4.3.2 Summarize the strength of a linear relationship with a correlation, r.
1. In discussing how its customers use online services, a bank manager noted “there seems to be a strong correlation between the use of the online bill paying feature and gender.” Comment on this statement.

4.2.1 Make a scatterplot to display the relationship between two quantitative variables.
2. For each of the following scenarios indicate which is the predictor variable and which is the response variable.
a. A study examined consumption levels of oil and carbon dioxide emissions for a sample of counties.
b. Data were collected on job performance rating and hours of training for a sample of employees at a telecommunications repair facility.
c. Salary data as well as years of managerial experience were collected for a sample of executives in the high tech industry.

4.3.2 Summarize the strength of a linear relationship with a correlation, r.
3. The following scatterplot shows monthly sales figures (in units) and number of months of experience on the job for a sample of 19 salespeople.

![Scatterplot of Sales vs Experience](image)

a. Describe the association between monthly sales and level of experience.
b. Do these data satisfy the conditions for computing a correlation coefficient? Explain.
c. Estimate the correlation.
4.3.2 Summarize the strength of a linear relationship with a correlation, r.

4. Shown below is a correlation table showing correlation coefficients between stock price, earnings per share (EPS), and price/earnings (P/E) ratio for a sample of 19 publicly traded companies.

**Correlations: Stock Price, EPS, PE**

<table>
<thead>
<tr>
<th></th>
<th>Stock Price</th>
<th>EPS</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS</td>
<td>0.875</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE</td>
<td>0.323</td>
<td>-0.111</td>
<td></td>
</tr>
</tbody>
</table>

a. What is the correlation between stock price and EPS? Interpret.
b. What is the correlation between stock price and PE? Interpret.
c. What is the correlation between EPS and PE? Interpret.

*Use the following to answer questions 5 – 8.*

To determine whether the cash bonus paid by a company is related to annual pay, data were gathered for 10 loan officers at AMC Bank who received cash bonuses in 2012. The data, scatterplot, and summary statistics are shown below.

<table>
<thead>
<tr>
<th></th>
<th><strong>ANNUAL PAY</strong></th>
<th><strong>CASH BONUS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$70609</td>
<td>$11225</td>
<td></td>
</tr>
<tr>
<td>$58487</td>
<td>$6238</td>
<td></td>
</tr>
<tr>
<td>$104561</td>
<td>$14194</td>
<td></td>
</tr>
<tr>
<td>$43922</td>
<td>$4188</td>
<td></td>
</tr>
<tr>
<td>$82613</td>
<td>$11863</td>
<td></td>
</tr>
<tr>
<td>$116250</td>
<td>$13671</td>
<td></td>
</tr>
<tr>
<td>$76751</td>
<td>$7758</td>
<td></td>
</tr>
<tr>
<td>$68513</td>
<td>$20760</td>
<td></td>
</tr>
<tr>
<td>$137000</td>
<td>$55000</td>
<td></td>
</tr>
<tr>
<td>$94469</td>
<td>$34368</td>
<td></td>
</tr>
</tbody>
</table>

**Mean**

<table>
<thead>
<tr>
<th></th>
<th><strong>ANNUAL PAY</strong></th>
<th><strong>CASH BONUS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$85318</td>
<td>$17927</td>
<td></td>
</tr>
</tbody>
</table>

**Standard Deviation**

<table>
<thead>
<tr>
<th></th>
<th><strong>ANNUAL PAY</strong></th>
<th><strong>CASH BONUS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$28077</td>
<td>$15618</td>
<td></td>
</tr>
</tbody>
</table>

**Correlation**

<table>
<thead>
<tr>
<th></th>
<th>0.735</th>
</tr>
</thead>
</table>
4.3.3 Model a linear relationship with a least squares regression model.
5. Comment on whether each of the following conditions for correlation / linear regression is met.
   a. Quantitative variable condition.
   b. Linearity condition.
   c. Outlier condition.

