Tasks	Ebb	1 Sales Register X 🗍 🗃 Mo	Explore	Winnices above	Categoroe	e 15000 ×		Sergie	+ X	Desperties
In the second	1	NV NO   INVOICE DATE	REGIO	N PRODUCT	CUSTOMER TYPE	EXEC NAME IN	V ANOUNT		^	rioperues
- 3a	1	1 01-04-2014	South	A.A	Familie	a.d.d.	43.430.00	5	10	· · Databass
24.00	2	4 02-54-2014	South		Create a Macro	X	72,712.56	5		J Data
P L	3	5 02-04-2014	Eest				58.001.00			I History
9 11	4	6 03-04-2014	West	What have of more	for some south the second set	QK	58,540.00	D		III Fald
other Person	5	7 03-04-2014	East	what type of Paced	op you were to treater	All and a local division of the local divisi	42,443.00	b		Statisti
in the second se	6	8 03-04-2014	East			Cauca	25,252.00	5		Contro
	7	9 03-04-2014	West	) Ymail Script	11117		40,459.00	2		Total
Accoun	8	10 04-04-2014	South				12,838.00	D		· Criteria
Custor	9	11 04-54-2014	North		1000		41,647.0			
Loan m	10	12 05-04-2014	North				32,299.00	5		· · Results
- Master	11	13 06-04-2014	West				72,524.0	0		H Gen
- Master	12	14 07-64-2014	East				24,420.00	D		Detecti
Paymer	13	15 08-04-2014	North	. Travente	Secondary Dear		49,808.00	b		
False Re	14	16 09-04-2014	East		Clevingel in Organi		62,178.00	D		Indicas
SALESP!	15	17 09-04-2014	South		Care		68,479.00			I + Comme
Cill Uner and	16	18 09-04-2014	South				83,904.00	P.		
Webkal	17	20 09-64-2014	North		Make: then rey claffac	it option	68,807.00	b		III Add
Vehical	18	21 09-04-2014	West				58,734.00	D .		comm
	10	22 09-04-2014	North	Air Conditioner	Foreign	Rathak	81,190.00	2		
	20	25 11-04-2014	North	Munic Systems	Fanign	Aamir	60,207.00	0		
	21	25 11-04-2014	North	Music Systems	Government	Hahak	75,996.00			
	22	28 11-04-2014	North	Arconditioner	Covernment.	Shaheukh	23,513.00			
	43	29 12-04-2014	South	Willie Systems	ronign	shannuan	2011400			
	100	20 12-04-2014	North	Tenvinon ses	Committe	Aame	17,200.00			
1 3	20	22 12.54.2014	Week	Air Candidianter	Desentia	Raha	43,733.04			
36/6712/1	-17	28 12 54 3014	Number	Maria Castante	Demotis	Culmon	56,900,00			
			CIRCUS.	indire ay com	a strike the	Construct.	10,000,000		*	-
SRunning Lad	10.10									
Managed Project	t Ade. IT	τ					.N	lot connected to IDEA Server   IDEAScript is running   N	umber of Records: 832 Disk?	Space: 240.61 G
	1	- 10m - 11	6	No. of Lot, No.	1 12 10 10 10 10 10 10 10 10 10 10 10 10 10		2.00	Construction in the state of th		12402
0⊻	- 08	J 🖂 X 🗄	e						* Da 18 au 4a	AL OC 3817

Fig: 4.9.7 Macro option after ending recording of macro

You can select either Visual Script or IDEA Script. Selecting either one will open the appropriate editor window and automatically paste the actions or codes in. When procedures are performed in IDEA, they are logged in the history found under the Properties window. You can copy all the procedures into IDEA Script or selected tasks. In our example in Fig 4.9.8, we select the "Copy the IDEA Script for the selected task(s)" option since we right-clicked on the record extraction area to



bring up the selection. Once the selection is made, the IDEA Script editor opens and the script is automatically created for you.



Fig: 4.9.8 Creating Macro from file History

You can make changes to both the source and output filenames along with the equation from within the IDEA Script editor to suit you. Alternatively, you can add dialog boxes and coding to allow for user input. The dialog box shown in Fig 4.9.4 was created in the IDEA Script editor's dialogs area, displayed in Fig 4.9.9.

9	IDEAScrip	A CONTRACTOR PRODUCTION PRICE			1 and 1			
Ella Edit Dahua Loomat Incet Tools	- Here	•						
	(1211)[2]=	-						
Project 9 APDelay	APDelayinPaymentTo/lendor.tss / NewDialog / NewDialog1 / NewDialog2 /							
E Wew Scription	-		2 ti					
- Source Code	SELECT THE FIELDS NEEDED:			E Appearance				
- Dialogs	+ School Manuface Rill Banching			Title				
- Source Code	Number field		OK	Position     Lett 130				
Dislogs				Width 100				
- I NewDialog				Top 170				
I NewDialog2	2. Select the Vendor Bill Booking Date Reld.		Cancel	E Misc	÷			
				Dialog Tools	7			
	3. Select the Payment Amount field.			A Select				
	1			Gall CAncel Button				
				Button				
				Radio Button     Scheck Rox				
	4. Select the Nemation field.			de Static Test				
				ab) Edit Box				
¢ >								
Viston				= 130 y= 170 w= 100 h= 1	1			
event.				- 120 9 - 110 11 - 100 11 - 11				

Fig: 4.9.9 Dialogue Editor

# **4.10 IMPORTING FILES WITH REPORT READER**

The information you want to audit is usually available as a printout. In most cases, the printout can be routed or "spooled" to disc, and then transferred to the auditor's personal computer. Occasionally, it may be necessary to use a print file interceptor, which will redirect the print file being sent to the printer to a file on disc.

The differences between an ASCII file and a Print Report file are mostly in the header and footer, and in the extra lines and totals that you have to include or remove in order to import the file into IDEA.

Report Reader extracts data from plain text reports saved to file or some Adobe Acrobat PDF files and translates it into the specialized file formats of IDEA.

### **Step-by-step procedure to import file with Report Reader:**

- 1. Access the print report or .PDF file.
- 2. Create the Base Layer.

Report Reader uses a set of layers to define the data to be imported into IDEA. The first layer defined is the Base Layer. The Base Layer tells Report Reader to write a record for each occurrence of the information defined by the layer. The Base Layer must be defined at the detail line, or transaction level. A detail line is the line or lines that contain the information to be identified as a single transaction or record in the resulting IDEA database. The detail line can be a single or multi-line entry in the report.

IDEA will be prompt to:

- Create a standard layer: Select this option when the report is in columnar format.
- **Create a floating layer**: Select this option if the data is not displayed in columns.
- **Exclude it from the output**: This option is only available after the Base Layer is created. This option allows you to exclude lines that contain data not required for import.
- 3. Create an Append Layer.

An Append Layer captures information that is associated with the Base Layer and that may be shared by more than one record. Data defined in the Append Layer will be appended to the records defined in the Base Layer. For example, in a report with transactions categorized by account number, it is possible to have one account number associated with a group of transactions. The individual transaction data is captured in the Base Layer and the Append Layer captures the account number that applies to the group of transactions.

If the information in the Append Layer is located before the transactions in the Base Layer, it is called a Pre-Append Layer. If the information in the Append Layer is located after the transaction in the Base Layer, it is called a Post-Append Layer. For example, in the following report, if the first entry (12-5-2017, Raj Malhotra) was selected as a sample for the Base Layer, the sample for the Append Layer should be Account 1023 and it will be a Pre-Append Layer because it is located before the sample line for the Base Layer

Account 1023		
12-5-2017	Raj Malhotra	539.00
12-5-2017	Arun Dubey	235.00
Account 1024		
3-5-2017	Praveen Jain	326.75
4-5-2017	Rahul Gupta	434.50

- 4. Preview the resulting database.
- 5. Scan the file for errors.
- 6. Import the file into IDEA.





### **Exercise: Import Print file into IDEA**

Step 1: On the Home tab, in the Import group, click Desktop to begin the import process.

Step 2: In order to import reports into IDEA, select Print Report and Adobe PDF, then select the samplefile.txt file.

