4.7 Exercises

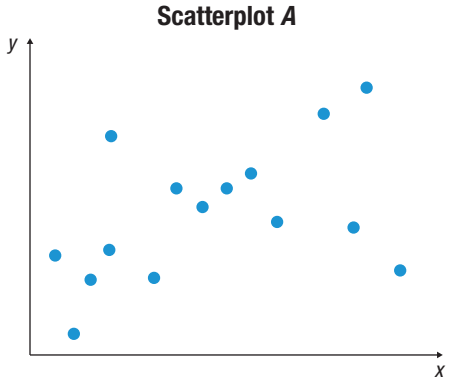
Basic Concepts

1. Give an example of a business situation in which knowledge of a relationship between two variables is desired.
2. If a relationship can be uncovered, what are the potential benefits?
3. What are bivariate data? How is bivariate data different from univariate data?
4. What graphical tool is often used in the discovery of relationships?
5. What are four common questions you should ask when studying a graphical representation of bivariate data?
6. If bivariate data exhibit an inverse relationship, what does that mean?
7. How do you construct exact relationships between two variables?
8. In what range is the value of r when bivariate data exhibit a positive relationship? A negative relationship?
9. If the value of r is small, does this always mean that no relationship exists? Explain.
10. What is confounding? Why is confounding a problem?

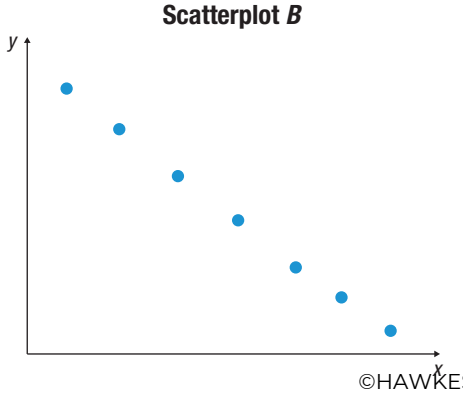
Exercises

11. Consider the following scatterplots and answer the following questions regarding the overall pattern of the data for each of the graphs.
    - Does the pattern roughly follow a straight line?
    - Is the pattern upward sloping or downward sloping?
    - Are the data values tightly clustered in the pattern or widely dispersed?
    - Are there significant deviations from the pattern?

a.



b.

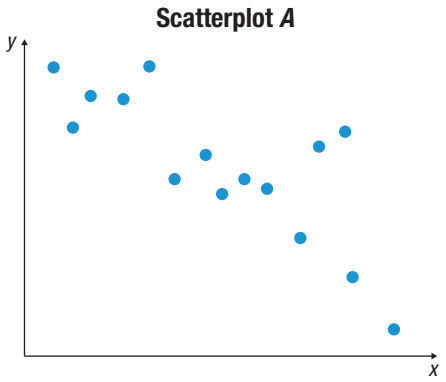


©HAWKES LEARNING

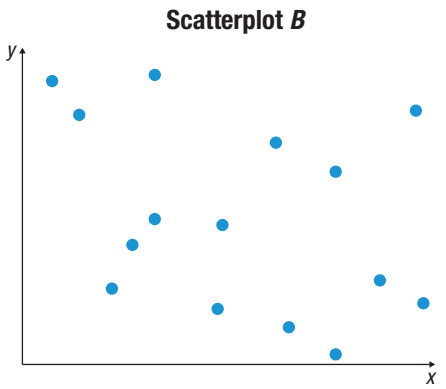
12. Consider the following scatterplots and answer the following questions regarding the overall pattern of the data for each of the graphs.

- Does the pattern roughly follow a straight line?
- Is the pattern upward sloping or downward sloping?
- Are the data values tightly clustered in the pattern or widely dispersed?
- Are there significant deviations from the pattern?

a.



b.



