Introduction to Statistics

Defn : Statistics is the science of collecting, organizing and analyzing the data

⇓

Decision Making Process

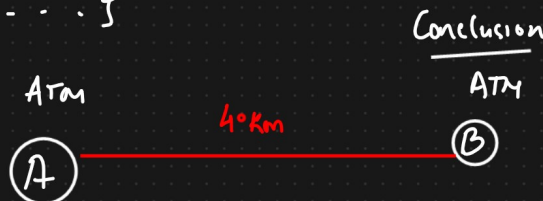Data ÷ "facts or pieces of information"

Eg : Heights of student in the class

$\{175cm, 180_2m, 190cm - - - -\}$

IQ of the student

$\{85, 90, 100 - - - - \}$

Eg :

Conclusion

ATM   ATM

Ⓐ ——————40 Km—————— Ⓑ

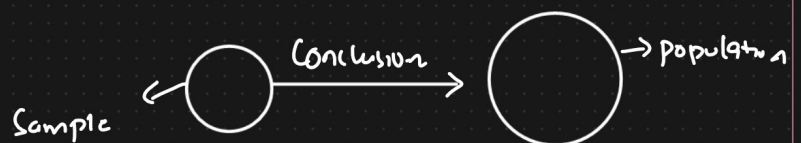## ② Types of Statistics

① Descriptive Stats

Defn : It consists of organizing and summarizing the data

① Measure of Central Tendency

[Mean, Median, Mode]

② Measure Of Dispersion

[Variance, Std]

② Inferential Stats

Defn : It consists of using data you have measured to form Conclusion

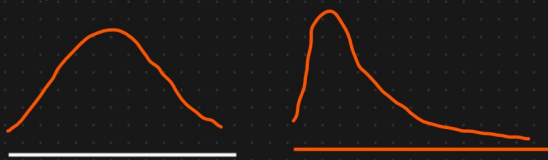Sample ⟵ Conclusion ⟶ ◯ → Population

① Z-test

② t-test

Hypothesis Testing, P value, Significance

② Histograms, Bar chart, pie charts

Eg:- Let say there are 50 students in a Maths class in the university. We have collected the height of the student in the class.

$$[175cm, 180cm, 160cm, 140cm, 130cm, 140cm, 140cm ---]$$

Descriptive Question

$$\left[\frac{175+180+160+140+130+140+140...}{50}\right] = \text{Average value}$$

"What is the average height of the students in the class"

"What is the common height of the students" $= \underline{140cms}$

Inferential Question

Sample data.

"Are the height of the students in the classroom similar to what you expect in the entire college"
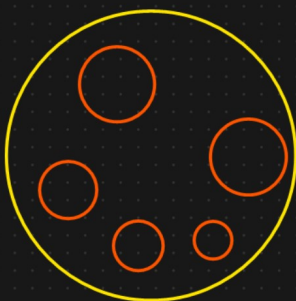
↑ Average

