What is data analytics ?

Data analytics is the process of analyzing raw data to uncover meaningful patterns, insights, and trends. It involves examining large datasets using various statistical and computational techniques to extract valuable information that can be used for decision-making and problem-solving. Data analytics encompasses descriptive, predictive, and prescriptive analytics methods, which help businesses and organizations understand past performance, predict future outcomes, and prescribe actions to achieve desired results.

Ways to use data analytics ?

Improved decision making
Better customer service
Efficient operations
Effective marketing
Etc..

Steps involved in data analytics

1. **Understand the problem**
   Understand the business problem, define the organizational goals and plan for lucrative solutions.

2. **Data collection**

   Gather the right data from various sources and other information based on your priorities.

3. **Data cleaning**

   Clean the data to remove unwanted, redundant and missing values and make it ready for analysis.

4. **Data exploration and analysis**

   Use data visualization and business intelligence tools, data mining techniques and predictive modeling to analyze data.

5. **Interpret the results**

   Interpret the results to find out hidden patterns, future treads and gain insights.

Data analytics tools

1. **Python**

   Python is an object-oriented open-source programming language that supports a range of libraries for data manipulation, data visualization and data modeling

2. **R language**

   R is an open-source programming language majorly used for numeric and statistical analysis. It provides a range of libraries for data analysis and visualization.

3. **Tableau**

   Tableau is a data visualization tool that can create a wide range of visualizations to interactively present the data, build reports and dashboards to showcase insights and trends.

4. **Power BI**

   Power BI is a business intelligence tool developed by Microsoft that has easy drag and drop functionality and supports multiple data sources with features that make data visually appealing.

5. **QlikView**

   QlikView provides interactive analytics with in-memory storage technology to analyze vast volumes of data and use data discoveries to support decision making.

6. **Apache Spark**

   Apache Spark is an open-source data analytics engine to process data in real-time and carry out complex analytics using SQL queries and machine learning algorithms.

7. **SAS**

   SAS is a statistical analysis software that can help you perform analytics, visualize data, write SQL queries and build machine learning models to make future predictions.

Data analytics applications

1. **Retail**

   Retailers use data analytics to understand their customer needs and buying habits to predict trends, recommend new products and boost their business.

2. **Health care**

   Healthcare industries analyze patient data to provide lifesaving diagnoses and treatment options. They also deal with healthcare plans, insurance information to derive key insights.

3. **Manufacturing**

   Using data analytics, manufacturing sectors can discover new cost saving and revenue opportunities. They can solve complex supply chain issues, labor constraints and equipment breakdowns.

4. **Banking**

   Banking institutions gather and access large volumes of data to derive analytical insights and make sound financial decisions. They find probable loan defaulters, customer churn out rate and detect frauds in transactions.

5. **Logistics**

   Logistics companies use data analytics to develop new business models, optimize routes, improve productivity and order processing capabilities as well as performance management.

