Least Square Method and Best Fit

Exercise 9(A)

1. Find the least-squares lines (regression line) for the following data set. Also show the data in a scatter diagram.

a)

x	2	4	6	8	10	12	14
y	2	5	6	7	10	11	15

Solution

Let the line of regression ~~y~~ of y on x be
$y = a + bx$ —— (I) where ~~a~~ b and ~~b~~ a are ~~obtai~~ constants and can be calculated by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2} \quad \text{—— (II)}$$

$$a = \frac{\Sigma y - b\Sigma x}{n} \quad \text{—— (III)}$$

CALCULATION TABLE

	x	y	x^2	xy
	2	2	4	4
	4	5	16	20
	6	6	36	36
	8	7	64	56
	10	10	100	100
	12	11	144	132
	14	15	196	210
$n = 7$	$\Sigma x = 56$	$\Sigma y = 56$	$\Sigma x^2 = 560$	$\Sigma xy = 558$

Substituting the values of n, Σx, Σy, Σx^2 & Σxy in (ii) & (iii), we get,

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2}$$

$$= \frac{56 \times 56 - 7 \times 558}{(56)^2 - 7 \times 560}$$

$$= \frac{3136 - 3906}{3136 - 3920}$$

$$\therefore b = 0.98$$

Also,

$$a = \frac{\Sigma y - b\Sigma x}{n}$$

$$= \frac{56 - 0.98 \times 56}{7}$$

$$= \frac{56 - 54.88}{7}$$

$$\therefore a = 0.16$$

Now, the regression line equation (i) becomes,

$$y = a + bx \text{ —— (i)}$$

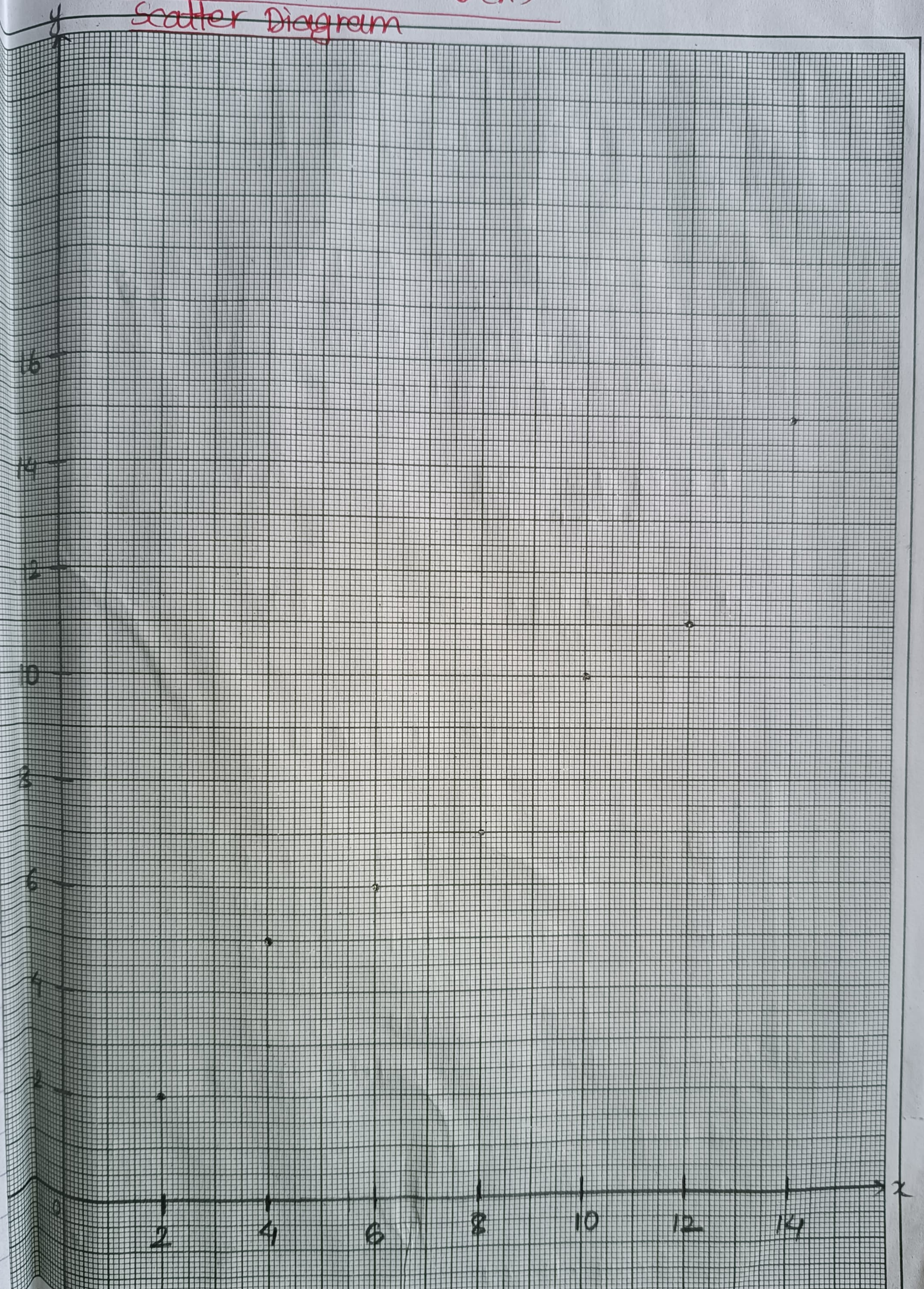
$$= 0.16 + 0.98 \times x$$

$$\therefore y = 0.16 + 0.98x$$

Therefore the regression line equation is
$y = 0.98x + 0.16$

Question Number 1(A)

Scatter Diagram



b)

x	60	63	72	75	80
y	125	110	130	135	150

Solution

let the line of regression of y on x be ~~$y = ax + b$~~ $y = a + bx$ ——— (I) where ~~a~~ b and ~~b~~ a are constants and can be calculated by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2} \quad \text{———} \quad (ii)$$

$$a = \frac{\Sigma y - b \Sigma x}{n} \quad \text{———} \quad (iii)$$

CALCULATION TABLE

	x	y	x^2	xy
	60	125	3600	7500
	63	110	3969	6930
	72	130	5184	9360
	75	135	5625	10125
	80	150	6400	12000
$n = 5$	$\Sigma x = 350$	$\Sigma y = 650$	$\Sigma x^2 = 24778$	$\Sigma xy = 45915$

