MEASURES OF CENTRAL TENDENCY

1. 3 Type of Data:

i) Raw Data : observation

eg. 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

n = no. of obsersation i.e n = 10

ii) Ungrouped / Discrete frequency distribution:

- Observations & frequencies

eg.

obs.	f
3	10
4	12
5	17
6	20
8	28

$N = \Sigma F$ = Total Frequency = 87

iii) Grouped / Continuous frequency distribution:

- classes & their frequencies.

eg.

Classes	freq.
0-10	10
10-20	20
20-30	20
	$\Sigma F = 50$

2. Types of classes:

Exclusive classes	Inclusive classes
(	

2. Types of classes

Exclusive / overlapping (exclude upper limit)

Inclusive / Non-overlapping (include upper limit)

eg. E

Exclusive classes	Inclusive classes
0 - 10	0 - 9
10 - 20	10 - 19
20 - 30	20 - 29
30 - 40	30 - 39
40 - 50	40 - 49

- Inclusive classes should be converted to exclusive classes.

$d = UL - LL$

(Add 0.5 / less 0.5)

LL UL

Exclusive →	exclusive classes
0 - 9	-0.5 - 9.5
10 - 19	9.5 - 19.5
20 - 29	19.5 - 29.5
30 - 39	29.5 - 39.5
40 - 49	39.5 - 49.5
Class limits	Class Boundaries

open-end classes	close end classes
Below 10	0 - 10
10 - 20	10 - 20
20 - 30	20 - 30
30 & above	30 - 40

3. Some formula (Class conversion imp.)
Incl → exclusive

Class	frequency	Mid value	Freq. width	Relative freq.
0-10	5	5	5/10	5/27
10-30	10	20	10/20	10/27
30-80	12	55	12/50	12/27
	N=27			

- Mid value = $\frac{UL + LL}{2}$
- Class width = UL - LL
- Frequency width / Density = $\frac{\text{Frequency}}{\text{class width}}$
- Relative frequency = $\frac{\text{Frequency}}{\text{Total Frequency (N)}}$

Sum of Relative Frequency is always 1

4. Cumulative Frequencies (C.F):

There are 2 types of C.F

less than C.F
- less than upper limit
- Total ↓

Graph (↑ upper)

More than C.F
- more than lower limit
- Total ↑

Graph (↓ lower)

Class	Freq.	less than C.F	More than C.F.
0-10	5	5 ↓	22
10-20	7	12	17
20-30	10	22	10 ↑
	N=22		

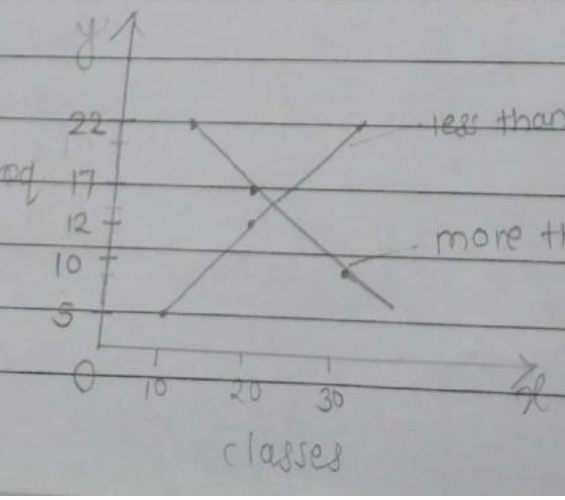
- Sum of less than c.f. of every class + more than c.f. of next class = Total Freq. ($N = \Sigma F$)

5. Ogive / C.F. Graph

Ogive

less than ogive
- plotted with upper limit and less than C.f.

More than ogive
- plotted with lower limit and more than C.f.



- X-cordinate ie Abscissa.
- Intersection of 2 ogive is Median.
- Ogive is S-shaped curve.
- Y-cordinate i.e ordinate that ordinate represent height.

