APPENDIX C: ANSWERS TO ODD-NUMBERED CHAPTER EXERCISES & SOLUTIONS TO PRACTICE TESTS

Answers to Odd-Numbered Chapter Exercises

CHAPTER 1

1. **a.** Interval **b.** Ratio **c.** Nominal **d.** Nominal **e.** Ordinal **f.** Ratio

3. Answers will vary.

5. Qualitative data are not numerical, whereas quantitative data are numerical. Examples will vary by student.

7. A discrete variable may assume only certain values. A continuous variable may assume an infinite number of values within a given range. The number of traffic citations issued each day during February in Garden City Beach, South Carolina, is a discrete variable. The weight of commercial trucks passing the weigh station at milepost 195 on Interstate 95 in North Carolina is a continuous variable.

9. **a.** Ordinal
b. Ratio
c. The newer system provides information on the distance between exits.

11. If you were using this store as typical of all Best Buy stores, then the daily number sold last month would be a sample. However, if you considered the store as the only store of interest, then the daily number sold last month would be a population.

13.

	Discrete Variable	Continuous Variable
Qualitative	b. Gender d. Soft drink preference g. Student rank in class h. Rating of a finance professor	
Quantitative	c. Sales volume of digital music players f. SAT scores i. Number of home video screens	a. Salary e. Temperature

	Discrete	Continuous
Nominal	b. Gender	
Ordinal	d. Soft drink preference g. Student rank in class h. Rating of a finance professor	
Interval	f. SAT scores	e. Temperature
Ratio	c. Sales volume of digital music players i. Number of home video screens	a. Salary

15. According to the sample information, 120/300 or 40% would accept a job transfer.

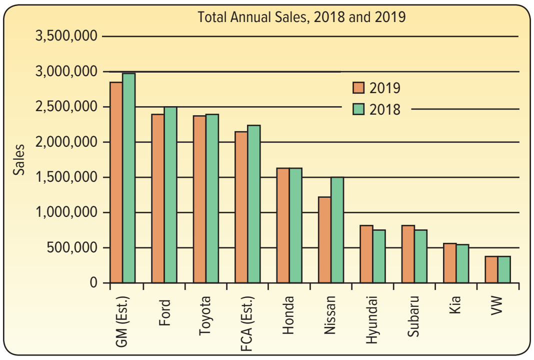
17. **a.**

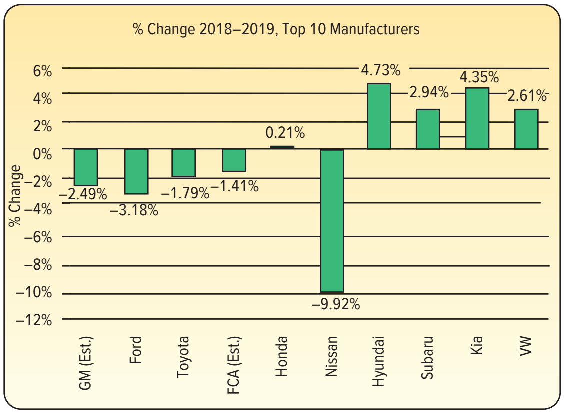
Manufacturer	Difference (units)
Tesla (Est.)	52,800
Hyundai	32,061
Kia	25,665
Subaru	19,982
BMW	13,812
Volvo	9,971
VW	9,258
Porsche	4,366
Mercedes (includes Sprinter)	3,593
Honda	3,342
Mitsubishi	2,972
Land Rover	2,593
Audi	788
Jaguar	568
smart	−596
Others	−2,622
MINI	−7,592
Mazda	−21,773
FCA (Est.)	−31,541
Toyota	−43,325
GM (Est.)	−73,610
Ford	−79,034
Nissan	−148,196

b. Percentage differences with top five and bottom five.

Manufacturer	% Change from 2018 to 2019
Tesla (Est.)	41.85%
Volvo	10.15%
Porsche	7.63%
Hyundai	4.73%
BMW	4.44%
Kia	4.35%
Subaru	2.94%
Land Rover	2.81%
VW	2.61%
Mitsubishi	2.52%
Jaguar	1.86%
Mercedes (includes Sprinter)	1.01%
Audi	0.35%
Honda	0.21%
FCA (Est.)	−1.41%
Toyota	−1.79%
GM (Est.)	−2.49%
Ford	−3.18%
Mazda	−7.25%
Nissan	−9.92%
Others	−13.55%
MINI	−17.38%
smart	−46.71%

c.





