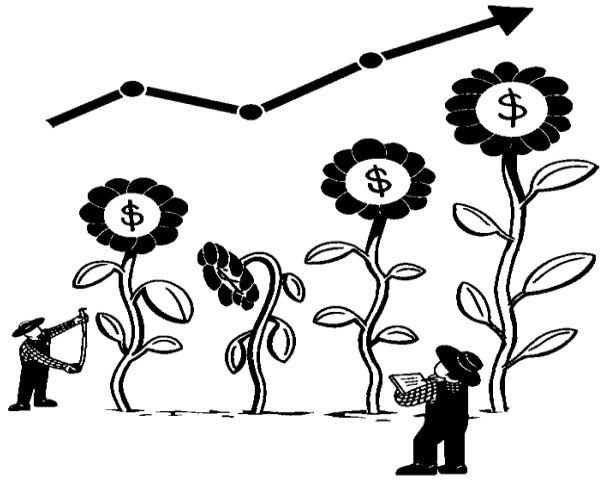
# DATA ANALYTICS BOOTCAMP

Uloh C. Armstrong

## ❑ MEANING AND SCOPE OF STATISTICS.

**Statistics** is a branch of mathematics and a scientific discipline concerned with collecting, organizing, analyzing, interpreting, and presenting data.

It involves methods and techniques used to gather, summarize, and draw conclusions or inferences from data.



**Statistics**

*[stə-'ti-stiks]*

A branch of applied mathematics that involves the collection, description, analysis, and inference of conclusions from quantitative data.

- Basic **STATISTICS**

- **TYPES OF STATISTICS.**

**Descriptive statistics:**
Describes and summarizes data through measures such as mean, median, mode, etc.

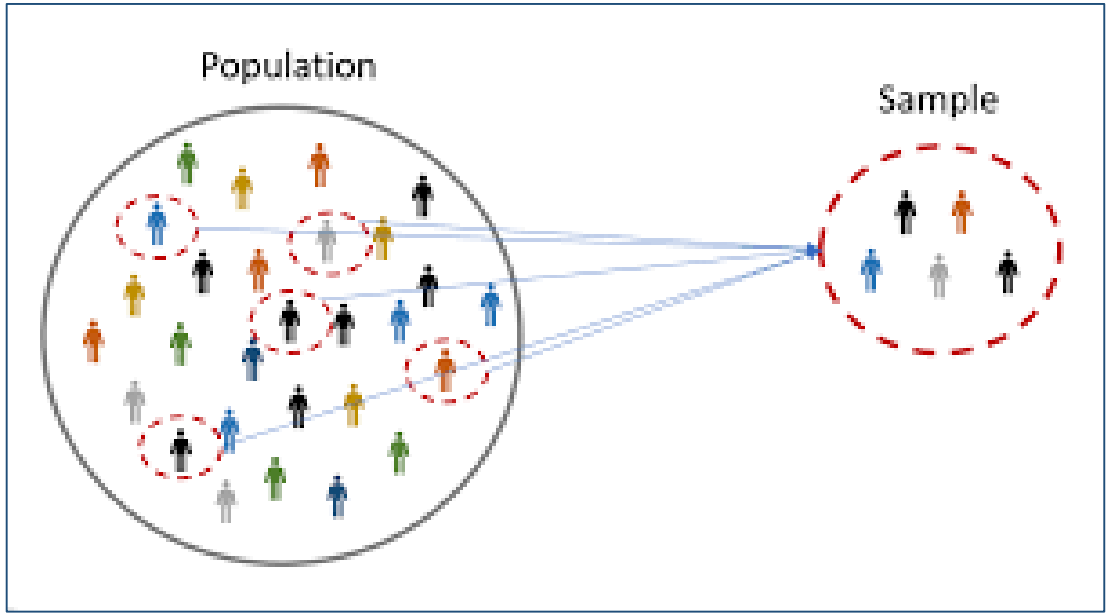
**Inferential statistics:**
Uses sample data to make inferences or predictions about a larger population

