

Data analysis for Internal Audit

Compiled by CA Ashwin Dedhia

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?

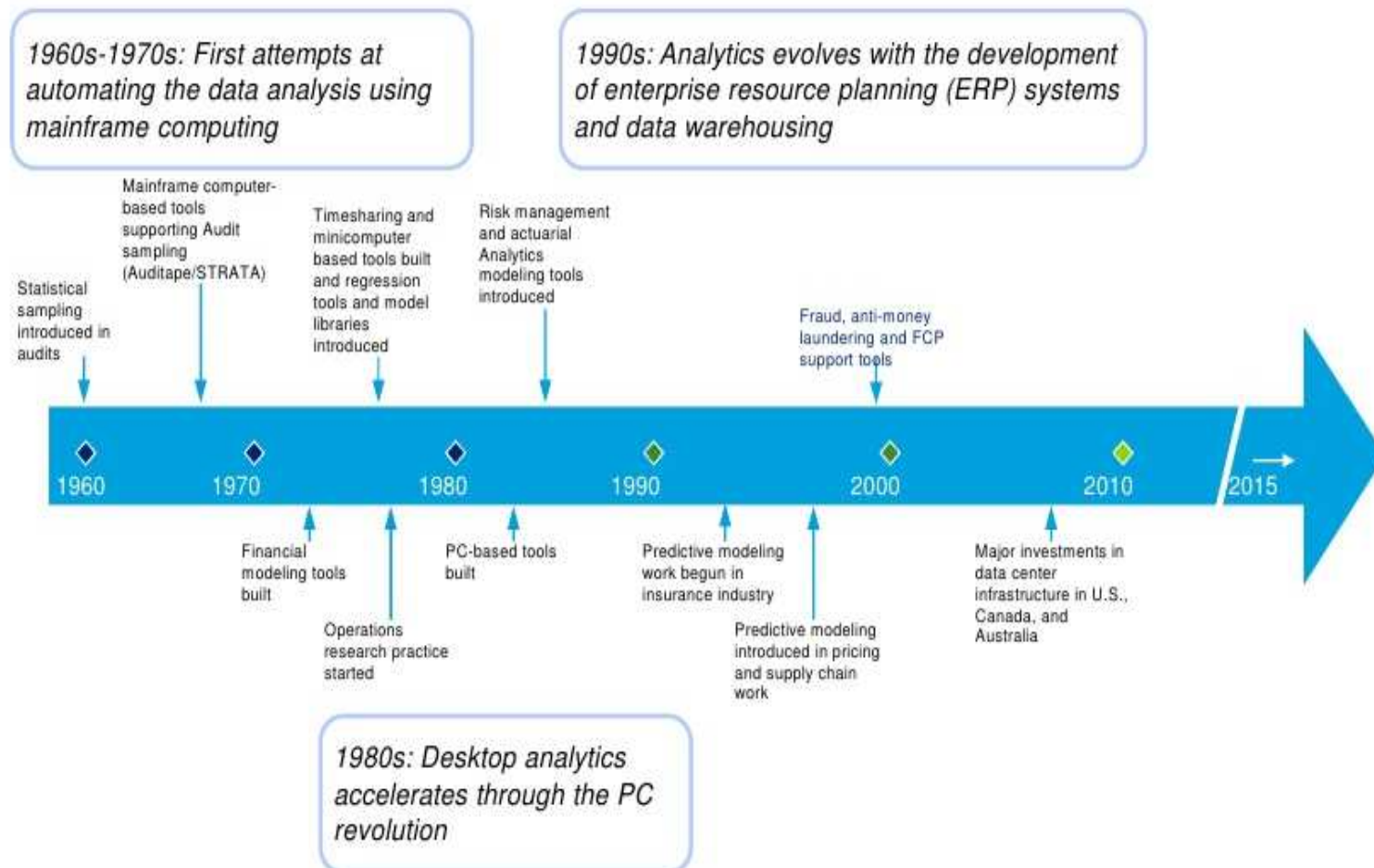
- Common View

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.