19. The graph shows a gradual increase for the years 2009 through 2012 followed by a decrease in earnings from 2012 through 2016. 2017 showed an increase over 2016. Between 2005 and 2017, the earnings ranged from less than $10 billion to over $40 billion. Recent changes may be related to the supply and demand for oil. Demand may be affected by other sources of energy generation (i.e., natural gas, wind, and solar).

21. **a.** League is a qualitative variable; the others are quantitative.
b. League is a nominal-level variable; the others are ratio-level variables.

CHAPTER 2

1. 25% market share.

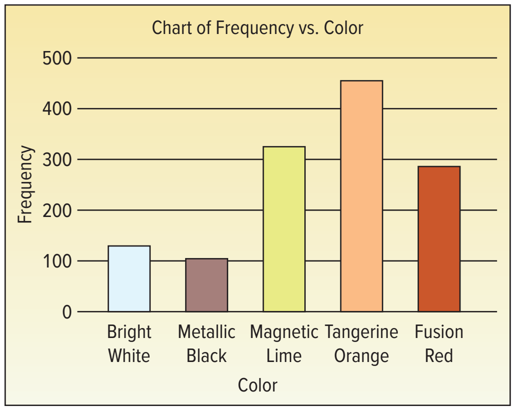
3.

Season	Frequency	Relative Frequency
Winter	100	.10
Spring	300	.30
Summer	400	.40
Fall	200	.20
	1,000	1.00

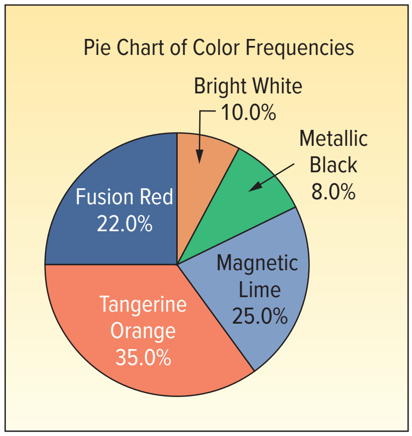
5. **a.** A frequency table.

Color	Frequency	Relative Frequency
Bright White	130	0.10
Metallic Black	104	0.08
Magnetic Lime	325	0.25
Tangerine Orange	455	0.35
Fusion Red	286	0.22
Total	1,300	1.00

b.



c.



d. 350,000 orange, 250,000 lime, 220,000 red, 100,000 white, and 80,000 black, found by multiplying relative frequency by 1,000,000 production.

7. $2^5 = 32$, $2^6 = 64$, therefore, 6 classes

9. $2^7 = 128$, $2^8 = 256$, suggests 8 classes

$i \geq \frac{\$567 - \$235}{8} = 41$ Class intervals of 45 or 50 would be acceptable.

11. **a.** $2^4 = 16$ Suggests 5 classes.

b. $i \geq \frac{31 - 25}{5} = 1.2$ Use interval of 1.5.

c. 24

d.

Units	*f*	Relative Frequency
24.0 up to 25.5	2	0.125
25.5 up to 27.0	4	0.250
27.0 up to 28.5	8	0.500
28.5 up to 30.0	0	0.000
30.0 up to 31.5	2	0.125
Total	16	1.000

e. The largest concentration is in the 27.0 up to 28.5 class (8).

13. **a.**

Number of Visits	*f*
0 up to 3	9
3 up to 6	21
6 up to 9	13
9 up to 12	4
12 up to 15	3
15 up to 18	1
Total	51

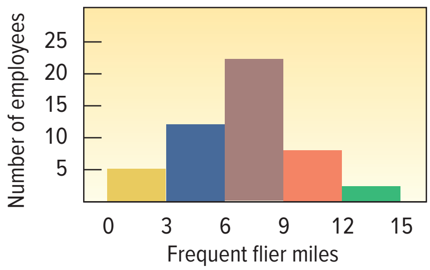
b. The largest group of shoppers (21) shop at the BiLo Supermarket 3, 4, or 5 times during a month period. Some customers visit the store only 1 time during the month, but others shop as many as 15 times.

c.

