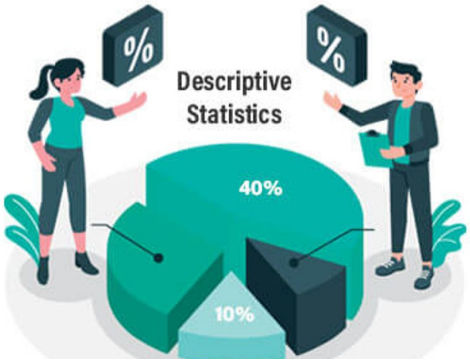
# Statistics-Task 3

By Menna Jaheen

**Statistics forms the core of data analytics, serving as the fundamental tool for identifying trends and patterns within vast numerical datasets. This mathematical discipline encompasses two main categories: Descriptive Statistics and Inferential Statistics.**

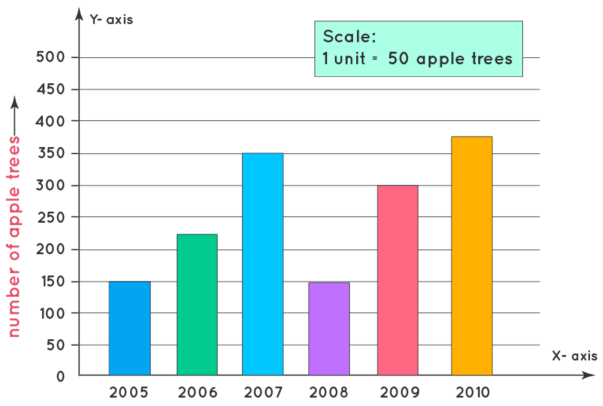
## Descriptive vs Inferential Statistics :

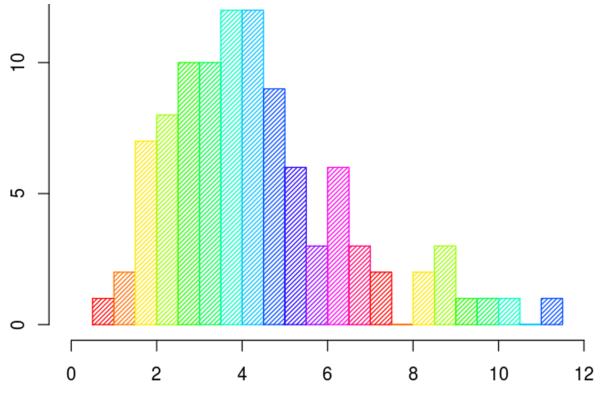




## Descriptive :

- Descriptive statistics is a branch of statistics that deals with summarizing and describing data using numbers and graphs.
- It provides methods for organizing, visualizing, and presenting data meaningfully and informally.
- Descriptive statistics describe the characteristics of the data set under study without generalizing beyond the analyzed data.



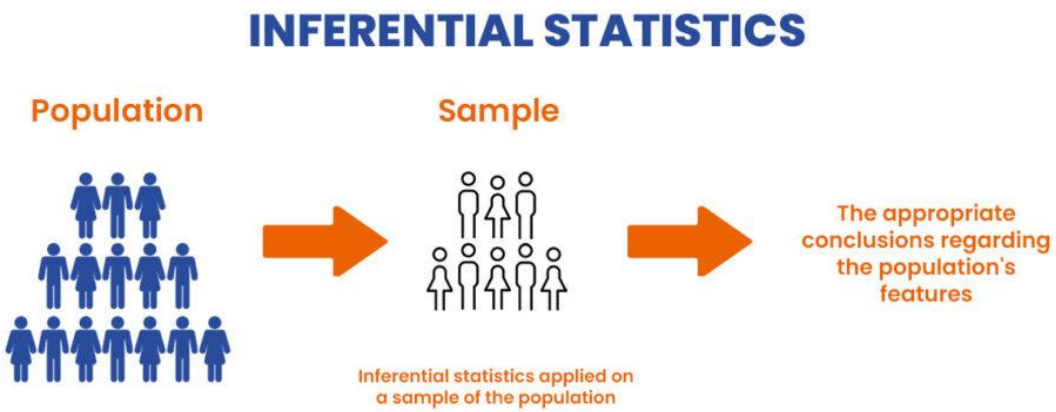

**Common measures and techniques in descriptive statistics :**

1. measures of **central tendency** (such as mean, median, and mode)
2. measures of **dispersion** (such as range, variance, and standard deviation)
3. **frequency distributions** (histograms, frequency tables), and graphical representations (box plots, bar charts, pie charts, etc.)

---

## Inferential :

- Involves making **inferences, predictions**, or generalizations about a larger population bas ed on data collected from a sample of that population.
- Extends the findings from a sample to the population from which the sample was drawn.
- Allow researchers to draw **conclusions, test hypotheses,** and make predictions about populations, even when it is impractical or impossible to study the entire population directly.