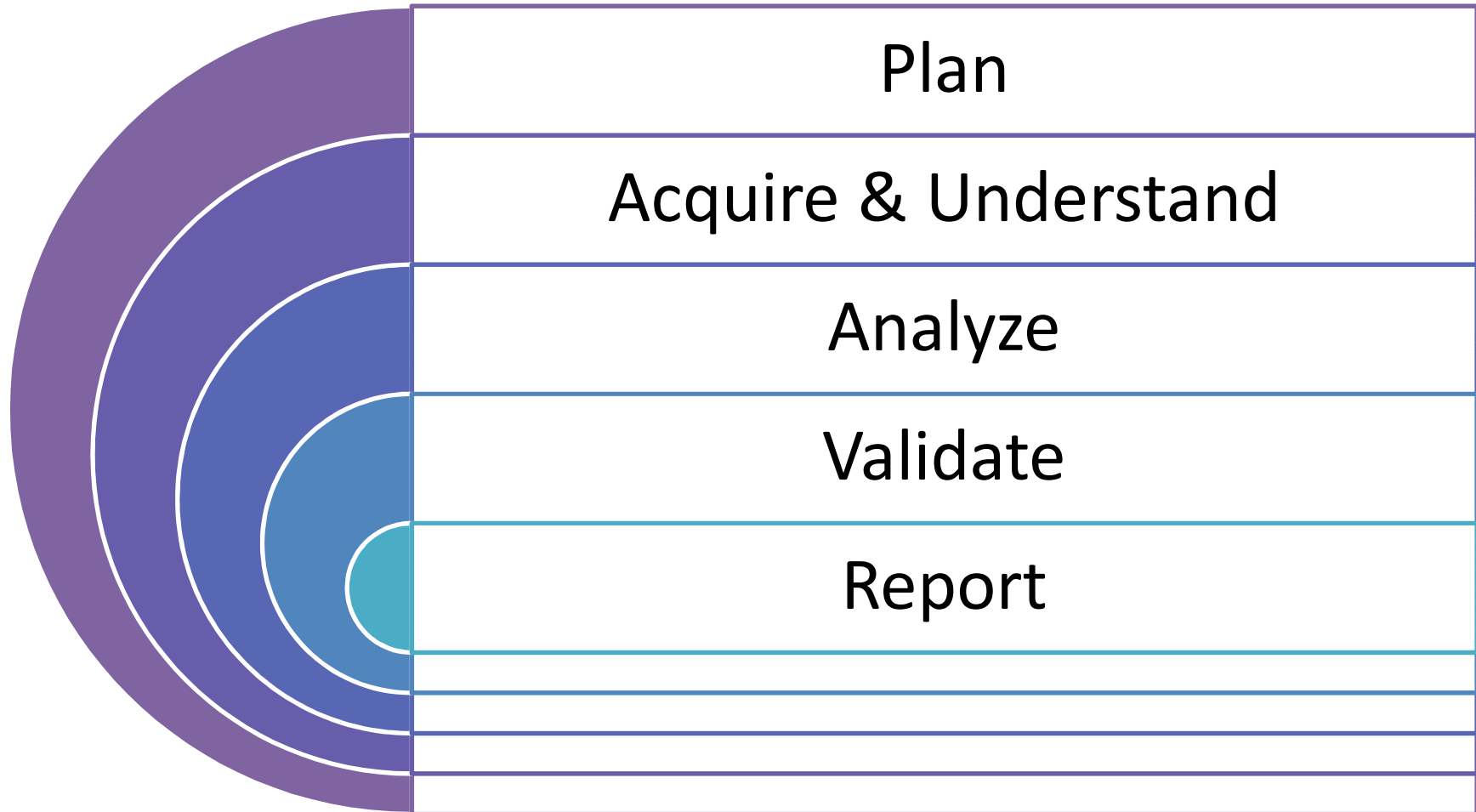
What is Data Analysis

- 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.