4.3.3 Model a linear relationship with a least squares regression model.
6. Estimate the linear regression model that relates the response variable (cash bonus) to the predictor variable (annual pay).
   a. Find the slope of the regression line.
   b. Find the intercept of the regression line.
   c. Write the equation of the linear model.

4.3.3 Model a linear relationship with a least squares regression model.
7. Find the value of $R^2$. Interpret its meaning in this context.

4.5.5 Examine the residuals from a linear model to assess the quality of the model.
8. Using the estimated regression equation,
   a. Estimate the cash bonus for an executive at Johnson Financial earning $82,613 a year.
   b. What is the residual for this estimate? What does it mean?
4.7.4 Recognize regression to the mean when it occurs in data.
9. A business executive taught his managers that praise is more effective than punishment. He was challenged by one of the managers who said that in his experience praising an employee for executing a task is typically followed by a lesser performance, whereas yelling at an employee for bad execution is typically followed by improved performance.

a. Explain this case in terms of a failure to appreciate regression toward the mean.
b. How might luck factor into the improvement or deterioration of performance?
Chapter 4: Correlation and Linear Regression – Quiz A – Key

1. In discussing how its customers use online services, a bank manager noted “there seems to be a strong correlation between the use of the online bill paying feature and gender.” Comment on this statement.

There may be an association between the use of online bill paying and gender, but these variables are both categorical so they cannot be correlated.

2. For each of the following scenarios indicate which is the predictor variable and which is the response variable.

a. A study examined consumption levels of oil and carbon dioxide emissions for a sample of counties.

Oil consumption is the predictor variable, and carbon dioxide emissions is the response variable.

b. Data were collected on job performance rating and hours of training for a sample of employees at a telecommunications repair facility.

Job performance is the response variable and hours of training is the predictor variable.

c. Salary data as well as years of managerial experience were collected for a sample of executives in the high tech industry.

Salary is the response variable, and years of experience is the predictor variable.
3. The following scatterplot shows monthly sales figures (in units) and number of months of experience on the job for a sample of 19 salespeople.

```
Scatterplot of Sales vs Experience
```

a. Describe the association between monthly sales and level of experience.

Positive and strong.

b. Do these data satisfy the conditions for computing a correlation coefficient? Explain.

Yes: variables are quantitative, the relationship is straight enough, and there are no apparent outliers.

c. Estimate the correlation.

.914 is the correlation; a value between .85 and .95 is acceptable.
4. Shown below is a correlation table showing correlation coefficients between stock price, earnings per share (EPS), and price/earnings (P/E) ratio for a sample of 19 publicly traded companies.

**Correlations: Stock Price, EPS, PE**

<table>
<thead>
<tr>
<th></th>
<th>Stock Price</th>
<th>EPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS</td>
<td>0.875</td>
<td></td>
</tr>
<tr>
<td>PE</td>
<td>0.323</td>
<td>-0.111</td>
</tr>
</tbody>
</table>

a. What is the correlation between stock price and EPS? Interpret.

.875, positive and strong.

b. What is the correlation between stock price and PE? Interpret.

.323, positive and weak.

c. What is the correlation between EPS and PE? Interpret.

-.111, negative and very weak.

*Use the following to answer questions 5 – 8.*

To determine whether the cash bonus paid by a company is related to annual pay, data were gathered for 10 account executives at Johnson Financial Group who received cash bonuses in 2007. The data, scatterplot, and summary statistics are shown below.

