**Learning Outcomes**

After completing this module a student will be able to:

1. Understand business analytics and develop business intelligence.

**Understand business analytics and develop business intelligence.**

In this section, we will read about:

* Introduction to business analytics and Concepts of business analytics
* Trends in business analytics
* Introduction to Big Data Analytics

**Introduction to business analytics and Concepts of business analytics**

**What is Business Analytics?**

Business Analytics is “the study of data through statistical and operations analysis, the formation of predictive models, application of optimization techniques, and the communication of these results to customers, business partners, and college executives”. Business Analytics requires quantitative methods and evidence-based data for business modeling and decision making; as such, Business Analytics requires the use of Big Data.

Business Analytics is a combination of Data Analytics, Business Intelligence and Computer Programming. It is the science of analysing data to find out patterns that will be helpful in developing strategies. Its usage can be found in almost every industry.

Business analytics is used by companies that are committed to making data-driven decisions.

**Big Data: An Overview**

SAS describes Big Data as “a term that describes the large volume of data –both structured and unstructured –that inundates a business on a day-to-day basis.” What’s important to keep in mind about Big Data is that the amount of data is not as important to an organization as the analytics that accompany it. When companies analyze Big Data, they are using Business Analytics to get the insights required for making better business decisions and strategic moves.

**Business Analytics Process**

**Business Problem Framing**

This is the first thing you do before you start your analysis. Even before you begin your analysis, you should understand the purpose of your analysis. Here you try to understand what the business is and what the business is trying to achieve. You formulate the business problem.

**Analytics Problem Framing**

Here you reformulate the business problem with respect to analytics. You develop a proposed set of factors and its relationship to output. Also, you define a metric ofsuccess of your model.

**Data**

Here you identify and select your data for analysis and its source. You work to clean the data and make it analysis ready. You also find relationships between data and report them.

**Methodology Selection and Model Building**

Once your data is worked on, you decide which method to use for your analysis. This is decided based on your data and the type of analysis you have to perform. You make multiple models and compare them based on the metrics you decided on.

**Deployment**

You validate your model to check if your model is giving accurate predictions. Once validated and reported, you deploy your model on the company's system which then will perform analysis on every new incoming data. When a model is deployed, it has to be constantly monitored for accuracy.

Diagram

Description automatically generated

Business Analytical Process

Reference -https://www.proschoolonline.com/certification-business-analytics-course/what-is-ba

**Understanding Business Analytics**

Business Analytics is the procedure through which information is dissected after studying past performances and issues, to devise a successful plan for the future. Big Data or large amounts of data is used to derive solutions.

This method of going about a business or this outlook towards building and sustaining a business is vital to the economy and industries that thrive in the economy

Graphical user interface, diagram

Description automatically generated

Understanding Business Analytical

Reference -https://www.edureka.co/blog/what-is-business-analytics

Diagram

Description automatically generated

Data Scientist

Reference -https://www.martinsights.com/?p=1049

**Components of Business Analytics**

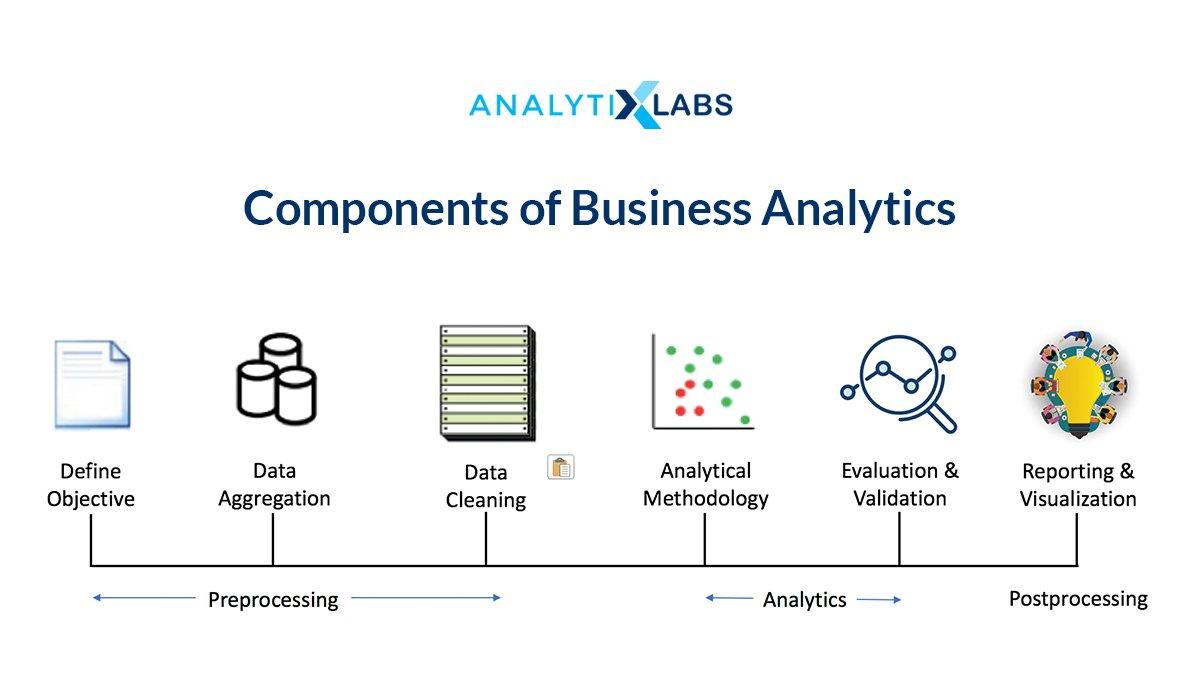
The process of business analytics starts from realizing that there is some problem or there is a scope of improvement. This is where good business acumen is required whereby understanding the inner working of the business, one can help and guide in the area of interest.

The next step is often retracing the steps and checking if the process is going in the rights direction or not. This is where business analysts must check that something that they are considering as a problem or scope of improvement actually exists or not.

Once all of this is cross-checked, then business analysts need to identify the relevant data required to do the job. This includes the identification of the data source, the type of data required, the amount of it, and the format in which it can become useful.

This is where Data Engineers also come in the scene as they are often responsible for creating the architecture to perform ETL operations that allow the Business Analysts to get hold of relevant data in a relatively short span of time.

Once the data is achieved, it is treated, cleaned, prepared, and finally is put to use. Here various methodologies are put in place that provide varied kinds of information from the data. While some are simply factual in nature, some provide predictive information.

All these aspects of the analytical process form the different components that are namely

Components of Business Analytics

Reference-https://www.analytixlabs.co.in/blog/what-is-business-analytics

* Define Objective
* Data Aggregation
* Data Cleaning
* Analytical Methodology
* Evaluation and Validation
* Reporting and Data Visualisation

**Define Objective**

This is the foremost step. Without having a clear understanding of business goals, questions we need to answer, and problems we ought to solve, none of the following steps will deliver. This also helps us to translate business objectives into analytics objectives and map data requirements.

