1. Three scatterplots are shown below. The calculated correlations are 0.628, –0.914, and –0.033. Determine which correlation goes with which scatterplot.

Match each scatterplot with its corresponding correlation.

(a) (1) __________
(b) (2) __________
(c) (3) __________

(1) ○ 0.628  ○ –0.914  ○ –0.033
(2) ○ 0.628  ○ –0.914  ○ –0.033
(3) ○ 0.628  ○ –0.914  ○ –0.033
2. Shown below is the output from a linear model predicting armspan (in cm) from height (in inches) and summary statistics. Assume that the association between armspan and height is linear. Use the output and summary statistics to complete parts a through d.

<table>
<thead>
<tr>
<th>LinReg</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y = a + bx$</td>
<td>Height, $x$</td>
<td>63.77</td>
</tr>
<tr>
<td>$a = 29.624114162$</td>
<td>Armspan, $y$</td>
<td>159.59</td>
</tr>
<tr>
<td>$b = 2.037968274$</td>
<td>$r^2 = 0.823933545$</td>
<td>$r = 0.907707852$</td>
</tr>
</tbody>
</table>

a. Report the regression equation, using the words "Height" and "Armspan," not x and y.

- A. Predicted Armspan = 29.62 + 2.04 Height
- B. Predicted Armspan = 2.04 + 29.62 Height
- C. Predicted Height = 2.04 + 29.62 Armspan
- D. Predicted Height = 29.62 + 2.04 Armspan

b. Determine the corresponding values and verify the slope by using the formula $b = \frac{s_y}{s_x}$.

- $r =$ ___________ (Round to three decimal places as needed.)
- $s_y =$ ___________ (Round to two decimal places as needed.)
- $s_x =$ ___________ (Round to two decimal places as needed.)
- $b =$ ___________ (Round to two decimal places as needed.)

c. Determine the corresponding values and verify the y-intercept by using the formula $a = \bar{y} - b\bar{x}$.

- $\bar{y} =$ ___________ (Round to two decimal places as needed.)
- $\bar{x} =$ ___________ (Round to two decimal places as needed.)
- $a =$ ___________ (Round to two decimal places as needed.)

d. Using the regression equation from part a, predict the armspan (in cm) for someone 60 inches tall.

Predicted Armspan = ___________ cm

(Round to one decimal place as needed.)
3. Measurements were made for a sample of adult men. A regression line was fit to predict the men's armspan from their height. The output from several statistical technologies is provided in the accompanying charts. The scatterplot confirms that the association between armspan and height is linear. Complete parts a and b.

Click the icon to see the scatterplot and the output from several statistical technologies.

a. Report the equation for predicting armspan from height. Use words such as armspan, not just x and y.

Predicted _______ = _______ + _______ (2) _______

(Round to three decimal places as needed.)

b. Explain how to find the slope and intercept for each of the provided outputs.

Simple linear regression results:
Dependent Variable: Armspan
Independent Variable: Height
Armspan = 11.6566038 + 2.417453 Height
Sample size: 15
R(correlation coefficient) = 0.8986
R-sq = 0.80744563
Estimate of error of standard deviation: 3.892501

C. In the fourth line, the slope is the constant, and the intercept is multiplied by the Height.

D. In the fourth line, the slope is multiplied by the Height, and the intercept is the constant.

Coefficients

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>11.656604</td>
</tr>
<tr>
<td>X Variable</td>
<td>2.417453</td>
</tr>
</tbody>
</table>

LinReg
\[ y = a + bx \]
\[ a = 11.656603788 \]
\[ b = 2.41745283 \]
\[ r^2 = 0.80744563 \]
\[ r = 0.898579785 \]
A: The a-value is the intercept; and the b-value is the slope.

