Probability & statistics

- Statistics - Deals with collection , analysis , interpretation and presentation of scientific data.
- Two types of statistics - (1) Descriptive -> Describes the data
                                          (2) Inferential -> To make predictions from the data

Basis for comparison	Descriptive statistics	Inferential statistics
Meaning	Descriptive Statistics is that branch of statistics which is concerned with describing the population under study.	Inferential Statistics is a type of statistics, that focuses on drawing conclusions about the population, on the basis of sample analysis and observation.
What it does?	Organize, analyze and present data in a meaningful way.	Compares, test and predicts data.
Form of final Result	Charts, Graphs and Tables	Probability
Usage	To describe a situation.	To explain the chances of occurrence of an event.
Function	It explains the data, which is already known, to summarize sample.	It attempts to reach the conclusion to learn about the population, that extends beyond the data available.

- Population < Sample
- Parameter -> Measure that describes whole population
- Statistic -> Measure that describes the sample
- Sampling-error -> Difference between a population parameter and a sample space
- Experiment -> Any process that generates data
  - Designed experiment
  - Observational experiment
  - Retrospective experiment

- Sampling Techniques -:
  - Probability Sampling ->
    - Simple random sampling
    - Cluster sampling
    - Systematic sampling
    - Stratified random sampling
  - Non- Probability sampling

- Measures of location describe the central tendency of the data ,includes mean, median and mode

Mean = Sum of favourable outcomes/No. of favourable outcomes

Trimmed mean –> k% trimmed mean is calculated by eliminating k% largest & smallest values and then calculate the mean.

Median = Middle point of the ordered data

Mode = Value of highest frequency element

- Measures of dispersion describes the spread of the data, includes range, interquartile range, standard deviation and variance.

Range = Smallest and largest observations

Interquartile range = Lower and upper quartiles
where Lower quartile is the median and the upper quartile is the dividing the data into 4 equal parts

Variance –> $S^2 = (E(x- m)^2)/ n-1$

Standard Deviation –> $S = \sqrt{(E(x- m)^2)/ N}$
           ((n-1) is the degree of freedom)

- Graphical representation of data -:
    - Bar graph
    - Pie chart
        - To calculate percentage -> (Frequency / Total frequency) * 100
        - To convert data into degrees -> (Given data / total value of data) * 360°
    - Line graph
    - Pictograph
    - Histogram
    - Stem and Leaf plot
    - Scatter plot
    - Point plot