**Data Aggregation**

The process of having a centralized location for the data, extracting and loading the relevant data by putting relevant filters, and creating subsets of data is the core aspect of Data Aggregation. This is where the data is transformed depending upon the business requirement. Also data pertaining from various sources are combined to have one large dataset. The format of the data is also sometimes changed at this step to make it compatible with the tool being used to achieve the objectives.

**Data Cleaning**

Data Cleaning is an extremely important component of business analytics because the data in its raw form sometimes is not directly usable. As the other components of Business Analytics use mathematics, statistics, and computer programming, the data must be compatible with these streams of study.

For example, for applying statistics, the data mustn’t have any extreme values (also known as outliers), while for mathematics, there should no blank cells or missing values (as matrix operations become difficult) while for programming, the concept of typecasting plays a role where the data is made sure to be in the right format (i.e. correct class or data type).

Also, the concepts of multicollinearity and curse of dimensionality come into play as the business analyst has to make sure that there are no implicit or explicit duplicate columns. The importance of getting rid of the unnecessary columns can only be understood once a good grasp of statistics is there. Other aspects include the resampling of data (under-sampling, oversampling, hybrid-sampling) removal of duplicate rows, etc.

**Analytical Methodology**

Having a detailed understanding of the different types of analytics out there dominate this component as this is where the analysts have to identify the method with which they will go to achieve their end goal. If the end goal is to understand what is the present situation of the business then that requires a different set of methods while if there is a need to identify which has happened in the past or what can happen in the future, then a different technique is required. Here, having the know-how of various procedures, methods, and algorithms is important, and knowing what to use, when one business analyst stands apart from the other. Different types of Analytics methods and tools are explained in the latter section.

**Evaluation and Validation**

Once the results come out, the next task is to understand if the result stands true given a different situation or not. This is where predictive models are used and their evaluation and validation are conducted whereas, for other methods, various simulation techniques are put to use to identify the most plausible outcome, thus providing a very reliable result. Here also, the business analyst needs to learn a range of techniques to identify the shortcoming in their method, work on it, improve it, and make their insights stable and valuable.

**Reporting and Data Visualisation**

Perhaps the most important and often overlooked component of the discipline is the aspect of communicating the results in an easy to understand way. This requires innovation and creativity and is the reason that this field is open to all and not only to mathematicians, statisticians, or computer programmers. To quickly make people understand the complex insights discovered over weeks or even months, reports or presentations are created that have simple tables, bullet points, etc. On top of all this, visualizing the data plays a major role here as it allows the people in the leadership positions to quickly view where the organization is coming from and perhaps where they are headed. Business Analytics ought to know the various ways of visualizing the data, the transformation required to be done on the data to make it possible, and finding innovative ways to string together different information ina smooth storytelling method.

**Types of Business Analytics Methods**

Big data analytics cannot be assumed as a one-size-fits-all strategy. In fact, what differentiates a finest data expert or data analyst from others, is their capability to categorize the type of analytics that can be used to benefit the business at the optimum level.

Fundamentally there are Four types of analytics.

1. **Descriptive analytics**-Describing or summarising the existing data using existing business intelligence tools to better understand what is going on or what has happened.
2. **Diagnostic analytics**-Focus on past performance to determine what happened and why. The result of the analysis is often an analytic dashboard.
3. **Predictive analytics**-Emphasizes on predicting the possible outcome using statistical models and machine learning techniques.
4. **Prescriptive analytics**-It is a type of predictive analytics that is used to recommend one or more courses of action on analyzing the data.

A picture containing diagram

Description automatically generated

Types of Business Analytics Methods

Reference-https://www.analyticsinsight.net/four-types-of-business-analytics-to-know/

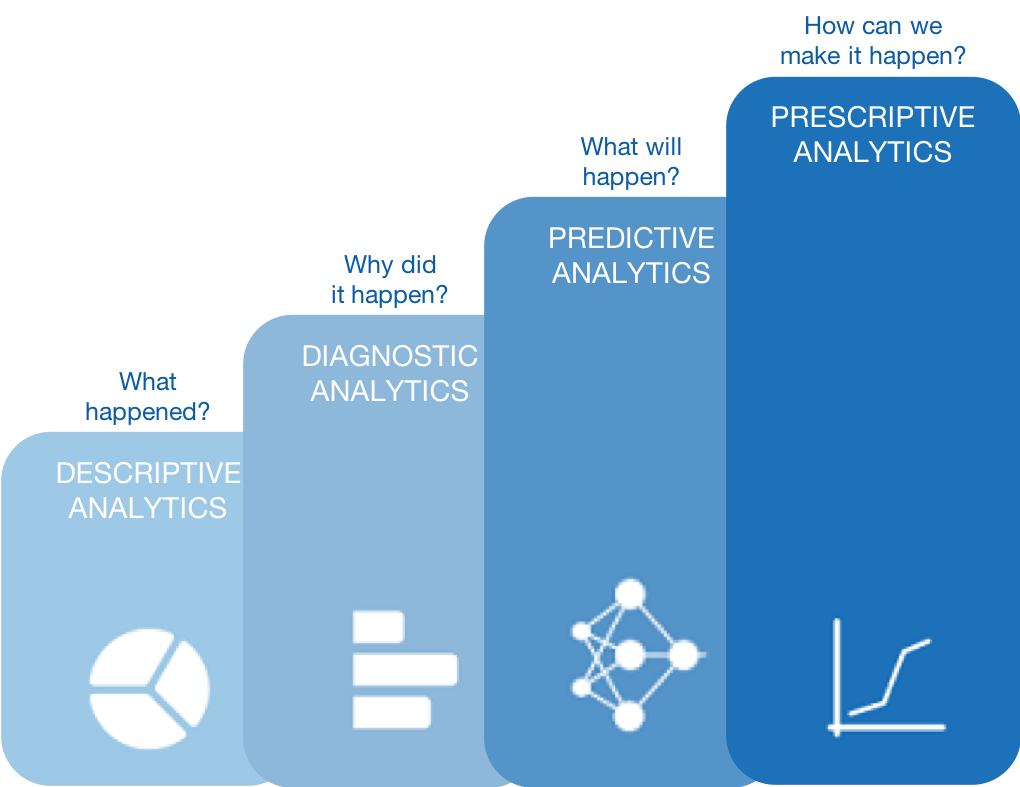
**Descriptive Analytics**

Descriptive analytics manipulates figures from multiple data sources to provide valuable information about the past. So, developers can predict important trends and signal the necessity of preventative or stimulative actions.

These describe what has already happened.

* For example, a retailer can learn the number of clients, the average bill, the most popular goods, etc.
* A medical company can evaluate the most common illnesses and susceptibility to disease.
* Governments can stimulate population growth by organising additional social support for families with several children if the overall number of pregnant women has been decreasing over several consistent years.

With the help of descriptive analysis, any company is able to group its customers by social factors, behaviour and other features, as well as monitoring peak activities according to seasonal or local factors.



