92 **A study is conducted to determine if one can predict the price of a stock based on the price-to-earnings ratio. The response variable in this study is:**

- **price of the stock.** *(True Answer)Correct*
- **the price-to-earnings ratio.** *Incorrect*
- **the researcher.** *Incorrect*
- **either the NASDAQ or the Dow Jones Industrial Average.** *Incorrect*

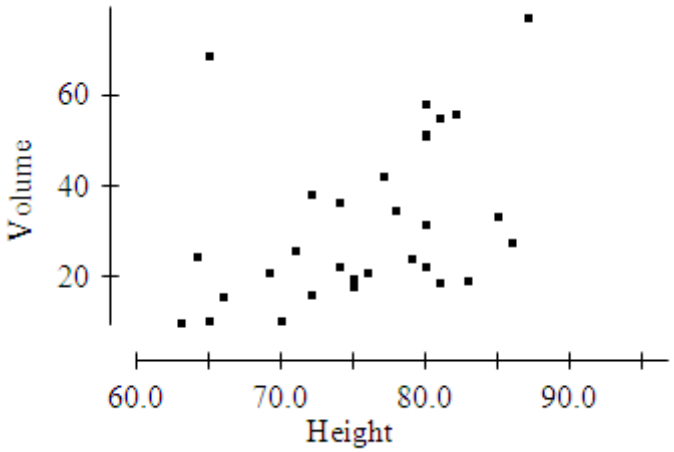
93 **A researcher is interested in determining if one could predict the score on a statistics exam from the amount of time spent studying for the exam. In this study, the explanatory variable is:**

- **the researcher.** *Incorrect*
- **the amount of time spent studying for the exam.** *(True Answer)Correct*
- **the score on the exam.** *Incorrect*
- **the fact that this is a statistics exam.** *Incorrect*

94 **When creating a scatterplot, one should:**

- **use the vertical axis for the response variable.** *(True Answer)Correct*
- **use the vertical axis for the explanatory variable.** *Incorrect*
- **use a different plotting symbol if the explanatory variable is categorical than if the response variable is categorical.** *Incorrect*
- **use a plotting scale that makes the overall trend roughly linear.** *Incorrect*

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



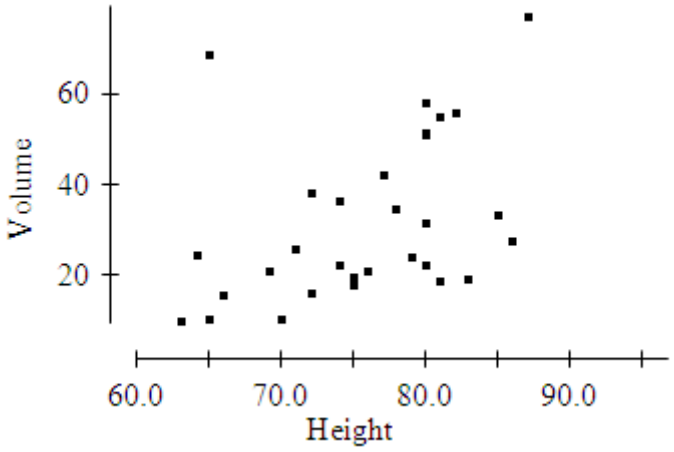
Reference: Ref 2-1

In this study, the response variable is:

- **height.** *Incorrect*
- **volume.** *(True Answer)Correct*
- **height or volume. It doesn't matter which is considered the response.** *Incorrect*
- **neither height nor volume. The measuring instrument used to measure height is the response variable.** *Incorrect*

---

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



Reference: Ref 2-1

The scatterplot suggests:

- **there is a positive association between height and volume.** *Incorrect*

- **there is an outlier in the plot.** *Incorrect*
- **both a) and b).** *(True Answer)Correct*
- **neither a) nor b).** *Incorrect*

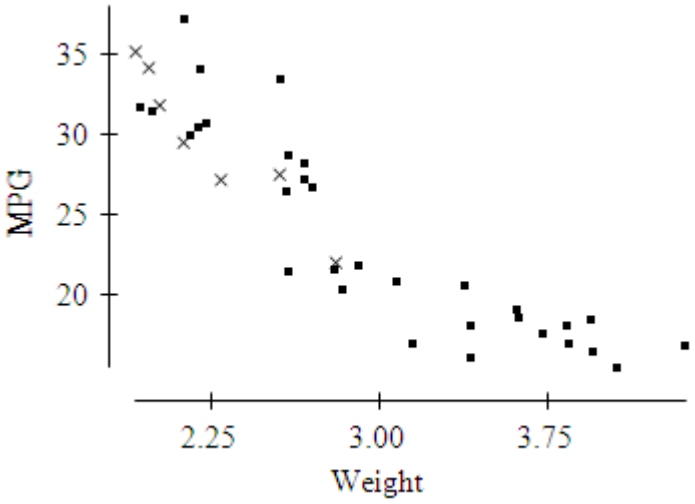
---

97 **At a large university, the office responsible for scheduling classes notices that demand is low for classes meeting before 10:00 AM or after 3:00 PM and is high for classes meeting between 10:00 AM and 3:00 PM. We may conclude which of the following?**

- **There is an association between demand for classes and the time the classes meet.** *(True Answer)Correct*
- **There is a *positive* association between demand for classes and the time the classes meet.** *Incorrect*
- **There is a *negative* association between demand for classes and the time the classes meet.** *Incorrect*
- **The time of the class meeting must cause changes in the demand for the classes.** *Incorrect*

---

98 **The graph below plots the gas mileage (miles per gallon, or MPG) of various 1978 model cars versus the weight of these cars in thousands of pounds.**



The points denoted by the plotting symbol × correspond to cars made in Japan. From this plot, we may conclude that:

- **there is little difference between Japanese cars and cars made in other countries.** *Incorrect*
- **Japanese cars tend to be lighter in weight than other cars.** *(True Answer)Correct*
- **Japanese cars tend to get poorer gas mileage than other cars.** *Incorrect*
- **the plot is invalid. A scatterplot is used to represent quantitative variables, and the country that makes a car is a qualitative variable.** *Incorrect*

99 **The stores of a large retail chain were divided into three groups. While customers were shopping, group 1 played light rock music, group 2 played classical music, and group 3 played show tunes. The daily sales for each day in a 30-day period were recorded. Suppose that, on average, sales were highest in those stores that played light rock music, second highest for those stores playing show tunes, and lowest for those stores playing classical music. We conclude:**

- **there is a positive association between sales and type of music played.** *Incorrect*
- **there is a negative association between sales and type of music played.** *Incorrect*
- **there is both positive and negative association present.** *Incorrect*
- **none of the above.** *(True Answer)Correct*

---

100 **A college newspaper interviews a psychologist about a proposed system for rating the teaching ability of faculty members. The psychologist says, "The evidence indicates that the correlation between a faculty member's research productivity and teaching rating is close to zero." A correct interpretation of this statement would be:**

- **good researchers tend to be poor teachers and vice versa.** *Incorrect*
- **good teachers tend to be poor researchers and vice versa.** *Incorrect*
- **good researchers are just as likely to be good teachers as they are bad teachers. Likewise for poor researchers.** *(True Answer)Correct*
- **good research and good teaching go hand in hand.** *Incorrect*

---

101 **A student wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches):**

Women	66	64	66	65	70	65
Men	72	68	70	68	74	69

Which of the following statements is true?

- **The variables measured are all categorical.** *Incorrect*
- **There is a strong negative association between the heights of men and women, because the women are always**

smaller than the men they date. *Incorrect*

- **There is a positive association between the heights of men and women.** *(True Answer)Correct*
- **Any height above 70 inches must be considered an outlier.** *Incorrect*

---

102 **Which of the following statements is true?**

- **The correlation coefficient equals the proportion of times two variables lie on a straight line.** *Incorrect*
- **The correlation coefficient will be +1.0 only if all the data lie on a perfectly horizontal straight line.** *Incorrect*
- **The correlation coefficient measures the fraction of outliers that appear in a scatterplot.** *Incorrect*
- **The correlation coefficient is a unitless number and must always lie between –1.0 and +1.0, inclusive.** *(True Answer)Correct*

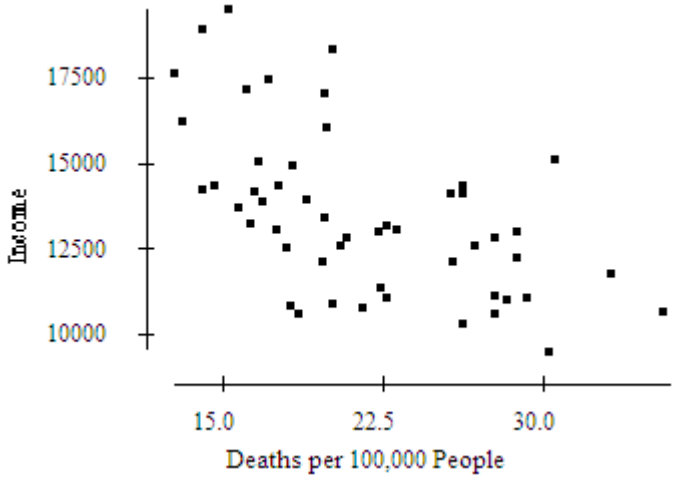
---

103 **A study found a correlation of $r = –0.61$ between the gender of a worker and his or her income. You may correctly conclude that:**

