Business Analyst

A business analyst uses her knowledge and experience, in combination with the insights generated by a data analyst to make decisions that affect the business. They work within a business or organization to identify and implement improvements to help a company achieve its goals.

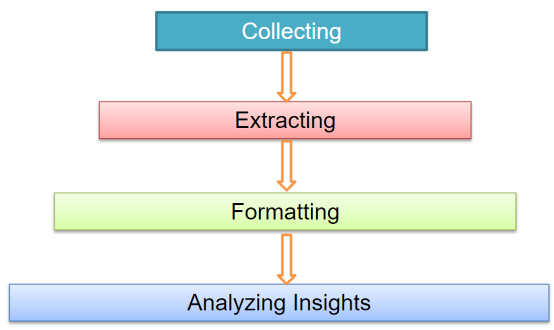
The distinction between Data Analyst and Business Analyst

Data analysts and business analysts both help drive data-driven decision-making in their organizations. Data analysts tend to work more closely with the data itself, while business analysts tend to be more involved in addressing business needs and recommending solutions.

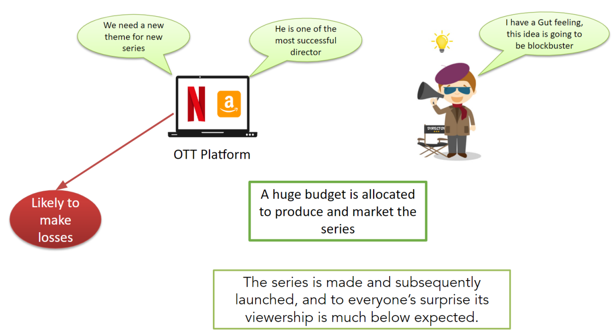
Data-Driven Decision Making

Data-driven decision-making means making decisions that are supported by data rather than decisions based on observations, gut, or even instinct. With the advancement of technology, the data available to organizations has become tremendous. This abundance of data provides organizations several opportunities to understand their performance to a greater degree, optimize their performance and leverage it for growth.

In simple terms, the process is – collecting, extracting, formatting, and analyzing insights.



Example



Let's consider a leading online video streaming (OTT) platform with several web series under its belt – some successful and some not. The platform is trying to decide which theme to make a new series on. A big director approaches the platform with an out-of-the-box idea that he knows in his gut is a blockbuster idea. The platform assumes that the director's past success is sufficient proof to back the project. A huge budget is allocated to produce and market the series.

The series is made and subsequently launched, and to everyone's surprise, its viewership is much below expected. The response is underwhelming, and the platform is likely to make losses.

How could the platform have avoided losses? Could the platform have used data points available to it to make an informed, data-based decision that has more weight than just a gut feeling?

Of course, yes. And this is exactly what more and more OTT platforms are doing.

There are various factors that the platform could have relied on, which are not limited to; but include analyzing similar series made in the past, the director & actors' past track record, and currently trending themes in the OTT space among others. The likeliness of the show underperforming would have been predictable before the platform took it on.

Advantages of using Data in the decision-making process

- **Risk assessment**

  - As often data-enabled insights help us identify that carpet bombs that can be attached to a project that we're working on. By this, what I mean is that often when you are working on a project, there are certain risks that you know about and you do whatever is within your means to mitigate them. However, there are certain risks that you do not know about, and often – using data insights in the decision-making process; helps you identify these risks.

- **Trend spotting and predicting outcomes**

  - Data insights can make it easier for you to gauge trends, which can help you predict outcomes. You can think about the failed web series example that we spoke about earlier to understand this more in detail. Taking a more practical view of that example, since the OTT platform commissioned the idea based on the credibility of the director, let's say that even data suggests the same and back his or her credibility and past track record. When we look at the actors' track records and trending themes in the OTT space, here is where data could paint a different picture and there could be a chance that the actors, who were part of the series, could be more successful in genres that were different from this particular genre. Or there could also be a chance that even though the content theme that the web series was based on is in trend in the OTT space, a plethora of content with the same theme on rival OTT platforms would have added a lot of competition to it and hence, this could be a red flag that could be suggested by data. Hence, as we can see from this; trend spotting becomes much easier with data and as a result taking decisions through it, by using it as a tool along with your knowledge; experience and understanding can enhance the precision of your decisions.

