①

Central tendency [Mean]

* Ideal central tendency
* least affected by Extreme values.

> ungrouped A.M $= \frac{\Sigma x_i}{N}$

> for grouped $= \frac{\Sigma fx}{N}$

→ combined mean $= \frac{N_1\bar{x}_1 + N_2\bar{x}_2}{N_1+N_2}$

→ weighted A.M $= \frac{\Sigma wx}{\Sigma w}$

> properties of A.M
- If each element is constant K the A.M is K.
- $\Sigma(x_i - \bar{x}) = 0$, $\Sigma(\bar{x} - x_i) = 0$
- sum of deviations from A.M is 0
  $\Sigma(x_i - \bar{x}) = 0$
- Sum of deviations of square is minimum only when deviation are taken from A.M
  $\Sigma(x - \bar{x})^2 =$ minimum.
- mean is affected due to sampling fluctuation

→ [Median]

> positional average, represent 50%, divides entire series into 2 equal parts.

Formulas (arrange in ascending order)

- individual / discreate series: $\left[\frac{n+1}{2}\right]^{th}$ term
- Continous series: $l + \left[\frac{\frac{N}{2} - cf}{f}\right] \times h$

> Properties
- Sum of Absolute deviation is minimum when deviations are taken from median.
  $\Sigma|x_i - M|$ is minimum
  $\Sigma|x_i - A|$ is minimum, A - Assumed median.
- for open-end classification median is the best measure of central tendency
- change in Extreme observations does not affect.
- median can find with help of graph called [ogive]

⇒ [Quartails]
- $Q_1 = 25\%$, $Q_2 = 50\%$, $Q_3 = 75\%$

→ Individual series / Descrete series.

$\boxed{Q_1} = \left[\frac{N+1}{4}\right]^{th}$ term, $\boxed{Q_3} = \left[3\left[\frac{N+1}{4}\right]\right]^{th}$

$\boxed{Q_2} =$ median

→ Continous series.

$\boxed{Q_1} = l + \left[\frac{\frac{N}{4} - cf}{f}\right] \times h$, $\boxed{Q_3} = l + \left[\frac{\frac{3N}{4} - cf}{f}\right] \times h$

$\boxed{Q_2} = l + \left[\frac{\frac{N}{2} - cf}{f}\right] \times h$

⇒ [Decile]
- divid series into 10 parts, D_5 = median.
- Individual / discreate:

$D_1 = \left[\frac{N+1}{10}\right]^{th}$, $D_3 = \left[3\left(\frac{N+1}{10}\right)\right]^{th}$, $D_8 \left\{8\left(\frac{N+1}{10}\right)\right]^{th}$

- continous series

$D_1 = l + \left[\frac{\frac{N}{10} - cf}{f}\right] \times h$ $D_6 = l + \left[\frac{\frac{6N}{10} - cf}{f}\right] \times h$

[∴ locate $\frac{N}{10}$ in cf and select D_1 class.]

⇒ [Percentile]
- divides entire series into 100 parts.

> $P_{50} = Q_2 = D_5$

Descrete / Individual

$P_1 = \left[\frac{N+1}{100}\right]^{th}$, $P_3 = \left[3\left(\frac{N+1}{100}\right)\right]^{th}$ term

Continous series

→ locate $= \frac{N}{100}$ in cf $P_{26} = \frac{26N}{100}$

$P_1 = l + \left[\frac{\frac{N}{100} - cf}{f}\right] \times h$

⇒ [mode]
- observation with highest frequency.

Individual: which number repeated more time

Discreate: which frequency is high.

Continous $= l + \left[\frac{f_1 - f_0}{2(f_1) - f_0 - f_2}\right] \times h$.

f_0 = before

f_2 = After.

∴ which has highest frequenty modal class.

- mode can be find by using graph.

Relation b/w mean, median, mode

3 median = mode + 2 mean.

symmentorical distribution (bell shaped curve)

mean = median = mode.

non-symmentorical

→ mean ≠ median ≠ mode

→ mean < median < mode

→ mean > median > mode.

→ Geomatric mean.

n^{th} root of the product of 'n' observations.

$GM = [x_1 \times x_2 \times x_3 \times \ldots x_n]^{1/n}$.

How to find root (trick)

√ = 12 times, -1, ÷n, +1, (x=) 12 times

- If the qustion is about percentage % the GM

⇒ Propeties
- If all observations are same the GM = k

$\log G = \frac{\Sigma \log x}{N}$, $G = \text{Antilog}\left[\frac{\Sigma \log x}{N}\right]$

- G.M mean of (xy) = GM of x × GM of y
- G.M of $\left(\frac{x}{y}\right) = \frac{\text{GM of } x}{\text{GM of } y}$

→ combined GM. : $\left[G_1^{N_1} \times G_2^{N_2}\right]^{\frac{1}{N_1+N_2}}$

⇒ [Harmonic mean]

- 'HM' is the reciprocal of the average of reciprocal of 'n' items'

$$HM = \frac{N}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\cdots\frac{1}{x_n}}$$

- Squence (x =)
- Reciprocal ($\div$ =)

> descrete = $HM = \frac{N}{\Sigma\left(\frac{f_i}{x_i}\right)}$

> continous, -11- (same)

→ properties :

- If all observations are same HM = K

- combined HM :- $\frac{N_1+N_2}{\left[\frac{N_1}{H_1}\right]+\left[\frac{N_2}{H_2}\right]}$