Substituting the values of n, Σx, Σy, Σx^2 and Σxy in (ii) & (iii) we get,

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2}$$

$$= \frac{350 \times 650 - 5 \times 45915}{(350)^2 - 5 \times 24778}$$

$$= \frac{227500 - 229575}{122500 - 123890}$$

$\therefore b = 1.49$

Also,

$$a = \frac{\Sigma y - b \Sigma x}{n}$$

$$= \frac{650 - 1.49 \times 350}{5}$$

$$= \frac{650 - 521.5}{5}$$

$\therefore a = 25.7$

Now, the regression line equation ① becomes,

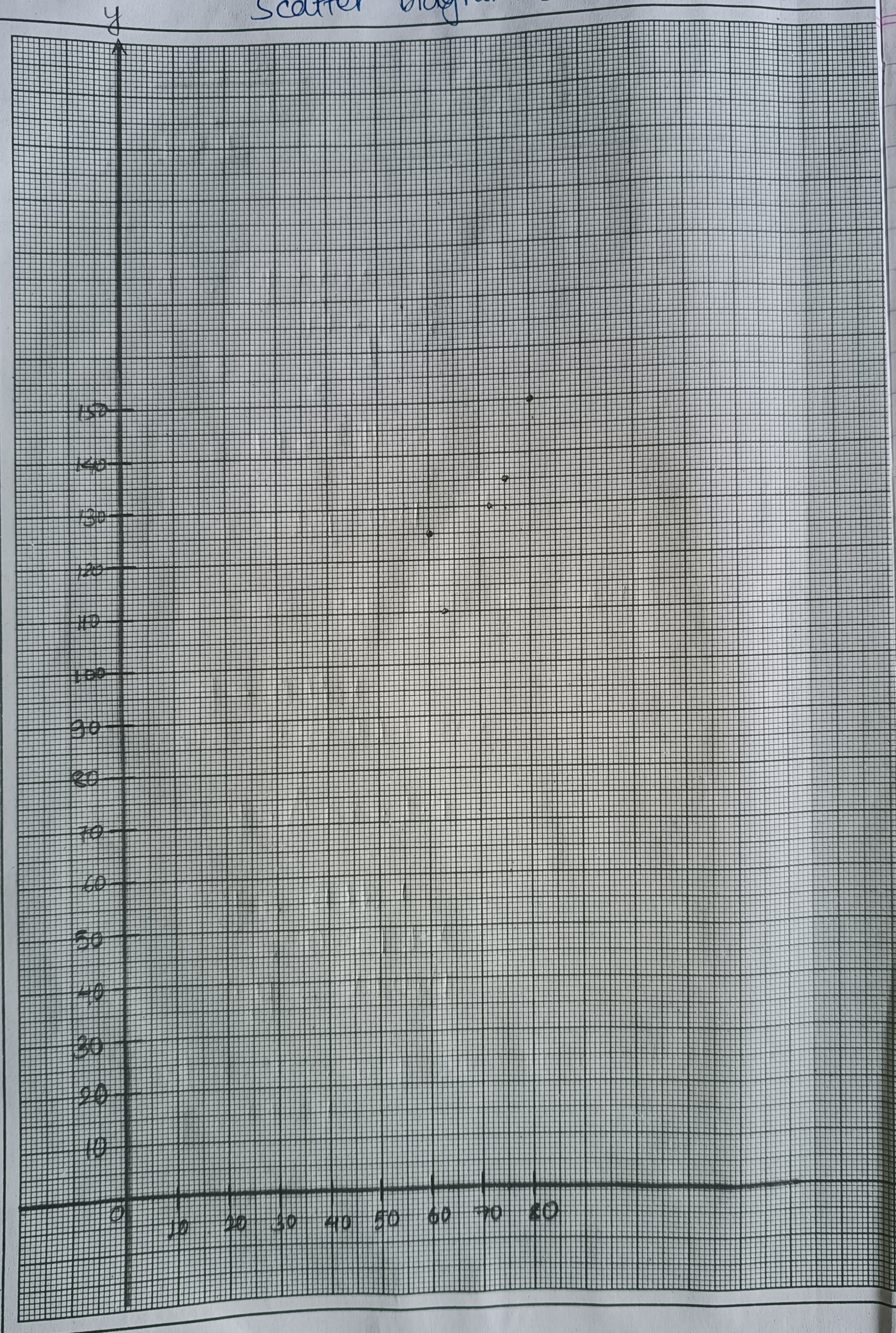
$y = a + bx$ — ①

$= 25.7 + 1.49x$

$\therefore y = 1.49x + 25.7$

Therefore, the regression line equation is $y = 1.49x + 25.7$

Scatter Diagram (1 B)



c)

x	1	2	3	4	5
y	1	3	4	3	6

Solution

Let the line of regression of y on x be $y = a+bx$ — (I) where b~~a~~ and ~~b~~a are constants and can be calculated by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2} \quad \text{——— (II)}$$

$$a = \frac{\Sigma y - b\Sigma x}{n} \quad \text{——— (III)}$$

CALCULATION TABLE

	x	y	x^2	xy
	1	1	1	1
	2	3	4	6
	3	4	9	12
	4	3	16	12
~~A=5~~	5	6	25	30
$n=5$	$\Sigma x = 15$	$\Sigma y = 17$	$\Sigma x^2 = 55$	$\Sigma xy = 61$

