Data analysis for Internal Audit

Compiled by CA Ashwin Dedhia

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

What is Data Analytics

- “Analytics is the process of obtaining an optimal or realistic decision based on existing data.”
  - Wikipedia
- “Data analytics is the science of examining raw data with the purpose of drawing conclusion about that information.”
  - Whatis.com
- “Analytics leverage data in a particular functional process (or application)to enable context-specific insight that is actionable.”
  - Gartner

What is Data Analytics?

- Common View

  Data analytics is defined as the process of

  inspecting,

  cleaning,

  transforming, and

  modeling data

  with the goal of highlighting

  useful information,

  suggesting conclusions, and

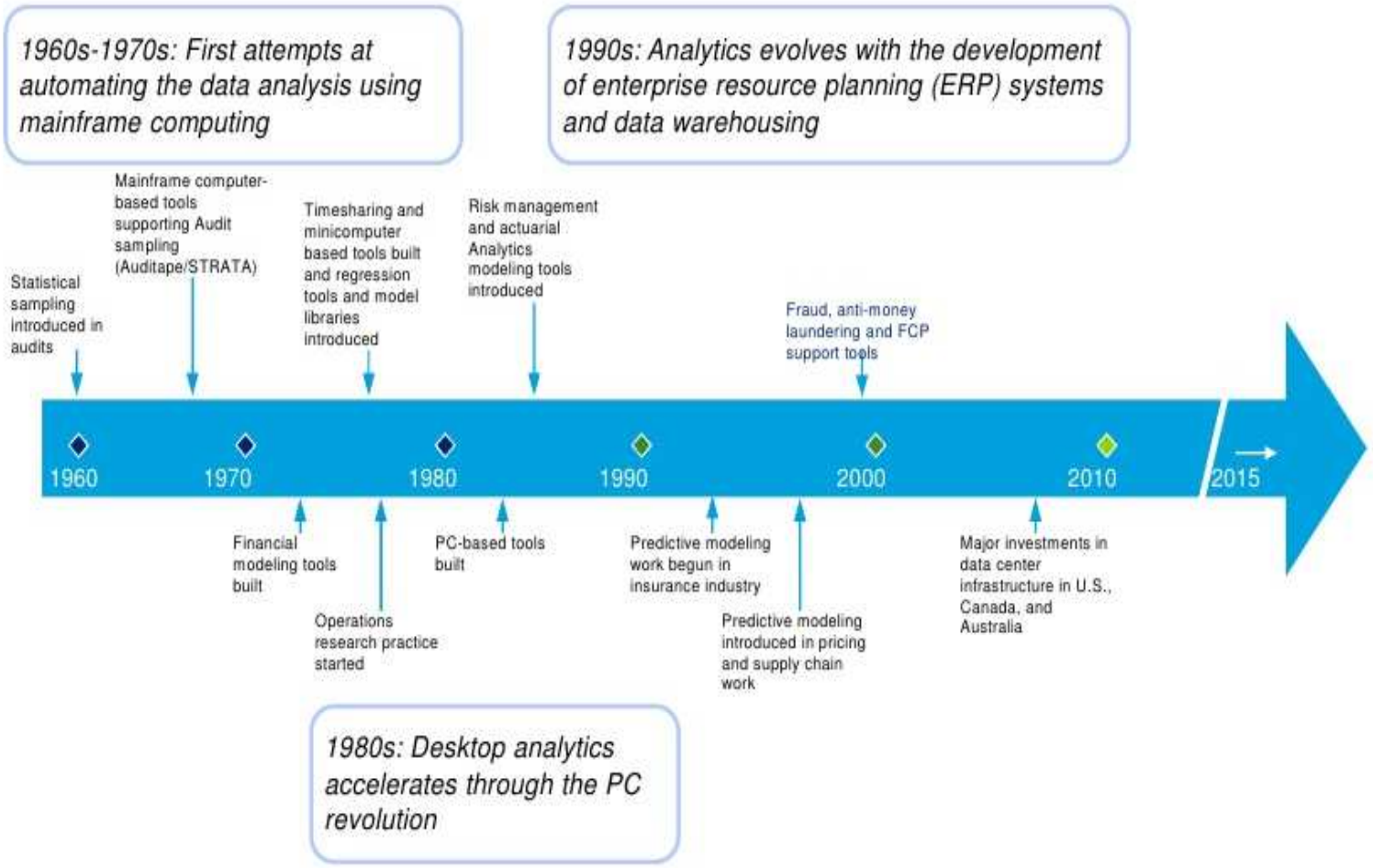
  supporting decision making.

What is Data Analysis

- Data analytics is an analytical process by which insights are extracted from operational, financial, and other forms of electronic data internal or external to the organization.
- These insights can be **historical, real-time or predictive** and can also be
- **risk-focused** (e.g. control effectiveness, fraud, waste, abuse, policy/regulatory non-compliance) or
- **performance focused** (e.g. increased sales, decreased costs, improved profitability)
- and frequently provide the “HOW?” and “WHY?” answers to the initial “WHAT?” questions frequently found in the information initially extracted from the data.

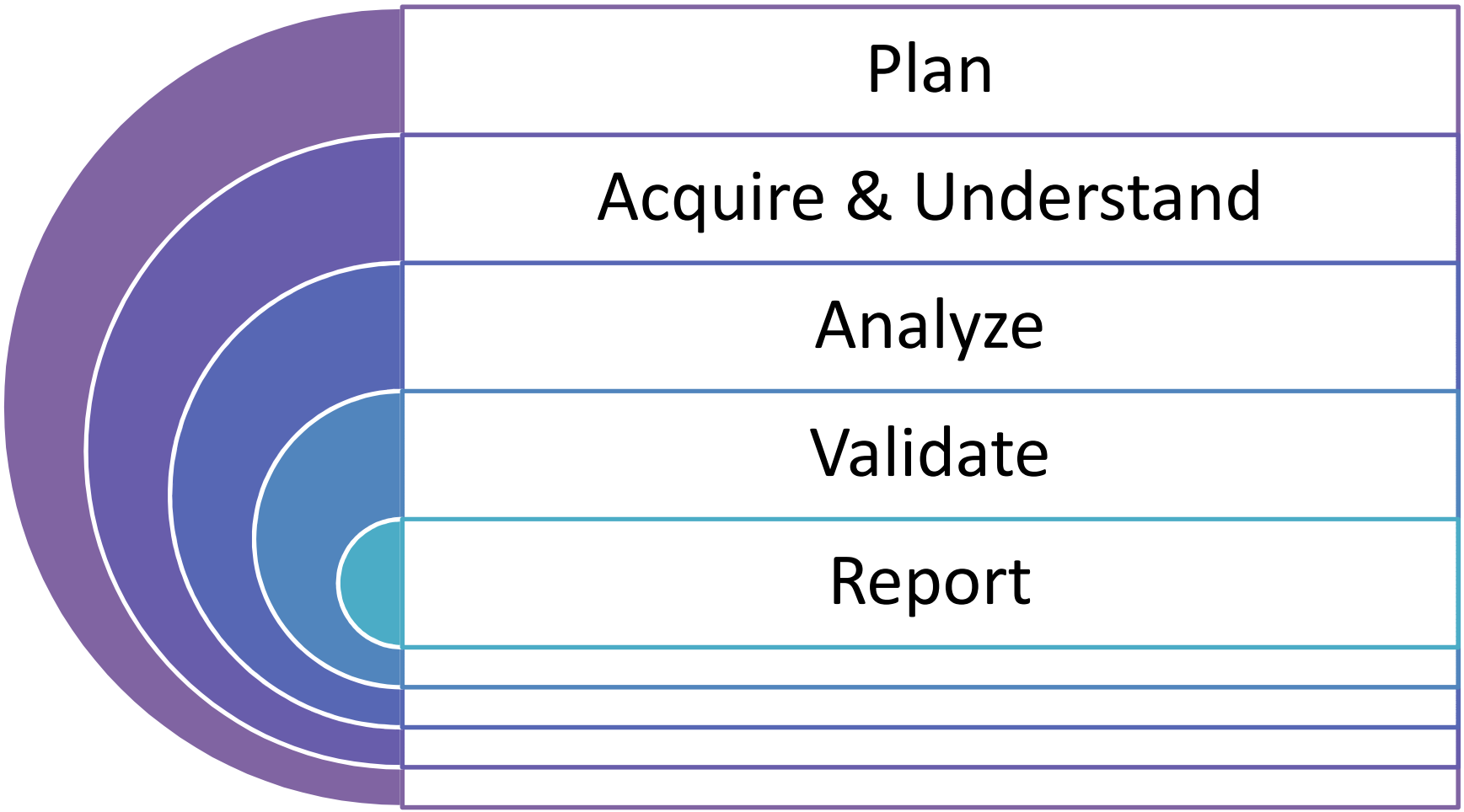
Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