C. The r-value is the intercept, and the a-value is the slope.

D. The r-value is the intercept, and the b-value is the slope.

1: Statistical Outputs

The regression equation is
Armspan = 11.7 + 2.42 Height

Simple linear regression results:
Dependent Variable: Armspan
Independent Variable: Height
Armspan = 11.6566038 + 2.417453 Height
Sample size: 15
R(correlation coefficient) = 0.8986
R-sq = 0.80744563
Estimate of error of standard deviation: 3.892501

<table>
<thead>
<tr>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>X Variable</td>
</tr>
</tbody>
</table>

LinReg
y = a + bx
a = 11.656603788
b = 2.41745283
r^2 = 0.80744563
r = 0.898579785

(1) ☐ Armspan (2) ☐ Armspan
☐ Height ☐ Height
4. Explain what makes this scatterplot hard to interpret. What should have been done differently?

Choose the correct answer below.

- A. The vertical axis starts at 52. To zoom out of the data, the vertical axis should have been started at about 0.
- B. The vertical axis starts at 52. To zoom in on the data, the vertical axis should have been started at about 55.
- C. The horizontal axis starts at 0. To zoom in on the data, the horizontal axis should have been started at about 60.
- D. The horizontal axis ends at 100. To zoom out of the data, the horizontal axis should have been ended at about 200.

5. The correlation between height and arm span in a sample of adult women was found to be $r = 0.935$. The correlation between arm span and height in a sample of adult men was found to be $r = 0.853$. Which association—the association between height and arm span for women, or the association between height and arm span for men—is stronger? Explain.

Choose the correct answer below.

- A. The association between height and arm span for men is stronger because the value of $r$ is closer to 0.
- B. The association between height and arm span for men is stronger because the value of $r$ is farther from 0.
- C. The association between height and arm span for women is stronger because the value of $r$ is farther from 0.
- D. The association between height and arm span for women is stronger because the value of $r$ is closer to 0.
6. The figure shows a scatterplot with the regression line. The equation of the regression line is shown below. The data are for the 50 regions in Country A. The predictor is the percentage of smoke-free homes. The response is the percentage of high school students who smoke. Complete parts a and b.

Predicted Pct. Smokers = 52.082 − 0.410(Pct. Smoke-free)

a. Explain what the trend shows. Choose the correct answer below.

☐ A. Approximately 41% of the people in Country A smoke.
☐ B. Approximately 41% of the people in Country A do not smoke.
☐ C. The higher the percentage of smoke-free homes in a region, the lower the percentage of high school students who smoke tends to be.
☐ D. The lower the percentage of smoke-free homes in a region, the lower the percentage of high school students who smoke tends to be.

b. Use the regression equation to predict the percentage of students in high school who smoke, assuming that there are 70% smoke-free homes in the region. Use 70 not 0.70.

___________ % (Round to the nearest integer as needed.)

7. The figure shows a graph of the death rate in automobile accidents and the age of the driver.

a. Explain what the graph tells us about drivers at different ages; state which ages show the safest drivers and which show the most dangerous drivers.

b. Explain why it would not be appropriate to use these data for linear regression.

a. Choose the correct answer below.

☐ A. The graph shows that drivers between about 40 and 60 years of age have more fatalities and that the safest drivers are drivers older than 80.
☐ B. The graph shows that teenage drivers have more fatalities and that the safest drivers are drivers older than 80.
☐ C. The graph shows that young drivers and old drivers have more fatalities and that the safest drivers are between about 40 and 60 years of age.
☐ D. The graph shows that drivers between about 40 and 60 years of age have more fatalities and that the safest drivers are teenage drivers.

b. Choose the correct answer below.

☐ A. It would not be appropriate for linear regression because the trend is not linear.
☐ B. It would not be appropriate for linear regression because there is not enough data.
☐ C. It would not be appropriate for linear regression because the trend is linear.
☐ D. It would not be appropriate for linear regression because there is too much data.
8. Answer the questions using complete sentences.
a. What is an influential point?

b. It has been noted that people who go to church frequently tend to have lower blood pressure than people who don't go to church. Does this mean you can lower your blood pressure by going to church? Why or why not? Explain.

a. Choose the correct answer below.

