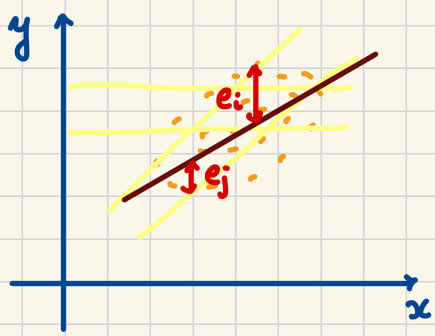
Simple Linear Regression.

The Model.

$Y = \beta_0 + \beta_1 X + \varepsilon$ w/ $\varepsilon \perp\!\!\!\perp X$
and $\varepsilon \sim N(0, \sigma^2)$

data set: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$



Every line of fit would be of the form

$$\boxed{y = b_0 + b_1 \cdot x}$$ w/ b_0 and b_1 "candidate" coeffs

e_k, $k = 1..n$ ... $\boxed{\text{RESIDUALS}}$

$$SSE = RSS = \sum_{k=1}^{n} e_k^2 \xrightarrow[b_0, b_1]{} \min$$

$$\sum_k (\underbrace{\boxed{y_k}}_{\text{observed}} - \underbrace{\boxed{\hat{y}_k}}_{\text{predicted}})^2 \longrightarrow \min$$

$$\sum_k (y_k - \boxed{b_0} - \boxed{b_1} x_k)^2 \xrightarrow[b_0, b_1]{} \min$$

For unbiasedness: $\sum_k e_k = 0$

$$\frac{\partial}{\partial b_0} RSS = -2 \sum_k (y_k - b_0 - b_1 x_k) = 0$$

$$\sum_k y_k = \sum_k (b_0 + b_1 x_k) = b_0 \cdot n + b_1 \sum_k x_k \quad / :n$$

$$\underbrace{\frac{1}{n} \sum_k y_k} = b_0 + b_1 \cdot \frac{1}{n} \sum_k x_k$$

$$\boxed{\bar{y} = \hat{b}_0 + \hat{b}_1 \cdot \bar{x}}$$ normal equation

$(\bar{x}, \bar{y})$ is on the least squares line

$$\frac{\partial}{\partial b_1} RSS = -2 \sum_{k=1}^{n} ((y_k - b_0 - b_1 x_k) \cdot x_k) = 0$$

$$\sum_k x_k y_k - b_0 \sum_k x_k - b_1 \sum_k x_k^2 = 0$$

$$\vdots$$

$$\boxed{\hat{b}_1 = \frac{\cancel{\frac{1}{n}} \sum (x_k - \bar{x})(y_k - \bar{y})}{\cancel{\frac{1}{n}} \sum (x_k - \bar{x})^2} = \frac{n \cdot \sum x_k \cdot y_k - \sum x_k \sum y_k}{n \cdot \sum x_k^2 - (\sum x_k)^2} = \frac{\sum x_k y_k - \frac{1}{n} \sum x_k \sum y_k}{\sum x_k^2 - \frac{1}{n} (\sum x_k)^2}}$$

Intuition: $$\hat{\beta}_1 = \frac{\text{"}Cov[X,Y]\text{"}}{\text{"}Var[X]\text{"}} = \frac{r_{X,Y} \cdot \cancel{S_X} \cdot S_Y}{S_X^{\cancel{2}}}$$

$$= r_{X,Y} \cdot \frac{S_Y}{S_X}$$

Search:
CAPM

$$\hat{\beta}_1 = \frac{S_{XY}}{S_X^2}$$

<u>Our estimators</u>

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{X}$$

$$\hat{\sigma}^2 = \frac{RSS}{n-2}$$

From $\varepsilon \perp\!\!\!\perp X$ requirement, we get the other normal equation

$$n \cdot \sum_k e_k x_k = \underbrace{\sum e_k}_{} \cdot \sum x_k = 0$$

$$\boxed{\sum_k e_k x_k = 0}$$