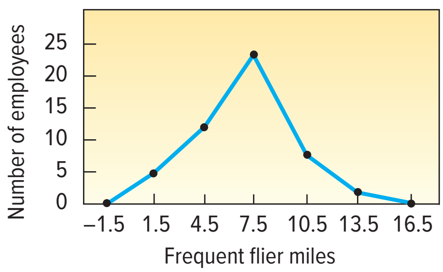
Number of Visits	Percent of Total
0 up to 3	17.65
3 up to 6	41.18
6 up to 9	25.49
9 up to 12	7.84
12 up to 15	5.88
15 up to 18	1.96
Total	100.00

15. **a.** Histogram
b. 100
c. 5
d. 28
e. 0.28
f. 12.5
g. 13

17. **a.** 50
b. 1.5 thousand miles, or 1,500 miles.
c.



d. $X = 1.5$, $Y = 5$
e.



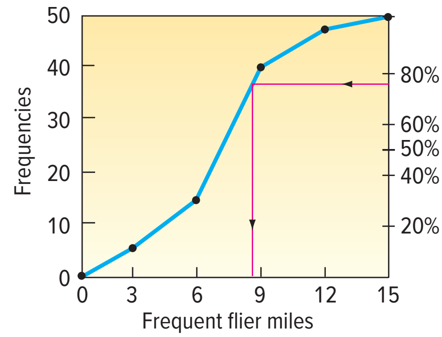
f. For the 50 employees, about half traveled between 6,000 and 9,000 miles. Five employees traveled less than 3,000 miles, and 2 traveled more than 12,000 miles.

19. **a.** 40
b. 5
c. 11 or 12
d. About $18/hr
e. About $9/hr
f. About 75%

21. **a.** 5
b.

Miles	*CF*
Less than 3	5
Less than 6	17
Less than 9	40
Less than 12	48
Less than 15	50

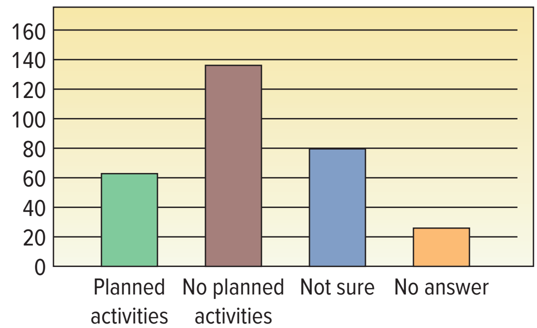
c.



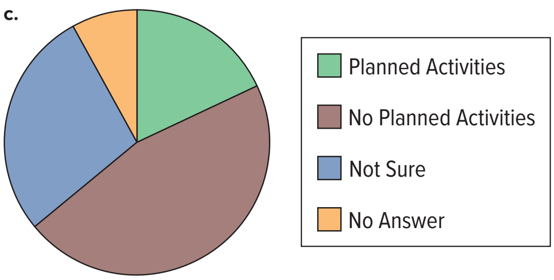
d. About 8.7 thousand miles

23. **a.** A qualitative variable uses either the nominal or ordinal scale of measurement. It is usually the result of counts. Quantitative variables are either discrete or continuous. There is a natural order to the results for a quantitative variable. Quantitative variables can use either the interval or ratio scale of measurement.
b. Both types of variables can be used for samples and populations.

25. **a.** Frequency table
b.



c.



d. A pie chart would be better because it clearly shows that nearly half of the customers prefer no planned activities.

27. $2^6 = 64$ and $2^7 = 128$, suggest 7 classes

29. **a.** 5, because $2^4 = 16 < 25$ and $2^5 = 32 > 25$
b. $i \geq \dfrac{48 - 16}{5} = 6.4$ Use interval of 7.
c. 15
d.

Class		Frequency
15 up to 22	|||	3
22 up to 29	~~||||~~ |||	8
29 up to 36	~~||||~~ ||	7
36 up to 43	~~||||~~	5
43 up to 50	||	2
		25

e. It is fairly symmetric, with most of the values between 22 and 36.

31. **a.** $2^5 = 32$, $2^6 = 64$, 6 classes recommended.

b. $i = \dfrac{10 - 1}{6} = 1.5$ use an interval of 2.

c. 0

d.

Class	Frequency
0 up to 2	1
2 up to 4	5
4 up to 6	12
6 up to 8	17
8 up to 10	8
10 up to 12	2

e. The distribution is fairly symmetric or bell-shaped with a large peak in the middle of the two classes of 4 up to 8.

33.

Number of Class	Frequency
4–15	9
16–27	4
28–39	6
40–51	1
Grand Total	**20**

This distribution is positively skewed with a "tail" to the right. Based on the data, 13 of the customers required between 4 and 27 attempts before actually talking with a person. Seven customers required more.