❑ Basic **STATISTICS**

## ❑ POPULATION AND SAMPLE

**Population:**
This is a set of all objects or units about which conclusions are to be drawn.

**Sample:**
A subset of a population.

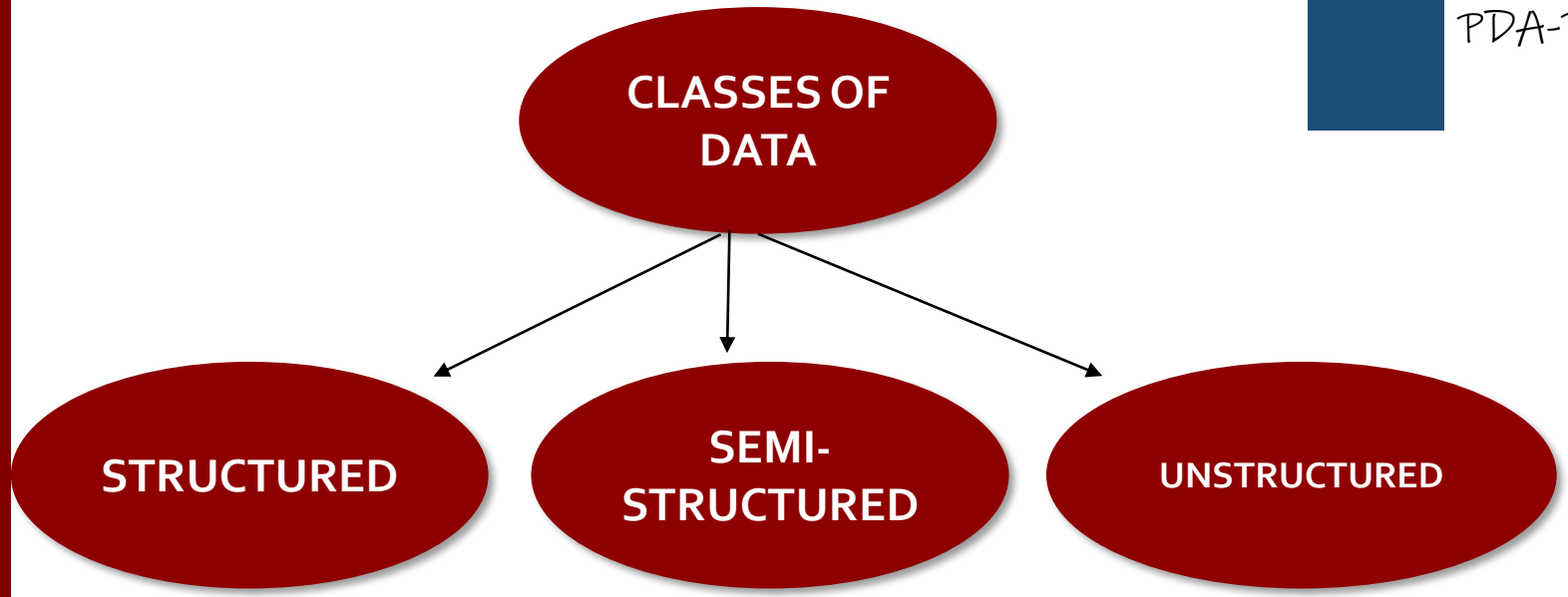


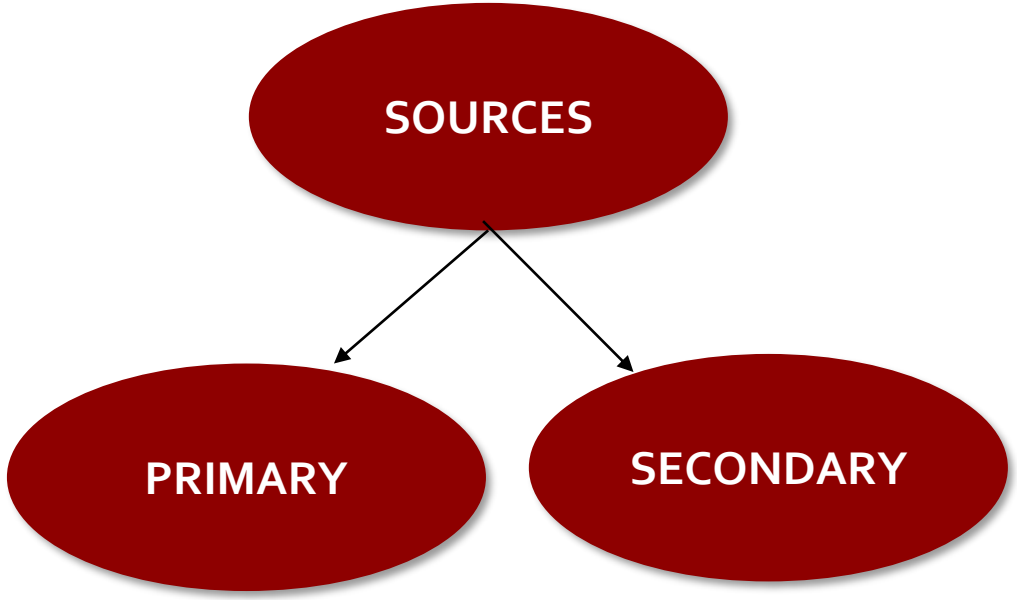
- Basic **STATISTICS**

- **STATISTICAL DATA.**
- **TYPES OF DATA.**

- Basic **STATISTICS**

## ❑ CLASSES AND SOURCES OF DATA.



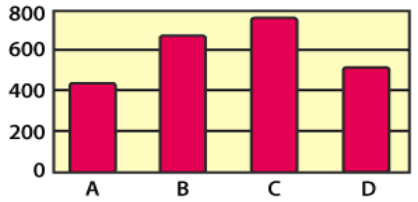
- Basic **STATISTICS**

# ❑ GRAPHICAL PRESENTATION OF DATA.

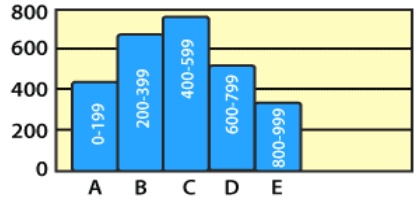
## TYPES OF GRAPHICAL REPRESENTATION

BYJU'S The Learning App

### Bar Graphs



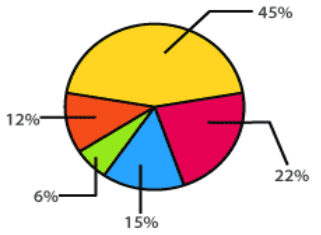

### Histograms



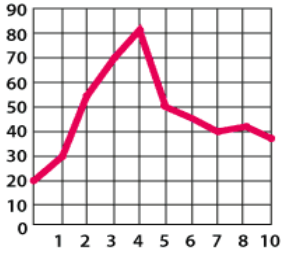

### Frequency Table

| Rulers of France | | |
|---|---|---|
| Reign (Years) | Tally | Frequency |
| 1-15 | 𝍸 𝍸 𝍸 III | 18 |
| 16-30 | 𝍸 𝍸 I | 11 |
| 31-45 | 𝍸 I | 6 |
| 46-60 | IIII | 4 |
| 61-75 | I | 1 |

