CLASS 18 15-06-2021

**QUESTIONS**

👉 What is Exploratory Data Analysis (EDA) ?  
👉 Why do we need exploratory data analysis ( EDA )?  
👉 What are the steps in exploratory data analysis ( EDA )?  
👉 What Is Bivariate Analysis?  
👉 What Is Univariate Analysis?  
👉 What is box plots in seaborn?

**ANSWERS**

1. **Exploratory data analysis (EDA)** is used by data scientists to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods. It can also help determine if the statistical techniques you are considering for data analysis are appropriate.

2. The main **purpose of EDA** is to help look at data before making any assumptions. It can help identify obvious errors, as well as better understand patterns within the data, detect outliers or anomalous events, find interesting relations among the variables.

3. **Steps in Data Exploration and Preprocessing:**

1. Identification of variables and data types.
2. Analyzing the basic metrics.
3. Non-Graphical Univariate Analysis.
4. Graphical Univariate Analysis.
5. Bivariate Analysis.
6. Variable transformations.
7. Missing value treatment.
8. Outlier treatment.

4. **Bivariate analysis:-** is performed to find the relationship between each variable in the dataset and the target variable of interest (or) using 2 variables and finding the relationship between them.

5. **Univariate analysis** is the technique of comparing and analyzing the dependency of a single predictor and a response variable. The prefix "uni" means one, emphasizing the fact that the analysis only accounts for one variable's effect on a dependent variable.

6. **Boxplots** are a measure of how well distributed the data in a data set is. It divides the data set into three quartiles. This graph represents the minimum, maximum, median, first quartile and third quartile in the data set.