Measures of Central Tendency

Central Tendency :-

The single expression can represent the whole group. This single expression in statistics is known as "Average". (or) central value.

Measures

- Mathematical Averages
  - ↓ Quantitative Data
  - ↓ Mean
  - ↓ 1) Arthemetic Mean (AM)
- positional Averages
  - 1) Median (M)
  - 2) Mode (Z)
- Special Averages
  - a) Moving Average
  - b) progressive

Average.

2) Geometric Mean (GM)

3) Harmonic Mean (HM)

patition values

- Two parts → Median
- four parts → Quartiles
- Ten parts → Deciles
- 100 parts → percentiles

1) Arthemetic Mean :-

Arthemetic mean is defined as "the sum of all observations by no of observations

Denoted by $\overline{X}$ (or) AM

Formulas for calculating A·M

Direct Method

Individual Series

$$\bar{x} = \frac{\sum x_i}{n}$$

$\sum x_i$ = Sum of all individual observations

n = no. of observations

Discrete Series

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$\sum f_i x_i$ = Sum of the product of observations & their respective frequencies

N: Sum of all the frequencies.

Continuous Series

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

x_i = Mid value of class interval

$\sum f_i x_i$

= Sum of the product of observations & their respecti[ve] frequencies

N: Sum of all the frequencies

Indirect Method

Individual Series

$$\bar{x} = A + \frac{\Sigma dx}{N}$$

Discrete Series

$$\bar{x} = A + \frac{\Sigma fidx}{N}$$

Contineous Series

A : assumed Mean

$dx = X_i - A$

n : no.. of observations

A = Assumed Mean

N = Sum of frequencies

$dx = X_i - A$

$$\bar{X} = A + \frac{\sum f_i dx}{N}$$

A = Assumed Mean

$dx = X_i - A$

N = sum of frequencies

X_i = Mid value of class interval

③ Step Deviation method

(contineous series)

$$AM \text{ (or) } \bar{x} = A + \frac{\Sigma f_i d_i}{N} \times C$$

A = Assumed Mean

$$d_i = \frac{x_i - A}{C}$$

N = Sum of frequencies

C = class width

x_i = Mid value of class interval

properties of AM :-

1) Sum of the deviations of all obsurvation from their A·M is "zero"

i.e $\Sigma(x_i - \bar{x}) = 0$

Eg:- Marks of student 'x'

		$(x_i - \bar{x})$
A	45	1.25
L	43	-0.75
E	39	-4.75
M	48	4.25
		0

Avg Marks of $X = \frac{45+43+39+48}{4}$

$= 43.75$

② Sum of the squares of deviations from their respective mean is minimum.

$$\sum (x_i - \bar{x})^2 = \text{Minimum}$$

③ If all the observations are say constant 'k' then mean is also → k

④ Mean is effected by both origin and scale.

$$Y = \underbrace{a}_{\text{origin}} \pm \underbrace{bx}_{\text{scale}}$$

$$\therefore \bar{Y} = a \pm b(\bar{x})$$

Eg:-

Tanuja(x_i)	Mounika(y_i)	Subjects
40	122	A
45	137	L

46	140	E
44	134	M

$$y_i = 2 + 3x$$

$$\bar{x}_1 = 43.75$$

$$\bar{y} = 133.25$$

$$\bar{y}_1 = 2 + 3(\bar{x})$$

$$= 2 + 3(43.75)$$

$$= 133.25$$

(5) Combined Mean

if $n_1 \rightarrow$ observations having $\overline{X_1}$ mean

if $n_2 \rightarrow$ observations having $\overline{X_2}$ Mean

then combined Mean of $n_1 + n_2 = n$ observations

is $$\boxed{\overline{X} = \frac{n_1 \overline{X_1} + n_2 \overline{X_2}}{n_1 + n_2}}$$

6) Corrected Arthemetic Mean:-

If some of the given observations wrongly taken as data and we correct the observations.

