Practice Tests 1 (Chapters 2-9)

**Identify the W's for the description of data.**

1) The State Athletic Association requires coaches to keep these records on all athletes: age, days absent, medical history, emergency contact, and any allergies the athlete may have.

   A) Who: Coaches; Cases: Each team is an individual case; What: Age, days absent, medical history, emergency contact, and allergy history; When: Current; Where: Not specified; Why: State requirement; How: Information is collected and stored as athletic records.
   
   B) Who: Athletes; Cases: Each athlete is an individual case; What: Years played, sport played, and position played; When: Current; Where: Not specified; Why: State requirement; How: Information is collected and stored as athletic records.

   C) Who: State Athletic Association; Cases: Each athlete is an individual case; What: Records; When: Current; Where: Not specified; Why: State requirement; How: Information is collected and stored as athletic records.

   D) Who: Athletes; Cases: Each athlete is an individual case; What: Records; When: Current; Where: Not specified; Why: Helpful information; How: Ask the athlete.

   E) Who: Athletes; Cases: Each athlete is an individual case; What: Age, days absent, medical history, emergency contact, and allergy history; When: Current; Where: Not specified; Why: State requirement; How: Information is collected and stored as athletic records.

**Classify the variable as categorical or quantitative.**

2) The color of your house

   A) Categorical
   
   B) Quantitative

3) The waiting time at a bus stop in minutes

   A) Quantitative
   
   B) Categorical

4) A person's political affiliation

   A) Categorical
   
   B) Quantitative

5) A person's height in feet

   A) Categorical
   
   B) Quantitative

**Name the variables in each description of data, then tell whether they are quantitative or categorical. For each quantitative variable name its unit of measure.**

6) A division I men's basketball coach is tracking player eligibility. He records each player's credits earned, the number of class absences each player has, and each player's GPA.

   A) Credits earned, quantitative, credits; number of class absences, quantitative, days; GPA, quantitative, scale from 0 to 4; player, categorical; coach, categorical.

   B) Credits earned, quantitative, decimals; number of class absences, quantitative, days; GPA, quantitative, decimals.

   C) Credits earned, quantitative, credits; number of class absences, quantitative, days; GPA, quantitative, scale from 0 to 4.

   D) Credits earned, quantitative, credits; number of class absences, quantitative, days.

   E) Credits earned, quantitative, decimals; number of class absences, quantitative, weeks; GPA, quantitative, GPA.
Provide an appropriate response.

7) The Centers for Disease Control lists causes of death for individual states in 2002. The mortality data for one state is given.

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Disease</td>
<td>29.2</td>
</tr>
<tr>
<td>Cancer</td>
<td>21.7</td>
</tr>
<tr>
<td>Circulatory diseases and stroke</td>
<td>8.6</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>5.7</td>
</tr>
<tr>
<td>Accidents</td>
<td>4.9</td>
</tr>
</tbody>
</table>

In this state, what percent of deaths were from causes not listed here?

A) 34.8%
B) 59.1%
C) 70.1%
D) 29.9%
E) The percent cannot be determined from the given percentages because the categories overlap.

8) The Centers for Disease Control lists causes of death for individual states in 2002. The mortality data for one state is given.

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Disease</td>
<td>28.1</td>
</tr>
<tr>
<td>Cancer</td>
<td>22.7</td>
</tr>
<tr>
<td>Circulatory diseases and stroke</td>
<td>7.4</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>6.5</td>
</tr>
<tr>
<td>Accidents</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Is it reasonable to conclude that, in this state, cancer or respiratory diseases were the cause of approximately 29% of deaths in 2002?

A) Yes, because the percentages can always be added in relative frequency tables.
B) No, because these categories overlap.
C) No, because the percentages do not add up to 100%.
D) No, because the percentages in relative frequencies tables can never be added.
E) Yes, because these categories do not overlap.

9) A magazine article reported on Springfield School District’s magnet school programs. Of the 1654 qualified applicants, 881 were accepted, 278 were wait-listed, and 495 were turned away for lack of space. Find the relative frequency distribution of the decisions made, and write a sentence describing it.

