* Descriptive Statistics:

- Measure of Central tendencies.
  - Mean, Median, Mode
- Summarization of data.
- Histogram:
- Pdf, cdf, probability
- Permutation
- SD, Variance
- Distribution
  - Normal Distribution
  - Gaussian, Bernoulli, Standard Normal, Transformation, Percentale, Outliers
  - 5 number theory

* Inferential Statistics:

- z test, t-test, anova
- Chi-Square, hypothesis testing (P value, Confidence Interval, Confidence level, z-table, t-table), Quantile, Range

Central limit Theorem:

Q2 What is Statistics?

Science of collecting, organizing and analzing data.

Q2 Whatis Descriptive Stats:

Organize and summarize

Q2 What is Inferential Stats:

Techniques to make conclusions / insights from data.

Q2 Whatis population data:

All available data in consideration is called population data.

"Jab data bohat hota hai, we can't analyze every record obtained from source."

Q2 Whatis sample data:

Part of data extracted from population data.

Q Sampling technique:

- Simple random Sampling
- Stratified Sampling
- Systematic Sampling
- Convenience Sampling

Simple random Sampling:

A type of probability Sampling in which the researcher randomly selects a subset of participant from a population.

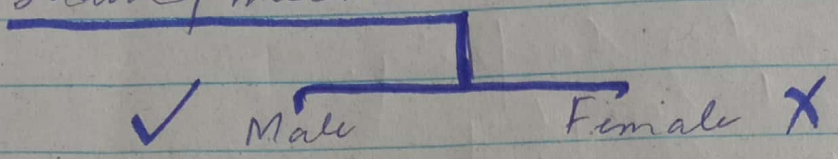
or

Individual Chosen randomly from a large set, all with the same probability.

" Want to conduct a sample from a university. The best place to obtain the sample is none other than cafeteria. As various people there to get efficient result. (unbiased)

Stratified Sampling: divide a population into homogenous sub population called Strata based on specific characteristic (eg color, gender, race, etc)

Shaving machine:



Systematic:

Randomly Taken sample but of same time interval.

1 – 4 – 7 – 10 (gap of 2 people)

Convenience:

Random collection of data Through Survey without setting the criteria.

Variables:

- Quantitative (Measured numerically)
- Qualitative / Categorical (Characteristic)

Quantitative:

- Continuous: Height, weight
- Discrete: Whole numbers (like no of children, no of accounts etc).

Variable Measurement Scales:

- Nominal = Categorical > color, gender, blood group
- Ordinal = Ordered number > age, height, Salary

- Interval = Ranges (1-10, 11-20)
- Ratio = having the same properties as interval data, with an equal and definitive ratio b/w data & absolute 'zero' being treated as a point of origin

Frequency Distribution:

No. of occurence of each element of the data.

Flower	Frequency	CF
Rose	3	3+0=3
Tulip	5	3+5=8
Sunflower	2	8+2=10

- **Bar Chart:** discrete value required.
- **Histogram:** Continuous value required.