It's not something New



Methodology



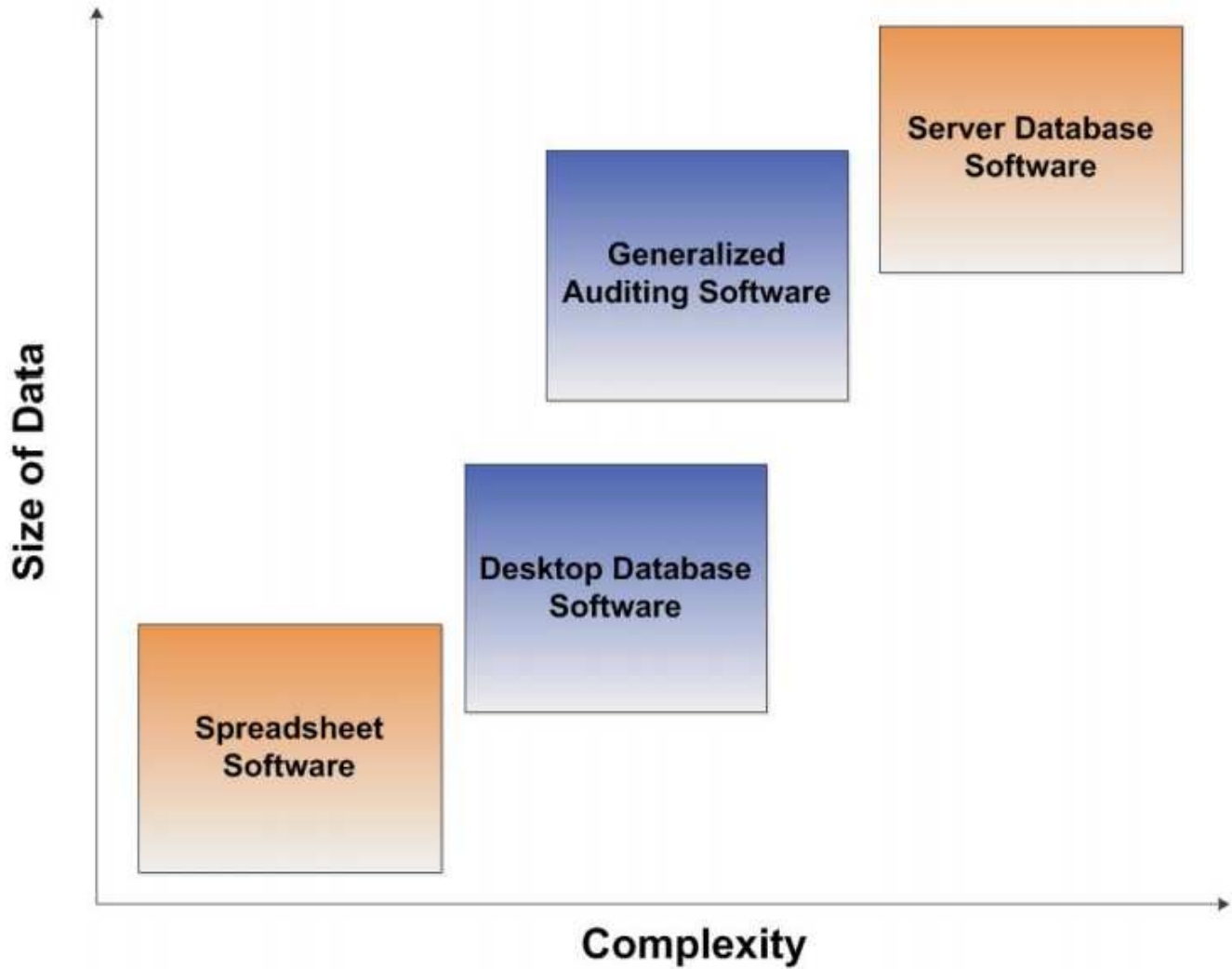
Desired Features of Analysis Tool

- 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

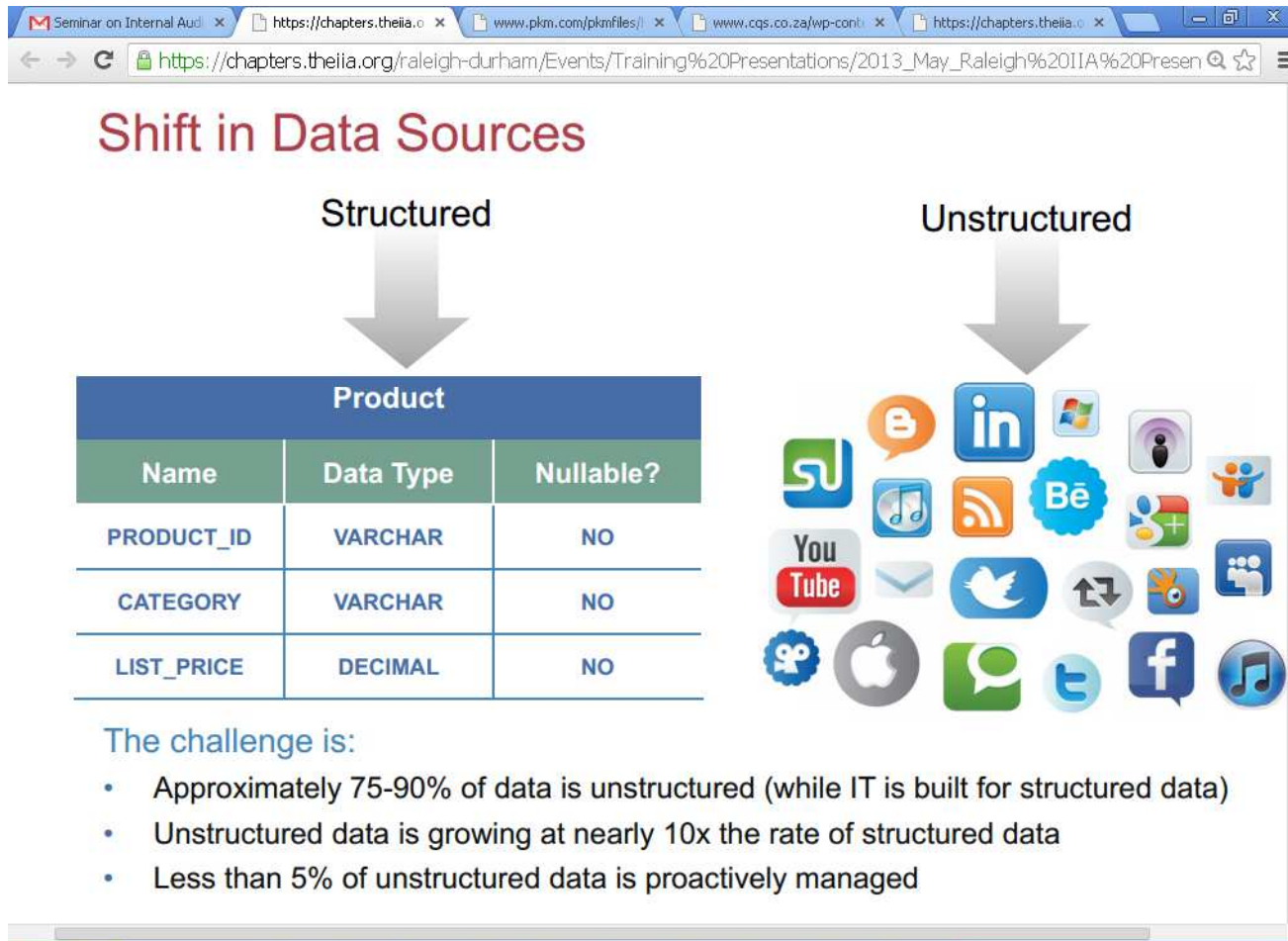
Trade-off



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




The slide is titled "Shift in Data Sources" and is presented within a browser window. It compares structured and unstructured data. On the left, under "Structured", a table lists product attributes. On the right, under "Unstructured", a collection of social media and application icons is shown. Below the icons, a list of challenges is provided.

Structured

Product		
Name	Data Type	Nullable?
PRODUCT_ID	VARCHAR	NO
CATEGORY	VARCHAR	NO
LIST_PRICE	DECIMAL	NO

Unstructured



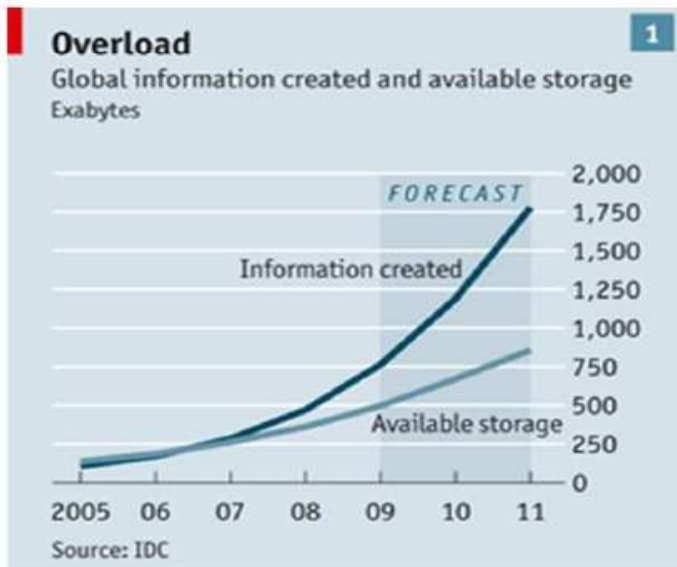
The challenge is:

