



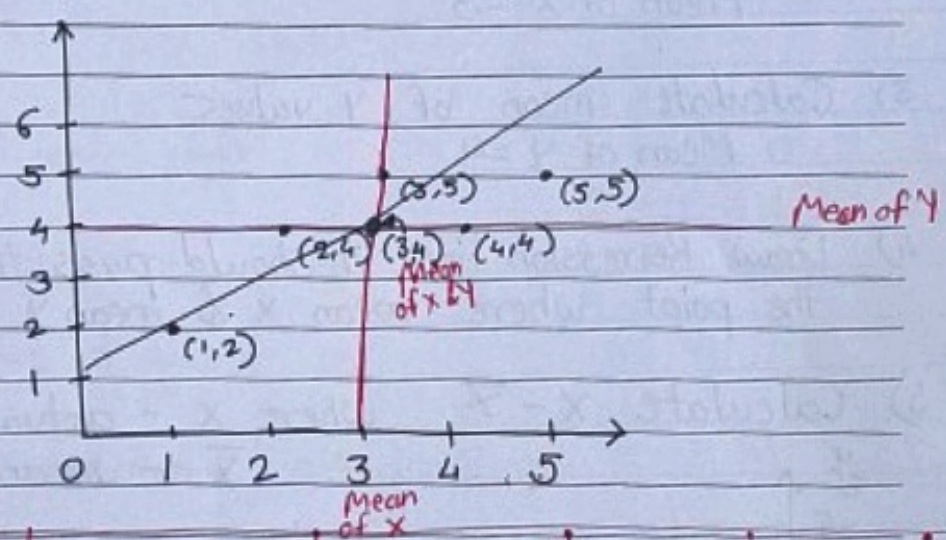
Semester : VI

Subject : CSC601 Data Analytics and Visualization

Academic Year: 2023- 2024

How to calculate linear regression using least square method?

$$\hat{y} = \theta_0 + \theta_1 x$$



Independent Variable	Dependent Variable	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
1	2	$(1-3) = -2$	$(2-4) = -2$	4	+4
2	4	$(2-3) = -1$	$(4-4) = 0$	1	0
3	5	$(3-3) = 0$	$(5-4) = 1$	0	0
4	4	$(4-3) = 1$	$(4-4) = 0$	1	0
5	5	$(5-3) = 2$	$(5-4) = 1$	4	2
Mean 3	4			10	06
				Sum	Sum



Semester : VI Subject : CSC601 Data Analytics and Visualization

Academic Year: 2023- 2024

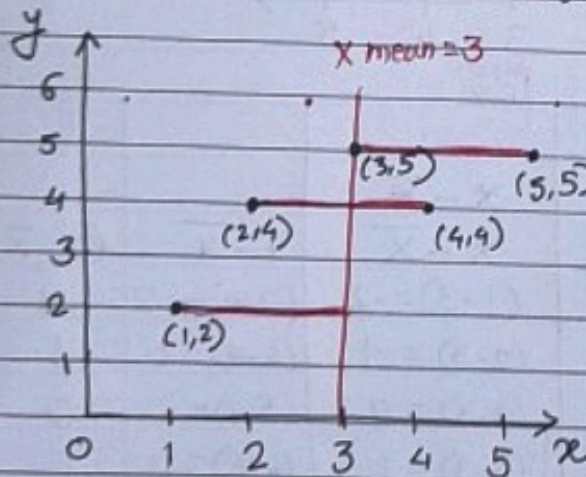
1) Plot the given values

2) Calculate mean of x values
Mean of $x = 3$

3) Calculate mean of y values
Mean of $y = 4$

4) Draw Regression line, it should pass through the point where mean x & mean y intersects

