Statistics

1 — Statistics
Branch of Mathematics deals with Data Collection, Organization, Analysis, Interpretation and Presentation

2 — Elements

- **1 Population** — Overall Data
  - **Example:** Population of India: 136.64 crores
- **2 Sample** — Certain Part/Group of Data
  - **Example:** Population of Delhi: 1.9 crores
- **3 Variables**
  - Quantitative Data — Age, Height, Weight
  - Categorical Data — Gender, State

3 — Types

- **Descriptive Statistics** — Describing Collected Data
  - **Example:** Data analysis done for, Total Covid test conducted for 136.64 Crore People or Certain Population
  - Includes
    - Collection
    - Organization
    - Classification
    - Summarization
    - Analysis
    - Interpretation
    - Presentation
  - Results
    - Charts
    - Graphs
    - Tables
- **Inferential Statistics** — Drawing conclusion based on data collected from sample of individuals from the whole data
  - **Example:** Data analysis done for sample test conducted upto now, 44 Crores gonna be matched to the overall 136.64 crores of Population
  - Includes
    - Comparision
    - Estimation
    - Relationship
    - Hypothesis testing
    - Predictions
  - Result — Probability Scores

4 — Descriptive Statistics

- **1 Measures of Center**

  Example : 1,2,3,4,4,5,2,4,10,9,12,8

Mean	Average of all values in sample	Mean = 1+2+3+4+4+5+2+4+10+9+12+8 / 12 = 64/12 = 5.3
Median	Measure Central value of sample set	Median = 5+2 / 2 = 3.5
Mode	Most recurrent value in sample set	Mode = 4

- **2 Measures of Spread**

  Example: 1 2 3 4 5 2 10 9 12 8

  - **Range** — Measure of How values are spread — Range = Max(data) - Min(data) = 12-1 = 11
  - **Inter Quartile Range** — Measure of Variability - Dividing into Quartiles
    - Steps
      - **1** Ascending Order — 1 2 3 4 5 8 9 10 12
      - **2**
        - Median — **5 is center** — 1 2 3 4 5 8 9 10 12
        - Quratile — **3 Quartile** — 1 2.5 4 | 5 | 8 9.5 12
      - IQR — 3rd Quartile - 1st Quartile — = 9.5 - 2.5 = 7
  - **Variance** — How much Current variable differs from Expected Value - Result in Large Numbers

    $$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$
    - σ^2 = population variance
    - Σ = sum of...
    - X = each value
    - μ = population mean
    - N = number of values in the population

    Population variance

    $$s^2 = \frac{\sum (X - \bar{x})^2}{n-1}$$
    - s^2 = sample variance
    - Σ = sum of...
    - X = each value
    - $\bar{x}$ = sample mean
    - n = number of values in the sample

    Sample variance
  - **Standard Deviation** — Measure of Dispersion of set of data from its Mean — Equation

    $$\text{Standard Deviation} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

    where:
    x_i = Value of the ith point in the data set
    $\bar{x}$ = The mean value of the data set
    n = The number of data points in the data set

5 — Additional Concepts

- **1 Logarithm** — It is an Inverse of Exponent
  - Example
    - We Know — 5 (5 Time Increment for Every Month) Next Month - 5*5 = 25 After Two Month - 5*5*5 = 125
    - We need to Find — If we have 125 om 5 Time increment & started from 5, how much month have taken to get this 125?
      - log5 125 = 3 Years
  - log10 100 = log10 10 2 = 2log10 10 = 2(1) = 2
- **2 Mean Absolute Deviation**

  Maths 90
  Science 92
  CS 95

  Avg = 92.33

  Maths 90 - 92.33 = 2.33
  Science 92 - 92.33 = 0.33
  CS 95 - 92.33 = 2.67

  Mean Absolute Deviation = 5.33