**Key methods in inferential statistics :**

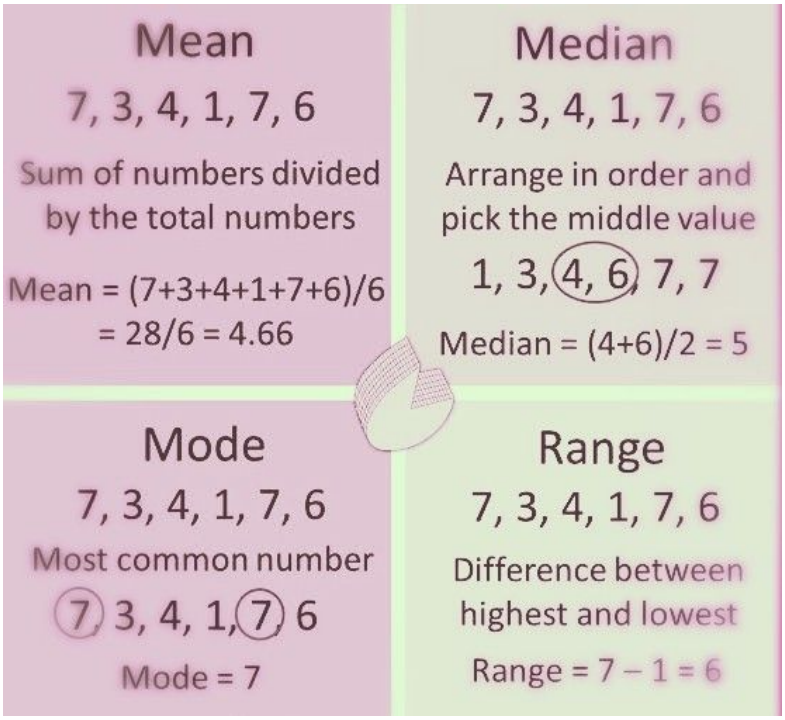
1. hypothesis testing, where researchers test hypotheses about population parameters using sample data
2. Regression analysis, where relationships between variables are examined and used to make predictions
3. Confidence intervals, which provide estimates of population parameters and their uncertainty levels.

---

# Mean, Median, Mode, and Range - How To Find It!

- **Mean :** an average which is found by adding up all the values in a set of data and dividing it by the total number of values you added together.

- **Median** : the middle number in the set of values. You find it by putting the numbers in order from the smallest to largest and covering up one number on each end until you get to the middle.
- **Mode** : the number or value, which appears most often in the set. To find the mode, you need to count how many times each value appears.
- **Range**: the difference between the lowest and the highest value. To work it out, simply subtract the lowest value from the highest.



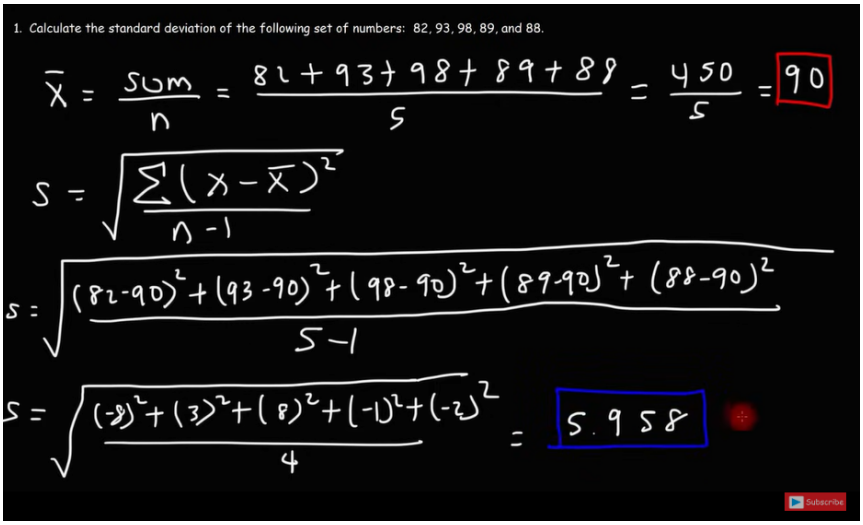

---

## How To Calculate The Standard Deviation :

- A standard deviation (or σ) is **a measure of how dispersed the data is in relation to the mean**. Low, or small
- Standard deviation indicates data are clustered tightly around the mean, and high, or large, standard deviation indicates data are more spread out.

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

**Example :**




---

## How To Calculate Variance :

- Variance is **a statistical measurement of the spread between numbers in a data set**.
- It measures how far each number in the set is from the mean (average), and thus from every other number in the set.
- Variance is often depicted by this symbol: σ2

$$\sigma^2 = \frac{\sum(xi - \bar{x})^2}{N}$$

**Example :**