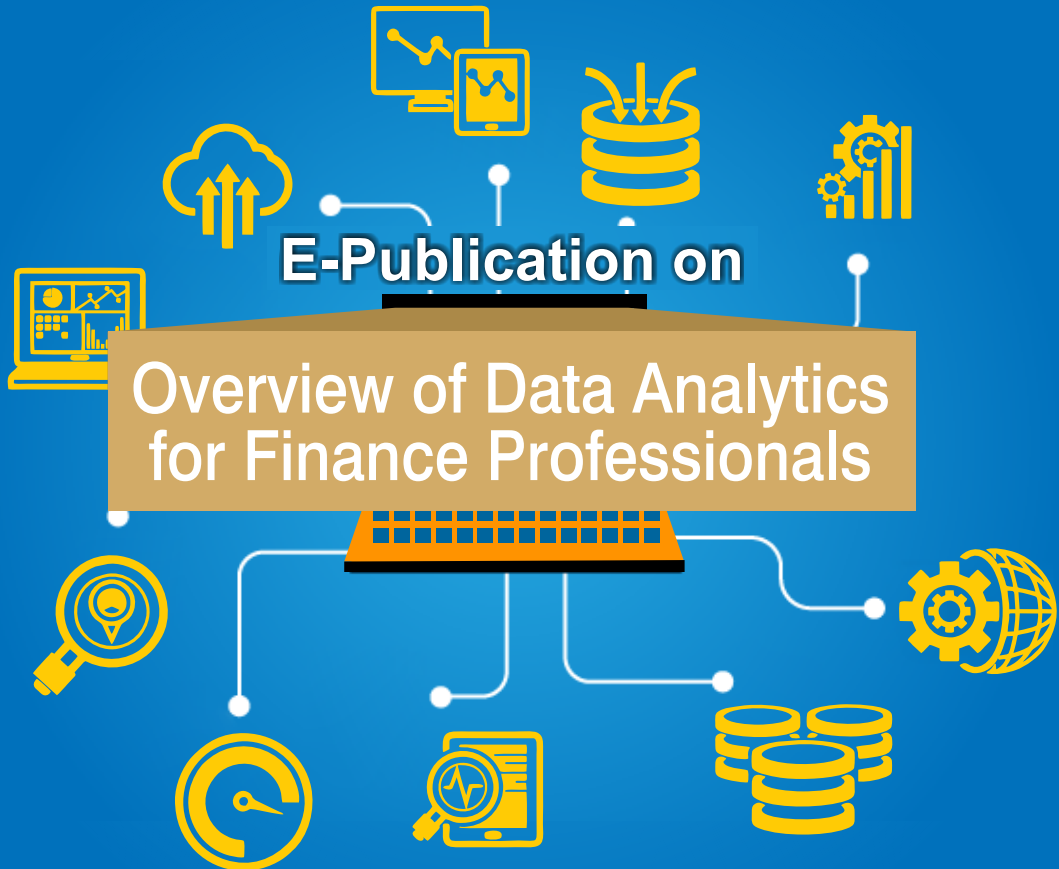




Western India Regional Council of The Institute of Chartered Accountants of India

(Set up by an Act of Parliament)



© WESTERN INDIA REGIONAL COUNCIL OF
THE INSTITUTE OF CHARTERED ACCOUNTANTS OF INDIA

Published by

**CA. Lalit Bajaj, Chairman, WIRC, Western India Regional Council of
The Institute of Chartered Accountants of India,**

ICAI Tower, Plot No. C-40, G Block, Opp. MCA Ground,
Next to Standard Chartered Bank, Bandra-Kurla Complex,
Bandra (East), Mumbai-400 051

Tel.: 022-336 71400 / 336 71500 • E-mail: wirc@icai.in • Web.: www.wirc-icai.org

Disclaimer

Opinions expressed in this book are those of the Contributors. Western India Regional Council of The Institute of Chartered Accountants of India, does not necessarily concur with the same.

While every care is taken to ensure the accuracy of the contents in this compilation, neither contributors nor Western India Regional Council of The Institute of Chartered Accountants of India is liable for any inadvertent errors or any action taken on the basis of this book.

Foreword



The year 2020 has seen a huge leap forward in adoption of digitalisation amongst professionals. All professionals, whether in industry or in practice, understand that as technology plays a bigger role in the day to day aspects of a business, we have to move with the times as technology and accountancy are slowly merging.

Words like blockchain, artificial intelligence, machine learning, and data analysis have become common amongst Chartered Accountants. While all of these buzzwords and their related concepts are important, none seems to be quite as important as data analytics.

Data Analytics is becoming commonplace in businesses as companies try to sift through huge amounts of information because interpreting data effectively can improve business intelligence and decision making significantly.

As Chartered Accountants we are trained to observe, analyse and even predict trends. Understanding the importance of future trends, this year WIRC has focused intensively on topics which impact the profession and generated publications which put our members and students on the path of future growth.

This publication namely, 'Overview of Data Analytics for Finance Professionals' is one such study which will help members and students become aware of big data analytics and its impact on the accounting profession, and every industry in general.

It will give you a closer look at data analytics and what it really means for us in this field. Accountants can use data analysis when reviewing financial information to ensure the company is running well, meeting goals, and maintaining or improving performance. This knowledge is essential to both a business's sustainability and survival.

For Members in practice, data analytics can be used to improve the client experience helping to retain and bring on new clients.

The importance of big data is only going to grow in the future. Learning about how data is used in accounting now will put you in a great position to become an asset in any role you play in the accounting profession.

I thank CA. Yashwant Kasar, RCM, for his dedication toward ensuring the publishing of this study as well as the team of contributors comprising CA. Anant Govande, CA. Alok Jajodia, CA. Dinesh Kumar Tejwani, CA. Mitesh Katira and CA. Shirish Padey.

I am confident the 'Overview of Data Analytics for Finance Professionals' publication will prove to be of tremendous benefit to all members and I am sure these insights will help everybody get a better grasp of this important topic of data analytics.

I congratulate all the contributors and members for this publication.

CA. Lalit Bajaj
Chairman, WIRC

Preface



As technology continues to evolve, it promotes changes to business models and in an increasingly data-driven world, Chartered Accountants need to be able to adapt to these technological changes.

Across the world and increasingly in India, Chartered Accountants find themselves requiring skills in data analytics as its use becomes widespread. Today, internal and external auditors use data analytics for continuous auditing, financial planning and to discover new consumer and market trends to drive company strategy.

In order to bring this subject to the forefront, WIRC is publishing this insightful book on 'Overview of Data Analytics for Finance Professionals'. It will highlight the use of analytics in auditing, by focusing on integrating analytics into the audit process and on defining how analytics can be used to enhance audit quality.

This book also highlights future trends, insights into usage of tools for data analysis, commonly used statistical models for practical data analytics and usage of data analytics.

I thank WIRC Chairman CA Lalit Bajaj for his encouragement and belief in the importance of practical publications for the education of our members. I would like to thank and appreciate the depth of work put forward by our contributors namely CA Anant Govande, CA Alok Jajodia, CA Dinesh Kumar Tejwani and CA Mitesh Katira for their commendable efforts to bring out this book within a short span of time. I also specially thank CA Shirish Padey for devoting time and reviewing the book with some very useful inputs.

I am sure that this book will be an asset to all the members looking to upgrade their skill-sets and ensuring they transform into future ready professionals.

CA Yashwant Kasar
RCM, WIRC

List of Contributors

CA Anant Govande
CA Alok Jajodia
CA Dinesh Kumar Tejwani
CA Mitesh Katira

Reviewer

CA Shirish Padey



**WESTERN INDIA REGIONAL COUNCIL OF
THE INSTITUTE OF CHARTERED ACCOUNTANTS OF INDIA
2020-21**

Chairman

CA. Lalit Bajaj

Vice-Chairman

CA. Vishal Doshi

Secretary

CA. Murtuza Kachwala

Treasurer

CA. Anand Jakhotiya

Members

CA. Abhijit Kelkar	CA. Arpit Kabra
CA. Arun Anandagiri	CA. Balkishan Agarwal
CA. Chintan Patel	CA. Drushti Desai
CA. Hitesh Pomal	CA. Jayesh Kala
CA. Kamlesh Saboo	CA. Manish Gadia
CA. Priti Savla	CA. Shilpa Shinagare
CA. Sushrut Chitale	CA. Rakesh Alshi
CA. Umesh Sharma	CA. Vikash Jain
CA. Vimal Agrawal	CA. Yashwant Kasar

Ex-Officio Members

CA. Nihar Jambusaria, *Vice-President, ICAI*

CA. Prafulla Chhajed, *Past President, ICAI*

CA. Aniket Talati	CA. Anil Bhandari
CA. Chandrashekhar Chitale	CA. Dheeraj Khandelwal
CA. Durgesh Kabra	CA. Jay Chhaira
CA. Nandkishore Hegde	CA. Shriniwas Joshi
CA. Tarun Ghia	

INDEX

Sr. No.	Particulars	Pg. No.
1.	Why Data Analytics	1
	Some Important Terms	2
	Types of Data Analytics	5
	Users of Data Analytics	7
	Data Analytics - Advantages	8
	Data Analytics - Challenges	8
2.	Data Analytics in Audit	9
	Changing audit scenario- Requirement for Fraud reporting	9
	Issues/challenges with the usage of Data Analytics in audit	13
3.	Data Analytics in Other Finance Functions	15
	Benchmarking	15
	Management Information	16
	Investor and Lender information	17
	Tax Function	17
	Financial Statement Analytics	18
4.	Tools for Data Analysis	20
	Data Smoothing/Transformation Tools	20
	Data Analysis Tools	21
	Visualisation tools	24

Sr. No.	Particulars	Pg. No.
5.	Data Analytics: Current & Future Technology Trends	27
	Algorithms	28
	Artificial Intelligence (AI) and Machine Learning (ML)	29
	Internet of Things (IoT)	31
	Big Data Analytics with data mining	32
	Blockchain	33
	Future of Data Analytics using new-age technologies	34
6.	Commonly used Statistical Models for Practical Data Analytics	37
	Mean	38
	Median	38
	Mode	39
	Standard Deviation	40
	Correlation	41
	Regression	43
7.	Data Analytics Usage	44
	Professional Firms	44
	Finance professionals in the Industry	45
	Finally A Word of Caution	47

Why Data Analytics



In today's information age, data is everywhere and ever-increasing. We are leaving digital footprints when we use the internet right from desktops, laptops, mobiles to wristbands!!

Traditionally businesses have been capturing data from internal accounting, and operations like CRM, payroll, inventory. But in the digital age external data like website traffic, social media feeds etc. have become equally critical.

