Prompt:

Consider a set of bivariate data $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$. Show that the line that minimizes the total squared error is given by $y = \hat{a}x + \hat{b}$ where

$$\hat{a} = \frac{n\sum(x_i y_i) - (\sum(x_i))(\sum(y_i))}{n\sum(x_i)^2 - (\sum(x_i))^2}$$

and

$$\hat{b} = \bar{y} - \hat{a}\bar{x}$$

Derivation:
We will consider the equation $E = \sum_{i=1}^{n}(y_i - ax_i - b)^2$. In order to minimize the equation we will take the derivative and then set it equal to zero for each of the variables, a and b. We will begin by taking the partial derivative with respect to b:

$$\frac{\partial E}{\partial b} = \sum_{i=1}^{n}\left[2\left(y_i - ax_i - b\right)(-1)\right] = -2\sum_{i=1}^{n}\left(y_i - ax_i - b\right) = -2\left(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} ax_i - \sum_{i=1}^{n} b\right)$$

$$= -2\left(n\bar{y} - an\bar{x} - b\right) = -2n\left(\bar{y} - a\bar{x} - b\right)$$

Setting $\frac{\partial E}{\partial b} = 0$ to minimize the error E yields:

$$0 = \frac{\partial E}{\partial b} = -2n\left(\bar{y} - \hat{a}\bar{x} - \hat{b}\right) \Rightarrow 0 = \bar{y} - \hat{a}\bar{x} - \hat{b} \Rightarrow \hat{b} = \bar{y} - \hat{a}\bar{x}$$

.
We have now shown that $\hat{b} = \bar{y} - \hat{a}\bar{x}$ when the total squared error E is minimized. Now we will take the partial derivative with respect to a:

$$\frac{\partial E}{\partial a} = \sum_{i=1}^{n}\left[2\left(y_i - ax_i - b\right)(-x_i)\right] = -2\left(\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} ax_i^2 - \sum_{i=1}^{n} bx_i\right)$$

$$= -2\left(\sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - bn\bar{x}\right)$$

Setting $\frac{\partial E}{\partial a} = 0$ to minimize the error E yields:

$$0 = \frac{\partial E}{\partial a} = -2\left(\sum_{i=1}^{n} x_i y_i - \hat{a}\sum_{i=1}^{n} x_i^2 - \hat{b}n\bar{x}\right)$$

Since $\hat{b} = \bar{y} - \hat{a}\bar{x}$, substitution yields:

$$0 = -2\left(\sum_{i=1}^{n} x_i y_i - \hat{a}\sum_{i=1}^{n} x_i^2 - (\bar{y} - \hat{a}\bar{x})\, n\bar{x}\right)$$

$$\Rightarrow 0 = -2\left(\sum_{i=1}^{n} x_i y_i - \hat{a}\sum_{i=1}^{n} x_i^2 - n\bar{x}\bar{y} + \hat{a}n\bar{x}^2\right)$$

$$\Rightarrow 0 = \sum_{i=1}^{n} x_i y_i - \hat{a}\sum_{i=1}^{n} x_i^2 - n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) + \hat{a}n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2$$

$$\Rightarrow 0 = n\sum_{i=1}^{n} x_i y_i - \hat{a}n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) + \hat{a}\left(\sum_{i=1}^{n} x_i\right)^2$$

$$\Rightarrow \hat{a}n\sum_{i=1}^{n} x_i^2 - \hat{a}\left(\sum_{i=1}^{n} x_i\right)^2 = n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$$

$$\Rightarrow \hat{a}\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right] = n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$$

$$\Rightarrow \hat{a} = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Thus we have shown that $\hat{a} = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$ when the total squared error E is minimized. Therefore, the line that minimizes the total squared error E is given by $y = \hat{a}x + \hat{b}$ where $\hat{a} = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$ and $\hat{b} = \bar{y} - \hat{a}\bar{x}$.