Step 3: Click Next. The Report Reader opens and displays the contents of the samplefile.txt file. The report will be displayed as in the following Fig 4.10.1



Fig. 4.10.1 Report reader Screen

Step 4: Scroll through the report and note the formatting within the report, including:

- General information (company name, report name)
- Report headings (range of accounts and period)
- Account information (number description)
- Detail information

Step 5: Creating a Base Layer

- Any detail line of the report can be used as a Base Layer 1.
- Drag your cursor across any detail line in the report to define it as your sample line.
- A message box appears, prompting you to create a standard layer or a floating layer
- Accept the default of Create a standard layer and then click Yes. Refer Fig 4.10.2
- The selected line is now copied to a section at the top of the page that is referred to as the Field Editor. Refer Fig 4.10.3.





Fig. 4.10.2. Selection of Layer Type

		x x x x 0 0 x x 1 0 1 x x 1 0 0 0 0 0 0				
1 0- 013				_	Field Details	
Þ	31	Cheque	2012-01-01	*	<b>₽</b> . <b>2</b> .	
				Â		
	1:30 PM Wy	vez Construction			1000	
	2012-12-31	Transaction Detail by	Account			
	Accrual Ba	sis January through De	cember 2012			
	Trans #	Type	Date			
	Chequin	g				
	31	Cheque	2012-01-01			
	43	Cheque	2012-01-24			
	155	Deposit	2012-01-28			
	156	Deposit	2012-01-30			
	32	Cheque	2012-02-01			
	157	Deposit	2012-02-16	+		
<				4		

### Fig. 4.10.3. Field Editor in Yellow

- The Field editor is the section between the two yellow lines with a red arrow in the left margin. In a multi-line entry, one of the lines must be the Anchor Line. The Anchor Line is highlighted with a red arrow in the left margin. By default, it is the first line, but a different line can be selected by clicking in the left margin of the line required to be the Anchor Line.
- The Field Details window also slides out from the right. When you define a field, this window will display the properties and attributes of the field.

Step 6: Define Trap:

• You must at this point define what distinguishes this detail line from the other lines in the report. It can be specific text, a combination of numbers, or a date. In Report Reader, these formats are

INFORMATION TECHNOLOGY





called Traps. For this report, you can see that the third field or column in the report is a Date field and it appears to be formatted the same way throughout the entire report (NNNN-NN-NN). Using this information, a trap can be defined in the Anchor Editor section (the line above the Field Editor).

• Place your cursor in the Anchor Editor directly above the first character of the date, just above the number 2. Since you want to import all the transactions (any entry with a date), click four times on the Numeric Trap button followed by a hyphen (-). Refer Fig 4.10.4.



Fig. 4.10.4. Anchor Editor

- Note that all the lines in the report that meet the criteria or follow the pattern are highlighted in blue. If you were to leave out the hyphen (-) you would end up trapping the header line above the field name row (Accrual Basis January through December 2012) because the 2012 lines up with your NNNN trap.
- The other options available to trap information are:
  - Text Trap
  - Space Trap
  - Non-Blank Trap
- You can browse through the report to ensure that all the detail lines and only the detail lines are highlighted. When satisfied that the required lines are selected, you must define the information you want to import into IDEA from the detail lines. This information will be brought into IDEA as separate fields.

Step 7: Define Fields:

- In the Field Editor (the row with the red arrow), highlight the Trans # information.
- The colour of the selected field will be orange and all the corresponding records of the detail lines in the report will be highlighted. Scroll through the report to verify that the field width

is sufficient to capture records with larger amounts of data and adjust the field width in the Field Editor if required. Refer Fig 4.10.5.



## Fig. 4.10.5. Define Fields

- In the Field Details window, enter the following details:
  - □ Name: TRANSNO
  - **Type:** Numeric
  - Decimals: 0
- In the Field Editor, highlight the Type information, leaving enough room for the larger data.
- In the Field Details window, enter the following details:
  - □ Name: TYPE
  - □ Type: Character
- In the Field Editor, highlight the Date information.
- In the Field Details window, enter the following details:
  - □ Name: DATE
  - Type: Date
  - □ Mask: YYYY-MM-DD
- In the Field Editor, highlight the Num information, leaving enough room for the larger data.
- In the Field Details window, enter the following details:
  - Name: NUM
- Type: Character (Note mix of numbers and character data, hence the Character type)

INFORMATION TECHNOLOGY

289



- In the Field Editor, highlight the Name information, leaving enough room for the larger data.
- In the Field Details window, enter the following details:
  - □ Name: NAME
  - □ Type: Character
- In the Field Editor, highlight the Memo information, leaving enough room for the larger data.
- In the Field Details window, enter the following details:
  - □ Name: MEMO
  - □ Type: Character
- In the Field Editor, highlight the Debit information, leaving enough room for a bigger number.
- In the Field Details window, enter the following details:
  - Name: DEBIT
  - **Type:** Numeric
  - Decimals: 2
- In the Field Editor, highlight the Credit information, leaving enough room for a bigger number.
- In the Field Details window, enter the following details:
  - Name: CREDIT
  - □ Type: Numeric
  - Decimals: 2

When you are done the fields should match the following table

Field Name	Туре	Parameter
TRANSNO	Numeric	Decimal:0
TYPE	Character	
DATE	Date	Mask: YYYY-MM-DD
NUM	Character	
NAME	Character	
MEMO	Character	
DEBIT	Numeric	Decimal:2
CREDIT	Numeric	Decimal :2

• Click the Save Layer button

Step 7: Creating an append layer

You now have to create an Append Layer to ensure you have an account name attributed to the correct detailed transaction information. In this report, you will be creating a Pre-Append Layer because the account name is located before the transactions of the Base Layer.

• Highlight the line under the report column titles that contains the account name, and then create a new standard layer. In the Anchor Editor, you need to make use of the Space Trap and

Text Trap buttons. The account names are indented a specified amount of spaces in this report so if you trap all the spaces up to when the text begins to identify the account name, you will eliminate any detail information you already pickup and header information in the report.

- Place your cursor in the Anchor Editor at the far left and click the Space Trap button to add spaces up to the start of the account names.
- Click the Text Trap button.



Fig. 4.10.6

- In the Field Editor, highlight the account name information.
- In the Field Details window, enter the following details:
  - □ Name: ACCTNAME
  - Type: Character
- Blank Cells: Use value from previous record.
- The field you just defined is located in the header which means that you would like to repeat the information it contains for each record (detail line) below it. To indicate to Report Reader that you want to repeat that information, the Use value from previous record option is selected in the Blank Cells box in the Attributes section of the Field Details window. Refer Fig 4.10.7
- Click save layer 🕑 button.





	SX M M B N S X B M H C					
SSSSSSSST				Field Details		J
Chegu	ing			10 A		
Torono a se			*	Layer Information	n	2
			Â	Layer Type	Standard	1
			-	Layer Name	Layer-2	
				Field Information	E.	
				Name	ACCTNAME	
1:30 PM	Wyvez Construction			Туре	Character	
2012-12-	31 Transaction Detail	by Account		Decimals	(none)	
Accrual	Basis January through	December 2012		Implied Decimals	(none)	
				Mask	(none)	
Trans #	Type	Date		Position		
				Offset	8	
Chegu	ing			Display Width	0	
3	1 Cheque	2012-01-01		Actual Width	28	
4	3 Cheque	2012-01-24		Attributes		
15	5 Deposit	2012-01-28		Blank Cells	Use value from p	re
15	6 Deposit	2012-01-30		Field On	NO Plank line	
3	Cheque	2012-02-01		Blank Cells	Diana inte	_
15	7 Depert	2012-02-01		Choose one of the o	ptions to fill blank	cells.
15	n neborro	2012-02-16				1919
15	o Transter	2012-02-16				
4	4 Cheque	2012-02-24				
	8 Bill Pmt -Cheque	2012-02-28				

Fig. 4.10.7 Attribute Section on right hand side

Step 8: Previewing the resulting database

• On the Report Reader toolbar, click the Preview Database button. The entire report will be displayed in a Preview Data Window.