### Circle Graph




### Line Graphs

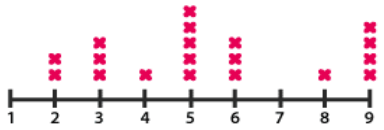



### Stem and Leaf Plot

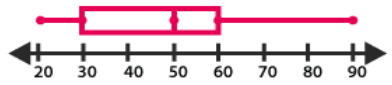
| Stem | Leaf |
|---|---|
| 0 | 1, 1, 2, 2, 3, 4, 4, 4, 4, 5, 8 |
| 1 | 0, 0, 0, 1, 1, 3, 7, 9 |
| 2 | 5, 5, 7, 7, 8, 8, 9, 9 |
| 3 | 0, 1, 1, 1, 2, 2, 2, 4, 5 |
| 4 | 0, 4, 8, 9 |
| 5 | 2, 6, 7, 7, 8 |
| 6 | 3, 6 |

Key : 6 | 3 = 63 Year

### Line Plot




### Box and Whisker Plot




© Byjus.com

- Basic **STATISTICS**

- **ARITHMETIC MEAN.**

The **arithmetic mean**, commonly known as the **average**, is a measure of central tendency used to represent a set of numerical data

Arithmetic Mean = x1 + x2 + x3 +..... + xn / n

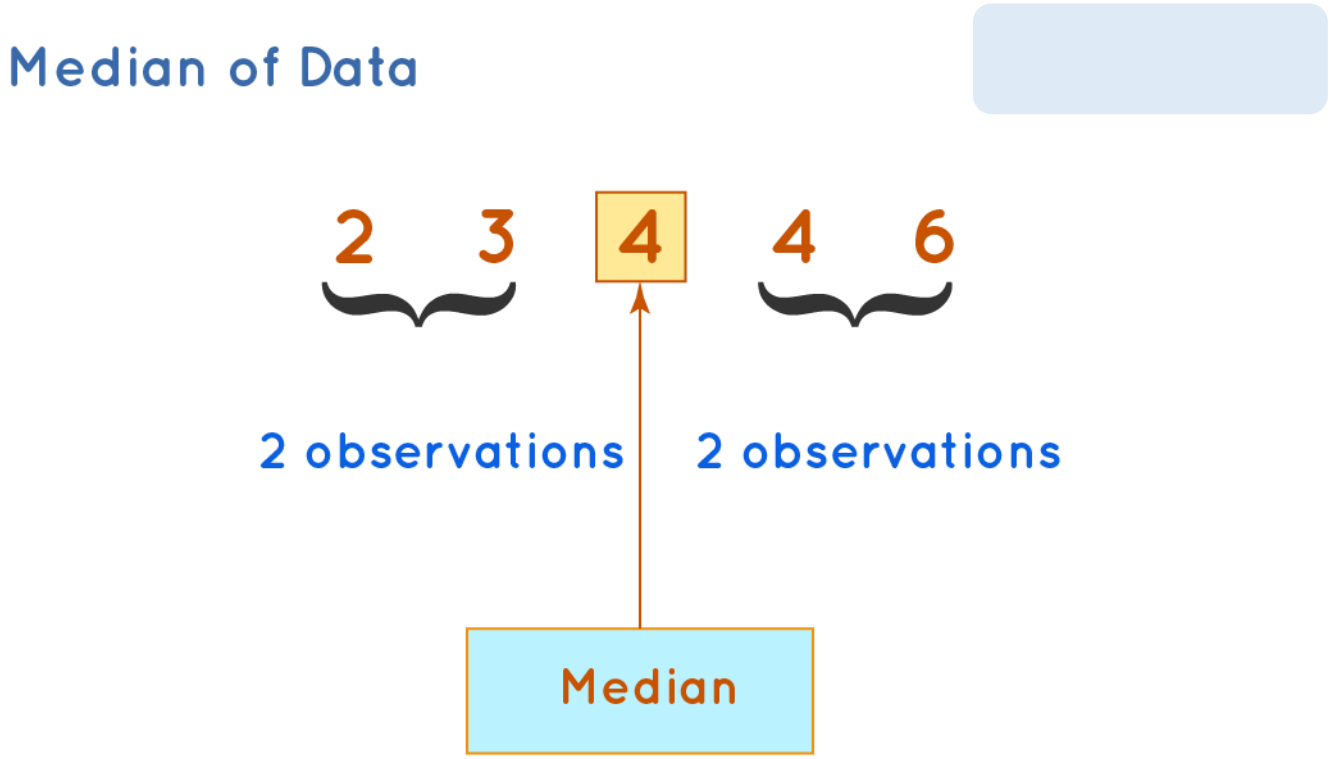
=(B3+B4+B5+B6+B7+B8+B9)/B11

| | A | B | D |
|---|---|---|---|
| 2 | Sr.No | Division A | |
| 3 | Student 1 | 56 | |
| 4 | Student 2 | 60 | |
| 5 | Student 3 | 56 | |
| 6 | Student 4 | 64 | |
| 7 | Student 5 | 70 | |
| 8 | Student 6 | 55 | |
| 9 | Student 7 | 50 | |
| 10 | | | |
| 11 | Number of Students (n) | 7 | |
| 12 | Arithmetic Mean | =(B3+B4+B5+B6+B7+B8+B9)/B11 | |
| 13 | | | |