$$\text{Then Corrected A.M} = \frac{\text{Correct Sum}}{\text{No. of observations}}$$

Correct Sum

= Wrong Sum − (Wrong observations sum) + (correct observation sum)

Wrong Sum = Wrong A.M $\times$ No. of observations

Solution:

No. of observations = 50

Wrong mean = 39.2

Wrong observations = 52, 23

Correct observations = 25, 32

Wrong Sum = Wrong Mean $\times$ No. of obser

= $39.2 \times 50 = 1960$

Correct Sum = wrong sum − (wrong) + (correct)

$= 1960 - (75) + (57)$

$= 1942$

$$\text{Correct AM} = \frac{\text{Correct Sum}}{\text{No. of}}$$

$$= \frac{1942}{50}$$

$$= \underline{\underline{38.84}}$$

Weighted Arthemetic Mean :

observations	weights (Based on Relative Importance)
x_1	1
x_2	2
x_3	3
x_4	4

$$\text{Weighted Arthemetic Mean} = \frac{\Sigma X_i W_i}{\Sigma W_i}$$

* Geometric Mean :- (G.M)

Defined as the n^{th} root of the product of all the observations in a given series.

$x_1, x_2, x_3, \ldots\ldots x_n$ are 'n' observations

$$\boxed{G.M = (x_1 \times x_2 \times x_3 \times \ldots\ldots \times x_n)^{1/n}}$$

for two observations

$$G.M = (a \times b)^{1/2}$$

$$G.M = \sqrt{ab}$$

for three observations

$$G.M = (a \times b \times c)^{1/3}$$

$$GM = \sqrt[3]{abc}$$

→ GM for individual series :-

$$G.M = \text{Antilog}\left[\frac{\sum \log x_i}{n}\right]$$

n = no. of observations.

→ G.M for discrete series :-

$$G.M = \text{Antilog}\left[\frac{\sum f_i \log x_i}{N}\right]$$

N = Sum of frequencies.

f_i = frequency of x_i observation

→ <u>G.M for Cortineous Series</u>

$$G.M = \text{Antilog}\left[\frac{\Sigma f_i \log x_i}{N}\right]$$

x_i : Mid value of class interval

N = Sum of frequencies

f_i : frequency of x_i observation

Finding out log of any value By using ordinary calculator:-

$\log(x)$

→ 15 times root for "x"

→ Subtract 1 from step 1

→ ÷ 0.000070271

Finding out Antilog of any value By using ordinary calculator:-

$\text{Antilog}(x)$

→ ÷ 227695

→ + 1

→ "x, =" Simultaneously

19 times.

* Finding out 'nth' root of any given number by using ordinary Calculator. $(x)^{1/n}$

→ "15 times root

→ "−1"

→ $\div n$

→ "+1"

→ "x" "=" simultaneously

15 times.

Properties of G.M :-

→ If one of the observation is zero then G.M is also '0'

→ If one of the observation is negative then G.M is not defined.

→ $G.M(X \cdot Y) = G.M(X) \cdot G.M(Y)$

→ $G.M\left(\frac{X}{Y}\right) = \frac{G.M(X)}{G.M(Y)}$

→ G.M is also effected by both origin and scale

$$\boxed{\begin{gathered} y = a + bx \\ G \cdot m \text{ of } y = a + b(G \cdot m \text{ of } x) \end{gathered}}$$

→ <u>Combined G.M</u>

If N_1 → observations → G_1 (G.M)

N_2 → observations → G_2

Combined G.M =

$$\text{Antilog}\left[\frac{n_1 \log G_1 + n_2 \log G_2}{n_1 + n_2}\right]$$

Harmonic Mean :- (H M)

The reciprocal Average of A.M is harmonic Mean.

→ **HM for individual Series :-**

$$H.M = \frac{n}{\sum \frac{1}{x_i}}$$

n = no. of observations.

