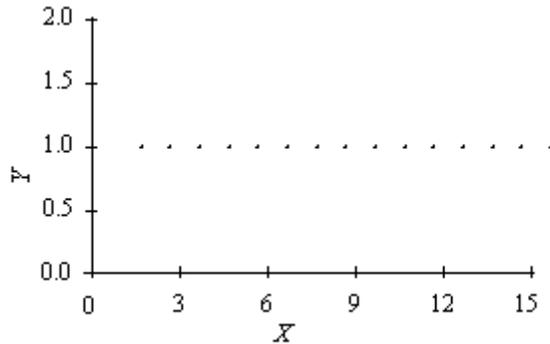


### CHAPTER 3: MULTIPLE CHOICE

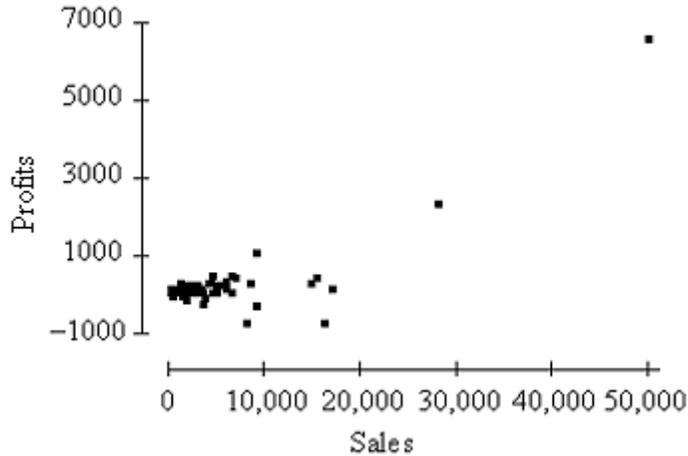
1. A scatterplot of a variable  $Y$  versus a variable  $X$  produced the results below. The value of  $Y$  for all values of  $X$  is exactly 1.0.



The correlation between  $Y$  and  $X$  is

- A) 1, because the points lie perfectly on a line.
  - B) either 1 or  $-1$ , because the points lie perfectly on a line.
  - C) 0, because  $Y$  does not change as  $X$  increases.
  - D) impossible to determine, since there is no slope to the data.
  - E) none of the above.
2. The fraction of the variation in the values of  $y$  that is explained by the least-squares regression of  $y$  on  $x$  is
- A) the correlation coefficient.
  - B) the slope of the least-squares regression line.
  - C) the square of the correlation coefficient.
  - D) the intercept of the least-squares regression line.
  - E) the residual.

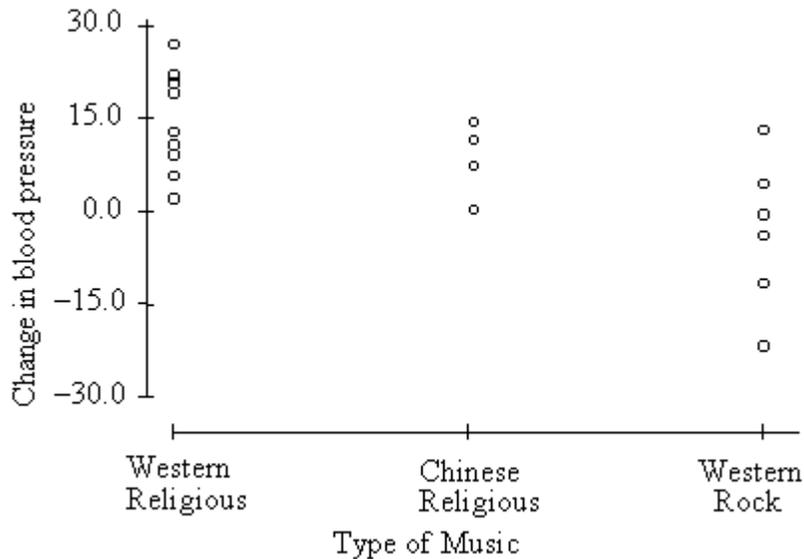
3. A sample of 79 companies was taken, and the annual profits ( $y$ ) were plotted against annual sales ( $x$ ). The plot is given below. All values in the plots are in units of \$100,000.



The correlation between sales and profits is found to be 0.814. Based on this information, we may conclude which of the following?

- A) If the sales were less than \$20,000, the equation of the least-squares regression line would predict the profits quite accurately.
- B) There are clearly influential observations present.
- C) If we group the companies in the plot into those that are small in size, those that are medium in size, and those that are large in size and compute the correlation between sales and profits for each group of companies separately, the correlation in each group will be about 0.8.
- D) Not surprisingly, increasing sales causes an increase in profits. This is confirmed by the large positive correlation.
- E) All of the above.