It’s not something New



Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Methodology



Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

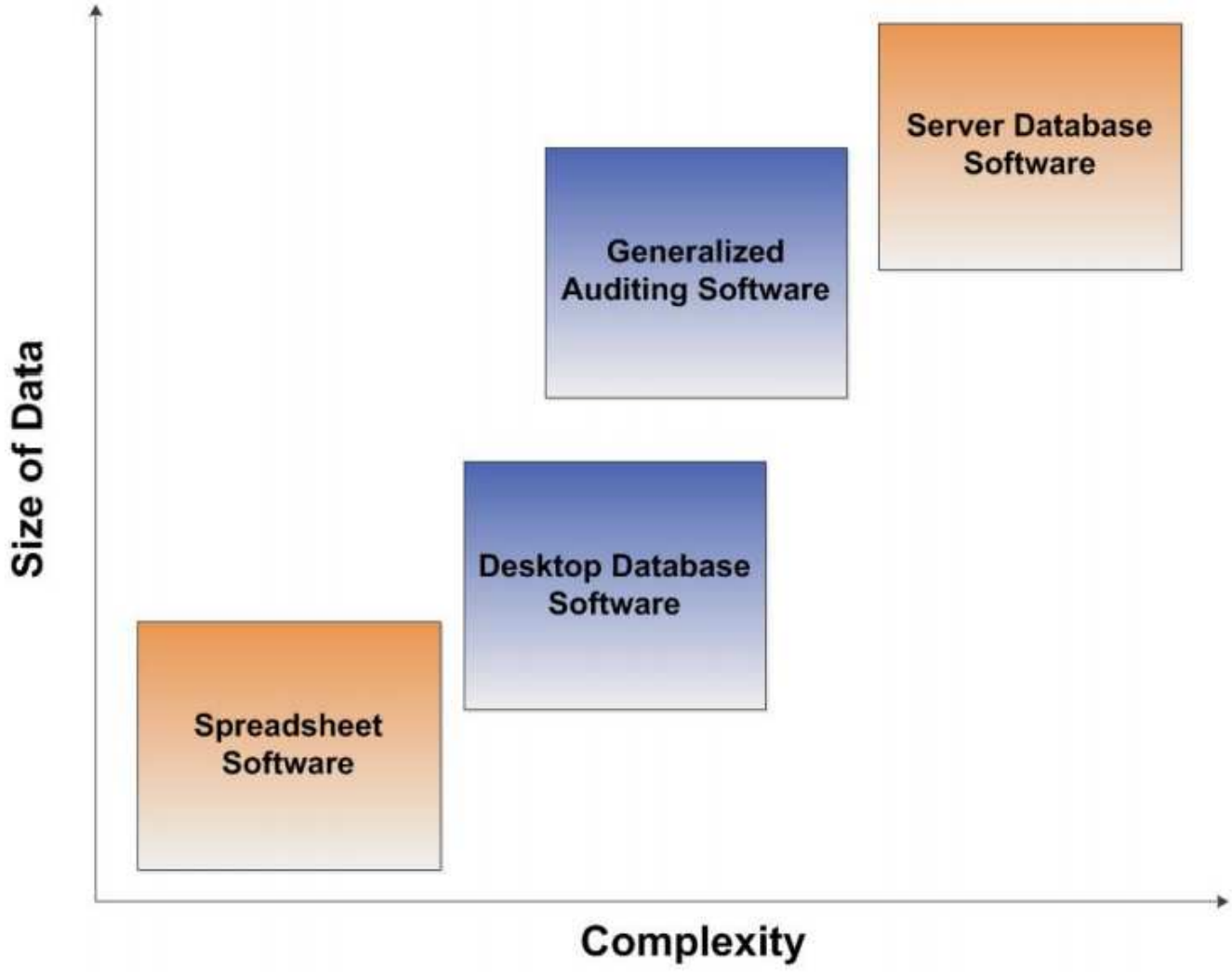
Desired Features of Analysis Tool

- Able to handle large data sets efficiently
- Wide array of analytical and statistical functions and procedures
- Programmability
- Logging of procedures performed on data
- Ability to easily re-run analysis with minor changes

Types of Tools

- Spreadsheet software (EXCEL)
- Databases
  - Desktop software (Microsoft Access)
  - Server-based (SQL/Oracle)
- Generalized auditing software
  - ACL
  - IDEA
- Other Tools
  - Tally
  - SAS
  - SPSS
  - Apache Hadoop

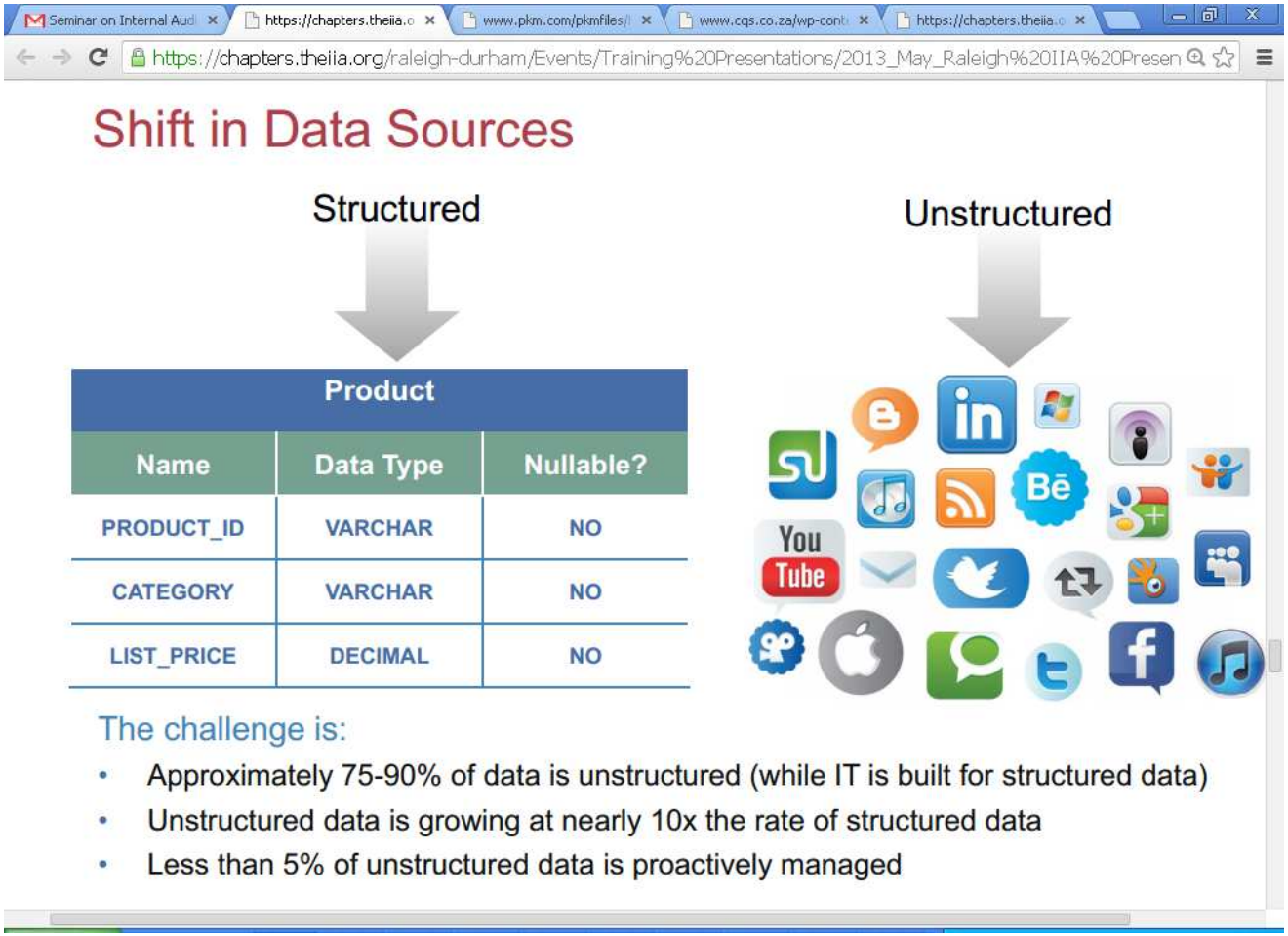
Trade-off


Server Database Software
Generalized Auditing Software
Desktop Database Software
Spreadsheet Software
Size of Data
Complexity

Benefits and Costs/Risk

- Benefits
  - Automation = saving time
  - Trending of transactions
  - Red flags
  - Master data issues
  - Control culture (you are being watched)
  - SOX, FCPA and other regulatory requirements
- Costs / Risks
  - Investment
  - Time for review, follow-up and communication
  - Insufficient understanding of source data
  - Lack of buy-in by management

Shift in Data Sources

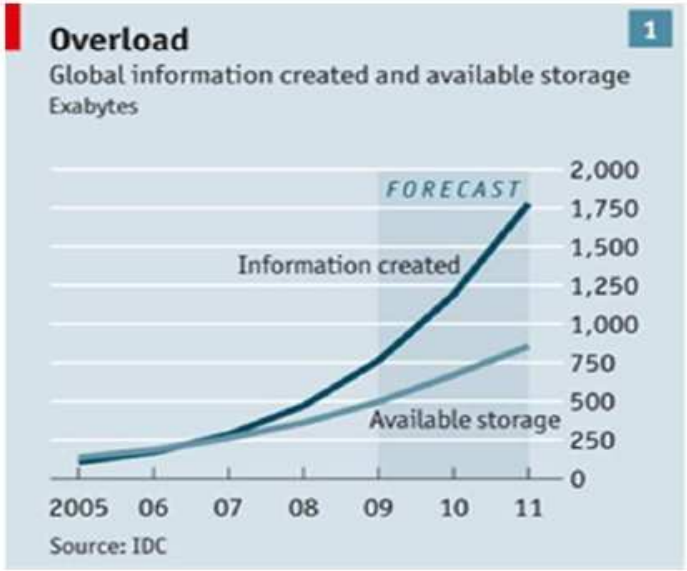


Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Data is growing Exponentially and there is pressure for companies to make faster and better decisions