	TRANSNO	TYPE	DATE	NUM	NAME	*	
1	31	Cheque	1/1/2012	1	Dobson's Exteriors		Bel
2	43	Cheque	1/24/2012	2	Insurance Experts		
3	155	Deposit	1/28/2012			Deposit	
4	156	Deposit	1/30/2012			Deposit	
5	32	Cheque	2/1/2012	3	Dobson's Exteriors		
6	157	Deposit	2/16/2012			Deposit	
7	158	Transfer	2/16/2012			Funds Tr	
8	44	Cheque	2/24/2012	4	Insurance Experts		
9	178	Bill Pmt -Cheque	2/28/2012	5	All Subs		
10	179	Bill Pmt - Cheque	2/28/2012	6	Blitz Marketing		
11	180	Bill Pmt - Cheque	2/28/2012	7	Roger's and Company CA's		
12	181	Bill Pmt - Cheque	2/28/2012	8	The Phone Co	Phone B	
13	182	Bill Pmt - Cheque	2/28/2012	9	Utility Corp		
14	239	Bill Pmt - Cheque	2/29/2012	10	Craig's Lumber		
15	240	Bill Pmt -Cheque	2/29/2012	11	Barata's Building Supplies		
16	270	Payment	2/29/2012	5454	Walker, Carol	Certified	
17			2 /1 /2012	111	n-L		

Fig. 4.10.8 Database Preview

- Click close to exit preview data window
- If you need to change information for a field, click on the field and make any necessary changes through the Field Details window. Changes can also be made by re-opening the layer for editing

through the Layer Manager.

• To access the Layer Manager, select Layers > Layer Manager. From the Layer Manager, you can edit, copy, or delete a layer. For more information on the Layer Manager refer to the Report Reader Help system

### Step 9: Importing into IDEA

- You are now ready to import the file into IDEA. Click the import button
- When prompted to proceed with the import, click Yes.
- Report Reader prompts you to save the template. Click OK.
- The Save As dialog box appears. You are prompted to save your template file .jpm in your current project folder. The template is similar to a record definition. It contains all the information you defined in Report Reader. You can reuse the template to import a similar report or you can open the template and make changes to it.
- In the File name field, enter sample file and click Save.
- The last import screen allows you to enter the name of the generated IDEA database.
- In the Database name field, enter name and click Finish
- Final output in IDEA format. Refer Fig 4.10.9.

$\square$	🕽 WC-TxnDt	I_by_Acct ×					▼ ×
	TRANSNO	TYPE	DATE	NUM	NAME	MEMO	*
1	31	Cheque	1/1/2012	1	Dobson's Exteriors		
2	43	Cheque	1/24/2012	2	Insurance Experts		
3	155	Deposit	1/28/2012			Deposit	
4	156	Deposit	1/30/2012			Deposit	
5	32	Cheque	2/1/2012	3	Dobson's Exteriors		
6	157	Deposit	2/16/2012			Deposit	
7	158	Transfer	2/16/2012			Funds Transfer	
8	44	Cheque	2/24/2012	4	Insurance Experts		
9	178	Bill Pmt - Cheque	2/28/2012	5	All Subs		
10	179	Bill Pmt - Cheque	2/28/2012	6	Blitz Marketing		
11	180	Bill Pmt - Cheque	2/28/2012	7	Roger's and Company CA's		
12	181	Bill Pmt - Cheque	2/28/2012	8	The Phone Co	Phone Bill	
13	182	Bill Pmt -Cheque	2/28/2012	9	Utility Corp		
14	239	Bill Pmt - Cheque	2/29/2012	10	Craig's Lumber		
15	240	Bill Pmt - Cheque	2/29/2012	11	Barata's Building Supplies		
16	270	Payment	2/29/2012	5454	Walker, Carol	Certified	
17	33	Cheque	3/1/2012	12	Dobson's Exteriors		
18	167	Payment	3/1/2012	112	Bentley, Karen:Roof and Attic	Certified	
19	261	Cheque	3/17/2012	13	Bank of Royals		
20	183	Bill Pmt - Cheque	3/18/2012	14	All Subs		
•		111					F

Fig. 4.10.9 Converted database

### **Reconciliation of data Imported**

Once the data has been defined and imported into IDEA, and before commencing audit testing, it is imperative that the data be reconciled back to the host system. There are several potential errors that may be made during the import process, such as:

- Requesting the wrong data
- Being supplied with the wrong data and/or for the wrong period
- Errors on the extraction of the data from the host system

INFORMATION TECHNOLOGY





- Problems with the media while transferring the data
- Errors when importing the data

The data can be reconciled back to the host system by reviewing Numeric field totals, the record count, or 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 that they are filed with the documentation from the download/import process.

There are different ways to reconcile the data. Certain IDEA tasks are particularly useful for reconciling data. First, it might be useful to browse the IDEA database. Use the Field Statistics and the Control Total properties, accessible from the Properties window, to reconcile totals and to ensure that the dates match the data you requested and do not include transactions from another period.

# **4.11 FILED MANIPULATION**

The Field Manipulation task is used to modify the layout of the fields in a database. There are two ways to access Field Manipulation:

• On the Data tab, in the Fields group, click the arrow in the bottom right corner



• Double-click any value in the Database window. The Field Manipulation dialog box appears.

	Field Name	Туре	Len	Dec	Parameter	Description	•	OK
3	SALESREP	Numeric	8	0				
4	CUSTOMER_NO	Character	5			Customer identification number	-	Append
5	PROD_CODE	Character	2					Delete
6	QTY	Numeric	8	0				Print
7	UNIT_PRICE	Numeric	8	2			E	
8	AMOUNT	Numeric	8	2				Сору
9	SALES_TAX	Numeric	8	2				Cancel
10	TOTAL SALE	Numeric	8	2				

### **MODIFYING FIELDS**

To rename a field double-click the field's name in the Field Name column. Any changes made to the field name will be recorded in the History. The information in the Type column is initially entered by IDEA during import. Four general types of fields: Character, Numeric, Date, Time and three special types of fields: Virtual, Editable and Multistate are available. You cannot change the length of a field or the number of decimals unless you have modified the information in the Type column. You can double-click in the Description column to change field descriptions

## EXERCISE 1

Make the following changes to the Sales Transactions-Database.

- Ensure that Sales Transactions-Database is the active database and that the Data property is selected in the Properties window.
- Double-click any record in the Database window.
- The Field Manipulation dialog box appears.
- Rename the CUSTOMER_NO field to CLIENT_NO.
- In the Description column for the SALES_TAX field, enter 8%.
- In the Description column for the INV_NO field, enter Invoice number.
- Click OK.
- Click Yes to make the changes to the fields in the database.

## **APPENDING FIELDS**

IDEA maintains data integrity upon import. The data cannot be changed. Instead, IDEA allows the creation of Virtual and Editable fields. For example, you can add a Virtual or Editable field that contains a calculation performed on an original data field

Virtual fields compute results "on the fly". The results are not stored in the database. However, any database created from a database with a Virtual field contains that field as a calculated value. Virtual fields can be Character, Numeric, Date, or Time fields.

Editable fields and regular fields are stored in the database and require some processing time in order to be physically added to the IDEA database.

## **EXERCISE 2: APPENDING VIRTUAL FIELDS**

Adding new fields might be useful when proving or verifying calculations, when creating a new field with calculated values, or when transforming existing data.

In this exercise, you will create a Virtual field called AMOUNT_CHECK in the Sales Transactions-Database. By creating a calculation, you will determine if the amount in the AMOUNT field is correct.

- Ensure that Sales Transactions-Database is the active database and that the Data property is selected in the Properties window.
- Double-click any record in the Database window.
- The Field Manipulation dialog box appears.
- Click Append.
- A new line is added to the Field Manipulation dialog box.
- Enter the following details for the new field:
  - □ Field Name: AMOUNT_CHECK
  - **Type:** Virtual Numeric
  - □ Length: n/a





- Decimals: 2
- □ Parameter: QTY * UNIT_PRICE
- Description: Calculated amount
- Click OK.
- □ If prompted, click Yes to make the changes to the database.

<b>I</b>	@ # <b>!</b> "	B¥2 C¥2		□ ┯				С	aseWare Il
File	Home	Data	Analysis	s Vi	ew	Macros	Smar	tAnalyzer	
ß		1	1?	₿				8	68
Clipboard	Index So	ort De Sorte Order	scribe ed Order	Append	Modify	Fields	Find	Action Field	Find *

Fig 4.11.1: Appending virtual field

$\square$	Sales Transac	tions-Database	×	<b>▼</b> ×
	SALES_TAX	TOTAL_SALE	AMOUNT_CHECK	NUM_AMOUNT_CHECK
1	23.96	263.56	307.20	307.20
2	1,105.82	12,164.00	11,058.18	11,058.18
3	9,175.98	100,935.78	91,759.80	91,759.80
4	508.31	5,591.40	5,083.09	5,083.09
5	457.80	5,035.80	4,578.00	4,578.00
6	83.26	915.87	1,063.35	1,063.35
7	1,000.00	11,000.00	10,000.00	10,000.00 👻
•			·	III F

## *Fig* 4.11.2: *Virtual field AMOUNT_CHECK*

Note the colour of the data in the AMOUNT_CHECK field, which indicates that this is a virtual field not part of original data but field appended by auditor

### **EXERCISE 3: APPENDING REGULAR FIELDS**

In IDEA, you can also add regular Numeric, Character, Date, and Time fields through Field Manipulation. These fields will appear as if they were added during import. They will retain the same colour as the other fields and will not have any parameters displayed. As with Virtual fields, any additions or changes made to these fields will be recorded in the History. The parameter of a regular field cannot be changed.