<table>
<thead>
<tr>
<th></th>
<th><strong>ANNUAL PAY</strong></th>
<th><strong>CASH BONUS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$ 70609</td>
<td>$ 11225</td>
<td></td>
</tr>
<tr>
<td>$ 58487</td>
<td>$ 6238</td>
<td></td>
</tr>
<tr>
<td>$ 104561</td>
<td>$ 14194</td>
<td></td>
</tr>
<tr>
<td>$ 43922</td>
<td>$ 4188</td>
<td></td>
</tr>
<tr>
<td>$ 82613</td>
<td>$ 11863</td>
<td></td>
</tr>
<tr>
<td>$ 116250</td>
<td>$ 13671</td>
<td></td>
</tr>
<tr>
<td>$ 76751</td>
<td>$ 7758</td>
<td></td>
</tr>
<tr>
<td>$ 68513</td>
<td>$ 20760</td>
<td></td>
</tr>
<tr>
<td>$ 137000</td>
<td>$ 55000</td>
<td></td>
</tr>
<tr>
<td>$ 94469</td>
<td>$ 34368</td>
<td></td>
</tr>
</tbody>
</table>

*Mean*  
$85318  
$17927

*Standard Deviation*  
$28077  
$15618

**Correlation**  
0.735
5. Comment on whether each of the following conditions for correlation / linear regression is met.

   a. Quantitative variable condition.
      Yes, both variables are quantitative.

   b. Linearity condition.
      Yes, appears straight enough.

   c. Outlier condition.
      Yes, no obvious outliers.

6. Estimate the linear regression model that relates the response variable (cash bonus) to the predictor variable (annual pay).

   a. Find the slope of the regression line.
      .409

   b. Find the intercept of the regression line.
      -16968

   c. Write the equation of the linear model.
      Cash Bonus = -16968 + .409 (Pay)
7. Find the value of $R^2$. Interpret its meaning in this context.

.54, which means that 54% of the variability in cash bonuses can be explained by pay.

8. Using the estimated regression equation,

a. Estimate the cash bonus for an executive at Johnson Financial earning $82,613 a year.
   $16,821.

b. What is the residual for this estimate?
   -4958

What does it mean?
The model over predicted the actual value.

9. A business executive taught his managers that praise is more effective than punishment. He was challenged by one of the managers who said that in his experience praising an employee for executing a task is typically followed by a lesser performance, whereas yelling at an employee for bad execution is typically followed by improved performance.

a. Explain this case in terms of a failure to appreciate regression toward the mean.

A manager’s performance, although based on considerable skill, will vary randomly from task to task.

b. How might luck factor into the improvement or deterioration of performance?

When a manager executes an excellent task, it is likely that he or she had a bit of luck in their favor in addition to their considerable skill. After the praise but not because of it, the luck component will probably disappear and the performance will be lower. Similarly, a poor performance is likely to be partly due to bad luck. After the criticism but not because of it, the next performance will likely be better.
4.3.2 Summarize the strength of a linear relationship with a correlation, r.
1. In commenting on the increase in home foreclosures, a news reporter stated “there appears to be a strong correlation between home foreclosures and job loss of the head of household.” Comment on this statement.

4.2.1 Make a scatterplot to display the relationship between two quantitative variables.
2. For each of the following scenarios indicate which is the predictor variable and which is the response variable.

a. A supermarket chain gathers data on the amount they spend on promotional material (specials, coupons, etc.) and sales revenue generated each quarter.
b. Government-sponsored research investigated the relationship between number of hours individuals spend on the Internet and age.
c. A real estate association conducted a study on home prices and economic strength for different regions of the United States.

4.3.2 Summarize the strength of a linear relationship with a correlation, r.
3. Shown below is a correlation table showing correlation coefficients between population (millions), migration rate of mobile subscriptions to smart phones (in %) and smart phone penetration per capita (in %) for a sample of 15 countries.

<table>
<thead>
<tr>
<th>Correlations: Population, Migration rate of mobile subscriptions, Smartphone penetration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migration rate</td>
</tr>
<tr>
<td>Smartphone penetration</td>
</tr>
</tbody>
</table>

a. What is the correlation between migration rate of mobile subscriptions and population? Interpret.
b. What is the correlation between smart phone penetration and population? Interpret.
c. What is the correlation between migration rate of mobile subscriptions and smart phone penetration? Interpret.
4.3.2 Summarize the strength of a linear relationship with a correlation, $r$.