Data Trends

A recent report by the Economist highlights that data-assets continue to grow exponentially



The Economist. Data, data everywhere. Feb 25th 2010

External and Internal Drivers

A recent Kennedy Report indicates that a variety of internal and external industry drivers are pushing our clients to embrace analytics

Top internal and external industry factors contributing to adoption:

External
1. External competitive pressure
2. Increased regulatory pressure
3. Technology advancement

Internal
1. Data proliferation and growth
2. Increasing sophistication of users
3. Maturation of ERP systems

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

What types of questions can analytics answer

Historical Perspective	• What happened • How Many • How often • Where
Current Perspective	• Where is the Problem • Why is this happening • What actions are needed
Future Perspective	• What if these trend Continue • What will happen next • What's the best that can be happen

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Can data tell me all

- Listening to the data is important… but so is experience and intuition. After all, what is intuition at its best but large amounts of data of all kinds filtered through a human brain rather than a math model ?
- The answer can come from
  - **Meaningful Data**
  - **Relevant Questions**
  - Beyond other aspects… technical, domain, etc.

Analytical Tests

- *Evaluations of financial information made by a study of plausible relationships among both financial and non-financial data to assess whether data appear reasonable*
  - Horizontal Analysis
  - Vertical Analysis
  - Ratios
  - Trend Analysis
  - Performance Measures
  - Statistics
  - Stratifications
  - Aging
  - Benford's Law, Regression, Monte Carlo Simulation

Horizontal Analysis

- Analyzes the increases and decreases in a given balance, normally financial statement items, over two or more periods or segments.
- Examples
  - Balance sheet
  - Income statement
  - Budget to actual
  - Multi Business
  - Multi Companies

	A	B	C	D	**Formula**
1	Asset Description	This Year	Last Year	Difference	
2	Cash	$1,000	900	(100)	**=B2-C2**
3	Accounts Receivable	1,000	900	(100)	**=B3-C3**
4	Fixed Assets	8,000	7,200	(800)	**=B4-C4**
5	Total Assets	**$10,000**	**9,000**	(1,000)	**=B5-C5**

Vertical Analysis

- Examines the elements of a financial statement for a single period whereby each balance sheet item is shown as a percentage of the total assets and every income statement item is shown as a percentage of the net sales. This can be completed for the following balance sheet information using the formula in the most right-hand column (calculating column C):

	A	B	C	**Formula**
1	Asset Description	Asset Balance	% of Balance	
2	Cash	$1,000	10%	**=B2/B$5**
3	Accounts Receivable	1,000	10%	**=B3/B$5**
4	Fixed Assets	8,000	80%	**=B4/B$5**
5	Total Assets	**$10,000**	**100%**	**=B5/B$5**

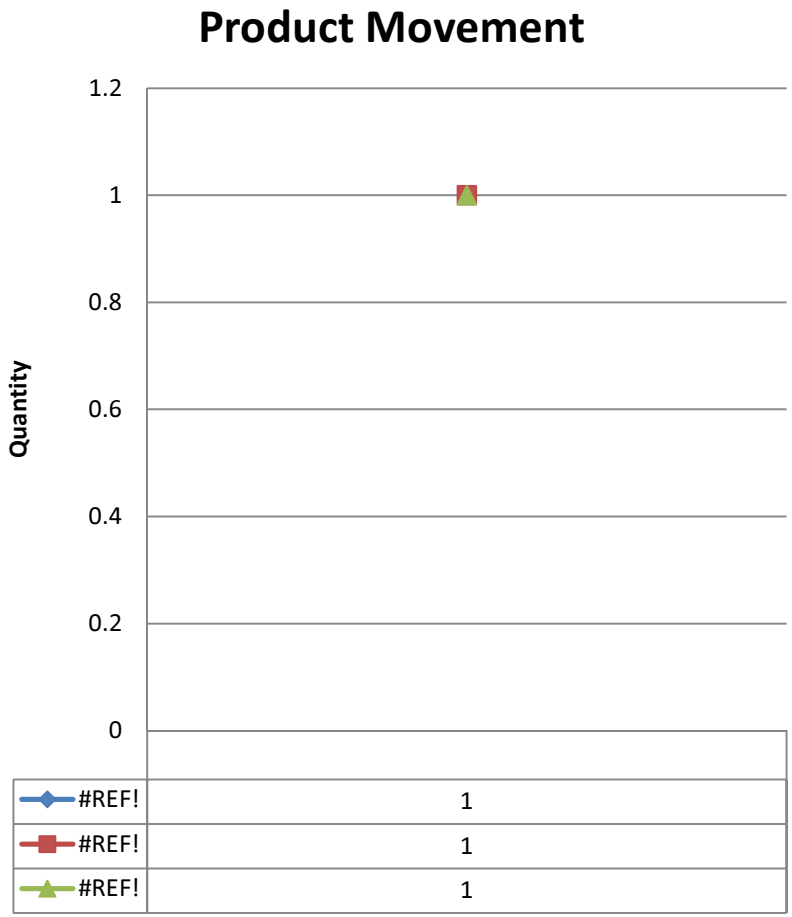
Ratios

- One or more balances is compared with one or more other balances such as the relation of total assets to the net sales of an organization. Ratios can be organized into broad categories of “Liquidity/Debt” and “Profitability”.

Name	Description
Working Capital	(Current Assets - Current Liabilities)
Working capital index	Current year WC - Prior year WC
Current Ratio	(Current Assets / Current Liabilities)
Days Payable Outstanding	(365 / (Sales / ((Beginning Accounts Payable + EndingAccounts Payable)/2)))
Days Sales Outstanding	(365 / (Sales / ((Beginning Accounts Receivable + Ending Accounts Receivable)/2)))
Inventory turnover	(Cost of Goods Sold / ((Beginning Inventory + Ending Inventory)/2))
Debt to Equity	Total Debt / Total Stockholders Equity
Operating Cash Flow	(Cash Flow From Operations / Current Liabilities)
Cash Flow Interest Coverage	((Cash Flow From Operations + Interest Paid + Taxes Paid) / Interest Paid)
Cash Flow to Capital Expenses	Cash Flow From Operations / Capital Expenses

Trend Analysis

- Comparing any of the analytical tests (horizontal, vertical, ratio, etc.) described above over two or more periods. Please note that the use of trend analysis is practically a given in doing any audit work as fraud and errors tend to create variances over time which would go undetected if only the single year was being analyzed.



Performance Measures

- The identification of critical success factors that can be tracked over time to assess progress made in achieving specific targets linked to an entity's vision. For example, the below represent a sampling of performance measures that could be used for accounts payable processing:
- The Excel calculations for the performance measures would be completed using formulas in Excel
  - Sum(), Count(), counta(), Average().

- Number of invoices processed
- Number of open invoices at period end
- Top 100 vendors purchases
- Average of top 25 max to min payments by vendor ratio
- % of adjustments to invoices processed
- Number of hours overtime worked by staff

Stratifications

- Counts the number and value of records of a population falling within specified intervals.
- Stratifications also provide a useful view into the largest, smallest, and average transactions.
- To create the above Stratification Report, follow the two-step process outlined here
- Aging Report can also be derived using Stratifications

- Step #1 – Complete a calculation of the strata using a multiple IF function in the right-hand column:

  =IF(A2>1000,"3",IF(A2>100,"2",I
  F(A2>0,"1","0")))