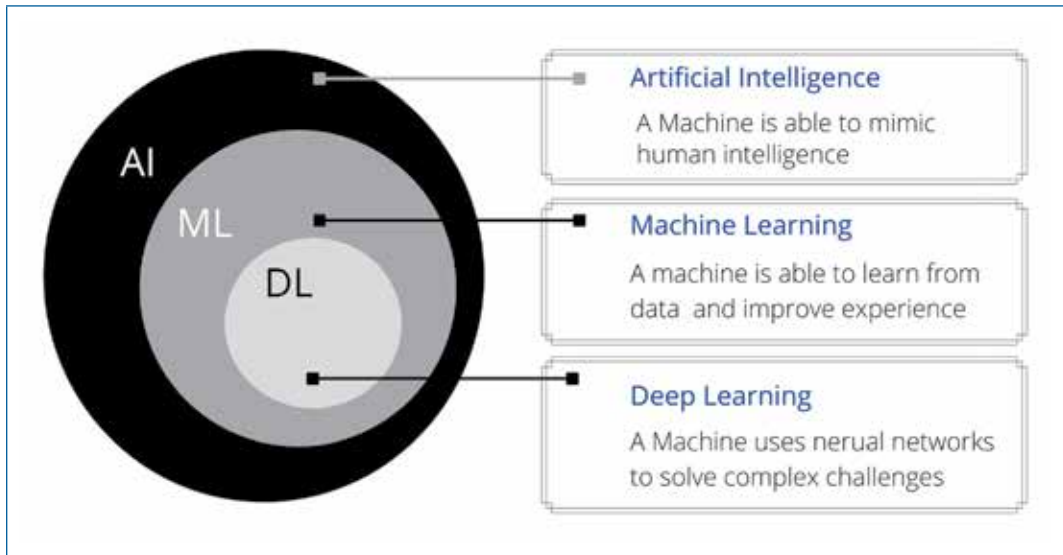
Increased data is not just limited to businesses; all government offices, defence offices, scientific research organisations and other arms of the modern society are also looking at ever-increasing data.

Data analytics can be defined as a science that analyses raw data to draw conclusions about the information contained therein.

For a chartered accountant, data analytics is a relatively new concept. To put it simply, data analytics enables an auditor to work on, assess, conclude, manipulate 100% of data, rather than just a sample of the population. However, the use is not just limited to audit and assurance, as we see in the following chapters.

Some Important Terms

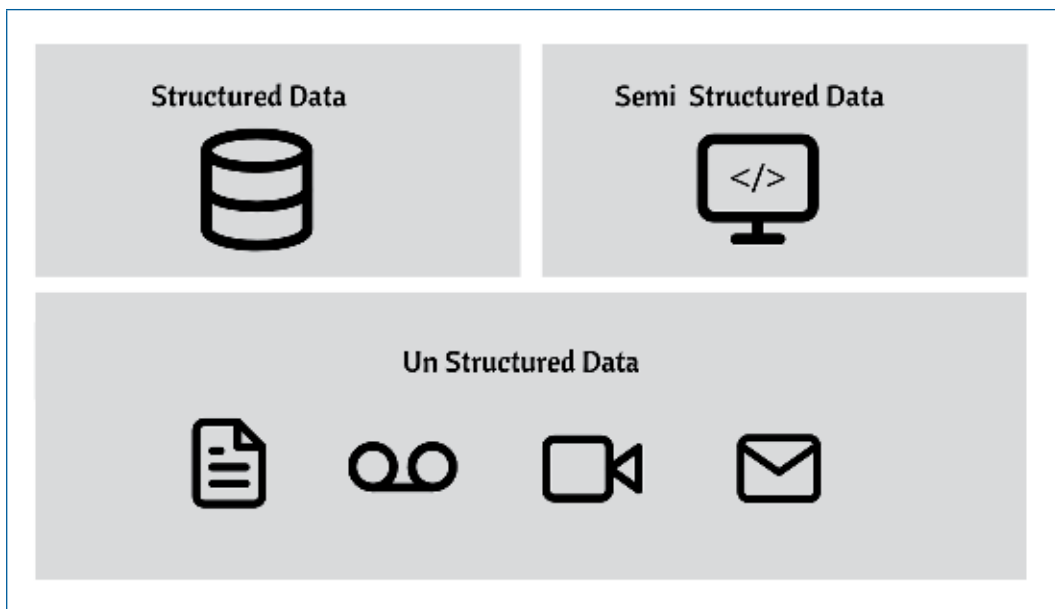
- A. **Database (DB):** Data stored in a database is in the form of tables, consisting of rows and columns.
- B. **Management Information Systems (MIS):** Management Information Systems are essentially Reports, designed to address standard and repetitive information required periodically by management for decision making. MIS typically execute predefined queries on the database.
- C. **Data Warehouse (DW):** Data stored in a database in the form of tables, may then be moved to a Staging Area for transformation, and then to a Data Warehouse. Data stored in a Data Warehouse is in the form of virtual multi-dimensional cubes (not tables). Typically, data in a Data Warehouse is never purged, as the more the data, the better will be the understanding of hidden trends, etc.
- D. **Artificial Intelligence (AI):** Artificial intelligence is a technology which enables a machine to simulate human behaviour. It is based on the experience of the tool, rather than being clearly programmed. So, unlike MIS, AI is expected to give answers to random questions on data.
- E. **Machine Learning (ML):** ML is also a subset of AI. Unlike BI, ML may be used for non-business purposes.
- F. **Deep learning (DL):** DL is a subset of ML, which imitates the working of the human brain without human supervision. It can process data and create patterns similar to the human brain and use them in decision making. It is also known as neural learning, used for detecting objects, recognizing speech and translating language.



- G. **Business Intelligence (BI)**: BI is essentially a subset of AI. BI are the tools and methodologies that make use of large business data to make intelligent business decisions. BI uses past and present data to make better decisions for current business operations.
- H. **Business Analytics (BA)**: Business analytics uses past and present data to predict future trends. BA tools are used to help make strategic decisions regarding new market opportunities, improving customer relationships etc. Unlike BI which takes a thorough look at past, present data, BA engages human intelligence and individual perspective to arrive at conclusions about the next plan of action. BA has to do with viewpoints and foresight, and that can be very subjective.
- I. **Advanced Analytics**: This data science uses scientific methods and processes to extract insights and knowledge from structured and unstructured data, generally large data. Advanced Analytics is the autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, typically beyond BI and BA, to discover deeper insights, make predictions, or generate recommendations.

Example: COMET Neowise was discovered by analysing data acquired by a space telescope.

J. Big Data



The term big data is frequently used while describing data analytics. So, what is big data?

Big data deals with all kinds of data:

- Structured data: data organised in a formatted database
- Unstructured data: data does not have a predefined data model. Examples: documents, images, videos, sensor data, audio, emails etc.
- Semi-structured data: mainly unorganised, but has some associated information like metadata tagging

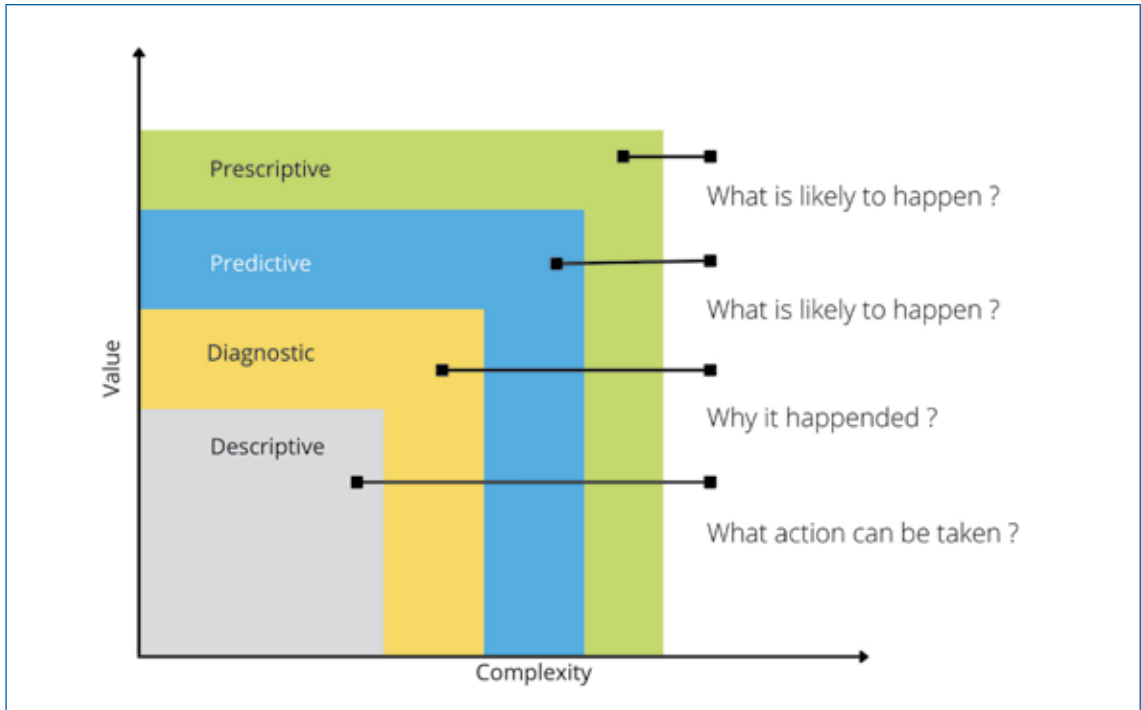
Although big data doesn't equate to any specific volume of data, big data deployments often involve terabytes (TB), petabytes (PB) and even exabytes (EB) of data captured over time.

Big data analytics can lead to positive business-related outcomes like:

- New revenue opportunities
- More effective marketing
- Better customer service

- Improved operational efficiency
- Competitive advantages over rivals

Types of Data Analytics



a. Descriptive Analytics

This is the most basic type of analytics and deals with numbers. It provides 'what happened' by quantitatively analysing data using statistical methods e.g. sum, mean, mode, percentage, frequency etc. Various revenue reports, KPI Dashboards are some of the common examples

b. Diagnostic Analytics

Diagnostics tend to provide an answer as to 'why it happened'. While a revenue report may show the actual figure of decline in sales over periods, the diagnostic analytics will look into the patterns and deviations to seek root cause analysis and provide reasons for the decline in the sales. Diagnostics involve co-relating two different datasets.

c. Predictive Analytics

Predictive analytics is concerned with the future and comes up with answers as to 'what is likely to happen'. Businesses use this to predict the demand for their products.

While description analysis shows sales data over quarters/years, diagnostic analytics provide the cause for decline/increase. Predictive analytics will co-relate these data to external data like demographics to guess future sales.

d. Prescriptive Analytics

This is the most advanced form of analytics and comes up with recommendations as to 'what action is to be taken'. This requires huge computing power as it tries to work out multiple scenarios, predict outcomes for each such scenario and then recommend the best one. Artificial Intelligence and Machine Learning use prescriptive analytics.