Substituting the values of n, Σx, Σy, Σx^2 and Σxy in (II) & (III) we ~~b~~ get,

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2}$$

$$= \frac{15 \times 17 - 5 \times 61}{(15)^2 - 5 \times 55}$$

$$= \frac{255 - 305}{225 - 275}$$

$$= \cancel{1.6}\ 1$$

Also,

$$a = \frac{\Sigma y - b\Sigma x}{n}$$

$$= \frac{17 - \cancel{1.6}1 \times 15}{5}$$

$$= \frac{17 - \cancel{24}15}{5}$$

$$= \cancel{1.4}\ 0.4$$

Now, the regression line equation ① becomes,

$$y = a + bx$$

$$= -1.4 + 1.6 \times x$$

$$\therefore y = 1.6x - 1.4$$

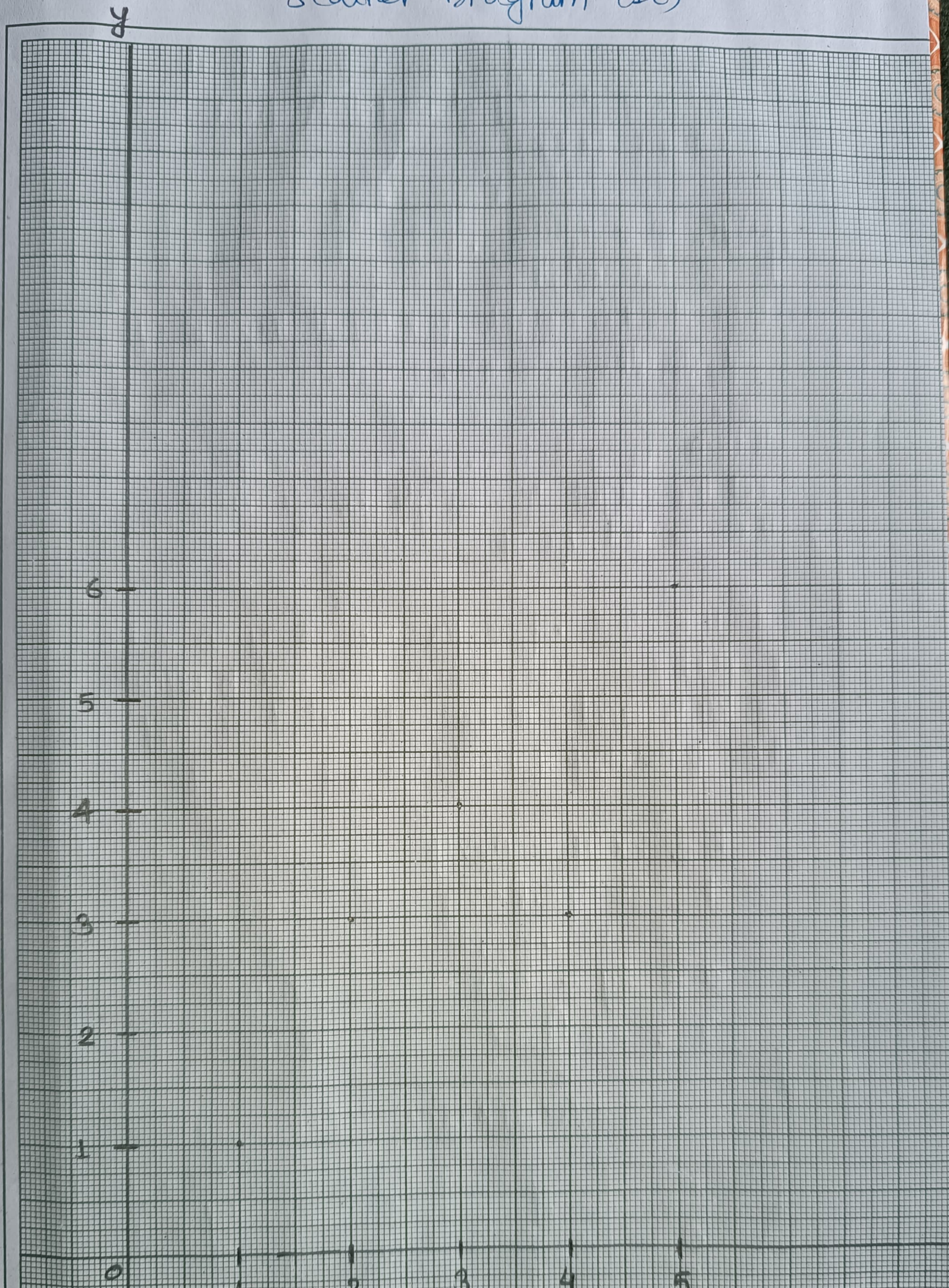
$$y = a + bx$$

$$= 0.4 + 1x$$

$$\therefore y = x + 0.4$$

Therefore, the regression line equation is $y = x + 0.4$

Scatter Diagram (se)



2. The following table gives the normal weight of a baby during the first six-month of life.

Age in month	0	2	3	5	6
weight in (kg)	5	7	8	10	12

Estimate the weight of baby at the age of 4 months

Solution

Let the age in month be x and the weight be y.

Let the line of regression of y on x be $y = a + bx$ —— (I) where ~~b~~ and ~~b~~ a are constants and can be calculated ~~as~~ by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2} \quad \text{—— (II)}$$

$$a = \frac{\Sigma y - b \Sigma x}{n} \quad \text{—— (III)}$$

CALCULATION TABLE

	x	y	x^2	xy
	0	5	0	0
	2	7	4	14
	3	8	9	24
	5	10	25	50
	6	12	36	72
$n = 5$	$\Sigma x = 16$	$\Sigma y = 42$	$\Sigma x^2 = 74$	$\Sigma xy = 160$

