Exploratory Data Analysis(EDA)

Definition

Exploratory Data Analysis (EDA) can be derived as a process of exploring or analyzing the data in different ways to derive insights from the data.

EDA plays a crucial role in data analytics and data science.

Once the EDA is performed on the dataset, it is used for statistical modelling and machine learning.

EDA is helpful to data scientists to find trends, patterns or hypothesis testing.

Basic components of EDA

Descriptive Statistics

Data Cleaning

Data Transformation

Relationship between variables

Data Visualization

Understanding Data

To understand data few functions are performed

Head

Tail

Shape

Describe

Info

Descriptive statistics

Mean

Median

Count

Variance

Standard Deviation

IQR

Mode

Range

Percentile

Data Cleaning

Handling Missing Values

Handling Null Values

Handling Duplicated

Deleting unnecessary column

Changing Columns format

Data Transformation

Normalization

Standardization

Categorical Encoding

Feature Engineering

Imputations

Relationship between variables

Univariate Analysis

Bivariate Analysis

Multivariate Analysis

Data Visualization

The visual display of the plots and charts of the data derives insights and complex data relationships which can be easily understood to non-technical clients also.

EDA can be performed using

Python

R

SQL

Excel

Tableau

Power BI

Thank You