Chapter 3: Describing Relationships

Key Vocabulary:

- response variable
- explanatory variable
- independent variable
- dependent variable
- scatterplot
- positive association
- negative association
- linear
- DOFS
- correlation
- r-value
- regression line
- mathematical model
- least-squares regression line
- $\hat{y}$ "y-hat"
- SSM
- SSE
- extrapolation
- r^2
- coefficient of determination
- residuals
- residual plot
- influential observation
- outliers
- lurking variable

3.1 Scatterplots and Correlation (pp.142-156)

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

2. What is the difference between a *response variable* and the *explanatory variable*?

3. How are response and explanatory variables related to *dependent* and *independent* variables?

4. When is it appropriate to use a *scatterplot* to display data?

5. A *scatterplot* shows the relationship between…

6. Which variable always appears on the horizontal axis of a scatterplot?

7. When examining a scatterplot, you can describe the overall pattern by its:

D_____ O_____ F_____ S_____

8. Explain the difference between a *positive association* and a *negative association*.

9. What is *correlation r*?

10. Answer the five questions for the *Check Your Understanding* on page 149.

11. What does correlation *measure*?

12. Explain why two variables must both be *quantitative* in order to find the *correlation* between them.

13. What is true about the relationship between two variables if the *r-value* is:
    a. Near 0?

    b. Near 1?

    c. Near -1?

    d. Exactly 1?

    e. Exactly -1?

14. Is *correlation* resistant to extreme observations? Explain.

15. 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 $\hat{y}$?

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?

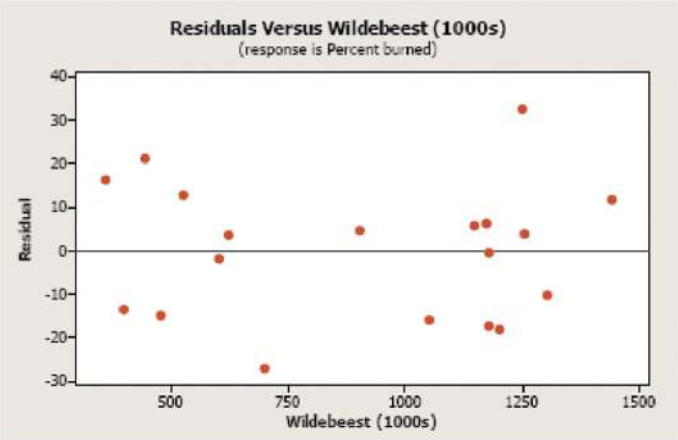
15. What is the formula for calculating the *coefficient of determination*?

16. If $r^2 = 0.95$, what can be concluded about the relationship between x and y?

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

17. When reporting a regression, should r or r^2 be used describe the success of the regression? Explain.

18. Identify the *slope*, the *y intercept*, *s* and r^2 on the computer output.



Predictor	Coef	SE Coef	T	P
Constant	92.29	10.06	9.17	0.000
Wildebeest (1000s)	-0.05762	0.01035	-5.56	0.000

S = 15.9880 R-Sq = 64.6% R-Sq(adj) = 62.5%

19. What are three limitations of *correlation* and *regression*?

20. What is an *outlier*?

21. What is an *influential point*?

22. Under what conditions does an outlier become an *influential observation*?

23. Why does *association* not imply *causation*?