Substituting the value of $n, \Sigma x, \Sigma y, \Sigma x^2$ and Σxy in (ii) & (iii), we get,

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \cdot \Sigma x^2}$$

$$= \frac{16 \times 42 - 5 \times 160}{(16)^2 - 5 \times 74}$$

$$= \frac{672 - 800}{256 - 370}$$

$\therefore b = 1.12$

Also,

$$a = \frac{\Sigma y - b \Sigma x}{n}$$

$$= \frac{42 - 1.12 \times 16}{5}$$

$\therefore a = 4.81$

Now the regression line (i) becomes,

$$y = 1.12x + 4.81$$

When $x = 4$,

$$y = 1.12 \times 4 + 4.81$$
$$= 4.48 + 4.81$$

$\therefore y = 9.29$

Therefore the weight of baby at the age of 4 months is 9.29 kg.

3. The following table shows the years of service and income of worker of a factory.

x	2	5	8	6	10	11	7
y	6	5	9	7	11	10	8

a) Determine the equation of the least-square lines for these data.

b) Draw a scatter diagram.

c) Estimate the income of worker who has served for 12 years.

Solution

Let the trend line equation be $y = a + bx$ — (I) where b and a are constants and can be calculated by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2} \quad \text{—— (II)} \ \&$$

$$a = \frac{\Sigma y - b\Sigma x}{n} \quad \text{—— (III)}$$

CALCULATION TABLE

	x	y	x^2	xy
	2	6	4	12
	5	5	25	25
	8	9	64	72
	6	7	36	42
	10	11	100	110
	11	10	121	110
	7	8	49	56
$n=7$	$\Sigma x = 49$	$\Sigma y = 56$	$\Sigma x^2 = 399$	$\Sigma xy = 427$

