Linear Regression Activity: Predicting Exam Scores

We want to predict a student's **Exam Score** (y) based on the number of **Hours Studied** (x).

Data

Student	Hours Studied (x)	Exam Score (y)
1	1	52
2	2	57
3	3	61
4	4	65
5	5	70

Task

We want to fit a **linear regression line** of the form:

$$y = mx + b$$

A new student studied **6 hours**. We want to predict the **Exam Score** using the regression equation.

Student	Hours Studied (x)	Exam Score (y)	xy	x^2
1	1	52	52	1
2	2	57	114	4
3	3	61	183	9
4	4	65	260	16
5	5	70	350	25

Summation Values

- **$\sum x$** = 1 + 2 + 3 + 4 + 5 = **15**
- **$\sum y$** = 52 + 57 + 61 + 65 + 70 = **305**
- **$\sum xy$** = (1×52) + (2×57) + (3×61) + (4×65) + (5×70)
  = 52 + 114 + 183 + 260 + 350 = **959**
- **$\sum x^2$** = $1^2 + 2^2 + 3^2 + 4^2 + 5^2$ = 1 + 4 + 9 + 16 + 25 = **55**

2. Compute the Slope m (5 points)

$$m = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$m = \ ?$$

$\Sigma x = 15 \quad \Sigma yy = 59$

$\Sigma y = 305 \quad \Sigma x^2 = 55$

$$\frac{5(959) - (15)(305)}{5(55) - (15)^2}$$

$m = 4.4$

---

4. Regression Equation (5 points)

Write the regression line:

$$y = mx + b$$

$$y = \ ?$$

4. REGRESSION LINE

y = 47.8

$\Sigma x = 15 \quad n = 5$

$\Sigma y = 305 \quad m = 4.4$

$$b = \frac{305 - (4.4)(15)}{5}$$

$b = 47.8$

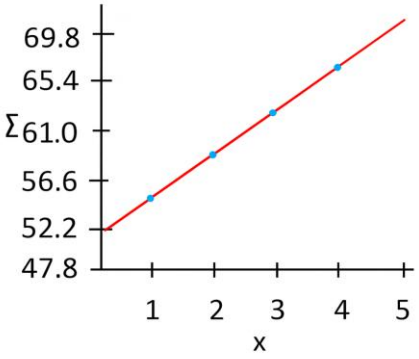
4. Regression Equation

Write the regression line:

m = 4.4 b = 47.8

y = (4.4x) + 47.8

y = 4.4x + 47.8

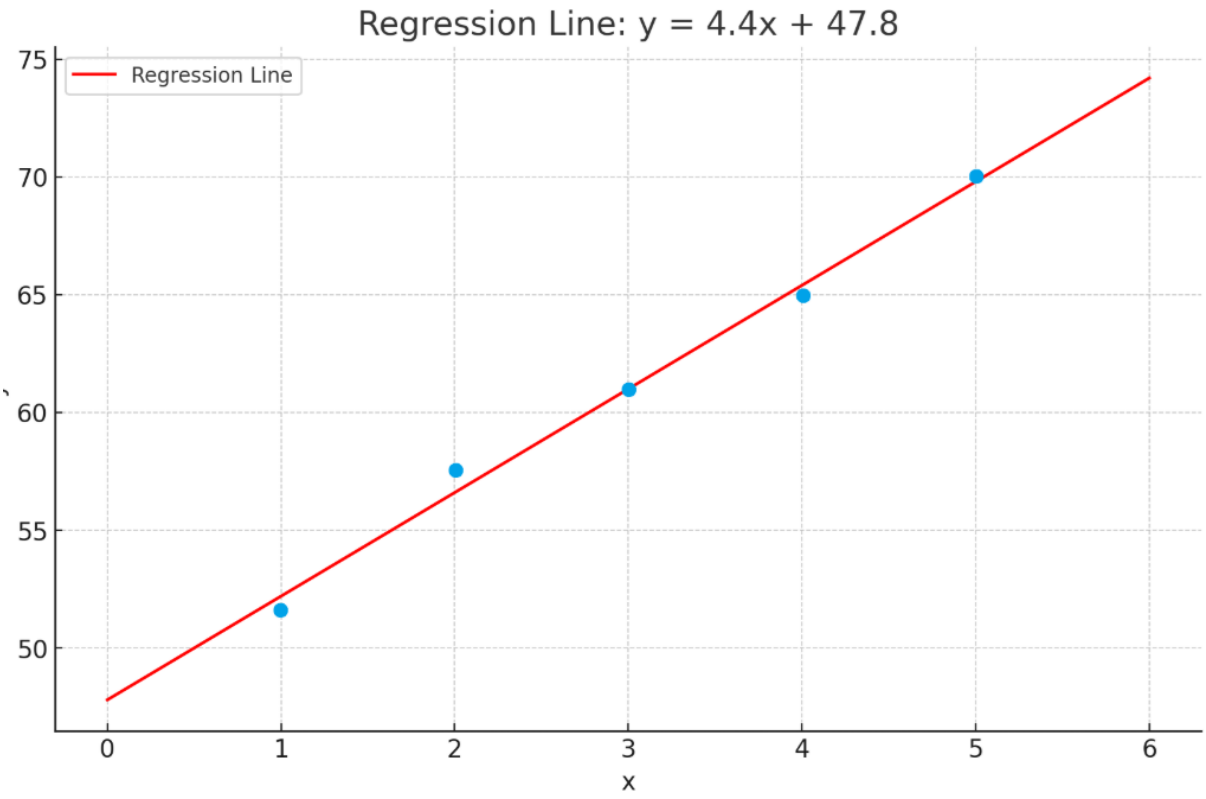


5. Draw the regression line using a scatter plot (10 points)

- Calculate $y_{predict}$ for each data points
- Draw a regression line using $y_{predict}$
- Use a circle 🔵 for all data points
- Use a red line for the regression line

Hours Studied (x)	Exam Score (y)
1	52
2	57
3	61
4	65
5	70

$y_{predict}$	$y_{predict}$
52.2	$y = 4.4(1) + 47.8 = 52.2$
56.6	$y = 4.4(2) + 47.8 = 56.6$
61.0	$y = 4.4(3) + 47.8 = 61.0$
65.4	$y = 4.4(4) + 47.8 = 69.8$



5. Draw the regression line using a scatter plot (10 points)

- Calculate $y_{predict}$ for each data points
- Draw a regression line using $y_{predict}$
- Use a circle 🔵 for all data points
- Use a red line for the regression line

Data

Student	Hours Studied (x)	Exam Score (y)
1	1	52
2	2	57
3	3	61
4	4	65
5	5	70

$y_{predict}$	$y_{predict}$
52.2	y = 4.4(1) + 47.8 = 52.2
56.6	y = 4.4(2) + 47.8 = 56.6
61.0	y = 4.4(3) + 47.8 = 61.0
65.4	y = 4.4(4) + 47.8 = 65.4
69.8	y = 4.4(5) + 47.8 = 69.8

- Get y using this formula :

$$y = \frac{\sum y}{41}$$

- Get 657 using this formula :

$$SST = \sum (y - \hat{y})^2$$

$$SST = \sum (y - \hat{y})$$

= 81 = 16 + 0 + 16 + 81 = 104

Student	Hours Studied (x)	Earm Score (x)	Mears (x)	$\hat{y} - \bar{y}$	$(y - \hat{y})$
1	1	52	61	8	91
2	2	37	61	4	16
3	3	61	61	0	0
4	4	60	61	4	16
5	5	70	61	9	31

Exam Srove 54	$x_i - 3$	$(x_i - \bar{x})^1$
56	52 - 61 = 1-8| = 8	9^1 = 0 x 0 = 81
57	57 - 61 = 1-4| = 8	4^9 = 4 x 4 = 16
61	61 - 61 = 0	9^9 = 0 x 0 = 0
62	65 - 61 = 4	4^5 = 4 x 4 = 16
70	70 - 61 = 9	9^9 = 9 x 9 = 81

8. Compute R^2 (20 points)

- Get R^2 using this formula :

$$R^2 = 1 - \frac{SSE}{SST}$$

$$R^2 = ?$$

SSE = 0.4

SST = 194

$$R = 1 - \frac{0.4}{194} = 0.99$$

m = 4.4, b = 47.8

9. Prediction (1 point)

Use your equation to predict the exam score for a student who studied 6 hours.

$m = 4.4, b = 47.8$

$y = m(6) + b$

$= 74.2$