→ **HM for discrete series :-**

$$H.M = \frac{N}{\sum \frac{f_i}{x_i}}$$

N = Sum of frequencies

→ Hm for contineous series

$$Hm = \frac{N}{\sum \frac{f_i}{x_i}}$$

x_i = Mid value of class interval

N = Sum of frequencies.

properties of H.M :-

* H.M is also effected by both origin and scale

$y = a + bx$

$HM \text{ of } y = a + b(HM \text{ of } x)$

* Combined H.M

→ n_1 → observations HM
↳ H_1

→ n_2 → observations HM
↳ H_2

$$\text{Combined H.M} = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

Relationship b/w AM, G.M & H.M

1) If all observations are equal

then $AM = G.M = HM$

2) If all observations are distinct positive integers

$$AM > GM > HM$$

3) from 1 & 2 $AM \geq GM \geq HM$

4) $GM^2 = AM \times HM$

$$GM = \sqrt{AM \times HM}$$

* Question: 1

The average salary of the whole employees in a company is 400/- per day. The Average salary of officers is 800 per day and that of clerks is 320/- per day. If the no. of officers is 40. Then no. of clerks in the Company?

Sol:- Combined Mean = 400 = $\bar{X}$

$\bar{X}_1 = 800 \quad n_1 = 40$

$\bar{X}_2 = 320 \quad n_2 = ?$

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$400 = \frac{40(800) + n_2(320)}{40 + n_2}$$

$$400(40 + n_2) = 32000 + 320n_2$$

$$16000 + 400n_2 = 32000 + 320n_2$$

$$80n_2 = 16000$$

$$\boxed{n_2 = 200}$$

Q:2 Average of 6 numbers is 30, If the avreraage of first four is 25 and that of last three is 35 then the fourth number is ?

Sol:-

$$\frac{x_1+x_2+x_3+x_4+x_5+x_6}{6}=30$$

$$\underline{x_1+x_2+x_3}+x_4+x_5+x_6=180$$

$$\frac{x_1+x_2+x_3+x_4}{4}=25$$

$$x_1+x_2+x_3+x_4=100$$

$$x_1+x_2+x_3=100-x_4$$

$$\frac{x_4+x_5+x_6}{3}=35$$

$$x_4+x_5+x_6=105$$

$$x_5+x_6=105-x_4$$

$$100-\cancel{x_4}+\cancel{x_4}+105-x_4=180$$

$$205-180=x_4$$

$$\boxed{x_4=25}$$

Question: 3 If the mean of squares of first 'n' natural numbers is 46 then $n=?$

Solution:-

$$\text{Mean} = \frac{\text{Sum of observations}}{\text{No. of observations.}}$$

$$46 = \frac{\frac{n(n+1)(2n+1)}{6}}{n}$$

$$46 = \frac{\cancel{n}(n+1)(2n+1)}{6\cancel{n}}$$

$$46 \times 6 = (n+1)(2n+1)$$

$$276 = (n+1)(2n+1)$$

$$2n^2 + 3n - 275 = 0$$

$$\frac{-3 \pm \sqrt{9 - 4(2)(-275}}{2(2)}$$

$$= \frac{-3 \pm \sqrt{2209}}{4}$$

$$= \frac{-3 \pm 47}{4}$$

$$= \frac{-50}{4}, \frac{44}{4}$$

$$= \cancel{-12.5}, 11$$

$$\boxed{n = 11}$$

* **Question : 4** Sum of deviations of 25 observations measured from 45 is -55 find out A.M of observations.?

$$\sum_{i=1}^{25}(x_i - 45) = -55$$

$$\sum_{i=1}^{25} x_i - 25(45) = -55$$

$$\sum x_i - 1125 = -55$$

$$\sum x_i = 1125 - 55$$

Sum of obser $\sum x_i = 1070$

$$\text{Mean} = \frac{\text{Sum}}{\text{Number}} = \frac{1070}{25}$$

$$\boxed{\bar{x} = 42.8}$$

Question: 5 The weighted A.M of first 'n' natural numbers whose weight are equal to corresponding numbers, is equal to ?