- **Gauge customer satisfaction**

  - All organizations strive for customer satisfaction. They may or may not achieve it. But it is definitely what they strive for. When we speak about customer satisfaction, a key challenge is – how do you gauge customer satisfaction? This is where data plays a key role. To quote an example, let's look at FMCG as a domain. In FMCG, customer loyalty is a key measure of customer satisfaction; and often this is defined by repeat purchases by a customer. If we look at the online sales channel, it is often easy to identify repeat purchases as the company's website or the e-commerce platform through which it is selling its product; often maintains a record of the details of its customers. But what about offline channels? Often, in India; offline channels like Kirana stores do not keep individual customer records of their purchases and the products they have purchased, and hence, it's more difficult to measure customer satisfaction. As a result, marketers have developed a parameter of 'same-store sales growth to gauge customer satisfaction in offline retail channels. Through this example, we can see how data measurement helps in gauging customer satisfaction through both online and offline mediums.

- **Enables innovation**

  - Often as a Business Analyst, you have to make decisions on both the market and the product. When we speak about the product, one of your key inputs would be in new product development. More often than not, the success or failure of a product depends on the market. Here is where data plays a critical role in the decision-making process. Using data and with the help of a data analyst, you would be able to study the market – its tastes and preferences and work backwards to develop the product. Often, this reverse engineering process of product development, when done correctly with data is something that helps in ensuring the success of a new product.

- And lastly, data-backed decisions help improve both external and internal processes in an organization. External processes include interactions of an organization with external stakeholders like its customers, vendors, and distribution & promotion partners among others; while internal processes include interactions of an organization with internal stakeholders, predominantly –

  its employees.

DATA

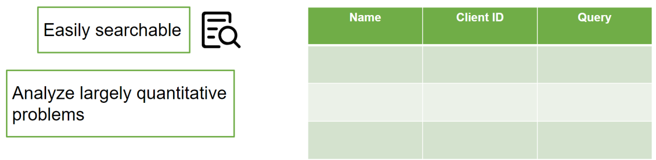
Broadly, there are two kinds of data – structured & unstructured data.

Structured data

is clearly defined data types with patterns that make them easily searchable. In simpler terms, when a layman looks at structured data, she can easily make out that this is data.

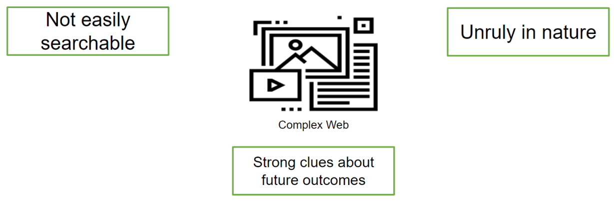
Structured data typically contains data types that are combined in a way to make them easy to search for in their data set. This means that structured data is easily detectable via search because it is highly organized information.

The image on your screen is an example of structured data. It's neatly segregated into fixed fields. It's the data that most of us are used to and like working with to analyze largely quantitative problems.



Unstructured data

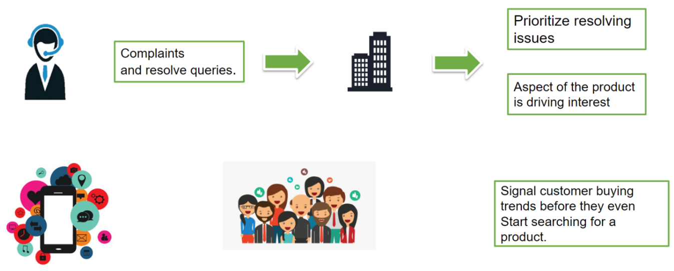
is comprised of data that is usually not as easily searchable. Think about an image or an audio or perhaps a video. When you think about them, they do not essentially seem like data or data types. However, a lot of data can be extracted from them, and hence this potential and possibility of extracting data make them a part of unstructured data.



Unstructured data, on the other hand, makes the capability to search much more difficult.