Examples

Description Analytics	Monthly State-wise sales during 2020-21 show sales down 24% during April-June, 12% during July-Sept 2% during Oct-Dec
Diagnostic Analytics	Lockdown caused by COVID-19 and consequent closure of retail outlets caused the decline
Predictive Analytics	With lockdown relaxed and normal activity rebounding, sales during Quarter IV will be 5% higher than the same during LY
Prescriptive Analytics	Shifting to a hybrid model of instore and online sale may result in an increase of 16% during 2021-22

e. Exploratory and Confirmatory Data Analytics

In Exploratory data analytics, no clear hypothesis is stated before analysing the data, and the data is searched for models that describe the data well. Exploratory data analytics starts by framing questions about what to do with data and then finding out the best way to manipulate data to find the answers. This involves understanding data structure, identifying missing or erroneous data, establishing a margin of error and figuring out hypotheses.

Confirmatory data analytics, on the other hand, sets out to test the stated hypothesis using statistical tools and conclude your findings.

Users of Data Analytics

- a. Businesses:** Businesses are using data analytics to improve operations, and strategize the future. Most modern e-commerce businesses are thriving only because of advance usage of data analytics.

Example: Netflix uses viewership data to recommend the shows you are likely to be interested in. More than that, it can use these large data to create its own content and direct it to the target audience, virtually guaranteeing success

- b. Governments:** Governments are using data analytics to predict major problems and take pre-emptive action before they turn into crises. The areas where they are being put to use are natural calamities, defence, internal security, health care etc.

Example: The Government of Indonesia uses sensors and public complaints to predict flood-prone areas and expedite response.

- c. Various scientific disciplines:** Use of data analytics in scientific research cannot be overemphasised. Few examples:

- Use of satellite imagery to analyse and predict climate models environmental changes
- Use of patient data to prevent epidemics, reduce healthcare cost, improve timely healthcare

- d. Tax Departments:** World over, tax authorities are collecting more data, sharing data among various other government departments and are using this big data to increase tax revenue, enforce compliance and bring transparency in tax administration.

In India, with the introduction of GST, an era of 100% digital tax has begun. What started as monthly online filings moved to e-way bills and online invoice upload in real-time.

Here are a few examples of how tax administration is using data analytics in India

- GST Department has a full network diagram of each assessee. This is made possible by correlating information collated from income-tax and customs. This has made it possible to pinpoint firms who have resorted to fake bill usage, even if they have resorted to a complex web of firms to route these

transactions. By Jan 2021, the department had detected over 7000 such cases of tax evasion via fake bills.

- The income tax department has started Project Insight and a dedicated centre called the Income Tax Transaction Analysis Centre (INTRAC). This centre uses data analytics in improving tax administration. Among other things, this project will use social media data of taxpayers and analyse expenditure patterns. This Rs. 1000 crore project is aimed at deterring non-compliance and improving tax collections.

Data Analytics - Advantages

The volumes, complexities and the variety of the data are increasing exponentially at all levels of the organisation. Analytics is the solution required to get the information of this ever-growing data, and various benefits can be derived on revenue, expenses, market share, and reputation.

- **Better decision-making.** The biggest benefit of using data analytics is the assistance it can provide in better decision-making capabilities, thus better enabling key strategic initiatives. Analytics plays an important role in driving business strategy.
- **Marketing and customers.** Data analytics assist in identifying and creating new product and service revenue streams. Marketing and sales groups invest in analytics helping them to tap new geographies, develop new business models, and generate higher revenues.

Data Analytics - Challenges

While there are huge benefits, the challenges in the implementation of analytical solutions are as follows:

- **Overall business structure.** Analytics is managed by a variety of functions within a company, and a wide range of functions benefit from analytics. More structure around coordination and alignment is needed to realize the impact and benefits of a company's data throughout the organization. Sometimes a lot of projects are done in silos by each function.
- **Data management is a key barrier.** A specific level of data management is required to make analytical projects successful. The GIGO principle applies everywhere. To get the best out of the project, proper data mining and data warehousing solution is required.

Data Analytics in Audit



In this chapter, we discuss how data analytics can enhance the quality of the audit, save time and enable the auditor to gain new insights. Data analytics is not mandatory as per current auditing standards.

Changing audit scenario- Requirement for Fraud reporting

Section 143(12) of the Companies Act, 2013 requires the auditor to report frauds/suspected frauds discovered during the audit to the Central Government (in case the amount involved is greater than Rs 1 crore) and to the Board/audit committee in case of frauds of lesser amounts. The auditor has to actively plan the audit procedures to comply with these reporting requirements. Thus, the expectations from auditors have increased significantly. Most large frauds involve senior management and the statutory auditor is not in a position to easily find the fraudulent collusion between senior management and third parties. Data analytics can help the auditor to meet such growing expectations from the audit.

The increasing role of ERP systems

Modern integrated computer systems are implemented across the organization. Financial transactions are generated not just by the accounting team, but automatically by each operating unit in the company during the course of their day-to-day activities. Such large enterprise resource planning (ERP) systems are expected to have built-in checks and controls to ensure the accuracy of the data. For example, it may not be possible to create an invoice from the ERP system unless there is adequate inventory in the system for the respective item. ERP systems are also expected to allow granular segregation of duties and have strong access controls and authorisation. This can enhance the quality of the data generated by the ERP system and make it more reliable. If the General Information Technology Controls (GITC) have passed the tests for design effectiveness and operating effectiveness for the period under audit, and if the automated controls, automated accounting procedures have been properly configured and controlled, the auditor can place a high degree of reliance on the output of such ERP systems.

But at the same time, ERP systems tend to increase the complexity in the transactions flow. The auditor has to understand the ERP transactions and controls to identify risks of material misstatements. It may not be possible to have a paper trail of approvals and authorisations. Controls compromised, especially cases of Management Override of Controls, cannot be identified easily by sample-based audit.

Data analytics enables the auditor to analyse the complete data set- 100% of the transactions in a given period. After some initial training, even non-technical persons can use the data analytics software to visualize the results graphically and identify anomalies for further investigation.

A word of caution, however, raised by some professional bodies globally, is that Big Data could result in Auditors becoming less independent and objective (Institute of Chartered Accountants of England and Wales)

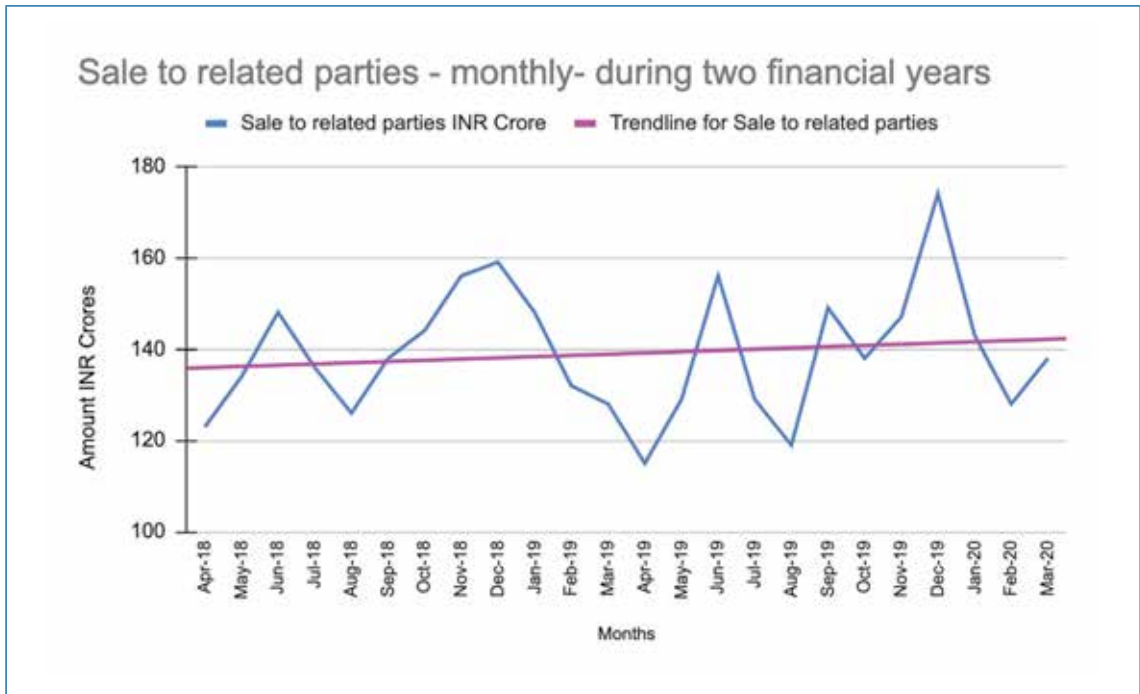
Evolution of CAATs into advanced Data Analytics

Computer Assisted Audit Techniques (CAATs) are used by auditors to query the data and find patterns or anomalies in it. Specialized software can be used to extract and sort data, summarise and stratify it. Data Analytics is, of course, "computer-assisted". The massive increase in available computing power and better software to handle Big Data has led to the evolution of CAATs into Data Analytics. Data Analytics goes much beyond CAATs. Data analytics can harness the organisation's semi-structured and unstructured data as well, which is often much bigger than the structured data of financial accounting systems. Data analytics enables the auditor to make more effective use of visualization techniques for meeting the audit objectives.