A) 1654 students applied for admission to the magnet schools program. 53% were accepted, 17% were wait-listed, and 30% were turned away.
B) 1654 students applied for admission to the magnet schools program. 53% were accepted, 32% were wait-listed, and 56% were turned away.
C) 1654 students applied for admission to the magnet schools program. 70% were accepted, and 30% were turned away.
D) 1654 students applied for admission to the magnet schools program. 53% were accepted, 32% were wait-listed, and 30% were turned away.
E) 1654 students applied for admission to the magnet schools program. 53% were accepted, and 47% were turned away.
10) A survey of autos parked in student and staff lots at a large university classified the brands by country of origin, as seen in the table.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Driver</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student</td>
<td>Staff</td>
</tr>
<tr>
<td>American</td>
<td>100</td>
<td>82</td>
</tr>
<tr>
<td>European</td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td>Asian</td>
<td>68</td>
<td>64</td>
</tr>
</tbody>
</table>

What is the conditional distribution of origin for staff?
A) 49% American, 18% European, 33% Asian
B) 27% American, 10% European, 18% Asian
C) 22% American, 5% European, 17% Asian
D) 50% American, 10% European, 39% Asian
E) 49% American, 15% European, 36% Asian

11) Students in a political science course were asked to describe their politics as "Liberal", "Moderate", or "Conservative." Here are the results:

<table>
<thead>
<tr>
<th>Politics</th>
<th>Liberal</th>
<th>Moderate</th>
<th>Conservative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>36</td>
<td>33</td>
<td>8</td>
<td>77</td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>37</td>
<td>26</td>
<td>112</td>
</tr>
<tr>
<td>Total</td>
<td>85</td>
<td>70</td>
<td>34</td>
<td>189</td>
</tr>
</tbody>
</table>

What percent of the females in the class consider themselves to be "Liberal"?
A) 43.8%  B) 42.9%  C) 45%  D) 46.8%  E) 42.4%

12) Students in a political science course were asked to describe their politics as "Liberal", "Moderate", or "Conservative." Here are the results:

<table>
<thead>
<tr>
<th>Politics</th>
<th>Liberal</th>
<th>Moderate</th>
<th>Conservative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>33</td>
<td>28</td>
<td>7</td>
<td>68</td>
</tr>
<tr>
<td>Male</td>
<td>43</td>
<td>48</td>
<td>19</td>
<td>110</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>76</td>
<td>26</td>
<td>178</td>
</tr>
</tbody>
</table>

What percent of the class considers themselves to be "Liberal"?
A) 48.5%  B) 18.5%  C) 74.5%  D) 42.7%  E) 24.2%

13) Students in a political science course were asked to describe their politics as "Liberal", "Moderate", or "Conservative." Here are the results:

<table>
<thead>
<tr>
<th>Politics</th>
<th>Liberal</th>
<th>Moderate</th>
<th>Conservative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>43</td>
<td>42</td>
<td>13</td>
<td>98</td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
<td>41</td>
<td>21</td>
<td>97</td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>83</td>
<td>34</td>
<td>195</td>
</tr>
</tbody>
</table>

What percent of all students in the class are males who consider themselves to be "Liberal"?
A) 22.1%  B) 17.9%  C) 36.1%  D) 40%  E) 44.9%
14) The Centers for Disease Control lists causes of death for individual states in 2002. The mortality data for one state is given.

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Disease</td>
<td>28.5</td>
</tr>
<tr>
<td>Cancer</td>
<td>21.4</td>
</tr>
<tr>
<td>Circulatory diseases and stroke</td>
<td>7.8</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>5.3</td>
</tr>
<tr>
<td>Accidents</td>
<td>4.7</td>
</tr>
</tbody>
</table>

Which of the following displays is/are appropriate for these data? (More than one display may be appropriate.)

I

II

III

A) I, II, III
B) I, III
C) I, II
D) I
E) None of these displays are appropriate.

15) The advantage of making a stem-and-leaf display instead of a dotplot is that a stem-and-leaf display
A) shows the shape of the distribution better than a dotplot.
B) preserves the individual data values.
C) satisfies the area principle.
D) A stem-and-leaf display has no advantages over a dotplot.
16) The five-number summary of credit hours for 24 students in a statistics class is:

<table>
<thead>
<tr>
<th>Min</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.0</td>
<td>15.0</td>
<td>16.5</td>
<td>18.0</td>
<td>22.0</td>
</tr>
</tbody>
</table>

From this information, we know that
A) there is at least one high outlier in the data.
B) there is at least one low outlier in the data.
C) there are no outliers in the data.
D) there are both low and high outliers in the data.
E) None of the above.

Describe the distribution (shape, center, spread, unusual features).

17) The histogram shows the sizes (in acres) of 169 farms in Lake County. In addition to describing the distribution, approximate the percentage of farms that are under 100 acres.

