Chapter 6

SCATTERPLOTS, ASSOCIATION, AND CORRELATION

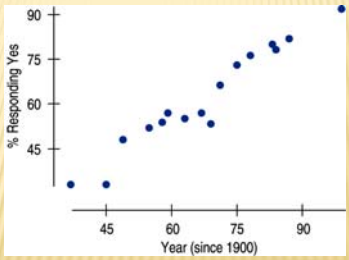
1

SCATTERPLOT

- A scatterplot is a graph in which the paired (*x*, *y*) sample data are plotted with a horizontal *x*-axis and a vertical *y*-axis.
- Each individual (*x*, *y*) pair is plotted as a single point.

2

SCATTERPLOT



3

RELATIONSHIPS

- We plot the explanatory variable (or outcome variable) on the *x*-axis and the response variable (or predictor variable) on the *y*-axis

4

RELATIONSHIPS

- The value of the explanatory variable is thought to partially explain the value of the response variable for that individual
- Example: People with higher education levels generally have higher incomes
  - Explanatory variable: education level
  - Response variable: income
- This relationship is not causal! The explanatory variable does not cause or determine the response variable

5

CORRELATION

- A correlation exists between two variables when one of them is related to the other in some way.
- The linear correlation coefficient, *r*, measures the strength of the linear relationship between the paired *x*- and *y*-values in a sample.
- The value of r^2 (often capitalized) is the proportion of the variation in *y* that is explained by the linear relationship between *x* and *y*.

6

CORRELATION COEFFICIENT

- If the value of *r* is close to 1, there is a strong positive correlation between the variables
- If the value of *r* is close to 0, there is no correlation between the variables
- If the value of *r* is close to -1, there is a strong negative correlation between the variables

7

CONDITIONS FOR CORRELATION

- Quantitative Variables Condition
- Straight Enough Condition
- Outlier Condition

8