

Stats.

Statistics is a branch of mathematics that involves collecting, analysing, interpreting, and presenting data. It provides tools and methods to understand and make sense of large amounts of data and to draw conclusions and make decisions based on the data.

In practice, statistics is used in a wide range of fields, such as business, economics, social sciences, medicine, and engineering. It is used to conduct research studies, analyse market trends, evaluate the effectiveness of treatments and interventions, and make forecasts and predictions.

— Business — Customer Behaviour —
Demand forecasting.

— medical — Clinical test.

— Government Policies. —

— Environmental Science —

Statistics

↓
Descriptive

Data - facts or Information -
eg - Height, Age, DOB, Weight,
Caste, Sex, Name

Collecting, organizing,
Interpretation & report.

↓
Inferential stats.

↓
Sample data
↓
Predict
↓
Population

Sample

Population

→
Qu-
Ques

= 20 math Class

→ Average Age of the
student in Math
Class ?

Age {21, 20, 21, 22, 23, 20, 19, 18, 19, 18, 20, 21}

→ Average age of student in University.

↓
Descriptive

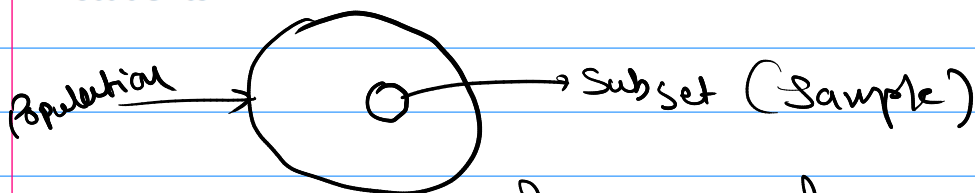
↓
Inferential

Descriptive statistics deals with the collection, organization, analysis, interpretation, and presentation of data. It focuses on summarizing and describing the main features of a set of data, without making inferences or predictions about the larger population.

Inferential statistics deals with making conclusions and predictions about a population based on a sample. It involves the use of probability theory to estimate the likelihood of certain events occurring, hypothesis testing to determine if a certain claim about a population is supported by the data, and regression analysis to examine the relationships between variables

Population refers to the entire group of individuals or objects that we are interested in studying. It is the complete set of observations that we want to make inferences about. For example, the population might be all the students in a particular school or all the cars in a particular city.

A **sample**, on the other hand, is a subset of the population. It is a smaller group of individuals or objects that we select from the population to study. Samples are used to estimate characteristics of the population, such as the mean or the proportion with a certain attribute. For example, we might randomly select 100 students.