- Approximately 75-90% of data is unstructured (while IT is built for structured data)
- Unstructured data is growing at nearly 10x the rate of structured data
- Less than 5% of unstructured data is proactively managed

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

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

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

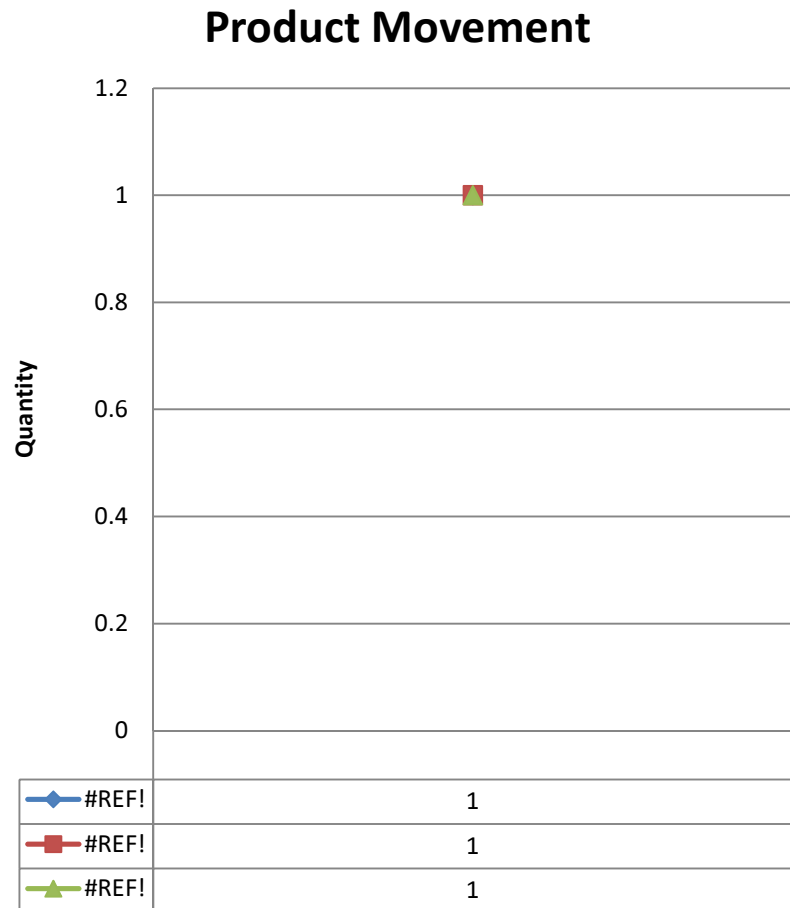
Ratios

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

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

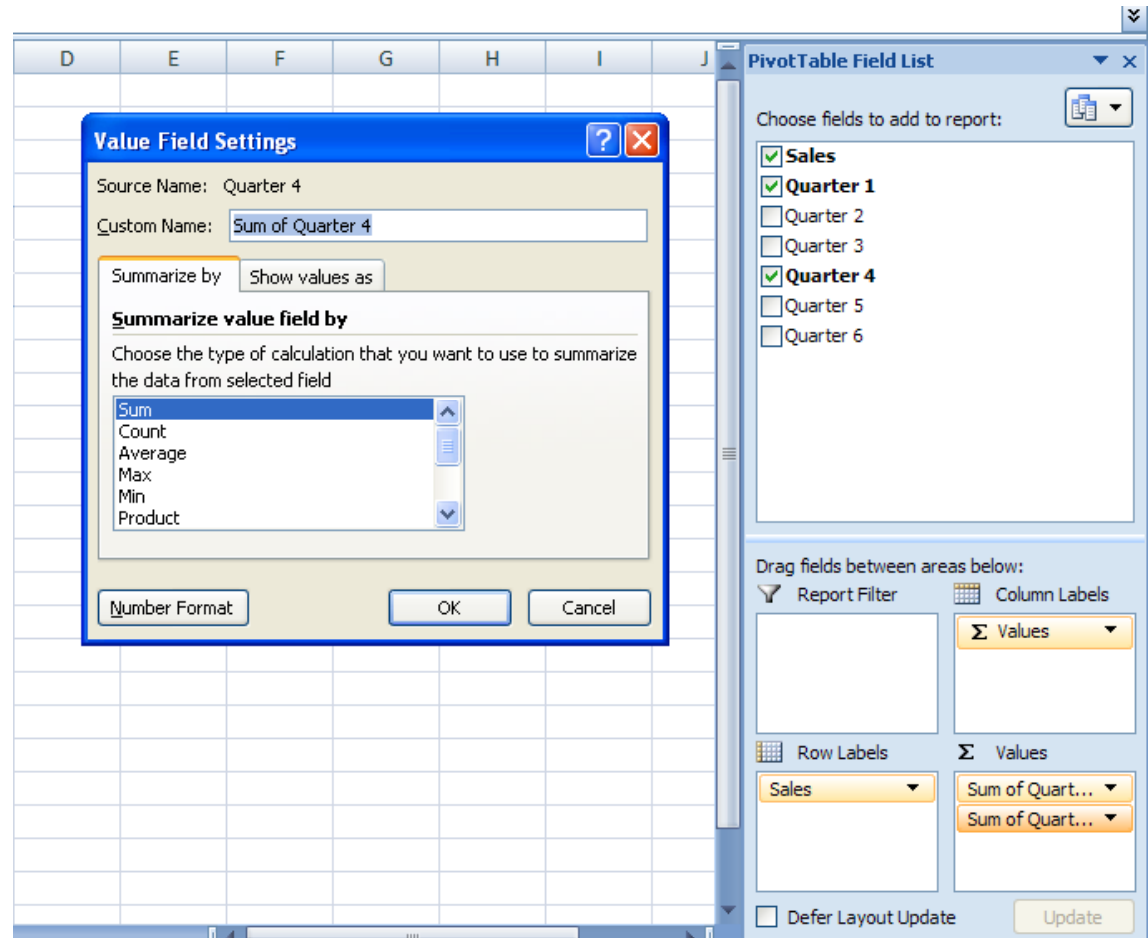
- The identification of critical success factors that can be tracked over time to assess progress made in achieving specific targets linked to an entity's vision. For example, the below represent a sampling of performance measures that could be used for accounts payable processing:
- The Excel calculations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