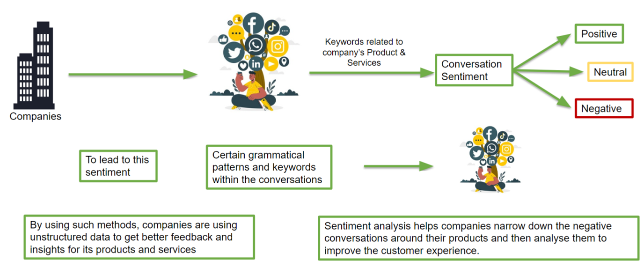
Unstructured data, as evident from its very name, does not conform neatly to a spreadsheet but may have its internal structure. It includes everything outside the boundaries of structured data.

While it is unruly, it is also incredibly valuable—unstructured data has the potential to depict a complex web of information that offers strong clues about future outcomes.



Unstructured data analysis is a crucial part of the data analytics process. Think of customer web chats, for example, a platform where customers commonly dish out their complaints and resolve queries. When analyzed as a whole, this web chat data can help guide companies on what to prioritize resolving or what aspect of the product is driving the most interest. Or social media data, which can signal customer buying trends before they even start searching for a product. If structured data could be considered a company's backbone, unstructured data is its competitive edge.

Example:

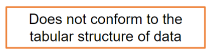


One interesting modern-day example of unstructured data is the use of it through web-scraping, which often leads to sentiment analysis, which helps companies gather insights on their products and services. Often, companies use web scraping as a method to analyze social media conversations. Through this method, a program is run on a social media site, which searches for keywords related to the company's products and services. Once the keywords bring out the relevant conversations, the program attempts to classify the conversation to find the sentiment that is present in it.

The sentiment can be classified as positive, negative or neutral, or even something else – as per the requirement of the company and the project. To lead to this sentiment, the program has searched for certain grammatical patterns and keywords within the conversations that it has derived from social media sites.

By using such methods, companies are using unstructured data to get better feedback and insights for their products and services. Sentiment analysis helps companies narrow down the negative conversations around their products and then analyzes them to improve the customer experience.

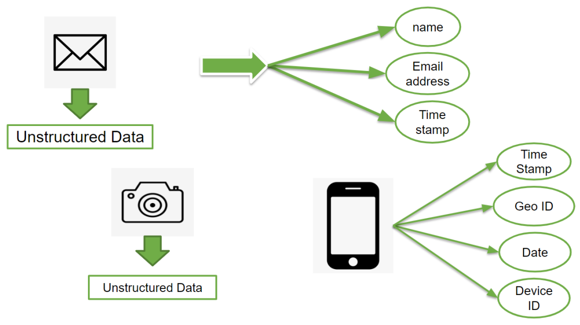
Semi-Structured



Semi-structured data is a form of structured data that does not conform to the tabular structure of data but does contain tags or other markers to differentiate elements and enforce hierarchies within the data.

Email messages are a good example of semi-structured data. While the actual content of the email is unstructured, this format does contain structured data such as the name and email address of the sender and recipient, and the time sent among others.

Another example is a digital photograph. The image itself is unstructured, but if the photo was taken on a smartphone, for example, it would be date and time stamped, geo-tagged, and would have a device ID. Once stored, the photo could also be given tags that would provide a structure, such as ‘dog’ or ‘pet.’



Making Decisions

As a business analyst, your primary job would be to make decisions based on your knowledge, understanding, and insights generated through data analytics.

Hence, it becomes really important for you to understand data analytics as an understanding of that domain can act as a competitive edge in your decision-making.

Data Analytics

Data analytics is the extraction of insights from data using statistical techniques and technologies.

A big challenge that is often faced by data analytics teams is the availability of quality data, which they can work with. Therefore, before moving on to data analytics, let’s see an example that deals with data collection and data cleaning.

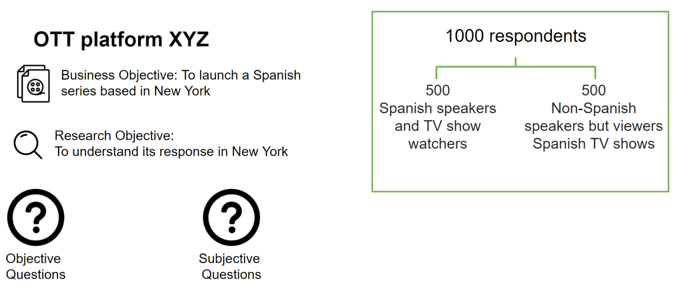
Example
Let’s take the help of a case study to understand this more in detail. Let’s say that an OTT platform, XYZ is looking to launch a Spanish series and wants to predict its response in New York.