A) The distribution of the size of farms in Lake County is symmetric, with farm sizes ranging from 0 to 450 acres. The mode of the distribution is between 100 and 150 acres. It appears that 118 of 169 farms are under 100 acres, approximately 70%.
B) The distribution of the size of farms in Lake County is skewed to the right. Most of the farms are smaller than 150 acres, with some larger ones, from 150 to 300 acres. Five farms were larger than the rest, over 400 acres. The mode of the distribution is between 0 and 50 acres. It appears that 62 of 169 farms are under 100 acres, approximately 37%.
C) The distribution of the size of farms in Lake County is skewed to the right. Most of the farms are smaller than 150 acres, with some larger ones, from 150 to 300 acres. Five farms were larger than the rest, over 400 acres. The mode of the distribution is between 0 and 50 acres. It appears that 118 of 169 farms are under 100 acres, approximately 70%.
D) The distribution of the size of farms in Lake County is symmetric, with farm sizes ranging from 0 to 450 acres. The mode of the distribution is between 0 and 50 acres. It appears that 118 of 169 farms are under 100 acres, approximately 70%.
E) The distribution of the size of farms in Lake County is skewed to the right. Most of the farms are smaller than 50 acres, with some larger ones, from 150 to 300 acres. Five farms were larger than the rest, over 400 acres. The mode of the distribution is between 0 and 50 acres. It appears that 118 of 169 farms are under 100 acres, approximately 70%.
Create a boxplot that represents the given data.

18) Here are the test scores of 32 students:

32 37 41 44 46 48 53 55
56 57 59 63 65 66 68 69
70 71 74 75 77 78 79
80 82 83 86 89 92 95 99

![Boxplots](image)

A) I  B) II  C) III  D) IV  E) V

Find the median of the data.

19) The stem-and-leaf display shows the results of the mathematic test of 30 students.

```
10 | 0
9  | 1 1
8  | 0 0 1 1 1 2 3 5 5 6 7 8 9
7  | 5 5 7 8 8 9
6  | 5 7 7 8
5  | 4 8
4  |
3  | 3
2  | 9
1  |
0  |
```

A) 79  B) 81  C) 80  D) 87.5  E) 88.5
20) A small company employs a supervisor at $1300 a week, an inventory manager at $800 a week, 5 stock boys at $400 a week each, and 3 drivers at $500 a week each.

   A) $560  B) $400  C) $500  D) $450  E) $800

**Identify potential outliers, if there are any, in the given data.**

21) The National Education Association collects data on the number of years of teaching experience of high-school teachers. A sample taken this year of 19 high-school teachers yielded the following data on number of years of teaching experience.

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>24</td>
<td>1</td>
<td>34</td>
<td>15</td>
<td>7</td>
<td>19</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>19</td>
<td>16</td>
<td>10</td>
<td>21</td>
<td>25</td>
<td>14</td>
<td>39</td>
<td>18</td>
</tr>
</tbody>
</table>

   A) 34, 39  B) 1, 39  C) 1  D) 1, 34, 39  E) None

**Solve the problem.**

22) Here are summary statistics for the normal monthly precipitation (in inches) for August for 20 different U.S. cities.

<table>
<thead>
<tr>
<th>Count</th>
<th>Mean</th>
<th>Median</th>
<th>StdDev</th>
<th>Min</th>
<th>Max</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>3.23</td>
<td>3.45</td>
<td>1.2</td>
<td>0.4</td>
<td>7.0</td>
<td>2.1</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Would you describe this distribution as symmetric or skewed?

   A) Skewed to the left, because the mean is smaller than the median and the upper quartile is farther from the median than the lower quartile.
   B) Skewed to the right, because the mean is smaller than the median and the lower quartile is farther from the median than the upper quartile.
   C) Skewed to the left, because the mean is smaller than the median and the lower quartile is farther from the median than the upper quartile.
   D) Skewed to the right, because the mean is larger than the median and the upper quartile is farther from the median than the lower quartile.
   E) Symmetric, because the mean and the median are close.

23) Here are the prices for 8 different CD players. Find the standard deviation.

<table>
<thead>
<tr>
<th>$195</th>
<th>$358</th>
<th>$201</th>
<th>$276</th>
<th>$161</th>
<th>$301</th>
<th>$387</th>
<th>$128</th>
</tr>
</thead>
</table>

   A) $503,506.1  B) $161  C) $94.1  D) $565,441.0  E) $144.5
24) Office workers were asked how long it took them to travel to work one morning. Here is the stem-and-leaf display.

```
2 | 0 0 0 2 3 4 4 5 7 8
3 | 0 2 5 7
4 | 1 2 7 8 9
5 | 0 2 8
6 | 0 5
```

Would you use the median or the mean to describe the center of this distribution?