In this exercise, you will create a Numeric field that is identical to the AMOUNT_CHECK Numeric field.

• Ensure that Sales Transactions-Database is the active database and that the Data property is selected in the Properties window.

- Double-click any record in the Database window.
- The Field Manipulation dialog box appears.
- Click Append.
- Enter the following details for the new field:
  - □ Field Name: NUM_AMOUNT_CHECK
  - **Type:** Numeric
  - Length: n/a (by default it will become 8)
  - Decimals: 2
- Parameter: AMOUNT_CHECK
- Description: Calculated amount
- Click OK.
- Click Yes to make the changes to the database

	Field Name	Туре	Len	Dec	Parameter	Description		ОК
5	PROD_CODE	<u>Character</u>	2					( Amound
6	QTY	Numeric	8	0				Append
7	UNIT_PRICE	Numeric	8	2			-	Delete
8	AMOUNT	Numeric	8	2				Print
9	SALES_TAX	Numeric	8	2		8%		
10	TOTAL_SALE	Numeric	8	2			=	Сору
11	AMOUNT_CHECK	Virtual Numeric		2	QTY * UNIT PRICE	Calculated amount		Cancel
12	NUM_AMOUNT_C	Numeric		2	AMOUNT CHECK	Calculated amount		

Fig 4.11.3: Regular field NUM_AMOUNT_CHECK

Note the colour of the data in the NUM_AMOUNT_CHECK field, which indicates that this is a regular protected field.

• Re-open the Field Manipulation dialog box.

	Field Name	Туре	Len	Dec	Parameter	Description	•	OK
5	PROD_CODE	Character	2					
6	QTY	Numeric	8	0				Append
7	UNIT_PRICE	Numeric	8	2				Delete
8	AMOUNT	Numeric	8	2				Print
9	SALES_TAX	Numeric	8	2		8%		
10	TOTAL_SALE	Numeric	8	2			E	Сору
11	NUM_AMOUNT_C	Numeric	8	2		Calculated amount		Cance
12	AMOUNT CHECK	Virtual Numeric		2	QTY * UNIT PRICE	Calculated amount		

Note that the parameter is not displayed for the NUM_AMOUNT_CHECK field.

• Click OK to return to the database.

## **EXERCISE 4: APPENDING EDITABLE FIELDS**

IDEA also allows for the creation of Editable fields. These fields will allow you to create a notation, or manually enter a value to add to a database. The creation and editing of these Editable fields is recorded in the History. Editable fields are also displayed as a different colour to distinguish them

INFORMATION TECHNOLOGY





from Virtual and Regular fields.

- With Sales Transactions-Database as the active database, append an Editable Character field to the database in order to add notes.
- You will need to add an empty value (represented by double quotes) in the Parameter column so IDEA knows you want to enter characters into the field at a later time.
- Enter the following details for the new Editable field:
  - □ Field Name: NOTES
  - □ Type: Editable Character
  - Length: 50
  - Decimal: n/a
  - Parameter:
  - Description: Notes on line items

Note that the Editable field created will wrap the text within the field as you type in order to display everything in the current column width.

## **EXERCISE 5: APPENDING SPECIAL FIELDS**

IDEA further allows for the creation of special fields. These fields are termed as Boolean and Multistate fields. They are used generally by auditor to provide remarks in sample file in a symbolic way. Refer Fig 4.11.4 to see types of fields options provided by IDEA



Fig. 4.11.4 – Types of fields

Boolean fields offer two options to user whereas multistate field offers three options. Refer Fig 4.11.5 below.



Capy Index Sert D sert Order	t? esorbe ed Deske	Append	Hadify Resource F		the shad	Find Go Search Ind To Search						
ile Explorer #	10	Account	Details × 🗑 Co	mbined de	talls visual conn	× Sales Register	х				- w X	Properti.
四二 二 四二 二 四二 二 四二 四二 四二 四二 四二 四二 四二 四二 四	1	NV_NO	INVOICE_DATE	REGION	PRODUCT	CUSTOWER_TYPE	EXEC_NAME	INV_AMOLINT	BOOLEAN	MULTISTATE		11 * Datal
22 MB	1	1	01-04-2014	South	Music Systems	Foreign	Shahrukh	43,430.00	~	V		£ 0.4
	2	2	01-04-2014	West	Refrignator	Domestic	Aamir	7,553.00	34	x		in He
esktop Project	2	1	01-04-2014	West	Music Systems	Government	Salman	5,724.00	34	2		In Fed.
Varna	4	4	02-04-2014	South	Air Conditioner	Domestic	Shahrukh	72,712.00	*			Sta
Benford Sect A	5	5	02-04-2014	East	Mobile Phones	Foreign	Shahrukh	\$8,801.00	**	4		E Co
Berford Sect	6	6	03-04-2014	West	Refrignator	Foreign	Salman	\$8,549.00	36	35		= 01
Benford Susr	7	7	03-04-2014	East	Air Conditioner	Foreign	Shahrukh	42,443.00	*	7		- Cit
Berlord Susp	1	8	63-04-2014	East	Refrignator	Government	Rithik	25,252.00	34			▼ Resu
- Benford Susp	2		03-04-2014	West	Mobile Phones	Demastic	Shahrukh	49,460.00	~	V		E Ga
Benford Susp	10	10	04-04-2014	South	Music Systems	Domestic	Aamir	32,838.00	34	22		Det
Combined d	11	11	04-04-2014	North	Music Systems	Demastic	Shahrukh	41,647.00	~	2		II windie
VisCon	12	12	06-04-2014	North	Music Systems	Domestic	Salman	32,258.00	26			
Customer Detail	13	13	06-04-2014	West.	Air Conditioner	Foreign	Fithik	72,524.00	24	4		A No
Loan master	14	14	07-04-2014	East	Refrignator	Domestic	Aarrir	24,420.00	36	2		E PRI
Master Data-Em	15	15	06-04-2014	North	Music Systems	Foreign	Aartúr	49,856.00	~	7		I REC
Paymert-Payme	18	15	00-04-2014	East	Mobile Phones	Government	Aarret	62,178.00	*			•
Payroll March 2	17	17	09-04-2014	South	Music Systems	Foreign	Shahrukh	69,479.00	~	Y		# N\
Sales Register	18	18	00-04-2014	South	Air Conditioner	Domentic	Shahrukh	83,904.00	~	×		∎ Đž
SALES95	19	-19	00-04-2014	East	Music Systems	Foreign	Aarrie	5,479.00	34	7		I INA
User_web_log-D	20	20	00-04-2014	North	Television Sets	Foreign	Rithik	63,807.00	ж			- PO
3	21	21	09-04-2014	West.	Mobile Phones	Foreign	Aarnir	\$8,734.00	4	10 2451		
File Es. Library	22	22	09-04-2014	North	Air Canditioner	Foreign	Rithik	81,190.00	~			🖶 CU
SAurring Tasks		-						Not	connected to IDEA	Server (	Number of Records: 999   Disk	Space: 240.5 (9.41

Fig. 4.11.5 – Boolean and multistate fields

# 4.12 JOIN DATABASES AND VISUAL CONNECTOR

IDEA has two main methods of joining databases together based on a key field:

- Join Databases
- Visual Connector

## JOIN DATABASES:

Data required for auditing is often spread across several files and perhaps across different computer systems. In order to carry out most audit tests and analyses, the data must be contained within a single file. IDEA provides a Join Databases task, which can be used to:

Combine fields from two databases into a single database for testing. For example, to carry out an inventory valuation, data is required from both the inventory master and pricing information databases.

Test for data, which matches or does not match across systems. For example, match the monthly payroll transactions database with the master payroll database to ensure there are no "ghost" employees, or that all former employees have been removed from the payroll.

Databases can be joined or matched if they contain a common link (referred to as the "key"), such as an employee number field.

Join Databases can perform a host of different join option on two databases, including all records in primary database, matches only, records with no secondary match, records with no primary match, as well as all records in both databases.