   O A. An influential point is a point that changes the regression equation by a large amount.
   O B. An influential point is used in the regression line to make predictions beyond the range of the data.
   O C. An influential point is a point that measures the strength of the linear association between two numerical variables.

b. Choose the correct answer below.

   O A. Going to church may not cause lower blood pressure. Just because two variables are related does not show that one caused the other.
   O B. Since the two variables are not related, going to church may not cause lower blood pressure.
   O C. Since the two variables are related, going to church may not cause lower blood pressure.

9. If the correlation between height and weight of a large group of people is 0.74, find the coefficient of determination (as a percent) and explain what it means. Assume that height is the predictor and weight is the response, and assume that the association between height and weight is linear.

Choose the correct answer below.

   O A. The coefficient of determination is 54.76%. Therefore, 54.76% of the variation in weight can be explained by the regression line.
   O B. The coefficient of determination is 74%. Therefore, 74% of the variation in weight can be explained by the regression line.
   O C. The coefficient of determination is 54.76%. Therefore, 54.76% of the variation in height can be explained by the regression line.
   O D. The coefficient of determination is 74%. Therefore, 74% of the variation in height can be explained by the regression line.
10. The equation for the regression line relating the salary and the year first employed is given above the figure.
   a. Report the slope and explain what it means.
   b. Either interpret the y-intercept of 4,255,424 or explain why it is not appropriate to interpret the y-intercept.

a. Choose the correct answer below.
   - A. A salary of $2099 would be the salary for a person who started in the year 0.
   - B. A salary of $2099 would be the salary for a person who started in the year 2010.
   - C. The average salary is $2099 less for each year later that the person was hired or an average of $2099 more for each year earlier.
   - D. The average salary is $2099 more for each year later that the person was hired or an average of $2099 less for each year earlier.

b. Choose the correct answer below.
   - A. The y-intercept of $4,255,424 would be the salary for a person who started prior to 1985.
   - B. The y-intercept of $4,255,424 would be the salary for a person who started after 2000.
   - C. The y-intercept of $4,255,424 would be the salary for a person who started in the year 0, which is not appropriate to interpret.
11. The figure shows a scatterplot with a regression line for teachers' average pay and the expenditure per pupil for each state for public schooling in 2007.

a. From the graph, is the correlation between teachers’ average pay and the expenditure per pupil positive or negative?
b. Interpret the slope.
c. Interpret the intercept or explain why it should not be interpreted.

a. Choose the correct answer below.

○ A. The correlation is positive because the graph shows a decreasing trend.
○ B. The correlation is negative because the graph shows an increasing trend.
○ C. The correlation is positive because the graph shows an increasing trend.
○ D. The correlation is negative because the graph shows a decreasing trend.

b. Choose the correct answer below.

○ A. For each additional dollar spent on teachers’ pay, the expenditure per pupil, on average, is about 23 cents lower.
○ B. For each additional dollar spent on teachers’ pay, the expenditure per pupil, on average, is about 14.24 dollars lower.
○ C. For each additional dollar spent on teachers’ pay, the expenditure per pupil, on average, is about 14.24 dollars higher.
○ D. For each additional dollar spent on teachers’ pay, the expenditure per pupil, on average, is about 23 cents higher.

c. Choose the correct answer below.

○ A. The intercept represents the average pay the teachers’ salary changes.
○ B. The intercept represents the mean cost of education excluding the cost of paying teachers.
○ C. The intercept might represent the cost of education when the teachers are paid nothing. However, doing so requires extrapolation which may not be valid.
12. Assume that in a political science class, the teacher gives a midterm exam and a final exam. Assume that the association between midterm and final scores is linear. The summary statistics shown below have been simplified for clarity. Also, \( r = 0.8 \) and \( n = 22 \).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midterm</td>
<td>73</td>
<td>10</td>
</tr>
<tr>
<td>Final</td>
<td>73</td>
<td>10</td>
</tr>
</tbody>
</table>

According to the regression equation, for a student who gets an 87 on the midterm, what is the predicted final exam grade? What phenomenon does this demonstrate? Explain.