A) Mean, because the data are skewed to the right.
B) Median, because the data are skewed to the right.
C) Median, because the data are skewed to the left.
D) Mean, because the data are symmetric.
E) Mean, because the data are skewed to the left.

25) Office workers were asked how long it took them to travel to work one morning. Here is the stem-and-leaf display.

```
2 | 0 0 0 2 3 4 4 5 7 8
3 | 0 2 5 7
4 | 1 2 7 8 9
5 | 0 2 8
6 | 0 5
```

Without actually finding the mean and the median, would you expect the mean to be higher or lower than the median?

A) Higher, because the data are skewed to the left.
B) Lower, because the data are skewed to the right.
C) Lower, because the data are skewed to the left.
D) Higher, because the data are skewed to the right.
E) Neither, because the mean would be equal to the median.
26) Here are boxplots of the points scored during the first 10 games of the basketball season for both Caroline and Alexandra. Summarize the similarities and differences in their performance so far.

\[ \text{Caroline} \quad \text{Alexandra} \]

A) The girls have a different average score per game. Caroline is much more consistent, because her IQR is about 4 points, while Alexandra's is over 15.
B) Both girls have a median score of about 18 points per game. Caroline is much more consistent, because her IQR is about 6 points, while Alexandra's is over 20.
C) The girls have a different average score per game, but the same median score of about 18 points per game. Their IQR are different, but this does not give anymore information on the girls' performance.
D) Both girls have a median score of about 18 points per game. Alexandra is much more consistent, because her IQR is about 15 points, while Caroline's is over 3.
E) Both girls have a median score of about 18 points per game. Caroline is much more consistent, because her IQR is about 4 points, while Alexandra's is over 15.

27) Adam played golf on Saturday and Sunday. He scored 82 both days. The scores of all golfers Saturday averaged 69 with a standard deviation of 18. The scores on Sunday averaged 82 with a standard deviation of 10. On which day did Adam do better compared with the other golfers? Explain. (Hint: in golf smaller scores are better)

A) Saturday. A score of 82 Saturday is \( \frac{13}{18} \) standard deviations from the mean while a score of 82 Sunday is 0 standard deviations from the mean.
B) He did not do better either day, he scored the same.
C) Sunday. A score of 82 Sunday is 0 standard deviations from the mean while a score of 82 Saturday is \( \frac{13}{18} \) standard deviations from the mean.
D) Sunday. A score of 82 Sunday is \( \frac{13}{18} \) standard deviations from the mean while a score of 82 Saturday is 0 standard deviations from the mean.
E) Saturday. A score of 82 Saturday is 0 standard deviations from the mean while a score of 82 Sunday is \( \frac{13}{18} \) standard deviations from the mean.

Draw the Normal model and use the 68-95-99.7 Rule to answer the question.
28) Assuming a Normal model applies, a town’s average annual snowfall (in inches) is modeled by $N(46, 4)$. Draw and label the Normal model. Then find the interval for the middle 95% of snowfall.

A)

\[ \begin{array}{c}
\text{Snowfall (in.)} \\
34 \quad 38 \quad 42 \quad 46 \quad 50 \quad 54 \quad 58 \\
\end{array} \]

B)

\[ \begin{array}{c}
\text{Snowfall (in.)} \\
34 \quad 38 \quad 42 \quad 46 \quad 50 \quad 54 \quad 58 \\
\end{array} \]

C)

\[ \begin{array}{c}
\text{Snowfall (in.)} \\
30 \quad 34 \quad 38 \quad 42 \quad 46 \quad 50 \quad 58 \quad 62 \\
\end{array} \]

D)

\[ \begin{array}{c}
\text{Snowfall (in.)} \\
30 \quad 34 \quad 38 \quad 42 \quad 46 \quad 50 \quad 54 \quad 58 \quad 62 \\
\end{array} \]

E)

\[ \begin{array}{c}
\text{Snowfall (in.)} \\
38 \quad 42 \quad 46 \quad 50 \quad 54 \quad 58 \quad 62 \\
\end{array} \]

Find the percent of a standard Normal model found in the given region. Round to the nearest hundredth of a percent.

29) $z > -1.82$

A) 96.56%  
B) 92.57%  
C) -3.44%  
D) 46.56%  
E) 3.44%

In a standard Normal model, state what value(s) of $z$ cuts off the described region.