- **women earn more than men on the average.** *Incorrect*
- **women earn less than men on the average.** *Incorrect*
- **an arithmetic mistake was made. Correlation must be positive.** *Incorrect*
- **this is incorrect because *r* makes no sense here.** *(True Answer)Correct*

---

104 **The following scatterplot displays the 1990 per capita income versus number of deaths due to traffic accidents per 100,000 people for each of the 50 states plus the District of Columbia.**

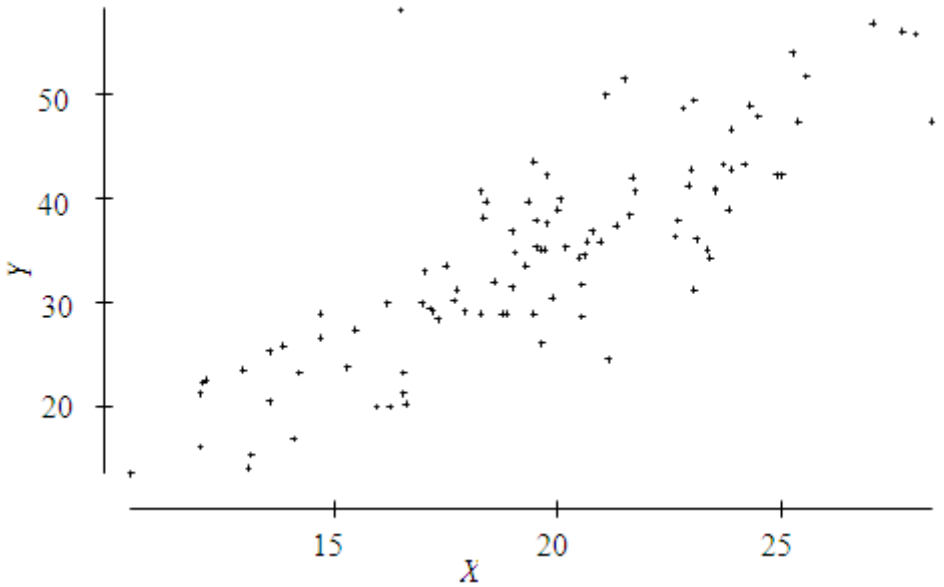


Which of the following is a plausible value for the correlation coefficient between weight and MPG?

- **+0.2** *Incorrect*
- **–0.5** *(True Answer)Correct*
- **+0.7** *Incorrect*
- **–1.0** *Incorrect*

10
5

Consider the following scatterplot.

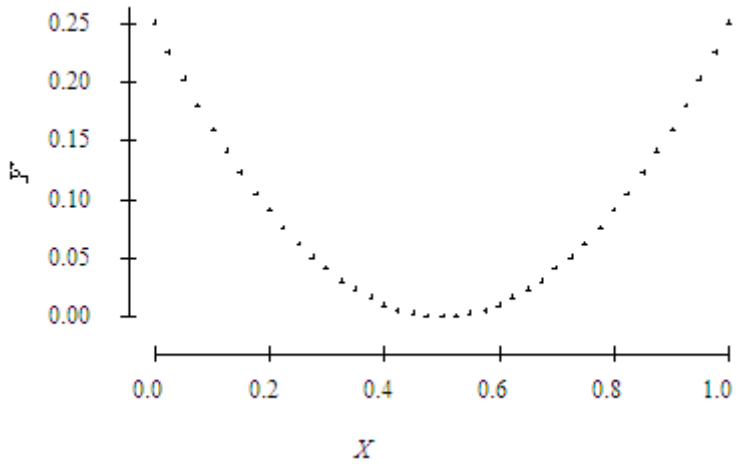


The correlation between *X* and *Y* is approximately:

- **0.999.** *Incorrect*
- **0.8.** *(True Answer)Correct*
- **0.0.** *Incorrect*
- **–0.7.** *Incorrect*

106

Consider the following scatterplot of two variables *X* and *Y*.



We may conclude that:

- **the correlation between *X* and *Y* must be close to 1**

because there is nearly a perfect relationship between them. *Incorrect*

- **the correlation between *X* and *Y* must be close to –1 because there is nearly a perfect relationship between them, but it is not a straight line relationship.** *Incorrect*
- **the correlation between *X* and *Y* is close to 0.** *(True Answer)Correct*
- **the correlation between *X* and *Y* could be any number between –1 and +1. Without knowing the actual values, we can say nothing more.** *Incorrect*

---

107 **A researcher wants to determine if a man's height and his foot length (measured from the heal to the big toe) are associated. Two men were measured and the results are below. The height and foot length are both measured in inches.**

	Male #1	Male #2
Height	70	75
Foot length	10	12

Reference: Ref 2-2

The correlation *r* computed from the measurements on these males is:

- **1.0.** *(True Answer)Correct*
- **positive, and between 0.25 and 0.75.** *Incorrect*
- **near 0, but could be either positive or negative.** *Incorrect*
- **exactly 0.** *Incorrect*

---

108 **A researcher wants to determine if a man's height and his foot length (measured from the heal to the big toe) are associated. Two men were measured and the results are below. The height and foot length are both measured in inches.**

	Male #1	Male #2
Height	70	75
Foot length	10	12

Reference: Ref 2-2

The correlation *r* would have units measured in:

- **inches.** *Incorrect*
- **square inches.** *Incorrect*
- **feet and inches.** *Incorrect*
- **no units. Correlation is a unitless quantity.** *(True Answer)Correct*

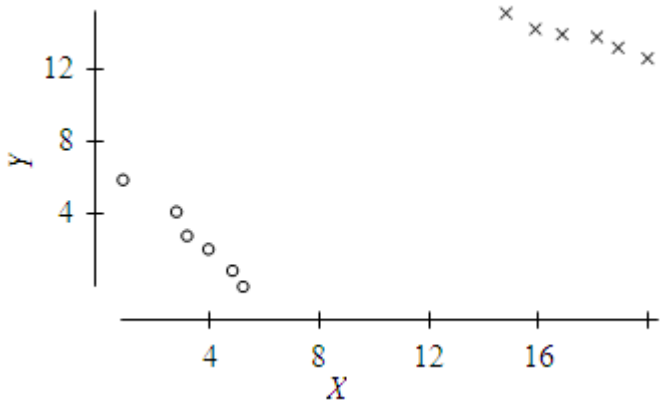
---

109 **Which of the following is true of the correlation coefficient *r*?**

- **It is a resistant measure of association.** *Incorrect*
- **–1 ? *r* ? 1.** *(True Answer)Correct*
- **If *r* is the correlation between *X* and *Y*, then –*r* is the correlation between *Y* and *X*.** *Incorrect*
- **The sign of the correlation tells us about the strength of the relationship between X and Y.** *Incorrect*

---

110 **The scatterplot below is from a small data set.**

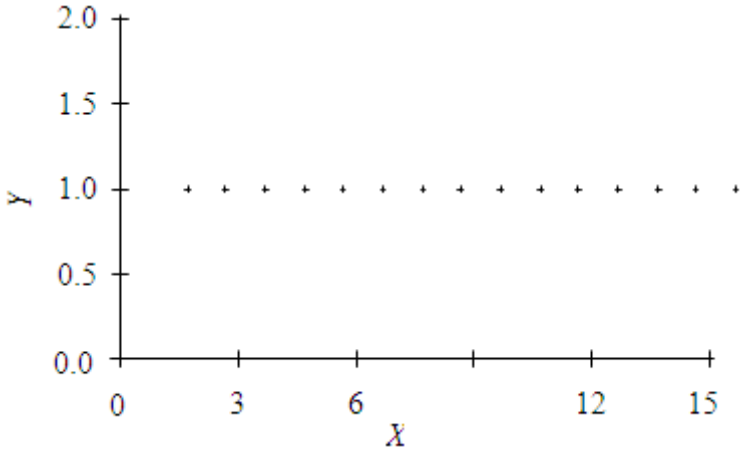


The data were classified as either of type 1 or type 2. Those of type 1 are indicated by o's and those of type 2 by x's. The overall correlation of the data in this scatterplot is:

- **positive.** *(True Answer)Correct*
- **negative, because the o's display a negative trend and the x's display a negative trend.** *Incorrect*
- **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.** *Incorrect*
- **impossible to compute for such a data set.** *Incorrect*