Sol:

$$\text{Weighted AM} = \frac{\sum x_i w_i}{\sum w_i}$$

x_i	w_i	$x_i w_i$
1	1	1^2
2	2	2^2
3	3	3^2
4	4	4^2
$\vdots$		$\vdots$

$\sum n = \frac{n(n+1)}{2}$

$\sum n^2 = \frac{n(n+1)(2n+1)}{6}$

$$= \frac{\sum x_i w_i}{\sum w_i} = \frac{\frac{\cancel{n}\cancel{(n+1)}(2n+1)}{\cancel{6}\,3}}{\cancel{n}\cancel{(n+1)}/\cancel{2}}$$

$$= 2n+1/3$$

Q:6 The H.M of A and B is $1/3$ and H.M of C and D is $1/5$ then H.M of A, B, C and D is ?

Sol:

Combined H.M = $\dfrac{n_1 + n_2}{\dfrac{n_1}{H_1} + \dfrac{n_2}{H_2}}$

$= \dfrac{2 + 2}{\dfrac{2}{1/3} + \dfrac{2}{1/5}}$

$= \dfrac{4}{6 + 10}$

$= 4/16 = 1/4$

Q:7 If H.M of 2 numbers is 4 and AM (A) and G.M (G) satisfies the equation $2A + G^2 = 27$ then the two numbers are ?

Solution:- If a, b are two numbers

$$A.M \rightarrow \frac{a+b}{2}$$

$$G.M \Rightarrow \sqrt{ab}$$

$$HM \rightarrow \frac{2ab}{a+b} = 4$$

$$\cancel{2}ab = \cancel{4}^{2}(a+b)$$

$$ab = 2\underline{(a+b)}$$

$$2A + G^2 = 27$$

$$\cancel{2}\left(\frac{a+b}{\cancel{2}}\right) + ab = 27$$

$$a+b + 2(a+b) = 27$$

$$3(a+b) = 27$$

$$a+b = \underline{9}$$

$$ab = 2(9)$$

$$ab = 18$$

$(a-b)^2 = (a+b)^2 - 4ab$

$= 81 - 4(18)$

$= 9$

$\cancel{a-b = 3}$ $\boxed{b = 3}$

$\cancel{a+b = 9}$

$2a = 12$

$\boxed{a = 6}$

Q:8 Mean of 'n' observations is $\bar{x}$, If the first observation is increased by 1, 2^{nd} observation is increased by

2 and so on then the new mean is ?

Sol:-

$x_1, x_2, x_3, \ldots\ldots x_n \Rightarrow \bar{X}$

$+1, +2, +3, \ldots\ldots +n \Rightarrow ?$

$$\frac{x_1+x_2+\cdots+x_n}{n} = \bar{X}$$

$x_1+x_2+\cdots+x_n = \underline{n\bar{X}}$

$x_1+1, x_2+2, x_3+3, \ldots\ldots x_n+n$

New mean = $\dfrac{x_1 + x_2 + x_3 + \cdots x_n + \dfrac{n(n+1)}{2}}{n}$

$$= \dfrac{n\bar{x} + \dfrac{n(n+1)}{2}}{n} \quad n$$

$$= \frac{\cancel{n}(\bar{x})}{\cancel{n}} + \frac{\cancel{n}(n+1)}{2\cancel{n}}$$

$$= \bar{x} + \left(\frac{n+1}{2}\right)$$

Positional Averages :-

Median :-

* In median the data must be arranged in ascending (or) descending order.

→ **Ungrouped (or) Individual Series :-**

If n is

odd

Median $= \left(\frac{n+1}{2}\right)^{th}$ term

even

Median $=$ Avg$\left(\frac{n}{2}, \frac{n}{2}+1\right)$ terms.

→ **Discrete series:-**

- Arrange the date into ascending or descending order according to their respective x_i' variables not frequency

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{term}$$

N = Sum of frequencies.

→ **Contineous series:-**

$$\text{Median} = l + \left(\frac{\frac{N}{2} - cf}{f}\right) \times c$$

l = lower class limit of of median class

N = Sum of frequencies.

