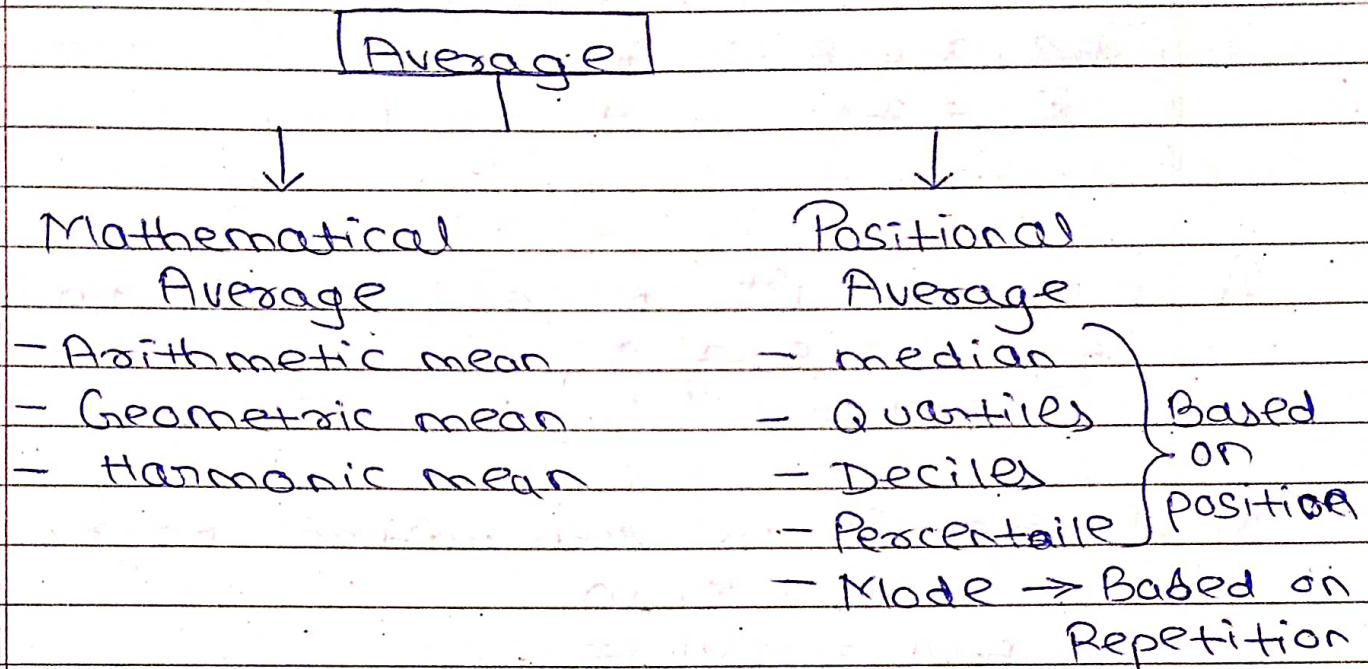


2. Measures of Central Tendency

* Average \rightarrow central value



* Homogeneous data (Equal values)

Avg useful

No

Heterogeneous data (unequal / distinct values)

Yes

* Arithmetic mean (\bar{X})

\rightarrow Sum of all observation divided by no. of observation.

$$\rightarrow \bar{X} = \frac{\sum x}{n} \quad \text{or} \quad \frac{\sum fx}{\sum f}$$

\rightarrow Best Avg. as compared to median & mode

\rightarrow Rigidly defined as compared to median & mode.

\rightarrow Affected by extreme values.

Date _____

* Shortcuts.

1] AM of 1st 'n' natural no $\rightarrow \frac{n+1}{2}$

$$1, 2, 3, 4, 5, 6, 7$$
$$\bar{x} = \frac{7+1}{2} = 4$$

2] AM of 1st 'n' odd natural no $\rightarrow n$

Eg. 1, 3, 5, 7, 9, 11

$$AM = 6$$

3] AM of 1st 'n' even natural no $\rightarrow \frac{n+1}{2}$

Eg. 2, 4, 6, 8, 10

$$AM = 5 + 1 = 6.$$

4] AM of equidistant values

$$= \frac{1^{st} \text{ value} + \text{last value}}{2}$$

1st value \rightarrow Smallest

last value \rightarrow largest

Note:-
Ascending order
required

* Combined AM

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

\rightarrow Avg. of all averages.

* Weighted AM

- weights considered as frequency
- It is used when all values are not equal importance.
- Weighted AM of 1st 'n' natural no. when values of X are equal to corresponding weights.

$= \frac{2n+1}{3} \leftarrow$

x	w	x	f
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6

OR

$$\bar{X}_w = \frac{\sum wx}{\sum w}$$

where w = weight.