---

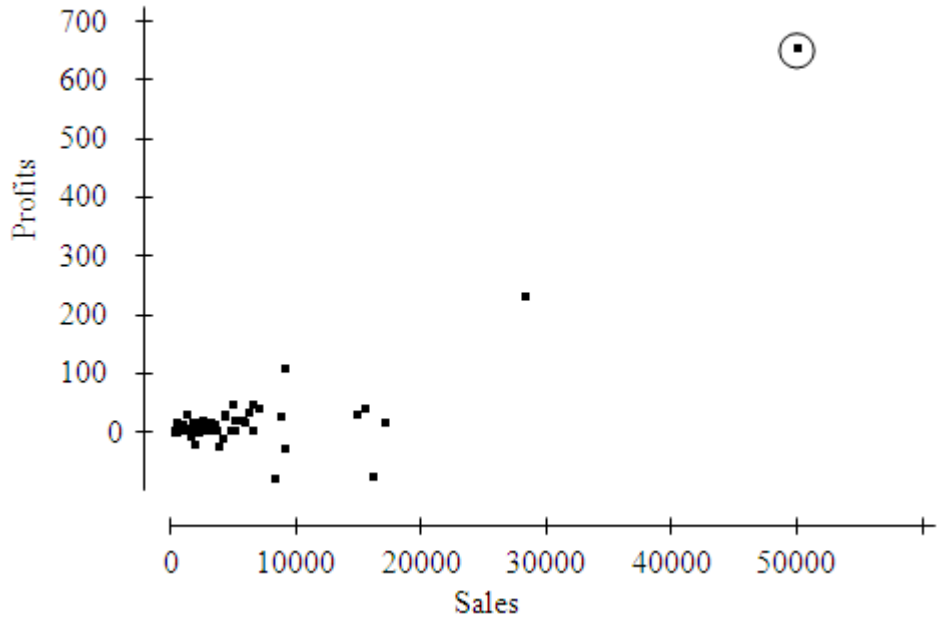
111 **A scatterplot of a variable *Y* versus a variable *X* produced the scatterplot below. The value of *Y* for all values of *X* is exactly 1.0. The correlation between *Y* and *X* is:**



- **+1.0 because the points lie perfectly on a line.** *Incorrect*
- **either +1.0 or –1.0, because the points lie perfectly on a line.** *Incorrect*
- **0 because *Y* does not change as *X* increases.** *(True Answer)Correct*
- **none of the above.** *Incorrect*

---

11 2 **The profits (in multiples of $100,000) versus the sales (in multiples of $100,000) for a number of companies are plotted below.**



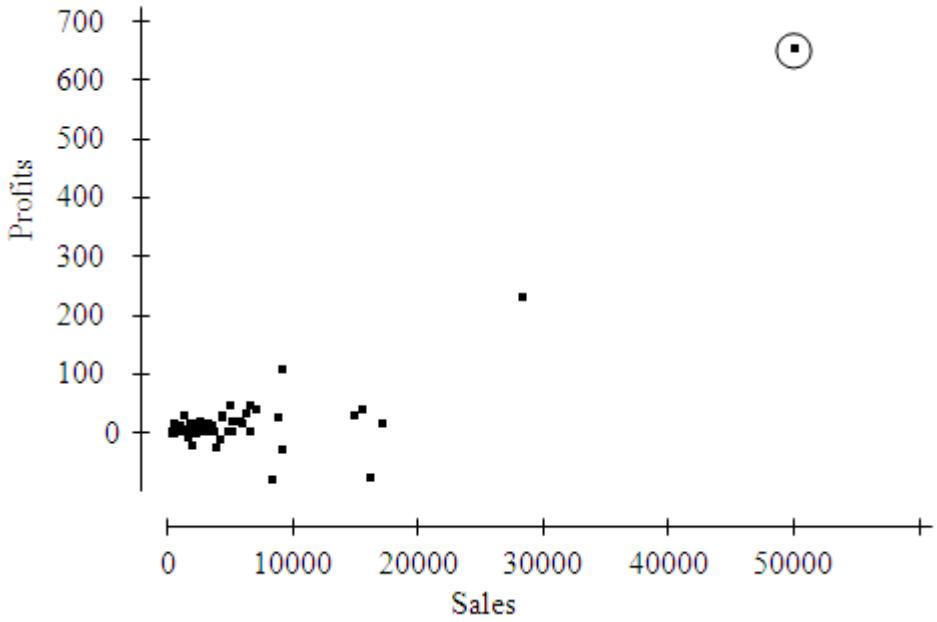
The correlation between profits and sales is 0.814. Suppose we removed the point that is circled from the data represented in the plot. The correlation between profits and sales would then be:

- **0.814.** *Incorrect*
- **larger than 0.814.** *Incorrect*
- **smaller than 0.814.** *(True Answer)Correct*
- **either larger or smaller than 0.814. It is impossible to say which.**

Incorrect

---

11 3 **The profits (in multiples of $100,000) versus the sales (in multiples of $100,000) for a number of companies are plotted below.**

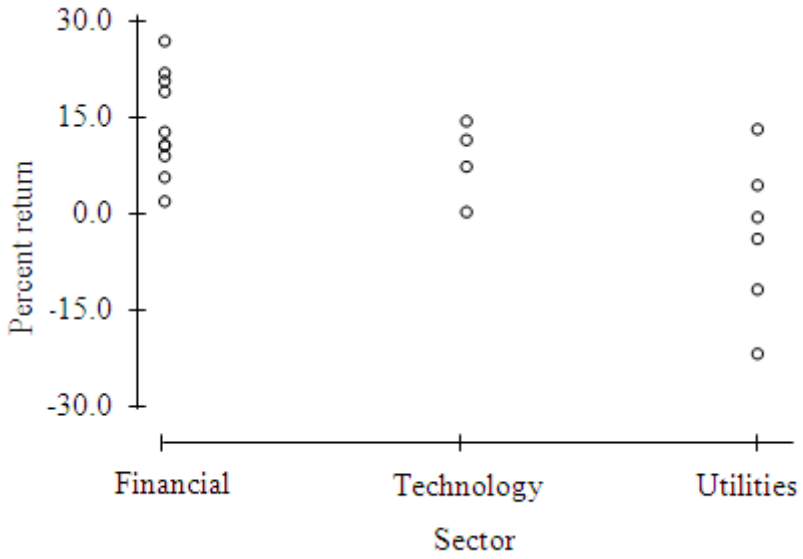


Notice that in the plot, profits is treated as the response variable and sales the explanatory variable. The correlation between profits and sales is 0.814. Suppose we had taken sales to be the response variable and profits to be the explanatory variable. In this case, the correlation between sales and profits would be:

- **0.814.** *(True Answer)Correct*
- **-0.814.** *Incorrect*
- **0.000.** *Incorrect*
- **Any number between -0.814 and 0.814, but one cannot state the exact value.** *Incorrect*

---

114 **A large mutual fund company offers sector funds. These sector funds concentrate their investments in narrow segments of the stock market. In this case, the sectors are financial services, technology, and utilities. The company offers a number of portfolios in each of these sectors. In the plot below, the percent return for these portfolios is plotted against the sector each represents.**

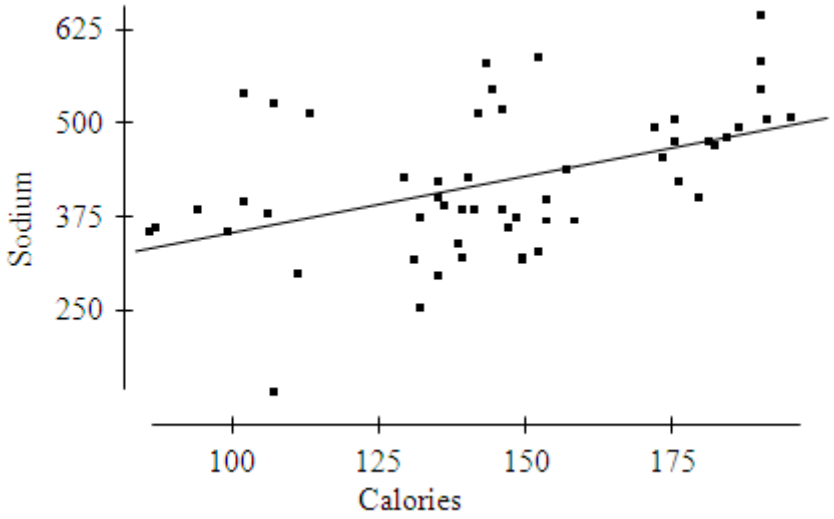


The correlation between percent return and sector is:

- **negative.** *Incorrect*
- **positive.** *Incorrect*
- **essentially 0 because the columns of points overlap in their values of percent return.** *Incorrect*
- **none of the above.** *(True Answer)Correct*

---

115 **The following is a scatterplot of the calories and sodium content of several brands of meat hot dogs. The least-squares regression line has been drawn in on the plot.**