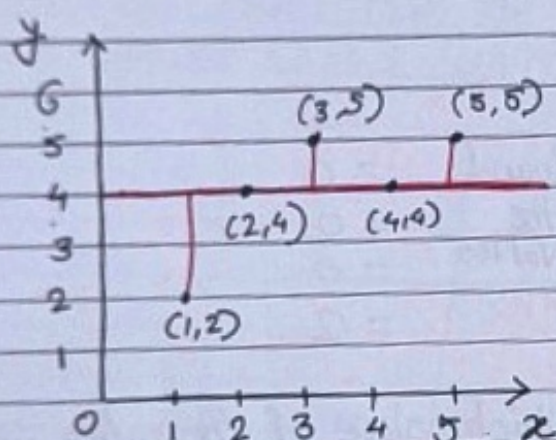
5) Calculate $x - \bar{x}$ where x - actual value
 \bar{x} - mean of x



$x - \bar{x}$
$1 - 3 = -2$
$2 - 3 = -1$
$3 - 3 = 0$
$4 - 3 = 1$
$5 - 3 = 2$



6) Calculate $y - \bar{y}$ where y - actual value
 \bar{y} - mean of y



$y - \bar{y}$
$2 - 4 = -2$
$4 - 4 = 0$
$5 - 4 = 1$
$4 - 4 = 0$
$5 - 4 = 1$

7) The equation $\hat{y} = \theta_0 + \theta_1 x$

Now we will find the value of θ_1

So square the values of $(x - \bar{x})$

$$(x - \bar{x})^2$$

4

1

0

1

4



8) Calculate $(x - \bar{x}) \cdot (y - \bar{y})$

$$(x - \bar{x}) \cdot (y - \bar{y})$$

$-2 \times -2 = -4$	$\left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{Round} \\ \text{the} \\ \text{Values} \end{array}$	$= 4$
$-1 \times 0 = 0$		$= 0$
$0 \times 1 = 0$		$= 0$
$1 \times 0 = 0$		$= 0$
$2 \times 1 = 2$		$= 2$

9) To calculate b_1 which slope of the line use following formula

$$b_1 =$$

1) First sum all the values of $(x - \bar{x})^2$ which is 10 (as denominator of eqⁿ)

2) Now sum all the values of $(x - \bar{x})(y - \bar{y})$ which is 6 (as numerator of the eqⁿ)

$$b_1 = \frac{6}{10} = 0.6 \text{ (slope)}$$

$$b_1 = \frac{(x - \bar{x})(y - \bar{y})}{(x - \bar{x})^2} \text{ slope}$$



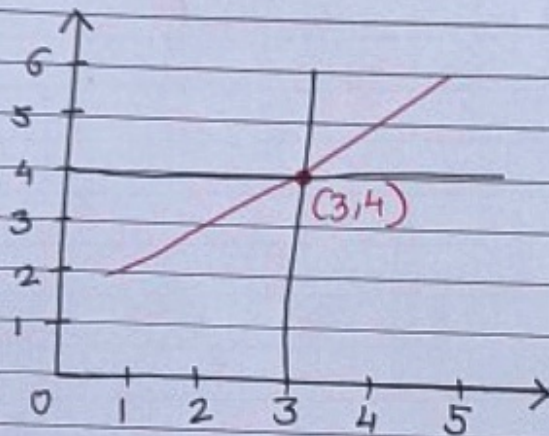
Semester : VI

Subject : CSC601 Data Analytics and Visualization

Academic Year: 2023- 2024

10) Now we need to find value of θ_0

We know that the regression line has to pass point $(3, 4)$ which is mean of X & mean of Y .



$$\hat{y} = \theta_0 + \theta_1 x$$

From this we know that of

$$\hat{y} = 4$$

$$x = 3$$

Also we have calculate θ_1 in previous step $\theta_1 = 0.6$

$$\hat{y} = \theta_0 + \theta_1 x$$

$$4 = \theta_0 + (0.6)(3)$$

$$4 = \theta_0 + 1.8$$

$$\theta_0 = 4 - 1.8$$

$$\theta_0 = 2.2$$

Intercept



Semester : VI

Subject : CSC601 Data Analytics and Visualization

Academic Year: 2023- 2024

$$11) \quad \hat{y} = \theta_0 = 2.2$$

$$\theta_1 = 0.6$$

$$\hat{y} = (2.2) + (0.6)x$$