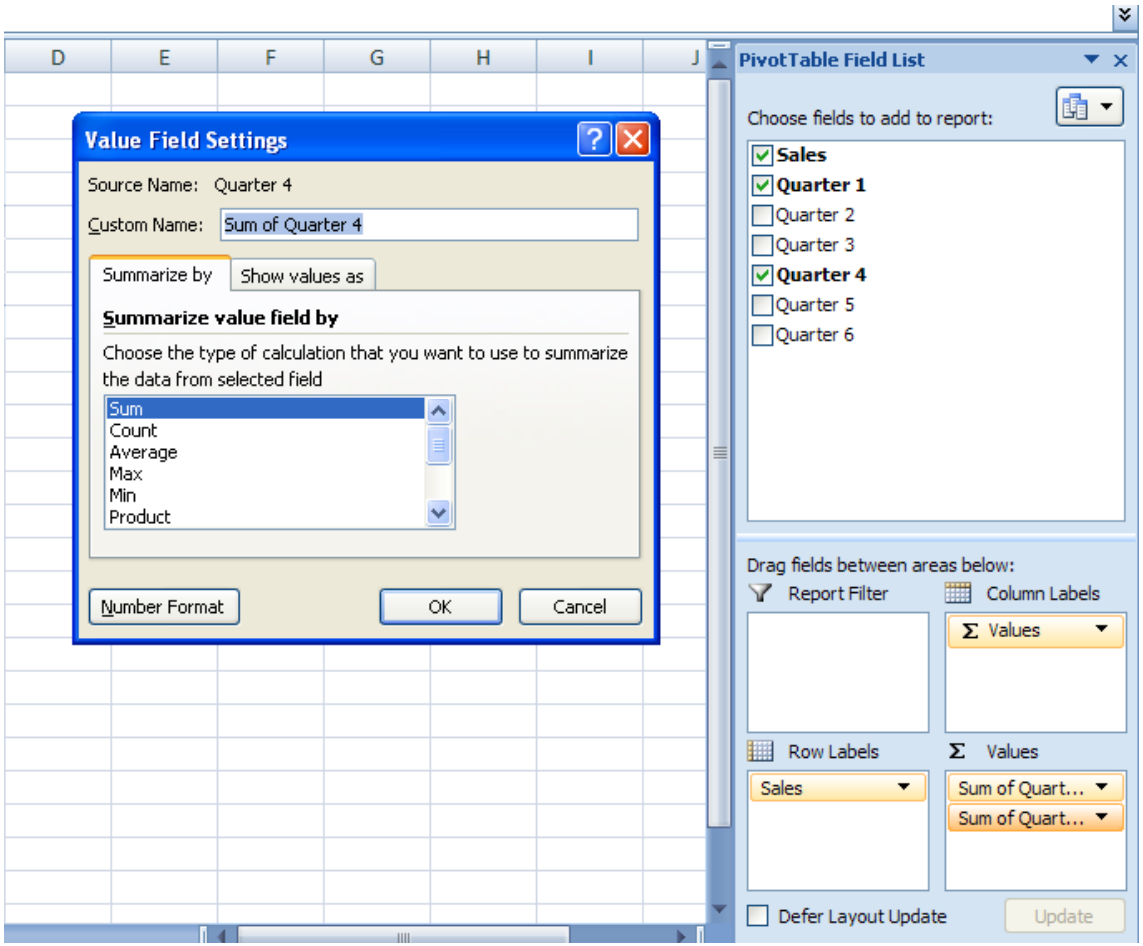
- Step #2 – Use the strata calculated above to populate the stratification table using the below COUNTIF () and SUMIF () formulas

  =COUNTIF(B2:B6,1)

  =SUMIF(B2:B6,1)

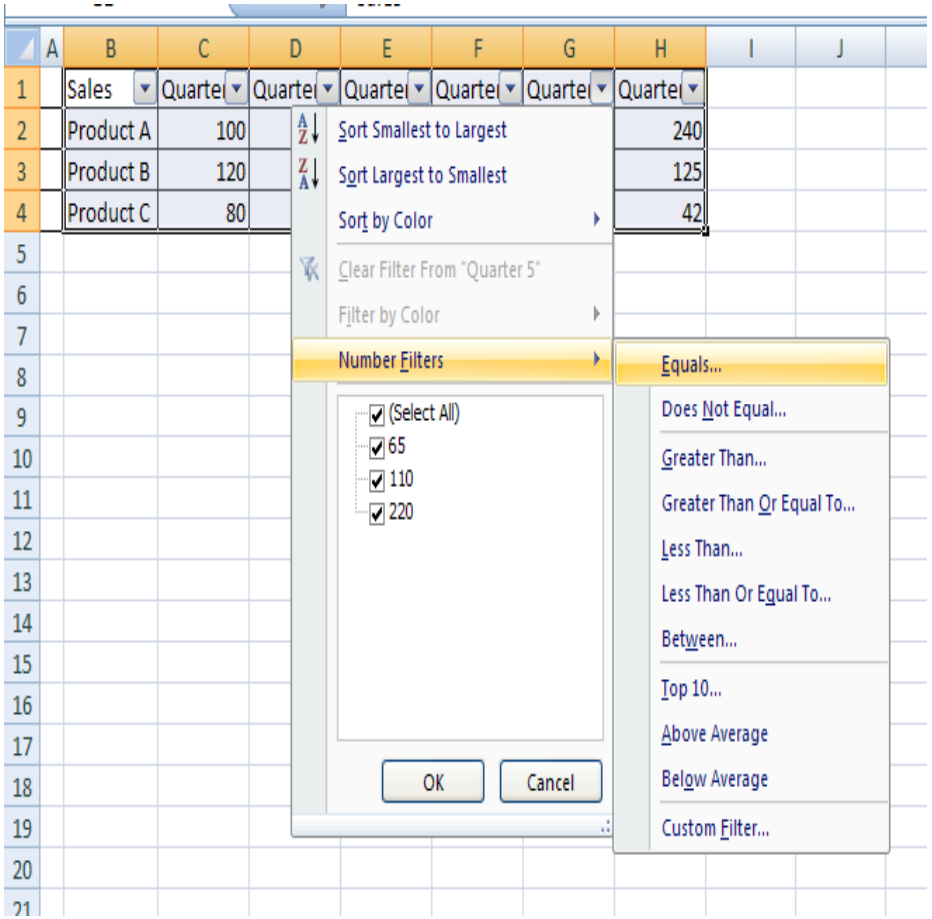
Cross Tabulate / Pivot

- Cross Tabulate lets you analyze character fields by setting them in rows and columns. By cross tabulating character fields, you can produce various summaries, explore areas of interest, and accumulate numeric fields.
- In Excel we can use
  - Pivot Tables
  - Pivot Charts



Extract & Filters

- Extracts specified items from one file and copies them to another file, normally using an “if” or “where” statement. Examples include extracting all balances over a predefined limit. This is most easily accomplished using the *Auto Filter command under the Filter option of the Data menu.*



GAPS

- Identifies gaps within a specified field in a file. For example, to identify any gaps in check sequence in the below table, the following formula (see cell C3) would be entered and copied down for the entire length of data in the worksheet (in this spreadsheet this would start at B3 and copied down to B6):
- Add Conditional Formatting to get highlights faster

Number	Skip Count	Formula (Text)
12023	#VALUE!	=IF(A3-A2<>1,(A3-A2)-1,0)
12024	0	=IF(A4-A3<>1,(A4-A3)-1,0)
12025	0	=IF(A5-A4<>1,(A5-A4)-1,0)
12026	0	
12027	0	
12028	0	
12029	0	
12030	0	
12031	0	
12032	0	
12033	0	
12035	1	
12037	1	
12038	0	
12039	0	
12040	0	
12041	0	
12042	0	
12043	0	
12044	0	
12046	1	
12047	0	
12048	0	
12051	2	
12052	0	
12053	0	

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

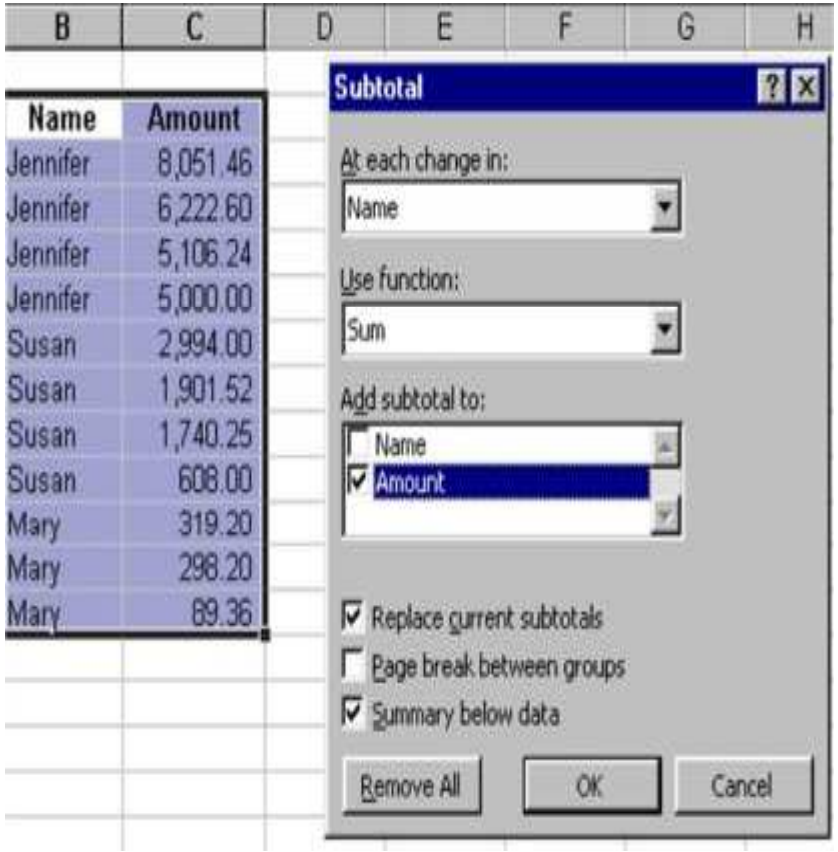
Join / Relate

- Combines specified fields from two different files into a single file using key fields. This function is used to create relational databases on key fields and mainly uses the *VLOOKUP() function in Excel.*
- *Please note that of all of the Excel uses as an audit tool, the joining of files is most difficult and better completed by an audit or database software. However, simple joining of files is possible and explained further below.*