6. Measures of Central Tendency and Dispersion
   (i) Measure of central tendency or average is average of observation in data.
   (ii) Absolute measure of dispersion (M.D): Measure of average variation in the data.
   (iii) Relative measure / coefficient of measure of dispersion: Used to compare variations in data sets.

$$\text{Coeff. of M.D} = \frac{\text{Measure of Dispersion}}{\text{Average}} \times 100$$

$$\left(\text{RMD} = \frac{MD}{MCT} \times 100 \right)$$

Relative measure / Coeff. of MD

7.	Averages (MCT)	M.D	Relative M.D.
•	AM	Range	Coeff. of Range
•	Mode	Q.D	Coeff. of Q.D.
•	Median	M.D (Mean Deviation)	Coeff. of M.D.
•	GM	about Mean, Mode, Median	
•	HM	Standard deviation(σ)	coeff. of variance (C.V)
•	Partition values	Median, Quartiles, Deciles, Percentiles	

8. Partition Values or fractiles:

- Median $= Q_2 = D_5 = P_{50}$
- $Q_1 = P_{25}$
- $Q_3 = P_{75}$
- $P_{70} = D_7$

9. Raw Data:

① $AM = \bar{x} = \frac{\Sigma x}{n}$

② Mode = obs. having max. frequency obs. (Repeated obs.)

③ Median that divides the data into 2 equal parts of 50% each
Median $= \left(\frac{n+1}{2}\right)^{th}$ obs.

④ 3 Quartiles i.e. Q_1, Q_2 & Q_3 those make 4 equal parts of 25% each.
$Q_i = i\left(\frac{n+1}{4}\right)^{th}$ obs.

⑤ 9 deciles i.e $D_1, D_2, \cdots, D_9$ those make 10 equal parts of 10% each.
$D_i = i\left(\frac{n+1}{10}\right)^{th}$ obs.

⑥ 99 percentiles divides data in 100 equal parts of 1% each i.e $P_1, P_2, \cdots, P_{99}$.
$P_i = i\left(\frac{n+1}{100}\right)^{th}$ obs.

Page No.
Date. / /

(7) Suppose $Q_1 = 1.75^{th}$ obs.
then find Q_1 value
$Q_1 = 1^{st}$ obs. $+ 0.75\,(2^{nd}$ obs. $- 1^{st}$ obs.$)$
$Q_1 = $ —.
$Q_i = D_i = P_i = i^{st}$ obs $+ 0. - ((i+1)^{th}$ obs $- (i)^{th}$ obs.$)$

(8) Range $= L - S$

(9) Coefficient of Range $= \frac{L-S}{L+S} \times 100$

(10) Quartile Deviation / Semi Inter Quartile Range
$= \frac{Q_3 - Q_1}{2}$

(11) Coeff. of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

(12) Mean Deviation about mean $= \frac{1}{n} \Sigma |x - \bar{x}|$

(13) M.D about Mode $= \frac{1}{n} \Sigma |x - \text{mode}|$

(14) M.D about Median $= \frac{1}{n} \Sigma |x - \text{median}|$

(15) Coeff. of M.D about mean $= \frac{\text{M.D about mean}}{\text{mean}} \times 100$

(16) Coeff. of M.D about mode $= \frac{\text{M.D about mode}}{\text{mode}} \times 100$

(17) Coeff. of M.D about median $= \frac{\text{M.D about median}}{\text{median}} \times 100$

(18) Variance $= \sigma^2 = \frac{1}{n} \Sigma (x - \bar{x})^2$

(19) Standard Deviation $= \sigma = \sqrt{var}$

(20) Coeff. of Variation $= \frac{\sigma}{\bar{x}} \times 100$

10. Ungrouped Data:

When Freq. are multiply by f inside Σ Replace n by N

(1) AM $= \bar{x} = \frac{\Sigma fx}{N}$ $\therefore N = \Sigma f$

(2) Mode = obs. having max. frequancy

(3) Median $= \left(\frac{N+1}{2}\right)^{th}$ obs.

(4) $Q_i = i\left(\frac{N+1}{4}\right)^{th}$ obs.

(5) $D_i = i\left(\frac{N+1}{10}\right)^{th}$ obs.

