Descriptive Statistics

Descriptive statistics are a collection of techniques used to summarize and describe the key features of a data set. They provide a high-level understanding of the data without going into complex analysis.

Quantitative and qualitative data are two fundamental categories used to classify data based on its characteristics

Quantitative Data:

- **Numerical:** Quantitative data is numerical in nature. It represents information that can be counted or measured and expressed in numbers.

- **Measurable:** You can quantify this data using units like meters, kilograms, seconds, percentages, etc.

- **Examples:** Examples of quantitative data include height, weight, temperature, income, number of customers, exam scores, etc.

Qualitative Data:

- **Descriptive:** Qualitative data is descriptive and non-numerical. It focuses on qualities, characteristics, or categories.

- **Non-Measurable:** You cannot quantify it with units.

- **Examples:** Examples of qualitative data include customer satisfaction ratings (excellent, good, fair, poor), colors (red, blue, green), product reviews (positive, negative, neutral), customer opinions, interview responses, etc.

Descriptive statistics can be broadly categorized into three main areas:

1. Measures of Central Tendency:

These statistics describe the "center" of the data set, indicating where most of the values tend to fall. Common measures include:

- **Mean:** The average of all the values in the data set.

- **Median:** The middle value when the data is ordered from least to greatest.

- **Mode:** The most frequent value in the data set.

2. Measures of Dispersion (Spread):

These statistics describe how spread out the data is from the central tendency. They indicate how much variation there is within the data set. Common measures include:
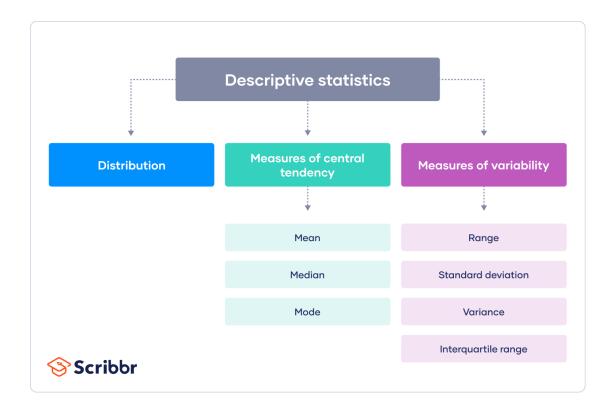
- **Range:** The difference between the highest and lowest values.

- **Variance:** The average squared deviation of each data point from the mean.

- **Standard Deviation:** The square root of the variance. It's measured in the same units as the original data and is generally considered a more interpretable measure of spread compared to variance.

3. Measures of Frequency Distribution:

These statistics describe how often each value (or range of values) appears in the data set. They help visualize the distribution of the data and identify patterns. Common measures include:

- **Frequency tables:** Tables that show the number of times each value or category appears in the data.

- **Histograms:** Bar charts that visualize the distribution of the data by dividing the data range into intervals and showing the frequency of data points within each interval.