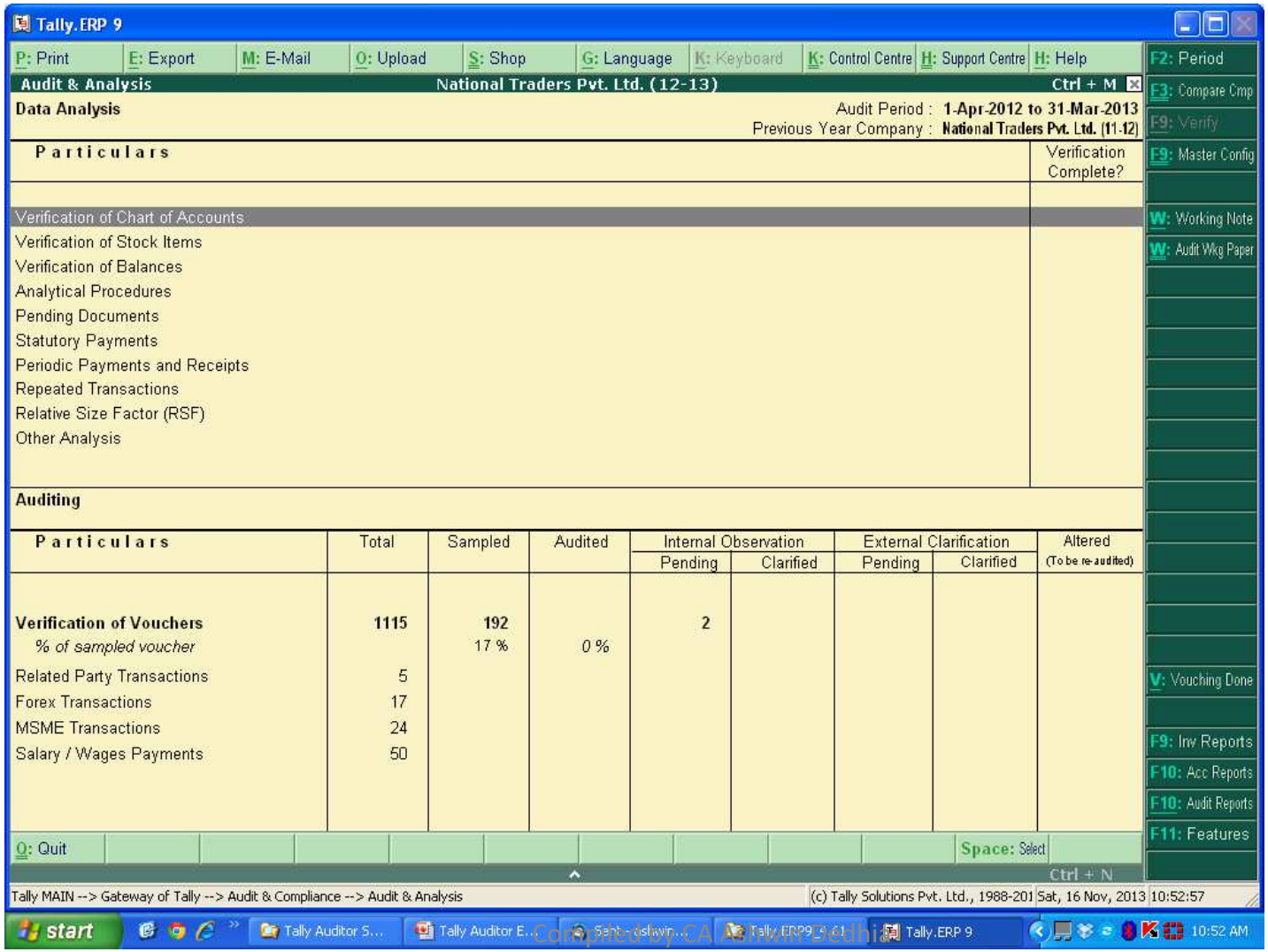
	A	B	C	D	E	F	G	H
1								
2				Most Likely				
3	**Vendor Address**		**Employee Address**	**Vendor Address**	**Forumla**	**Match?**	**Forumla**	
4	120 Berger Place		20816 HardwoodCourt	145 Ridge Road	=VLOOKUP(C4,A$3:A$8,1)	N	=IF(D4=C4,"Y","N")	
5	145 Ridge Road		22 Waverly Place	22 Waverly Place	=VLOOKUP(C6,A$3:A$8,1)	Y	=IF(D5=C5,"Y","N")	
6	22 Waverly Place		35 Waldorf Avenue	2412 Maple Ave	=VLOOKUP(C6,A$3:A$8,1)	N	=IF(D6=D6,"Y","N")	
7	2300 Pilgrim Lane		4 Meadowland Court	2412 Maple Ave	=VLOOKUP(C7,A$3:A$8,1)	N	=IF(D7=D7,"Y","N")	
8	2412 Maple Ave		42 Park Avenue	2412 Maple Ave	=VLOOKUP(C8,A$3:A$8,1)	N	=IF(D8=D8,"Y","N")	
9								
10								
11								

Summarize

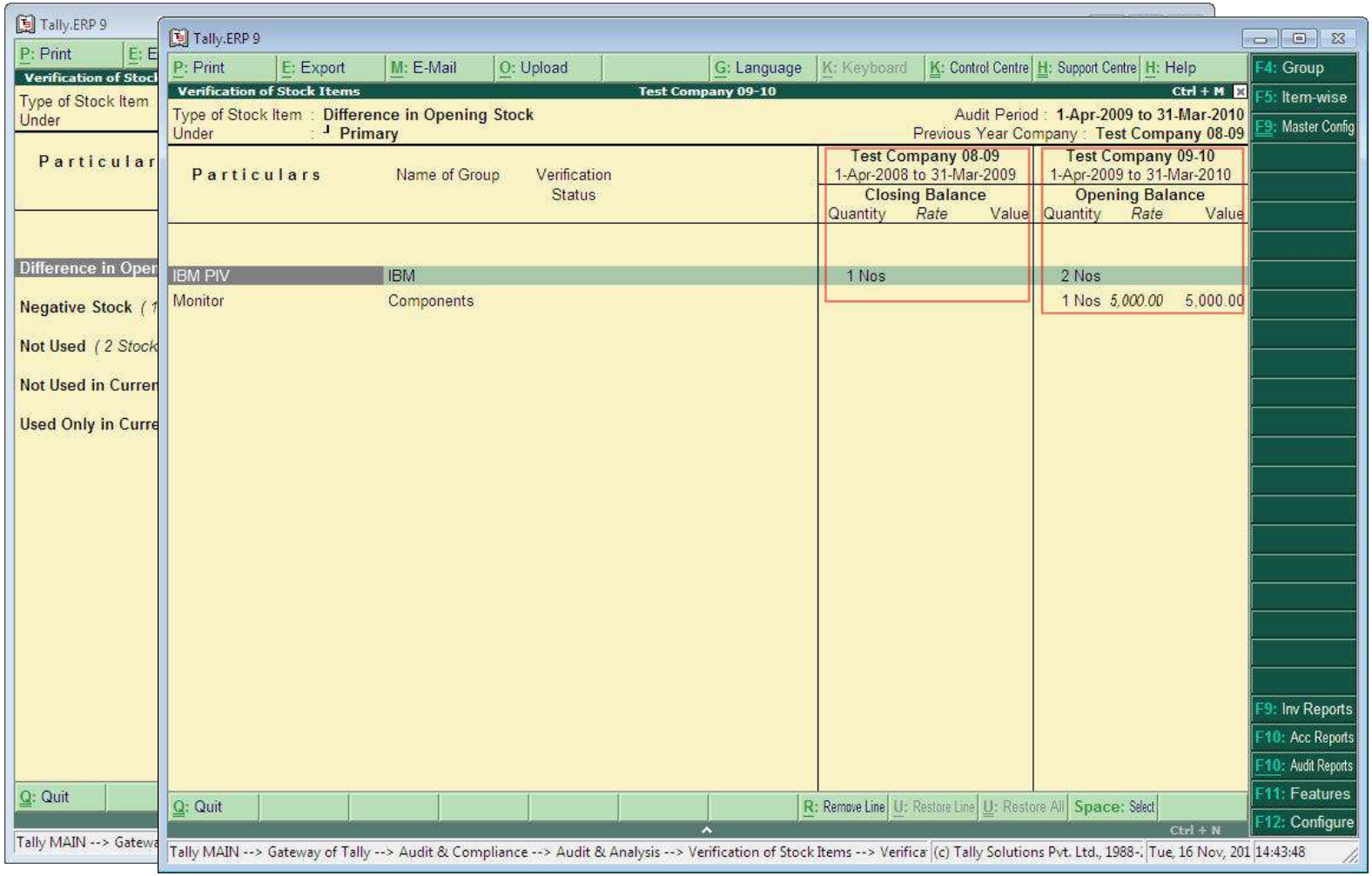
- Accumulates numerical values based on a specified key field.
- For example one potential test could be summarizing travel and entertainment expense amounts by employee to identify unusually high payment amounts. To complete this test, first sort the data using the Sort command explained above. Then, highlight the data to subtotal and then select *Subtotals from the Data menu item to be presented*



