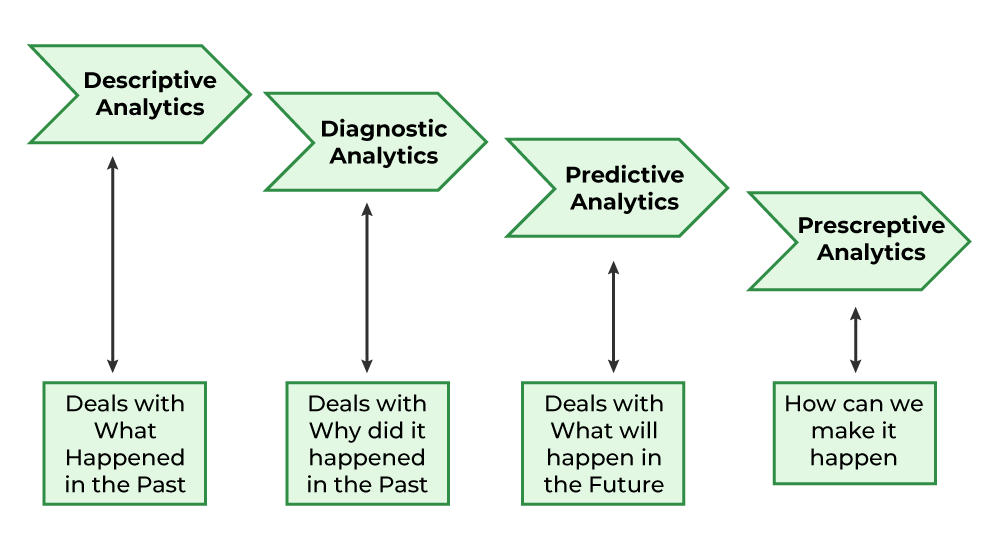
**What is Data Analytics :**

***Data analytics***is an important field that involves the process of collecting, processing, and interpreting data to uncover insights and help in making decisions. Data analytics is the practice of examining raw data to identify trends, draw conclusions, and extract meaningful information. This involves various techniques and tools to process and transform data into valuable insights that can be used for decision-making.

**Types of Data Analytics**

There are four major types of data analytics:

1. **Predictive (forecasting)**
2. **Descriptive (business intelligence and data mining)**
3. **Prescriptive (optimization and simulation)**
4. **Diagnostic analytics**



### **Predictive Analytics**

Predictive analytics turn the data into valuable, actionable information. predictive analytics uses data to determine the probable outcome of an event or a likelihood of a situation occurring. Predictive analytics holds a variety of statistical techniques from modeling, [machine learning](https://www.geeksforgeeks.org/machine-learning/), [data mining](https://www.geeksforgeeks.org/data-mining/), and [game theory](https://www.geeksforgeeks.org/game-theory/) that analyze current and historical facts to make predictions about a future event. Techniques that are used for predictive analytics are:

* Linear Regression
* Time Series Analysis and Forecasting
* Data Mining

### **Descriptive Analytics**

Descriptive analytics looks at data and analyze past event for insight as to how to approach future events. It looks at past performance and understands the performance by mining historical data to understand the cause of success or failure in the past. Almost all management reporting such as sales, marketing, operations, and finance uses this type of analysis.

**Common examples of Descriptive analytics are company reports that provide historic reviews like:**

* Data Queries
* Reports
* Descriptive Statistics
* Data dashboard

### **Prescriptive Analytics**

Prescriptive Analytics automatically synthesize big data, mathematical science, business rule, and machine learning to make a prediction and then suggests a decision option to take advantage of the prediction.

For example, [Prescriptive Analytics](https://www.geeksforgeeks.org/what-is-prescriptive-analytics-in-data-science/) can benefit healthcare strategic planning by using analytics to leverage operational and usage data combined with data of external factors such as economic data, population demography, etc.

### **Diagnostic Analytics**

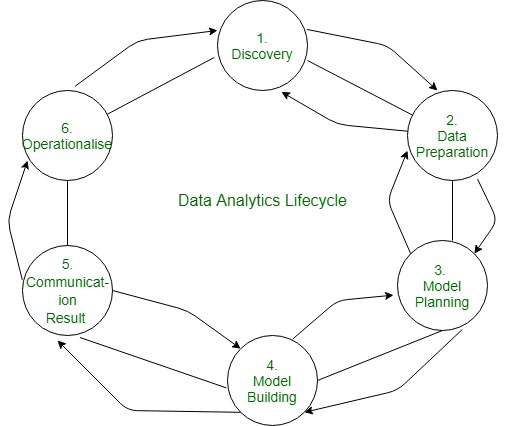
In this analysis, we generally use historical data over other data to answer any question or for the solution of any problem. We try to find any dependency and pattern in the historical data of the particular problem

**For example**, companies go for this analysis because it gives a great insight into a problem, and they also keep detailed information about their disposal otherwise data collection may turn out individual for every problem and it will be very time-consuming.  Common techniques used for Diagnostic Analytics are:

* Data discovery
* Data mining
* Correlations

**Life Cycle of Data Analytics :**  
The [Data analytic](https://www.geeksforgeeks.org/data-analytics-and-its-type/) lifecycle is designed for Big Data problems and data science projects. The cycle is iterative to represent real project. To address the distinct requirements for performing analysis on Big Data, step–by–step methodology is needed to organize the activities and tasks involved with acquiring, processing, analyzing, and repurposing data.

* **Phase 1: Discovery –**
* The data science team learns and investigates the problem.
* Develop context and understanding.
* Come to know about data sources needed and available for the project.
* The team formulates the initial hypothesis that can be later tested with data.
* **Phase 2: Data Preparation –**
* Steps to explore, preprocess, and condition data before modeling and analysis.
* It requires the presence of an analytic sandbox, the team executes, loads, and transforms, to get data into the sandbox.
* Data preparation tasks are likely to be performed multiple times and not in predefined order.
* Several tools commonly used for this phase are – Hadoop, Alpine Miner, Open Refine, etc.
* **Phase 3: Model Planning –**
* The team explores data to learn about relationships between variables and subsequently, selects key variables and the most suitable models.
* In this phase, the data science team develops data sets for training, testing, and production purposes.
* Team builds and executes models based on the work done in the model planning phase.
* Several tools commonly used for this phase are – Matlab and STASTICA.
* **Phase 4: Model Building –**
* Team develops datasets for testing, training, and production purposes.
* Team also considers whether its existing tools will suffice for running the models or if they need more robust environment for executing models.
* Free or open-source tools – Rand PL/R, Octave, WEKA.
* Commercial tools – Matlab and STASTICA.
* **Phase 5: Communication Results –**
* After executing model team need to compare outcomes of modeling to criteria established for success and failure.
* Team considers how best to articulate findings and outcomes to various team members and stakeholders, taking into account warning, assumptions.
* Team should identify key findings, quantify business value, and develop narrative to summarize and convey findings to stakeholders.
* **Phase 6: Operationalize –**
* The team communicates benefits of project more broadly and sets up pilot project to deploy work in controlled way before broadening the work to full enterprise of users.
* This approach enables team to learn about performance and related constraints of the model in production environment on small scale which make adjustments before full deployment.
* The team delivers final reports, briefings, codes.
* Free or open source tools – Octave, WEKA, SQL, MADlib.



**Roadmap of Data analytics :**

