Unit I
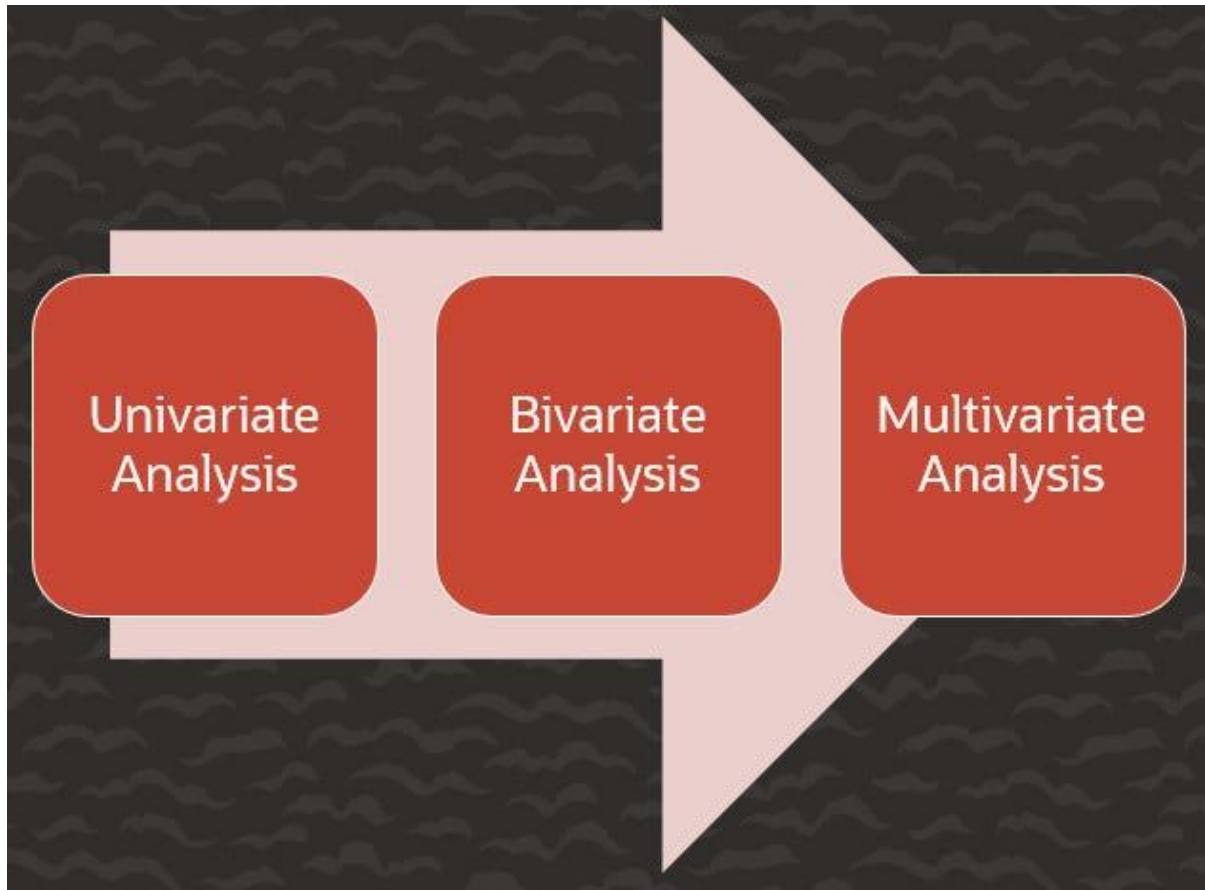
Descriptive Univariate Analysis:

Descriptive univariate analysis involves examining and summarizing the characteristics of a single variable in a dataset. Common measures used in descriptive univariate analysis include:

1. Measures of Central Tendency: These statistics indicate the center or average of the data.

   - Mean: The sum of all values divided by the number of data points.

   - Median: The middle value of the data when arranged in ascending or descending order.

   - Mode: The most frequently occurring value in the data.

2. Measures of Dispersion: These statistics quantify the spread or variability of the data.

   - Range: The difference between the maximum and minimum values in the data.

   - Variance: The average of the squared differences between each data point and the mean.

   - Standard Deviation: The square root of the variance.

3. Frequency Distribution: A table or chart that shows the number of occurrences of each value in a dataset.

Data Visualization in Descriptive Univariate Analysis:

For data visualization in descriptive univariate analysis, the choice of diagrams and plots depends on the scale type of the variable:

- For nominal and ordinal data: Bar charts, pie charts, and stacked bar charts are commonly used to show the distribution of categories.

- For interval and ratio data: Histograms, box plots, line charts, and scatter plots are often employed to visualize the distribution and characteristics of the numerical data.

These visualizations help data analysts gain insights into the data's distribution, identify patterns, and detect potential outliers or anomalies, enabling them to make informed decisions and draw meaningful conclusions.