4. Volunteers for a research study were divided into three groups. Group 1 listened to Western religious music, group 2 listened to Western rock music, and group 3 listened to Chinese religious music. The blood pressure of each volunteer was measured before and after listening to the music, and the change in blood pressure (blood pressure before listening minus blood pressure after listening) was recorded. A scatterplot of change in blood pressure versus type of music listened to is given below.



The correlation between change in blood pressure and type of music listened to is

- A) negative.
  - B) positive.
  - C) first negative, then positive.
  - D) nearly 0.
  - E) none of the above.
5. Which of the following is true of the least-squares regression line?
- A) The slope is the change in the response variable that would be predicted by a unit change in the explanatory variable.
  - B) It always passes through the point  $(\bar{X}, \bar{Y})$ , the means of the explanatory and response variables, respectively.
  - C) It will only pass through all the data points if  $r = \pm 1$ .
  - D) No more than 50% of the residual values will be positive.
  - E) All of the above.

6. A least-squares regression line is fitted to a set of data. If one of the data points has a positive residual, then
- the correlation between the values of the response and explanatory variables must be positive.
  - the point must lie above the least-squares regression line.
  - the slope of the least-squares regression line must be positive.
  - the point must lie near the right edge of the scatterplot.
  - all of the above.
7. A study found a correlation of  $r = -0.61$  between the gender of a worker and his or her income. We may correctly conclude that
- women earn more than men on the average.
  - women earn less than men on the average.
  - an arithmetic mistake was made, since correlation must always be positive.
  - this result is incorrect, because computing  $r$  makes no sense in this situation.
  - on average, women earn 61% less than men.
8. A researcher wishes to determine whether the rate of water flow (in liters per second) over an experimental soil bed can be used to predict the amount of soil washed away (in kilograms). The researcher measures the amount of soil washed away for various flow rates and from these data calculates the least-squares regression line to be
- $$\widehat{\text{amount of eroded soil}} = 0.4 + 1.3 \times (\text{flow rate})$$
- The correlation between amount of eroded soil and flow rate would be
- 1/1.3.
  - 0.4.
  - 1.3.
  - positive, but we cannot say what the exact value is using the information given.
  - either positive or negative. It is impossible to say anything about the correlation from the information given.
9. In a study of 1991 model cars, a researcher computed the least-squares regression line of price (in dollars) on horsepower. He obtained the following equation for this line.
- $$\widehat{\text{Price}} = -6677 + 175 \times \text{horsepower}$$
- Based on the least-squares regression line, we would predict that a 1991 model car with horsepower equal to 200 would cost
- \$41,677.
  - \$35,000.
  - \$34,175.
  - \$28,323.
  - \$13,354.

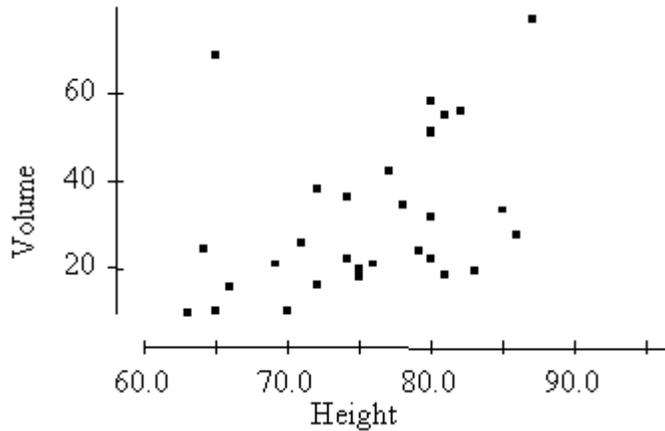
10. A study is conducted to determine if one can predict the yield of a crop based on the amount of yearly rainfall. The response variable in this study is
- A) the yield of the crop.
  - B) the amount of yearly rainfall.
  - C) the experimenter.
  - D) either bushels or inches of water.
  - E) the month the crop is harvested.
11. A researcher wishes to study how the average weight  $Y$  (in kilograms) of children changes during the first year of life. He plots these averages versus the children's age  $X$  (in months) and decides to fit a least-squares regression line to the data with  $X$  as the explanatory variable and  $Y$  as the response variable. He computes the following quantities.
- $r =$  correlation between  $X$  and  $Y = 0.9$
  - $\bar{X} =$  mean of the values of  $X = 6.5$
  - $\bar{Y} =$  mean of the values of  $Y = 6.6$
  - $s =$  standard deviation of the values of  $X = 3.6$
  - $s_m =$  standard deviation of the values of  $Y = 1.2$

The slope of the least-squares line is

- A) 0.30.
- B) 0.88.
- C) 1.01.
- D) 2.7.
- E) 3.0.

Use the following to answer question 12:

A researcher measures the height (in feet) and volume of usable lumber (in cubic feet) of 32 cherry trees. The goal is to determine if the volume of a tree's usable lumber can be estimated from the height of the tree. The results are plotted below.