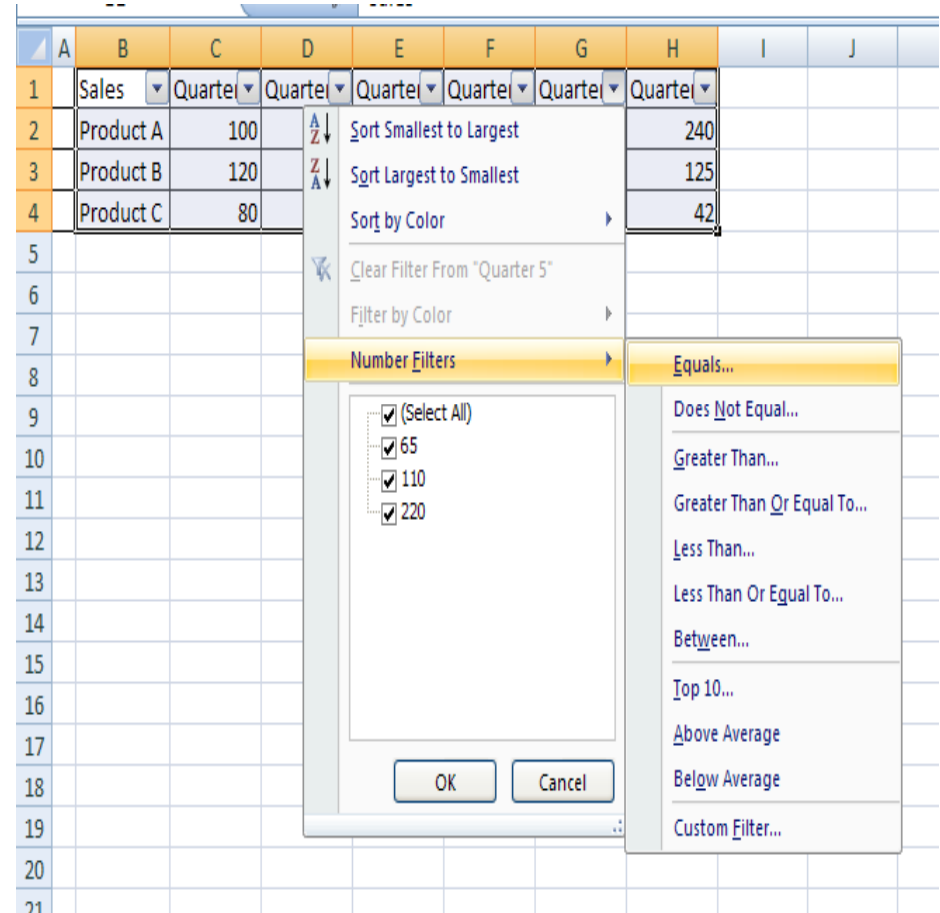
Cross Tabulate / Pivot

- 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	

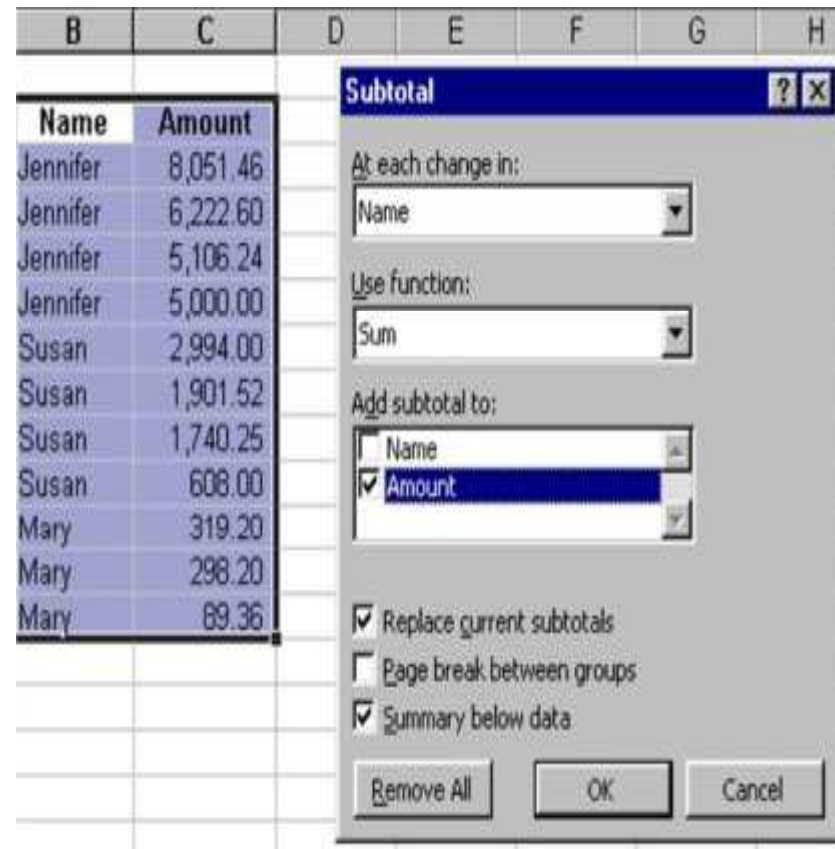
Join / Relate

- 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

Tally.ERP 9

P: Print E: Export M: E-Mail O: Upload S: Shop G: Language K: Keyboard K: Control Centre H: Support Centre H: Help

Audit & Analysis National Traders Pvt. Ltd. (12-13) Ctrl + M

Data Analysis Audit Period : 1-Apr-2012 to 31-Mar-2013
Previous Year Company : National Traders Pvt. Ltd. (11-12)

Particulars Verification Complete?