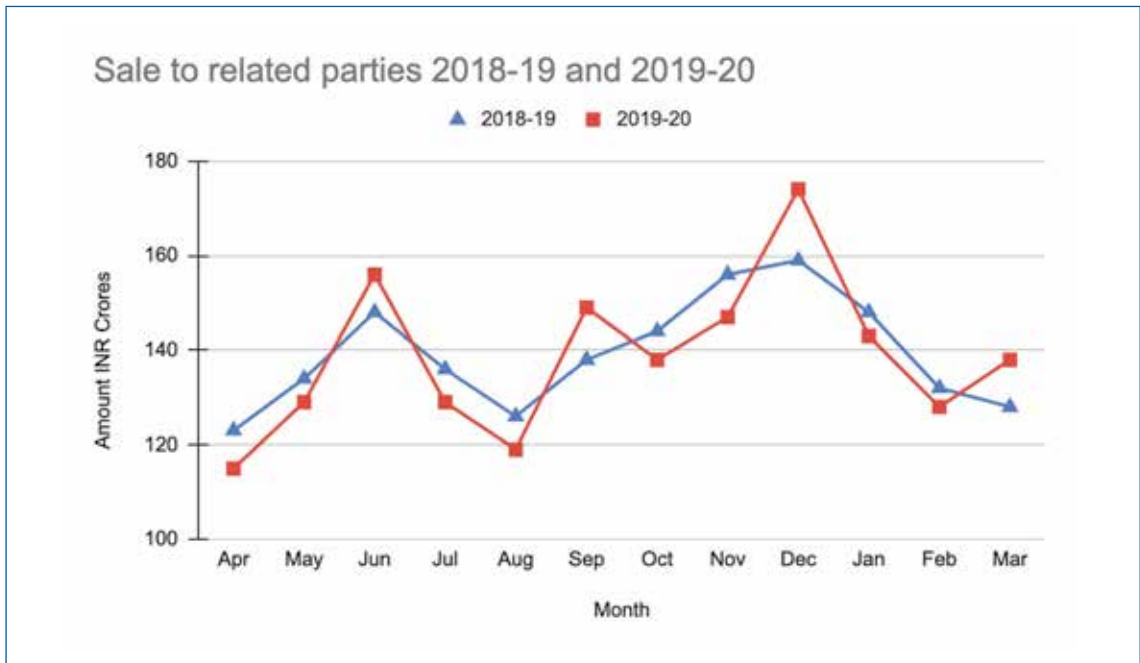
A few major types of analyses that can be performed using data analytics software-

- As the number of transactions and size of data increases, it becomes increasingly difficult to perform traditional procedures such as Ledger Scrutiny. Appropriate criteria can be defined to identify non-standard entries, considering the concepts of Ledger Scrutiny for relevant accounts, based on risk assessment.
 - e.g. -
 - Accounts used infrequently
 - An unusual combination of accounts
 - Debit Cash, Credit Revenue
 - Entered and approved by the same person
 - Persons passing entries, outside their function
 - Unusual time, holidays etc.
 - Amounts having round numbers or same ending numbers
 - Small value but many entries
 - Accounts having significant estimates or period-end adjustments
 - Accounts identified with the risk of material misstatement
 - Entries passed after the closure of months
- **Duplicate analysis and gap analysis:** This can be performed on data such as invoice numbers or bank transactions. The software can easily detect duplicate invoices or missing dates in the bank statement. It can also be used to find duplicate payments to vendors.
- **Trend analysis:** This involves a comparison of current period values with the corresponding values of one or more previous periods. Using visual representation, trend analysis can bring out significant patterns in the data.

For example, this chart shows the trend analysis of sales to related parties by a company in two financial years. The auditor can do trend analysis and plot the trend line. This can highlight the outliers from the trend line which can be taken up for further scrutiny.



However, when the same data is presented visually in a different way, it can bring out some interesting patterns-



In this second chart, it appears that in quarter-end months (June, September, December and March), the sale to related parties in the current year is much higher than the corresponding values of the previous year. In other months of the year, the sales to related parties are lower than corresponding values of the previous year. If this company is required to disclose quarterly financial results, then this anomaly will have to be taken up for further investigation.

- **Ratio analysis:** Auditor can analyse key ratios such as gross margin, working capital turnover, inventory days and compare them against similar ratios in previous periods. This can bring out anomalies that need further investigation.
- **Regression analysis.** This is an advanced statistical technique used for data analysis. A time-series regression uses data from previous periods to predict the amounts for future periods. Regression analysis helps in the identification of potential for material misstatement.

Issues/challenges with the usage of Data Analytics in audit

- Current audit standards do not mandate the usage of data analytics during the various stages of the audit, although data analytics may enable the auditor to gain new insights into client's business and speed up the audit tests.
- Identification of non-standard (usually Manual) entries in ERP environments, which is Fraud Risk, requires proper understanding and technical expertise. Only proper identification can lead to appropriate conclusions.
- Data availability and quality. A lot of efforts have to be spent to get the data in a proper shape for analysis. ETL procedures (Extract, Transform and Load) are followed to obtain "clean" data for analysis. However, ETL procedures are time-consuming and require manual efforts if the source data has a lot of anomalies.
- Data retention. The data belongs to the client, and not the audit firm. Retaining such huge data belonging to third parties is expensive and can be viewed as unnecessary to comply with audit standards.
- Privacy, confidentiality and security issues. Companies are reluctant to share their 'complete' data sets with their auditors for analysis. They fear the potential leakage of data and loss of competitive advantage. Various governments have been increasing the privacy laws related to capturing and retaining personal information. Auditors have to be careful in ensuring compliance with such regulatory requirements. This can add to the compliance cost.

- Availability of trained personnel for data analytics. It is not easy to get trained people to do insightful data analysis.
- If scripts are to be used to analyse data, the scripts must first be tested for all test cases and modified till expected results for all test cases give correct results. Documentation of testing of the scripts for all versions needs to be archived and maintained.
- Caution on Freeware, especially if scripts are to be executed on the client environment. They may have malware/trojans.

Data analytics in an audit can provide a lot of value to auditors and their clients. It is an emerging field and auditors should gear up to harness its potential.

Data Analytics in Other Finance Functions



Benchmarking

Most businesses want to measure the performance of their company's products or services against their competition, other complementing products or against another complementing industry. Hence benchmarking with 'best in class' or comparison is one of the most demanded exercises in the analysis. Tax departments and auditors have started using benchmarking extensively to compare profitability percentage with industry-standard or compare the price charged to a sister associate concern with the market price.

Benchmarking can be done from the publicly available data at various levels. E.g. at the company level; at the aggregated industry level; aggregated at the city or state level; etc. Source of data for benchmarking could be

- Data are available with the government portals like the GST, Income Tax, Ministry of company affairs.
- Industry-level studies and research carried by various ministries
- Data filed with the stock exchanges
- Private data mining companies make various types of data available on an analysable platter. For example, Capitaline, Prowess, Zauba Corp provide company-level data.
- Shipment data from the customs is available for exports and imports benchmarking.

Types of Benchmark analysis

- Absolute numbers like Turnover, quantity sold, production capacity, etc.
- Other variables like price charged market cap, etc.
- Ratios like gross margins ratio, working capital ratio, debtor days, creditor days, price to earnings ratio, etc.

Management Information

Trends (growth, profitability, turnover), comparisons (against budgets, forecasts), ratios (working capital, gross margin, debt), etc. are the type of information required by the management to make informed and timely decisions. Benchmarking (discussed above) is one of the evaluation factors that management uses regularly.

Benchmarking is the comparison with the external environment while the Management Information is both, evaluation of internal key performance indicators and comparison with the external environment.

Management Information is like dashboards in a car without which it is not a good idea to drive although one can easily drive without the same. It shows indicators of the cashflow like a fuel gauge, the growth like a speedometer and so on. Along with the dashboard of the car, this is also the rear-view mirror of the car which is as critical for one to decide whether it is safe to take a turn or change the lane or to overtake.

Guidelines-

- Creation of most relevant and representative matrices at various levels for the executives to track.
- Create an intuitive, interactive and easy to understand dashboards and reports indicating the green shoots and red flags right in time.
- Implement tech automation for faster and accurate information.

Investor and Lender information

Investors and lenders are in the continuous hunt for new opportunities to invest, track the health of their existing investments, and exit strategies from their matured existing investments.

Benchmarking and MIS are very relevant from the investors' standpoint as well. Investors also use the data from various other sources like

- Credit Information Bureaus for finding out the investees/borrowers' overall credibility, timeliness and reliability.
- Private and government players such as CIBIL, Equifax, Experian, CRIF Highmark, etc.
- Credit scores analysed by the reputed private and government rating agencies like the CRISIL, CARE, ICRA, FITCH, etc.

These credit bureaus do a lot of analytics and arrive at the readymade credit scores. These are completely driven by the data from the company, quality of data, external industry benchmarking and so on. Technology and automated data analytics have made the process matured and objective oriented over a period. The decisions about lending or investing also call for a lot of other data apart from the financial data. There are a lot of data aggregators who keep analysing and make data analytics available to investors. The data analysed are:

- The character of key managerial personnel
- Past and ongoing litigations (criminal and civil)
- Checking of information worldwide and not just restricted to India.

Tax Function

The role of the tax function is changing rapidly. Typically, it was expected of the tax department to look at historical and present data to ensure compliance and tax planning.

Now with data analytics tools, the tax department is expected to dig deep into historical data to find insights and come up with predictive analysis. Some examples:

- Detecting potential errors and non-compliance risks
- Modelling tax implications of certain events like the sale of a business unit
- Predicting tax implications on changes in external or internal conditions
- Interpreting tax laws
- A better understanding of the tax environment

Financial Statement Analytics

Different people do financial analysis for different purposes, but overall three objects of financial statement analysis: financial position, operating results and cash flow. Three major financial statements analysed are the Balance Sheet, the Profit & Loss Account and the Cash Flow Statement. Based on this, the general framework of financial statement analysis consists of the solvency analysis, profitability analysis, operational capability analysis and capital structure.

For example:

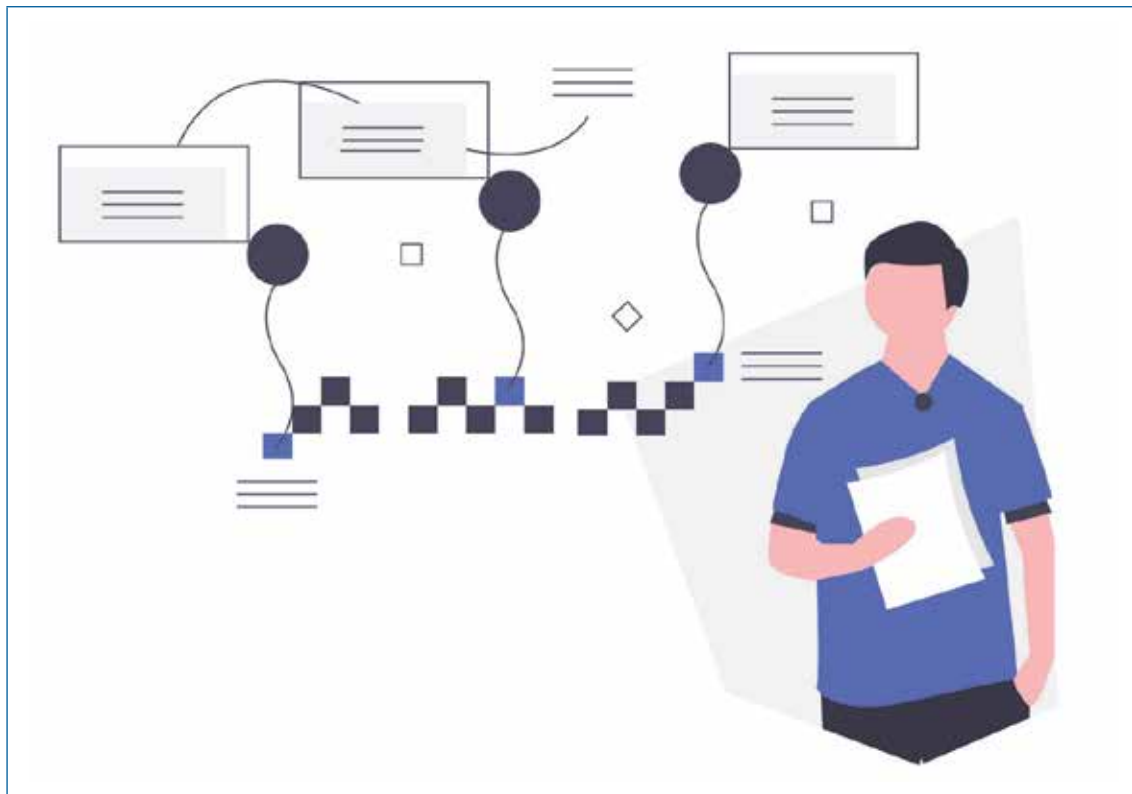
- The internal department of the company focuses on analysing compliance with financial policies and standards, capital appreciation, capital maintenance, allocation of resources effectively, etc.
- Investors focus on analysing the profitability, use of funds, understanding investment returns, investment risks, etc.
- Creditors focus on analysing solvency of the company, degree of financial security, etc.