Substituting the values of n, Σx, Σy, Σx^2 and Σxy in (ii) & (iii), we get,

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2}$$

$$= \frac{49 \times 56 - 7 \times 427}{(49)^2 - 7 \times 399}$$

$$= \frac{2744 - 2989}{2401 - 2793}$$

$$= 0.625$$

Also,

$$a = \frac{\Sigma y - b \Sigma x}{n}$$

$$= \frac{56 - 0.625 \times 49}{7}$$

$$= 3.625$$

a) Now the trend line equation ① becomes,

$$y = a + bx$$
$$= 3.625 + 0.625 \times x$$
$$\therefore y = 0.625x + 3.625$$

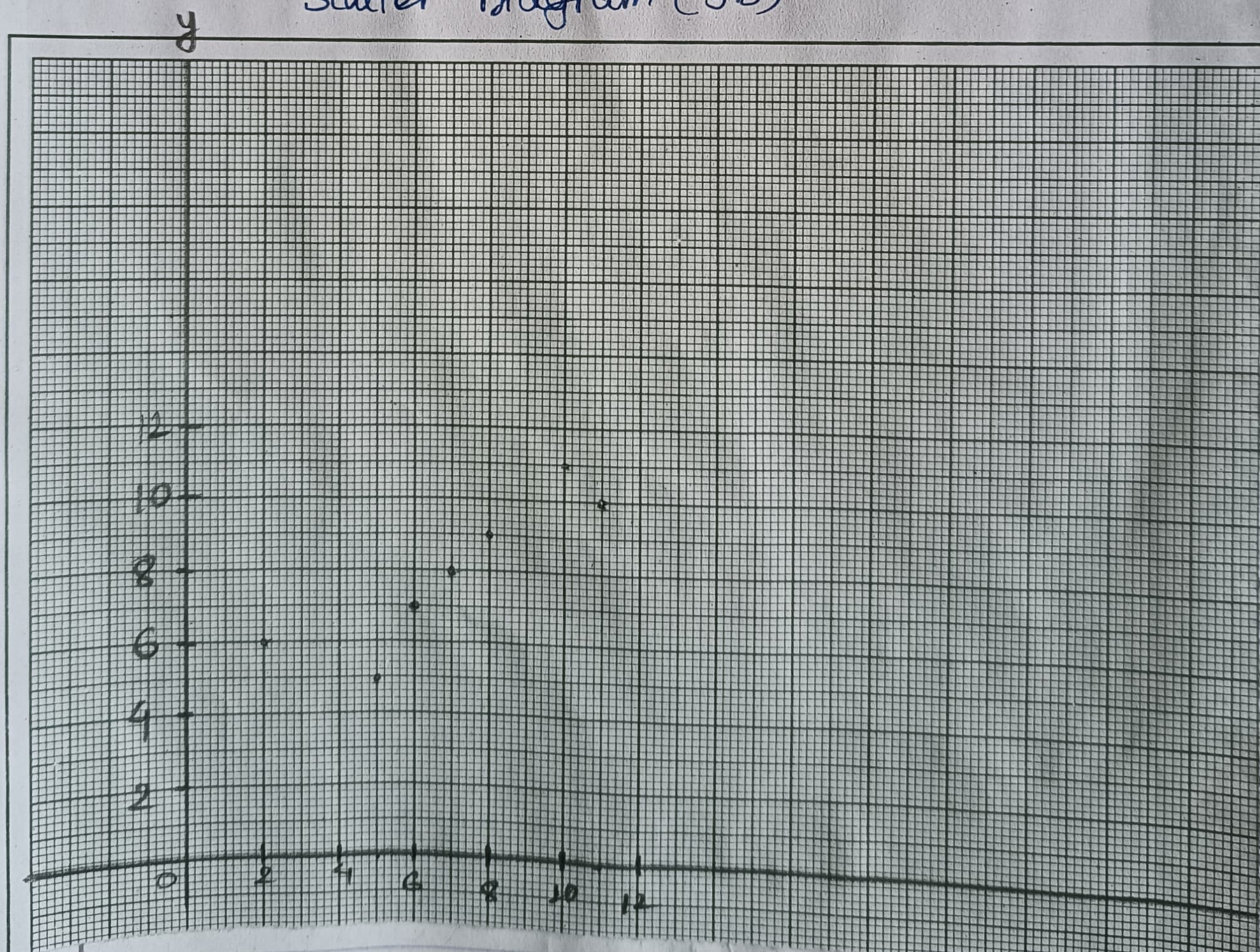
c) When $x = 12$,

$$y = 0.625 \times 12 + 3.625$$
$$= 7.5 + 3.625$$
$$\therefore y = 11.125$$

Therefore the income of the worker who has served for 12 years is 11.125.

b)

Scatter Diagram (9B)



4. Following are the data relating to the annual profit of the company to its annual advertising expenditure in lakh rupees.