- Basic **STATISTICS**

- **MEDIAN.**

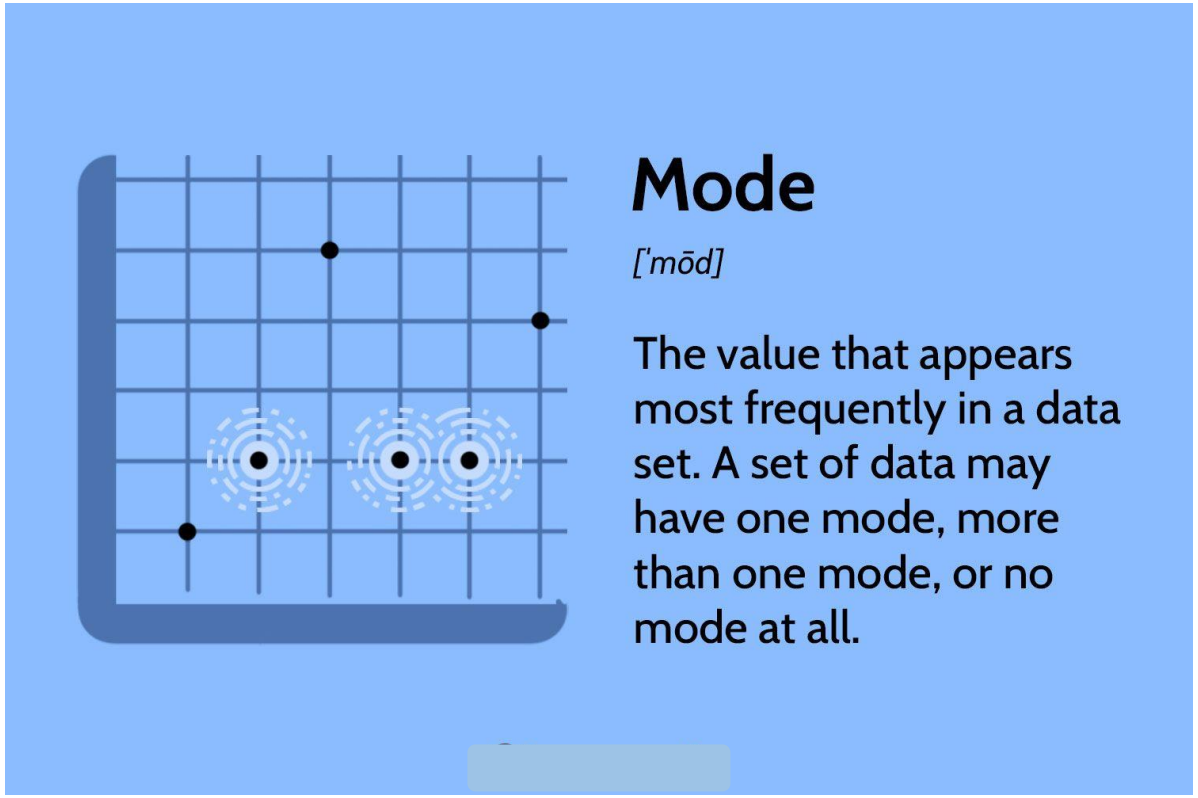
The **median** represents the middle value in a dataset when arranged in ascending or descending order.



- Basic **STATISTICS**

- **MODE.**

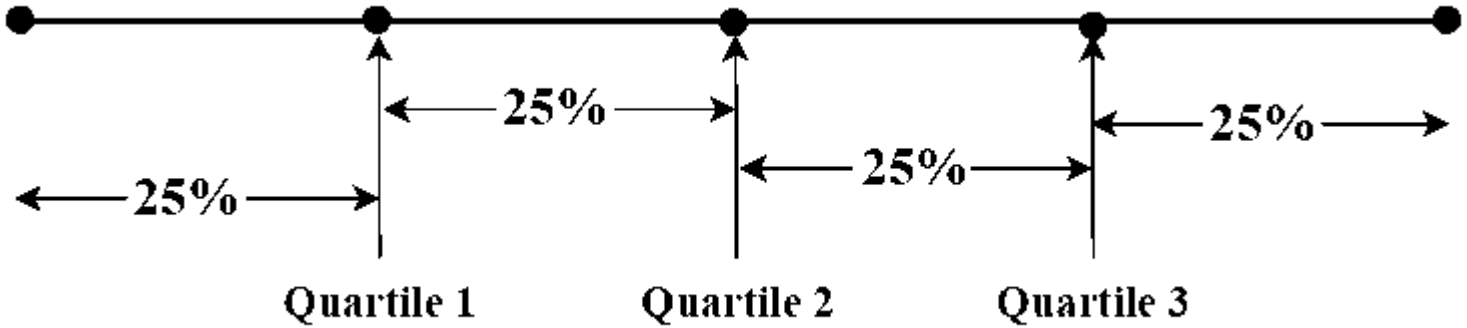
The **mode** is the value/values within a dataset that occur most frequently.



