Linear Regression Formulas

$\bar{x}$ is the mean of x values

$\bar{y}$ is the mean of y values

s_x is the sample standard deviation for x values

s_y is the sample standard deviation for y values

r is the regression coefficient

The line of regression is:

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

where $b_1 = (r \cdot s_y)/s_x$

and $b_0 = \bar{y} - b_1 \bar{x}$

Things to remember

- Standardizing variables has **no effect** on the regression coefficient “r”, neither does changing units of the variables.
- R^2 tells you how much of the variability of the y variable is predicted by the x variable
- r ranges from -1 to 1
- Correlation **does not** imply causation!