(Just greater than $\frac{N}{2}$ is the median class) cf = Cumulative frequency just above median class

f = frequency of median class.

c = class length / class width.

properties of Median :-

* Median is effected by both change of origin and scale.

$$y = a + bx$$

$$\boxed{\text{Median of } y = a + b(\text{Median of } x)}$$

* Sum of the absolute deviations of all observations from their median is equal to minimum.

$$\Sigma \left(x_i - \text{Median} \right) = \text{Minimum}$$

* If all the observations are equal then median is also equal (or) same.

* Quartiles:-

In partition values, they can be divided into 4 equal parts → Quartiles.

Q_1 ↓ (first (or) lower quartile)

Q_2 ↓ Median = Me ↓ (Second Quartile)

Q_3 → (third (or) upper quartile)

for individual series :-

after arranging the data into ascending (or) descending order.

$$Q_1 = \left(\frac{n+1}{4}\right)^{th} \text{ term}$$

$$Q_2 = 2\left(\frac{n+1}{4}\right)^{th} \text{ term}$$

$$Q_3 = 3\left(\frac{n+1}{4}\right)^{th} \text{ term.}$$

where n = no. of observations

Eg :- Following are the wages of the Laboures 82, 56, 90, 120, 75, 75, 80, 130, 65, find Q_1, Q_2 and Q_3.

Sol :- 56, 65, 75, 75, 80, 82, 90, 120, 130

$$Q_1 = \left(\frac{9+1}{4}\right)^{th} \text{ term}$$

$$= \left(\frac{10}{4}\right)^{th} \text{ term}$$

$$= 2.5^{th} \text{ term}$$

$$= 2^{nd} \text{ Term} + 0.5(3^{rd} - 2^{nd} \text{ term})$$

$$= 65 + 0.5(75 - 65)$$

$$= 65 + 5$$

$$\boxed{Q_1 = 70} \quad \boxed{Q_2 = 80}$$

$$Q_3 = 3\left(\frac{N+1}{4}\right)^{th} \text{ term}$$

$$= 3(2.5)$$

$$= 7.5^{th} \text{ term}$$

$$= 7^{th} \text{ term} + 0.5(8^{th} - 7^{th})$$

$$= 90 + 0.5(120 - 90)$$

$$= 90 + 0.5(30)$$

$$\boxed{Q_3 = 105}$$

for discrete series:-

Arrang the data into assending (or) discending order according to their

x_i: variables (Not frequencies)

$$Q_1 = \left(\frac{N+1}{4}\right)^{th} \text{term}$$

$$Q_2 = 2\left(\frac{N+1}{4}\right)^{th} \text{term}$$

$$Q_3 = 3\left(\frac{N+1}{4}\right)^{th} \text{teaarm}$$

where N = Sum of frequencies.

for contineous series :-

$$Q_n = L + \left(\frac{\frac{nN}{4} - cf}{f}\right) \times c.$$

where n : 1, 2, 3

$\frac{nN}{4}$ is the Quartile class $\begin{cases} \frac{N}{4} \\ 2N/4 \\ 3N/4 \end{cases}$

cf = Just greater than Cumulative frequency of quartile class

f = frequency of quartile class

* Deciles and percentiles :-

Deciles → into 10 equal parts → There are 9 deciles

$D_1, D_2, \ldots D_9$

Percentiles → into 100 equal parts → There are 99 percentiles.

$P_1, P_2, \ldots, P_{99}$

Individual Series :- [after arranging the data into ascending (or) descending order]

$$D_n = n\left(\frac{n+1}{10}\right)^{th} \text{ term } (n: 1, 2, 3 \ldots 9)$$

$$P_n = n\left(\frac{n+1}{100}\right)^{th} \text{ term } (n: 1, 2, 3 \ldots 99)$$

Discrete Series :- after arranging the data into ascending (or) descending order according to x_i values (No frequency values)

$$D_n = n\left(\frac{N+1}{10}\right)^{th} \text{ term } (n: 1, 2, 3 \ldots 9)$$

$$P_n = n\left(\frac{N+1}{100}\right)^{th} \text{ term } (n: 1, 2, 3, \ldots 99)$$

N = Sum of frequencies.