- Basic **STATISTICS**

## ❑ QUARTILES.

**Quartiles** are values that divide a dataset into four equal parts, each representing 25% of the data when arranged in ascending or descending order.

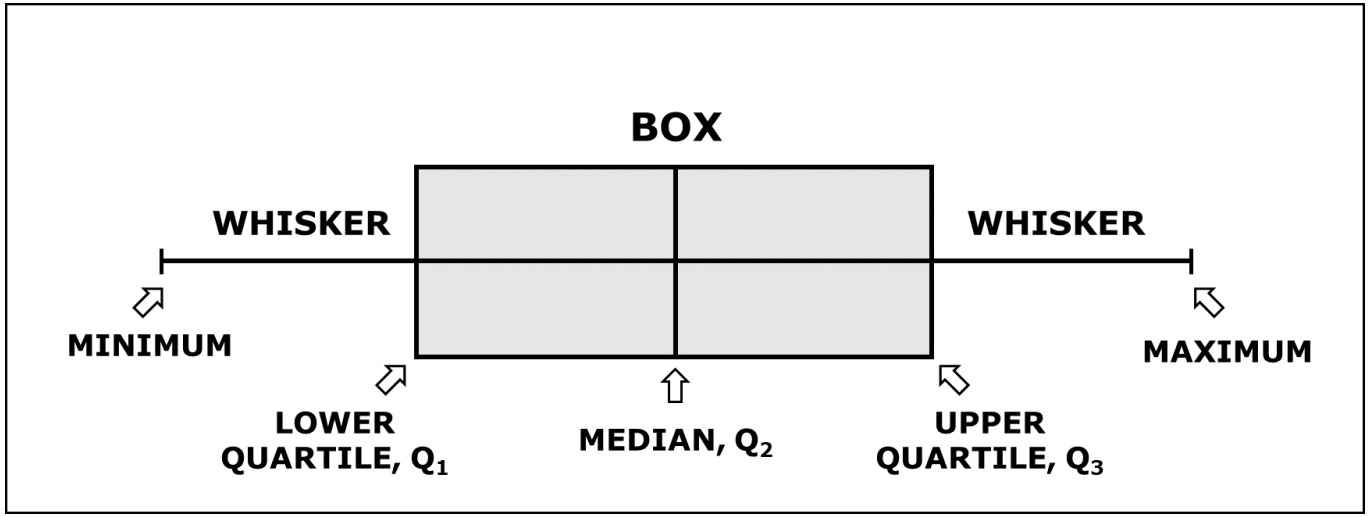


- Basic **STATISTICS**

## ❑ BOX & WHISKERS PLOT (BOX PLOT).

A **box plot**, also known as a **box-and-whisker plot**, is a graphical representation that displays the distribution, central tendency, and variability of numerical data through five summary statistics:

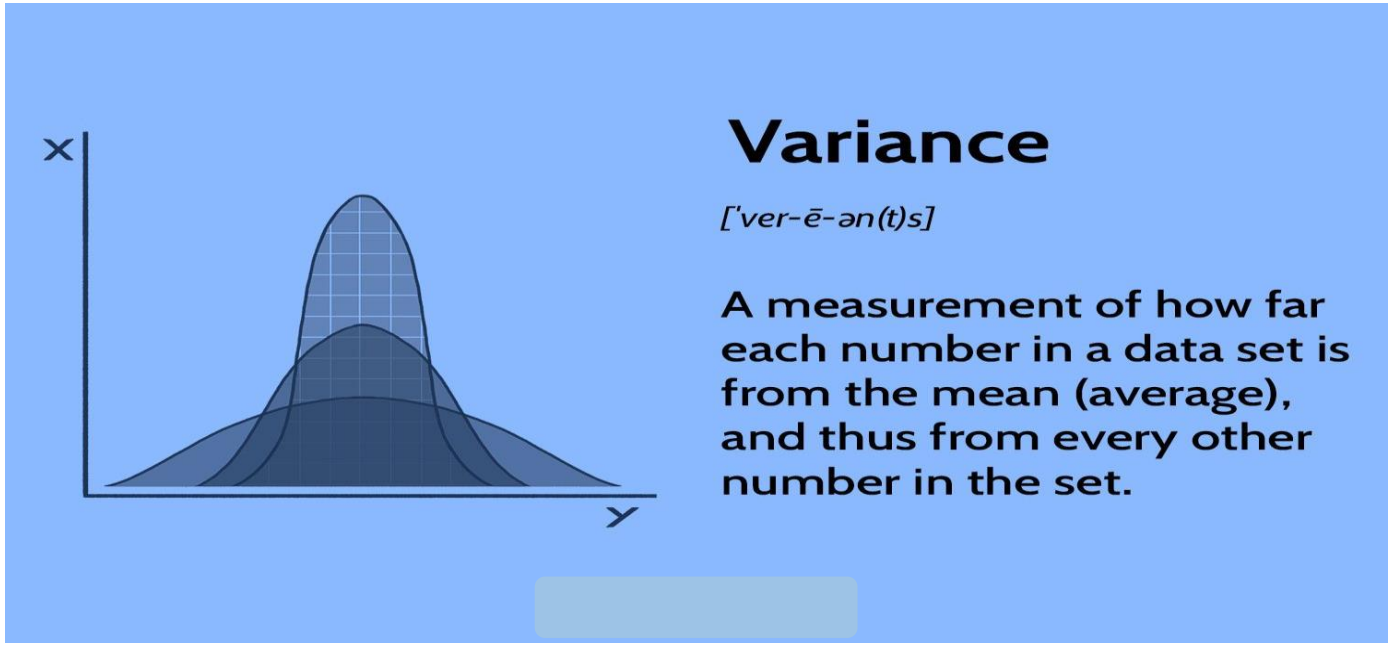
*minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum*



