

Linear Regression

Basic Formula: -

$$y = mx + c$$

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

$$c = y - mx$$

x' = mean of x

y' = mean of y

| x | y | $x - x'$ | $y - y'$ | $(x - x')^2$ | $(x - x')(y - y')$ |
|----------|--------------|----------|----------|---------------|--------------------|
| 2 | 17 | -1 | -9.75 | 1 | 9.75 |
| 4 | 35 | 1 | 8.25 | 1 | 8.25 |
| 5 | 46 | 2 | 19.25 | 4 | 38.5 |
| 1 | 9 | -2 | -17.75 | 4 | 35.5 |
| | | | | | |
| $x' = 3$ | $y' = 26.75$ | | | $\Sigma = 10$ | $\Sigma = 92$ |

$$M = 92/10 = 9.2$$

$$C = 26.75 - (9.2 * 3) = 26.75 - 27.6 = -0.85$$

Best fit line: $y = mx + c$

$$Y = 9.2X + (-0.85)$$

$$Y = 9.2X - 0.85$$

Error: -

$$R^2 = \frac{\sum (y_p - y')^2}{\sum (y - y')^2}$$

y_p = predicted y

y' = mean of y

| x | y | y_p | $y - y'$ | $y_p - y'$ | $(y - y')^2$ | $(y_p - y')^2$ |
|----------|--------------|-------|----------|------------|-------------------|-------------------|
| 2 | 17 | 17.55 | -9.75 | -9.2 | 95.06 | 84.64 |
| 4 | 35 | 35.95 | 8.25 | 9.2 | 68.06 | 84.64 |
| 5 | 46 | 45.15 | 19.25 | 18.4 | 370.56 | 338.56 |
| 1 | 9 | 8.35 | -17.75 | -18.4 | 315.06 | 338.56 |
| | | | | | | |
| $x' = 3$ | $y' = 26.75$ | | | | $\Sigma = 848.72$ | $\Sigma = 846.40$ |

$$R^2 = 846.40 / 848.72 = 0.997$$

$R^2 \propto$ Model Quality