

Linear Regression is a supervised machine learning and statistical technique used to model the relationship between a continuous dependent variable (y) and one or more independent variables.

Comparison: Linear vs Logistic Regression

Feature	Linear Reg	Logistic Reg
Core Purpose	Used to predict a continuous value	Used to classify an observation
Output Type	Continuous value (eg salary temp)	Probability 0 and 1 mapped
Equation/Fraction	Straight line $y = ax + b$	Sigmoid function $p = \frac{1}{1 + e^{-ax + b}}$
Graph Shape	Linear (Straight line)	Sigmoid function is S-shaped

2 Solving the Problem.

x = Years of Experience and y = Salary in k.

$$\sum x = 68, \quad \sum y = 409, \quad \sum x^2 = 729$$
$$\sum xy = 4128$$

No. of observ (n) = 9.

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

$$a = \frac{9(4128) - (68)(409)}{9(729) - (68)^2}$$

$$a = \frac{37152 - 27812}{6561 - 4624} = \frac{9340}{1937}$$

$$a = 4.822$$

Calculating the intercept

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

$$b = \frac{409}{9} - (4.822 \times 68) = \frac{409 - 327.896}{9} \\ = \frac{81.104}{9} = 9.012$$

Regression eqn

$$y = 4.822x + 9.012$$

For every additional year of exp, salary increases by approximately 4822 with a baseline starting salary roughly 9012

Is Linear Regression Best Fit?

To evaluate if linear Regression is an appropriate fit; we can calculate the coefficient of Determination (R^2) which measures the proportion of variance in the dependent variable predictable from the independent variable.

Yes linear Regression is an excellent fit for this data.