30) the lowest 40%

A) 0.25  
B) 0.50  
C) -0.25  
D) 0.57  
E) -0.57
31) The volumes of soda in quart soda bottles can be described by a Normal model with a mean of 32.3 oz and a standard deviation of 1.2 oz. What percentage of bottles can we expect to have a volume less than 32 oz?

A) 9.87%  
B) 59.87%  
C) 47.15%  
D) 38.21%  
E) 40.13%

Solve the problem. Round to the nearest tenth.

32) Based on the Normal model for snowfall in a certain town N(57, 8), how many inches of snow would represent the 25th percentile?

A) 65 inches  
B) 51.6 inches  
C) 14.3 inches  
D) 62.4 inches  
E) 49 inches

33) A correlation of zero between two quantitative variables means that

A) there is no linear association between the two variables.  
B) we have done something wrong in our calculation of $r$.  
C) re-expressing the data will guarantee a linear association between the two variables.  
D) there is no association between the two variables.  
E) None of the above.

Determine whether the scatterplot shows little or no association, a negative association, a linear association, a moderately strong association, or a very strong association (multiple associations are possible).

34)  

A) Negative association, linear association  
B) nonlinear association  
C) Little or no association  
D) Moderately strong association  
E) Linear positive association, moderately strong association
35) A) Positive association, moderately strong association, linear association
B) Nonlinear association
C) Linear association, moderately weak association
D) Negative association
E) Little or no association

36) A) Negative association, linear association
B) Little or no association
C) Positive association, linear association
D) Moderately strong association
E) Positive association, moderately strong association

Find the correlation.

37) The paired data below consist of the test scores of 6 randomly selected students and the number of hours they studied for the test.

<table>
<thead>
<tr>
<th>Hours</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>86</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
</tr>
<tr>
<td>10</td>
<td>59</td>
</tr>
<tr>
<td>9</td>
<td>87</td>
</tr>
</tbody>
</table>

A) -0.678   B) 0.678   C) 0.157   D) 0.224   E) -0.224
Several scatterplots are given with calculated correlations. Which is which?

38) 1) \[ \begin{array}{c}
\text{A) 1d, 2a, 3b, 4c} \\
\text{B) 1b, 2c, 3a, 4d} \\
\text{C) 1a, 2b, 3c, 4d} \\
\text{D) 1b, 2c, 3d, 4a} \\
\text{E) 1c, 2a, 3d, 4c}
\end{array} \]


Solve the problem.

39) Soda is often considered unhealthy because its content is high in both caffeine and refined sugar. But are the two variables related? The correlation between caffeine and refined sugar is 0.187. What does it tell about the association between the two variables? Describe the association.

A) Weak nonlinear relation in a positive direction
B) Weak curved relation in a positive direction
C) Strong linear relation
D) Weak linear relation
E) No evidence of linear relation.
Tell what the residual plot indicates about the appropriateness of the linear model that was fit to the data.

40) 

A) Model is not appropriate. The relationship is nonlinear.  
B) Model is appropriate.  
C) Model may not be appropriate. The spread is changing.

41) 

A) Model may not be appropriate. The spread is changing.  
B) Model is not appropriate. The relationship is nonlinear.  
C) Model is appropriate. 

Use the model to make the appropriate prediction.

42) A golf ball is dropped from 15 different heights (in inches) and the height of the bounce is recorded (in inches.)  

The regression analysis gives the model \( \text{bounce} = 0.5 + 0.69 \text{drop} \). Predict the height of the bounce if dropped from 70 inches.  
A) 47.8 inches  
B) 71.19 inches  
C) 48.8 inches  
D) 48.3 inches  
E) 100.72 inches

Answer the question.
43) Which of the following scatterplots of residuals suggests that a linear model may not be applicable?

A) 

B) 

C) 

D)
44) Which of the labeled points below are outliers?

A) Points A, B, and D  
B) Point D  
C) Points B and D  
D) Points A, B, C, and D  
E) Points A and B

Answer the question.

45) Which of the labeled points below will exert the largest leverage on a linear model of the data?