Comprehensive financial statement analysis consists of the following:

Solvency analysis	Current Ratio
	Quick ratio
	Networking capital turnover
	Asset-liability ratio
	Cash ratio
	Interest coverage ratio
	Net cash flow to current liability ratio

	Net cash flow to total debts ratio
	Property ratio
Profitability analysis	Sales margin
	Sales expense ratio
	Return on net worth capital
	ROA (return on assets)
	ROE (return on equity)
	Net profit margin ratio
	Cost of goods sold ratio
	EBIDTA
Operational capability analysis	Accounts receivable turnover days
	Accounts receivable turnover rate
	Accounts payable turnover days
	Accounts payable turnover rate
	Inventory turnover days
	Inventory turnover ratio
	Working capital turnover ratio
	Working capital turnover days
	Overall business cycle
Capital structure analysis	Total debt-equity ratio
	Long term debt to equity ratio

Tools for Data Analysis



Data Smoothing/Transformation Tools

Big datasets can be analysed only if the data is made available to the software in certain cleansed format and ironed up in specific field types and structures. But most datasets have a lot of issues in terms of the formats, structures, spellings, etc.

The data cleansing and aligning is the most critical part of the data analytics project, but, unfortunately, is less thought of while strategising the project.

Hence, the role of tools called 'ETL' i.e, tools to Extract, Transform and Load the data for further analysis becomes very critical. There is another school of thought which believes in Extract, Load and Transform the data (ELT) tools.

The ETL/ELT tools help to build pipelines through which the data from multiple sources is brought together, cleansed and stored for analysis in a place called a data warehouse or data lake (in case of inclusion of unstructured or semi-structured data).

Many software for visualisation and analysis also have ETL/ELT tools inbuilt. However, there are many specialised tools which can intuitively and easily handle this cleansing process of the data. These specialised tools also handle the pipeline of the aggregation of the relevant data from different sources, cleaning it and keeping it ready for analysis in a data warehouse.

Xplenty, Stitch, ABS Glue, Skyvia etc are some of the examples of the ETL/ELT tools commonly used for straightening of the data.

Data Analysis Tools

Data analysis tools refer to the software which enables the process of data analysis.

Broadly these tools can be classified as under

- a. Spreadsheets
- b. Business Intelligence
- c. Audit Analytical Tools
- d. Programming languages
- e. Industry-Specific Tools

For the purpose of this book, we will limit ourselves to the tools that are relevant to the chartered accountants.

a. Spreadsheets

There are two things that a chartered accountant cannot do without Accounting/ERP system and spreadsheet. While spreadsheets are great tools for entering and presenting financial information, they have evolved over a while to provide sophisticated analytical tools.

There are several spreadsheets tools available but Microsoft Excel (MS Excel) remains the most common. Few Examples of data analytical formulae being used in excel are:

- VLOOKUP and HLOOKUP remains the most common function to connect information between two different datasets
- COUNTIF, COUNTIFS, SUMIF and SUMIFS for getting conditional results quickly.

These tools are useful for internal auditors, external auditors, fraud investigators, accountants and compliance professionals.

These tools, like BI tools, allow multiple data to be collated. The key functionalities covered are

- Filter and extract
- Merge, match and reconcile
- Track duplicates
- Extract samples

Some of such tools are :

Tool	Details
IDEA	IDEA is a user-friendly audit analytics tool, commonly used by chartered accountants.
ACL	ACL is a comprehensive solution to control audit, risk and compliance
TeamMate Analytics	TeamMate Analytics is an easy-to-use tool which operates within MS Excel.

d. Programming Languages

Data scientists use several programming languages in their work. Programming allows the creation of specific analytical solutions which may not be available in other ready to use software. Two most popular languages for this purpose are: Python and R

- **Python** is a general-purpose programming language and since its inception in 1991, it has become very popular. What has made it suitable for data analytics is the availability of several libraries which make data manipulation powerful. Some of these are - NumPy: for scientific computation, Pandas: for structured data with rows and columns

Python can achieve everything that a data analytics needs

- Reading and writing a vast variety of file formats
- Cleaning, normalising and extracting data
- Apply mathematical operations to such data

- o Connecting data to statistical models
- o Summarising data via visualisation
- **R Programming** is an open-source programming language primarily meant for data analysis. According to the official web site of R Project, "R is a language and environment for statistical computing and graphics."

All libraries in R make data analysis easy. Data cleansing and normalising is a very lengthy and messy process, which is made easy by several packages available in R. Example: dplyr, data.table.tidyr, ggplot2

R can be easily learnt by someone without prior coding experience. After a while, one can even write a new statistical model

e. Industry-Specific Tools

There are also several industry-specific tools available or built into industry-specific ERPs, such as SAP, Oracle, Microsoft Dynamics, Sage, Tally

A summary of various tools is as follows:

Basic Tools	Microsoft Excel, Google Spreadsheet, Microsoft Access,
Specialised Audit Tools	IDEA, ACL, TeamMate
Report Writers	Crystal Reports, Business Objects, Cognos
Server Based	My SQL, Microsoft SQL
Business Intelligence	Microsoft Power BI, Tableau, QlikView, Zoho Reports
Programming Languages	'R' Programming, Python, SAS
Integrated Analytics Tools within ERP/ Accounting Applications	SAP, Oracle, Microsoft Dynamics, Sage, Tally

Visualisation tools

One of the ways to present the data is by using visualisation tools. The visualisation makes the data communicate the required message to its reader. It makes a lot of sense to see the visual trends, charts, graphs than to go through huge chunks of data tables.

The sales dashboard, CEO's dashboard, etc are examples of the visualisation of relevant data.

Following are the common features of the data visualisation tools:

- **Multiple databases**

Most visualisation tools give options for picking the data from multiple sources like excel, csv, sql server, my sql, etc. This creates a lot of ease for the users to pick up the relevant data without much of transformation.

- **Relationship building**

The visualisation tools come with inbuilt capabilities of establishing relationships between various data sets. This gives the user ease of making the datasets interactive and generates a variety of reports from the interconnected datasets.

- **Colourful**

Visualisation tools allow easy depiction of the information to the reader by using various colours. It grabs the attention of the reader on the appropriate information.

- **Interactive**

The charts have now become dynamic in visualisation tools and they interact as and when the user hovers over them. They also allow users to drill information to different dimensions dynamically. The users can drag and drop the items in their dashboard dynamically.

- **Graphical user interface**

The visualization tools come with the options to view the data in different forms like a geographical map, an area chart, a histogram, etc. This helps users to get more insights from the data and data conveys important messages.

Following are some of the well-accepted list of visualization tools:

- **Power Bi**

Power Bi is developed by Microsoft. It is the tool born from the combination of the power query and power pivot, popular analytical tools in later versions of excel. Power Bi comes with very good visualization support with the ease of excel (finance professionals are most used to).

Power Bi comes with a mobile application as well. This makes it easy for publishing the data on the go to the users and interacting with them.

This can be a great starting point for the journey into data analytics.

- **Tableau**

Tableau is a very advanced data visualization tool and one of the early entrants in the market.

Tableau is better than Power Bi for the professional developers who want to create complex visualizations for publishing. Tableau is used as a reporting tool by many corporates and it can integrate easily with the complex databases. Tableau comparatively has a complex user interface and needs a bit of training to start operating.

- **Zoho Reports**

This is a lightweight and completely cloud-based analytics solution. This can be used in a "pay as you go" model. It is very simple and intuitive. This is made with the intent to adopt analytics by non-technical persons.

Data Analytics: Current & Future Technology Trends



Our smartphones carry all the data needed (and much more) for us to be productive all day long, from calendars to calculators, home energy notifications to transportation options, and food delivery services to weather forecasts. We take these conveniences for granted. But when some of them do not work, we feel the absence most.

With the increase in social media in terms of posts, images, videos, audios, location tagging, etc. and with the increase in proactiveness by the business for improving customer experience, customer satisfaction, getting customer feedback; the scale of big data is also huge advancing into both personal and professional space.

Increase in transaction volumes (including skyrocket volumes in eCommerce), increase in frauds, cyber-attacks, machines generating data themselves, it would practically become impossible for humans to manually analyse the data and reach to conclusions, take timely decisions.

In such a situation, how does an enterprise drive business effectively? The answer lies in harnessing this data challenge via new-age technologies. The world's biggest technological breakthroughs are happening in the areas of Big Data, Business Intelligence, Cloud, Internet of Things, and Artificial Intelligence.

Algorithms

An algorithm is a set of processes and calculations that can create a model from the data. To create a model, the algorithm first analyzes the data provided, looking for specific types of patterns or trends. The algorithm uses the results of this analysis over many iterations to find the optimal parameters for creating the mining model. These parameters are then applied across the entire data set to extract actionable patterns and detailed statistics.

All the new-age technologies like Machine learning, Data mining, Statistical models, Predictive analytics, Internet of Things, etc. use algorithms in some form. There are various types of algorithms and depending upon the specific requirements either one or combination of many algorithms are used. Some of the algorithm types are -

- **Classification algorithms** predict one or more discrete variables, based on the other attributes in the dataset.
- **Regression algorithms** predict one or more continuous numeric variables, such as profit or loss, based on other attributes in the dataset.
- **Segmentation algorithms** divide data into groups, or clusters, of items that have similar properties.
- **Association algorithms** find correlations between different attributes in a dataset. The most common application of this kind of algorithm is for creating association rules, which can be used in a market analysis.
- **Sequence analysis algorithms** summarize frequent sequences or episodes in data, such as a series of clicks in a web site, or a series of log events preceding machine maintenance.