The predicted final exam grade is _________.
(Round to the nearest integer as needed.)

What phenomenon does this demonstrate? Explain. Choose the correct answer below.

- A. Regression toward the mean, because the student's predicted final score is closer to the mean than was their midterm score.
- B. Regression toward the mean, because the student's predicted final score is farther from the mean than was their midterm score.
- C. Extrapolation, because the score of 87 is outside the range of the data.
- D. Extrapolation, because the predicted score is lower than the midterm score.

13. For what types of associations are regression models useful?

Choose the correct answer below.

- Linear
- Non-linear
- Both linear and non-linear
- For all types of associations

14. What type of effect can outliers have on a regression line?

Choose the correct answer below.

- A. No effect
- B. A small and insignificant effect
- C. Outliers are never included in a regression line.
- D. A big effect

15. When one has influential points in their data, how should regression and correlation be done?

Choose the correct answer below.

- A. Remove the influential points from your data set before doing regression and correlation
- B. Do regression and correlation with and without these points and comment on the differences
- C. Don't use regression or correlation on data sets containing influential points
- D. Always include the influential points in your data set when doing regression and correlation
16. Fill in the blank.

Since outliers can greatly affect the regression line they are also called _______ points.

Since outliers can greatly affect the regression line they are also called (1) ____________ points.

(1)  
- effective
- impact
- influential
- regression

17. Under what conditions can extrapolation be used to make predictions beyond the range of the data?

Choose the correct answer below.

- A. When there is a strong positive linear association in the data.
- B. When the correlation coefficient is close to −1 or +1.
- C. When the data set contains a large number of pairs of data.
- D. Never

18. Fill in the blank.

The value that measures how much variation in the response variable is explained by the explanatory variable is called the _______.

The value that measures how much variation in the response variable is explained by the explanatory variable is called the (1) ____________

(1)  
- regression coefficient.
- correlation coefficient.
- coefficient of causality.
- coefficient of determination.
19. The accompanying calculator screenshots show the scatterplot and the correlation coefficient between the number of days absent and the final grade for a sample of college students in a general education statistics course at a large community college. The relationship between "days absent" and "final grade" can be described as __________.

2. Click the icon to view the calculator screenshots.

- A. A strong positive linear relationship
- B. A moderate positive linear relationship
- C. A weak negative linear relationship
- D. A strong negative linear relationship

2: Calculator Screenshots

$$y = a + bx$$

$$8 
eq 0 \text{ and } p 
eq 0$$

$$b = -2.827272727$$

$$s = 3.067320418$$

$$r^2 = 0.8975572812$$

$$r = -0.9473949974$$

20. The following linear regression model can be used to predict ticket sales at a popular water park.

Ticket sales per hour = −631.25 + 11.25(current temperature in °F)

What is the predicted number of tickets sold per hour if the temperature is 86°F? Round to the nearest whole ticket.

- A. About 276 tickets
- B. About 301 tickets
- C. About 252 tickets
- D. About 336 tickets

21. Suppose it has been established that "Annual Income" and "Years of College" are linearly related, and that the relationship can be modeled using the equation below.

Annual Income = $23,400 + $7200(Years of College)

In this model, "Annual Income" is the ___ variable, and "Years of College" is the ___ variable. The two variables have a ___ linear relationship.