* Geometric Mean

- nth root of product of all observation.
- If any value is '0' or '-ve' then GM cannot be determined.
- Mainly used to find Avg. rates (or) ratio's of quantities which are changing at a cumulative rates.
- It is less affected by sampling fluctuation compared to mode & median.

→ It is less affected by extreme values compared to AM.

→ GM is extensively used to construct index no.

$$\rightarrow GM = \sqrt{x_1 \cdot x_2 \cdot x_3 \cdots x_n}$$

$$\textcircled{OR} \sqrt[N]{(x_1)^{f_1} \cdot (x_2)^{f_2} \cdots (x_n)^{f_n}}$$

where $N = \sum f$.

→ If $z = xy$

then GM of $z = (\text{GM of } x)(\text{GM of } y)$

→ If $z = x/y$

then GM of $z = \frac{\text{GM of } x}{\text{GM of } y}$

→ GM is used when change is related to immediate preceding data.

E.g. Avg. rate of depreciation

\textcircled{OR}

Avg. rate of growth of population

\textcircled{OR}

Avg. ROI (Compound interest)

→ Combined GM $= \sqrt[n_1+n_2]{(G_1)^{n_1} (G_2)^{n_2}}$

*** How to find Roots on calc.**

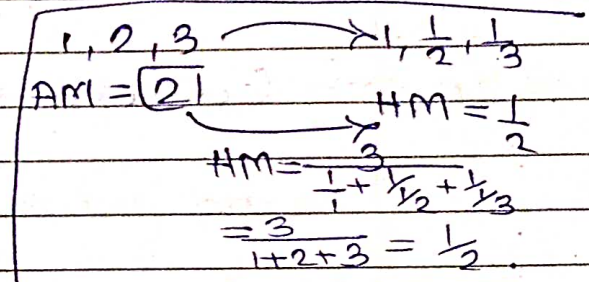
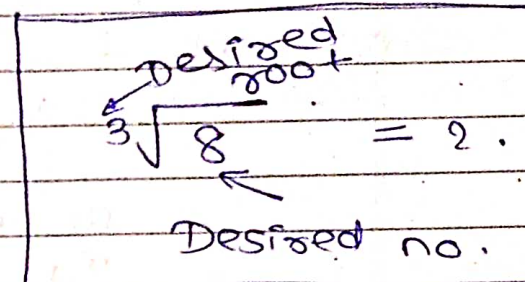
• Take desired ~~root~~ no. & press $\sqrt{\quad}$ 12 times.

• $-1 =$

• \div desired root $=$

• $+1 =$

• Press $x =$
 $x =$
 $x =$
... } 12 times.



*** Harmonic Mean (HM)**

\rightarrow Reciprocal of AM \rightarrow

Note
Data Bhi reciprocal
hona chahiye

\rightarrow No. of observation divided by sum of reciprocal of all observation.

\rightarrow If any value is '0' then reciprocal of '0' is undefined, so HM cannot be calculated.

\rightarrow HM has a very restricted use & they are usually used for calculating avg. speed & avg. rates of quantities etc.

→ less affected by extreme values & sampling fluctuation compare to AM and GM.

$$\rightarrow HM = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}}$$

(OR)

$$\frac{\sum f}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \frac{f_3}{x_3} + \dots + \frac{f_n}{x_n}}$$

$$\rightarrow \text{Combined HM} = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

* Relation betⁿ AM, GM, HM

1) Homogeneous data → $AM = GM = HM$

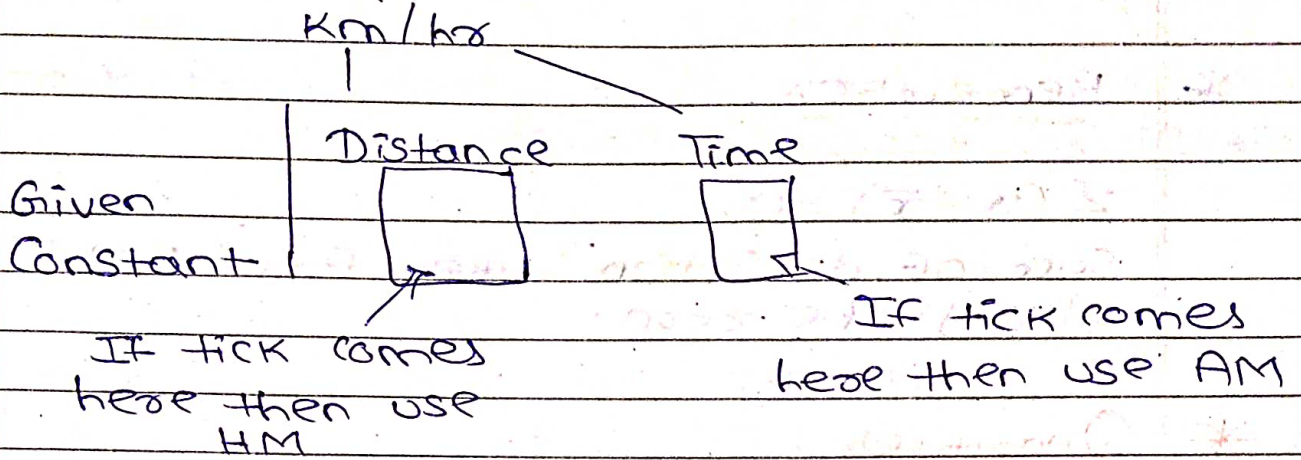
Heterogeneous data / Unequal values → $AM > GM > HM$