4. A consumer research group investigating the relationship between the price of meat (per pound) and the fat content (grams) gathered data that produced the following scatterplot.

![Scatterplot of Fat Grams vs Price/lb](chart.png)

a. Describe the association between the price of meat and fat content.
b. Estimate the correlation.
c. If the point in the lower left hand corner ($2.00 per pound, 6 grams of fat) is removed, would the correlation become stronger or weaker or remain the same? Explain.
Use the following to answer questions 5 – 8.

A small independent organic food store offers a variety of specialty coffees. To
determine whether price has an impact on sales, the managers kept track of how many
pounds of each variety of coffee were sold last month. The data, scatterplot, and
summary statistics are shown below.

<table>
<thead>
<tr>
<th>PRICE PER POUND</th>
<th>POUNDS SOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ 3.99</td>
<td>75</td>
</tr>
<tr>
<td>$ 5.99</td>
<td>60</td>
</tr>
<tr>
<td>$ 7.00</td>
<td>65</td>
</tr>
<tr>
<td>$ 12.00</td>
<td>45</td>
</tr>
<tr>
<td>$ 4.50</td>
<td>80</td>
</tr>
<tr>
<td>$ 7.50</td>
<td>70</td>
</tr>
<tr>
<td>$ 15.00</td>
<td>25</td>
</tr>
<tr>
<td>$ 10.00</td>
<td>35</td>
</tr>
<tr>
<td>$ 12.50</td>
<td>40</td>
</tr>
<tr>
<td>$ 8.99</td>
<td>50</td>
</tr>
</tbody>
</table>

Mean $ 8.75 54.50
Standard Deviation $ 3.63 18.33
Correlation -0.927

Scatterplot of Pounds Sold vs Price/Pound
4.3.3 Model a linear relationship with a least squares regression model.
5. Comment on whether each of the following conditions for correlation / linear regression is met.

a. Quantitative variable condition.
b. Linearity condition.
c. Outlier condition.

4.3.3 Model a linear relationship with a least squares regression model.
6. Estimate the linear regression model that relates the response variable (monthly sales) to the predictor variable (price per pound).

a. Find the slope of the regression line.
b. Find the intercept of the regression line.
c. Write the equation of the linear model.

4.3.3 Model a linear relationship with a least squares regression model.
7. Find the value of $R^2$. Interpret its meaning in this context.

4.5.5 Examine the residuals from a linear model to assess the quality of the model.
8. Using the estimated regression equation,

a. Estimate the monthly sales for a variety of coffee that costs $12 per pound.
b. What is the residual for this estimate? What does it mean?

4.7.4 Recognize regression to the mean when it occurs in data.
9. Some people believe that there is a Sports Illustrated cover jinx; that is, individuals or teams who appear on the cover of Sports Illustrated will subsequently experience bad luck.

a. Explain this case in terms of a failure to appreciate regression toward the mean.
b. Give an explanation for the deterioration in performance after the featured cover?
Chapter 4: Correlation and Linear Regression – Quiz B – Key

1. In commenting on the increase in home foreclosures, a news reporter stated “there appears to be a strong correlation between home foreclosures and job loss of the head of household.” Comment on this statement.

There may be an association between home foreclosures and the head of household losing his/her job, but these variables are both categorical so they cannot be correlated.

2. For each of the following scenarios indicate which is the predictor variable and which is the response variable.

a. A supermarket chain gathers data on the amount they spend on promotional material (specials, coupons, etc.) and sales revenue generated each quarter.

Amount spent on promotional material is the predictor variable, sales revenue the response variable.

b. Government-sponsored research investigated the relationship between number of hours individuals spend on the Internet and age.

Number of hours spent on the Internet is the response variable and age the predictor variable.

c. A real estate association conducted a study on home prices and economic strength for different regions of the United States.

Home price is the response variable and economic strength is the predictor variable.

