It's a best fit line if the sum of squared vertical distances (from the point to the line) is as small as it can be. I.e., find b_0 and b_1 that minimize the sum $f(b_0, b_1)$.

$$f(b_0, b_1) = \sum_{i=1}^{n} \left(y_i - [b_0 + b_1 x_i]\right)^2$$

Where:

- y_i: observed value of the dependent variable for i-th data point. I.e., the label.
- $b_0 + b_1 x_i$: predicted value of y_i, where b_1 is the slope, i.e., the rate of change per 1 unit of x.
- $(y_i - b_0 - b_1 x_i)^2$: error squared for each data point. can be seen as `y_actual - y_predicted`.

We then take the partial derivative with respect to both b_0 and b_1 and equating them to 0 to find the values that minimize the sum. Using the chain rule, we get:

$$\frac{\partial f(b_0, b_1)}{\partial b_0} = \sum 2(y_i - b_0 - b_1\, x_i)\,(-1) = 0 \tag{1}$$

$$\frac{\partial f(b_0, b_1)}{\partial b_1} = \sum 2(y_i - b_0 - b_1 x_i)(-x_i) = 0 \tag{2}$$

—

From (1), we get:

$$\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i) \;=\; 0.$$

Distribute the summation:

$$\sum_{i=1}^{n} y_i \;-\; \sum_{i=1}^{n} b_0 \;-\; \sum_{i=1}^{n} b_1 x_i \;=\; 0.$$

Consider the summation:

$$\sum_{i=1}^{n} b_0.$$

Because b_0 is a constant, we have:

$$b_0 + b_0 + \cdots + b_0 \;=\; n\, b_0.$$

Thus, it follows that

$$\sum_{i=1}^{n} b_0 \;=\; n\, b_0.$$

Because b_0 and b_1 are constants (with respect to the sum over i), we have:

$$\sum_{i=1}^{n} y_i \;-\; b_0 \sum_{i=1}^{n} 1 \;-\; b_1 \sum_{i=1}^{n} x_i \;=\; 0.$$

$$\sum_{i=1}^{n} y_i \;-\; b_0\, n \;-\; b_1 \sum_{i=1}^{n} x_i \;=\; 0.$$

Hence,

$$b_0\, n \;=\; \sum_{i=1}^{n} y_i \;-\; b_1 \sum_{i=1}^{n} x_i.$$

Divide by n:

$$b_0 \;=\; \frac{1}{n} \sum_{i=1}^{n} y_i \;-\; b_1\, \frac{1}{n} \sum_{i=1}^{n} x_i.$$

so we obtain

$$\boxed{b_0 \;=\; \bar{y} \;-\; b_1\, \bar{x}.}$$

Step 2: Using partial derivative of $f(b_0, b_1)$ w.r.t. b_1

Earlier from (2), we got:

$$\frac{\partial f(b_0, b_1)}{\partial b_1} = \sum 2(y_i - b_0 - b_1 x_i)(-x_i) = 0$$

Setting this to zero for the critical point:

$$-2 \sum_{i=1}^{n} x_i\, (y_i - b_0 - b_1 x_i) \;=\; 0.$$

$$\sum_{i=1}^{n} x_i\, (y_i - b_0 - b_1 x_i) \;=\; 0.$$

$$\sum_{i=1}^{n} (x_i y_i - x_i b_0 - b_1 x_i^2) \;=\; 0.$$

Distribute inside the summation:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; \sum_{i=1}^{n} x_i\, b_0 \;-\; \sum_{i=1}^{n} x_i\, (b_1\, x_i) \;=\; 0.$$

Because b_0 and b_1 are constants:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; b_0 \sum_{i=1}^{n} x_i \;-\; b_1 \sum_{i=1}^{n} x_i^2 \;=\; 0.$$

Now substitute $b_0 = \bar{y} - b_1\bar{x}$ (found in the previous step):

$$\sum_{i=1}^{n} x_i\, y_i \;-\; (\bar{y} - b_1\bar{x}) \sum_{i=1}^{n} x_i \;-\; b_1 \sum_{i=1}^{n} x_i^2 \;=\; 0.$$

Distribute the middle term:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; \bar{y} \sum_{i=1}^{n} x_i \;+\; b_1\, \bar{x} \sum_{i=1}^{n} x_i \;-\; b_1 \sum_{i=1}^{n} x_i^2 \;=\; 0.$$

Group together the terms that contain b_1:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; \bar{y} \sum_{i=1}^{n} x_i \;+\; b_1 \Bigl(\bar{x} \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} x_i^2\Bigr) \;=\; 0.$$

Isolate the b_1 terms:

$$b_1 \Bigl(\bar{x} \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} x_i^2\Bigr) \;=\; \bar{y} \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} x_i\, y_i.$$

