

ASSIGNMENT #2

LINEAR REGRESSION ACTIVITY: PREDICTING EXAM SCORES

- WE WANT TO PREDICT A STUDENT'S EXAM SCORE (y) BASED ON THE NUMBERS OF HOURS STUDIED (x).

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.

1) FILL IN THE TABLE

- COMPUTE x^2 FOR EACH STUDENT.
- COMPUTE xy FOR EACH STUDENT.
- FIND THE TOTALS: $\sum x$, $\sum y$, $\sum x^2$, AND $\sum xy$.

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
	$\sum x = 15$	$\sum y = 305$	$\sum xy = 959$	$\sum x^2 = 55$

2) COMPUTE THE SLOPE m

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

$$n = 5$$

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

$$m = \frac{220}{50} = 4.4$$

$$= \frac{4795 - 4575}{275 - 225}$$

3) COMPUTE THE INTERCEPT b

$$b = \frac{\sum y - m \sum x}{n} = \frac{305 - (4.4)(15)}{5} = \frac{289}{5} = 47.8$$

A) REGRESSION EQUATION

$$y = 4.4x + 47.8$$

S) DRAW THE REGRESSION LINE USING A SCATTER PLOT

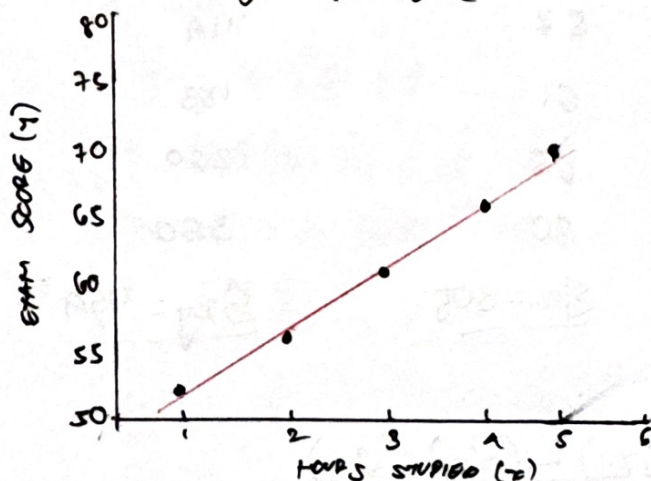
- CALCULATE y_{predict} FOR EACH DATA POINTS
- DRAW A REGRESSION LINE USING y_{predict}
- USE A CIRCLE ● FOR ALL DATA POINTS
- USE A RED LINE FOR THE REGRESSION LINE

STUDENT	HOURS STUDIED (x)	EXAM SCORE (y)	PREDICTED EXAM SCORE (y_{predict})
1	1	52	52.2
2	2	57	56.6
3	3	61	61.0
4	4	65	65.4
5	5	70	69.8

SOL:

$$\begin{aligned}
 1) y &= 4.4(1) + 47.8 = 52.2 & 3) y &= 4.4(3) + 47.8 = 61.0 & 5) y &= 4.4(5) + 47.8 = 69.8 \\
 2) y &= 4.4(2) + 47.8 = 56.6 & 4) y &= 4.4(4) + 47.8 = 65.4
 \end{aligned}$$

TO ~~CALCULATE~~ THE ~~SUM~~ OF ~~SQUARED~~ ~~ERRORS~~



4) CALCULATE THE SUM OF SQUARED ERRORS

STUDENT	HOURS STUDIED (x)	EXAM SCORE (y)	PREDICTED EXAM SCORE (y_{predict})	$y_i - y_{\text{predict}}$	$(y_i - y_{\text{predict}})^2$
1	1	52	52.2	-0.2	0.04
2	2	57	56.6	0.4	0.16
3	3	61	61.0	0.0	0.00
4	4	65	65.4	-0.4	0.16
5	5	70	69.8	0.2	0.04

$$SSE = 0.40$$

7) CALCULATE THE SUM OF SQUARED TOTAL

- GET \bar{y} USING THIS FORMULA: $\bar{y} = \frac{\sum y_i}{n} = \frac{305}{5} = 61$

- GET SST USING THIS FORMULA: $SST = \sum (y_i - \bar{y})^2$

STUDENT	x	y	\bar{y}	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
1	1	52	61	-9	81
2	2	57	61	-4	16
3	3	61	61	0	0
4	4	65	61	4	16
5	5	70	61	9	81

$$SST = 194 //$$

8) COMPUTE R^2

- GET R^2 USING THIS FORMULA:

$$R^2 = 1 - \frac{SSR}{SST} = \frac{0.40}{194} \approx 0.00206 //$$

9) PREDICTION

- USE YOUR EQUATION TO PREDICT THE EXAM SCORE FOR A STUDENT WHO STUDIED 6 HOURS

$$y = m(x) + b$$

$$y = 4.4(6) + 47.8$$

$$y = 26.4 + 47.8 = 74.2$$

$$y = 74.2 //$$