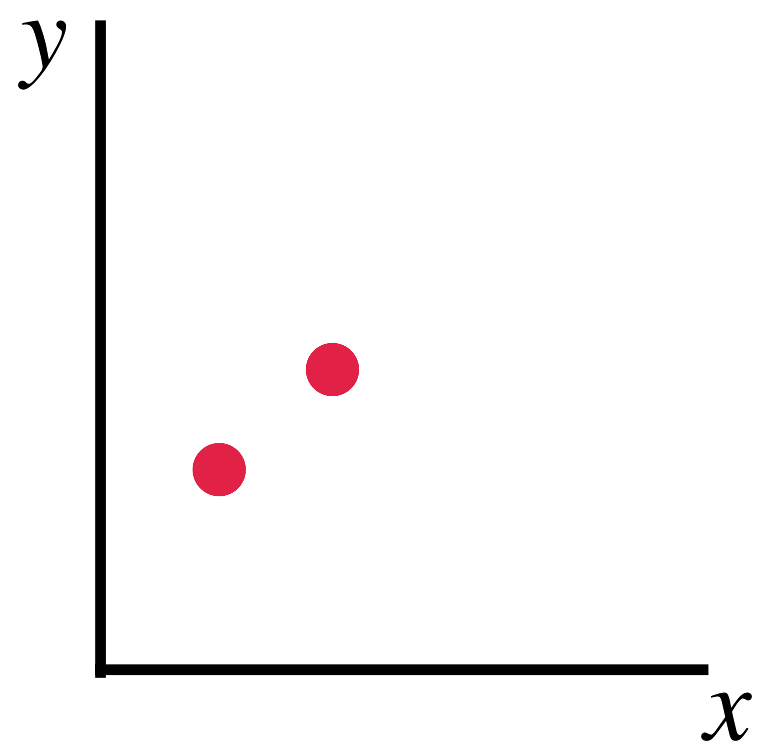
Example: Very Simple Linear Regression

$\hat{y} = b_0 + b_1 x$



Loss function:

$$RSS = \sum_{i}^{N} (\text{actual - predicted})^2 = \sum_{i}^{N} (y_i - \hat{y}_i)^2$$

$$= \sum_{i}^{N} (y_i - (b_0 + b_1 x))^2$$

$$= \sum_{i}^{N} (y_i - b_0 - b_1 x)^2$$

Assume only 2 data points: $(x_1, y_1) = (1,2), (x_2, y_2) = (2,3)$

Example: Very Simple Linear Regression

Gradient:

$$\begin{bmatrix} \frac{\partial RSS}{\partial b_0} \\ \frac{\partial RSS}{\partial b_1} \end{bmatrix} = \begin{bmatrix} -2\sum_i^N (y_i - (b_0 + b_1 x_i)) \\ -2\sum_i^N x_i(y_i - (b_0 + b_1 x_i)) \end{bmatrix}$$

Initialize the gradient algorithm at **(0,0)**

$$\implies \begin{bmatrix} -2\sum_i^N (y_i - (0 + 0x_i)) \\ -2\sum_i^N x_i(y_i - (0 + 0x_i)) \end{bmatrix} = \begin{bmatrix} -2\sum_i^N (y_i) \\ -2\sum_i^N x_i(y_i) \end{bmatrix}$$

plug in our actual values