Statistics!

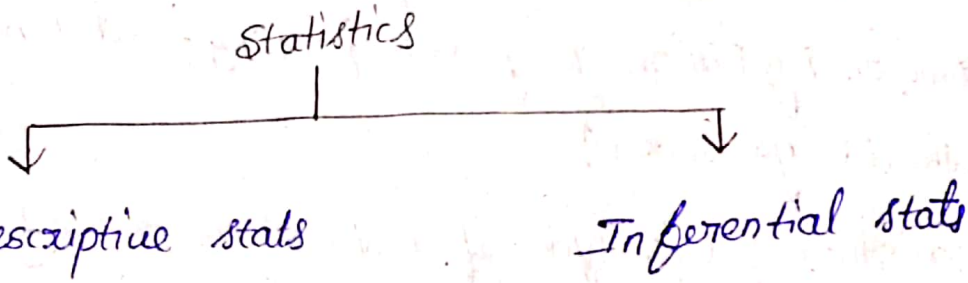
Definition :- Statistics is the science of collecting, organising and analyzing the data.

Data : "facts or pieces of information"

Eg! Ages of students in classroom

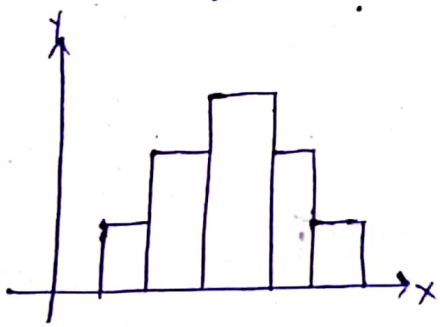
weights of students in classroom

* Statistics divided into two types.

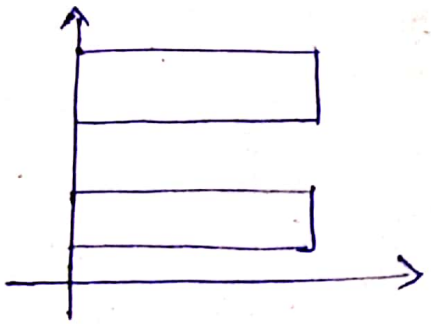


1). Descriptive stats!

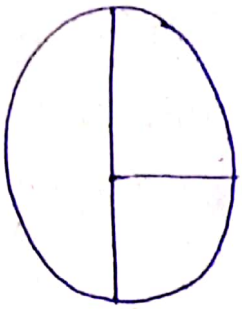
It consists of organising and summarizing the data.



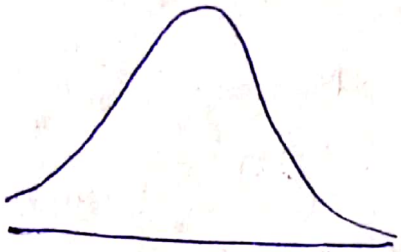
Histogram



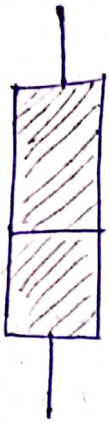
Bar chart



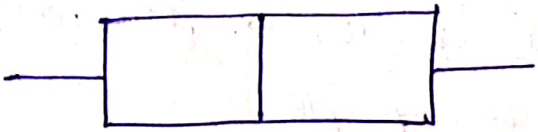
Pie chart



Distribution



candle stick



Box plot

2. Inferential Stats :

It consists of collecting sample data and making conclusion about population data using some experiments. (Hypothesis Testing)

Ex:-

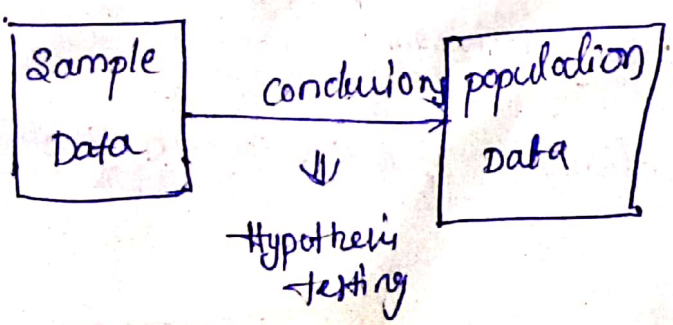
university → 500 people

class A → 60 people

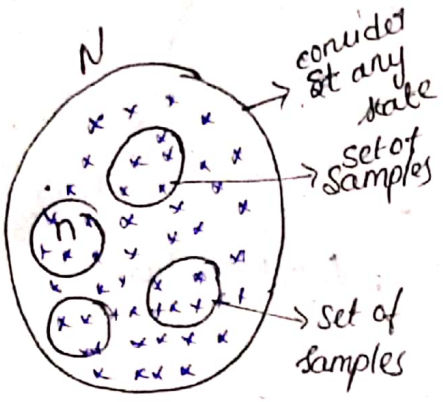
↓

Sample data ⇒ Age ⇒ Average age of the entire university

*



Sample Data Vs Population Data:



population data represented by (N)

Sample data is represented by (n).

