Lesson Learning Objectives

LO 1. Define the explanatory variable as the independent variable (predictor), and the response variable as the dependent variable (predicted).

LO 2. Plot the explanatory variable (x) on the x-axis and the response variable (y) on the y-axis, and fit a linear regression model

$$y = \beta_0 + \beta_1 x,$$

where β_0 is the intercept, and β_1 is the slope.

- Note that the point estimates (estimated from observed data) for β_0 and β_1 are b_0 and b_1, respectively.

LO 3. When describing the association between two numerical variables, evaluate

- direction: positive ($x \uparrow, y \uparrow$), negative ($x \downarrow, y \uparrow$)
- form: linear or not
- strength: determined by the scatter around the underlying relationship

LO 4. Define correlation as the *linear* association between two numerical variables.

- Note that a relationship that is nonlinear is simply called an association.

LO 5. Note that correlation coefficient (R, also called Pearson's R) has the following properties:

- the magnitude (absolute value) of the correlation coefficient measures the strength of the linear association between two numerical variables
- the sign of the correlation coefficient indicates the direction of association
- the correlation coefficient is always between -1 and 1, -1 indicating perfect negative linear association, +1 indicating perfect positive linear association, and 0 indicating no linear relationship
- the correlation coefficient is unitless
- since the correlation coefficient is unitless, it is not affected by changes in the center or scale of either variable (such as unit conversions)
- the correlation of X with Y is the same as of Y with X