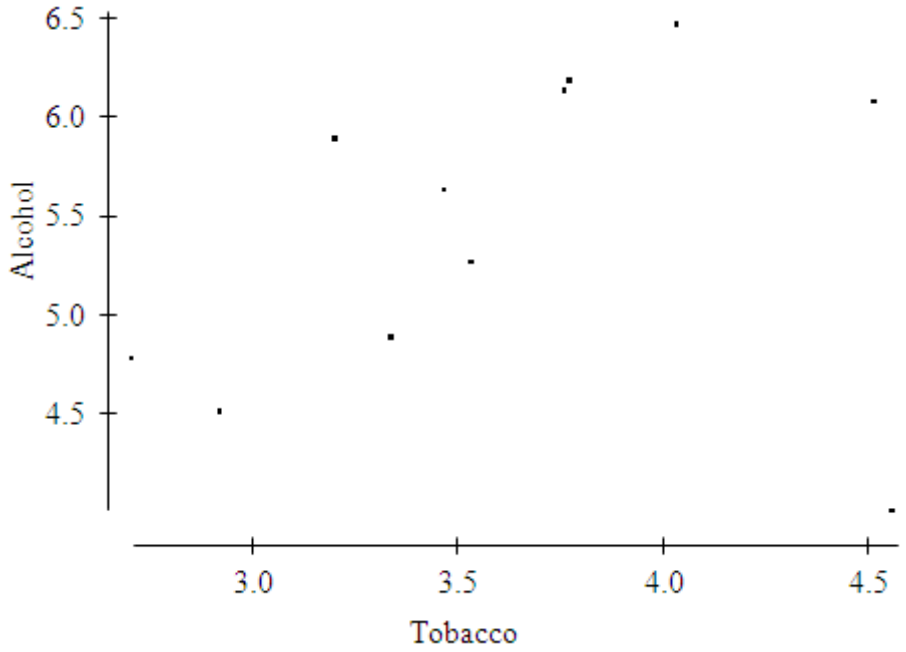


Referring to the scatterplot above, based on the least-squares regression line one would predict that a hot dog containing 100 calories would have a sodium content of about:

- **70.** *Incorrect*
- **350.** *(True Answer)Correct*
- **400.** *Incorrect*
- **600.** *Incorrect*

---

116 **The British government conducts regular surveys of household spending. The average weekly household spending on tobacco products and alcoholic beverages for each of 11 regions in Great Britain was recorded. A scatterplot of spending on tobacco versus spending on alcohol is given below.**



Which of the following statements holds?

- **The observation in the lower right corner of the plot is influential.** *(True Answer)Correct*
- **There is clear evidence of negative association between spending on alcohol and tobacco.** *Incorrect*
- **The equation of the least-squares line for this plot would be approximately $y = 10 - 2x$.** *Incorrect*
- **The correlation coefficient for this data is 0.99.** *Incorrect*

---

117 **The fraction of the variation in the values of y that is explained by the least-squares regression of y on x is:**

- **the correlation coefficient.** *Incorrect*
- **the slope of the least-squares regression line.** *Incorrect*
- **the square of the correlation coefficient.** *(True Answer)Correct*
- **the intercept of the least-squares regression line.** *Incorrect*

---

118 **In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the first test. The equation of the least-squares regression line was:**

$$y = 8.5 + 0.85x$$

where *y* represents the final exam score and *x* is the score on the first exam. Suppose Joe scores a 90 on the first exam. What would be the predicted value of his score on the final exam?

- **90.** *Incorrect*
- **85.** *(True Answer)Correct*
- **80.** *Incorrect*
- **The value cannot be determined from the information given. We also need to know the correlation.** *Incorrect*

---

119 **John's parents recorded his height at various ages up to 66 months. Below is a record of the results.**

Age (months)	36	48	54	60	66
Height (inches)	35	38	41	43	45

Reference: Ref 2-3

Which of the following is the equation of the least-squares regression line of John's height on age? (NOTE: You do not need to directly calculate the least-squares regression line to answer this question.)

- **Height = 12 × (Age)** *Incorrect*
- **Height = Age/12** *Incorrect*
- **Height = 60 – 0.22 × (Age)** *Incorrect*
- **Height = 22.3 + 0.34 × (Age)** *(True Answer)Correct*

---

120 **John's parents recorded his height at various ages up to 66 months. Below is a record of the results.**

Age (months)	36	48	54	60	66
Height (inches)	35	38	41	43	45

Reference: Ref 2-3

John's parents decide to use the least-squares regression line of John's height on age to predict his height at age 21 years (252 months). We conclude:

- **John's height, in inches, should be about half his age, in months.** *Incorrect*
- **that the parents will get a fairly accurate estimate of his height at age 21 years because the data are clearly correlated.** *Incorrect*
- **such a prediction could be misleading, because it involves extrapolation.** *(True Answer)Correct*
- **all of the above.** *Incorrect*

---

12 1 **A researcher wished to determine whether a company's profits can be used to predict the market value of the company. Based on data from a sample of over 80 companies from the Forbes 500 list, the researcher calculated the equation of the least-squares line for predicting market value from profits to be:**

Market value = 388.2 + 13.7(Profits)

The correlation between market value and profits would be:

- **1/13.7.** *Incorrect*
- **13.7/388.2.** *Incorrect*
- **positive, but we cannot say what the exact value is.** *(True Answer)Correct*
- **either positive or negative. It is impossible to say anything about the correlation from the information given.** *Incorrect*

---

122 **The least-squares regression line is:**

- **the line that makes the square of the correlation in the data as large as possible.** *Incorrect*
- **the line that makes the sum of the squares of the vertical distances of the data points from the line as small as possible.** *(True Answer)Correct*
- **the line that best splits the data in half, with half of the points above the line and half below the line.** *Incorrect*
- **all of the above.** *Incorrect*

---

123 **Which of the following is true of the least-squares regression line?**

- **The slope is the change in the response variable that would be predicted by a unit change in the explanatory variable.** *Incorrect*
- **It always passes through the point ($\bar{x}$, $\bar{y}$), the means of the explanatory and response variables, respectively.** *Incorrect*
- **It will only pass through all the data points if $r = \pm 1$.** *Incorrect*
- **All of the above.** *(True Answer)Correct*

---

124 **A pediatrician 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 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* = 0.84**
$\bar{x}$ = mean of the values of *X* = 5.69
$\bar{y}$ = mean of the values of *Y* = 6.26

s_x = standard deviation of the values of X = 3.23

s_y = standard deviation of the values of Y = 2.04

The slope of the least-squares line is:

- **0.53.** *(True Answer)Correct*
- **0.64.** *Incorrect*
- **0.84.** *Incorrect*
- **2.04.** *Incorrect*

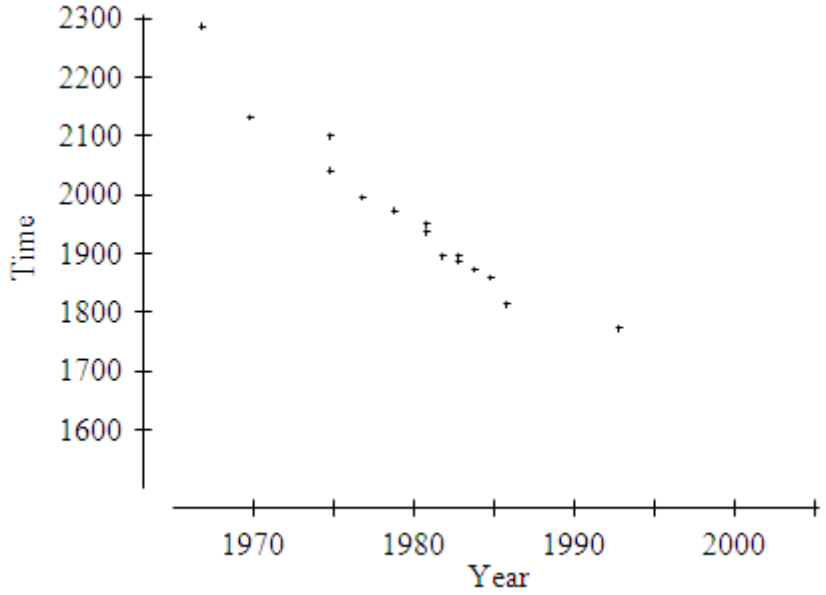
---

125 **Recall that when we standardize the values of a variable, the standardized value has mean 0 and standard deviation 1. Suppose we measure two variables, X and Y, on each of several subjects. We standardize both variables and then compute the least-squares regression line of Y on X for these standardized values. Suppose the slope of this least-squares regression line is –0.44. We may conclude:**

- **the intercept will be 1.0.** *Incorrect*
- **the intercept will also be –0.44.** *Incorrect*
- **the correlation will be 1.0.** *Incorrect*
- **the correlation will be –0.44.** *(True Answer)Correct*

---

126 **Below is a scatterplot of the world record time for women in the 10,000-meter run versus the year in which the record was set. Note that time is in seconds and the data are for the period 1965 to 1995.**



Based on this plot, we can expect:

- **by 2005, the world record time for women will be well below 1500 seconds.** *Incorrect*
- **that about every decade, the world record time will decrease**

by at least 100 seconds. *Incorrect*

- **that about every decade, the world record time will decrease by about 50 seconds.** *Incorrect*
- **none of the above.** *(True Answer)Correct*

---

127 **In a study of 2005 model cars, a researcher found that the fraction of the variation in the car's miles per gallon (MPG) which was explained by the least-squares regression on weight was about 0.64. For the cars in this study, the correlation between the car's MPG and its weight was found to be positive. The actual value of the correlation is:**