Data Analysis in Tally

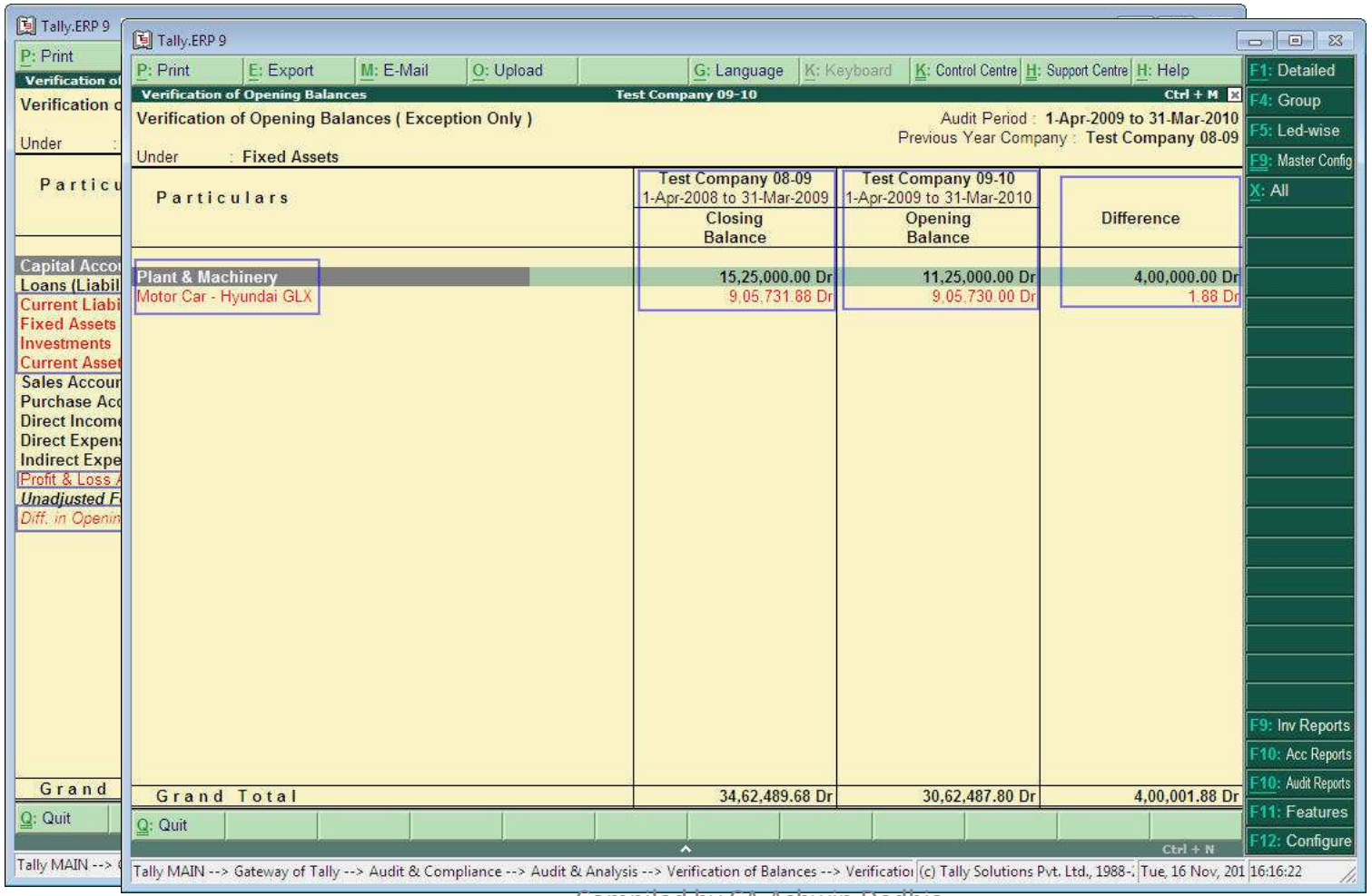


Verification of Stock Items

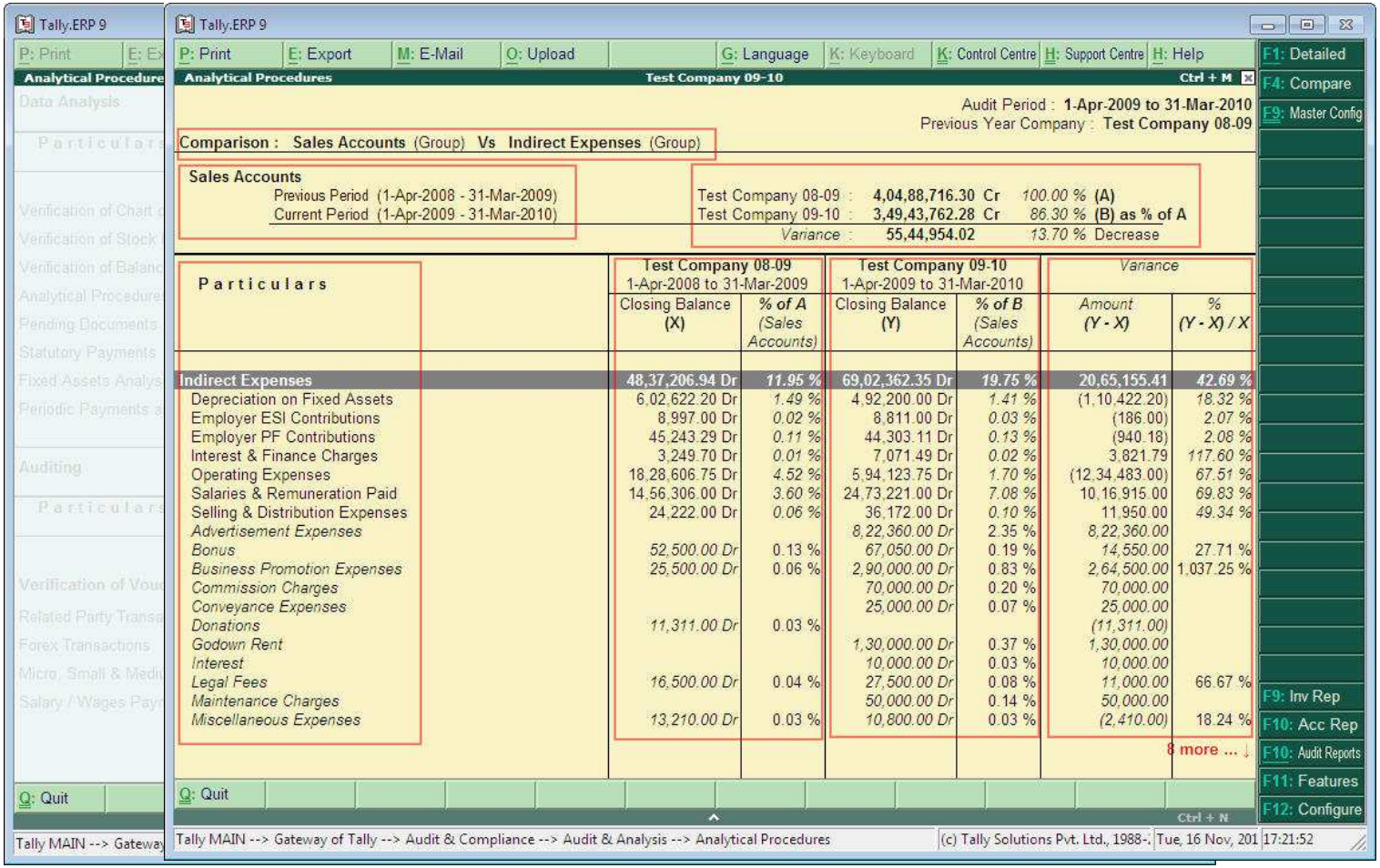


Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Verification of Balances

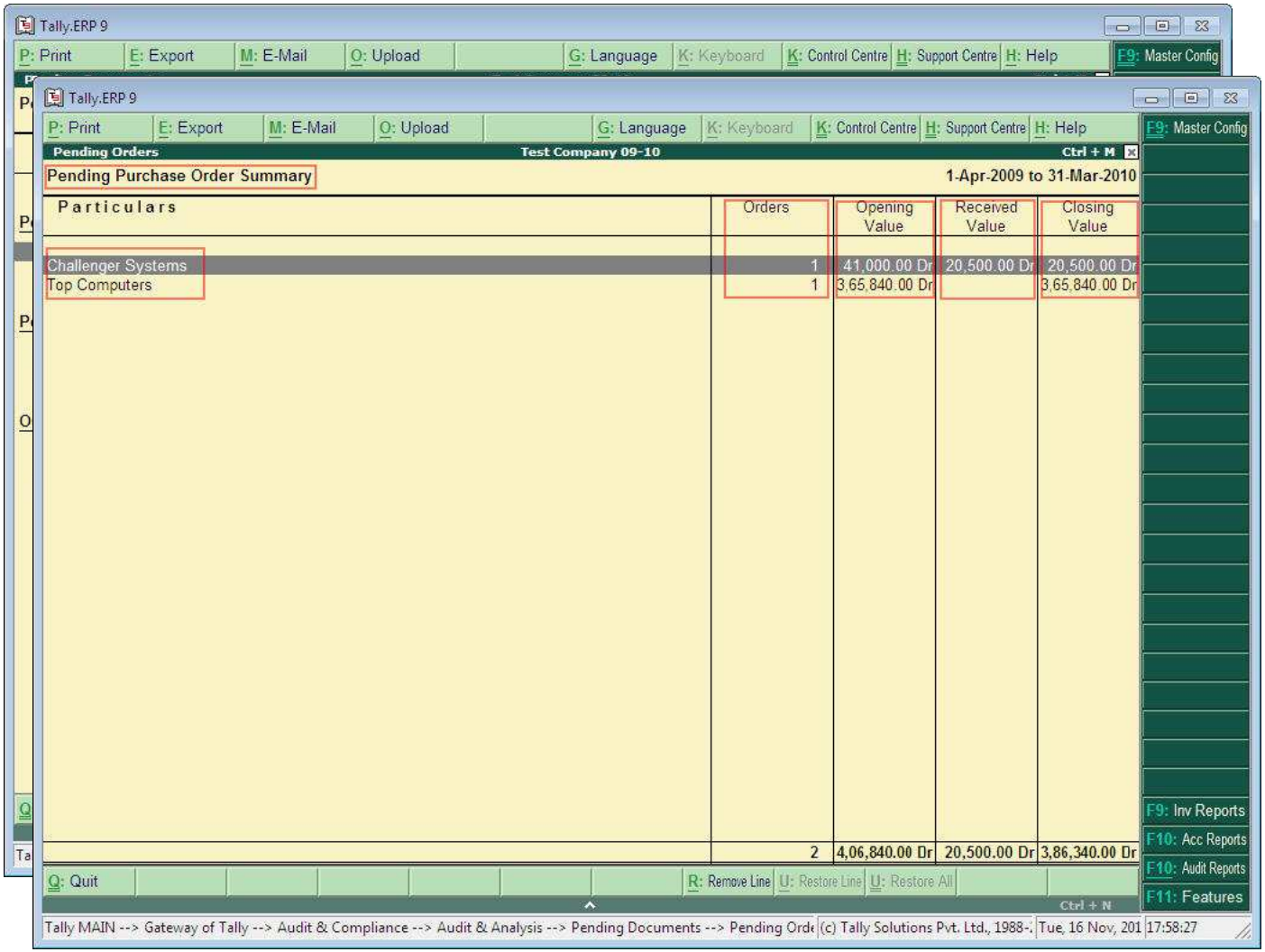