Descriptive Analytics

Reference -https://magora-systems.com/types-of-data-analytics/

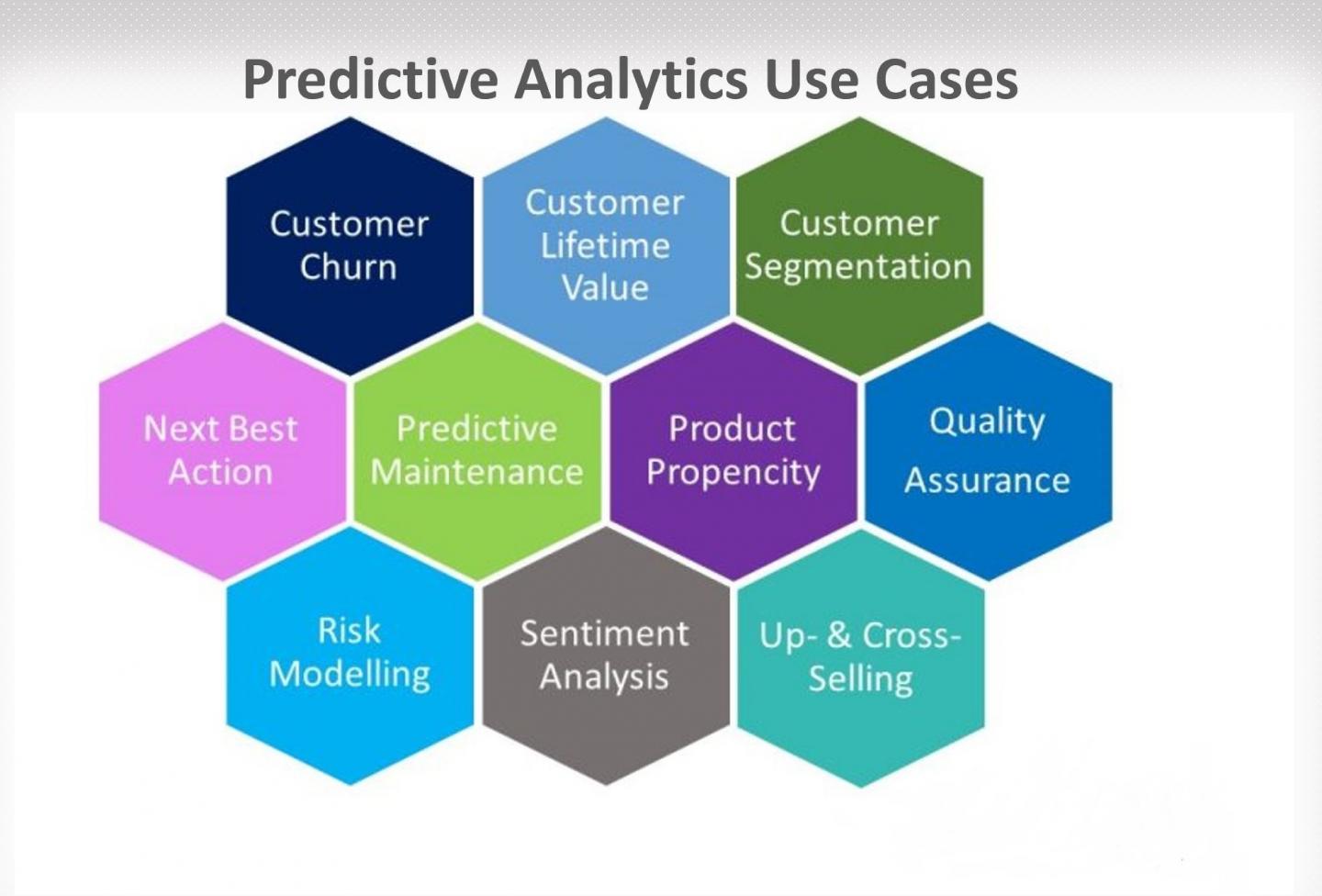
**Diagnostic analytics**

This type doesn’t simply state but aims to find the reason behind why something happened. We use it to identify the patterns and consequences of our policies and actions. It provides a deeper understanding of any given problem.

In IT, for this purpose, we use business intelligence (BI), to be more specific -it’s machine learning techniques again. These are computer-based methods of analysing and reporting valuable information. For example, using ML in healthcare, medics can diagnose a person’s susceptibility to cancer-based on medical screenings.

**Predictive analytics**

Predictive analytics forecasts the possibility of future events using statistical models and machine learning techniques.



Predictive analytics address what might happen.

Predictive Analytics

Reference -https://magora-systems.com/types-of-data-analytics/

To be able to predict trends and see into the future, this type of analytics uses the results of the previous two –i.e. it bases its results on true facts of the past.

* With the help of predictive analysis, an entrepreneur can optimise the raw material storage and the warehouse stock. Computer systems predict stock exchanges, market fluctuations and currency exchange rates. This is specifically useful in finance, production, logistics and banking.

However, it’s important to understand that all the results this type of analytics provides you with are approximate. The accuracy of data and the stability of the situation have a significant influence on the result. It requires careful processing and constant optimisation.

Combining the approaches gives the best, most relevant results.

**Prescriptive analytics**

Prescriptive analysis is based on mathematical modelling. Its mission is to show the consequences of certain actions based on possible changes to data and conditions.

