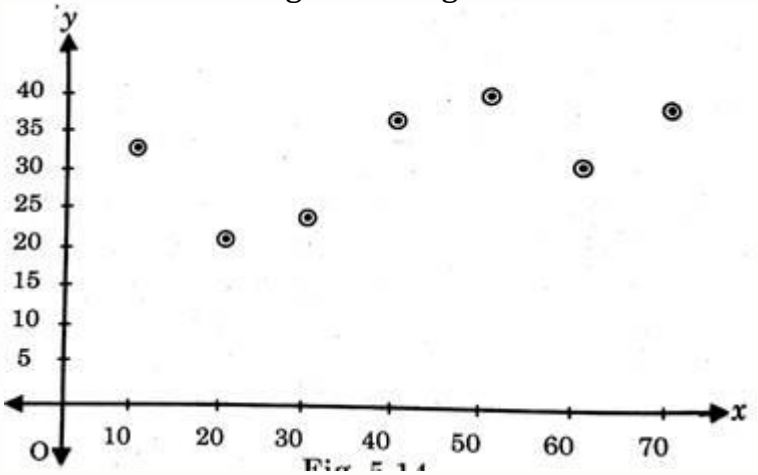
5. CORRELATION

EXERCISE 5.1

1) Draw scatter diagram for the data given below and interpret it.

x	10	20	30	40	50	60	70
y	32	20	24	36	40	28	38

Sol: The scatter diagram is as given below:



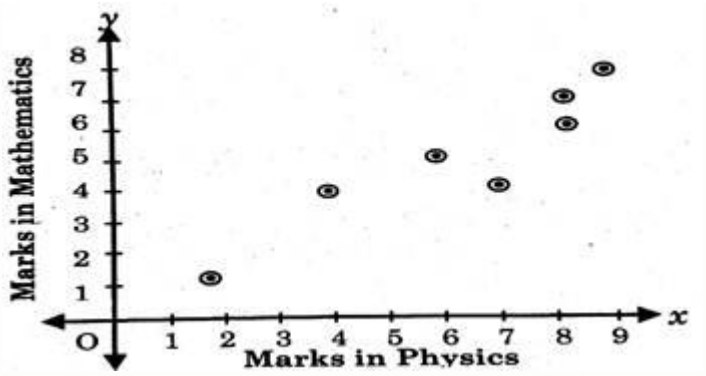
∴ There is positive correlation

2) For the following data of marks of 7 students in physics (x) and Mathematics (y), draw scatter diagram and state the type of correlation.

x	8	6	2	4	7	8	9
y	6	5	1	4	4	7	8

Sol:

Physics (x)	8	6	2	4	7	8	9
Mathematics (y)	6	5	1	4	4	7	8



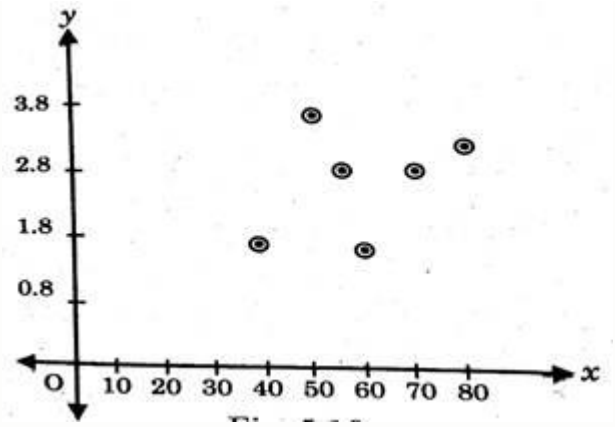
∴ There is positive correlation

3) Draw scatter diagram for the data given below. Is there any correlation between Aptitude score and Grade points?

Aptitude score	40	50	55	60	70	80
Grade points	1.8	3.8	2.8	1.7	2.8	3.2

Sol:

Aptitude score(x)	40	50	55	60	70	80
Grade points (y)	1.8	3.8	2.8	1.7	2.8	3.2



∴ There is positive correlation

4) Find correlation coefficient between x and y series for the following data.

