Chapter 3: Describing Relationships

3.1 Scatterplots and Correlation (pp.142-156)

1. Why do we study the relationship between two quantitative variables?

1. What is the difference between a *response variable* and the *explanatory variable*?
2. How are response and explanatory variables related to *dependent* and *independent* variables?
3. When is it appropriate to use a *scatterplot* to display data?
4. A *scatterplot* shows the relationship between…
5. Which variable always appears on the horizontal axis of a scatterplot?
6. When examining a scatterplot, you can describe the overall pattern by its:

**D\_\_\_\_\_ O\_\_\_\_\_ F\_\_\_\_\_ S\_\_\_\_\_**

1. Explain the difference between a *positive association* and a *negative association*.
2. What is *correlation r*?
3. Answer the five questions for the *Check Your Understanding* on page 149.
4. What does correlatio*n* *measure*?
5. Explain why two variables must both be *quantitative* in order to find the *correlation* between them.
6. What is true about the relationship between two variables if the *r-value* is:
7. Near 0?
8. Near 1?
9. Near -1?
10. Exactly 1?
11. Exactly -1?
12. Is *correlation* resistant to extreme observations? Explain.
13. What do you need to know in order to *interpret* correlation?

3.2 Least-Squares Regression (pp.164-188)

1. What is a *regression line*?
2. In what way is a *regression line* a *mathematical model*?
3. What is the general form of a *regression equation*? Define each variable in the equation.
4. What is the difference between *y* and **?
5. What is *extrapolation* and why is this dangerous?
6. Answer the four questions for the *Check Your Understanding* on page 167.
7. What is a *residual*? How do you interpret a residual?
8. What is a *least-squares regression line*?
9. What is the formula for the equation of the *least-squares regression line*?
10. The *least-squares* *regression line* always passes through the point ...
11. What is a *residual plot*? Sketch a graph of a residual plot.
12. If a *least-squares regression line* fits the data well, what two characteristics should the *residual plot* exhibit?
13. What is the standard deviation of the residuals? How is it interpreted?
14. How is the *coefficient of determination* defined?
15. What is the formula for calculating the *coefficient of determination*?
16. If *r2* = 0.95, what can be concluded about the relationship between *x* and *y*?

\_\_\_\_\_\_% of the variation in (response variable) is accounted for by the regression line.

1. When reporting a regression, should r or *r2* be used describe the success of the regression? Explain.
2. Identify the *slope*, the *y intercept*, *s* and *r2* on the computer output.

 

1. What are three limitations of *correlation* and *regression*?
2. What is an *outlier*?
3. What is an *influential point*?
4. Under what conditions does an outlier become an *influential observation*?
5. What is a *lurking variable*?
6. Why does *association* not imply *causation*?