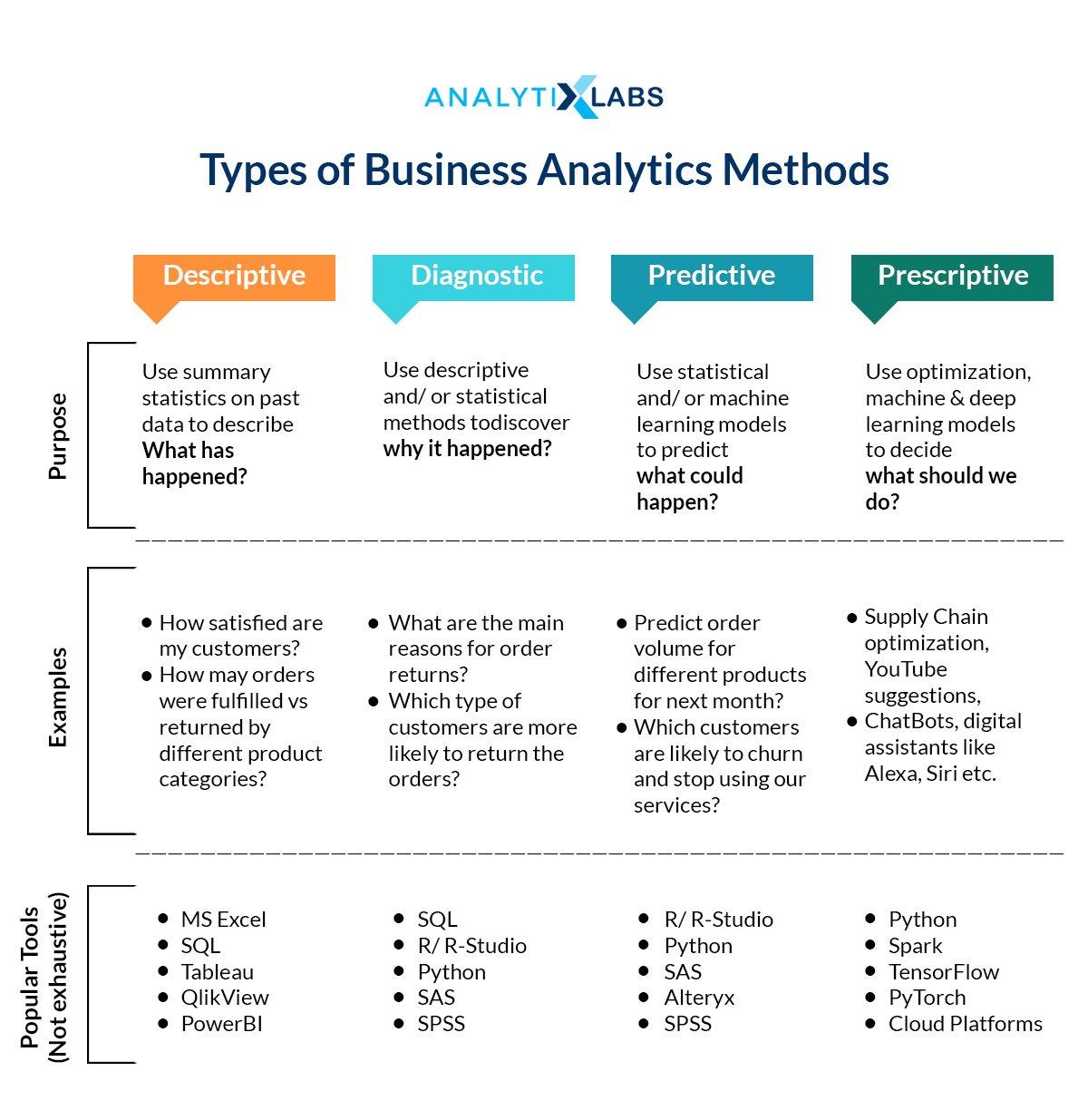


Prescriptive Analytics

Reference -https://magora-systems.com/types-of-data-analytics/

When is it time to use prescriptive analysis?

* Prescriptive analytics is a branch with a high degree of responsibility: it utilises top-notch tools and technologies, such as machine learning and data mining.
* It helps in decision-making by constructing a potential future and estimating the probability and extent of any given factor’s influence. (Remember “Back to the Future” –they were realising the prescriptive analysis in a series of events).
* It’s complicated and expensive. If you are working with vital factors, it lets you save much more than you spend. But for some cases, it’s worth using mathematical modelling and statistical analysis to get value for money solution.



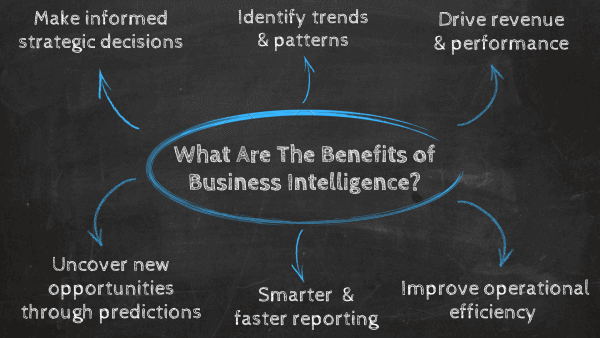
Types of Business Analytics Methods

Reference-https://www.analytixlabs.co.in/blog/what-is-business-analytics/

**Uses and Benefits of Business Analytics**

Business Analytics is adopted by companies to facilitate data-driven decisions. The insight that is arrived at with the help of business analytics enables companies to optimize their various processes to provide even better results and thus achieve a competitive advantage over others.

* To carry out data mining and exploring new data to find new patterns and relationships.
* To carry out statistical and quantitative analysis to provide explanations for certain occurrences.
* Test previous decisions are taken with the help of A/B testing and multivariate testing.
* Deploy predictive modeling to predict future outcomes.



Uses and benefits of Business Analytics

Reference-https://www.datapine.com/blog/benefits-of-business-intelligence-and-business-analytics

**Business Analytics Tools**

What is business analytics can also be answered by understanding the tools used in this field. As the discipline of business analytics is of a unique kind as it covers various types of analytics and this is the reason that there is a range of business analytics tools to accomplish the objectives.

As business analysts grapple with a number of problems that range from sourcing data to create predictive models, tools that specialize in different fields are put to use and a good business analyst must know how to use all of them in harmony. Some of the common Business Analytics related tools include:

* SQL
* Tableau/ QlikView/ Power BI
* Birt
* Python
* R
* MS Excel
* Sisense
* Clear Analytics
* Pentaho BI
* MicroStrategy

**SQL**

It is among the most important tools as SQL queries allow the user to easily filter out and create subsets of an otherwise large dataset. By having the relevant amount of data, the analyst can quickly start working on the cleaning of the data and then creating models out of it. SQL sometimes is used with other tools such as Omni Sci or Zeppelin by Apache where PostgreSQL is used.

**Tableau/ QlikView/ Power BI**

The most important tool for report generation through the means of visualization. Tableau allows the user to quickly create interesting, complex, and detailed graphs that can magnify the impact of a report. The good aspect of this tool is that it is easy to use and requires less data preparation in order to get the desired output.

**Birt**

Another useful report based tool allows us to create graphs and dashboards, however, it is relatively more complex than tableau as the user needs to have a decent knowledge of Java to make the most out of it.

**Python**

One of the most advanced tools, python allows the user to perform multiple things. Python can be used to perform basic steps such as data cleaning to a complex aspect of analytics that includes the development of various kinds of models. The development of highly complex machine learning and deep learning models is particularly effective through this tool. Python also allows us to create reports and has libraries for visualization but it is up to the user to use them or use dedicated visualization tools.

**R**

This statistical tool created “by the statisticians for the statisticians”, allows a business analyst to perform all the descriptive and inferential statistics along with the development of statistical models. If compared to python it has a bit of a steep learning curve but this eventually pays off as it has a large community of users and is respected in the world of corporate as well as academia.

**MS Excel**

One of the most basic yet widely used and effective tools. The importance of MS Excel in the field of Business Analytics can be understood from realizing the difference between a sword and a needle. While performing the complex operation of data extraction, model development, and report generation, there are several heavyweight tools. MS Excel on the other hand sometimes is used at the very end or sometimes in the very beginning to provide an easy, user interactive experience to gain quick insights regarding the data or the final output and this is the reason that still many times the final output is in the form of an Excel.

**Sisense**

Sisense is an agile business intelligence (BI) solution that provides advanced tools to manage and support business data with analytics, visuals and reporting.

**Clear Analytics**

Clear Analytics is a data analytics solution that enables small to midsize business users to perform a variety of self-service analytics within an Excel-based environment. ... Clear Analytics can publish data to Microsoft's Power BI Cloud Portal to make dashboards available online or on mobile devices.

**Pentaho BI**

Pentaho is a business intelligence system designed to help companies make data-driven decisions, with a platform for data integration and analytics. The platform includes extract, transform, and load (ETL), big data analytics, visualizations, dashboards, reporting, data mining, and predictive analytics.

**MicroStrategy**

MicroStrategy is a Business Intelligence software, which offers a wide range of data analytics capabilities. As a suite of applications, it offers Data Discovery, Advanced Analytics, Data Visualizations, Embedded BI, and Banded Reports and Statements.