13. A manufacturing company which produces laminate for countertops is interested in studying the relationship between the number of hours of training an employee receives and the number of defects per countertop produced. Ten employees are randomly selected. The number of hours of training which each employee has received is recorded and the number of defects on the most recent countertop produced is determined. The results are as follows.

Employee Training	
Hours of Training	Defects per Countertop
1	1
4	4
7	0
3	3
2	5
2	4
5	3
5	2
1	5
6	1

©HAWKES LEARNING

a. Analyze the data collected for the study by answering the following questions.

i. Do the variables selected for measurement seem appropriate for answering the question that the manufacturing company is interested in?

ii. What biases or errors might be present in the data?

iii. How are the data collected – through observation or controlled experiment?

b. Plot the data points on a scatterplot.

c. Based on the scatterplot in part **b.**, answer the following questions regarding the overall pattern of the data.

i. Does the pattern roughly follow a straight line?

ii. Is the pattern upward sloping or downward sloping? Are the data values tightly clustered in the pattern or widely dispersed?

iii. Are there significant deviations from the pattern?

14. Illustrate, using a scatterplot, a data set that would have a correlation coefficient of 1.

15. Illustrate, using a scatterplot, a data set that would have a correlation coefficient of –1.

16. Describe the relationships indicated by the correlation coefficients as tightly clustered in a positive linear fashion, tightly clustered in a negative linear fashion, loosely clustered in a positive linear fashion, loosely clustered in a negative linear fashion, or no linear relationship.

a. $r = 0.9$ **c.** $r = -0.9$ **e.** $r = 0$

b. $r = 0.5$ **d.** $r = -0.5$

17. Describe the relationships indicated by the correlation coefficients as tightly clustered in a positive linear fashion, tightly clustered in a negative linear fashion, loosely clustered in a positive linear fashion, loosely clustered in a negative linear fashion, or no linear relationship.

a. $r = 0.8$ **c.** $r = -0.8$ **e.** $r = 0.1$

b. $r = 0.4$ **d.** $r = -0.4$

18. A sample of 10 female swimmers, all 17 years old, is selected from a local swim league. Each swimmer's best time (in seconds) in the 50-yard freestyle and in the 100-yard individual medley are obtained. The 100-yard individual medley consists of swimming 25 yards with each of the four major strokes. The data are given in the following table.

Best Times

Freestyle	27.4	27.0	26.8	30.7	28.5	28.6	29.6	30.8	31.5	29.8
Medley	66.3	66.4	66.7	78.7	69.4	72.0	73.5	81.1	78.6	73.5

a. Construct a scatterplot of the data.

b. Does there appear to be a negative or positive relationship between the variables?

c. Compute the correlation coefficient.

©HAWKES LEARNING

19. A personnel director is interested in studying the relationship (if any) between age and salary. Sixteen employees are randomly selected and their ages and salaries are recorded.

Ages and Salaries			
Age	**Salary ($)**	**Age**	**Salary ($)**
25	22,000	49	39,000
55	45,000	37	45,000
27	43,000	62	60,000
30	30,000	40	35,000
22	24,000	35	34,000
33	53,000	29	30,000
19	18,000	58	73,000
45	38,000	52	42,000

a. Plot the data points on a scatterplot.

b. Determine the correlation coefficient.

c. Describe the relationship indicated by the correlation coefficient and the scatterplot.

20. The following variables have high positive linear correlations. Is it reasonable to conclude that an increase in one variable causes an increase in the other variable? Explain what could be causing this apparent relationship.

a. Height and vocabulary

b. Absenteeism from school and sale of cough syrup

c. Sale of turkey and sale of toys

21. The following variables have high positive linear correlations. Is it reasonable to conclude that an increase in one variable causes an increase in the other variable? Explain what could be causing this apparent relationship.

a. Sale of air conditioners and sale of tomatoes

b. Sale of greeting cards and sale of chocolates

c. The number of wrecks on a local highway and absenteeism from work

©HAWKES LEARNING