- Basic **STATISTICS**

- **THE VARIANCE.**

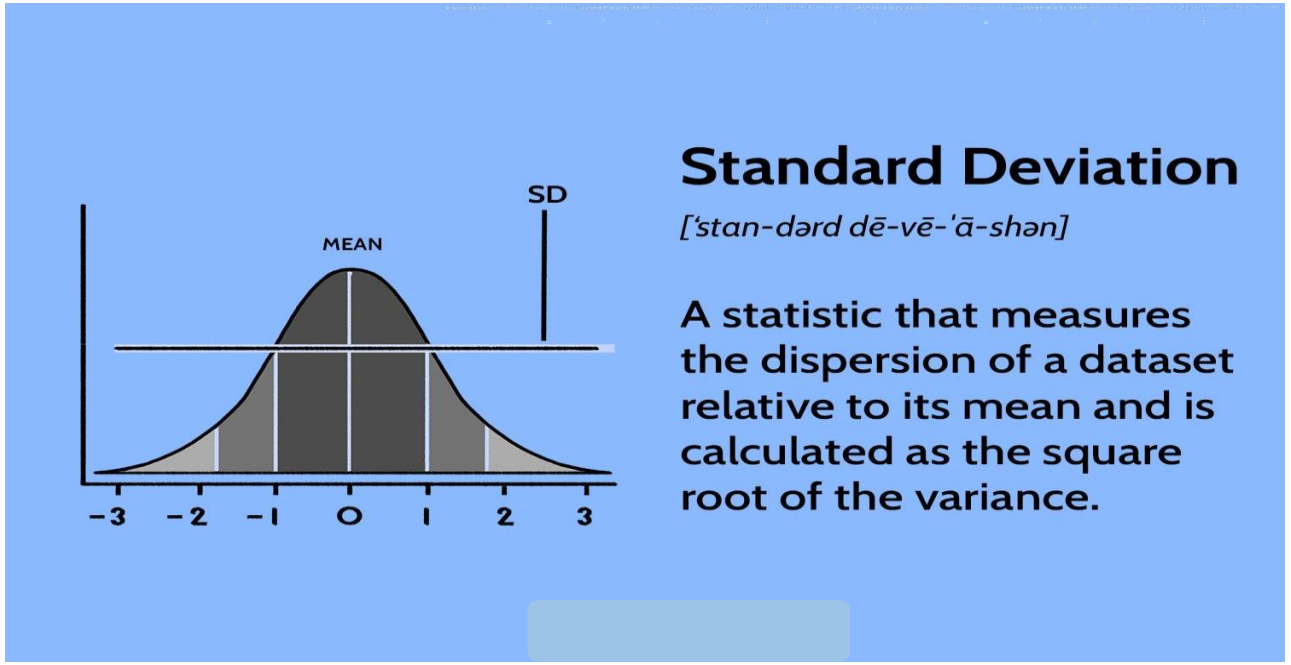
The **Variance** is a measure of how spread out the values in a data set are. It is closely related to the standard deviation, but it is expressed in a different unit of measurement.