### Join Databases Precautions

- Only two databases can be joined at one time. If more than two databases must be joined, you must combine two databases, then join a third to the resultant database, and so on. You should also consider the Visual Connector task if you need to join more than two databases.
- The common fields (the "key") do not need to have the same name, but they must be of identical field type.

INFORMATION TECHNOLOGY





- The databases must be in the same location. For example, a database in the Samples project cannot be joined to a database in the test project.
- When using the Join Databases task, care must be taken to select the primary and secondary databases in the correct order as the secondary database is joined to the primary database. Transaction file should be the Primary file and Master file should be the secondary file. There should be Many to one relationship.

### IDEA provides five join options:

### Matches only



In this option, IDEA processes each record in the primary database (in "key" order), cross matching with the secondary database on the key. If there is a match on the key, a record is written to the output database. If there is no match, the next primary database record is processed. Matches only will search for the keys in the secondary database for matches against each record.

### Records with no secondary match



In this option, IDEA processes each record in the primary database in turn, cross matching with the secondary database on the key. It writes a record to the output database if the key exists in the primary database only.

### Records with no primary match

IDEA processes each record in the secondary database in turn, cross matching the key with the primary database. IDEA writes a record to the output database if any key exists in the secondary database only.

### All records in primary file



IDEA processes each record in the primary database (in "key" order) in turn, cross matching with the secondary database on the key.

Each record in the primary database is written to the resultant output database including the selected fields from the secondary database if the key exists.



#### All records in both files



In this option, IDEA processes the database in key order, writing a record to the output database for every record in each database where there is a key match.

If there is more than one record for a particular key in both the primary and secondary databases, each of the primary database records are matched against the first record for the key in the secondary database as in the options Matches only and All records in primary file. However, all additional records for the key in the secondary database are also written to the output database but all primary fields will be blank for Character fields, 0000/00/00 for Date fields and zero for Numeric fields.

### JOIN Vs VISUAL CONNECTOR

Join Databases has two important limitations:

- It can only join two databases at a time.
- The Matches only option is similar to Visual Connector but has the following limitations:
  - □ When a many-to-many relationship exists, only the first matched record from the secondary database is included in the result.
  - □ The resulting joined database does not exceed the number of records in the primary database, regardless of the matching ability of the key fields. In other words, there is only one primary match for every secondary match. With the exception of matching types such as all records in secondary database and all records from both files that can and will, respectively, have more records than the primary database.
- On the other hand, Visual Connector is an easy-to-use, visual method of connecting two or more databases together by drawing lines between key fields on a canvas, and does not have any limitations on how many records are joined.
- Visual connector can join more than two files at a time
- Visual connector unlike Join option which has five different options has only two options matches only and all records in primary file

## EXCERCISE:1 - Join Customer details.imd with account details.imd

Account details file has fields – ACCNO, STATUS, BR_CODE, BALTODAY, BALYESTDAY

Customer details file has fields - ACCNO, CUST_NAME, CUST_ADD, LOCALITY, STATE

Join both the file using common field ACCNO

Step 1: Open primary file account details and ensure it is active file in IDEA

Step 2: Go to analysis tab and click on join button





Hone Hone	Ce	ta: Analys	sis Viet	e Macros	Snartkreiv	15						a (
N Bactor	) (法	Tigt Riscords Key Value	篇 Gap D 到 Tarfs	intention 🛅 D. Intention 🛄 Si	plcate Key + aterica +	SE Summarization (1) John Statistication (1) Priot Table (2) Overt	불= Join Mil Vacal Connector 봄-s Appartel 또 Connector	HAmbute Elitandon	Monetary Unit *			
93	Đ01			Explore		Categoriae	Relate		Sarpie			
Expla., #	18	Account D	letals ×								* X	Properti
n Ha		ACCNO		BR_CODE	BALTODAY	BALYESTDW'					.C.	11 - Databa
1.510		43808686		65	473.53	473.53						
194		\$3207396		-48	11,814.52	12,236.48						i Data
t mar i		43808500 /	N	87	596.91	432.94		inie i	Tetalence:	*		H He
1000	4	72907580		45	43,35	13,35		2010-1	Varia Diapez.	mildi		E Fold
top Project	3			79	2,169.67	1,916.50				0		State
THE	- 3	B2130e68 L		11	1.5.0.48	1,156,45	Prevery database: Acc	ount Details	Pields	0		Total
Account		82120005 1		-45	33.23	53.23	Number of records: 300	10		Cancal		10 Cal.
Custome		81618305.1		77	6,023.37	6.351.09	Otteria:		100			in chin
Loan ma	- 10	210/2145			1,750,79	0,231,300			1.0000	Help		T Result
Master D		31673000		90	1,620,34	2.247.00			10000			
Master D		713003466		25	40.12	2512	Secondary database:		select			I + Indica
Payment		62213682		36	258.64	1 452 30	Number of records:		Paide			
Payroll_A	34	0513505		32	15 000 40	10 712 80						li a como
Sales Reg		72232216		41	615.77	995.68	Nicosene: Inter Database		Madela			···· Com
SALE595	16	82729400 0	91	88	96.56	68.56						Add Miles
User_wet		85205003		-51	390.84	361.34	Matches only		All records in primary file			com
Vehical C	C1E	12354887		32	775.60	1,041,65	C Records with no second	dary match	All records in both files			
Vehical P	10	3430395		51	22.65	34.90	O Records withing primary	visith				
		32006868		36	102.14	48.62						
		92702604 8	4	.67	186.52	160.02	Create a vehial databa					
		\$2120005		37	416.72	1,333,24						
		2608199		91	290.94	290.94						
	24	22099996		87	1.000.00	1,517,121,70						
		92013884		97	74,326,13	93,176.12						
COLUMN TWO	21	13205004 /	NC N	22	343.75	335.87						
	27	23306290		70	12.01	10.62						
Running Tasks	100											
neged Project .	Ade, IT	τ.						Note	onnected to IDEA Server	Number of Ru	icondic 1,000   Disk:	Space: 240.56
Color Street		100	1	. 6	ALC: NOT THE	the second se						58.50

Fig: 4.12.1 – Join options

Step 3: Primary database automatically selected is Account Details. Select Secondary database customer details

Step 4: Select option 'All records in both the files" since we want to join both the files

Step 5: Provide file name as 'combined customer details". Refer Fig 4.12.2

310 PH	1° N	第三日:	2- <b>7</b>			Ca	seWare IDEA - Account Det	ails					- 6
Tim Hone	- 01	Analys	sis Vie	w Mecros	Snatkrely	35							~ <b>0</b>
Refut 2-255	1 If 11 A	Top Records	illi Gert Si teris	Intectors (20) onto Loss (20) Contenes	apicate Key + attrica +	SESumerication III Aphy Sistetification III Priot Table III Ovint	Service MS Valuel Deresector	H Attribute	Pronetary Unit *				
The Post of H	1	T Account D	winter w	1			10016					* X	
		ACICNIC 43808686	<b>ATATUS</b>	ER_CODE 63	BALTODAY 473.53	84LYESTDW 473.53						â	II * Databas
回日	2 11 4	43806599 / 72967596	N.	40 87 45	11,814.52 596.91 43.35	12,236.40 432.04 13,35		Join D	atabases				E Heter E Field
Name Account	87	82130688 0 82120005 1 82429892	2	70 10 40 81	2,169.65 1,320.40 53.23 2,725.57	1,155.45 53.23 1,714.35	Prenary database: Acc Number of records: 300	ount Details 10	D	ita O	al		Total
Loan ma B Master D Master D	9 10 11	83628205 1 42917103 32622000		77 96 82	6,351.00 1,750.21 1,630.10	6,351.98 4,550.21 3,247.90	Otteria:	rkarser Details		He	0		🛛 🕶 Results
Poyment Poynol JA	12 13 14	73707296 62213602 2513596 7233236		73 36 37	40.12 2,543.64 10,999.41	40.12 1,953.32 10,712.89	Number of records: 300	10	n	Ada			······································
	10	82729409 0 81205003 12354887 0	0	88 51 50	96.50 390.64 775.60	68.56 361.34	Records withins accord	derg metch	) 셴records in prima ) 셰records in prima	ry file Nea			E Add comm
- J Vehical P	10 20 21	2430395 32506868 92702694 /	4	51 36 67	22.65 103.54 166.52	34.90 48.62 160.02	O Records withing griner	y match					
	22 22 24	\$2120005 2608199 22000008 (		37 91 87	416.72 290.94 1,000.00	1,333,24 290.04 1,517,121.70							
	25	\$2013884 13205964 7 23306299	¥C.	97 32 70	74,336,11 343,75 53,91	93,176.12 335.87 53.01						4	
Samering Tau	Adre I							Net co	reserved to DFA Se		Marridge of B	unerter 1.000 à Dirk	Same Mill 56 GR
02	(	9	a ×	ı e	wB	3		That Co				- Nr 18 al 4a	09.01 07-06-2017