It forms a research team, which decides that they want to conduct a primary study of 1000 respondents, 500 of whom are Spanish speakers and TV show watchers; while the other 500 do not speak Spanish but watch Spanish TV shows.

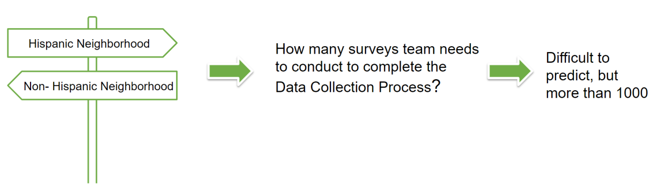
They prepare a detailed questionnaire consisting of both objectives as well as subjective questions that they want all respondents to answer.

The team splits itself, with some members going to Hispanic neighbourhoods; while others going to non-Hispanic neighbourhoods to conduct the interviews.

Now, how many surveys do you think the team needs to conduct to complete the data collection process for this research?



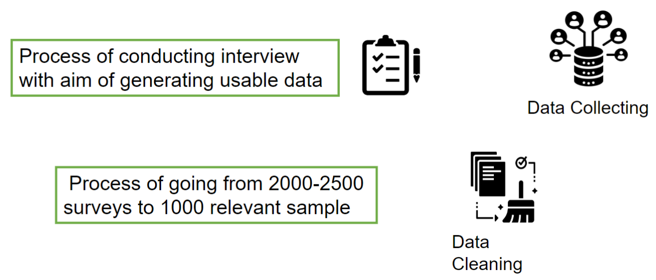
The exact answer cannot be predicted. However, we’re pretty sure that it would be more than 1000.



The team going to Hispanic neighbourhoods may share the questionnaire with respondents, who speak Spanish but do not watch TV in that language.

The team going to non-Hispanic neighbourhoods will definitely share the questionnaire with respondents, who do not speak Spanish and do not watch TV in that language.

To get to a defined target of 1000 respondents, with 500 of them speaking Spanish and watching TV in that language, while 500 of them not speaking Spanish and watching TV in that language, there is a chance that the team may have to conduct 2000 – 2500 surveys.



The process of going out and conducting interviews to generate usable data is known as

data collection.

The process of going from 2000 or 2500 surveys that the team may have had to conduct as part of data collection to 1000 relevant ones is known as data cleaning. In this process, the team is going from a large data set to a relevant data set that fits the objective of its research.

Types of Data Analytics

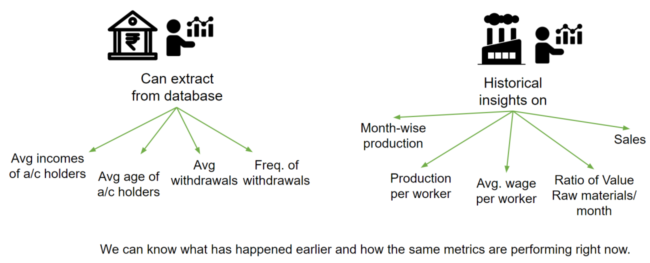
There are 4 predominant types of data analytics - Descriptive, Diagnostic, Predictive & Prescriptive.

Descriptive analytics

Descriptive analytics answers the question: What has happened and what is happening right now?

Descriptive analytics uses historical and current data from multiple sources to describe the present state by identifying trends and patterns. Descriptive analytics uses basic arithmetic like average, sum, and percentage changes among others. Many times in data analysis, descriptive analytics is the first step that helps organizations understand the facts of what has already happened.

Example:



An analyst looking at a large database of account holders of a bank can extract information about the average incomes of account holders, their average age, their average withdrawals, and their frequency of withdrawals among many other such parameters that can be defined by basic arithmetic calculations.

Another example could be a manufacturing unit, where an analyst can get historical insights regarding the company's month-wise production, production per worker, the average wage per worker, the ratio of the value of raw materials purchased per month, and sales among other such parameters.