- A. Independent; Dependent; Positive
- B. Dependent; Independent; Positive
- C. Dependent; Independent; Negative
- D. Independent; Dependent; Negative
22. Suppose it has been established that "Home value" and "Years of College" are linearly related, and that the relationship can be modeled using the equation below.

\[ \text{Home value} = 75,000 + 12,500(\text{Years of College}) \]

In this model, "Years of College" is the \_\_\_ variable, and "Home value" is the \_\_\_ variable. The two variables have a \_\_\_ relationship.

\[ \text{A. Independent; Dependent; Positive} \]
\[ \text{B. Independent; Dependent; Negative} \]
\[ \text{C. Dependent; Independent; Positive} \]
\[ \text{D. Dependent; Independent; Negative} \]

23. The following linear regression model can be used to predict ticket sales at a popular water park.

\[ \text{Ticket sales per hour} = -631.25 + 11.25(\text{current temperature in } ^\circ\text{F}) \]

Choose the statement that best states the meaning of the slope in this context.

\[ \text{A. The slope says that a one degree increase in temperature is associated with an average increase in ticket sales of 11.25 tickets.} \]
\[ \text{B. The slope says that if ticket sales are decreasing there must have been a drop in temperature.} \]
\[ \text{C. The slope says that high temperatures are causing more people to buy tickets to the water park.} \]
\[ \text{D. None of the above} \]

24. A random sample of 30 married couples were asked to report the height of their spouse and the height of their biological parent of the same gender as their spouse. The output of a regression analysis for predicting spouse height from parent height is given. Assume that the conditions of the linear regression model are satisfied. What is the slope of the regression line? Choose the statement that is the correct interpretation of the slope in context.

\[ \text{Click the icon to view the regression analysis output.} \]

\[ \text{A. The slope is 48.40. On average, for each inch taller a parent is, the spouse is about 48.40 inches taller, in the sample.} \]
\[ \text{B. The slope is 48.40. On average, for each inch taller a parent is, the spouse is about 0.25 inches taller, in the sample.} \]
\[ \text{C. The slope is 0.25. On average, for each 0.25 inches taller a parent is, the spouse is about 1 inch taller, in the sample.} \]
\[ \text{D. The slope is 0.25. On average, for each inch taller a parent is, the spouse is about 0.25 inches taller, in the sample.} \]

3: Regression Analysis Output

<table>
<thead>
<tr>
<th>Regression Analysis: Spouse versus Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>The regression equation is</td>
</tr>
<tr>
<td>spouse = 48.40 + 0.25 parent</td>
</tr>
<tr>
<td>Predictor: Constant</td>
</tr>
<tr>
<td>Parameter Estimate: 48.398</td>
</tr>
<tr>
<td>Standard Error: 39.695</td>
</tr>
<tr>
<td>T-statistic: 1.219</td>
</tr>
<tr>
<td>p-value: 0.277</td>
</tr>
<tr>
<td>S = 7.794899045</td>
</tr>
<tr>
<td>R-sq = 0.036858791</td>
</tr>
<tr>
<td>r = 0.1919864344</td>
</tr>
</tbody>
</table>

| Predictor: Parent                        |
| Parameter Estimate: 0.247                |
| Standard Error: 0.566                    |
| T-statistic: 0.437                       |
| p-value: 0.680                           |
25. A random sample of 30 couples who were also new home owners were asked to report the cost of their first house and their combined age when they married. The output of a regression analysis for predicting home cost from combined age is given. Assume that the conditions of the linear regression model are satisfied. If the slope were 1, what would that say about the association?

4 Click the icon to view the regression analysis output.

- A. If the slope were 1, it would mean that on average, for every additional year in combined age, the home cost would be 1 dollar more.
- B. If the slope were 1, it would mean that on average, for every additional year in combined age, the home cost would be $2,122.75 higher.
- C. If the slope were 1, it would mean that on average, for every additional year in combined age, the home cost would be $2,122.75 lower.
- D. None of the above are true.

