Statistics

Introduction to Statistics:-

→ To check the quality of the data (general statement)

high quality, high performance accuracy.

[Descriptive statistics]

→ To make statement / conclusion.

[Inferential statistics]

What is Statistics?

It is the science of collecting, Organizing and analyzing data (for better decision making).

What is Data?

Data [raw Data] means raw facts or piece of information, that also can be measured (numerical data).

→ Information will contain data, but data doesn't contain information.

→ Facts mean non false statements.

Ex:- The IQ of a class students.

{98, 97, 68, 57, 110} avg_IQ, min_IQ, max_IQ

Descriptive Statistics:-

It Consists of Organizing and summarizing Data.

Inferential statistics:-

Techniques where we used the data that we have measured to form conclusion.

(or)

To make a statement / Conclusion on a Descriptive statistics we use inferential statistics.

Ex:- ① Are the avg marks of the java class students is same as python class students in the Besant. [inferential]

② What is the avg marks SQL students [Descriptive]

Population (N) and Sample (S):-

Population (N):-

The Entire group of data we call it as a population.

Ex:- All peoples in India.

Sample (S):-

As a subset of a population we call it as a Sample.

Ex:- one lakh people from different region of India.

Key Points:-

-> Populations are larger than Samples.
-> Samples should be representative of the population.
-> Samples allow for easier, faster and less costly Data collection

Types of Sampling techniques:-

-> Simple Random Sampling
-> Stratified Sampling
-> Systematic Sampling
-> Convenience Sampling.

Simple Random Sampling:-

Every member of a population has an Equal chance being selected for our sample.

(or)

Every individual is chosen purely by chance.
Ex:- avg ratio of a bike. married people in Bangalore.

Stratified Sampling:-

Where the population split into non overlapping groups.

(or)

population is divided into ~~clusters~~ groups like age or gender, and random samples are taken from each group.

Ex:- the person is alive or dead
the avg ratio of a men or women who are

taking a course in Python.

Systematic Sampling:-

From the population [Split into non o/x every n^{th} sample is called it as a systematic sample.

(or)

Selecting every n^{th} individual from a list.

Ex:- Every 10^{th} person.

→ while doing survey in the mall on the topic of modernization collecting information of every 5^{th} person who is coming out from mall.

Convenience Sampling:-

The sample is collected based on our Convenience (Easiest to reach) from the particular Domain Experts.

Ex:- friends, classmates.

Collecting data from your friends, or nearby people because they are easy to contact.