Note :-

- If all items are same = AM = HM = GM
- If all observations are diffrent
  AM > GM > HM
  AM ≥ GM ≥ HM.
- $GM^2 = AM \times HM$
- weighted A.M = $\frac{\Sigma wx}{\Sigma w}$
- weighted H.M = $\frac{\Sigma w}{\Sigma\left(\frac{w}{x}\right)}$
- -11- G.M = Antilog $\left[\frac{\Sigma w(\log x)}{\Sigma w}\right]$
- $1+2+3+\cdots+n = \frac{n(n+1)}{2}$
- $1^2+2^2+3^2+\cdots+n^2 = \frac{n(n+1)(2n+1)}{6}$.
- A.M of first 'n' natural no = $\frac{n+1}{2}$
- HM of 1, 1/2, 1/3, 1/4 ... 1/n is = $\frac{2}{n+1}$

Note	change in origin (+ or -)	change of scale (× or ÷)
mean	$\bar{x}+K$	$K\times\bar{x}$
median	median+K	K× median
mode	mode+K	K× mode

- weighted H.M of $1^2, 2^2, 3^2, \ldots n^2$ = $\frac{2n+1}{3}$ when wight is 1, 2, 3 ... n

[DESPERSION]

> dispersion: The degree of the scatterness or spread or variation of the variable about a central value is called dispersion.

→ measure of dispersion.

- Absolute measure of dispersion :- They are measured and expressed in the units of data like cm, kg, km.
  > Range
  > quartile deviation
  > mean deviation
  > standard deviation.
- Relative measure of dispersion :- used for comparing two variable with diffrent units
  > coefficient of range
  > -11- of QD
  > -11- of MD
  > -11- variance.

→ [Range] :- The range is the diffrence b/w the largest and the smallest values in the distribution.

- $\boxed{\text{Range} = L-S}$, $\boxed{\text{coefficient of Range} = \frac{L-S}{L+S}\times 100}$

→ Properties

- If all the observations are same then Range = 0
- no effect → change of origin
- change in range when there a change in scale. $\boxed{|K| \times (\text{old range})}$

→ [Quartile deviation]

It is defined as half of the difference b/w the third quartile and the first quartile in a given data set.

- Inter quartile range = $Q_3 - Q_1$
- quartile deviation or Semi interquartile range = $\frac{Q_3-Q_1}{2}$
- co-efficient of QD = $\frac{Q_3-Q_1}{Q_3+Q_1}\times 100$

> If the distribution is symentrical $\frac{Q_3-Q_1}{2}$ = median

→ formulas same as median.

→ [mean deviation]

The average of the absolute deviations from a central value is called mean deviation.

> $d = (x-\bar{x})$, $\frac{\Sigma d}{N}$ = MD.

> $MD_{\bar{x}} = \frac{\Sigma f|x-\bar{x}|}{N}$ $\qquad$ $MD_M = \frac{\Sigma f|x-M|}{N}$

> coefficient = $\frac{MD}{\bar{x}}\times 100$ $\qquad$ coefficient = $\frac{MD}{\text{mode}}\times 100$

② → [Standard deviation]

Square root of the average of the square of deviations of all observations from arithmetic mean.

$S.D(\sigma) = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}}$ (or) $\sqrt{\frac{\Sigma x^2}{N} - \left(\frac{\Sigma x}{N}\right)^2}$

discreate:

$S.D = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}}$ (or) $\sqrt{\frac{\Sigma f x^2}{N} - \left(\frac{\Sigma f x}{N}\right)^2}$

- If the observations are in decimals like 2.5, 5.5, then use this!

$SD = \sqrt{\frac{\Sigma f d^2}{N} - \left[\frac{\Sigma f d}{N}\right]^2}$ where :- A is assumed mean

$[d = x - A]$.

- Another formula :- $\sqrt{\frac{\Sigma f v_i^2}{N} - \left[\frac{\Sigma f v_i}{N}\right]^2}$

where $v_i = \frac{x-A}{h}$

\# $SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}}$ variance $= \frac{\Sigma(x-\bar{x})^2}{N}$

$SD = \sqrt{\text{variance}}$ variance $= (SD)^2$ σ^2

- change in origin: no change
- Change in scale: $|K| \times$ old SD.

variance $= (K)^2 \times$ old variance.

formulas :

* AM of a & b $= \frac{a+b}{2}$

* SD of a & b $= \frac{|a-b|}{2}$

* AM of first 'n' numbers $= \frac{n+1}{2}$

* SD of first 'n' numbers $= \sqrt{\frac{n^2-1}{12}}$

* combined S.D or pooled S.D)

$\sqrt{\frac{N_1(\sigma_1^2+d_1^2)+N_2(\sigma_2^2+d_2^2)}{N_1+N_2}}$.

→ $d_1 = \bar{x}_{12} - \bar{x}_1$, → $d_2 = \bar{x}_{12} - \bar{x}_2$

→ $\bar{x}_{12}$ = combined mean $= \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$

→ co-efficient of variance $= \frac{SD}{mean} \times 100$

$C.V = \frac{\sigma}{\bar{x}} \times 100$

∴ used when question is about which one is more variance and which is more consistent.

∴ more C.V = more variable (less consistancy)

∴ less C.V = less variable (more consistancy)

Formula

\# $QD : MD : SD = 10 : 12 : 15$

\# $6QD = 5MD = 4SD$

\# If all data of same ~~vari~~ value then the variance is 0.

\# change of scale and origin.

	origin $(x+K)$	scale $K(x)$
mean	✓ $\bar{x}+K$	✓ $K \times \bar{x}$
median	✓ median$+K$	✓ $K \times$ median
mode	✓ mode$+K$	✓ $K \times$ mode
Range	× (not affect)	✓ $\lvert K \rvert \times$ Range
Q.D	×	✓ $\lvert K \rvert \times$ QD
M.D	×	✓ $\lvert K \rvert \times$ MD
S.D	×	✓ $\lvert K \rvert \times$ SD.