- **0.80.** *(True Answer)Correct*
- **0.64.** *Incorrect*
- **0.41.** *Incorrect*
- **impossible to determine from the information given.** *Incorrect*

---

128 **In the same study of 2005 model cars, a researcher computed the least-squares regression line of miles per gallon (MPG) on weight (in pounds). He obtained the following equation for this line:**

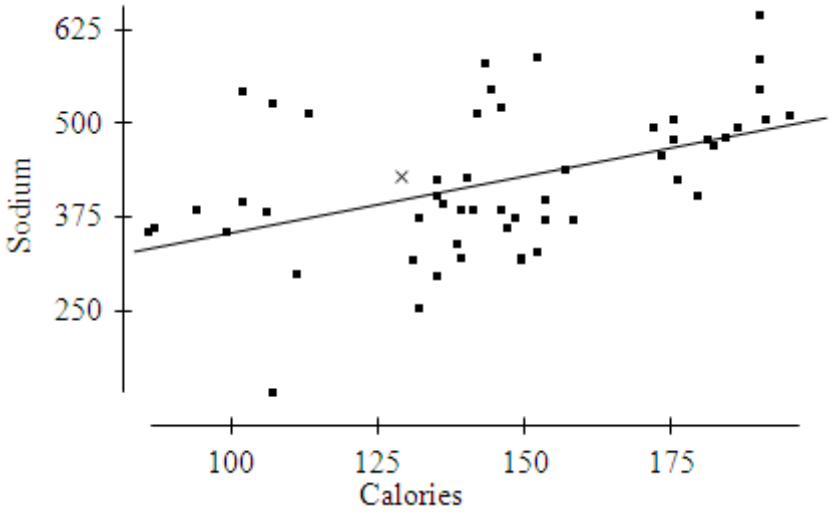
$$\text{MPG} = 34.941 - .004 \times \text{weight}$$

Based on the least-squares regression line, we would predict that a 2005 model car with weight equal to 3000 pounds would have an MPG of:

- **34.941.** *Incorrect*
- **30.931.** *Incorrect*
- **22.941.** *(True Answer)Correct*
- **26.921.** *Incorrect*

---

129 **The following is a scatterplot of the calories and sodium content of several brands of meat hot dogs. The least-squares regression line has been drawn in on the plot.**



Referring to the scatterplot above, the value of the residual for the point labeled x:

- **is about 40.** *(True Answer)Correct*
- **is about 1300.** *Incorrect*
- **is about 425.** *Incorrect*
- **cannot be determined from the information given.** *Incorrect*

---

130 **Researchers studied a sample of 100 adults between the ages of 25 and 35 and found a strong negative correlation between the annual salary and the number of pounds the individual was overweight. We may conclude which of the following?**

- **This is strong, but not conclusive, evidence that being overweight results in lower salaries.** *Incorrect*
- **If the annual salary and the number of pounds overweight for each individual in this study was plotted on a scatterplot, the points would lie close to a negatively sloping straight line.** *(True Answer)Correct*
- **If a larger sample of adults between the ages of 25 and 35 had been studied, the correlation would have been even stronger.** *Incorrect*
- **All of the above.** *Incorrect*

---

131 **A researcher wished to determine whether a company's profits can be used to predict the market value of the company. Based on data from a sample of 80 companies from the Forbes 500 list, the researcher calculated the equation of the least-squares regression line for predicting market value from profits to be:**

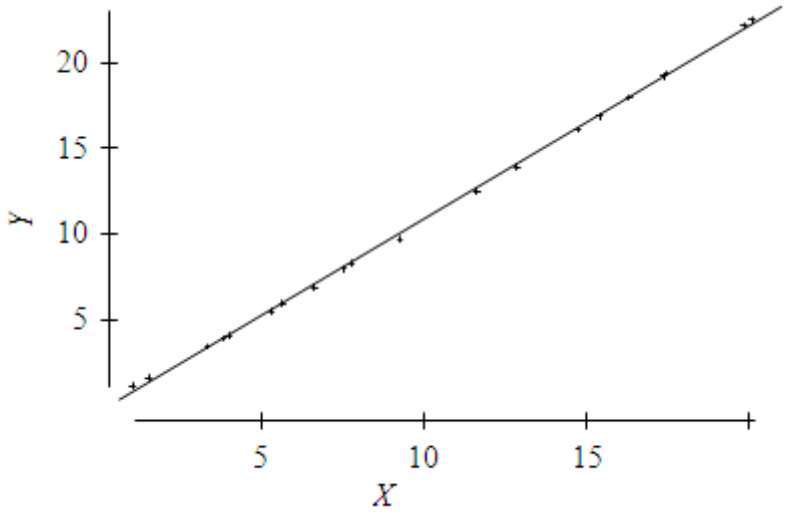
Market value = 388.2 + 13.7(Profits)

One company used by the researcher had profits of $36 million dollars and the market value of this company was $885 million. These values were used in the calculation of the least-squares regression line. The residual corresponding to these values is:

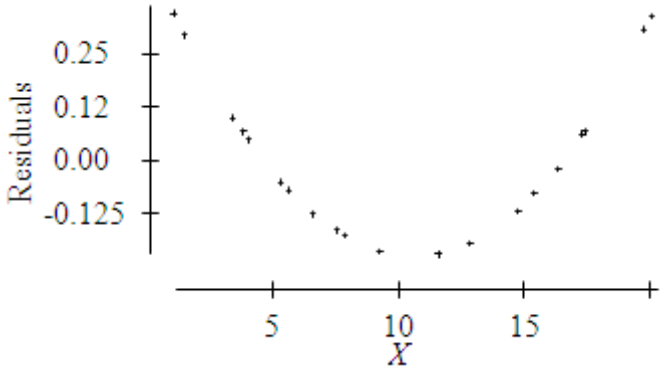
- **$3.6 million.** *(True Answer)Correct*
- **–$3.6 million.** *Incorrect*
- **$13.7 million.** *Incorrect*
- **–$849 million.** *Incorrect*

---

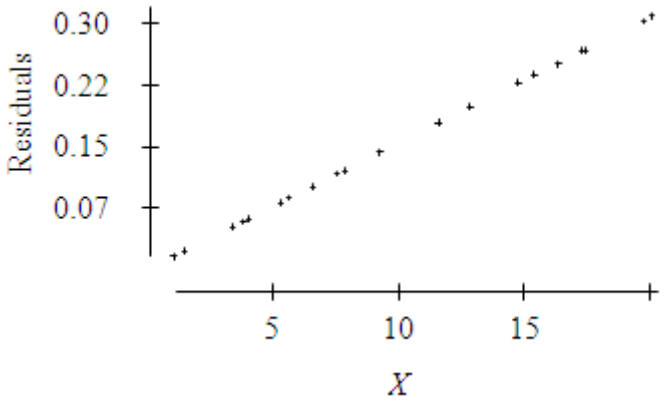
132 **A response *Y* and explanatory variable *X* were measured on each of several subjects. A scatterplot of the measurements is given below. The least-squares regression line is shown in the plot.**



Which of the following is a plot of the residuals for the above data versus *X*?

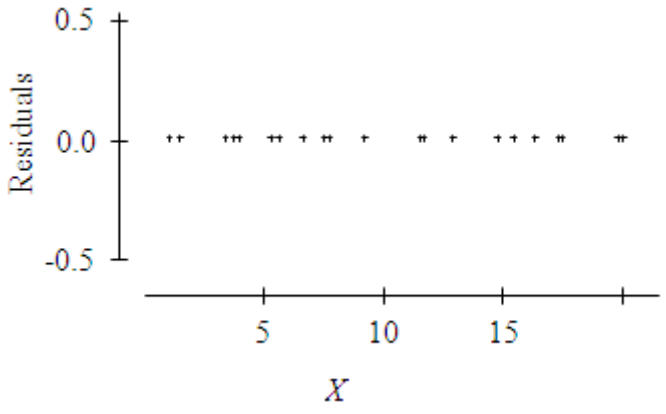
- 

(True Answer)Correct

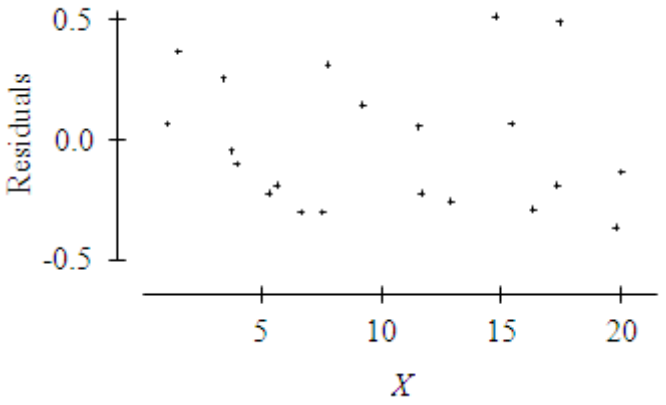
- 

Incorrect

-



Incorrect

- 

Incorrect

---

133 **The least-squares regression line is fit to a set of data. If one of the data points has a negative residual, then:**

