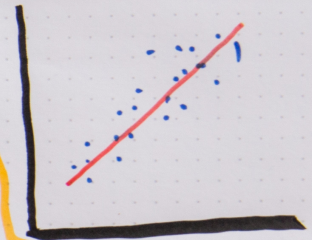
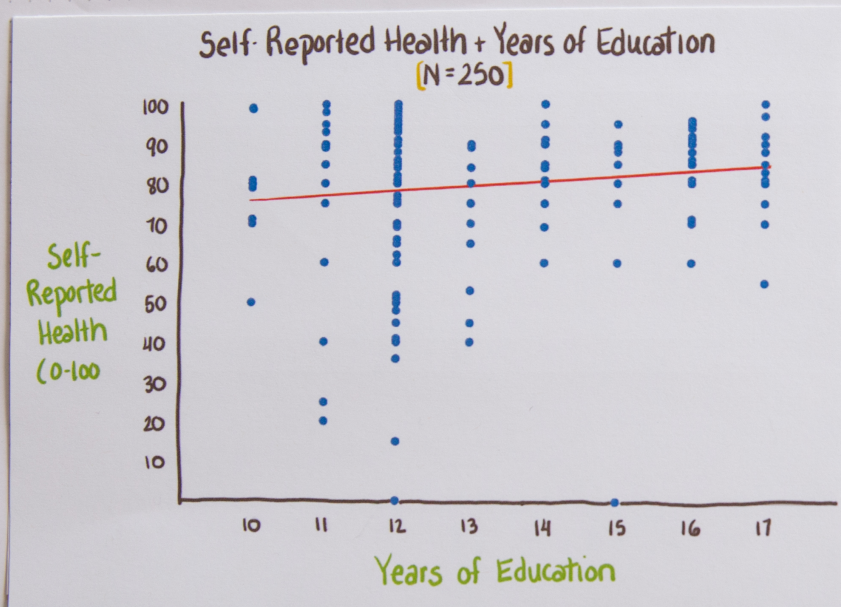


Relationship Between 2 Variables



What is the 'best' line

1. Choose a candidate line.
2. Find the distance between each point and the line (e_i)
3. Square the e_i and add: $\sum_i e_i^2$
4. Choose a, b to minimize the sum of squared e_i 's.

\hat{a}, \hat{b} - estimator

Regression Line

$$\text{Health} = 64 + 1.2 * \text{Years Ed.}$$

$$\text{Health} = a + b \cdot \text{Years Ed}$$

$$Y_i = a + b \cdot X_i + e_i$$