Generally $AM \geq GM \geq HM$

2) $GM = \sqrt{(AM)(HM)}$ $GM^2 = (AM)(HM)$

3) If only 2 observations → a & b
 $AM = \frac{a+b}{2}$, $GM = \sqrt{ab}$, $HM = \frac{2ab}{a+b}$

* 2 unit concept



* Median.

→ Middle most value of entire observation.

→ divides data into 2 equal parts.

→ Half of the data lies above median & half of the data lies below median.

→ Best avg. after AM.

→ Best avg for [open class interval]

↓ means

Lower limit of class not given	⊗	Upper limit of class not given	⊗	Both are not given
--------------------------------	---	--------------------------------	---	--------------------

→ Sum of absolute deviation taken from median always be minimum.

$\sum |x - \text{median}| = \text{minimum}$

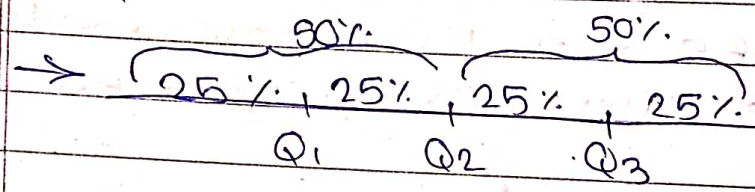
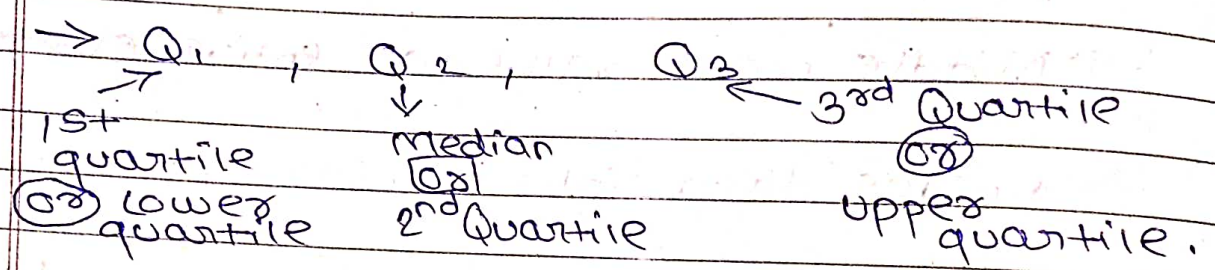
* Remember

$\sum (x - \bar{x}) = 0$

Sum of deviation taken from AM always be zero.

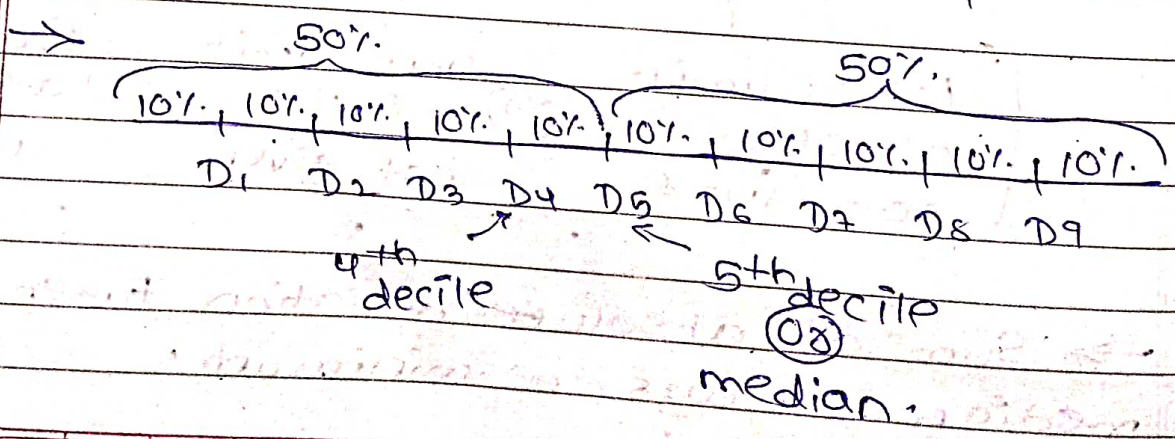
* Quartiles

→ divides data into 4 equal parts:



* Deciles

→ Divides data into 10 equal parts



* Percentiles

→ $P_1, P_2, \dots, P_{50}, \dots, P_{99}$
 \uparrow 50th percentile (or) Median.