- **the correlation between the values of the response and explanatory variables must be positive.** *Incorrect*
- **the point must lie below the least-squares regression line.** *(True Answer)Correct*
- **the point must lie near the left edge of the scatterplot.** *Incorrect*
- **all of the above.** *Incorrect*

---

134 **Which of the following statements concerning residuals is true?**

- **The sum of the residuals is always 0.** *Incorrect*
- **A plot of the residuals is useful for assessing the fit of the least-squares regression line.** *Incorrect*
- **The value of a residual is the observed value of the response minus the value of the response that one would predict from the least-squares regression line.** *Incorrect*
- **All of the above.** *(True Answer)Correct*

---

135 **The owner of a chain of supermarkets notices that there is a positive correlation between the sales of beer and the sales of ice cream over the course of the previous year. Seasons when sales of beer were above average, sales of ice cream also tended to be above average. Likewise, during seasons when sales of beer were below average, sales of ice cream also tended to be below average. Which of the following would be a valid conclusion from these facts?**

- **There must be an error. There should be no association between beer and ice cream sales.** *Incorrect*
- **Evidently, for a significant proportion of customers of these supermarkets, drinking beer causes a desire for ice cream or eating ice cream causes a thirst for beer.** *Incorrect*
- **A scatterplot of monthly ice cream sales versus monthly beer sales would show that a straight line describes the pattern in the plot, but it would have to be a horizontal line.** *Incorrect*
- **None of the above.** *(True Answer)Correct*

---

136 **A researcher studies the relationship between the total SAT score (Math SAT score plus Critical Reading SAT score plus Writing SAT score) and the grade point average (G.P.A.) of college students at the end of their freshman year. In order to use a relatively homogeneous group of students, the researcher examines only data of high school valedictorians (students who graduated at the top of their high school class) who have completed their first year of college. The researcher finds the correlation between total SAT score and G.P.A. at the end of the freshman year to be very close to 0. Which of the following would be a valid conclusion from these facts?**

- **Because the group of students studied is a very homogeneous group of students, the results should give a very accurate estimate of the correlation the researcher would find if all college students who have completed their freshman year were studied.** *Incorrect*
- **The correlation we would find if all college students who have completed their freshman year were studied would be even smaller than that found by the researcher. By restricting to valedictorians, the researcher is examining a group that will be more informative than those students who have only completed their freshman year.** *Incorrect*
- **The researcher made a mistake. Correlation cannot be calculated (the formula for correlation is invalid)**

unless all students who completed their freshman year are included. *Incorrect*

- **None of the above.** *(True Answer)Correct*

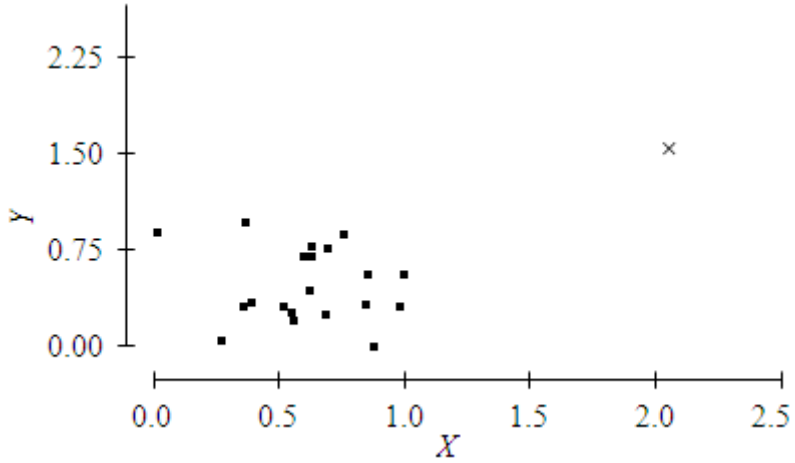
---

137 **When exploring very large sets of data involving many variables, which of the following is true?**

- **Extrapolation is safe because it is based on a greater quantity of evidence.** *Incorrect*
- **Associations will be stronger than would be seen in a much smaller subset of the data.** *Incorrect*
- **A strong association is good evidence for causation because it is based on a large quantity of information.** *Incorrect*
- **None of the above.** *(True Answer)Correct*

---

138 **Consider the scatterplot below.**

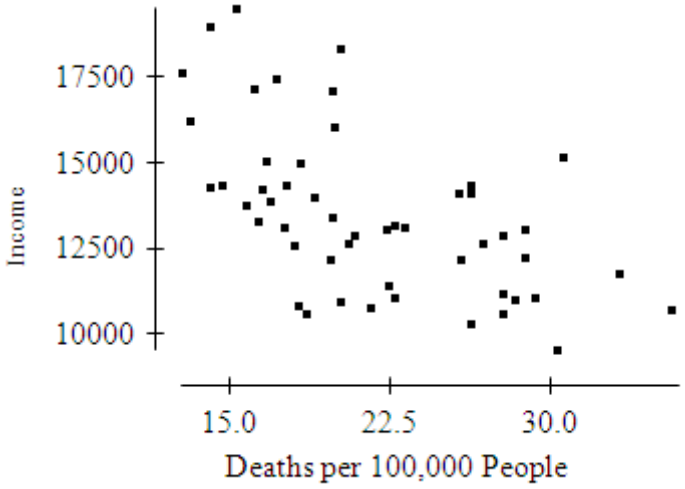


The point indicated by the plotting symbol x would be:

- **a residual.** *Incorrect*
- **influential.** *(True Answer)Correct*
- **a *z*-score.** *Incorrect*
- **a least-squares point.** *Incorrect*

---

139 **The following scatterplot displays the 1990 per capita income versus number of deaths due to traffic accidents per 100,000 people for each of the 50 states plus the District of Columbia.**



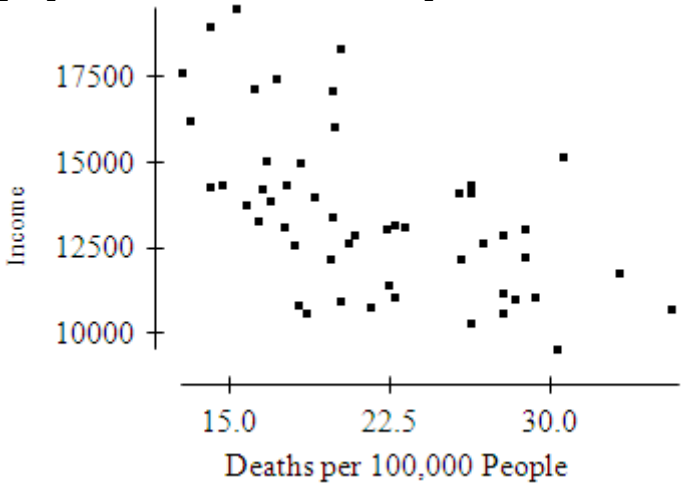
The correlation between income and deaths per 100,000 people is $r = -0.57$.
Reference: Ref 2-4

If instead of plotting these variables for each of the 50 states and the District of Columbia, we plotted the values of these variables for each county in the United States, we would expect the value of the correlation r to be:

- **exactly the same.** *Incorrect*
- **smaller.** *(True Answer)Correct*
- **+0.57 (the magnitude is the same, but the sign should change).** *Incorrect*
- **much higher and probably near 1 because there are many more counties than states.** *Incorrect*

---

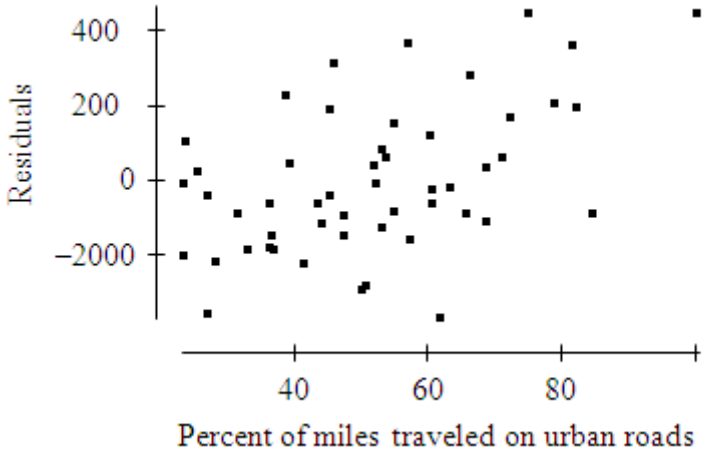
140 **The following scatterplot displays the 1990 per capita income versus number of deaths due to traffic accidents per 100,000 people for each of the 50 states plus the District of Columbia.**



The correlation between income and deaths per 100,000

people is $r = -0.57$.
Reference: Ref 2-4

The least-squares regression line was fit to the data in the scatterplot and the residuals computed. A plot of the residuals versus the percent of miles traveled on urban roads in the state is given below.