**Applications of Business Analytics**

Marketing –Business Analytics helps in the marketing field as it correctly studies consumer behavior and market trends. In addition, companies can thus base their strategies on this vital information and identify their target audience as well as identify new markets to penetrate.

Finance –In the field of finance, business analytics tools can help companies to uncover vital insights on stock performances by processing vast amounts of data.

Human Resources –HR professionals are turning to business analytics tools to conduct background checks on candidates and get relevant valuable information.

Manufacturing –The data that is collected can be subjected to these tools to provide information regarding inventory management, supply chain management, performance insight, and risk mitigation methods. Companies can also work on their operations skills through these tools.

Logo, company name

Description automatically generated

Business Analytics Tools

Reference-https://sigma4sap.com/?page\_id=466

**Trends in business analytics**

**Business Analytics Trends For 2021**

**Data Quality Management (DQM)**

The analytics trends in data quality grew greatly this past year. The development of business intelligence to analyze and extract value from the countless sources of data that we gather at a high scale, brought alongside a bunch of errors and low-quality reports: the disparity of data sources and data types added some more complexity to the data integration process.

**Data Discovery/Visualization**

Data discovery has increased its impact in the last year. The already mentioned survey conducted by the Business Application Research Center listed data discovery in the top 3 business intelligence trends by the importance hierarchy. BI practitioners steadily show that the empowerment of business users is a strong and consistent trend.

**Artificial Intelligence**

Artificial intelligence is already widely used in business applications, including automation, data analytics, and natural language processing. Across industries, these three fields of AI are streamlining operations and improving efficiencies. Automation alleviates repetitive or even dangerous tasks.



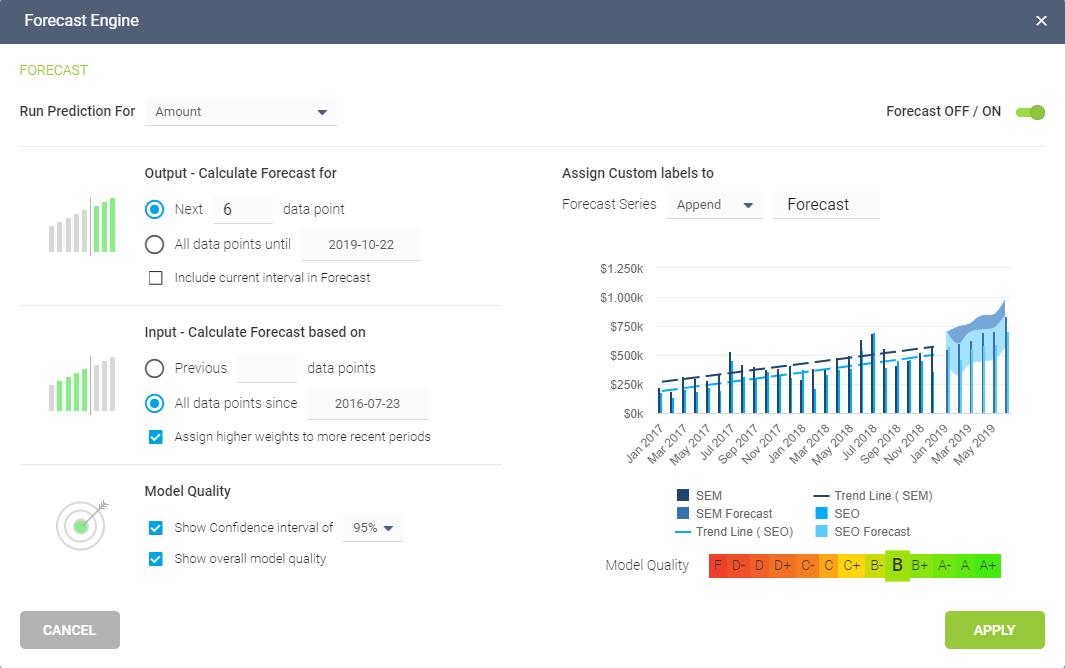
ArtificialIntelligence

Reference-https://feedsportal.com/artificial-intelligence-is-the-future-of-business-analytics/

**Predictive And Prescriptive Analytics Tools**

Predictive analytics is the practice of extracting information from existing data sets in order to forecast future probabilities. It’s an extension of data mining which refers only to past data. Predictive analytics includes estimated future data and therefore always includes the possibility of errors from its definition, although those errors steadily decrease as software that manages large volumes of data today becomes smarter and more efficient. Predictive analytics indicates what might happen in the future with an acceptable level of reliability, including a few alternative scenarios and risk assessment. Applied to business, predictive analytics is used to analyze current data and historical facts in order to better understand customers, products, and partners and to identify potential risks and opportunities for a company.

Prescriptive analytics goes a step further into the future. It examines data or content to determine what decisions should be made and which steps taken to achieve an intended goal. It is characterized by techniques such as graph analysis, simulation, complex event processing, neural networks, recommendation engines, heuristics, and machine learning. Prescriptive analytics tries to see what the effect of future decisions will be in order to adjust the decisions before they are actually made. This improves decision-making a lot, as future outcomes are taken into consideration in the prediction. Prescriptive analytics can help you optimize scheduling, production, inventory, and supply chain design to deliver what your customers want in the most optimized way.



Predictive And Prescriptive Analytics Tools

Reference-https://www.datapine.com/blog/business-intelligence-trends/

**Collaborative Business Intelligence**

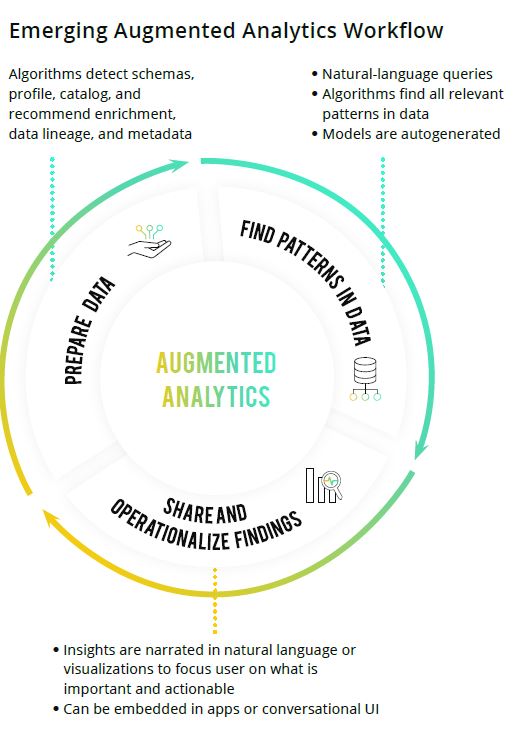
Collaborative BI (collaborative business intelligence) is the merging of business intelligence software with collaboration tools, including social and Web 2.0 technologies, to support improved data-driven decision making.

**Data-driven Culture**

The concept of a data-driven culture treats data as the main resource for leveraging insights in every department of the organization. While companies have always been interested in their numbers, the extent of data use is exercised at a higher level within a data-driven culture.

**Augmented Analytics**

Augmented analytics is the use of statistical and linguistic technologies to improve data management performance, from data analysis to data sharing and business intelligence. It is somehow connected to the ability to transform big data into smaller, more usable, datasets.