Verification of Chart of Accounts
Verification of Stock Items
Verification of Balances
Analytical Procedures
Pending Documents
Statutory Payments
Periodic Payments and Receipts
Repeated Transactions
Relative Size Factor (RSF)
Other Analysis

Auditing

Particulars	Total	Sampled	Audited	Internal Observation		External Clarification		Altered (To be re-audited)
				Pending	Clarified	Pending	Clarified	
Verification of Vouchers	1115	192		2				
% of sampled voucher		17 %	0 %					
Related Party Transactions	5							
Forex Transactions	17							
MSME Transactions	24							
Salary / Wages Payments	50							

Q: Quit Space: Select Ctrl + N

Tally MAIN --> Gateway of Tally --> Audit & Compliance --> Audit & Analysis (c) Tally Solutions Pvt. Ltd., 1988-2013 Sat, 16 Nov, 2013 10:52:57

start Tally Auditor S... Tally Auditor E... Tally.ERP 9 10:52 AM

Verification of Stock Items

Tally.ERP 9

Verification of Stock Items

Test Company 09-10

Audit Period : 1-Apr-2009 to 31-Mar-2010
Previous Year Company : Test Company 08-09

Particulars	Name of Group	Verification Status	Test Company 08-09 1-Apr-2008 to 31-Mar-2009 Closing Balance			Test Company 09-10 1-Apr-2009 to 31-Mar-2010 Opening Balance		
			Quantity	Rate	Value	Quantity	Rate	Value
IBM PIV	IBM		1 Nos			2 Nos		
Monitor	Components					1 Nos	5,000.00	5,000.00

Particulars

Type of Stock Item : Difference in Opening Stock
Under : Primary

IBM PIV

Monitor

IBM

Components

1 Nos

2 Nos

1 Nos 5,000.00 5,000.00

Q: Quit

R: Remove Line U: Restore Line U: Restore All Space: Select

Tally MAIN --> Gateway of Tally --> Audit & Compliance --> Audit & Analysis --> Verification of Stock Items --> Verifica (c) Tally Solutions Pvt. Ltd., 1988- Tue, 16 Nov, 201 14:43:48

Verification of Balances

Tally.ERP 9

P: Print E: Export M: E-Mail O: Upload G: Language K: Keyboard K: Control Centre H: Support Centre H: Help

Verification of Opening Balances Test Company 09-10 Ctrl + M

Verification of Opening Balances (Exception Only) Audit Period : 1-Apr-2009 to 31-Mar-2010
Previous Year Company : Test Company 08-09

Under : Fixed Assets

Particulars	Test Company 08-09 1-Apr-2008 to 31-Mar-2009	Test Company 09-10 1-Apr-2009 to 31-Mar-2010	Difference
	Closing Balance	Opening Balance	
Plant & Machinery	15,25,000.00 Dr	11,25,000.00 Dr	4,00,000.00 Dr
Motor Car - Hyundai GLX	9,05,731.88 Dr	9,05,730.00 Dr	1.88 Dr
Grand Total	34,62,489.68 Dr	30,62,487.80 Dr	4,00,001.88 Dr

Q: Quit Ctrl + N

Tally MAIN --> Gateway of Tally --> Audit & Compliance --> Audit & Analysis --> Verification of Balances --> Verification of (c) Tally Solutions Pvt. Ltd., 1988-; Tue, 16 Nov, 201 16:16:22

Analytical Procedures

Tally.ERP 9

P: Print E: Export M: E-Mail O: Upload G: Language K: Keyboard K: Control Centre H: Support Centre H: Help

Analytical Procedures Test Company 09-10 Ctrl + M

Audit Period : 1-Apr-2009 to 31-Mar-2010
Previous Year Company : Test Company 08-09

Comparison : Sales Accounts (Group) Vs Indirect Expenses (Group)

Sales Accounts
Previous Period (1-Apr-2008 - 31-Mar-2009) Test Company 08-09 : 4,04,88,716.30 Cr 100.00 % (A)
Current Period (1-Apr-2009 - 31-Mar-2010) Test Company 09-10 : 3,49,43,762.28 Cr 86.30 % (B) as % of A
Variance : 55,44,954.02 13.70 % Decrease

Particulars	Test Company 08-09 1-Apr-2008 to 31-Mar-2009		Test Company 09-10 1-Apr-2009 to 31-Mar-2010		Variance	
	Closing Balance (X)	% of A (Sales Accounts)	Closing Balance (Y)	% of B (Sales Accounts)	Amount (Y - X)	% (Y - X) / X
Indirect Expenses	48,37,206.94 Dr	11.95 %	69,02,362.35 Dr	19.75 %	20,65,155.41	42.69 %
Depreciation on Fixed Assets	6,02,622.20 Dr	1.49 %	4,92,200.00 Dr	1.41 %	(1,10,422.20)	18.32 %
Employer ESI Contributions	8,997.00 Dr	0.02 %	8,811.00 Dr	0.03 %	(186.00)	2.07 %
Employer PF Contributions	45,243.29 Dr	0.11 %	44,303.11 Dr	0.13 %	(940.18)	2.08 %
Interest & Finance Charges	3,249.70 Dr	0.01 %	7,071.49 Dr	0.02 %	3,821.79	117.60 %
Operating Expenses	18,28,606.75 Dr	4.52 %	5,94,123.75 Dr	1.70 %	(12,34,483.00)	67.51 %
Salaries & Remuneration Paid	14,56,306.00 Dr	3.60 %	24,73,221.00 Dr	7.08 %	10,16,915.00	69.83 %
Selling & Distribution Expenses	24,222.00 Dr	0.06 %	36,172.00 Dr	0.10 %	11,950.00	49.34 %
Advertisement Expenses			8,22,360.00 Dr	2.35 %	8,22,360.00	
Bonus	52,500.00 Dr	0.13 %	67,050.00 Dr	0.19 %	14,550.00	27.71 %
Business Promotion Expenses	25,500.00 Dr	0.06 %	2,90,000.00 Dr	0.83 %	2,64,500.00	1,037.25 %
Commission Charges			70,000.00 Dr	0.20 %	70,000.00	
Conveyance Expenses			25,000.00 Dr	0.07 %	25,000.00	
Donations	11,311.00 Dr	0.03 %			(11,311.00)	
Godown Rent			1,30,000.00 Dr	0.37 %	1,30,000.00	
Interest			10,000.00 Dr	0.03 %	10,000.00	
Legal Fees	16,500.00 Dr	0.04 %	27,500.00 Dr	0.08 %	11,000.00	66.67 %
Maintenance Charges			50,000.00 Dr	0.14 %	50,000.00	
Miscellaneous Expenses	13,210.00 Dr	0.03 %	10,800.00 Dr	0.03 %	(2,410.00)	18.24 %
					8 more ...	