Fig. 4.12.2 – User Input to join options

Step 6: Click on Match option

Step 7: Provide primary and secondary key which is common in both the files ACCNO. Refer Fig 4.12.3



Tards         Direct         Extract         Color         Direct         Color         Mage           Ref Explor.         ***         Account beaks         ***         None         ***         Page           Ref Explor.         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.*         ***         Account beaks         ***         Page         ***         Page           Ref Explor.*         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.*         ***         Account beaks         ***         Account beaks         ***         Page           Ref Explor.*         ***         Account beaks         ***         Account beaks         ***         Page         ***         Page         ***         Page         ****         ****	ena protect	新見	Tigi Rakorda Key Value	illi Gan D Sil terfo	ninten 🖪 D. refetare 🗐 D	picate Key + stelica +	SE Summarization (10) Aping Statestication (10) Privat Table (20) Overt	Seller Mill Vale Bei Appartit S Care	l Caresector are	H Attribute			
Care Expl., *         Concordinations         Concordination         Concordination         Concordination	Taska I	BUD			Explore		Categorae	Relate		Serple			
Image: Second	ile Expla., #	18	Account D	etals ×			and the second			kin Da	tabases	100	• × Properti
1         4         4         473.3         473.3         473.3           1         4         4         173.3         473.3         473.3         473.3           1         4         4         173.3         473.3         473.3         473.3           1         4         4         173.3         473.3         473.3         473.3           1         4         4         173.3         473.3         473.3         473.3           1         4         4         173.4         123.6         4         174.6         6         6         6         6         6         6         6         6         6         6         7         6         6         6         7         6         7         6         7         6         7         6         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7 <th7< th=""> <th7< th=""> <th7< th=""> <th7< td=""><td>- PL</td><td></td><td></td><td></td><td>ER_CODE</td><td>BALTODAY</td><td>BALYESTDAY</td><td></td><td></td><td>1011110</td><td>an energy and</td><td></td><td>1 T T Databa</td></th7<></th7<></th7<></th7<>	- PL				ER_CODE	BALTODAY	BALYESTDAY			1011110	an energy and		1 T T Databa
Image: Source Market Source         Source Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source         Market Source	- 22		43808686		65	473.53	473.53					OK .	
Image: State of the	28	2	53207306		48	11,814.52	12,236.40		Prine	vy database: Account Details	Eelds		A Data
Bit	100		43808500 N		87	596.91	432.94		Nambs	et afrecarda: 3000		Cancel	iii Hetz
Distribution         Distribution<	9 ==	4	72967580		45	43,35	13,35		OIN	isia	10	-	ili Fald
Name         1         21/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         0         1         12/2881         1         0         0         12/2881         0         1         12/2881         0         1         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         <	inktop Project	3	2719391		79	2,169.67	1,\$16.50					map.	Statis
Addate         7         61/300         1         4         1.3.3         1.2.3         PRAVENT of State of State         Field           Content         1         22.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2	Name	8	82130608 0		- 11	1,320,48	1,856,45		10000	and and a second second second	Salart		a Contr
Control         Mathematical State         Mathematical Stat	n Accourt		\$2120005 1		-49	5123	\$3.23		Seconde	ry databaset custories becaus			Total
Listant         J         Listant         Listant	Custome	1	82429892		81	2,725.57	1,704.35		rames	er di necercii: Jului	Fields		M Criter
All 197     All 197	- Loan ma	9	82628205 1			6,351.04	6,351.98						T Reads
■ Match         ™ Match         <	E- Master D		42917193		90	1,790,21	4,650,21		File rat	me: combined customer details	Match		
Provide (1)         201208         31         34,45         11,15         Provide (1)         Of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1) of the and (1)	- Master D		32522000		32	1830.15	3,247.309				All contracts in columns (ife		1 miletion
Image: Property in the second with the	- Payment		63113603		13	3504	1 0 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		(C) Pana		Distances of heater A res		
State         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0         0 <td>- Payroll_A</td> <td></td> <td>02110002</td> <td></td> <td>- 30</td> <td>45,000,45</td> <td>1,833,56</td> <td></td> <td>Reco</td> <td>ids with no secondary match</td> <td>All records in goth files</td> <td></td> <td></td>	- Payroll_A		02110002		- 30	45,000,45	1,833,56		Reco	ids with no secondary match	All records in goth files		
Image: Second and the secon	🗉 🗐 Sales Reg 📄		T113296			615.77	545 G0		Recor	da with no grimary match			···· Comm
Image: Note of the secondary index of the se	- SALES95	16	87736495 0			91.4.77	68.56			Match Key	Fielde -	× 1	- Add
Openant Debts/R         111         202400 (I)         102         77.64         (1941)(6)         Primary         Odar         Secondary         OK           20         30000 (I)         30000 (I)         312         22.64         3400         Accinity         Accinity         OK           30         30000 (I)         31         23.64         3400         Accinity         Accinity         OK           30         30000 (I)         31         23.64         3400         Accinity         Accinity         OK           30         30000 (I)         31         31.64         44.02         Accinity         Accinity         Accinity         OK           32         30000 (I)         37         416.22         1303.64         Accinity	- User_wet		#1205003		-51	200.64	36134			HIBOLD IVEY	(Velius)		comr
Ling         Velocal         110         Accord         Accord <td>- Vehical C</td> <td>TE</td> <td>32354897</td> <td></td> <td>22</td> <td>775.68</td> <td>1.041.65</td> <td>p</td> <td>rimany:</td> <td>Order</td> <td>Secondary</td> <td>OK</td> <td></td>	- Vehical C	TE	32354897		22	775.68	1.041.65	p	rimany:	Order	Secondary	OK	
30         200666         36         00.44         45.0         20067           21         200666         27         45.02         45.02         45.02           22         47.03661         27         45.52         15.002         Grand           22         47.03661         27         45.52         15.002         Grand           23         200666         17         45.52         15.002         HMB           24         200666         17         10.002         17.17.02.70         HMB           24         200666         17         10.002         10.77.02.70         HMB           21         10.0044         10         10.33.00         10.17.02.70         HMB           21         10.0044         10         10.33.00         10.17.02.70         HMB           21         10.00444         10         10.33.00         10.17.02.70         HMB           21         10.00444         10         10.33.00         10.17.02.70         HMB           21         10.004444         10         10.39.01         HMB         HMB	Vehical P	10			51	22.65	34.90	ACCN0 (N)		Ascending	ACCNO (N)	v	
21         127/10244         0         0         160.0         Cenal           22         1200864         10         200.04         200.04         100         100           101         2000064         101         200.06         101         100.00         100           22         2000064         101         100.00         101,101         100         100           4         22         100.004         101,101         100,000         100,101         100           4         22         100.004         101,101         100,000         100,101         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100,000         100			32006868		36	102,54	48.52					Delete	
22         \$10006         37         465.27         \$1003.4           23         \$20006         91         300.4         200.4         200.4           34         \$200064         107         90.04         200.4         200.4         200.4           34         \$200064         107         90.00         57.10         107.10         140           37         \$1001044         107         \$100.20         \$107.10.70         140         140           37         \$1001044         107         \$100.20         \$107.10.70         140         140           37         \$1001044         107         \$100.20         \$107.10.70         140         140           37         \$1001044         107         \$100.20         \$107.10.70         140           37         \$1001044         108         \$10.07         \$100.20         \$100.20         \$100.20           38         \$100.20         \$100.20         \$100.20         \$100.20         \$100.20         \$100.20         \$100.20			\$2702604 F		.67	186.52	160.02					Carcel	
101         360390         01         300.54         240.54         140           > 22         100104         07         36.25         161.75           < > 22         100104         07         36.25         161.75           < > 22         100104         07         36.25         161.75			\$2120005		37	416.72	1,333.24						
3#         20000001         37         0000000         (17) (12) 70           >         >         300000000         107 (12) 70         10000000         10000000           >         320         100000000         107 (12) 70         10000000         10000000         10000000           QUE         320         100000000         10000000         10000000         10000000         10000000           QUE         320         100000000         100000000         10000000         10000000         1000000000000         1000000000000000000000000000000000000			2608199		91	290.94	290.94					Help	
4         32         101014         97         34,56,10         91,06,12           276         132,05644         32         34,35         33,87           276         272         132,0564         32         34,35		24	22099998 1		却	1.000.00	1,517,121.70						
28 10066 N 32 3635 3157 27 2056 7 13 131 501			92013894		97	74,326,13	93,176.12						
1. 1 22262W 70 124 5.01	And Intellige	21	13205004 N		22	343.75	335.87						
		27	23306290		70	3.64	\$1.0t						