12. The scatterplot above suggests that
- A) there is a positive association between height and volume.
  - B) there is an outlier in the plot.
  - C) both A and B.
  - D) neither A nor B.
  - E) the relationship between height and volume is nonlinear.

Use the following to answer questions 13-14:

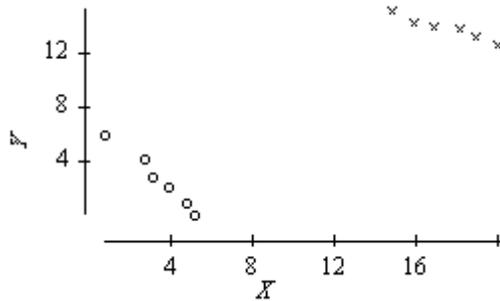
I wish to determine the correlation between the height (in inches) and weight (in pounds) of 21-year-old males. To do this, I measure the height and weight of two 21-year-old men. The measured values are

	<u>Male #1</u>	<u>Male #2</u>
Height	70	75
Weight	160	200

13. Referring to the information above, which of the following units would the correlation coefficient  $r$  have?
- A) Inches.
  - B) Pounds.
  - C) Pounds per inch.
  - D) None, because  $r$  has no units.
  - E) Inches-pounds.
14. Referring to the information above, the correlation  $r$  computed from the measurements on these males is
- A) equal to 1.
  - B) positive and between 0.25 and 0.75.
  - C) near 0, but could be either positive or negative.
  - D) exactly 0.
  - E) Meaningless, since the slope is greater than 1.
15. At a large university, the office responsible for scheduling classes notices that demand is low for classes that meet before 10:00 AM or after 3:00 PM and is high for classes that meet between 10:00 AM and 3:00 PM. Which of the following may we conclude?
- A) There is an association between demand for classes and the time the classes meet.
  - B) The association between demand for classes and time for classes is linear.
  - C) There is a *negative* association between demand for classes and the time the classes meet.
  - D) There is no association between demand for classes and the time the classes meet.
  - E) There is a *positive* association between demand for classes and the time the classes meet.
16. Volunteers for a research study were divided into three groups. Group 1 listened to Western religious music, group 2 listened to Western rock music, and group 3 listened to Chinese religious music. The blood pressure of each volunteer was measured before and after listening to the music, and the change in blood pressure (blood pressure before listening minus blood pressure after listening) was recorded. To explore the relationship between type of music listened to and change in blood pressure, we could
- A) see if blood pressure decreases as type of music increases by examining a scatterplot.
  - B) make a histogram of the change in blood pressure for all of the volunteers.
  - C) make side-by-side boxplots of the change in blood pressure, with a separate boxplot for each group.
  - D) make a pie chart displaying the distribution of type of music listened to for all of the volunteers.
  - E) do all of the above.

17. The least-squares regression line is
- A) the line that makes the square of the correlation in the data as large as possible.
  - B) the line that makes the sum of the squares of the vertical distances of the data points from the line as small as possible.
  - C) the line that passes through the greatest number of data points.
  - D) the line that best splits the data in half, with half of the points above the line and half below the line.
  - E) all of the above.

18. The scatterplot below is from a small data set.



The data were classified as either type 1 or type 2. Those of type 1 are indicated by o's, those of type 2 by x's. The overall correlation of the data in this scatterplot is

- A) positive.
  - B) near 0, since the overall data do not show a distinct pattern.
  - C) near 0, because the o's display a negative trend and the x's display a negative trend, but the trend from the o's to the x's is positive. The different trends cancel.
  - D) impossible to compute for such a data set.
  - E) negative, since the o's display a negative trend and the x's display a negative trend.
19. A researcher wishes to determine whether the rate of water flow (in liters per second) over an experimental soil bed can be used to predict the amount of soil washed away (in kilograms). The researcher measures the amount of soil washed away for various flow rates and from these data calculates the least-squares regression line to be

$$\text{amount of eroded soil} = 0.4 + 1.3 \times (\text{flow rate})$$

One of the flow rates used by the researcher was 0.3 liters per second; for this flow rate, the amount of eroded soil was 0.8 kilograms. These values were used in the calculation of the least-squares regression line. The residual corresponding to these values is

- A) 0.01.
- B) -0.01.
- C) 0.5.
- D) -0.5.
- E) -3.5.

20. In a study of 1991 model cars, a researcher found that the fraction of the variation in the price of cars that was explained by the least-squares regression on horsepower was about 0.64. For the cars in this study, the correlation between the price of the car and its horsepower was found to be positive. The actual value of the correlation
- A) is 0.80.
  - B) is 0.64.
  - C) is 0.41.
  - D) is  $-0.80$ .
  - E) cannot be determined from the information given.

## Answer Key

1. C
2. C
3. B
4. E
5. E
6. B
7. D
8. D
9. D
10. A
11. A
12. C
13. D
14. A
15. A
16. C
17. B
18. A
19. A
20. A