(6) $P_i = i\left(\frac{N+1}{100}\right)^{th}$ obs.

(7) Range $= L - S$

(8) Coeff. of Range $= \frac{L-S}{L+S} \times 100$

(9) $QD = \frac{Q_3 - Q_1}{2}$

(10) Coeff. of $QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

(11) M.D about Mean/mode/median
$$= \frac{1}{N} \Sigma F(x - \text{mean/mode/median})$$

(12) Coeff. of MD about Mean/mode/median
$$= \frac{\text{MD about Mean/mode/median}}{\text{Mean/mode/median}} \times 100$$

(13) Variance $= \frac{1}{N} \Sigma f x^2 - (\bar{x})^2$

(14) $SD = \sigma = \sqrt{\text{var}}$.

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

11. Grouped Data: 1st to find $x = \frac{UL + LL}{2}$

(1) $AM = \frac{\Sigma Fx}{N}$

- Classes must be exclusive
- Class have inclusive to convered exclusive.

(2) Mode: classes having max. freq is modal class.
$$\text{Mode} = l + \frac{f_m - f_1}{2 \times f_m - f_1 - f_2} \times h$$

l = lower limit
f_m = freq. of model class
f_1 = Freq. of pre model class
f_2 = Freq. of post model class
h = width = UL - LL

③ Median :

Median class of first c.f $\geq \frac{N}{2}$

$$\text{Median} = l + \frac{\frac{N}{2} - c.f}{F} \times h$$

c.f = freq. of pre median class

④ Quartile :

Class having c.f First time $\geq \frac{iN}{4}$

$$Q_i = l + \frac{\frac{iN}{4} - c.f}{F} \times h$$

⑤ Decile :

class having c.f First time $\geq \frac{iN}{10}$

$$D_i = l + \frac{\frac{iN}{10} - c.f}{F} \times h$$

⑥ Percentile :

Class having c.f. First time $\geq \frac{iN}{100}$

$$P_i = l + \frac{\frac{iN}{100} - c.f.}{F} \times h$$

Note 1 :

1. The coeff. of variance $c.v = \frac{\sigma}{\bar{x}} \times 100$ is used to compare variation in the 2 data sets larger than is the c.v., lesser is the consistency, reliability, uniformity or the homogeneity of the data.

Page No.
Date. / /

Note 2: • Combined / Pooled mean $= \bar{x}_c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

• Combined variance $= \sigma^2_c = \frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}$

$\therefore d_1 = \bar{x}_1 - \bar{x}_c$

$\therefore d_2 = \bar{x}_2 - \bar{x}_c$

• Combined GM cannot be formed.

• $HC = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$

Note 3: Computation of GM & HM

1. Raw Data

$AM = \bar{x} = \frac{\Sigma x}{n}$

$HM = \frac{n}{\Sigma \frac{1}{x}}$

$GM = \sqrt[n]{x_1 x_2 x_3 \cdots x_n} \;/\; (x_1 \cdot x_2 \cdots x_n)^{1/n}$

$GM = \text{antilog}\left[\frac{1}{n}\Sigma \log x\right]$

Note: If $Z = X \cdot Y$
$GM(Z) = GM(X) \cdot GM(Y)$
If $Z = X/Y$
$GM(Z) = \frac{GM(X)}{GM(Y)}$

• In case of only 2 obs.

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

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

Page No.
Date. / /

2. For Frequency distribution

1. $AM = \frac{\Sigma fx}{N}$

2. $HM = \frac{N}{\Sigma \frac{f}{x}}$

3. $GM = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdots x_n^{f_n}}$

Note: i) $AM = \frac{\text{Variable}}{\text{constant}}$ eg. marks, Rains, etc

ii) $HM = \frac{\text{constant}}{\text{Varies}}$ eg. speed, etc.

iii) $GM = \frac{\text{Varies}}{\text{Varies}}$ eg. Ratio, Proportion, %, Rate of change.

Generally, in case of Ratio or percentage, we use Geometric mean.