3. Shown below is a correlation table showing correlation coefficients between population (millions), migration rate of mobile subscriptions to smart phones (in %) and smart phone penetration per capita (in %) for a sample of 15 countries.

| Correlations: Population, Migration rate of mobile subscriptions, Smartphone penetration |
|---------------------------------|-----------|-----------|
| Migration rate                  | Population| Migration rate |
| Migration rate                  | -0.266    | 0.885     |
| Smartphone penetration          | -0.419    |            |

a. What is the correlation between migration rate of mobile subscriptions and population? Interpret.

-0.266, negative and weak.

b. What is the correlation between smart phone penetration and population? Interpret.

-0.419, negative and moderate.

c. What is the correlation between migration rate of mobile subscriptions and smart phone penetration? Interpret.

0.885, positive and strong.
4. A consumer research group investigating the relationship between the price of meat (per pound) and the fat content (grams) gathered data that produced the following scatterplot.

![Scatterplot of Fat Grams vs Price/lb](image)

a. Describe the association between the price of meat and fat content.
Negative, moderately strong.

b. Estimate the correlation.
-.719, estimate between -.65 and -.75 acceptable.

c. If the point in the lower left hand corner ($2.00 per pound, 6 grams of fat) is removed, would the correlation become stronger or weaker or remain the same? Explain.
It would become stronger (the actual correlation would be -.867). Because this point is not quite in line with the rest of the data, removing it would strengthen the linear association and make the correlation stronger.
Use the following to answer questions 5 – 8.

A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month. The data, scatterplot and summary statistics are shown below.

<table>
<thead>
<tr>
<th>PRICE PER POUND</th>
<th>POUNDS SOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ 3.99</td>
<td>75</td>
</tr>
<tr>
<td>$ 5.99</td>
<td>60</td>
</tr>
<tr>
<td>$ 7.00</td>
<td>65</td>
</tr>
<tr>
<td>$ 12.00</td>
<td>45</td>
</tr>
<tr>
<td>$ 4.50</td>
<td>80</td>
</tr>
<tr>
<td>$ 7.50</td>
<td>70</td>
</tr>
<tr>
<td>$ 15.00</td>
<td>25</td>
</tr>
<tr>
<td>$ 10.00</td>
<td>35</td>
</tr>
<tr>
<td>$ 12.50</td>
<td>40</td>
</tr>
<tr>
<td>$ 8.99</td>
<td>50</td>
</tr>
</tbody>
</table>

Mean $ 8.75 54.50
Standard Deviation $ 3.63 18.33
Correlation -0.927
5. Comment on whether each of the following conditions for correlation / linear regression is met.

a. Quantitative variable condition.
Yes, both variables are quantitative.

b. Linearity condition.
Yes, appears straight enough.

c. Outlier condition.
Yes, no obvious outliers.

6. Estimate the linear regression model that relates the response variable (monthly sales) to the predictor variable (price per pound).

a. Find the slope of the regression line.
-4.681

b. Find the intercept of the regression line.
95.459

c. Write the equation of the linear model.
Pounds Sold = 95.459 - 4.681 Price per pound

7. Find the value of $R^2$. Interpret its meaning in this context.

.859, which means that 85.9% of the variability in the number of pounds of coffee sold per week is explained by price.

8. Using the estimated regression equation,

a. Estimate the monthly sales for a variety of coffee that costs $12 per pound.
39.29 pounds

b. What is the residual for this estimate?
5.71 pounds

What does it mean?
The model under predicted the actual value.
9. Some people believe that there is a Sports Illustrated cover jinx; that is, individuals or teams who appear on the cover of Sports Illustrated will subsequently experience bad luck.

a. Explain this case in terms of a failure to appreciate regression toward the mean.

Regression toward the mean plays a role in the so-called "Sophomore Slump," a good example of which is that a player who wins "rookie of the year" typically does less well in his second season. It is nearly impossible for most athletes to keep up the level of performance that gave them any distinguished honor.

b. Give an explanation for the deterioration in performance after the featured cover?