Some of the functions in which algorithms are used are as follows:

Related to IT Infrastructure	<ul style="list-style-type: none">● Calculate the probability that a server will fail within the next 6 months.● Analyze the factors leading to server failure.● Identify servers that have similar usage characteristics.
------------------------------	--

Healthcare Industry	<ul style="list-style-type: none"> ● Categorize patient outcomes and explore related factors. ● Capture and analyze sequences of activities during outpatient visits, to formulate best practices around common activities. ● Create patient risk profiles groups based on attributes such as demographics and behaviours.
Sales and Marketing	<ul style="list-style-type: none"> ● Use market analysis and target audience to determine product placement. ● Forecast next year's sales. ● Suggest additional products to a customer for the purchase. ● Analyze survey data from visitors to an event, to find which activities or booths were correlated, to plan future activities.
Risk Analytics	<ul style="list-style-type: none"> ● Flag the customers in a prospective buyers' list as good or poor prospects. ● Generate a risk score given demographics.
Website Analytics	<ul style="list-style-type: none"> ● Predict visitors on the website given past historical and seasonal trends. ● Perform clickstream analysis of a company's Web site. ● Analyze users by browsing and buying patterns.

Artificial Intelligence (AI) and Machine Learning (ML)

Using ML algorithms, AI solutions can continuously improve based on the information they collect. AI and ML can analyse massive amounts of data, finding patterns that a human would never see, and distribute data-driven insights throughout the organization.

Examples of data analytics using AI and ML are as follows:

- In our daily life, ML and AI play an intrinsic role. Email spam filtering, search engine results, music and media streaming services, online ads network, social media feeds, maps navigation while travel, etc. are examples of ML & AI.

- **Minimise fraud** – ML allows for creating auto algorithms that process large data and help find hidden correlations between user behaviour and the probability of fraudulent actions. The financial services companies use this technology to track the social media feeds and the social behaviour of the customers and the solution predicts fraudulent intentions.

Such analytics can help in reviewing expense and vendor trends. For example, a specific airline and a hotel are always booked by a particular employee/s can get a red flag to the finance head and necessary training or actions can be planned to minimise the fraudulent intentions if any.

- **Creditworthiness assessment** – ML can help lenders determine the creditworthiness of potential customers. By analysing past spending behaviour and patterns, a system could identify how much credit should be extended to a given customer. This helps in lowering the risk to the financial institution.
- **Predicting bad debts** – AI helps finance professionals to predict the status of the collection well in advance. The multivariate analysis of customers data such as industry type, credit rating, product purchase, salesperson, order approver, history, etc... will provide the forecast of the probability that the customer will pay on time, or will be delayed payment or will not pay at all.

This was used effectively by the companies during the pandemic outbreak this year for planning their cash flow.

- **Process Automation** – ML, AI clubbed with Robotic process automation helps to replace and automate repetitive tasks through intelligent process automation. Some examples of process automation in finance functions are
 - **Bank Reconciliation** – based on the data from the bank, the system can automatically reconcile the data in the accounting solution. If no one-to-one entry is found, then it can prompt the possible entries against which the entry can be reconciled.
 - **Auto entries** – the technology has the capability of getting the info from the bank accounts and passing the accounting entries automatically related to bank charges, the collections received, etc...
 - **Invoice Approval** – when the invoice is received from the vendor, based on the rules and the history of the PO against which the invoice is received, the approvals can be made automated. This saves a lot of time for the senior team members.

Internet of Things (IoT)

The Internet of things describes the network of physical objects- “things” - that are embedded with sensors, software, and other technologies to connect and exchange data with other devices and systems over the Internet. It enables devices to interact, collaborate and learn from each other’s experiences just like humans do.

According to Statista (<https://www.statista.com/>), the number of IoT connected devices is expected to reach 31 billion by this year-end and to grow to 75 billion by 2025. These connected devices will produce huge amounts of data, which can be analysed and a lot of insights can be drawn. With connected devices, an organization will be able to collect more data than ever before, which can then be mined for trends offering insight into aspects of the business-like customer purchasing habits and inventory.

Examples of data analytics using IoT are as follows:

Daily Life Usages

- **Traffic Camera Sensors** - they sense the traffic conditions and report to the central server, who can then control the signals based on real-time analytics. The camera senses the speed of the moving vehicle and an e-challan is sent to the vehicle owner in case the speed limit is crossed.
- **IoT Application wearables** - this is most commonly seen and an early adopter of the technology. Fit-bits and smartwatches are the best examples of the IoT wearables. The data is transmitted to smartphones and the phones do various kinds of analytics like heart rate variability, active hours, running and walking pace, etc. Healthcare industry is also adopting IoT wearables. For example, glucose monitoring devices.
- **Connected home appliances** - there are smart appliances available, who would analyse the surrounding and adjust. Air Conditioners adjust the temperatures accordingly to the surrounding temperature. Televisions adjust the brightness based on the light in the surroundings.

Industry-specific

- **Smart Farming** - A greenhouse farming technique enhances the yield of crops by controlling environmental parameters. A greenhouse is equipped with IoT enabled devices for monitoring weather, soil composition, soil moisture levels, crop health and growth. The data then can be analyzed to determine the best time to harvest plants and create fertilizer profiles and schedules. It, then, processes the data and applies a control action.

- **Smart Cities** - This is a concept prevailing getting popular. A smart city is a city that uses technology to provide services and solve city problems.

One of the examples of the Smart City is Palo Alto, San Francisco - the first city of its kind. They took a whole new approach to traffic management. They analysed and found, most cars on the streets go around and round the same block, in search of parking spots. That's the main reason for traffic congestion in the city. They installed sensors at all the parking spots around the city. These sensors pass the occupancy status of each spot to the cloud. Any number of applications can consume that data. It can guide the drivers through the shortest route to an open spot.

- **Industrial Security & Safety** - Sensors and cameras can be used to monitor the perimeter of restricted areas and detect trespassers in non-authorized areas. Upon real-time analysis, alerts can be sent to smart devices.

Small leaks of hazardous chemicals or pressure build-ups can be identified and fixed before they become serious problems. Identifying and fixing leaks of fluids reduces corrosion and minimizes maintenance costs.

IoT-enabled detection systems are also used to monitor chemical factories, nuclear facilities, and mining operations.

Finance function

The data is analysed and it saves a lot of costs, breakages, enhancements in the supply chain, maintenance improvements, etc. Few examples are:

- Real-time production monitoring data
- Automated inventory monitoring and tracking data
- Asset monitoring data (equipment, products, etc.)
- Predictive maintenance data
- Fleet monitoring data (monitoring vehicles and other mobile assets)

Big Data Analytics with data mining

Big data after being stored in a data lake, it can be used for advanced analytics for the following:

- **Data mining** - sifts through data in search of patterns and relationships
- **Predictive analytics** - builds models to forecast customer and people behaviour and other future analytics

- **Machine learning** - taps algorithms to analyse large data sets
- **Deep learning** - more advanced offshoot of machine learning
- **Text mining** - sifts through unstructured text data identifying concepts, patterns, topics, keywords and other attributes

Examples of big data analytics

Big data comes from myriad different sources, such as business transaction systems, customer records, internet clickstream logs, mobile applications, social networks, machine-generated data, real-time data sensors and many others.

Customer insights - using customer records, various analytics that can be done could be -

- **Comparative analysis** - user behaviour metrics, real-time customer engagement comparing the company's products and services, brand authority compared with the competition, etc.
- **Social media listening** - analysing customer feedback on social media about the company's products and services going beyond general surveys and polls, identifying the target audience, marketing campaign planning, etc.
- **Marketing analysis** - analysing the success of the promotion of new products and services, more information to be more innovative and informed, etc.
- **Customer satisfaction & sentiment analysis** - analysis of the customers' feeling of the company brand, innovative ideas to preserve brand loyalty, analysis of customer service levels, improvement of customer service, etc.

Real-time Stock market insights - ML is changing trade and investments. Instead of simply analyzing stock prices, big data can take into account political and social trends that may affect the stock market. ML monitors trends in real-time, allowing analysts to compile and evaluate the appropriate data and make smart decisions.

Blockchain

A blockchain is a distributed database that hosts a continuously growing number of records. The database stores record in blocks rather than collating them in a single file. Each block is then 'chained' to the next block in linear, chronological order using a cryptographic signature; as a result, records cannot be revised, and any attempted changes are visible to all participants.

Although blockchain technology is the foundation for a cryptocurrency (such as bitcoin), there are a variety of financial and accounting applications beyond the realm

of cryptocurrency. With its unalterable blocks and secured record-keeping capabilities, blockchain can offer a valuable and trustworthy forensic trail if quality issues arise after the delivery of an item.

- Healthcare companies can analyse products to identify signs of tampering or careless handling,
- Grocery stores can analyse and identify sources of contamination quickly in event of a recall,
- Forensic data is readily available and transparent, eliminating the need to manually sift through paper-based records. The analysis is real-time.
- Payment processing and cross-border money transfers not only eliminates the need for third-party intermediaries but also reduces the costs of processing money transfer, increases efficiency and improve security. [Arab Jordan Investment bank is using this].
- Payment dispute resolution
- Improved regulatory compliance

Future of Data Analytics using new-age technologies

Data Analytics is getting used for a long time now, but, with the new-age technologies, the analytics will change dramatically. Given the business dynamics, volatility and the volume, there is a lot the future will unwrap in terms of usage of technology, upgradation of personnel skills and the advent of new analytics. Some of the future analytics could be as follows:

Product and Sales Recommendations

Although there are various applications of automated financial product recommendation existing even today, some of them involve rule-based systems (instead of machine learning) where data still goes through manual resources to be able to recommend trades or investments to customers.

- The future will see ML & AI technologies being used by insurance recommendation sites to suggest customers a particular home or vehicle insurance policy.
- The future will see the car suggesting the particular route to the owner of the car among the various routes to save on insurance premium and claims.
- The future might also see the digital assistants suggesting changes in the portfolio

- The future might see an offering of ML-based objective & reliable advisory via personalised apps and personal digital assistants.

Enhanced Security

Data security in banking, finance and other individual intensive industries is a critically important area. With all the information available online, companies find it increasingly challenging to keep all the usernames, passwords and security questions safe.

The future will see a dramatic shift in security, where passwords, usernames and security questions may no longer be the norm of user security. ML applications will transform future security within the industry to adopt voice recognition, facial recognition or similar biometric data.

Customer Sentiment Analysis

It is a well-known fact that human factors drive the stock market, and customer sentiments are the most important factor.

- ML can be of great help to finance companies when it comes to analysing current market trends, predicting the change and social media usage of every customer. Businesses can learn from the financial activity of the customers continuously and position their products accordingly.
- Customer sentiment analysis can also complement current information on different types of commercial and economic developments.

Automated Data Analytics

Amount of data continues to grow and not every data that comes in is useful for business. So, businesses have to extract the useful ones and they spend a lot of time, money and resources doing this.

- Data analysis automation will scan the useful data to be analysed. It will lead to improved efficiency and effectiveness of data, enabling companies to get insights faster.
- Data analytics solutions will be more advanced and user-friendly. This will help in generating the reports automatically without much skill. It will also be much easier for the data scientist to carry out complex analytical operations.
- Fast data allows for processing in real-time streams. Because of this stream processing, data can be analyzed promptly, within as little as just one millisecond. This would bring more value to organizations that will be able to make business decisions and take actions immediately when data arrives.