Contineous series :-

$$D_n = l + \left(\frac{\frac{nN}{10} - cf}{f} \right) \times c$$

l = lower class limit of Decile class

n = 1, 2, 3, ...9

N = Sum of frequencies

cf = Cumulative frequency of just greaterthan decile class

f = frequency of decile class

c = class length/width

$$P_n = L + \left(\frac{\frac{nN}{100} - cf}{f} \right) \times c$$

$n = 1, 2, 3 \ldots\ldots 99$

L = lower class limit of percentile class

N = Sum of frequencies

$\frac{nN}{100}$ ⇒ value that just greater than in the frequency coloumn is percentile class

cf = cumulative frequency just greater than percentile class

f : frequency of percentile class

c = class length.

Note :-

$$\boxed{\text{Median} = Q_2 = D_5 = P_{50}}$$

* all portion values are effected by both origin and scale

$$\boxed{Y_p = a + b X_p}$$

* find Q_3 of $x, \frac{x}{2}, \frac{x}{5}, \frac{x}{3}$ if $Q_2 = 10$.

Solution:- $\frac{x}{5}, \frac{x}{3}, \frac{x}{2}, \frac{x}{1}$

$Q_2 = 2\left(\frac{N+1}{4}\right)^{th}$ observation

$= 2\left(\frac{5}{4}\right)^{th}$

$10 = 2.5^{th}$ observation

$10 = 2^{nd}$ term $+ 0.5(3^{rd} - 2^{nd}$ term$)$

$10 = \frac{x}{3} + 0.5\left(\frac{x}{2} - \frac{x}{3}\right)$

$10 = \frac{x}{3} + \frac{1}{2}\left(\frac{x}{6}\right)$

$10 = \frac{4x + x}{12}$

$120 = 5x$

$\boxed{x = 24}$

$24/5, \; 24/3, \; 24/2, \; 24$

$Q3 = 3\left(\frac{N+1}{4}\right)^{th}$ observation

$= 3\left(\frac{5}{4}\right)$

$= 3.75^{th}$ term

$= 3^{rd} + 0.75(4^{th} - 3^{rd})$

$= 12 + 0.75(24 - 12)$

$= 12 + 0.75(12)$

$= 12 + 9$

$\boxed{Q_3 = 21}$

Mode:

→ It is the most repeated value in the given data.

* If the data has just one modal value it is called "Unimodal distribution"
* If the data has just 2 modal values it is called "Bimodal distribution"

Mode for Individual Series :-

Most repeated value in ungrouped data (Individual Series)

Mode for discrete series :-

Value of x_i corresponding highest frequency

Mode for contineous series :-

$$Z = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c$$

l = lower class limit of modal class

f_1 = frequency of modal class

f_0 = frequency of pre-modal class

f_2 = frequency of post-modal class

c = Class length.

" highest frequency of class ⇒ modal class "

Properties of Mode (Z)