Analytical Procedures



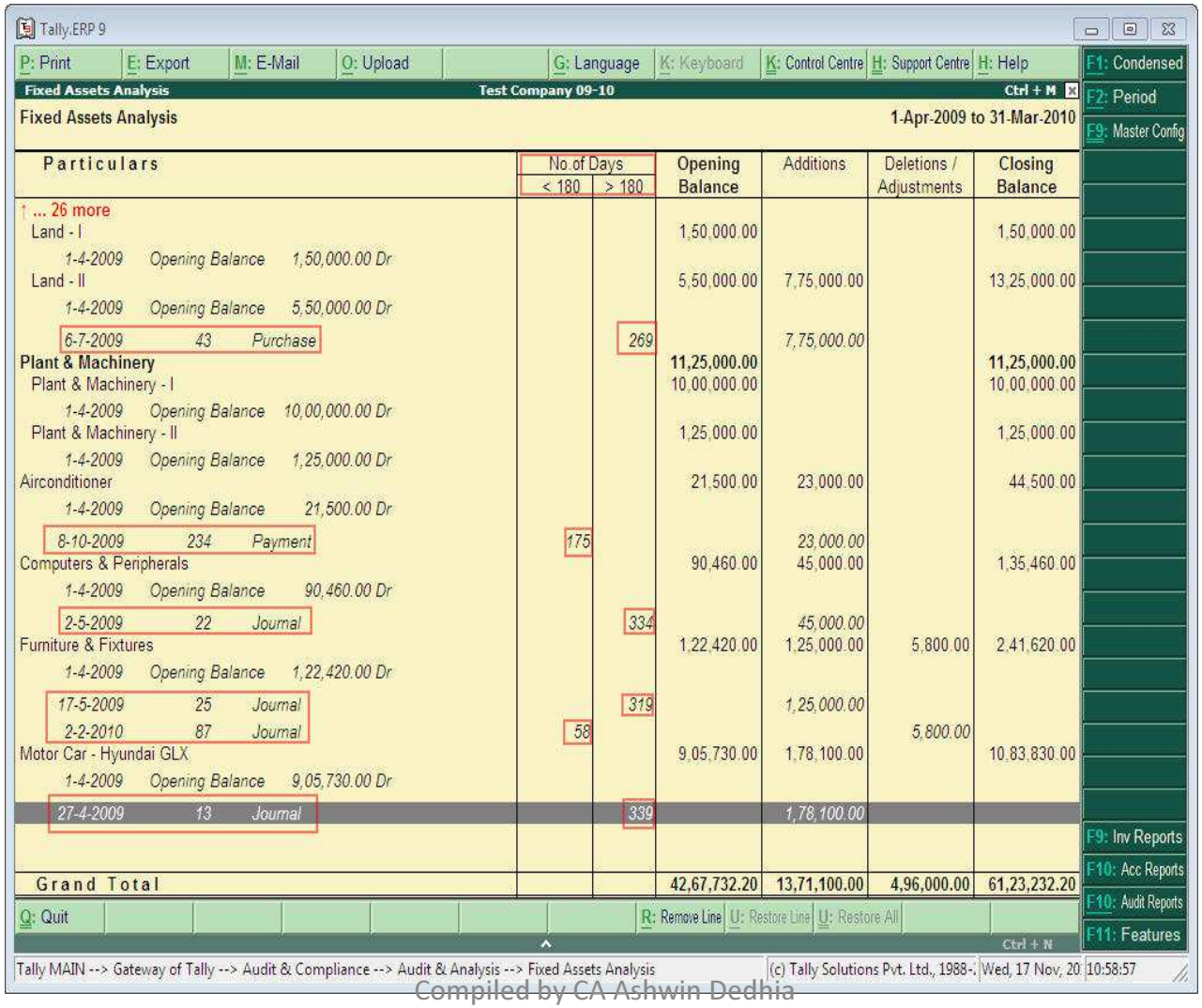
Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Pending Documents



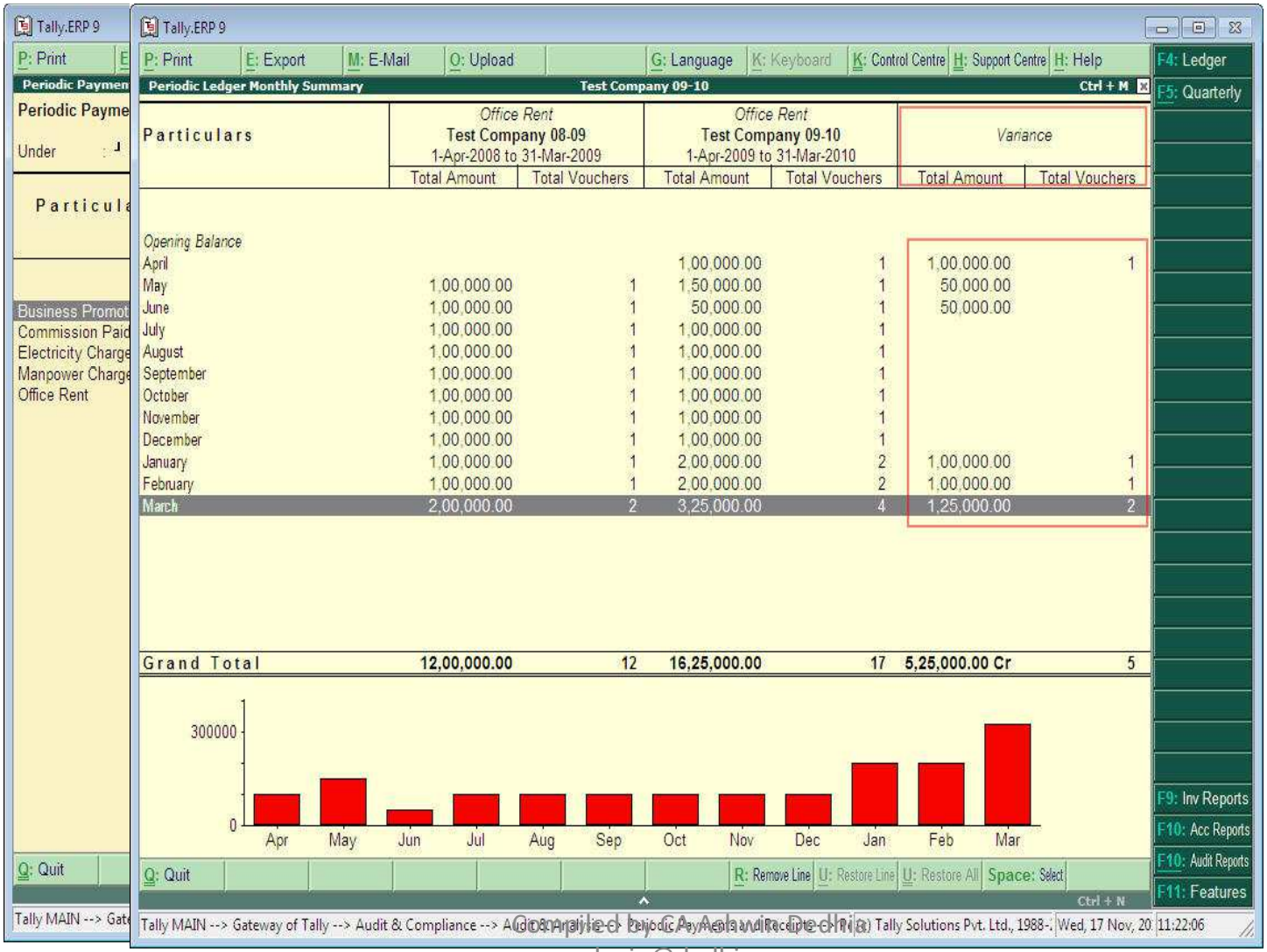
Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Fixed Assets