Q: Quit

Tally MAIN --> Gateway of Tally --> Audit & Compliance --> Audit & Analysis --> Analytical Procedures (c) Tally Solutions Pvt. Ltd., 1988- Tue, 16 Nov, 201 17:21:52

Pending Documents

The screenshot displays the Tally.ERP 9 interface with a 'Pending Purchase Order Summary' report. The report is for 'Test Company 09-10' and covers the period from 1-Apr-2009 to 31-Mar-2010. The report is organized into a table with the following columns: Particulars, Orders, Opening Value, Received Value, and Closing Value. Two items are listed: Challenger Systems and Top Computers. The summary row at the bottom shows a total of 2 orders with an opening value of 4,06,840.00 Dr, a received value of 20,500.00 Dr, and a closing value of 3,86,340.00 Dr.

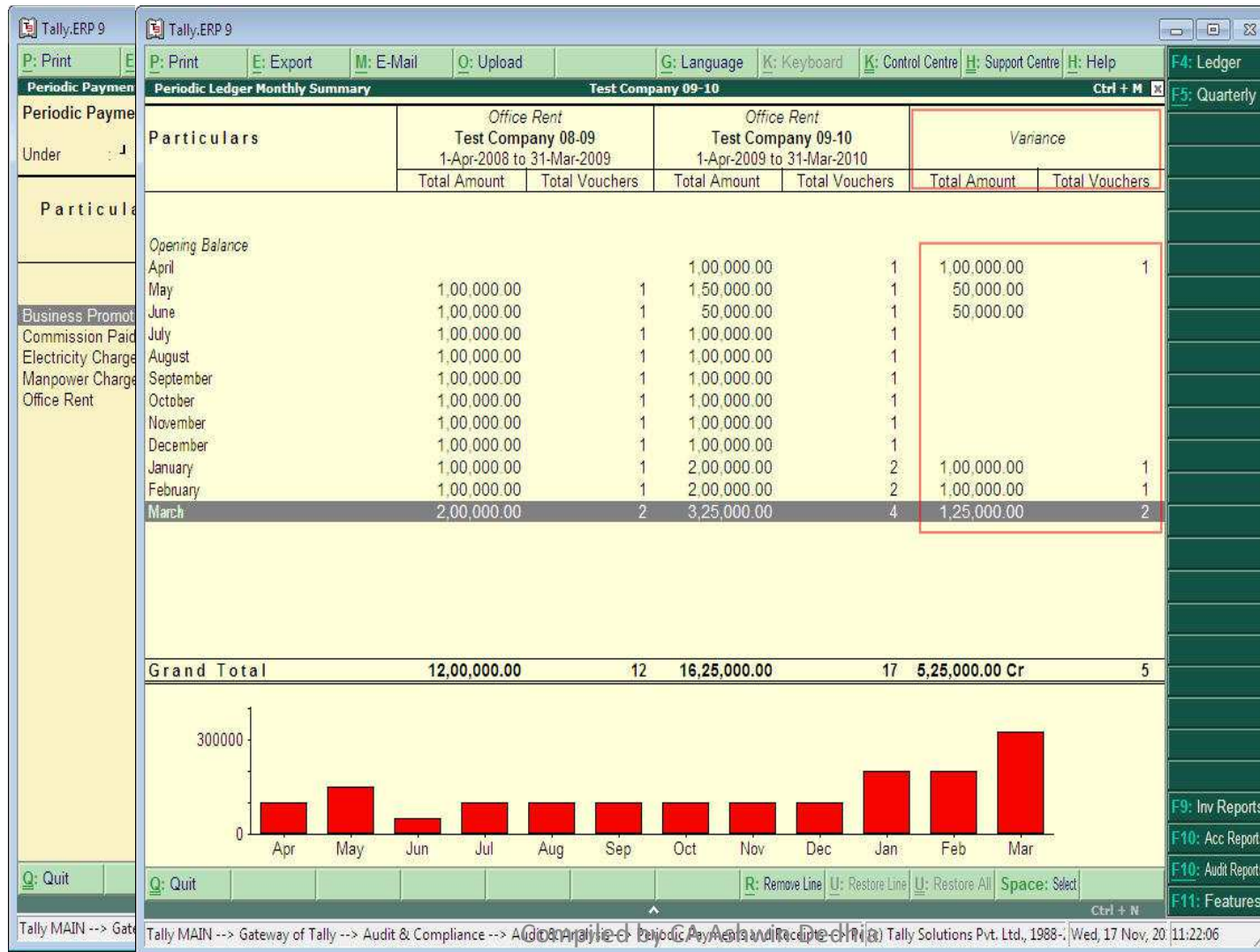
Particulars	Orders	Opening Value	Received Value	Closing Value
Challenger Systems	1	41,000.00 Dr	20,500.00 Dr	20,500.00 Dr
Top Computers	1	3,65,840.00 Dr		3,65,840.00 Dr
	2	4,06,840.00 Dr	20,500.00 Dr	3,86,340.00 Dr

