

## # DESCRIPTIVE STATISTICS :

### I. MEASURE OF CENTRAL TENDENCY

It gives the representative value around which the entire dataset tends to cluster.

1. MEAN

2. MEDIAN

3. MODE

#### 1. MEAN (or AVERAGE) :-

$$\bar{X} = \frac{\text{Sum of Observations}}{\text{Total no. of observations}}$$

Eg. :- OBSERVATIONS,  $X = \{2, 4, 6, 8, 10\}$

$$\bar{X} = \frac{(2+4+6+8+10)}{5} = 6$$

∴ We calculate mean for NUMERICAL VARIABLES.

#### \* LIMITATION :-

Affected by OUTLIERS or EXTREME VALUES.

Eg. :-  $\{2, 4, 6, 8, 10, 100\}$ , now mean,  $\bar{X} = 21.67$

OUTLIER affecting previous mean BADLY.

## 2. MEDIAN

- Middle value of the dataset when it is ordered.
- Applicable  $\rightarrow$  NUMERICAL VARIABLES.
- Less influenced by OUTLIERS.

MEDIAN

$\rightarrow n$  is ODD:

$$\text{MEDIAN} = \left( \frac{n+1}{2} \right)^{\text{th}} \text{ OBSERVATION}$$

$\rightarrow n$  is EVEN:

$$\text{MEDIAN} = \left[ \frac{\left( \frac{n}{2} \right)^{\text{th}} + \left( \frac{n}{2} + 1 \right)^{\text{th}}}{2} \right] \text{ OBSERVATION}$$

or

\* The average of the middle two values.

$$\text{E.g. : } x = \{3, 1, 7, 5, 9\}$$

arrange  $\rightarrow \{1, 3, 5, 7, 9\} \Rightarrow 5 \text{ terms} \Rightarrow \text{ODD}$

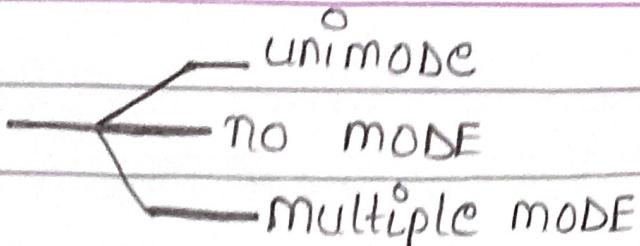
$$\therefore \text{median} = 3^{\text{rd}} \text{ observn} = 5$$

$$\text{E.g. : } x = \{1, 3, 5, 7, 9, 11\} \Rightarrow 6 \text{ terms} \Rightarrow \text{EVEN}$$

$$\therefore \text{median} = \left( \frac{5+1}{2} \right) = \left( \frac{12}{2} \right) = 6$$

Date:

### 3. MODE



SURVEY → Favourite FRUIT among

APPLE, BANANA, ORANGES :-

∴ MODE = FRUIT mentioned ↑Frequently

- MOST Frequent value.
- Applicable Best FOR CATEGORICAL VARIABLES.
- In case of NUMERICAL VARIABLES → Unique Count SHOULD be less.

$$\text{E.g. } \therefore X = \{2, 4, 3, 4, 2, 6, 7, 2\}$$

⇒ 2 = MODE.