Augmented Analytics

Reference-https://www.sisense.com/whitepapers/augmented-analytics-the-future-of-business-intelligence/

**Mobile BI**

Mobile BI refers to the recent trend of business users accessing their data and dashboards on mobile and tablet devices. This change is forcing designers to re-evaluate the user experience (UX) and replicate the same relevant content usually seen solely on desktop devices on mobile.



Mobile BI

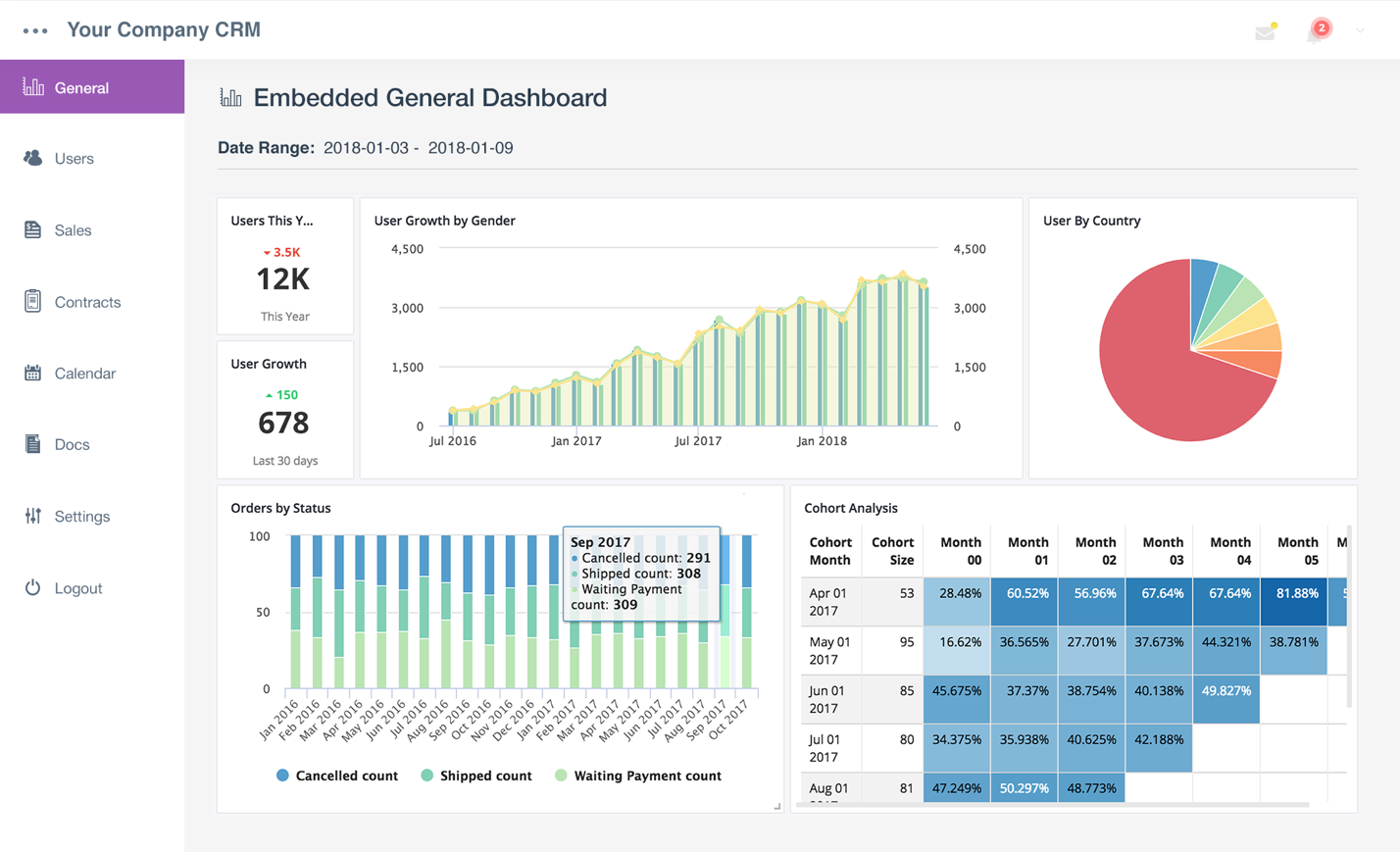
<https://matillion.com/wp-content/uploads/2015/09/mobile-business-intelligence-solution-needs-to-be.jpg>

**Data Automation**

Business intelligence topics wouldn’t be complete without data (analysis) automation. In the last decade, we saw so much data produced, stored, and ready to process that companies and organizations were seriously looking for modern data automation solutions to tackle massive volumes of information that has been collected. Gartner predicts that next year more than 40% of data science tasks will be automated, hence, this is one of the trends in business intelligence that we need to keep an eye on.

**Embedded analytics**

Embedded analytics is the integration of analytic content and capabilities within business process applications. It provides relevant information and analytical tools designed so users can work smarter and more efficiently in the applications they use every day.



Embedded analytics

Reference-https://www.holistics.io/features/embedded-analytics/

**Natural language processing**

Natural language processing (NLP) is a subfield of linguistics, computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data.

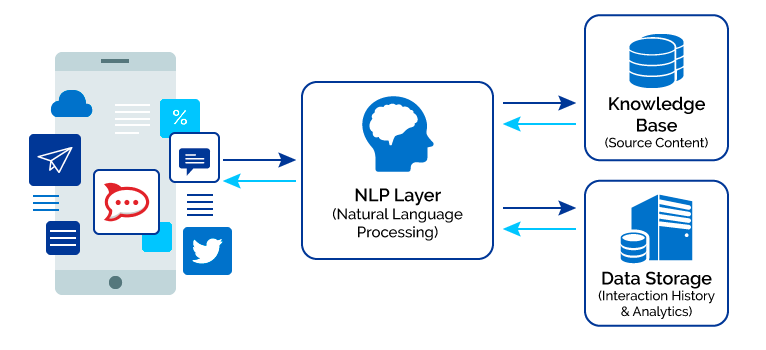
Speech Recognition: Converting spoken words into data a computer can understand. This is the NLP technology you use every time you ask Siri, Cortana, Echo or Google Voice a question.

Machine Translation: Translating text from one language to another. This is the tech that underlies translation apps like Google Translate.

Natural Language Generation: Outputting information as a human language. This is the tech you use every time Siri or Cortana answers your question.

Semantic Search: Closely linked to speech recognition, as above, this allows you to ask natural questions of an app like Siri, rather than having to formulate your question in a particular, unnatural way.

Machine Learning: Machine learning is a whole other topic, but essentially, it uses the data that NLP interprets to “teach” itself about future actions.



NLP

Reference-https://marutitech.com/what-nlp-reasons-everyone-retail-use-it/

**Descriptive analytics**

**What is Descriptive Analytics?**

Descriptive analytics is the interpretation of historical data to better understand changes that have occurred in a business. Descriptive analytics describes the use of a range of historic data to draw comparisons.

Descriptive analytics is a statistical method that is used to search and summarize historical data in order to identify patterns or meaning.