* If x, y are related by

$$y = a + bx$$

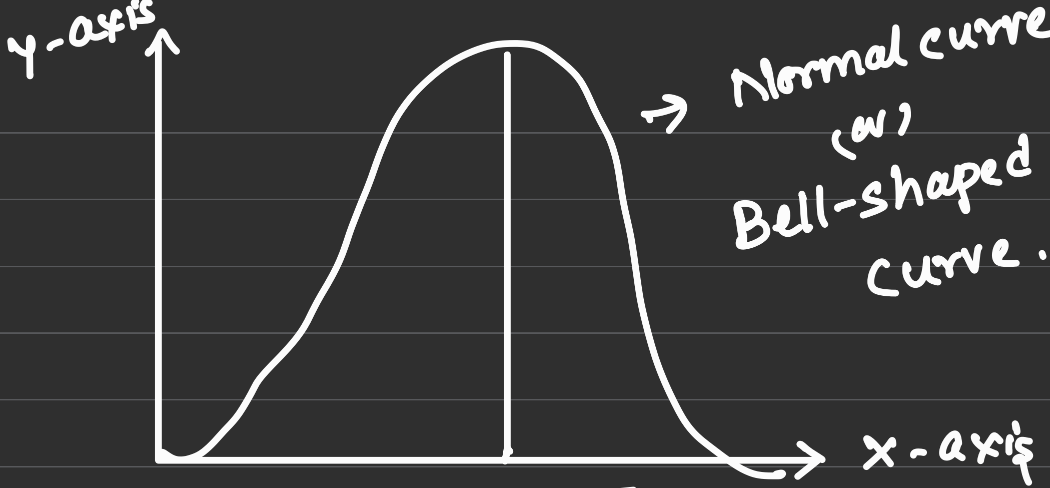
$$\text{Mode of } y = a + b(\text{Mode of } x)$$

* If all observations are constant 'k' mode is also 'k'

Relationship between Mean, Median & Mode :-

* If mean = Median = Mode

Then the distribution is symmetrical (or) Normal (or) Unskewed distribution



$M = Z = \bar{X}$

* If mean $\neq$ Median $\neq$ Mode then the distribution is asymmetrical (or)
Non-normal (or)
Skewed distribution.

Positively skewed distribution

Negatively skewed distribution

↓

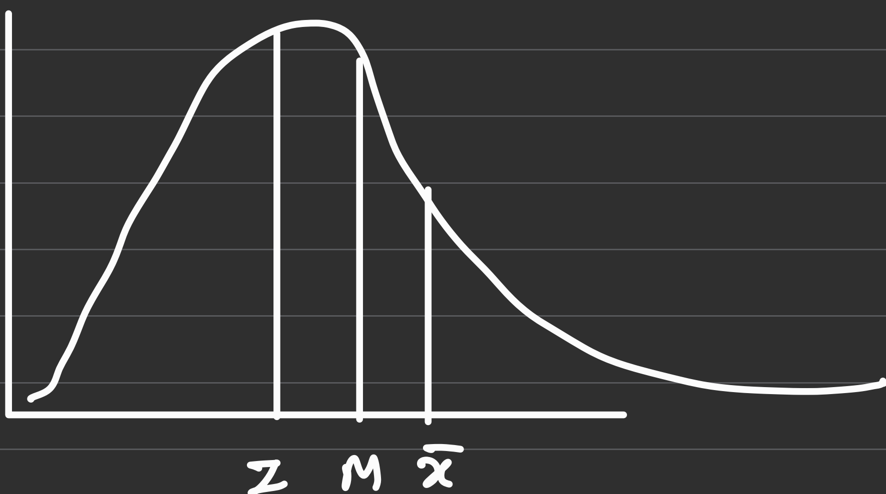
This means the data has a few high

values that pull the mean to the right making the mean greater than the median, which is greater than mode

↓

Mean > Median > Mode

↓



Negatively skewed distribution:

* A negatively skewed distribution, also known as a left-Skewed distribution The mean is pulled towards the left i.e the data has few lowest values.

Mean < Median < Mode

* formula

$$\boxed{\text{Mode} = 3\,\text{Median} - 2\,\text{Mean}}$$

$$\boxed{\text{Mean} - \text{Mode} = 3(\text{Mean} - \text{Median})}$$

Q:- 9 for a moderately Skewed distribution the median is twice the mean, the mode is ____ times the median

Sol:-

$$\text{Median} = 2\,\text{Mean}$$

$$\text{Mean} = \text{median}/2$$

$$\text{Mode} = 3\,\text{Median} - \cancel{2}\left(\frac{\text{Median}}{\cancel{2}}\right)$$

$$\boxed{\text{Mode} = 2\,\text{Median}}$$