This plot suggests:

- **a high number of deaths per 100,000 people implies low per capita income, but only for states with a high percentage of miles traveled on urban roads.** *Incorrect*
- **a high number of deaths per 100,000 people implies low per capita income, but only for states with a low percentage of miles traveled on urban roads.** *Incorrect*
- **percentage of miles traveled on urban roads may be a lurking variable in understanding the association between income and deaths per 100,000 people.** *(True Answer)Correct*
- **none of the above.** *Incorrect*

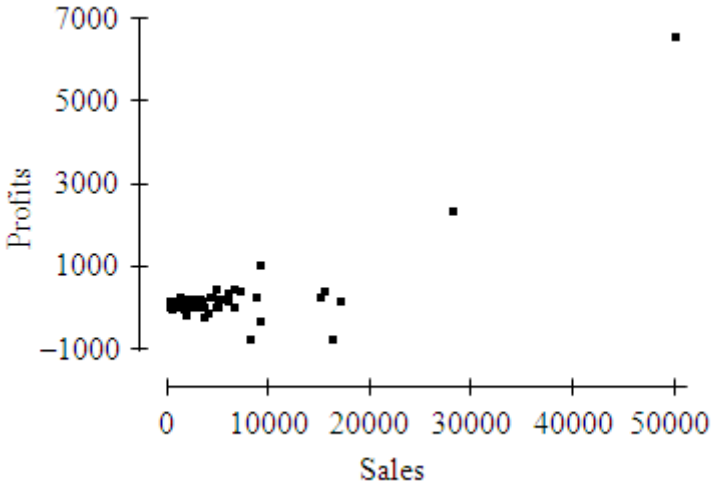
---

141 **Two variables, x and y, are measured on each of several individuals. The correlation between these variables is found to be 0.88. To interpret this correlation we should do which of the following?**

- **Compute the least-squares regression line of y on x and consider whether the slope is positive or negative.** *Incorrect*
- **Interchange the roles of x and y (i.e., treat x as the response and y as the predictor variable) and recompute the correlation.** *Incorrect*
- **Plot the data.** *(True Answer)Correct*
- **All of the above.** *Incorrect*

---

142 **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?

- **Not surprisingly, increasing sales causes an increase in profits. This is confirmed by the large positive correlation.** *Incorrect*
- **There are clearly influential observations present.** *(True Answer)Correct*
- **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.** *Incorrect*
- **All of the above.** *Incorrect*

---

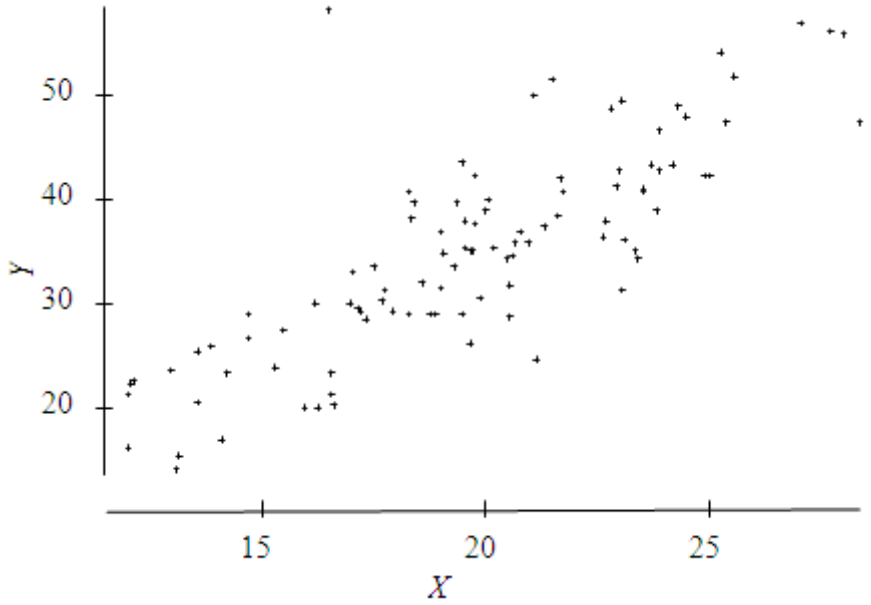
143 **A researcher computed the average Math SAT score of all high school seniors who took the SAT exam for each of the 50 states. The researcher also computed the average salary of high school teachers in each of these states and plotted these average salaries against the average Math SAT scores. The plot showed a distinct negative association between average Math SAT scores and teacher salaries. A second researcher conducted a similar study, but computed the average Math SAT score for each school district in the nation and plotted these against the average salary of high school teachers in the district. The association between average Math SAT score and average teacher salaries in the plot of the second researcher will most likely be:**

- **about the same as the association seen by the first researcher.** *Incorrect*

- **much stronger than that seen by the first researcher.** *Incorrect*
- **much stronger than that seen by the first researcher, but with the opposite sign.** *Incorrect*
- **weaker than that seen by the first researcher.** *(True Answer)Correct*

---

144 **Consider the following scatterplot.**



From this plot we can conclude:

- **there is evidence of a modest cause-and-effect relation between *X* and *Y* with increases in *X* causing increases in *Y*.** *Incorrect*
- **there is an outlier in the plot.** *(True Answer)Correct*
- **there is a strongly influential point in the plot.** *Incorrect*
- **all of the above.** *Incorrect*

---

145 **When possible, the best way to establish that an observed association is the result of a cause-and-effect relation is by means of:**

- **the least-squares regression line.** *Incorrect*
- **the correlation coefficient.** *Incorrect*
- **examining *z*-scores rather than the original variables.** *Incorrect*
- **a well-designed experiment.** *(True Answer)Correct*

---

146 **Which of the following would be necessary to establish a cause-and-effect relation between two variables?**

- **Strong association between the variables.** *Incorrect*
- **An association between the variables is observed in many different settings.** *Incorrect*
- **The alleged cause is plausible.** *Incorrect*

- **All of the above.** *(True Answer)Correct*

---

147 **Recent data show that states that spend an above average amount of money *X* per pupil in high school tend to have below average mean Critical Reading SAT scores *Y* of all students taking the SAT in the state. In other words, there is a negative association between *X* and *Y*. This is particularly true in states that have a large percentage of all high school students taking the exam. Such states also tend to have larger populations. The most plausible explanation for this association is:**

- ***X* causes *Y*. Overspending generally leads to extra, unnecessary programs, diverting attention from basic subjects. Inadequate training in these basic subjects generally leads to lower SAT scores.** *Incorrect*
- ***Y* causes *X*. Low SAT scores create concerns about the quality of education. This inevitably leads to additional spending to help solve the problem.** *Incorrect*
- **changes in *X* and *Y* are due to changes in other lurking variables. If a higher percentage of students take the exam, the average score will be lower. Also, states with larger populations have large urban areas where the cost of living is higher and more money is needed for expenses.** *(True Answer)Correct*
- **the association between *X* and *Y* is purely coincidental. It is implausible to believe the observed association could be anything other than accidental.** *Incorrect*

---

148 **A researcher observes that, on average, the number of businesses in cities with major league baseball teams is larger than in cities without major league baseball teams. The most plausible explanation for this observed association is:**

- **the presence of a major league baseball team causes the number of divorces to rise because the city becomes an attractive location for businesses.** *Incorrect*
- **the high number of businesses is responsible for the presence of a major league baseball team because cities with a healthy business environment are attractive to major league baseball teams.** *Incorrect*
- **the association is due to a lurking variable (major league teams tend to be in large cities with more people, hence a greater number of businesses).** *(True Answer)Correct*
- **the observed association is purely coincidental. It is implausible to believe the observed association could be**

anything other than accidental. *Incorrect*

---

149 **An article in the student newspaper of a large university had the headline "A's swapped for evaluations?" The article included the following:**

According to a new study, teachers may be more inclined to give higher grades to students, hoping to gain favor with the university administrators who grant tenure. The study examined the average grade and teaching evaluation in a large number of courses given in 1997 in order to investigate the effects of grade inflation on evaluations. "I am concerned with student evaluations because instruction has become a popularity contest for some teachers," said Professor Smith, who recently completed the study. Results showed higher grades directly corresponded to a more positive evaluation.
Reference: Ref 2-5

The underlined statement indicates the study found:

- **course grade is positively associated with teaching evaluation.** *(True Answer)Correct*
- **teaching evaluation is negatively associated with course grade.** *Incorrect*
- **there was a perfect positive correlation between course grade and teaching evaluation.** *Incorrect*
- **all of the above.** *Incorrect*

---

150 **An article in the student newspaper of a large university had the headline "A's swapped for evaluations?" The article included the following:**

According to a new study, teachers may be more inclined to give higher grades to students, hoping to gain favor with the university administrators who grant tenure. The study examined the average grade and teaching evaluation in a large number of courses given in 1997 in order to investigate the effects of grade inflation on evaluations. "I am concerned with student evaluations because instruction has become a popularity contest for some teachers," said Professor Smith, who recently completed the study. Results showed higher grades directly corresponded to a more positive evaluation.
Reference: Ref 2-5

Which of the following would be a valid conclusion to draw from the study?