- Basic **STATISTICS**

- **STANDARD DEVIATION.**

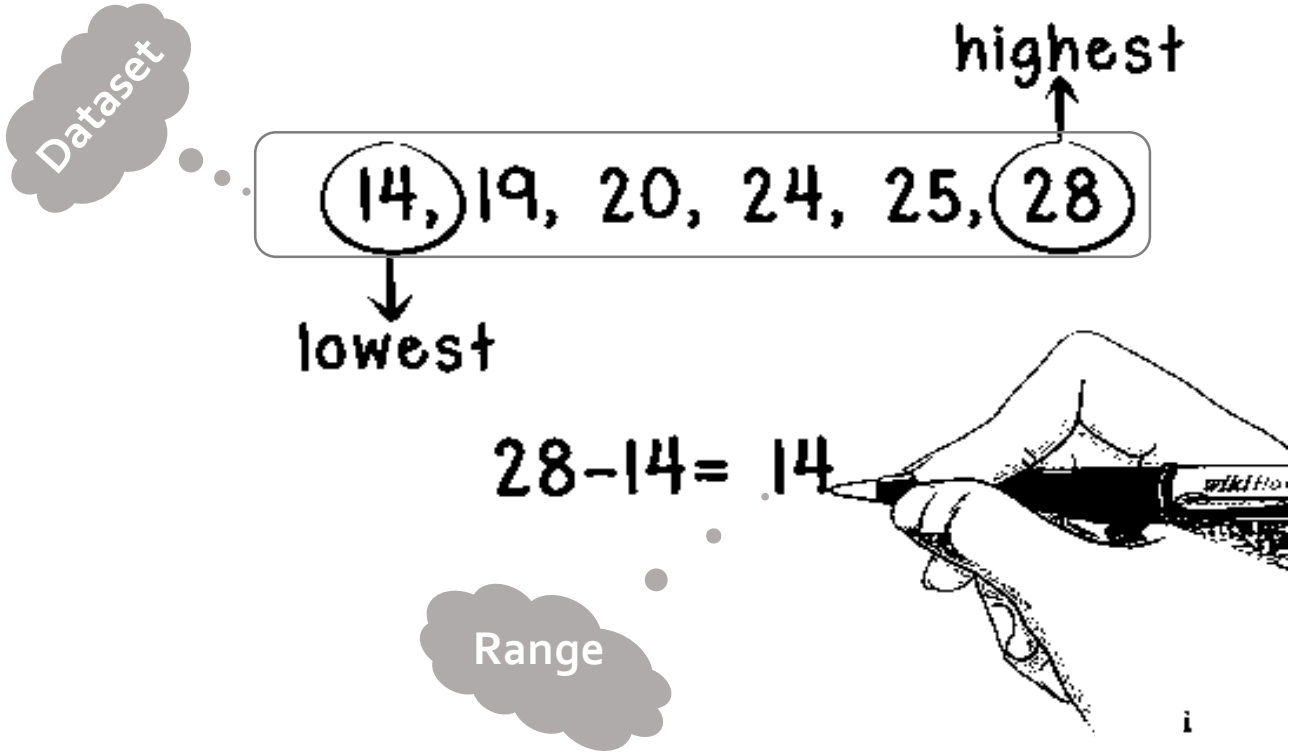
The **Standard deviation** is a measure of how spread out the values in a data set are.
It is a way of quantifying the amount of variation in a set of data



- Basic **STATISTICS**

- **RANGE.**

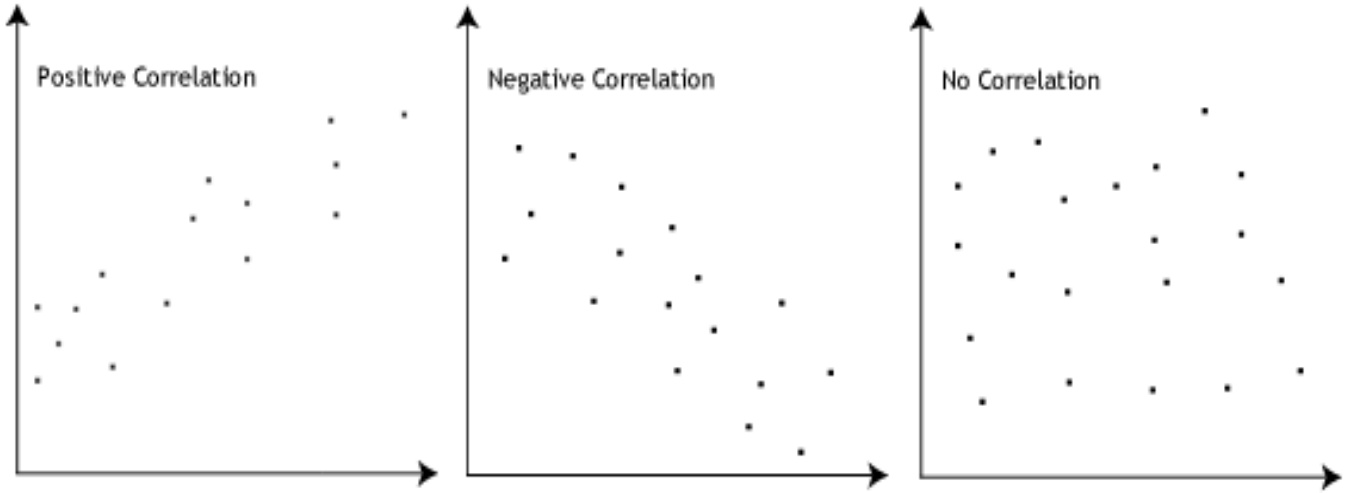
The **range** in statistics refers to the difference between the highest and lowest values within a dataset.