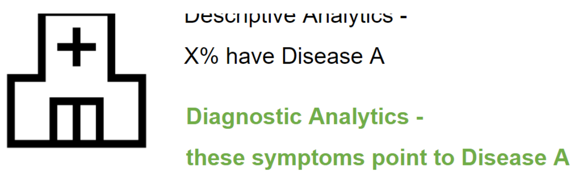
As we can see, in both the examples of descriptive analytics, we can know what has happened earlier and how the same metrics are performing right now.

Diagnostics Analysis

Diagnostic analytics answers the question: Why did this happen?

Diagnostic analytics uses data, which is often built on descriptive analytics to discover the reasons or causes for past performance. It is oftentimes referred to as root cause analysis.

Example:
Let's say that an unusually high number of people come to the emergency room of a hospital complaining of similar symptoms. After testing them, it is found that about 70% of them have Disease A. Descriptive analytics helps us get to the final number of people having Disease A. However, it is Diagnostic analysis that can help to make correlations between the symptoms and determine what symptoms point to Disease A.

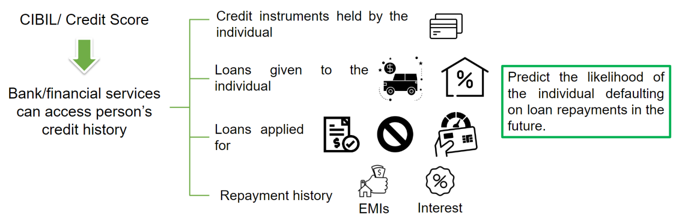


Predictive Analysis

Predictive analytics answers the question: What is likely to happen in the future?

Predictive analytics applies techniques such as statistical modelling and forecasting to the output of descriptive and diagnostic analytics to make predictions about future outcomes. Predictive analytics is often considered a type of “advanced analytics,” and frequently depends on technology.

Example:



Whenever an individual applies for a loan, the bank or financial services company where she has applied, asks her for documents like her Aadhar Card, and PAN Card among others. The reason for this is that they want to access the individual’s CIBIL, which is a credit score, and along with it comes the credit history of the individual. Through the CIBIL report, the bank can access various aspects of the individual’s credit history, which include: The types of credit instruments held by the individual – like active credit cards, past credit cards, active loans, and past loans among others

1. **The types of loans given to the individual – like whether the individual has accessed an auto loan, home loan, personal loan, consumer loan, education loan, or say a loan against property.**
2. **The loans that the individual has applied for – you would not believe it but even if you apply for a loan that does not get approved, it still shows up on your CIBIL report and affects your credit score.**
3. **The repayment history of the individual – which includes insights on whether the individual has paid her EMIs on time, does the individual pay her credit card on time, or whether has there ever been a delay in EMI, interest, or credit card payment.**

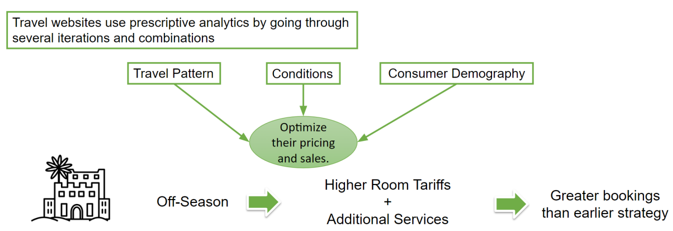
The core idea behind accessing an individual’s CIBIL is to predict the likelihood of the individual defaulting on loan repayments in the future.

Prescriptive Analysis

Prescriptive analytics answers the question: What do we need to do?

Prescriptive analytics is a type of advanced analytics that involves the application of testing and other techniques to recommend specific solutions that will deliver desired outcomes. In business, prescriptive analytics uses machine learning, business rules, and algorithms.

Example:



Travel websites use prescriptive analytics by going through several iterations and combinations of factors such as travel patterns, conditions, and consumer demographics to optimize their pricing and sales.

An interesting example of this could be in the hospitality industry. Using a prescriptive analytics model, a luxury hotel may find that during the off-season, increasing their room tariffs and offering free additional services such as free spa appointments, and a couple of complimentary meals is likely to boost bookings compared to their initial strategy of lower tariffs with few additional services.