The most common explanation for the perceived effect is that athletes are generally featured on the cover after an outlier performance; their future performance is likely to be less impressive by comparison. This decline in performance would then be misperceived as being related to, or even possibly caused by, the appearance on the magazine cover.
Chapter 4: Correlation and Linear Regression – Quiz C – Multiple Choice

Name_________________________

4.2.1 Make a scatterplot to display the relationship between two quantitative variables.
1. A study examined consumption levels of oil and carbon dioxide emissions for a sample of counties. The response variable in this study is

A. oil.  
B. oil consumption.  
C. carbon dioxide emissions.  
D. countries.  
E. none of the above.

4.2.1 Make a scatterplot to display the relationship between two quantitative variables.
2. A supermarket chain gathers data on the amount they spend on promotional material (e.g., coupons, etc.) and sales revenue generated each quarter. The predictor variable is

A. sales revenue.  
B. amount spent on promotional material.  
C. number of coupons offered.  
D. supermarket chains.  
E. none of the above.

4.3.2 Summarize the strength of a linear relationship with a correlation, $r$.
3. The scatterplot shows monthly sales figures (in units) and number of months of experience for a sample of salespeople.

![Scatterplot of Sales vs Experience](image-url)
The association between monthly sales and level of experience can be described as

A. positive and weak.
B. negative and weak.
C. negative and strong.
D. positive and strong.
E. nonlinear.

4.3.2 Summarize the strength of a linear relationship with a correlation, $r$.
4. The scatterplot shows monthly sales figures (in units) and number of months of experience for a sample of salespeople.

The correlation between monthly sales and level of experience is most likely

A. -.235.
B. 0.
C. .180.
D. -.914.
E. .914.
4.3.2 Summarize the strength of a linear relationship with a correlation, r.

5. Shown below is a correlation table showing correlation coefficients between stock price, earnings per share (EPS) and price/earnings (P/E) ratio for a sample of 19 publicly traded companies. Which of the following statements is false?

**Correlations: Stock Price, EPS, PE**

<table>
<thead>
<tr>
<th></th>
<th>Stock Price</th>
<th>EPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS</td>
<td>0.875</td>
<td></td>
</tr>
<tr>
<td>PE</td>
<td>0.323</td>
<td>-0.111</td>
</tr>
</tbody>
</table>

A. EPS is the best predictor of stock price.
B. The strongest correlation is between EPS and stock price.
C. There is a weak negative association between PE and EPS.
D. PE is the best predictor of stock price.
E. The weakest correlation is between PE and EPS.

4.2.1 Make a scatterplot to display the relationship between two quantitative variables.

6. A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month. Based on the scatterplot shown below, which of the following statements is true?

A. The quantitative variable condition is satisfied.
B. The linearity condition is satisfied.
C. There are no obvious outliers.
D. All of the above.
E. None of the above.
4.3.2 Summarize the strength of a linear relationship with a correlation, r.

7. A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month.

Based on the scatterplot, the linear relationship between number of pounds of coffee sold per week and price is

A. strong and positive.
B. strong and negative.
C. weak and negative.
D. weak and positive.
E. nonexistent.
4.3.3 Model a linear relationship with a least squares regression model.

8. A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month.

<table>
<thead>
<tr>
<th>PRICE PER POUND</th>
<th>POUNDS SOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3.99</td>
<td>75</td>
</tr>
<tr>
<td>$5.99</td>
<td>60</td>
</tr>
<tr>
<td>$7.00</td>
<td>65</td>
</tr>
<tr>
<td>$12.00</td>
<td>45</td>
</tr>
<tr>
<td>$4.50</td>
<td>80</td>
</tr>
<tr>
<td>$7.50</td>
<td>70</td>
</tr>
<tr>
<td>$15.00</td>
<td>25</td>
</tr>
<tr>
<td>$10.00</td>
<td>35</td>
</tr>
<tr>
<td>$12.50</td>
<td>40</td>
</tr>
<tr>
<td>$8.99</td>
<td>50</td>
</tr>
</tbody>
</table>

Mean $8.75 54.50
Standard Deviation $3.63 18.33
Correlation -0.927

Based on the data and summary statistics shown below, the slope of the estimated regression line that relates the response variable (monthly sales) to the predictor variable (price per pound) is

A. 95.459.
B. .858.
C. -4.681.
D. -.858.
E. -8.999.
4.3.3 Model a linear relationship with a least squares regression model.
9. A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month.