Rearrange it more cleanly:

$$b_1 \;=\; \frac{\bar{y} \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} x_i\, y_i}{\bar{x} \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} x_i^2}.$$

$$b_1 \;=\; \frac{\sum_{i=1}^{n} \bar{y}x_i - x_i y_i}{\sum_{i=1}^{n} \bar{x}x_i - x_i^2} \;=\; \frac{\sum_{i=1}^{n} x_i(\bar{y} - y_i)}{\sum_{i=1}^{n} x_i(\bar{x} - x_i)}.$$

Properties of Mean-Centered Data

First, note an important property of mean-centered terms:

$$\sum_{i=1}^{n}(x_i - \bar{x}) \;=\; \sum_{i=1}^{n} x_i \;-\; \sum_{i=1}^{n} \bar{x}.$$

Since $\bar{x}$ is a constant,

$$\sum_{i=1}^{n} \bar{x} \;=\; \bar{x} + \bar{x} + \cdots + \bar{x} \;=\; n\,\bar{x}.$$

But by definition of $\bar{x}$, $\sum_{i=1}^{n} x_i = n\bar{x}$ (expand $\bar{x}$ if you want to see it). Hence,

$$\sum_{i=1}^{n}(x_i - \bar{x}) \;=\; (n\,\bar{x}) \;-\; (n\,\bar{x}) \;=\; 0.$$

Similarly,

$$\sum_{i=1}^{n}(y_i - \bar{y}) \;=\; 0.$$

Therefore, these $(x_i - \bar{x})$ and $(y_i - \bar{y})$ terms are “centered” around **zero**.

Additionally,

$$x_i = (x_i - \bar{x}) + \bar{x}$$

Substituting this into the **numerator**, we get:

$$= \sum_{i=1}^{n}[(x_i - \bar{x}) + \bar{x}](\bar{y} - y_i)$$

$$= \sum_{i=1}^{n}(x_i - \bar{x})(\bar{y} - y_i) + \bar{x}(\bar{y} - y_i)$$

$$= \sum_{i=1}^{n}(x_i - \bar{x})(\bar{y} - y_i) + 0$$

Therefore, we get:

$$S_{xy} \;=\; \sum_{i=1}^{n}(x_i - \bar{x})(\bar{y} - y_i)$$

Denominator:

$$S_{xx} = \sum_{i=1}^{n} x_i(\bar{x} - x_i)$$

$$S_{xx} = \sum_{i=1}^{n} [(x_i - \bar{x}) + \bar{x}](\bar{x} - x_i)$$

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Finally,

$$b_1 = \hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}}$$

Numerator: Expand the product

$$S_{xy} \;=\; \sum_{i=1}^{n} \Bigl[x_i y_i \;-\; x_i \bar{y} \;-\; \bar{x}\, y_i \;+\; \bar{x}\, \bar{y} \Bigr].$$

Distribute the summation:

$$S_{xy} \;=\; \sum_{i=1}^{n} x_i\, y_i \;-\; \bar{y} \sum_{i=1}^{n} x_i \;-\; \bar{x} \sum_{i=1}^{n} y_i \;+\; \sum_{i=1}^{n} \bar{x}\, \bar{y}.$$

Expanding the bar's:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; 2\, \frac{(\sum x_i)(\sum y_i)}{n} \;+\; \sum_{i=1}^{n} \bar{x}\, \bar{y}.$$

Since $\bar{x}\bar{y}$ is constant, we get:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; 2\, \frac{(\sum x_i)(\sum y_i)}{n} \;+\; n(\bar{x}\, \bar{y}).$$

Hence, combining everything:

$$\sum_{i=1}^{n} x_i\, y_i \;-\; 2\, \frac{(\sum x_i)(\sum y_i)}{n} \;+\; \frac{(\sum x_i)(\sum y_i)}{n} \;=\; \sum_{i=1}^{n} x_i\, y_i \;-\; \frac{(\sum x_i)(\sum y_i)}{n}.$$

That matches the usual final form. In other words,

$$\boxed{S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \;=\; \sum_{i=1}^{n} x_i\, y_i \;-\; \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n}.}$$

Denominator

We start with:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Expand each term:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i^2 - 2\,x_i\,\bar{x} + \bar{x}^2).$$

Distribute the summation:

$$\sum_{i=1}^{n} x_i^2 \;-\; 2\,\bar{x}\sum_{i=1}^{n} x_i \;+\; \sum_{i=1}^{n} \bar{x}^2.$$

Recall $\sum_{i=1}^{n} x_i = n\,\bar{x}$ and $\sum_{i=1}^{n} \bar{x}^2 = n\,\bar{x}^2$. Hence,

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 \;-\; 2\,\bar{x}\,(n\,\bar{x}) \;+\; n\,\bar{x}^2.$$

$$= \sum_{i=1}^{n} x_i^2 \;-\; 2n\bar{x}^2 \;+\; n\,\bar{x}^2.$$

Thus,

$$= \sum_{i=1}^{n} x_i^2 \;-\; n\left(\frac{\sum x_i}{n}\right)^2.$$

Therefore:

$$\boxed{S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \;=\; \sum_{i=1}^{n} x_i^2 \;-\; \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}.}$$

These forms remove the need to do any calculations with the mean of X or Y.