Let us consider a state which is consisting of 10 crore people.

⇓

These 10 crore people is called population data.

⇓

To select some set of people in some places with 1000 people ⇒ Sample size

⇓

These data data is called Sample data.

Eg: Let's Say there are 20 classrooms in a university and you have collected the age of students in one classroom.

Ages { 21, 20, 18, 34, 17, 22, 24, 25, 26 }

weight { 60, 70, 75, 55, 85, 72, 73, 68, 71 }

Based on this example the types type of questions may ask.

Descriptive stats: What is the average age of students in the classroom?

Relationship between age & gender?

Inferential Stats:

* Are the average age of the students in the classrooms less than / greater / equal the average age of the students in the university?

Sampling Techniques:

Sampling → choosing a sample.

1). Simple Random Sampling:

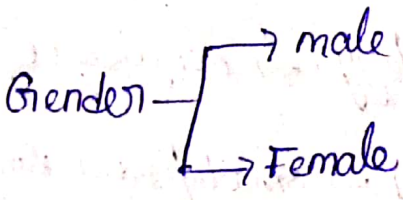
Every member of the population (N) has an equal chance of being selected for your sample (n)

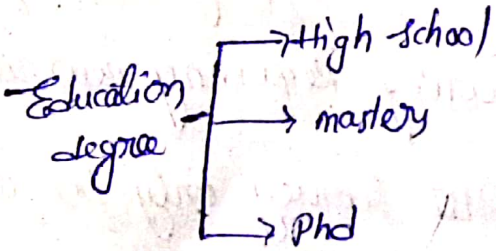
Ex: Asking the reviews in the exit poll of a movie theatre.

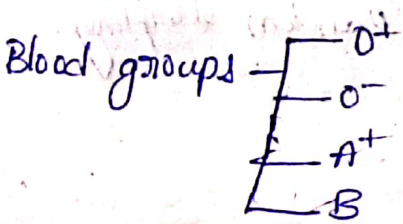
→ Here there is an equal chance of being selected for sample (n) in the movie theatre.

2). Stratified Sampling:

Stratified sampling means the method in which the total population is divided into smaller groups to complete the sampling process.







3). Systematic Sampling:-

The sample members from a larger population are selected according to a random starting point with sampling interval (fixed, periodic).

* Selecting every nth individual out of population(N).

4). Convenience Sampling:-

Only those who are interested in Survey will only participate.

Ex: 1. Data Science Survey form

⇓

Those whose are interested in AI, Data Science they can only fill the form.

2). Job application form for a Software Engineering role.

⇓

Those whose are interested in the role of Software engineering will fill the form.

⇒ Try to find the below examples are belongs to which sampling Techniques.

① Survey Regarding new Technology. → convienice Sampling

② RBI Survey only for women → stratified + Random sampling

③ Credit card: Stratified + Random sampling.

Variable: A variable is a property that can take any values.

Eg: age = 14
age = 25
age = 100

Variables
Ages = [24, 25, 26, 27]

* Variables are two types.

① Quantitative variables → measured Numerically {mathematical operation}

Ex: age, weight, height, rainfall (cm), temp, distance.

② Qualitative variables → also called categorical variables.

{Based on some characteristics they are grouped together}.

Ex: Gender, Types of flowers, Types of movies

Quantitative variable

Discrete variable

Eg: whole numbers

Eg: No. of Bank accounts
{1, 2, 3, 4, 5} 2.5 X

→ No. of children:
whole number

continuous variable

Eg: Decimal values (continuous)

Eg: Height → 6.2 weight 65.5
Rainfall (cm), speed

* Try to find below examples are belongs to which variables.

1). what kind of variable is marital status?

Ans: Categorical variable.

2). what kind of variable is Ganga river length?

Ans: continuous variable

3). what kind of variable is movie duration?

Ans: continuous variable

4). what kind of variable is pincode?

Ans: Discrete variable

5). what kind of variable is IQ?

Ans: Discrete variable