Advertising expenditure in lakh	20	25	28	32	38	45	50
Profit in lakh Rs.	40	50	65	90	115	130	180

a) Determine the equation of the least square method.

b) Predict the profit of the company when advertising expenditure is Rs. 75 lakh.

c) Draw a scatter diagram.

Solution

Let advertising expenditure be x & profit be y.

Let the trend line equation be $y = a + bx$ — (I)

Where b and a are constants and can be calculated by the following formulae:

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2} \text{ — (II) \&}$$

$$a = \frac{\Sigma y - b\Sigma x}{n} \text{ — (III)}$$

CALCULATION TABLE

x	y	x^2	xy
20	40	400	800
25	50	625	1250
28	65	784	1820
32	90	1024	2880
38	115	1444	4370 ~~4870~~
45	130	2025	5850
50	180	2500	9000
$n=7$ $\Sigma x = 238$	$\Sigma y = 670$	$\Sigma x^2 = 8802$	$\Sigma xy = 25970$

Substituting the value of n, Σx, Σy, Σx^2 and Σxy in (i) & (ii), we get,

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2}$$

$$= \frac{238 \times 670 - 7 \times 25970}{(238)^2 - 7 \times 8802}$$

$$= \frac{-22330}{-4970}$$

$$= 4.4929$$

Also,

$$a = \frac{\Sigma y - \cancel{\Sigma bx}\, b\Sigma x}{n}$$

$$= \frac{670 - 4.4929 \times 238}{7}$$

$$= \frac{670 - 1069.3102}{7}$$

$$= -57.04$$

Now the trend line (i) becomes,

$$y = a + bx$$

$$\cancel{a} = -57.04 + 4.4929 \times x$$

$$\therefore y = 4.4929x - 57.04$$

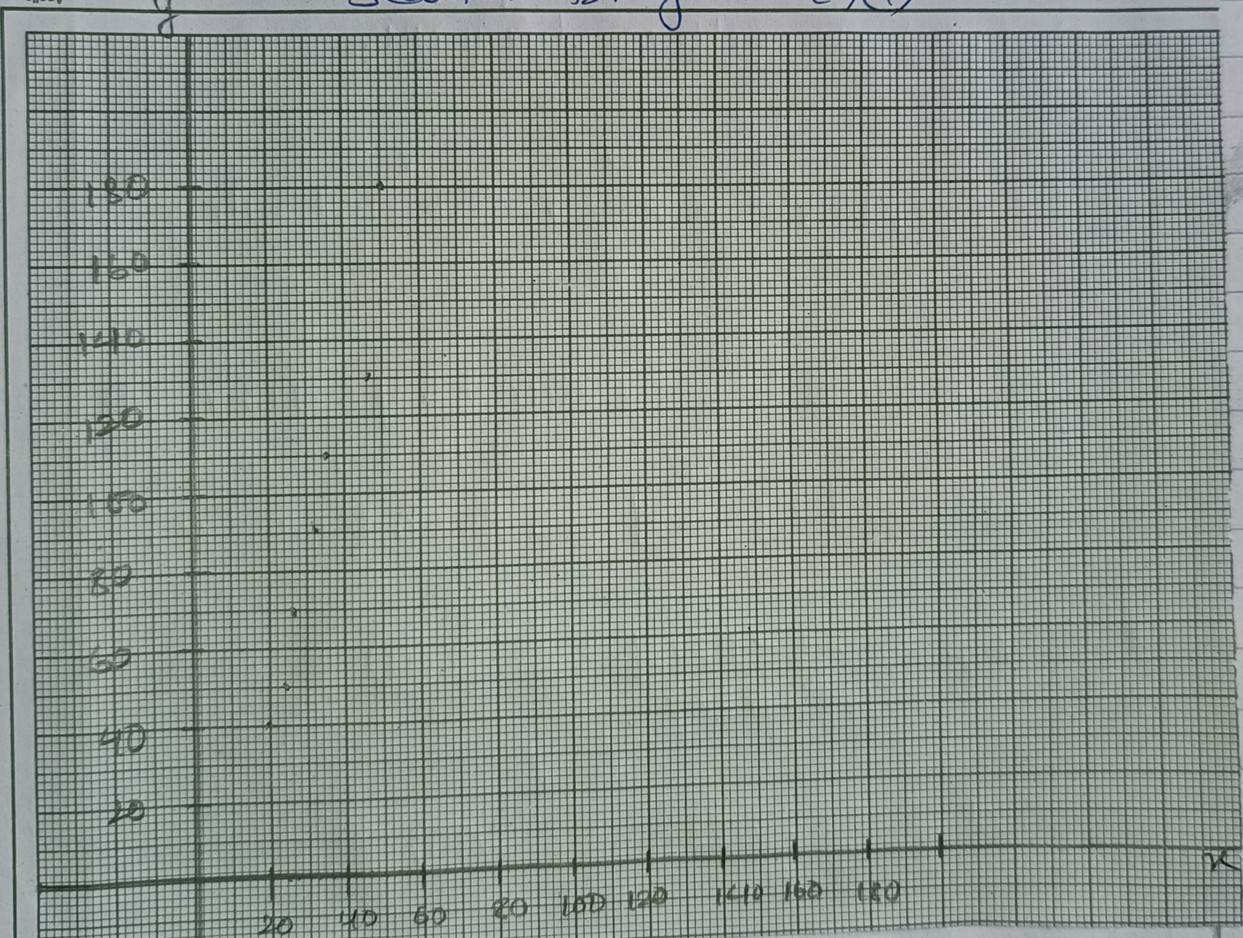
when $x = 75$,

$y = 4.4929 \times 75 - 57.04$

$\therefore y = 279.92$