$n = 15, \bar{x} = 25, \bar{y} = 18, \sigma_x = 3.01, \sigma_y = 3.03$
$\Sigma(x_i - \bar{x})\,(y_i - \bar{y}) = 122$

Sol: Given: $n = 15, \bar{x} = 25, \bar{y} = 18, \sigma_x = 3.01, \sigma_y = 3.03, \Sigma(x_i - \bar{x})\,(y_i - \bar{y}) = 122$
To find: r
$\text{cov}\,(x,y) = \frac{1}{n}\Sigma(x_i - \bar{x})\,(y_i - \bar{y})$

$= \frac{122}{15}$
$= 8.133$
Karl Pearson's coefficient of correlation,

$r = \frac{\text{cov}\,(x,y)}{\sigma_x \sigma_y}$

$= \frac{8.133}{3.01 \times 3.03}$

$= 0.892$

$\therefore r = 0.892$

5) The correlation coefficient between two variables x and y is 0.48. The covariance is 36 and the variance of x is 16. Find the standard deviation of y.

Sol: Given: $r = 0.48, \text{cov}\,(x,y) = 36, v_x = 16,$

To find: σ_y

$v_x = 16$

$\therefore \quad \sigma_x = \sqrt{v_x} \quad = \sqrt{16} \quad = 4$

Karl Pearson's coefficient of correlation,

$r = \frac{\text{cov}\,(x,y)}{\sigma_x \sigma_y}$

$\therefore 0.48 = \frac{36}{4 \times \sigma_y}$

$\therefore \sigma_y = \frac{36}{4 \times 0.48}$

$\therefore \sigma_y = 0.892$

6) In the following data one of the value of y is missing. Arithmetic means of x and y series are 6 and 8 respectively.

($\sqrt{2} = 1.4142$)

x	6	2	10	4	8
y	9	11	?	8	7

(I) Estimate missing observation

(II) Calculate correlation coefficient

Sol: Let the missing observation be a

x	6	2	10	4	8
y	9	11	a	8	7

$\bar{x} = 6, \quad \bar{y} = 8, \quad n = 5$

$\bar{y} = \frac{\Sigma y}{n} = \frac{35+a}{5}$

$\therefore 8 = \frac{35+a}{5}$

$\therefore a + 35 = 40$

$\therefore a = 5$

$\therefore$ Missing frequency $= 5$

x	y	x_2	y_2	xy
6	9	36	81	54
2	11	4	121	22
10	5	100	25	50
4	8	16	64	32
8	7	64	49	56
30	40	220	340	214

Here, $\Sigma x = 30$; $\Sigma y = 40$

$\Sigma x2 = 220$; $\Sigma y2 = 320$

$\Sigma xy = 214$; $n = 5$

∴ Karl Pearson's coefficient of correlation,

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x2 - (\Sigma x)^2}\sqrt{n\Sigma y2 - (\Sigma y)^2}}$$

$$= \frac{5(214) - 30(40)}{\sqrt{5(220) - (30)^2}\sqrt{5(340) - (40)^2}}$$

$$= \frac{-130}{\sqrt{200}\sqrt{100}}$$

$$= r = -$$

∴ = 0.919

∴ There is high degree negative correlation.

7) Find correlation coefficient from the following data, [Given = $\sqrt{3}$ = 1.732]

x	3	6	2	9	5
y	4	5	8	6	7

Sol:

x	y	x_2	y_2	xy
3	4	9	16	12
6	5	36	25	30
2	8	4	64	16
9	6	81	36	54
5	7	25	49	
25	30	155	190	147

Here, $\Sigma x = 25$; $\Sigma y = 30$

$\Sigma x2 = 155$; $\Sigma y2 = 190$

$\Sigma xy = 147$; $n = 5$

∴ Karl Pearson's coefficient of correlation,

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x2-(\sum x)^2}\sqrt{n\sum y2-(\sum y)^2}}$$

$$= \frac{5(147)-25(30)}{\sqrt{5(155)-(25)^2}\sqrt{5(190)-(30)^2}}$$

$$= \frac{-15}{\sqrt{150}\sqrt{50}}$$

$\therefore r = -0.173$

$\therefore$ There is high degree negative correlation.

8) Correlation coefficient between x y 0.3 and their covariance3 is 12. The variance of x is 9, Find the standard deviation of y.

Sol: Given: $r = 0.3$, $\text{cov}(x,y) = 12$, $v_x = 9$,
To find: σ_y

$v_x = 9 \quad \sigma_x = 3$

Karl Pearson's coefficient of correlation,

$$r = \frac{\text{cov}(x,y)}{\sigma_x \cdot \sigma_y}$$

$$\therefore 0.3 = \frac{12}{3 \times \sigma_y}$$

$$\therefore \sigma_y = \frac{12}{3 \times 0.3}$$

$\therefore \sigma_y = 13.33$