29/10/25

Introduction to Statistics

* To check the quality of the data
  → Descriptive Statistics

* To make a statement / conclusion
  → Inferential Statistics

What is Statistics?

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

What is data?

Data means raw facts or pieces of information that also can be measured.

Eg: The IQ of a class students

{98, 97, 68, 57, 110}

avg, min, max

Descriptive Statistics

It consists of organizing & summarizing

Inferential statistics

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

To make a statement/conclusion on a descriptive statistics

Eg. Qw.

* Are the avg marks of the Java class students are same as that of python class?
  → Inferential.
* What is the avg marks of SQL students?
  → Descriptive

Population (N) and Sample (n)

- The entire group of the data is population.
  eg: All people in India
  - A subset of a population is sample.
    eg: One Lakh people from different region of India

* Populations are larger than samples.
* Samples should be representative of population.
* Samples allow for easier, faster & less costly data collection.

Types of sampling techniques

1. Simple random sampling
   - Every member of a population has an equal chance of being selected for our sample.
   - eg: Avg mileage of any 120cc bike.
     Avg ~~married~~ rate of married people in banglore

2. Stratified sampling
   - where the population is split into ~~not~~ non-overlapping groups.
   - eg: ~~Avg ratio of~~ Person is alive or dead.

3. Systematic sampling
   - From the population every n^{th} sample we will take.
   - eg: While doing survey in the mall on the topic of modernization, collecting information of every fifth person who is coming out from mall.

4. Convenience sampling
   - The sample is collected based on our convenience from the particular domain experts.
   eg: A ~~factory~~ teacher surveys only students in her class to know opinions about online learning.

Note:- Sampling technique selection always depends on problem statement.

Variable

A variable is a property that can take any value.

Two Types

1) Quantitative (Numerical) variable
2) Qualitative (Categorical) variable.