$Q_1 = P_{25}$	$Q_3 = P_{75}$
$Q_2 = D_5 = P_{50}$	
$D_1 = P_{10}$	$D_4 = P_{40}$ $D_7 = 70$
$D_2 = P_{20}$	$D_5 = P_{50}$ $D_8 = 80$
$D_3 = P_{30}$	$D_6 = P_{60}$ $D_9 = 90$

	Raw Data	Discrete series	Continuous series																
	3, 5, 8, 7, 10, 9	<table border="1"> <tr><th>X</th><th>R</th></tr> <tr><td>10</td><td>3</td></tr> <tr><td>20</td><td>8</td></tr> <tr><td>30</td><td>5</td></tr> </table>	X	R	10	3	20	8	30	5	<table border="1"> <tr><th>Class</th><th>f</th></tr> <tr><td>10-20</td><td>3</td></tr> <tr><td>20-30</td><td>8</td></tr> <tr><td>30-40</td><td>5</td></tr> </table>	Class	f	10-20	3	20-30	8	30-40	5
X	R																		
10	3																		
20	8																		
30	5																		
Class	f																		
10-20	3																		
20-30	8																		
30-40	5																		
Quartile (Q_k)	$k \left(\frac{N+1}{4} \right)^{th}$ obs.	$k \left(\frac{N+1}{4} \right)^{th}$ obs.	$L_i + \left(\frac{kN - CF}{f} \right) x_h$																
Decile (D_k)	$k \left(\frac{N+1}{10} \right)^{th}$ obs.	$k \left(\frac{N+1}{10} \right)^{th}$ obs.	$L_i + \left(\frac{kN - CF}{f} \right) x_h$																
Percentile (P_k)	$k \left(\frac{N+1}{100} \right)^{th}$ obs.	$k \left(\frac{N+1}{100} \right)^{th}$ obs.	$L_i + \left(\frac{kN - CF}{f} \right) x_h$																
Median	$\left(\frac{N+1}{2} \right)^{th}$ obs.	$\left(\frac{N+1}{2} \right)^{th}$ obs.	$L_i + \left(\frac{N - CF}{f} \right) x_h$																

Notes

Arrange data into ascending order

- 1) X should be in ascending order
- 2) Find less than c.f.

- 1) Data should be continuous series
- 2) Find less than CF
- 3) $L_1 \rightarrow$ Lower limit
 $L_2 \rightarrow$ Upper limit
CF \rightarrow Cumulative frequency
 $h \rightarrow$ class width ($L_2 - L_1$)
- 4) To determine class always find first N value
part $\frac{N}{2}, \frac{KN}{4}, \frac{KN}{10}, \frac{KN}{100}$

* Mode

- \rightarrow Value repeats maximum no. of times.
- \rightarrow Most unstable but quickest method to find avg.
- \rightarrow It is not affected by extreme observation but most affected by sampling fluctuation.
- \rightarrow It can be located graphically with the help of histogram (08) Area diagram (08) frequency diagram.
- \rightarrow For raw data
Mode \rightarrow Value repeats
Zyada bar

→ For discrete series

Mode → value from x (jiski frequency highest hai)

→ Continuous series

$$\text{Mode} = L_1 + \left(\frac{f_m - f_1}{2f_m - f_1 - f_2} \right) \times h$$

f_m → maximum frequency

f_1 → above f_m

h → class width

f_2 → Below f_m

$l_2 - l_1$

* Relation betⁿ AM, Median & Mode.

1) Homogeneous data [Equal value]
(Symmetrical data)

$$\boxed{AM = \text{Median} = \text{Mode}}$$

2) Asymmetrical data [Heterogeneous]

↓
Positively
skewed

$$\text{Mean} > \text{Median} > \text{Mode}$$

↓
Negatively
skewed

$$\text{Mean} < \text{Median} < \text{mode}$$

3) $\text{Mean} - \text{mode} = 3(\text{mean} - \text{median})$

(OR)

$$\text{Mode} = 3\text{median} - 2\text{mean}$$