<table>
<thead>
<tr>
<th>PRICE PER POUND</th>
<th>POUNDS SOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3.99</td>
<td>75</td>
</tr>
<tr>
<td>$5.99</td>
<td>60</td>
</tr>
<tr>
<td>$7.00</td>
<td>65</td>
</tr>
<tr>
<td>$12.00</td>
<td>45</td>
</tr>
<tr>
<td>$4.50</td>
<td>80</td>
</tr>
<tr>
<td>$7.50</td>
<td>70</td>
</tr>
<tr>
<td>$15.00</td>
<td>25</td>
</tr>
<tr>
<td>$10.00</td>
<td>35</td>
</tr>
<tr>
<td>$12.50</td>
<td>40</td>
</tr>
<tr>
<td>$8.99</td>
<td>50</td>
</tr>
</tbody>
</table>

Mean $8.75  54.50
Standard Deviation $3.63  18.33
Correlation -0.927

Based on the data and summary statistics, the intercept of the estimated regression line that relates the response variable (monthly sales) to the predictor variable (price per pound) is

A. 95.459.
B. .858.
C. -4.684.
D. -.858.
E. -8.999.

4.3.3 Model a linear relationship with a least squares regression model.
10. A small independent organic food store offers a variety of specialty coffees. To determine whether price has an impact on sales, the managers kept track of how many pounds of each variety of coffee were sold last month.

<table>
<thead>
<tr>
<th>Mean</th>
<th>$8.75</th>
<th>54.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>$3.63</td>
<td>18.33</td>
</tr>
<tr>
<td>Correlation</td>
<td>-0.927</td>
<td></td>
</tr>
</tbody>
</table>

Based on the summary statistics shown below, what percent of the variability in the number of pounds of coffee sold per week can be explained by price?

A. 95.47%
B. 100%
C. 85.9%
D. 55.6%
E. 4.68%
4.6.11 Analyze and transform data in a nonlinear relationship.
11. A scatterplot shows a curved relationship between 2 variables. What methods can be used to analyze the relationship between these variables?

A. Fit an exponential or quadratic function
B. Summarize the strength of the association by using correlation and $R^2$.
C. Transform one or both of the variables by using square root, logarithm, or reciprocal functions
D. A and C
E. All of the above
Chapter 4: Correlation and Linear Regression - Quiz C – Key

1. C
2. B
3. D
4. E
5. D
6. D
7. B
8. C
9. A
10. C
11. D
Chapter 4: Correlation and Linear Regression – Quiz D – Multiple Choice

Name_________________________

4.1.1 Make a scatterplot to display the relationship between two quantitative variables.
1. A consumer research group examining the relationship between the price of meat (per pound) and fat content (in grams) gathered data that produced the following scatterplot.

If the point in the lower left hand corner (2 grams of fat; $3.00 per pound) is removed, the correlation would most likely

A. remain the same.
B. become positive.
C. become weaker negative.
D. become stronger negative.
E. become zero.

4.3.6 Conditions for correlation; define and interpret linear regression; make predictions.
2. Data were collected on monthly sales revenues (in $1,000s) and monthly advertising expenditures ($100s) for a sample of drug stores. The regression line relating revenues (Y) to advertising expenditure (X) is estimated to be \( \hat{y} = -48.3 + 9.00x \). The correct interpretation of the slope is that for each additional

A. $1 spent on advertising, predicted sales revenue increases by $9,000.
B. $100 spent on advertising, predicted sales revenue increases by $9,000.
C. $100 spent on advertising, predicted sales revenue decreases by $9,000.
D. $1,000 in sales revenue, advertising expenditures decrease by $48.30.
E. $100 in sales revenue, advertising expenditures decrease by $48.30.
4.3.3 Model a linear relationship with a least squares regression model.