① $AM \geq GM \geq HM$ (In General)
② $AM = GM = HM$ (obs. are equal)
③ $AM > GM > HM$ (obs. are distinct)

If 2 are equal, then All 3 equal.
eg. AM = 10, GM = 10 i.e HM = 10

Page No.
Date. / /

Note 5: Change of Origin and Scale:

(+, −) → change of origin

(×, ÷) → change of scale

- MD is never negative
- $y = ax + b$ — a: change of scale; b: change of origin

$$MCT(y) = a\,MCT(x) + b$$

$$MD(y) = |a|\,MD(x)$$

- MCT can be 0, +ve.
- MCT depend on change of origin & change of scale both.

But MD is independent of change of origin & depend on change of scale.

Note 6: Wrong and Correct observation given

$$\Sigma x_c = \Sigma x_\omega - x_\omega + x_c$$

- Subtracted wrong
- Add correct

$$\Sigma x^2_c = \Sigma x^2_\omega - x^2_\omega + x^2_c$$

$$\Sigma \frac{1}{x}_c = \Sigma \frac{1}{x}_\omega - \frac{1}{x}_\omega + \frac{1}{x_c}$$

$$P_c = \frac{a \cdot b \cdot c \cdot d}{a} \times x$$

$$P_c = \frac{P_\omega}{x_c} \cdot x_c$$

Page No.
Date. / /

Note 7: Mean − Mode = 3 (mean − median)
— skewed curve.

Note 8: A less than type of ogive, then

Abscissa of point on ogive having ordinate

① $\frac{N}{2}$ is median $= Q_2 = D_5 = P_{50}$

② $\frac{1N}{4}$ is $Q_1 = P_{25}$

③ $\frac{3N}{4}$ is $Q_3 = P_{75}$

④ $\frac{3N}{10}$ is $D_3 = P_{30}$

Note 9:
1) Sum of deviation from mean is 0.
2) Sum of absolute deviation from median is minimum
i.e smallest among mean, mode & median
3) If all obs. are equal then,
Mean = k, GM = k, HM = k
But mode (k) does not exist
std. dev = 0 MD = 0 R = 0
i.e Variance = 0
4) GM cannot be found if obs. have opposite sign or some observation is 0.
HM cannot be found if one obs. is 0.
Weighted mean $= \frac{\Sigma WX}{\Sigma W}$
5) If one obs. have equal freq. then mode doesn't exist.

	Mean	Median	Mode
①	Most commonly use. Based on all obs.	Not based on all obs.	
②	unique (Rigidly)	unique/Rigidly	Not unique
③	Not suitable for open end classes	Most suitable for open end classes	Not most suitable for open end classes
④	Affected by sampling flue.	Not affected	Affected

	Range	QD	MD	Std. dev.
①	very easy	Best for open	End classes	Most commonly use
②	Not based on all obs.	Not based on all obs. (50% only)	Based on all obs.	Based on all obs.
③	most affected by sampling	Not affected	Affected	Affected
④	Not possible for open end class.	Most suitable for open end class	Not suitable	Not

① Mean = Mode = Median — Symmetic data
Mean > median > mode — positively skewed
Mean < median < mode — -ve skewed

(2) $AM \geq GM \geq HM$ — General
$AM = GM = HM$ — equal obs.
$AM > GM > HM$ — distinct obs.

(3) For symmetric data
$MD = 0.8\sigma$
$QD = 0.675\sigma$
Hence, standard error (MD) $= \frac{MD}{0.8}$

$SE(QD) = \frac{QD}{0.675}$

$MD = 1.2\,QD$

(4) For 2 obs.
$\sigma = \frac{1}{2}$ Range / Range $= 2\sigma$

If freq. are equal ingore them

(5) Natural no. $= \frac{n(n+1)}{2}$

Square Natural no. $= \frac{n(n+1)(2n+1)}{6}$

Cube Natural no. $= \frac{n^2(n+1)^2}{4}$

AM of n natural No. $= \frac{(n+1)}{2}$

HM of n natural no. $= \frac{2}{n+1}$

Variance of n natural no $= \frac{n^2-1}{12}$

S.D. of n natural no. $= \sqrt{\frac{n^2-1}{12}}$