Companies using data analytics

Amazon, Accenture, Cigna, Cerner, target, McAfee, Rapido, Flipkart, Walmaet

Walmart case study

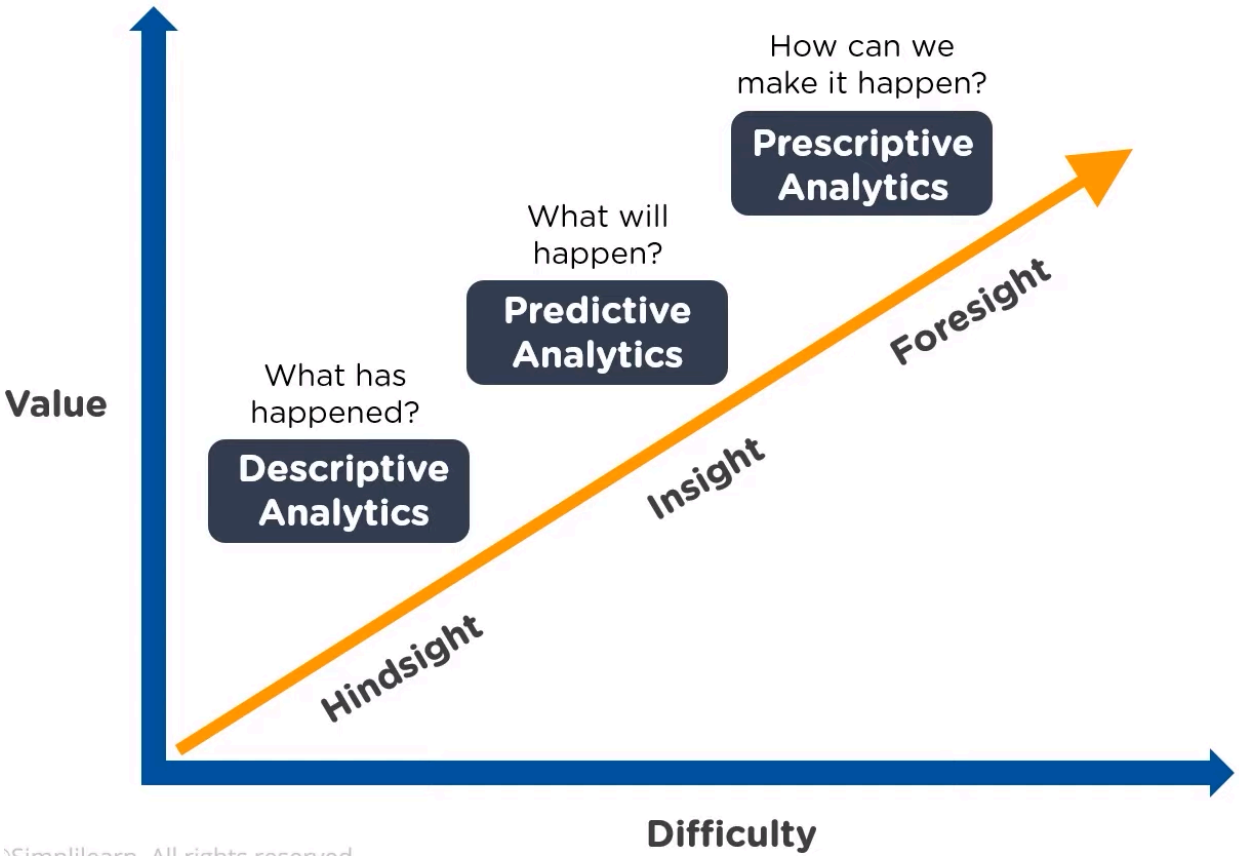
Walmart is an American multinational retail company with over 11,500 stores in 27 countries and eCommerce websites in 10 countries. It has over 2.2 million employees around the world, 1.5 million U.S. alone. More than 240 million customers shop at Walmart each week online and at its banner stores. Walmart collects 2.5 petabytes of data from 1 million customers every hour. To make sense of all this information, Walmart has created Data cafe - a state-of-the-art analytics hub. Over 200 streams of internal and external data, including 40 petabytes of recent transactional data, can be modified , manipulated and visualized. Walmart constantly analyzes over 100 million keywords to know what people near each store are saying each on social media. Walmart uses modern tools and technologies to derive business insights and improve customer satisfaction.

Using all this technologies walmart can

- Better management of the supply chain.

- Optimize product assortment.
- Personalize the shopping experience.
- Give relevant products recommendations.
- Optimize and analyze transportation lanes and routes for its fleet of trucks.

Types of data analytics



Why python for data analytics

- Easy to learn with simple syntax.
- Scalable and flexible.
- Huge collection of libraries.
- Graphics and visualization.
- Community support.

Python libraries for data analytics

- **NumPy**
    - Supports n-dimensional arrays
    - Provides numerical computing
    - Useful for linear algebra and fourier transform
- **Pandas**
    - Useful for handling missing data
    - Perform mathematical operations
    - Provides functions to manipulate data
- **Matplotlib**
    - Plotting library in python
    - Several toolkits extend matplotlib functionality
    - Creates interactive visualizations
- **Scipy**
- **Scikit**