- **A teacher can improve his or her teaching evaluations by giving good grades.** *Incorrect*

- **A good teacher, as measured by teaching evaluations, helps students learn better, resulting in higher grades.** *Incorrect*
- **Teachers of courses in which the mean grade is above average apparently tend to have above average teaching evaluations.** *(True Answer)Correct*
- **All of the above.** *Incorrect*

---

151 **A researcher computed the average Math SAT score of all high school seniors who took the SAT exam for each of the 50 states. The researcher also computed the average salary of high school teachers in each of these states and plotted these average salaries against the average Math SAT scores for each state. The plot showed a distinct negative association between average Math SAT scores and teacher salaries. The researcher may legitimately conclude which of the following?**

- **Increasing the average salary of teachers will cause the average of Math SAT scores to decrease, but it is not correct to conclude that increasing the salaries of individual teachers causes the Math SAT scores of individual students to increase.** *Incorrect*
- **States that pay teachers highly tend to do a poor job of teaching mathematics, on average.** *Incorrect*
- **States whose students tend to perform poorly in mathematics probably have a higher proportion of problem students and thus need to pay teachers higher salaries in order to attract them to teach in those states.** *Incorrect*
- **The data used by the researcher do not provide evidence that increasing the salary of teachers will cause the performance of students on the Math SAT to get worse.** *(True Answer)Correct*

---

152 **The average number of home runs hit by major league baseball players is greater now than it was three decades ago. A researcher suspects that the reason may be that baseballs are livelier now than 30 years ago. To check this he tested two baseballs, one that was manufactured 30 years ago (but never used) and one that was new. He noticed that the new baseball bounced higher than the older ball when both were dropped from the same height; i.e., the new baseball was livelier than the old one. The researcher can legitimately conclude:**

- **there is a positive association between the liveliness of the balls tested and the average number of home runs hit in the year that the ball was used.** *(True*

Answer)Correct

- **baseballs have been gradually getting livelier over the last three decades.** *Incorrect*
- **there is good evidence that the increase in the liveliness of baseballs has caused the increase in home run. This is because there is a positive association between liveliness of baseballs and average number of home runs hit, and because there is a plausible theory for the observed association.** *Incorrect*
- **all of the above.** *Incorrect*

---

153 **A researcher notices that in a sample of adults, those who take larger amounts of vitamin C have fewer illnesses. However, those who take larger amounts of vitamin C also tend to exercise more. As explanations for having fewer illnesses, the variables "amount of vitamin C taken" and "amount of exercise" are:**

- **skewed.** *Incorrect*
- **confounded.** *(True Answer)Correct*
- **lurking variables.** *Incorrect*
- **symmetric.** *Incorrect*

---

154 **In 1982, Kennesaw, Georgia passed a law requiring all citizens to own at least one gun. Although the law was never enforced, six months after the law was passed, the number of burglaries in that month was less than in the month prior to passage of the law. We may conclude which of the following?**

- **Gun ownership and burglary rates are negatively associated.** *Incorrect*
- **Gun ownership causes a reduction in crime. This is because there is a negative association between gun ownership and burglary rates, and because there is a plausible explanation for this association (gun ownership acts as a deterrent to crime).** *Incorrect*
- **Both a) and b).** *Incorrect*
- **Neither a) nor b).** *(True Answer)Correct*

---

155 **A study of the salaries of full professors at Upper Wabash Tech shows that the median salary for female professors is considerably less than the median male salary. Further investigation shows that the median salaries for male and female full professors are about the same in every department (English, physics, etc.) of the university. This apparent contradiction is an example of:**

- **extrapolation.** *Incorrect*

- **Simpson's paradox.** *(True Answer)Correct*
- **causation.** *Incorrect*
- **correlation.** *Incorrect*

---

156 **The reversal of the direction of an association when lurking variables are taken into account is called:**

- **Simpson's paradox.** *(True Answer)Correct*
- **least-squares regression.** *Incorrect*
- **negative association.** *Incorrect*
- **a residual plot.** *Incorrect*

---

157 **A survey asked each individual to choose their favorite sport from the list of baseball, football, basketball, and hockey (or another sport). The following two-way table categorizes a person's favorite sport by the gender.**

Sport	Male	Female
Baseball	13,959	2641
Football	3148	2469
Basketball	3222	709
Hockey	1457	690
Other	2039	1783

Which of the following statements is consistent with the table?

- **There is absolutely no evidence of a relationship between the gender of the person and their favorite sport.** *Incorrect*
- **More women were surveyed than men.** *Incorrect*
- **Men display a greater tendency to enjoy baseball than do women.** *(True Answer)Correct*
- **The correlation between these two variables is clearly positive.** *Incorrect*

---

158 **In a study of the link between high blood pressure and cardiovascular disease, a group of white male businessmen aged 35 to 64 was followed for five years. At the beginning of the study, each man had his blood pressure measured and it was classified as either "low" systolic blood pressure (less than 140 mmHg) or "high" blood pressure (140 mmHg or higher). The following table gives the number of men in each blood pressure category and the number of deaths from cardiovascular disease during the five-year period.**

Blood pressure	Deaths	Total
Low	10	2000
High	50	3500

Based on these data, which of the following statements is correct?

- **These data are consistent with the idea that there is a link between high blood pressure and death from cardiovascular disease.** *(True Answer)Correct*
- **The mortality rate (proportion of deaths) for men with high blood pressure is five times that of men with low blood pressure.** *Incorrect*
- **These data probably understate the link between high blood pressure and death from cardiovascular disease because men will tend to understate their true blood pressure.** *Incorrect*
- **All of the above.** *Incorrect*

---

159 **_X_ and _Y_ are two categorical variables. The best way to determine if there is a relationship between them is to:**

- **calculate the correlation between _X_ and _Y_.** *Incorrect*
- **draw a scatterplot of the _X_ and _Y_ values.** *Incorrect*
- **make a two-way table of the _X_ and _Y_ values.** *(True Answer)Correct*
- **do all of the above.** *Incorrect*

---

160 **A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $30,000 or $50,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the two tables below.**

	Managers			Workers	
	Male	Female		Male	Female
$100,000	80	20	$30,000	30	20
$200,000	20	30	$50,000	20	80

Reference: Ref 2-6

The proportion of male workers who make $50,000 per year is:

- **0.067.** *Incorrect*
- **0.133.** *Incorrect*
- **0.200.** *Incorrect*
- **0.400.** *(True Answer)Correct*

---

161 **A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $30,000 or $50,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the two tables below.**

	Managers			Workers	

	Male	Female		Male	Female
$100,000	80	20	$30,000	30	20
$200,000	20	30	$50,000	20	80

Reference: Ref 2-6

The proportion of female managers who make $200,000 per year is:

- **0.100.** *Incorrect*
- **0.200.** *Incorrect*
- **0.400.** *Incorrect*
- **0.600.** *(True Answer)Correct*

---

16 2 **A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $30,000 or $50,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the two tables below.**

	Managers			Workers	
	Male	Female		Male	Female
$100,000	80	20	$30,000	30	20
$200,000	20	30	$50,000	20	80

Reference: Ref 2-6

We may conclude:

- **the mean salary of female managers is greater than that of male managers.** *Incorrect*
- **the mean salary of males in this business is greater than the mean salary of females.** *Incorrect*
- **the mean salary of female workers is greater than that of male workers.** *Incorrect*
- **all of the above.** *(True Answer)Correct*

---

163 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of males that are registered as Democrats is:

- **0.19.** *Incorrect*
- **0.37.** *(True Answer)Correct*
- **0.43.** *Incorrect*

- **0.71.** *Incorrect*

---

164 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of registered Democrats who are male is:

- **0.19.** *Incorrect*
- **0.37.** *Incorrect*
- **0.43.** *(True Answer)Correct*
- **0.71.** *Incorrect*

---

165 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of all voters who are male and registered Democrats is:

- **0.19.** *(True Answer)Correct*
- **0.37.** *Incorrect*
- **0.43.** *Incorrect*
- **0.71.** *Incorrect*

---

166 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of females that are registered as

Republican is:

- **0.36.** *(True Answer)Correct*
- **0.38.** *Incorrect*
- **0.65.** *Incorrect*
- **0.71.** *Incorrect*

---

167 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of registered voters in the "Other" category who are female is:

- **0.54.** *Incorrect*
- **0.46.** *(True Answer)Correct*
- **0.10.** *Incorrect*
- **0.79.** *Incorrect*

---

168 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436
Republican	456	285
Other	94	79

Reference: Ref 2-7

The proportion of all voters who are Democrats is:

- **0.47.** *Incorrect*
- **0.37.** *Incorrect*
- **0.45.** *(True Answer)Correct*
- **0.54.** *Incorrect*

---

169 **A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.**

	Male	Female
Democrat	325	436

Republican	456	285
Other	94	79

Reference: Ref 2-7

The total number of all voters who are male is:

- **800.** *Incorrect*
- **761.** *Incorrect*
- **875.** *(True Answer)Correct*
- **1675.** *Incorrect*