

# DAY - 1

## Introduction to Statistics.

Two Types of Statistics.

- ① Descriptive Stats.
- ② Inferential Stats.

- ① Descriptive Stats.

→ ① Measure of Central Tendency.

→ ② Measure of Dispersion.

Anything that is used for Summarising the data.

This Includes:- Histograms, PDF, CDF, Probability, Permutation, mean, median, mode, Variance, Standard Deviation

Distribution that we cover.

- Gaussian Distribution.
- Log Normal Distribution.
- Binomial Distribution.
- Bernoulli's Distribution.
- Pareto Distribution. (Power law Distribution)
- Standard Normal Distributions.

## ② INFERENTIAL STATS.

- Z Test
- T Test
- ANOVA Test (F Test)
- CHI SQUARE Test.
- HYPOTHESIS Testing

You will learn how to see

- Z Table, T table, CHISQUARE Table.

What is Statistics?

- Statistics is the Science of collecting, organising and analyzing data.

What is Data?

- Data:- Facts or pieces of information that can be measured.

Types of Statistics:-

① Descriptive STATS:- It consists of organising and summarising of DATA.

② Inferential STATS:- It is a Technique where in we used the data that we have measured to form conclusion.

\* What are the Sampling Techniques.

① Simple Random Sampling. - Every member of the population ( $N$ ) has an equal chance of being selected for your sample ( $n$ )

② Stratified Sampling - It is a technique where the population ( $N$ ) is split into non-overlapping groups (strata)

③ Systematic Sampling :-

From population ( $N$ ) - we pick up every  $n^{\text{th}}$  Individual.

④ Convenience Sampling:- In this survey only the person who are expert in the domain, will participate in this survey.

⑤ Random Sampling:-

### - Variables:

A variable is a property that can take on any value.

Two Kinds of variables.

① Quantitative Variable - Measured Numerically.

② Qualitative / Categorical variables -

↳ Gender / Blood Group / T-shirt Size.

Quantitative Variable.

↓  
Discrete Variable

↓  
Whole number.

Eg:- No of Bank Acc.

No of Children in Family

↓  
Continuous Variable.

Eg Height :- 172.5, 162.5cm etc.

weight :- 100kg, 99.75kg etc.

Rain fall :- 1.12 inch

1.5 inch.

Examples:-

⑥ Gender - Categorical Variable.

⑦ Marital Status - Categorical Variable.

⑧ River length - Continuous Variable.

⑨ Population of State - Discrete Variable.

⑩ Song length - Continuous Variable.

⑪ Blood pressure - Continuous Variable.



## Variable Measurement Scales:-

4 Types of Measured Variable

- ① Nominal    ② Ordinal    ③ Interval    ④ Ratio.

①

Nominal Data - (Categorical data) - Different classes.

Eg:- Color, Genders, Type of Flower.

②

Ordinal Data :- Order of the data matters, value does not

Eg:- Students marks      Rank.

100

96

57

85

44

1
2
4
3
5

→ This is called ordinal Data.

③

Interval Data - Order matters, values also matters,  
natural zero is not present.

Eg:- Fahrenheit :- 70-80, 80-90, 90-100

0 Fahrenheit has no meaning.

④ Ratio Data :-



## Frequency Distribution:-

Sample Dataset:- Rose, Lilly, Sunflower, Rose, Lilly, Sunflower,  
Rose, Lilly, Lilly.

Flower	Frequency	Cumulative Frequency
Rose	3	3
Lilly	4	7 (3+4)
Sunflower	2	9 (7+2)

### NOTE:-

If the Variable is discrete, we can draw a bar chart, if the Variable is continuous then we can draw a Histogram.

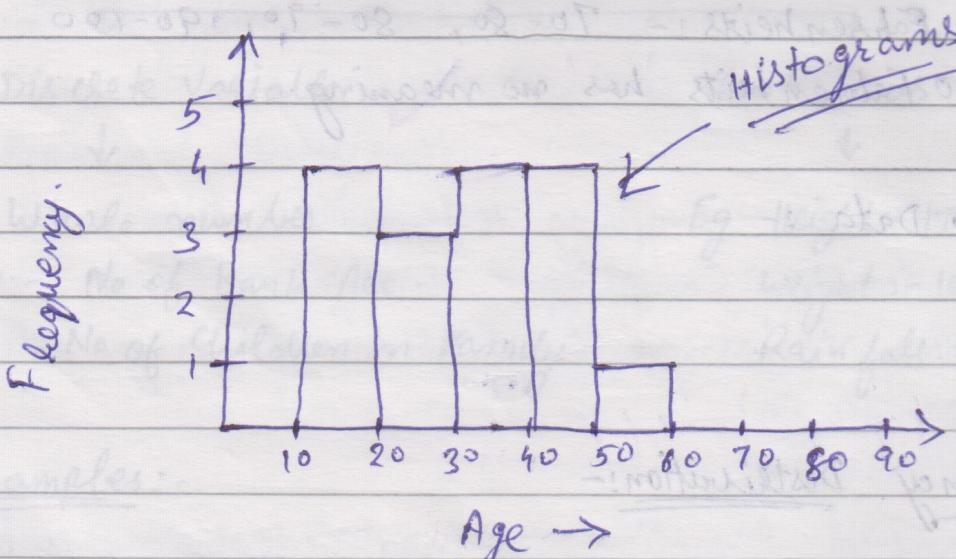
Eg:- Bar chart



Histograms: → Continuous Variables.

Eg:- Age - (10, 12, 14, 18, 24, 26, 30, 35, 36, 37, 40, 41, 42, 43, 50, 51)

Bins - 10



\* What is Pdf? -

Pdf is smoothing of "Histogram".

PDF - Probability Density Function.