Compiled by CA Ashwin Dedhia
ashwin@dedhia.com

Fixed Assets

Particulars		No. of Days		Opening Balance	Additions	Deletions / Adjustments	Closing Balance
		< 180	> 180				
↑ ... 26 more							
Land - I				1,50,000.00			1,50,000.00
1-4-2009	Opening Balance	1,50,000.00	Dr				
Land - II				5,50,000.00	7,75,000.00		13,25,000.00
1-4-2009	Opening Balance	5,50,000.00	Dr				
6-7-2009	43 Purchase			269	7,75,000.00		
Plant & Machinery				11,25,000.00			11,25,000.00
Plant & Machinery - I				10,00,000.00			10,00,000.00
1-4-2009	Opening Balance	10,00,000.00	Dr				
Plant & Machinery - II				1,25,000.00			1,25,000.00
1-4-2009	Opening Balance	1,25,000.00	Dr				
Airconditioner				21,500.00	23,000.00		44,500.00
1-4-2009	Opening Balance	21,500.00	Dr				
8-10-2009	234 Payment			175	23,000.00		
Computers & Peripherals				90,460.00	45,000.00		1,35,460.00
1-4-2009	Opening Balance	90,460.00	Dr				
2-5-2009	22 Journal			334	45,000.00		
Furniture & Fixtures				1,22,420.00	1,25,000.00	5,800.00	2,41,620.00
1-4-2009	Opening Balance	1,22,420.00	Dr				
17-5-2009	25 Journal			319	1,25,000.00		
2-2-2010	87 Journal			58		5,800.00	
Motor Car - Hyundai GLX				9,05,730.00	1,78,100.00		10,83,830.00
1-4-2009	Opening Balance	9,05,730.00	Dr				
27-4-2009	13 Journal			339	1,78,100.00		
Grand Total				42,67,732.20	13,71,100.00	4,96,000.00	61,23,232.20

Periodic Payments and Receipts

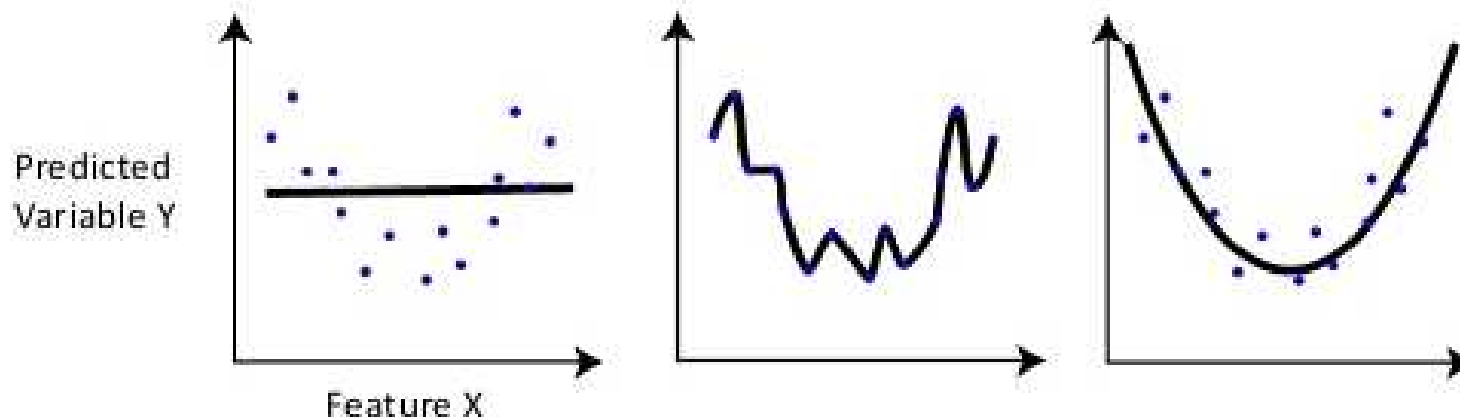


Verification of Chart of Accounts

Verification of Chart of Accounts		Test Company 09-10		Audit Period : 1-Apr-2009 to 31-Mar-2010	
Type of Ledger : Not Used in Current Year		Previous Year Company : Test Company 08-09		Test Company 08-09	
Under : Primary		1-Apr-2008 to 31-Mar-2009		Opening Balance	Closing Balance
Particulars	Name of Group	Verification Status	Last Voucher Date		
Not Used (8 Ledgers)	Anind Kumar	South Debtors	28-1-2009		12,000.00 Cr
Not Used in Current Year	Batliwala & Co.	West Debtors	23-7-2008		
Only Balances, No	Bonus Paid	Salaries & Remuneration Paid			
Regrouped (4 Ledgers)	Donations	Indirect Expenses			
Revenue Ledgers h	Group Term SuperAnnuation Exp	Indirect Expenses	20-2-2009		
Used Only in Current Year	Interest & Others	Indirect Expenses	1-1-2009		
	Rent	Indirect Expenses	2-3-2009		
	Repairs & Maintenance - Buildings	Repairs & Maintenance	15-3-2009		
	TDS Duty	TDS Payable	30-3-2009		3,399.00 Cr
	TDS on Commission	TDS Payable	20-4-2008		226.00 Cr
	TDS on Contracts	TDS Payable	31-7-2008		3,673.00 Cr

What is a good Model / Options?

Essentially all models are wrong, but some are useful



➤ Undersmoothing



➤ Overfitting



➤ Good level of complexity



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