4: Regression Analysis Output

<table>
<thead>
<tr>
<th>Regression Analysis: Combined age versus home cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>The regression equation is</td>
</tr>
<tr>
<td>home cost = 73.74 + 2122.75 combined age</td>
</tr>
<tr>
<td>Predictor: Constant</td>
</tr>
<tr>
<td>Parameter Estimate: 73.74</td>
</tr>
<tr>
<td>Standard Error: 24.655</td>
</tr>
<tr>
<td>T-statistic: 0.003</td>
</tr>
<tr>
<td>p-value: 0.998</td>
</tr>
<tr>
<td>S=19837.70325 R-sq=0.8562890509 r=0.9253588768</td>
</tr>
</tbody>
</table>

26. A horticulturist conducted an experiment on 110 thirty-six inch plant boxes to see if the amount of plant food given to the plant boxes was associated with the number of tomatoes harvested from the plants. The average amount of plant food given was 27.8 milliliters with a standard deviation of 2.1 milliliters. The average number of tomatoes harvested was 7.5 with a standard deviation of 1.5. The correlation coefficient was 0.7691. Use the information to calculate the slope of the linear model that predicts the number of tomatoes harvested from the amount of plant food given. Round to the nearest hundredth.

- A. 1.08
- B. -7.50
- C. 0.55
- D. The slope cannot be determined without the actual data.

27. In the NHL, the correlation between "Goals scored per game" and "Minutes on the ice" for a team of players is found to be 0.8178. Choose the statement that is true about the coefficient of determination.

- A. The coefficient of determination, $r^2$, is equal to approximately 0.6688.
- B. When given as a percent, the coefficient of determination is always between 0 and 100%.
- C. The coefficient of determination states that about 66.88% of the variation in goals scored per game is explained by minutes on the ice.
- D. All of these are true statements.
28. In the NBA, the correlation between "steals per game" and "blocked shots per game" is found to be 0.8045. Choose the statement that is true about the coefficient of determination.

- A. The coefficient of determination states that about 64.72% of the variation in blocked shots per game is explained by steals per game.
- B. The coefficient of determination, $r^2$, is equal to approximately 0.6472.
- C. When given as a percent, the coefficient of determination is always between 0 and 100%.
- D. All of these are true statements.

29. It is determined that a positive linear association exists between age (for children between the ages of 3 and 9 years) and attention span (measured in minutes). The accompanying scatterplot shows the association. The prediction equation is also given. A college instructor uses the model to predict the attention span of the students in her class who have an average age of 29. Choose the best statement to summarize why this is not an appropriate use for the model.

- A. This is an inappropriate use of the model because a 29-year-old person would be an outlier in this context.
- B. This is an inappropriate use of the model because age does not cause attention span to increase. Correlation does not mean causation.
- C. This is an inappropriate use of the model because the model was used to make predictions beyond the scope of the data. The college instructor is extrapolating.

5: Age and Attention Span

\[
\text{attention span} = 4.68 + 3.40(\text{age})
\]
The data in the accompanying table represent the amount of pressure (psi) exerted by a stamping machine \( (x) \), and the amount of scrap brass shavings (in pounds) that are collected from the machine each hour \( (y) \). Also given are the outputs from two different statistical technologies. A scatterplot of the data confirms that there is a linear association. Report the equation for predicting scrap brass shavings using words such as scrap, not \( x \) and \( y \). State the slope and intercept of the prediction equation. Round all calculations to the nearest thousandth.

6 Click the icon to view the data table and the technology outputs.

- **A.** scrap = \(-2.019 + 2.134\) (pressure); slope = \(-2.019\) and the intercept is \(2.134\).
- **B.** scrap = \(2.134 - 2.019\) (pressure); slope = \(2.134\) and the intercept is \(-2.019\).
- **C.** scrap = \(2.134 - 2.019\) (pressure); slope = \(-2.019\) and the intercept is \(2.134\).
- **D.** scrap = \(-2.019 + 2.134\) (pressure); slope = \(2.134\) and the intercept is \(-2.019\).