For learning analytics, this is a reflective analysis of learner data and is meant to provide insight into historical patterns of behaviors and performance in online learning environments.

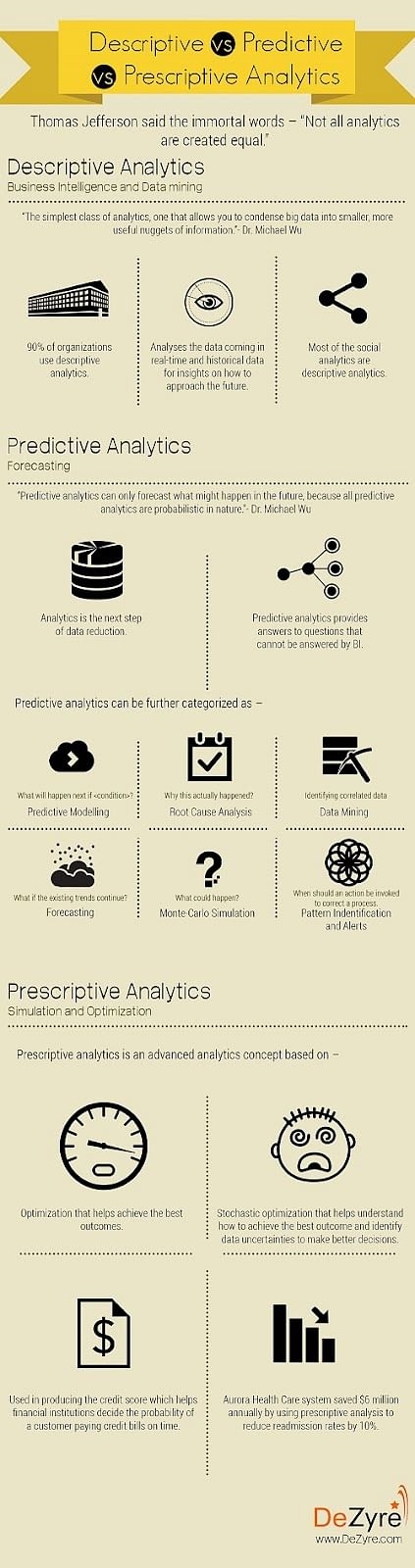
For example, in an online learning course with a discussion board, descriptive analytics could determine how many students participated in the discussion, or how many times a particular student posted in the discussion forum.

Descriptive approach is considered to be the foundation of research. Its logical design is based on statistics of the research analysis. Since this analysis doesn’t explain the cause of result, hence it can’t take into account the validity of research results. Here are the common methods used in descriptive analytics:

**Observation Method:** In this method data is observed in both natural and artificial ways in order to draw meaningful conclusions. It is an effective way of inferring from natural observation since we can obtain original results of the research. On the other hand, in the artificial method results would depend on the quantity of the data provided for observation.

**Case study Method:** It involves a deep study on all the problems discussed. It makes us understand a particular situation more closely.

**Survey Method:** In this method questionnaires are prepared and given to the participants. After receiving the answers the research is preceded and results are concluded.



Descriptive Analytics

Reference-https://www.dezyre.com/article/types-of-analytics-descriptive-predictive-prescriptive-analytics/209

**How does descriptive analytics work?**

Data aggregation and data mining are two techniques used in descriptive analytics to discover historical data. Data is first gathered and sorted by data aggregation in order to make the datasets more manageable by analysts.

Data mining describes the next step of the analysis and involves a search of the data to identify patterns and meaning. Identified patterns are analyzed to discover the specific ways that learners interacted with the learning content and within the learning environment.

**What can descriptive analytics tell us?**



Descriptive Analytics

Reference-https://www.valamis.com/hub/descriptive-analytics

The kind of information that descriptive analytics can provide depends on the learning analytic capability of the learning management system (LMS) being used and what the system is reporting on specifically.

Some common indicators that can be identified include learner engagement and learner performance. With learner engagement, analysts can detect the participation level of learners in the course and how and when course resources were accessed.

Performance data provides analysts with insight into how well learners succeeded on the course; this information could come from data taken from assessments or assignments. It’s important to note that insights learned from descriptive analysis are not used for making inferences or predictions about a learner’s future performance.

The analytical method is meant to provide strategic insight into where learners, or a specific learner, may have needed more support. It can also help course designers improve the design of learning by providing insight into what went well and what did not go well on the course.

**Examples of descriptive analytics**

Many LMS platforms and learning systems offer descriptive analytical reporting with the aim of helping businesses and institutions measure learner performance to ensure that training goals and targets are met.

The findings from descriptive analytics can quickly identify areas that require improvement - whether that be improving learner engagement or the effectiveness of course delivery.

Here are some examples of how descriptive analytics is being used in the field of learning analytics:

* Tracking course enrollments, course compliance rates,
* Recording which learning resources are accessed and how often
* Summarizing the number of times a learner posts in a discussion board
* Tracking assignment and assessment grades
* Comparing pre-test and post-test assessments
* Analyzing course completion rates by learner or by course
* Collating course survey results
* Identifying length of time that learners took to complete a course

Graphical user interface, application

Description automatically generated

Example of Descriptive Analytics

Reference-https://www.sisense.com/glossary/descriptive-analytics/

The functions delivered by descriptive analytics solutions fall broadly into five categories:

**State business metrics:** Determine which metrics are important for evaluating performance against business goals. Some goal examples would be to increase revenue, reduce costs, improve operational efficiency, and measure productivity. Each goal must have associated KPIs to help monitor achievement.

**Identify data required:** Business data is located in many different sources within the enterprise, including systems of record, databases, desktops, and shadow IT repositories. To measure accurately against KPIs, companies must catalog and prepare the correct data sources to extract the needed data and calculate metrics based on the current state of the business.

**Extract and prepare data:** Data must be prepared for analysis. Deduplication, transformation, and cleansing are a few examples of the data preparation steps that need to take place prior to analysis. Often, this is the most time-consuming and labor-intensive step, requiring up to 80% of an analyst’s time, but it is critical for ensuring accuracy.

**Analyze data:** Data analysts can create models and run analyses such as summary statistics, clustering, and regression analysis on the data to determine patterns and measure performance. Key metrics are calculated and compared with stated business goals to evaluate performance based on historical results. Data scientists often use open source tools like R and Python to programmatically analyze and visualize data.

**Present data:** Results of the analytics are usually presented to stakeholders in the form of charts and graphs. This is where the data visualization mentioned earlier comes into play. BI tools give users the ability to present data visually in a way that non-data analysts can understand. Many self service data visualization tools also enable business users to create their own visualizations and manipulate the output.

**Advantages of descriptive analytics**

When learners engage in online learning, they leave a digital trace behind with every interaction they have in the learning environment.

This means that descriptive analytics in online learning can gain insight into behaviours and performance indicators that would otherwise not be known.

Here are some advantages to utilizing this information:

* Quickly and easily report on the Return on Investment (ROI) by showing how performance achieved business or target goals.
* Identify gaps and performance issues early - before they become problems.
* Identify specific learners who require additional support, regardless of how many students or employees there are.
* Identify successful learners in order to offer positive feedback or additional resources.
* Analyze the value and impact of course design and learning resources.