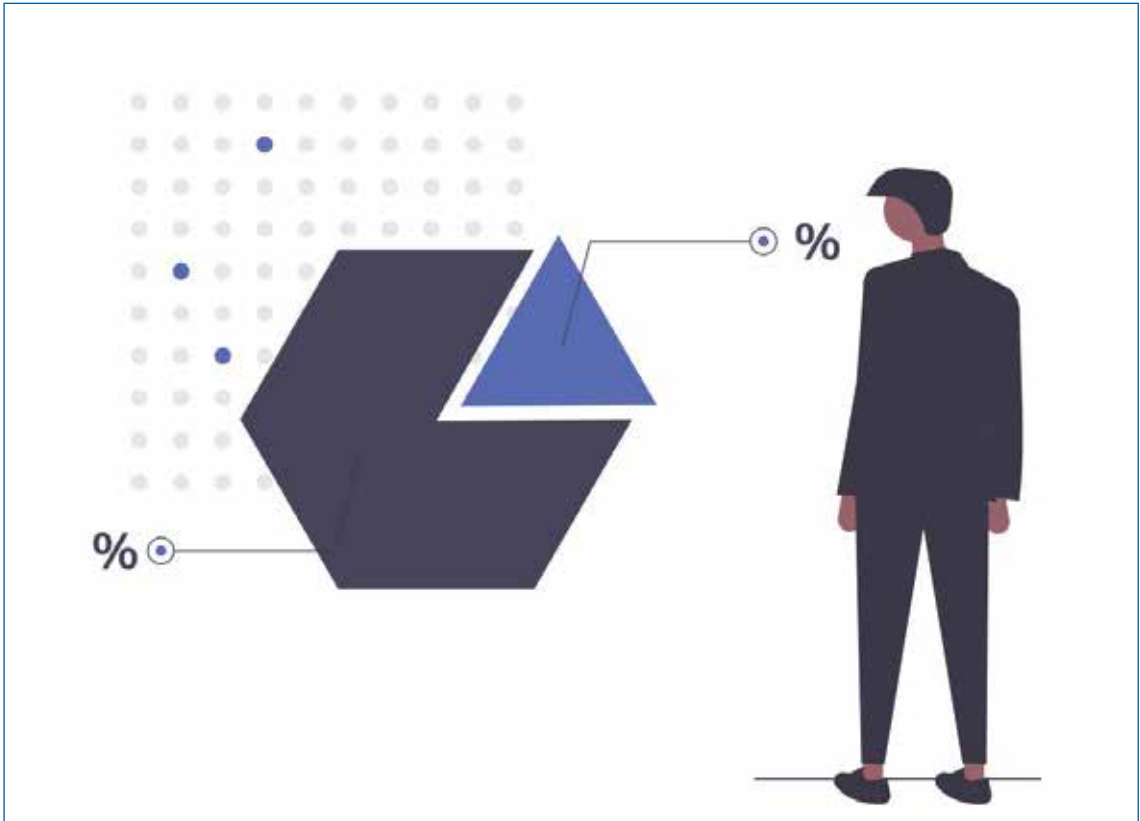
Internet of Behaviour (IoB)

After IoT, then the next upcoming technology trend is IoB. IoB is about using data to change behaviours. With an increase in technologies that gather the 'digital dust' of daily life; the data that spans the digital world and physical world; the information can be used to change behaviours.

The IoB can gather, combine and process data from many sources including commercial customer data, social media, public domain deployments of facial recognition, and location tracking, wearable devices distributing data, etc.

- For commercial vehicles, telematics can monitor driving behaviour, from sudden braking to aggressive turns. Companies can use the data for driver performance, routing and safety.
- The same wearables that health insurance companies use to track physical activities to reduce premiums could also be used to monitor grocery purchases; too many unhealthy purchases could also increase premiums.

Commonly used Statistical Models for Practical Data Analytics



Understanding the statistical models and concepts is required for data analytics. One of the very well-known and important concepts of statistics is '**Measure of central tendency**'. Although this is a very basic model, it helps to create a lot of data models and representative trends.

The measure of central tendency helps to derive a representative from a population which describes the population very closely. So we can say that a measure of central tendency attempts to describe a dataset by a single value. This single value is the central position within the data set.

Some of these measures are described below:

Mean

Arithmetic Mean helps a lot in the data analysis. The formula for deriving arithmetic mean is to 'divide the sum of all the observations with the number of observations'.

Example:

It is important for the auditor to find out the average monthly sales to understand the size and scale of the company. This average is also called the "mean".

Month	Sales in INR	Month	Sales in INR
Jan	1004050	Jul	1200561
Feb	1100506	Aug	1109281
Mar	1234055	Sep	1009821
Apr	988569	Oct	1057211
May	1026789	Nov	1094781
Jun	1086410	Dec	1209815
Total	13121849		
Total Months	12		
Mean	1093487		

To find the mean, add the monthly total sales and divide by the total no of months in the data

Median

Median is the middle number when the data is sorted in ascending or descending order.

Example:

As explained above, for getting a better representative of the monthly sales from the total annual sales of a company, we use the "mean". If the sales are very sporadic due to the seasonality, (e.g. Sales for Diwali (Oct) and Christmas (Dec) for crackers are five times that of a normal month), then we should use median

Sales are organised in the ascending order. Sales from the exact middle of the list will be picked up as the representative of the monthly sales. If the number of observations is an even number, then the average of the middle two observations in the sorted dataset is taken as the median

Month	Sales in INR
Jul	369021
Jan	400102
Jun	486410
Apr	560914
Aug	580237
Nov	594781
May	626789
Feb	640910
Mar	678000
Sep	709821
Dec	1109281
Oct	1200561
Total	7956827

In this case, since the total observations are 12 (even number), the middle 2 observations are highlighted in the above table. Average of the two highlighted observations is 610785.

Mode

Most commonly occurring number in the observation is called mode. There can be one or more than one mode in the data set.

Example:

In a service complex, several companies are working with a varying number of employees. The table below shows the number of employees in each company

Company	No. of Employees
A	5
B	6
C	8
D	10
E	5
F	7
G	10
H	12
I	9
J	11
K	1
L	10
Total	94

The mode in the above case would be 10 as it occurs in most of the observations in the dataset. This represents that most companies in the service complex have 10 employees.

Mean, Mode or Median?

- If the values are close to each other, 'mean' works fine as a representation of the data.
- If the values are too different from one another, 'median' will be a better representation of the data.
- If the dataset has many repetitive values, 'mode' suits best to find the most representative sample.

Standard Deviation

The standard deviation is the statistic that measures the dispersion of the dataset relative to its mean. If the data points are further from the mean, there is a higher deviation from the data set. In stock market parlance - A volatile stock has a high standard deviation, while the deviation of a stable blue-chip stock is usually rather low.

Example

A standard deviation is a useful tool in investing and trading strategies as it helps measure market and security volatility and predict performance trends. An index fund is likely to have a low standard deviation versus its benchmark index, as the fund's goal is to replicate the index. On the other hand, one can expect aggressive growth funds to have a high standard deviation from relative stock indices, as their portfolio managers make aggressive bets to generate higher-than-average returns.

Correlation

Correlation helps in finding relationships between two or more things. For data analysis, it is very critical to find out the relationship between certain parameters. For e.g. stock market changes vis-a-vis a particular stock or the relation of change in the gold rates vis-a-vis stock market indices.

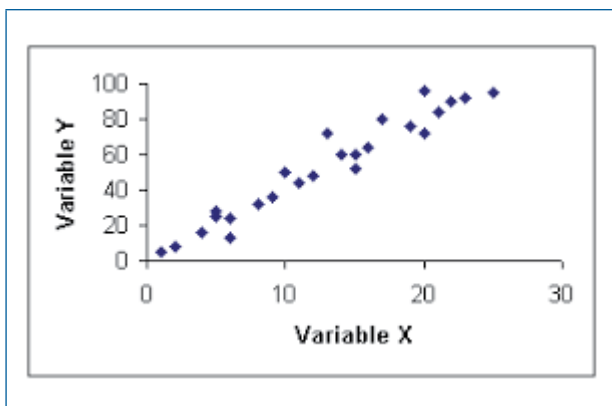
The correlation can be positive, negative or no relationship.

- Positive correlation - the variable changes in line with the other e.g. advertising cost vis-a-vis sales.
- Negative correlation - the variable changes in the opposite direction of the other e.g. reduction in turnaround time of delivery resulting in an increase in sales.
- No relationship - the variables do not have any relationship e.g. advertisement cost and delivery turnaround time

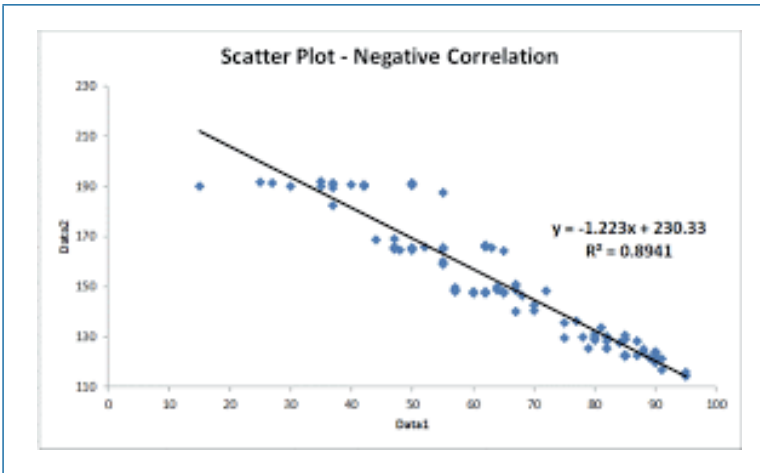
This correlation between the data sets can be established by the graphical representation.

Example:

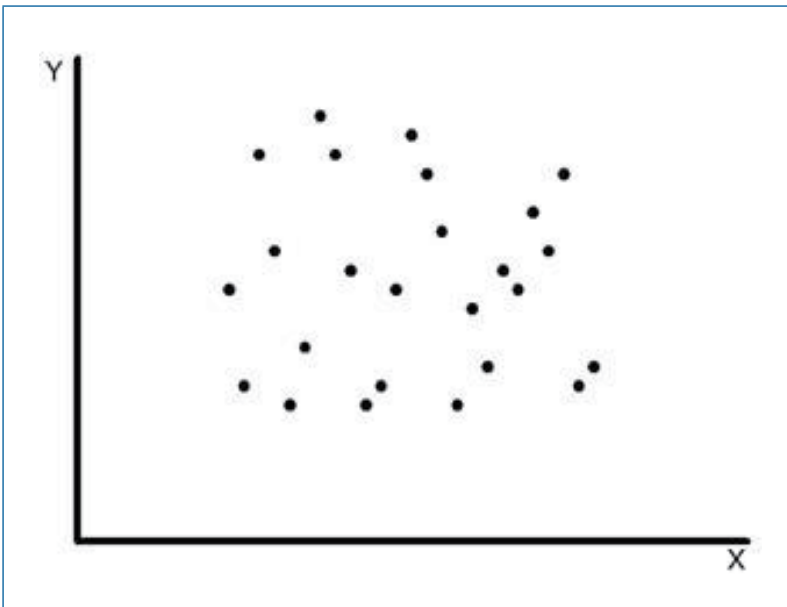
Plotting a positive correlation: Increase in the number of people in an office is linked with the increase in the electricity bill.