A) Point C  
B) Point B  
C) Point E  
D) Point D  
E) Point A
46) An economist noticed that nations with more TV sets have higher life expectancy. He established a strong positive association between length of life and number of TV sets. Describe three different possible cause- and- effect relationships that might be present.
   A) Perhaps more TV sets cause higher life expectancy, higher life expectancy cause more TV sets, or both could be caused by a lurking variable such as the TV sets brand.
   B) Perhaps more TV sets cause higher life expectancy, higher life expectancy cause more TV sets, or both could be caused by a lurking variable such as the wealth of the nation.
   C) There is only one cause- and- effect relationship: more TV sets cause higher life expectancy.
   D) There is only one cause- and- effect relationship: higher life expectancy cause more TV sets.
   E) There are only two cause- and- effect relationships: more TV sets cause higher life expectancy, or higher life expectancy cause more TV sets.

47) All but one of the statements below contain a mistake. Which one could be true?
   A) The correlation between the breed of a dog and its weight is 0.435.
   B) The correlation between height and weight is 0.568 inches per pound.
   C) If the correlation between blood alcohol level and reaction time is 0.73, then the correlation between reaction time and blood alcohol level is -0.73.
   D) The correlation between height and weight is 0.568.
   E) The correlation between gender and age is -0.171.

48) A residuals plot is useful because
   I. it will help us to see whether a linear model makes sense.
   II. it might show a pattern in the data that was hard to see in the original scatterplot.
   A) I only                    B) Neither I nor II                   C) I and II                    D) II only

49) A regression analysis of students’ college grade point averages (GPAs) and their high school GPAs found $R^2 = 0.311$. Which of these is true?
   I. High school GPA accounts for 31.1% of college GPA.
   II. 31.1% of college GPAs can be correctly predicted with this model.
   III. 31.1% of the variance in college GPA can be accounted for by the model
   A) II only                     B) I and II                       C) III only                    D) I only                    E) None

50) When using midterm exam scores to predict a student’s final grade in a class, the student would prefer to have a
   A) positive residual, because that means the student’s final grade is higher than we would predict with the model.
   B) positive residual, because that means the student’s final grade is lower than we would predict with the model.
   C) negative residual, because that means the student’s final grade is lower than we would predict with the model.
   D) negative residual, because that means the student’s final grade is higher than we would predict with the model.
   E) residual equal to zero, because that means the student’s final grade is exactly what we would predict with the model.

51) Which is true?
   I. Random scatter in the residuals indicates a model with high predictive power.
   II. If two variables are very strongly associated, then the correlation between them will be near +1.0 or -1.0.
   III. The higher the correlation between two variables the more likely the association is based in cause and effect.
   A) I and II only                B) I only                  C) I, II, and III              D) none                    E) II only
52) A regression analysis of company profits and the amount of money the company spent on advertising found $r^2 = 0.72$. Which of these is true?

I. The model can correctly predict the profit for 72% of companies.
II. On average, about 72% of a company’s profit results from advertising.
III. On average, companies spend about 72% of their profits on advertising.

A) III only  B) none  C) I and III  D) I only  E) II only

Provide an appropriate response.

53) Is there a relationship between the raises administrators at State University receive and their performance on the job? A faculty group wants to determine whether job rating (x) is a useful linear predictor of raise (y). Consequently, the group considered the straight-line regression model, $\hat{y} = \beta_0 + \beta_1 x$. Using the method of least squares, the faculty group obtained the following prediction equation, $\hat{y} = 14,000 - 2,000 x$. Interpret the estimated slope of the line.

A) For an administrator with a rating of 1.0, we estimate his/her raise to be $2,000.
B) For a 1-point increase in an administrator’s rating, we estimate the administrator’s raise to increase $2,000.
C) For a $1 increase in an administrator’s raise, we estimate the administrator’s rating to decrease 2,000 points.
D) For a 1-point increase in an administrator’s rating, we estimate the administrator’s raise to decrease $2,000.

54) A large national bank charges local companies for using its services. A bank official reported the results of a regression analysis designed to predict the bank’s charges (y), measured in dollars per month, for services rendered to local companies. One independent variable used to predict service charge to a company is the company’s sales revenue (x), measured in millions of dollars. Data for 21 companies who use the bank’s services were used to fit the model, $\hat{y} = \beta_0 + \beta_1 x$. The results of the simple linear regression are provided below.

$\hat{y} = 2,700 + 20x, s = 65, 2$-tailed p-value = 0.064 (for testing $\beta_1$)

Interpret the estimate of $\beta_0$, the y-intercept of the line.

A) About 95% of the observed service charges fall within $2,700 of the least squares line.
B) All companies will be charged at least $2,700 by the bank.
C) There is no practical interpretation since a sales revenue of $0$ is a nonsensical value.
D) For every $1 million increase in sales revenue, we expect a service charge to increase $2,700.