Fig. 4.12.3 – User Input to match option

### Step 8: Click on OK

Step 9: See the output in combined customer details file. Refer Fig 4.12.4. you would observe that joined file contains all the fields from customer details and account details files

Povel	(F) (元)	Fop Records Gry Value	E Cap C	letection 🗐 Du rdfe Law 🔆 Sta	plicate Key + atetica +	BESummerication [] BEStratification [] Diart	] Aging ] Privot Table	문제 Append 프 Company	ter HAttribute III Hone ERandon III Varial III Other •	tary Unit + Nes +			
2492	Extrac			Explore		Categora		Relate	Sarple				
le Expla., #	10	Account D	etais ×	Combined	customer del	talk x						X	Properti
- 69		ACCNO	BTATUS	BR_CODE	BALTODAY	BALYESTDAY	ACCN01	CUBT_NAME	CUST_ADD	LOCALITY	BTATE		II w Databa
- 22	1	2004190		95	1,218,21	P 1,218.29	2004190	FRANCIS J. TAYLOR	140 1 AVE.	Malad	Maharantera		
24 (24)	2	2006393		73	7,449.20	6,753.16	3006313	WILLIAM WIECKOWSKI	43-191 67 AVE.	Mahalasmi	Maharashtra		J Data
	1	2008092		85	231.44	4 231,44	2008092	JOHN MAXWELL	<b>B60 KENSIGNTON AVE.</b>	Andheri	Maharashtra		E Histo
9 11	.4	2010999		87	8,509.00	8,500.00	2010000	DICK SMITH	151 MAGNUS AVE.	Rajkot	Gujarat		E Feld
rabing Depart	5	2014595		94	746.00	0 746.00	2014505	CATHE AUER	957 QUEEN ST.	Surat	Gujent		Statis
land a land	8	2018299		91	4,542,18	8 3,753.22	2018299	RIC POST	509 23 E AVE.	Vadodana	Gujarat		E Conte
and a second	7	2021392		37	4,730.57	7 4,630.57	2021302	RANDY SCOTT	533 DOMINIC ST.	Malad	Maharashtra		Total
ACCOL A	-8	2026599		99	20,114,5	7 20,534.85	2026500	CAROL INNELL	455 39 E AVE.	Mabalaumi	Maharashtra		E Criter
Be	2	20,26689		.99	752.75	5 802.75	2026689	ROGER SMN/THE	141 58 ST.	Andheri	Meharashtra		
	10	2035590		80	10.8	8 10.88	2035590	CHRIS DVSON	72-930 QUEEN ST.	Rajkot	Gujanat		Results
	11	2103497		75	1434.44	6 1,461.66	2103407	LINDA JONES	140 11 AVE.	Sunt	Gujarat		
	12	2114296		79	2,840.6	7 2,641.98	2114296	TOM OZENBERGER	19-956 MAIN ST.	Vadodana	Gujarat		I + Indicas
	13	2117497		70	1,797.13	3 1,787.13	2117407	KATHY MANCINELLI	635 59 AVE.	Maled	Maharashtra		
	14	2126092		71	180.1	9 180.19	2126092	MRS, TOM KAY	67.62.5 AVE.	Mahalaomi	Mehanshins		No No
-U Be	15	2132496	0	37	234.1	0 234.19	2132406	DAVID WITCHER	556 24 ST.	Andheri	Mahanashtra		0.000
- Br	16	2136091		82	0.00	00.0	2136001	LOUISE BAUMAN	731 83 AVE.	Rajkot	Gujarat		U + Comm
-100 Br	17	2207090	0	51	528.13	3 528.13	2207090	RAY CASEY	336 KING ST.	Sunt	Gujarat		
	18	2209195		69	299.4	E 299.48	2209195	MRS. RON HART	45-333 12 S AVE	Vadodana	Gujanat		H Add
1.2	19	2218295		96	4,607.86	6 4,607.96	2218295	HENRY HEITZMAN	332 RIVERSIDE DRIVE	Malod	Maharashtra		comr
100 C	20	2221299		51	26,082.18	6 20,080.14	2221200	JEFF HIRSH	26-216 80 AVE.	Mahalaomi	Maharashtra		
Custo	21	2223196		85	9,729.83	7, 4,725,45	2223196	JOHN LAGANA	400 BASELINE RD.	Andheri	Maharashtra		
Loan	22	2225597		32	-1,370.50	852.45	2225507	ARTHUR C. LINN	770 87 E AVE.	Rejkot	Gujanat		
- Misste	23	2226892		81	523.60	2 523.02	2226802	GLORIA GRANT	605 FOUNTAIN ST.	Suret	Gujanat		
Maste	24	2230191		37	10,122.21	1 9,760.56	2230191	LEWIST. BUTTS	57-S3 GARDEN RD.	Vadodene	Gujarat		
Paym v	25	2231397		37	47,055.00	8 45,448,42	2231307	GARY RISHER	6-745 MAIN ST.	Malad	Maharashtra		
3	26	2233799		37	2,330.5	0 2,377.90	2233790	ROCHELLE MONTGOMERY	13-786 RIVERSIDE DRIVE	Mahalaumi	Maharashtra		
	27	2234797		88	1,762.81	1,664.96	2234797	CHUCK BRODERICK	170 60 S AVE.	Andheri	Maharashtra		
SRunning Taulo lanaged Project: /	Ade. IT			1 6	wB				Not connected	to IDEA Server		umber of Records: 1,000   Disk	Specie 248.54 09.05

Fig: 4.12.4 – Joined database using join function

EXCERCISE: 2 – Join using visual connector option - Customer details.imd with account details. imd

Account details file has fields - ACCNO, STATUS, BR_CODE, BALTODAY, BALYESTDAY

Customer details file has fields - ACCNO, CUST_NAME, CUST_ADD, LOCALITY, STATE

Join both the file using common field ACCNO

Step 1: Open primary file account details and ensure it is active file in IDEA

Step 2: Go to analysis tab and click on visual connector button. Refer Fig 4.12.5

Step 3: From the list of databases on the left panel, select file customer details

INFORMATION TECHNOLOGY

303



Step 4: Create a link between both the files by dragging the mouse starting from ACCNO field in account details file and dropping at ACCNO filed at customer details file. This will establish a connection between two files. Refer Fig 4.12.6

				Visual Connec	tor : Create / Edit	Relation				- 0 ×
5.2										
Desktop Project		La contra								
Name	Re	Account Details	10							
<ul> <li>Anno Santa Santa</li></ul>		<ul> <li>≥ acoso</li> <li>⇒ anue.</li> <li>⇒ anue.</li></ul>								
							Design at 1	Hale	- CW	Court 1
							Provide A8	нер	UK.	Laicei
📑 📴 🧔		XI e	W3 🖬	3					- N 8 al	09:10

Fig: 4.12.5 – Visual Connector Options

		Visual Connector : Create / Edit Relation	- 0 ×
5.4			
sktop Project	Account Details	Customer Details	
Accord Fairly Catinet Fairly Lion nutre Material Data Figure Payment Program Sates Regist Sates		E Adovo (M) E outravane (R) E outravane (R) E outravane (R) E otrate (R)	
· .			

