

## Vocabulary: Correlation



### Vocabulary

- **Correlation** – a measure of the relationship between two variables.
  - If the variables vary together (as one goes up, the other tends to go up), they are *positively correlated*.
  - If the variables vary in opposite directions (as one goes up, the other tends to go down), they are *negatively correlated*.
  - If the variables are unrelated, they have *no correlation*.
- **Correlation coefficient** – a number between  $-1$  and  $1$ , represented by the letter  $r$ , that indicates how strongly two variables are correlated and the type of correlation.
  - A positive value of  $r$  indicates a positive correlation, while a negative value of  $r$  indicates a negative correlation.
  - The closer  $r$  is to  $1$ , the stronger the positive correlation between the variables.
  - The closer  $r$  is to  $-1$ , the stronger the negative correlation.
  - A value of  $r$  near  $0$  indicates little to no correlation between the variables.
- **Least-squares fit line** – the line that best represents the linear relationship between two variables.
  - The least-squares fit line is also called the *least-squares regression line*.
  - The least-squares fit line is the line that minimizes the total sum of the squares of the *residuals*, where the residuals are the vertical distances from the data points to the least-squares fit line, as shown to the right.
- **Scatter plot** – a graph of  $(x, y)$  points that shows the general relationship between two variables,  $x$  and  $y$ .
  - Generally, the variable on the horizontal ( $x$ ) axis is the *independent variable*, and the variable on the vertical ( $y$ ) axis is the *dependent variable*.

