Introduction to statistics

Definition :- Statistics is a mathematical science including methods of collecting, organizing and analysing data in a such a way that meaningful conclusions can be drawn from them.

data : facts or pieces of information that can be stored and measured.

example1: Scores made by Virat Kohli in last five matches
$$[102, 33, 0, 45, 46]$$

Example2 - Heights of students in the classrooms
$$[110cm, 150cm, 120 cm ----]$$

Motivation | Uses of Statistics

1. Weather forecast
2. Sports Analytics — $[102, 33, 0, 45, 45]$ —— $min - 0$
   $max - 102$
   $Avg -$
3. Election Campaign
4. FMCG | e-commerce
6. Medical | genetics

Types of statistics

① Descriptive Statistics
↓
Describe

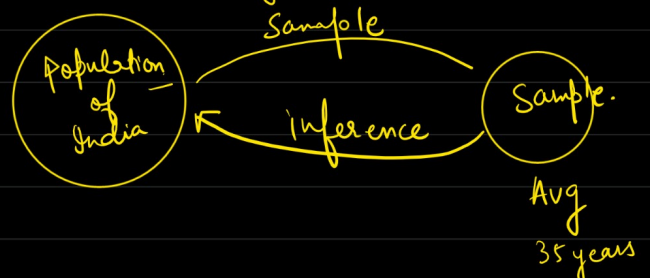**Defn** → It consists of organising and summarizing the data | population.

Virat kohli — $[45, 62, 65, 81 ----]$

strike rate

② Inferential Statistics
↓
Defn Inference.

It consists of using data that has been measured to form conclusion about population.

→ With the given sample data can we conclude Something about the population.

Sample

Population of India ← inference → Sample.

Avg
35 years

① Descriptive Statistics

Example-1 - Virat kohli Scores

[45, 60, 80, 80, 70, 80, 91 ---]

Avg score/strike rate

Example-2 Avg Height/weight of students in class.

[150 cm, 130 cm, 130 cm, ----]

Q. what is Avg height of students in class? → $\frac{150+130+ ---}{50}$ → some avg value.

Q. What is the most frequent height of students in class? → 130cm.

Q. what is avg score scored by Vk in his career?
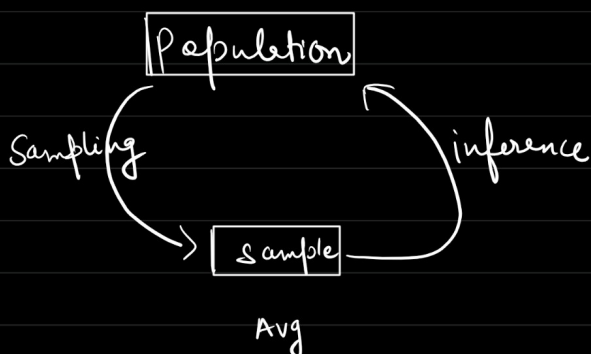
$\frac{45+ 60+ \cdots}{n}$ = avg score

Techniques of Descriptive statistics

① Measures of Central tendency (mean, median, mode)
② Measures of Symmetricity (skewness, kurtosis)
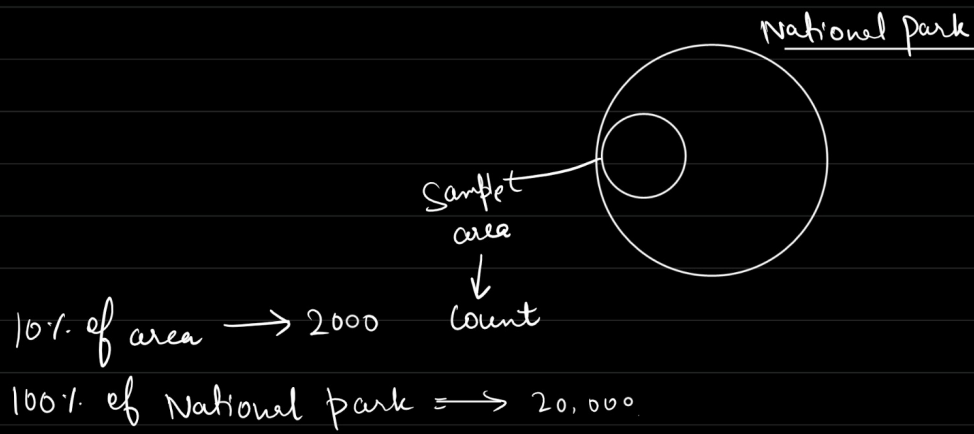③ Measures of dispersion (standard deviation, variance)

② inferential statistics

Why? → Population is large
→ time and resource constraints

Example -1   Avg age of Population of India.

$$\boxed{\text{Population}}$$

Sampling   inference

$$\boxed{\text{sample}}$$

Avg

Example-2 → No of trees in a National Park

National park

Sample
area
↓
Count

10% of area ⟶ 2000
100% of National park ⟹ 20,000

Descriptive Question	Inferential Question

* Classroom

→ What is Avg age/height of students in class ?

→ Are the avg height of students in the class is what you expect in the entire school ?