Therefore the profit of the company is 279.92 lakh when advertising expenditure is Rs. 75 lakh.

Scatter Diagram (4a)



5. The table gives advertisement expenditure in '000' & sales of heater in each units.

Expenditure	12	8	10	4	9	5	14
Sales	9	5	8	5	10	6	12

a) Determine the equation of least square line for these data.

b) Estimate sales when expenditure is Rs. 20,000.

c) Draw a scatter diagram.

Solution

Let the expenditure be x, ~~so~~ and sales be y.

Let the trend line equation be $y = a + bx$ — (I)
Where b and a are constants and can be calculated by using the given formulae:

$$b = \frac{\Sigma x \Sigma y - n \Sigma xy}{(\Sigma x)^2 - n \Sigma x^2} \text{ —— (II)}$$

$$a = \frac{\Sigma y - b \Sigma x}{n} \text{ —— (III)}$$

CALCULATION TABLE

x	y	x^2	xy
12	9	144	108
6	5	36	30
10	8	100	80
4	5	16	20
9	10	81	90
5	6	25	30
14	12	196	168
$n=7$ $\Sigma x = 60$	$\Sigma y = 55$	$\Sigma x^2 = 598$	$\Sigma xy = 528$

Substituting the value of n, Σx, Σy, Σx^2 & Σxy in (II) & (III) we get,

$$b = \frac{\Sigma x \Sigma y - n\Sigma xy}{(\Sigma x)^2 - n\Sigma x^2}$$

$$= \frac{60 \times 55 - 7 \times 526}{(60)^2 - 7 \times 598}$$

$$= \frac{-382}{-586}$$

$$= 0.65$$

Also,

$$a = \frac{\Sigma y - b\Sigma x}{n}$$

$$= \frac{55 - 0.65 \times 60}{7}$$

$$= 2.28$$

Trend line eqn (1) becomes,

$y = a + bx$

$y = 2.28 + 0.65x$

$\therefore y = 0.65x + 2.28$

~~when $x = 20{,}000$~~

~~$y = 0.65 \times 20000 + 2.28$~~

~~$= 13000 + 2.28$~~

When $x = 20$,

$y = 0.65 \times 20 + 2.28$

$= 13 + 2.28$

$\therefore y = 15.28$.

SCATTER DIAGRAM