6: Data Table and Technology Outputs

<table>
<thead>
<tr>
<th>( x )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>2.30</td>
</tr>
<tr>
<td>7.80</td>
<td>15.14</td>
</tr>
<tr>
<td>14.51</td>
<td>28.65</td>
</tr>
<tr>
<td>2.80</td>
<td>4.15</td>
</tr>
<tr>
<td>4.01</td>
<td>6.35</td>
</tr>
<tr>
<td>6.21</td>
<td>10.52</td>
</tr>
<tr>
<td>11.84</td>
<td>24.05</td>
</tr>
<tr>
<td>5.11</td>
<td>8.75</td>
</tr>
<tr>
<td>11.67</td>
<td>22.22</td>
</tr>
<tr>
<td>8.70</td>
<td>17.02</td>
</tr>
</tbody>
</table>

T-Test

\[ y = a + bx \]

\[ a = -2.018775528 \]
\[ b = 2.134464237 \]
\[ r^2 = 0.96074796 \]
\[ r = 0.998035467 \]

<table>
<thead>
<tr>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>XVariable 1</td>
</tr>
</tbody>
</table>
1. (1) − 0.033
   (2) 0.628
   (3) − 0.914

2. A. Predicted Armspan = 29.62 + 2.04 Height
   0.908
   8.24
   3.67
   2.04
   159.59
   63.77
   29.50
   152.0

3. (1) Armspan
   11.657
   2.417
   (2) Height

D. The slope is the number being multiplied by Height, and the intercept is the constant.
D. In the fourth line, the slope is multiplied by the Height, and the intercept is the constant.
C. The slope is the X Variable value, and the intercept is the Intercept value.
A. The a-value is the intercept, and the b-value is the slope.

4. C. The horizontal axis starts at 0. To zoom in on the data, the horizontal axis should have been started at about 60.

5. C. The association between height and arm span for women is stronger because the value of r is farther from 0.

6. C.
   The higher the percentage of smoke-free homes in a region, the lower the percentage of high school students who smoke tends to be.
   23

7. C.
   The graph shows that young drivers and old drivers have more fatalities and that the safest drivers are between about 40 and 60 years of age.
   A. It would not be appropriate for linear regression because the trend is not linear.
8. A. An influential point is a point that changes the regression equation by a large amount.
   
   A.
   Going to church may not cause lower blood pressure. Just because two variables are related does not show that one caused the other.

9. A.
   The coefficient of determination is 54.76%. Therefore, 54.76% of the variation in weight can be explained by the regression line.

10. C.
    The average salary is $2099 less for each year later that the person was hired or an average of $2099 more for each year earlier.

    C.
    The y-intercept of $4,255,424 would be the salary for a person who started in the year 0, which is not appropriate to interpret.

11. C. The correlation is positive because the graph shows an increasing trend.
    
    D. For each additional dollar spent on teachers' pay, the expenditure per pupil, on average, is about 23 cents higher.

    C.
    The intercept might represent the cost of education when the teachers are paid nothing. However, doing so requires extrapolation which may not be valid.

12. 84
    
    A.
    Regression toward the mean, because the student's predicted final score is closer to the mean than was their midterm score.

13. Linear

14. D. A big effect

15. B. Do regression and correlation with and without these points and comment on the differences

16. (1) influential

17. D. Never

18. (1) coefficient of determination.

19. D. A strong negative linear relationship
20. D. About 336 tickets

21. B. Dependent; Independent; Positive

22. A. Independent; Dependent; Positive

23. A.
   The slope says that a one degree increase in temperature is associated with an average increase in ticket sales of 11.25 tickets.

24. D. The slope is 0.25. On average, for each inch taller a parent is, the spouse is about 0.25 inches taller, in the sample.

25. A.
   If the slope were 1, it would mean that on average, for every additional year in combined age, the home cost would be 1 dollar more.

26. C. 0.55

27. D. All of these are true statements.

28. D. All of these are true statements.

29. C.
   This is an inappropriate use of the model because the model was used to make predictions beyond the scope of the data. The college instructor is extrapolating.

30. D. scrap = -2.019 + 2.134(pressure); slope = 2.134 and the intercept is -2.019.