↓

Population

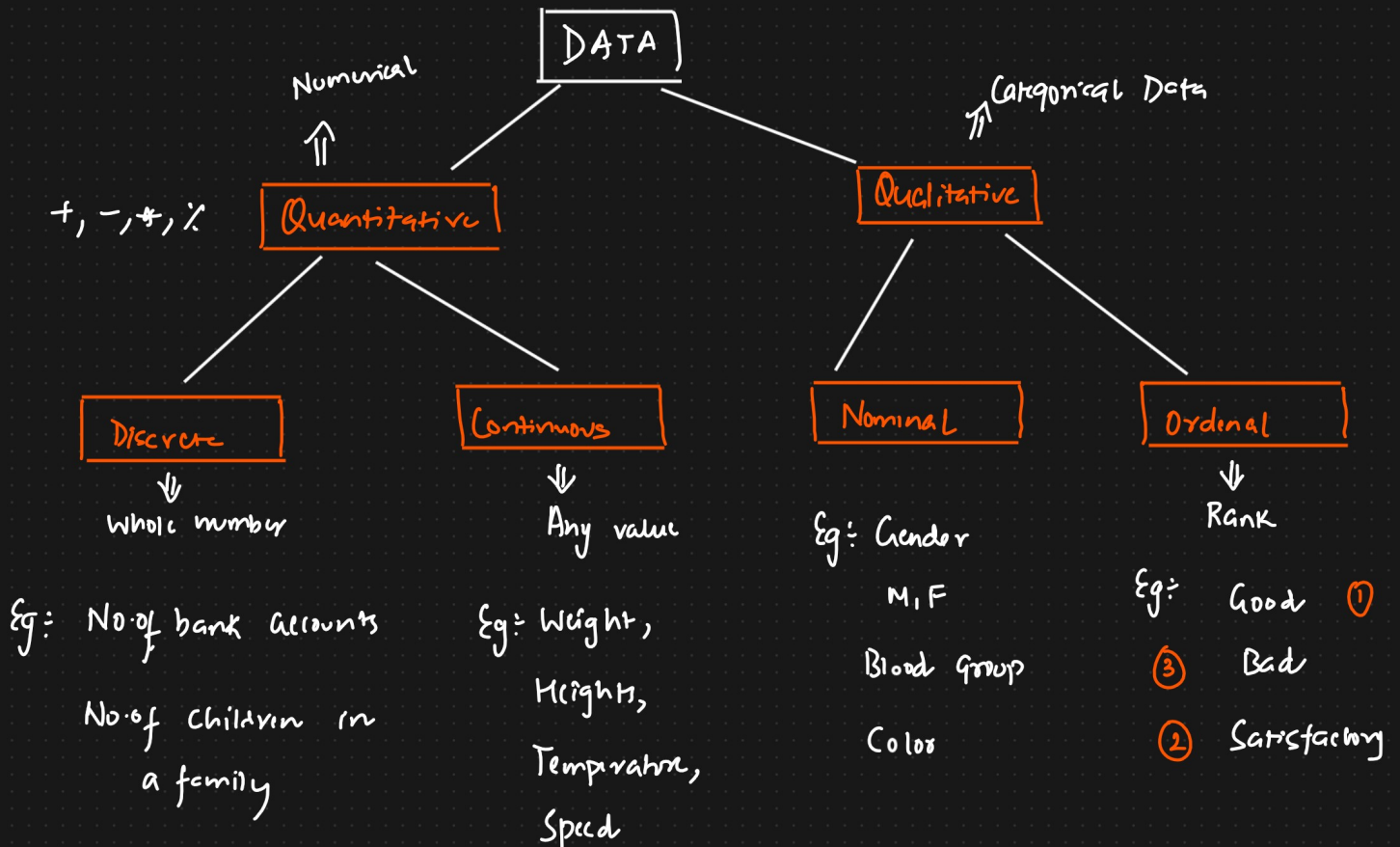Sample Data And Population DATA

Eg:- Exit Poll

Party A   56%
Party B   46%

③ Types Of Data

DMC	DC	ISI	BUI	FWI	Classes	Region
3.4	7.6	1.3	3.4	0.5	not fire	0
4.1	7.6	1	3.9	0.4	not fire	0
2.5	7.1	0.3	2.7	0.1	not fire	0
1.3	6.9	0	1.7	0	not fire	0
3	14.2	1.2	3.9	0.5	not fire	0
5.8	22.2	3.1	7	2.5	fire	0
9.9	30.5	6.4	10.9	7.2	fire	0
12.1	38.3	5.6	13.5	7.1	fire	0

DATA

Numerical ⇧   ⇧ Caregonical Data

+, −, *, %   Quantitative   Qualitative

Discrete   Continuous   Nominal   Ordinal
  ⇩   ⇩   ⇩
Whole number   Any value   Eg: Gender   Rank

Eg: No.of bank accounts   Eg: Weight,   M, F   Eg: Good   ①

No.of children in   Height,   Blood Group   ③   Bad
a family   Temperature,   Color   ②   Satisfactory
                         Speed

Assignment:

① What kind of variable Marital Status is?

② What kind of variable Nile Rive length is?

③ What "   "   "   is Movie duration?

④ Scale of Measurement of Data

① Nominal Scale Data

② Ordinal Scale Data

③ Interval Scale Data

④ Ratio Scale Data

① Nominal Scale Data :

① Qualitative / Categorical variable

② Eg : Gender, Colors, labels

③ Order does not matter

Eg: Survey [10]

$\begin{cases} \text{Red} \longrightarrow 5 \longrightarrow 50\% \\ \text{Blue} \longrightarrow 3 \longrightarrow 30\% \\ \text{Yellow} \longrightarrow 2 \longrightarrow 20\% \end{cases}$

② Ordinal Scale Data

① Ranking and order matter

② Difference cannot be measured

Eg : Qualification   Marks

$\begin{bmatrix} \text{PHd} & 1^{st} \\ \text{BE} & 3^{rd} \end{bmatrix}$
$\begin{bmatrix} \text{Markrs} & 2^{nd} \\ \text{Blom} & 5^{th} \end{bmatrix}$
Bsc 4th

$\begin{bmatrix} 100 & 1^{st} \\ 90 & 2^{nd} \\ 65 & 4^{th} \end{bmatrix} \Rightarrow$
70   3rd
30   5th

Eg : Temperature $\boxed{0\,F} \leftarrow$

③ Interval Scale Data

① The rank and order matters ✓

② Difference can be measured (excluding ratio) ✓

③ Does not have "0" starting value

$\dfrac{60}{30} \div \dfrac{2}{1}$
$\begin{bmatrix} 30F \\ 60F \end{bmatrix}$ $60-30 \Rightarrow 30$

80 F
90F $90-30 = 60$

−10F
−20F
−30F
40F
50 F

④ Ratio Scale Data

Grades

① Order and Rank matter

0, 100, 90, 60, 35, 45, 50

② Differences and rato are measurable

100, 90, 60, 50, 45, 35

③ It does have a "0" starting

100 : 50 ⇒ 2:1

Examples

① Length of Different rivers in the world?

② Marital status?

③ IQ measurement?