Plotting a negative correlation: Decrease in climatic temperature results in an increase in sales of the woollen garments.



No Correlation - The relationship between the two variables does not exist. Like the population of a country and its climatic temperature.

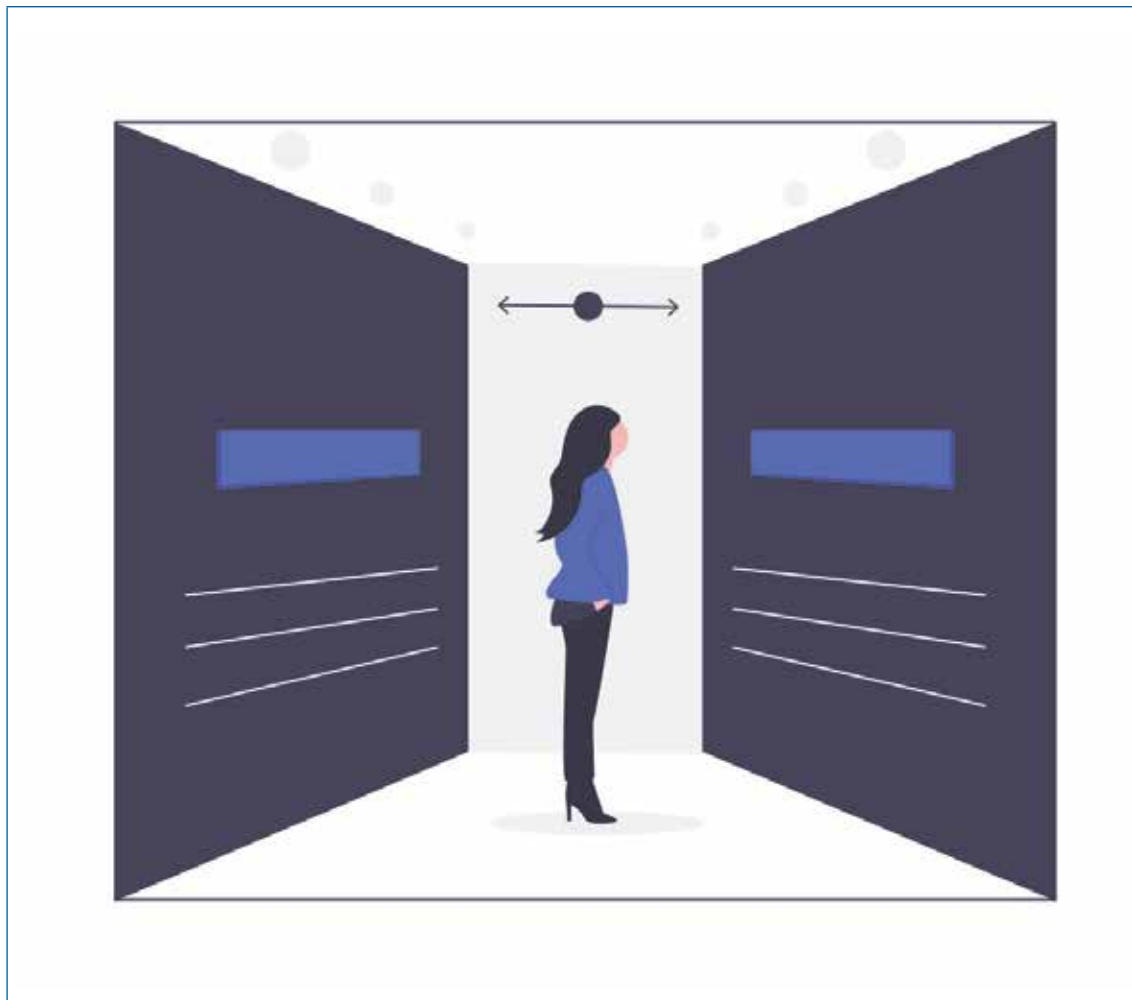


Regression

Regression is a further extension of the correlation. In a relationship, there is one variable dependent on an independent variable. Independent variable is one that can be changed by choice, e.g. no. of salespersons. The dependent variable would be e.g. sales (sales may increase with the increase in the number of salespersons). Regression is a technique used for calculating the impact of an independent variable on the dependent variable.

Regression analysis is one of the foundation stones of predictive analysis. This is one of the tools to do 'What if' analysis, 'scenario analysis' and so on useful for decision making.

Data Analytics Usage



Professional Firms

It is observed that small professional firms continue to rely heavily on Microsoft Excel and to some extent, inbuilt tools within accounting software/ERP systems for data analytics. The use of data analytics is also limited to audit functions.

Whereas large accounting/consulting firms are investing heavily in adopting data analytics, artificial intelligence and machine learning. Such technology is being used

across tax, auditing, consulting and risk management. For example, A large audit firm has an AI-enabled document review platform that can review and extract relevant information from contracts, reducing man-hours spent on this process by up to 50%.

While large firms are creating their own platforms for effective data analytics, small and medium firms have access to off-the-shelf solutions. As these tools and software become more affordable, tech-savvy chartered accountants have started using them to have an edge.

With the accounting moving in-line with the business process by implementation of the ERP systems, as the focus on Corporate and IT Governance increases, the issues around accuracy should go down. Also, the skill around checks and balances which was considered a premium one for the accountants could now be taken over by the computer. This makes analytical skills on the available data, a significant one for the professionals to acquire, so as to differentiate themselves from the competition.

There is a growth in the number of businesses using data analytics, in the same way, an auditor would do. With the help of such tools, businesses can get more insights into risk and performance. In such cases, they would have greater expectations from their auditors for advice on potential risks. Without the use of data analytics tools, it will be difficult for the auditor to meet client expectations.

Use of data analytics is very necessary to enhance audit quality, find potential risks and communicate audit findings more effectively.

Small firms can leverage data analytics to generate insights from accounting data and provide value-added services to enhance their revenue.

Finance professionals in the Industry

Finance function and finance professionals have always been the custodian and curator of financial data. Whether it is bringing the data together or distributing the information, many reporting processes have been involving a great deal of effort of information assembly, validation, and reporting. The activity has become mundane and never-ending over time.

But, with the advent of new-age technologies, regular reporting can become comparatively easy. Finance professionals can use data-analytics and report a lot more insights. Some of the finance professionals using data-analytics with the help of new-age technologies have taken a completely new role in the organisation.

Strategy

The finance professional is now seen as a strategist. The data-analytics allow finance to track Key Performance Indicators (KPIs), monitor ongoing results and performance,

and inform the business or make recommendations to the business. Finance professionals can bridge the gap between strategic and operational decision-making with analytics. It's the difference between 'managing the business' (post-mortem) and 'running the business' (proactive and future insights).

- Finance professionals can exercise more '**centralized control of operational business decision**' making. The analytics can help finance function to help businesses making decisions such as:
 - the price point should be used for this customer on this day.
 - the inventory products should be pulled forward or out of the supply chain.
- Assisting organisations in setting '**financial goals**' of the organisation. Based on the trend analysis using forecasting models, gives insights into cash flow, profits for the next 15 to 24 months. The business can make informed decisions making sure to minimise cash strain.

Risk Management

Strategic '**risk management**' of the organisation. Analytics are used to mitigate various risks not only in real-time. Predictive analytical technologies can be used to predict the risks and avoid in advance. Some of the examples are -

- **Prevent repetitive losses** - If the claim of the expense of the similar nature/ item is being done, again and again, the analytics can show a red flag. The root cause of the claim can be identified and solved.
- **Monitor performance** - Analytics will be able to hold units or departments accountable for exceeding or failing to meet goals, recognize red flags that may indicate something needs to be changed, or discover why a business strategy isn't working out as well as planned
- **Improve reporting** - With in-depth data and analytics, exceptional reporting can be generated. Data will help to identify actionable, easy for anyone to understand, and will be able to support any mitigation strategy.
- **Forecasting and decision making** - Without effective analytics, it's difficult for one to learn from the past or prepare for the future. Analytics comprehends what went wrong historically and allows one to prepare for future potential incidents, thus assisting to mitigate risks in advance.

Other Functions

Finance-supported analytics can drive value outside the finance function like

- **procurement** function assisting in spend analysis and vendor management, vendor performance analysis.
- **business units** for margin-erosion analysis, pricing analytics, customer profitability alignment.
- **sales and marketing** function helping in price point, revenue leakage, revenue drivers, demand/price elasticity, customer retention & churn analysis.
- **the supply chain** for sales, finance and requirement linked forecasting, new product introduction profitability.
- **information technology** function for decisions on technology investment planning and prioritization.

A Case Study

Leading telecom finance professional with the help of analytics helped the company (outside the finance function), targeting the company's most profitable customers to achieve:

- **High Net-worth customers specific telecom plans:** To calculate the lifetime value of individual customers and develop scenario models for discounts, marketing, and rewards to guide interactions with the customer.
- **Higher profitability to the company:** deploying analytical solutions to front-line employees and marketers to derive individual customer profitability from each transaction, in real-time. The sales team was able to change plans in real-time giving them the view of the benefits to the customer and profitability to the company over the long term.

Finally A Word of Caution

The new-age technology like Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) are the backbones of most of the aforesaid analytic technologies. There will naturally be a huge reliance on the outputs provided by these tools by auditors and management. There should always be a question that should prompt the users of the analytics **“What is the assurance that the outputs are complete and accurate?”**

AI and ML both are like the brain of a child. It learns. It also learns incorrect things. It also unlearns what it knew. It may develop biases. Testing of these technologies is a very different ball-game, not at all as straightforward as building a laundry list of Test Cases and confirming outputs. There may be unstructured questions asked, to which answers will pop-up. Should these answers be accepted at face value?

As an example, per the applicable BASEL norms for determining Capital Adequacy for a bank, Data Mining tools are run on the Data Warehouse to discover hidden trends. If the outputs are incorrect, so also will be the determination of Capital Adequacy.

Chartered Accountants are trained to be sceptical. Take the help of experts, if required (SA 620 – Using the work of Auditors' Expert).