Data

ungrouped data
- → Avg
- → variance
- → S.D

grouped data
- → Avg
- → variance
- → S.D.

Ungrouped data

Mean

→ Random data. Ex :- 5, 10, 12, 30, 60, 50.

if $x_1, x_2 - - x_n$, then avg of given data is $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, where n is total number of quantities.

→ Variance :- σ^2 = variance $= \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$.

where $\bar{x}$ is the mean of sample

x_i is variable. , where $i = 1, 2, 3 - -$

σ = Standard deviation = $\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$

Grouped data

heregiven data is arranged in asending or desending order.

Ex:-

Marks Scored	f_i
0 - 5	5
5 - 10	15
10 - 15	20

→ **grouped data and discontinuity**

Marks	f_i
1 - 5	10
6 - 10	15
11 - 15	20

→ When U.B of previous data = L.B of the present then the data is said to be continuous else discontinuous

→ We can add and sub some value from ~~U.f~~ U.B and L.B resp to make discontinuous data continuous.

→ **Average of grouped data**

$$A.m = \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum f_i}, \quad \sum f_i = N, \quad x_i = \text{Midpoint}$$

$$\frac{U.B + L.B}{2} = x_i$$

→ Variance $= \sigma^2 = \frac{1}{N}\left(\sum_{i=1}^{n} f_i (x_i - \bar{x})^2\right)$

→ S.D $= \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}$

→ Mean Deviation $= \frac{1}{N}\sum f_i |x_i - \bar{x}|$

KARL PEARSONS COFFICIENT OF SKEWNESS.

$$SKP = \frac{3(\text{mean} - \cancel{\text{mode}}\ \text{median})}{\sigma} \quad \text{or} \quad \frac{\text{mean} - \text{mode}}{\sigma}$$

$$\text{mean} = \frac{\sum f_i x_i}{\sum f_i}, \quad \text{mean} = \bar{x} = A + \frac{\sum fd}{N} \times i$$

$d = \left(\frac{m - A}{10}\right)$ where m is midpoint.

Q_2 = Median

→ cofficient of skeweness by Bowley's method (SKB)

$$SKB = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

~~Q_3~~ = SKB is involved by quartile deviations.

$Q_1 = \frac{N}{4}$ ↗ see in c.f $= \frac{\sum f}{4}$ → Quartaige deviation

$Q_2 = \frac{N}{2} = \frac{\sum f}{2}$ → two - " "

$Q_3 = \frac{3N}{4} = \frac{3\sum f}{4}$ → three - " "

$$Q_1 = L_1 + \left(\frac{\frac{N}{4} - \text{P.c.f or (mi)}}{f_1}\right) \times c$$

f_1 → corresponding frequency.

(P.c.f = Preceding term of cumulative frequency)

$$Q_2 = L_2 + \left(\frac{\frac{N}{2} - \text{P.c.f or } (m_2)}{f_2}\right) \times c$$

c → width of class interval.

$$Q_3 = L_3 + \left(\frac{\frac{3N}{4} - p.c.f(m_3)}{f_3} \right) \times c$$

L_1	$\frac{N}{4}$	f_1	m_1	$cf \to Q_1$
L_2	$\frac{N}{2}$	f_2	m_2	$cf \to Q_2$
L_3	$\frac{3N}{4}$	f_3	m_3	$c.f \to Q_3$

SKB negative indicates negative skewness.

→ curve fitting.

for st line normal equations are

$\Sigma y = na + b\Sigma x \to$ ① , for $y = a + bx$

$\Sigma xy = a\Sigma x + b\Sigma x^2$ ② normal equations

a and b and be found out by solving equations ① & ②

x	y	x^2	xy	

Ex:- ~~Find~~ Fit a Secondary eqn for given data.

Sol Let the secondary equation can be written as $y = a + bx + cx^2$, where a, b, c are arbitary contants which are determined by solving normal equations.

Normal equations:-

apply summations w.r.t to independentent variable 'x' together eq I

$\Sigma y = \Sigma a + b\Sigma x + c\Sigma x^2$

$\Sigma y = na\ b\Sigma x + c\Sigma x^2$ — ① → Normal eq

Illy multiply independent variable 'x' together given eq I and take sumation on both side of

$yx = ax + bx^2 + cx^3$

$\Sigma yx = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$ — ② → Normal equation

Illy multiply x again and apply summation

$\Sigma yx^2 = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$ — ③ → Normal equation

Now solve for a, b, c.

x	y	x^2	x^3	x^4	xy	x^2y

other curves

4 $y = ab^x$, 5 $y = ae^{bx}$ 6 $y = ax^b$

→ fit a expontial function for the following data

$y = ab^x$

Since given equation is exponential, It reduces to linear form by applying log on both sides

$$\log y = \log a + x \log b \quad \text{---} (1)$$

Put $\log y = Y$, $\log a = A$, $\log b = B$

then, the eq (1) becomes

$Y = A + Bx$ (a) → this is a Linear equation in the form of st line

We know the normal equations for above equation is $\sum_{i=1}^{n} Y = nA + b\sum_{i=1}^{n} x$ --- (2)

$\underbrace{\sum_{i=1}^{n} Yx}_{\sum_{i=1}^{n} x(\log y)} = A\sum_{i=1}^{n} x + B\sum_{i=1}^{n} x^2$ --- (3)

equations 2 and 3 and Normal equation to eq (a) ∴ We can find A and B,

Where $A = \log a$, $B = \log b$.

$a = \text{antilog}(A)$, $b = \text{antilog}(B)$

$\log y = Y$ Now substitute a and b in given eqn of curve. $y = ab^x$

⑤ $y = ae^{bx}$

$\log y = \log a + bx \log e$ — ①

Now $\log y = Y$

$\log a = A$

$b \log e = B$

Now equation ① becomes $Y = A + Bx$ — ②

Now find A and B.

Normal equations of eq ② are $\sum_{i=1}^{n} y = nA + B\sum_{i=1}^{n} x$

$\sum_{i=1}^{n} yx + A\sum_{i=1}^{n} x + B\sum_{i=1}^{n} x^2$

After finding A and B

$a = \text{antilog}(A)$

$b = \cancel{\text{antilog}}\ (\cancel{b\log e}\ \frac{B}{\log e}$

$$\text{mode} = l + \frac{(f_K - f_{K-1}) \times c}{2f_K - f_{K-1} + f_{K+1}}$$

$f_K \rightarrow$ is the Mode frequency

$f_{K+1} \rightarrow$ frequency of next class

$f_{K-1} \rightarrow$ frequency of previous class

$l \rightarrow$ lower bound of mode class

Note:-> if $Q_1+Q_3=2Q_2$ then $SKB=0$

> if $Q_1+Q_3 > 2Q_2$ → +ve Skewness

> if $Q_1+Q_2 < 2Q_2$ → −ve skewness

> if $Q_3=Q_1$ => SKB is infinity.