**Introduction to Big Data Analytics**

## What is Data?

The quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.

## What is Big Data?

**Big Data** is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

In this Big Data analytics, you will learn,

* [What is Data?](https://www.guru99.com/what-is-big-data.html#5)
* [What is Big Data?](https://www.guru99.com/what-is-big-data.html#6)
* [What is an Example of Big Data?](https://www.guru99.com/what-is-big-data.html#1)
* [Types Of Big Data](https://www.guru99.com/what-is-big-data.html#2)
* [Characteristics Of Big Data](https://www.guru99.com/what-is-big-data.html#3)
* [Advantages Of Big Data Processing](https://www.guru99.com/what-is-big-data.html#4)



What is Big Data?

Reference: https://www.guru99.com/images/Big\_Data/061114\_0759\_WhatIsBigDa1.jpg

## What is an Example of Big Data?

Following are some of the Big Data examples-

The **New York Stock Exchange** is an example of Big Data that generates about **one terabyte** of new trade data per day.



Reference: https://www.guru99.com/images/Big\_Data/061114\_0759\_WhatIsBigDa2.jpg

**Social Media**

The statistic shows that **500+terabytes** of new data get ingested into the databases of social media site **Facebook**, every day. This data is mainly generated in terms of photo and video uploads, message exchanges, putting comments etc.



Reference: https://www.guru99.com/images/Big\_Data/061114\_0759\_WhatIsBigDa3.jpg

A single **Jet engine** can generate **10+terabytes** of data in **30 minutes** of flight time. With many thousand flights per day, generation of data reaches up to many **Petabytes.**



Reference: https://www.guru99.com/images/Big\_Data/061114\_0759\_WhatIsBigDa4.jpg

## Types Of Big Data

Following are the types of Big Data:

1. **Structured**
2. **Unstructured**
3. **Semi-structured**

### **Structured**

Any data that can be stored, accessed and processed in the form of fixed format is termed as a ‘structured’ data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data (where the format is well known in advance) and also deriving value out of it. However, nowadays, we are foreseeing issues when a size of such data grows to a huge extent, typical sizes are being in the rage of multiple zettabytes.

**Do you know? 1021 bytes** equal to **1 zettabyte** or **one billion terabytes** forms **a zettabyte**.

Looking at these figures one can easily understand why the name Big Data is given and imagine the challenges involved in its storage and processing.

**Do you know?** Data stored in a relational database management system is one example of a **‘structured’** data.

**Examples Of Structured Data**

An ‘Employee’ table in a database is an example of Structured Data

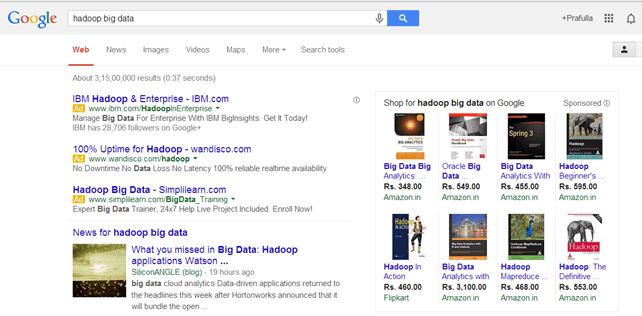
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Employee\_ID** | **Employee\_Name** | **Gender** | **Department** | **Salary\_In\_lacs** |
| 2365 | Rajesh Kulkarni | Male | Finance | 650000 |
| 3398 | Pratibha Joshi | Female | Admin | 650000 |
| 7465 | Shushil Roy | Male | Admin | 500000 |
| 7500 | Shubhojit Das | Male | Finance | 500000 |
| 7699 | Priya Sane | Female | Finance | 550000 |

### **Unstructured**

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don’t know how to derive value out of it since this data is in its raw form or unstructured format.

**Examples Of Un-structured Data**

The output returned by ‘Google Search’



Example Of Un-structured Data

Reference: https://www.guru99.com/images/Big\_Data/061114\_0759\_WhatIsBigDa5.png

### **Semi-structured**

Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational [DBMS](https://www.guru99.com/what-is-dbms.html). Example of semi-structured data is a data represented in an XML file.

Examples Of Semi-structured Data

Personal data stored in an XML file-

<rec><name>Prashant Rao</name><sex>Male</sex><age>35</age></rec>

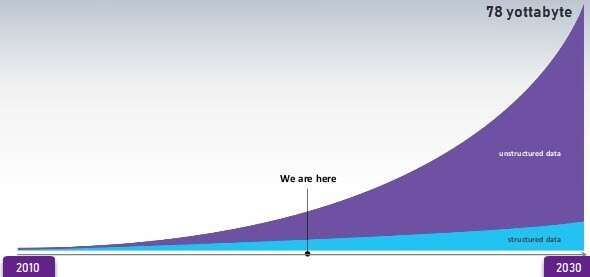
<rec><name>Seema R.</name><sex>Female</sex><age>41</age></rec>

<rec><name>Satish Mane</name><sex>Male</sex><age>29</age></rec>

<rec><name>Subrato Roy</name><sex>Male</sex><age>26</age></rec>

<rec><name>Jeremiah J.</name><sex>Male</sex><age>35</age></rec>

**Data Growth over the years**



Data Growth over the years

Reference: <https://www.guru99.com/images/1/big-data-growth.jpg>

Please note that [web application](https://www.guru99.com/difference-web-application-website.html) data, which is unstructured, consists of log files, transaction history files etc. OLTP systems are built to work with structured data wherein data is stored in relations (tables).

## Characteristics Of Big Data

Big data can be described by the following characteristics:

* Volume
* Variety
* Velocity
* Variability

**(i) Volume –** The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data. Hence, **‘Volume’** is one characteristic which needs to be considered while dealing with Big Data solutions.

**(ii) Variety –** The next aspect of Big Data is its **variety**.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

**(iii) Velocity –** The term **‘velocity’** refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors,[Mobile](https://www.guru99.com/mobile-testing.html)devices, etc. The flow of data is massive and continuous.

**(iv) Variability –** This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

## Advantages Of Big Data Processing

Ability to process Big Data in DBMS brings in multiple benefits, such as-

* Businesses can utilize outside intelligence while taking decisions

Access to social data from search engines and sites like facebook, twitter are enabling organizations to fine tune their business strategies.

* Improved customer service

Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

* Early identification of risk to the product/services, if any
* Better operational efficiency

Big Data technologies can be used for creating a staging area or landing zone for new data before identifying what data should be moved to the [data warehouse](https://www.guru99.com/data-warehousing.html). In addition, such integration of Big Data technologies and data warehouse helps an organization to offload infrequently accessed data.

## Summary

* Big Data definition : Big Data meaning a data that is huge in size. Bigdata is a term used to describe a collection of data that is huge in size and yet growing exponentially with time.
* Big Data analytics examples includes stock exchanges, social media sites, jet engines, etc.
* Big Data could be 1) Structured, 2) Unstructured, 3) Semi-structured
* Volume, Variety, Velocity, and Variability are few Big Data characteristics
* Improved customer service, better operational efficiency, Better Decision Making are few advantages of Bigdata