- Basic **STATISTICS**

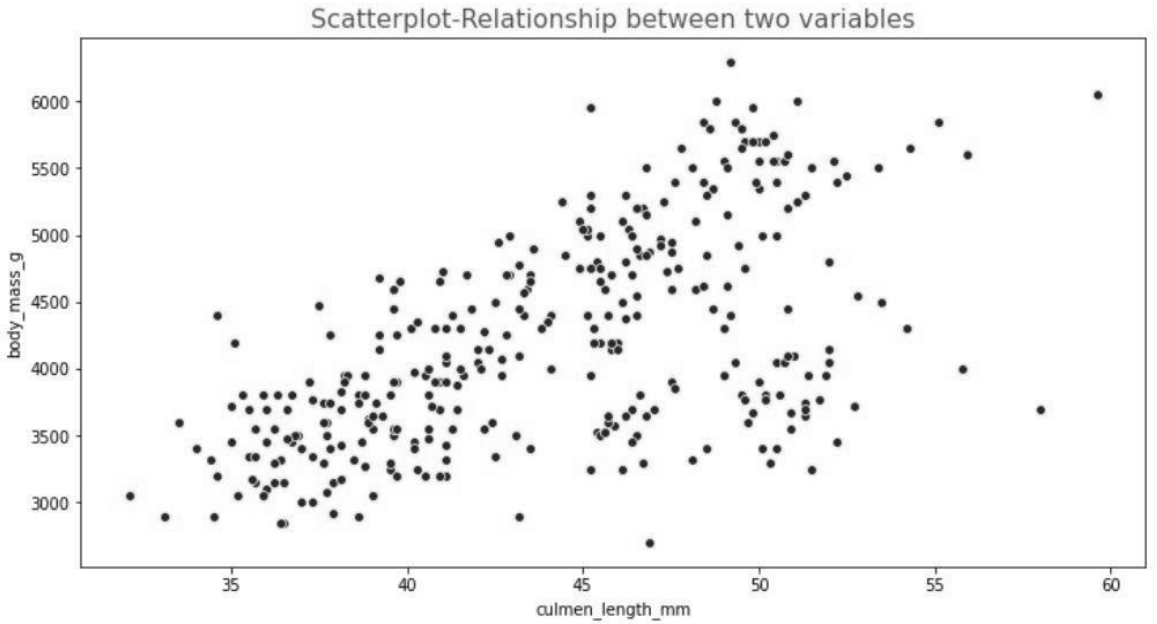
- **CORRELATION ANALYSIS.**

**Correlation analysis** is a statistical technique used to measure and describe the strength and direction of the relationship between two variables.



- Basic **STATISTICS**
- Meaning and scope of statistics.
- Types of statistics.
- Statistical data.
- Types of data.
- Classes and sources of data.
- Graphical presentation of data.
- Arithmetic mean.
- Median.
- Mode.
- Quartiles.
- Box & Whiskers plot (Box plot).
- The Variance
- Standard deviation.
- Range.
- Correlation analysis.
- **SCATTER DIAGRAM.**
- Regression Analysis

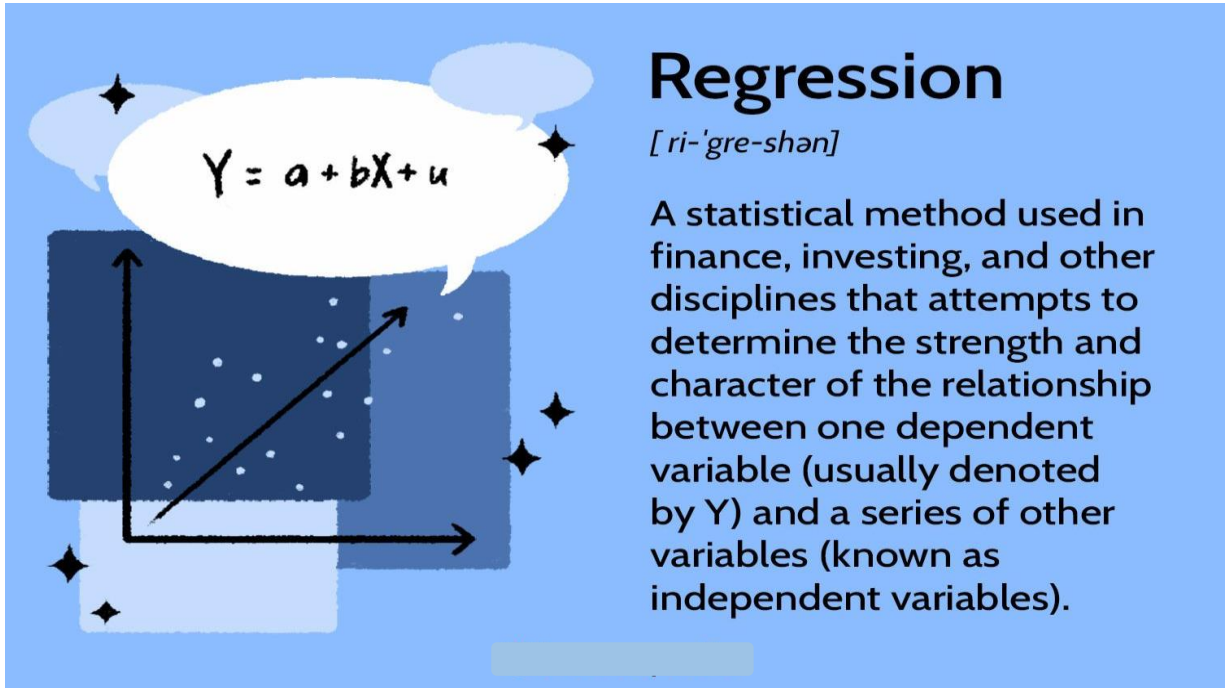
A **scatter plot**, also known as a **scatter diagram** or **scatter graph**, is a type of data visualization that displays the relationship or association between two numerical variables.
It helps to visualize how changes in one variable are related to changes in another variable.



- Basic **STATISTICS**

- **Regression Analysis**

**Regresssion analysis** is a statistical method used to model relationship between two or more variables. It is a way of determining how a change in one variable is associated with a change in another variable.



# DATA ANALYTICS BOOTCAMP

Uloh C. Armstrong