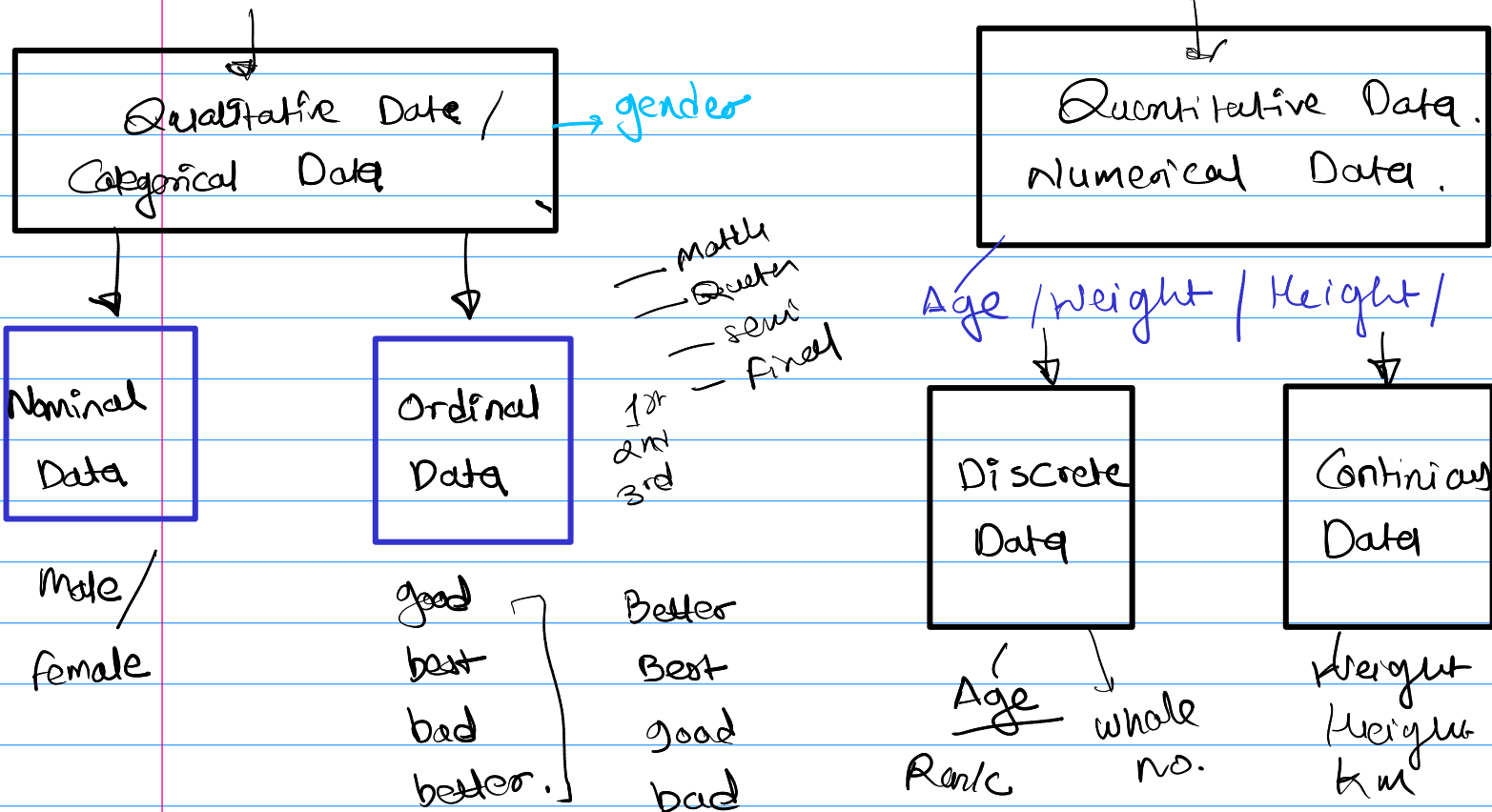


ex- All cricket fans vs fans who present in the stadium.

- All students vs who visit in College.

→ Data →

Type of Data



Measure of Central Tendency -

A measure of central tendency is a statistical measure that represents a typical or central value for a dataset. It provides a summary of the data by identifying a single value that is most representative of the dataset as a whole.

Mean Median Mode

mean → The mean is the sum of all values in the dataset divided by the number of values.

$N=5$ - (Total number)

$$\frac{3+4+1+2+5}{5} = \frac{15}{5} = 3 = \text{mean.}$$

Dataset
3
4
1
2
5

Population (N)

Population mean

μ -

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

Sample (n)

Sample mean

\bar{x}

Variable

Name → Age Roll No DOB

3	x_1
4	x_2
1	x_3
2	x_4
5	x_5

$$\mu = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5(N)}$$

$$= \frac{\sum_{i=1}^N x_i}{N}$$

Σ - summation
Addition

$$\text{Sample Mean } (\bar{x}) = \frac{\sum_{i=1}^n x_i}{n}$$

$n \rightarrow$ number of sample.

$N \rightarrow$ number of population.

Dis → It is Robust to outlier

$$\text{ex} = \{1, 2, 5, 4, 9, 11, 13\} \Rightarrow \mu = \frac{1+2+5+4+9+11+13}{7} = \frac{36}{7}$$

$$\mu = 5.14$$

$$\text{ex} = \{1, 2, 5, 4, 9, 11, 13, 99\} \Rightarrow \text{mean} = \frac{36+99}{8} = \frac{135}{8} = 16.875$$

outlier

② Median

The median is the middle value in dataset when data is arranged in order.

3 4 1 2 5 6

$$1\ 2\ \boxed{3\ 4}\ 5\ 6 = \left(\frac{3+4}{2} \right) = 7/2 = 3.5$$

$n =$ median

1 2 $\boxed{3}$ 4 5
 ↘ median.

1 2 $\boxed{3\ 4}$ 5 600000

③ Mode - \rightarrow Categorical Data

The mode is the value that appears the most frequently in dataset.

data = { 1, 2, 3, 1, 4, 1, 3, 1, 1, 0 }