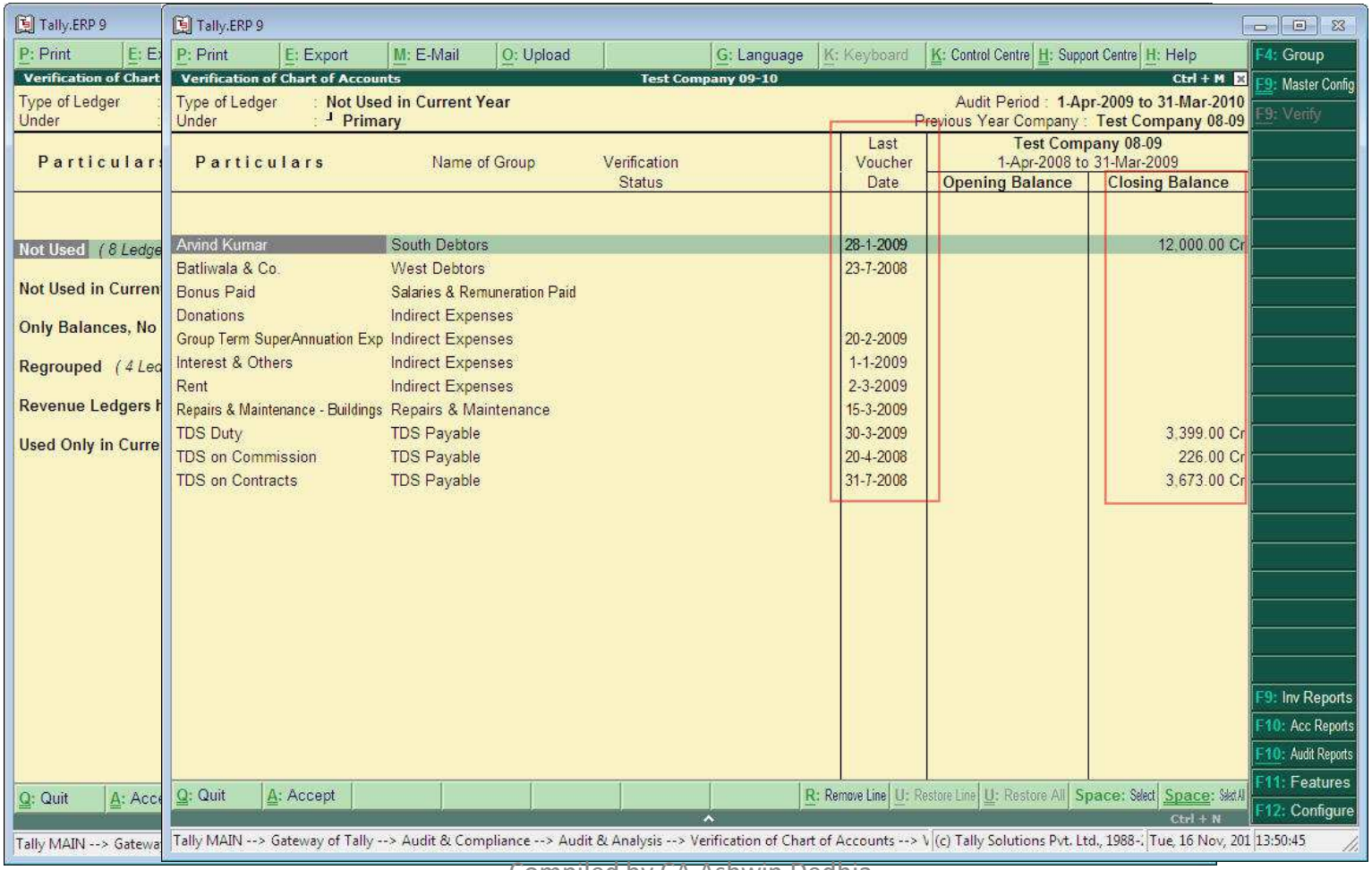


Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Periodic Payments and Receipts

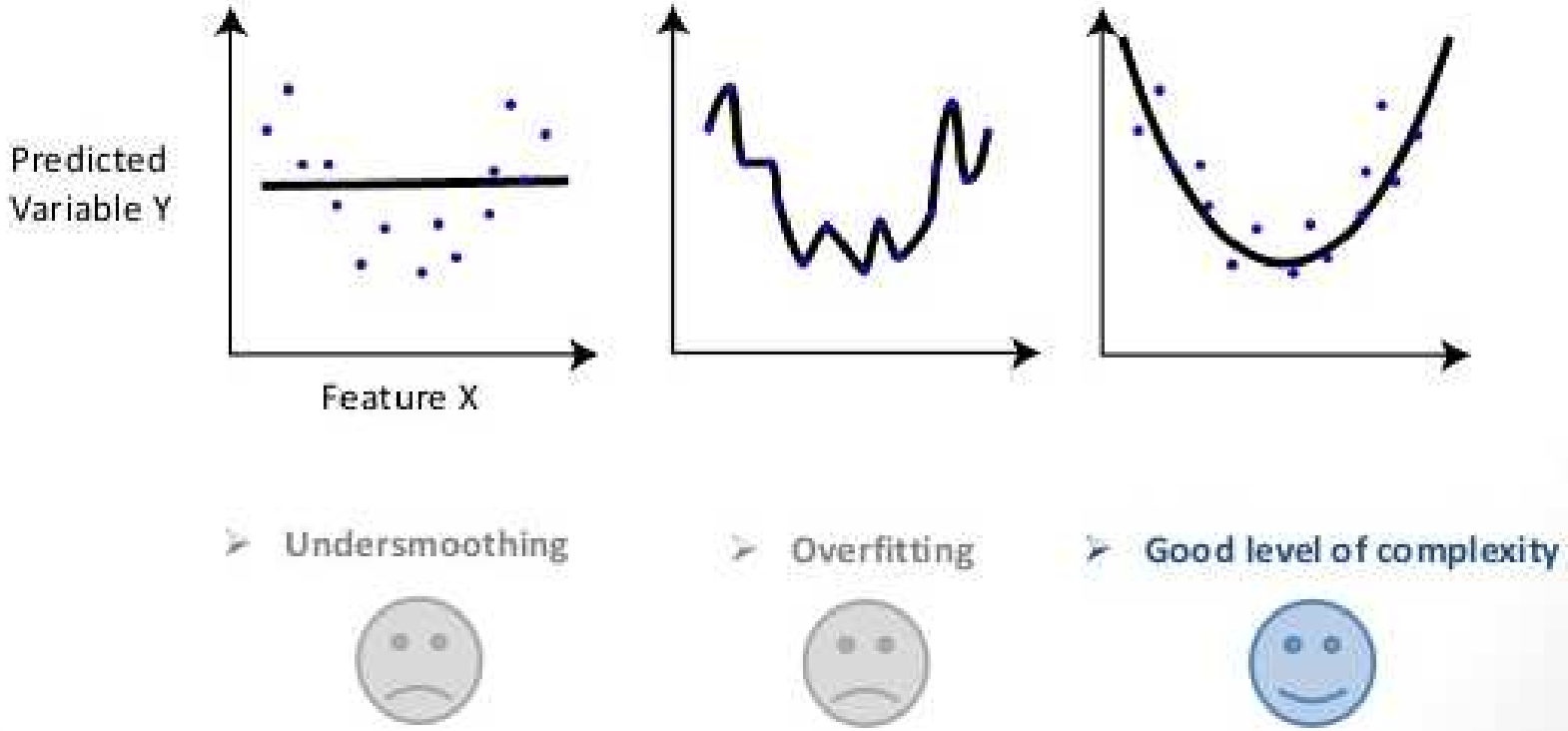


Verification of Chart of Accounts



What is a good Model / Options?

Essentially all models are wrong, but some are useful



Thank You

It’s easy to lie with statistics.
It’s hard to tell the truth without it.
A Dunkels

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com