3. Data were collected on monthly sales revenues (in $1,000s) and monthly advertising expenditures ($100s) for a sample of drug stores. The regression line relating revenues (Y) to advertising expenditure (X) is estimated to be $\hat{y} = -48.3 + 9.00x$. The predicted sales revenue for a month in which $\$1,000$ was spent on advertising is

A. $50,000.
B. $851.70.
C. $8,951.70.
D. $41,700.
E. $90,000.

4.3.3 Examine the residuals from a linear model to assess the quality of the model.

4. A company studying the productivity of its employees on a new information system was interested in determining if the age (X) of data entry operators influenced the number of completed entries made per hour (Y). The regression equation is $\hat{y} = 14.374 - 0.145x$. Suppose the actual completed entries per hour for an operator who is 35 years old was 8. The residual is

A. -1.3
B. 2.6
C. -3.5
D. 1.3
E. -2.2

4.3.6 Summarize the strength of a linear relationship with a correlation, r.

5. A company studying the productivity of their employees on a new information system was interested in determining if the age (X) of data entry operators influenced the number of completed entries made per hour (Y). The regression equation is $\hat{y} = 14.374 - 0.145x$. If $s_x = 14.04$ and $s_y = 2.61$, then the correlation coefficient between age and productivity is

A. .779
B. -.236
C. .575
D. -.929
E. -.779

4.3.6 Summarize the strength of a linear relationship with a correlation, r.

6. Suppose the correlation, r, between two variables x and y is -0.44. What would you predict about a y value if the x value is 2 standard deviations above its mean?

A. It will be .88 standard deviations below its mean.
B. It will be .88 standard deviations above its mean.
C. It will be 2 standard deviations below its mean.
D. It will be .44 standard deviations below its mean.
E. It will be .44 standard deviations above its mean.
4.3.9 Model a linear relationship with a least squares regression model.
7. Suppose the correlation, \( r \), between two variables \( x \) and \( y \) is -0.44. What percentage of the variability in \( y \) cannot be explained by \( x \)?

A. 19%
B. 44%
C. 81%
D. 88%
E. 12%

4.8.5 Examine the residuals from a linear model to assess the quality of the model.
8. Linear regression was used to describe the trend in world population over time. Below is a plot of the residuals versus predicted values. What does the plot of residuals suggest?

A. An outlier is present in the data set.
B. The linearity condition is not satisfied.
C. A high leverage point is present in the data set.
D. The data are not normal.
E. The equal spread condition is not satisfied.
4.8.5 Examine the residuals from a linear model to assess the quality of the model.
9. Based on the following residual plot, which condition / assumption for linear regression is not satisfied?

A. Linearity.
B. Quantitative Variables.
C. Equal Spread.
D. Outlier.
E. None of the above; all conditions are satisfied.
4.3.2 Summarize the strength of a linear relationship with a correlation, r.

10. For the following scatterplot,

![Scatterplot of y vs x](image)

The likely correlation coefficient is

A. +0.35  
B. +0.90  
C. +0.77  
D. -0.89  
E. -1.00

4.11.6 Analyze and transform data in a nonlinear relationship.

11. The disadvantage of re-expressing variables is that

A. We have to explain the association in terms of the transformed variables.  
B. We cannot use the standard regression models  
C. It is not commonly used  
D. It can be difficult  
E. A and D
Chapter 4: Correlation and Linear Regression - Quiz D – Key

1. D
2. B
3. D
4. A
5. E
6. A
7. C
8. B
9. C
10. A
11. E