Fig: 4.12.6 – Establishing connection using key field

Step 5: Click on Ok and select appropriate option. In this case your objective is to join the file, select option "all records in primary file, all matches". Refer Fig 4.12.7

Step 6: Provide file name – combined details visual connector and click on OK. Refer Fig 4.12.7

Step 7: Observe the output in combined file created using visual connector option. Refer Fig 4.12.8



→ → ビ 計 計 目 : → → ビ 計 計 目 : → → → ビ 計 計 : → → → ビ 計 計 :	ga <del>z</del> ss Vev Marca Snartkreb	CaseWare IDEA - VisCon or	- 0
Re run Tools E-toolse	II fan Defection III belente fan -	SEfumerication III Apro Harsen et Una Consector III Albude Procedory Unit * Visual Connector : Create / Edit Relation - II K	
File Explan. *  Account 0  Accoun	Destituto Project Name Re Customer Datais Customer Datais Date master Minister Data-Sup Paymaret-Payma	Account Cetals II Contract Cetals II II Contract Cetals II	× X     Properti #     I = Databas     Vata     Field     Stricti     Centn
Partine         C 220002           →         Accore         B           →         B rc         B </td <td>Payell March 26     Sain Register     SalE505     User yweb Jog-Da     Vehical Paircy M</td> <td>Append djackee reverse to staped felds     OK     Gebend and     Gebend and     Gebend featile insulationmentar     Generic     Gebend dische insulationmentar     Generic system disblare</td> <td>Total ■ Criteria II ▼ Results II ▼ Indices No index II ▼ Comme</td>	Payell March 26     Sain Register     SalE505     User yweb Jog-Da     Vehical Paircy M	Append djackee reverse to staped felds     OK     Gebend and     Gebend and     Gebend featile insulationmentar     Generic     Gebend dische insulationmentar     Generic system disblare	Total ■ Criteria II ▼ Results II ▼ Indices No index II ▼ Comme
Br         HI         2556407	¢3	Percer-Al Heb OK Cancel	iii Add comm Datab han bees trunca The maxim runda of
C.F., C.L. 27 230620	70 52.0	51.01 WELLAND, IBMERN 596 HARANE. Sund Disjonet	recon
Managad Project Ade. III	× e -	Not connected to DEA Server Number of Recode 1	09-22 ↓ -41 4+ 09-22 -07-06-2017

Fig. 4.12.7 – visual connection option

P. Indees	e 用	Top Records Key Value	E Cap C	Intection 200 refe Law 2005to European	plcate Key + stelica +	E Summerication	E Aging Trivot Table	Berlan M Berlappend D	Company	副 Attribut 副 Attribut 圖 Other •	e 🔤 Honetary	(Unit * *		
ala al	1	Account 1	letaits X	Combined	details visual i	man_ X							* X	Descent
pages 1	<u> </u>	ACONO I	PLITATP	-88.COOF	BALTODAY	BU VERTOWN	CI	ST NUME	CURT	400	LOCALITY	STATE		riupen
ia	1	1 30139406		65	473.53	473.93	DANEL LEE	ANDRACK	308 25 F AVE	- mere	Raibre	Ginarat		11 * Data
	2	53207396		43	11.814.52	12,236,48	H. ALAN HE	AD	194 HENT CU	UB RD	Reiket	Guiarat		10
7 II	1	43908500	N	87	596.01	412.04	PAUL WOO	DS	71-426 DOMP	UCST.	Malari	Maharashtra		1.1
8	4	72907590		45	43.35	11.35	LOS BRUNK	2	62-763 KING S	T.	Rekt	Guerat		1 2 2
_	5	2719391		79	2.169.67	1,916.50	PAUL PIETS		728 BASELINE	RD.	Andheri	Maharashtra		
Pance	8	82130608	0	81	1320.48	1.856.45	TONY UKIG	Al	909 87 S AVE.		Sanat	Guarat		1 4 2
	7	92120095	1	40	53.23	51,23	MICHAEL PORTELLA		401 37 AVE.		Malad Maharashtra		1	
Accos A	8	82429892		81	2,725.57	1,784.35	ADBERT B.	RHODES	535 LELICO A	NE.	Mahalarra	Maharashtra		= 0
3 Br	2	82628295	1	77	6.351.98	6,351.98	CAROLYNY	WHY RANK	549-42 ST.		Reknt	Guieret		
0.0x	10	42917193		95	1,750.21	4,550,21	JAMES A. VI	ERMEYDEN	49-853 QUEEN	IST.	Maled	Maharashtra		T Res
3 Be	11	32622090		82	1,630.38	3,247.99	PAT THOM	AS	767 KENSIGNT	TON AVE.	Andhen	Maharashtra		
Be	12	73707296		73	40.12	40.12	TOM LEDES	MA	457 KING ST.		Andheri	Maharasittra		il + ind
8 R F	13	62213692		36	2,543,64	1,053.32	BROCK HIN	ES	40-141 MAIN:	ST.	Rejket	Gujerat		
8 Be	14	2513596		37	10,989,46	10,712.89	WILMA YOU	INGSMA	40-229 89 W A	WE.	Vedodara	Gujarat		1 + 00
6e	15	72232296		41	615.77	505.68	PATRICIAA	NN MARKO	16-875 31 AVE		Sunat	Gujarat		
Ð Ør	16	82729499	0	88	36,50	68.56	AL RECKEL		842-90 ST.		Vadodara	Gigarat		H A
( <b>1</b> 6)	17	93205093		51	390.84	361,34	KEWIN RITCI	HIE	852 7 N AVE.		Sanat	Gujeret		4
9 B K	18	32304897	1.1	32	775.60	1,041.65	BOB DUNLE	AVY	593 1 S AVE.		Sanat	Gujarat		
2 PO	19	2430395		51	22.65	34.90	FRANK ZUR	BORG	612 RIVERSIDE	DRIVE	Mahalaorra	Moharasittea		
<u> </u>	20	32006898		36	109.14	48.62	SARA WINT	ERSTEIN	44-475-91 AVE		Vadodara	Gigarat		
2.53	21	92702694	N	67	166,52	160.02	MR. W. W. B	LADEL .	892 KENT 5T.		Andheri	Maharashtra		
	22	92120095		37	416.72	1,333.24	MICHAELP	ORTELLA	401.37.AVE		Melad	Maharashtra		
.0510	23	2603199		91	290.94	290.94	DAVID MC I	DANIEL.	439 BANK ST.		Mahalaorre	Moharasittra		
Dan .	24	22099998	1	87	1,000.00	1,517,121.70	EDWARD CI	DOLIDGE	277 74 E AVE.		Andheri	Moharashtra		
WIDH V	25	92013894		97	74,326,10	93,176.12	<b>JIM STEWA</b>	RT	44-505 LILLICO	O AVE.	Malad	Maharashtra		
-	26	13205994	N.	32	343.75	335.87	MANNY CA	RMICHAEL	19-960 53 W A	WE.	Mahalasmi	Moharashtra		
	27	23206299		70	53,91	53,91	WILLIAM L.	<b>JENSEN</b>	596 14 N AVE.		Sunit	Gujarat		

Fig. 4.12.8 – Joined database using visual connection option

# 4.13 COMPARE DATABASES

IDEA can compare the financials from one year to the next, and it's called Compare Databases. This task requires two databases to be defined, a field in each database to be designated the "total" field, and a field in each database designated as key fields (match fields). The function generates the number of records in the primary database (P_NRECS), the number of records in the secondary database (S_NRECS), the sum of the total field defined in the primary database (P_TOTAL), the sum of the total field defined in the secondary database (S_TOTAL), and a DIFFERENCE column which subtracts the TOTAL fields from each other.

## EXERCISE: Compare two purchase databases

- Open the Purchases-L4-2010 database.
- You need to add a Virtual Numeric field with two decimal places called TOTAL_GROSS. The field must use a criterion that calculates the sum of all the gross monthly purchases for the year (JAN_GROSS + FEB_GROSS...etc.)

INFORMATION TECHNOLOGY

305



- Open the Purchases-L4-2011 database.
- You need to add a Virtual Numeric field with two decimal places called TOTAL_GROSS. The field must use a criterion that calculates the sum of all the gross monthly purchases for the year (JAN_GROSS + FEB_GROSS...etc.)
- On the Analysis tab, in the Relate group, click Compare.
- The Compare Databases dialog box appears.
- Click Select and select the Purchases-L4-2010 database as the secondary database.
- Click Match to select key fields that are common between both databases.
- Select SUPPNO for both and click OK.
- From the Total field drop-down list for both databases, select TOTAL_GROSS.
- In the File name field, enter Purchases 2011 vs 2010 for L4.
- Click OK.

Primary database:	ED-Purchases-L4-2011		ОК
Total field:	TOTAL_GROSS		Cancel Help
Secondary database:	ED-Purchases-L4-2010	Select	
Total field:	TOTAL_GROSS		
File ermer	Purchases 2011 vs 2010 for L4		

### Fig. 4.13.1 - Compare database options