35. **a.** 56

b. 10 (found by 60 – 50)

c. 55

d. 17

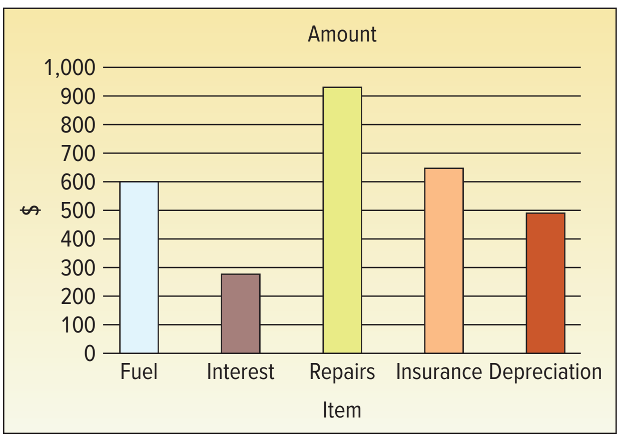
37. **a.** Use $35 because the minimum is ($265 – $82)/6 = $30.5.

b.

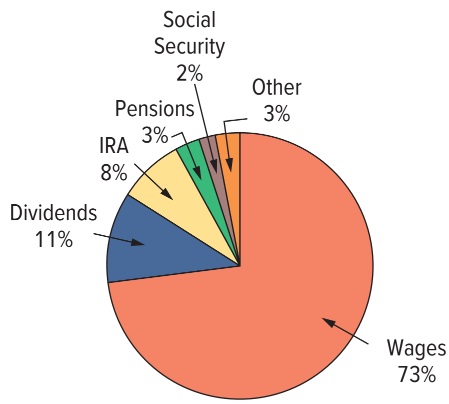
$ 70 up to $105	4
105 up to 140	17
140 up to 175	14
175 up to 210	2
210 up to 245	6
245 up to 280	1

c. The purchases range from a low of about $70 to a high of about $280. The concentration is in the $105 up to $140 and $140 up to $175 classes.

39. Bar charts are preferred when the goal is to compare the actual amount in each category.



41.

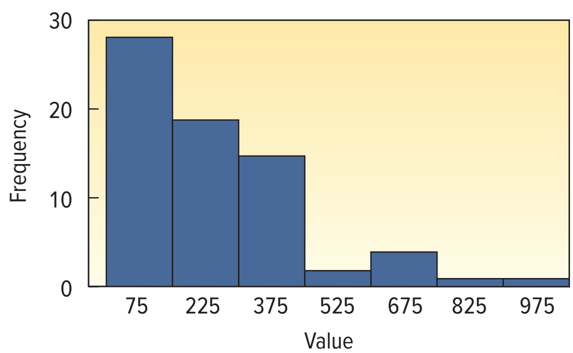


SC Income	Percent	Cumulative
Wages	73	73
Dividends	11	84
IRA	8	92
Pensions	3	95
Social Security	2	97
Other	3	100

By far the largest part of income in South Carolina is wages. Almost three-fourths of the adjusted gross income comes from wages. Dividends and IRAs each contribute roughly another 10%.

43. **a.** Since $2^6 = 64 < 70 < 128 = 2^7$, 7 classes are recommended. The interval should be at least (1,002.2 – 3.3)/7 = 142.7. Use 150 as a convenient value.

b. Based on the histogram, the majority of people have less than $500,000 in their investment portfolio and may not have enough money for retirement. Merrill Lynch financial advisors need to promote the importance of investing for retirement in this age group.

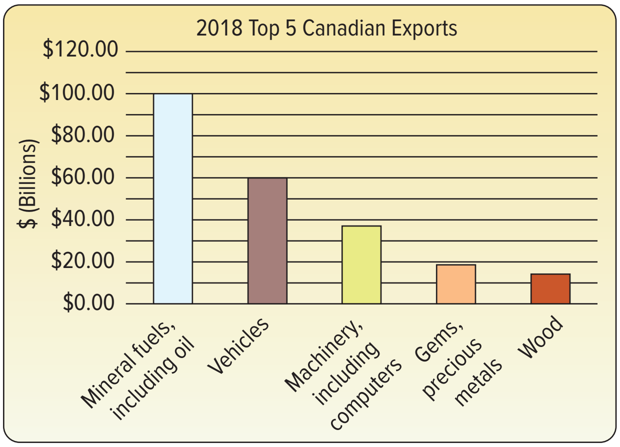


45. **a.** Pie chart

b. 700, found by 0.7(1,000)

c. Yes, 0.70 + 0.20 = 0.90

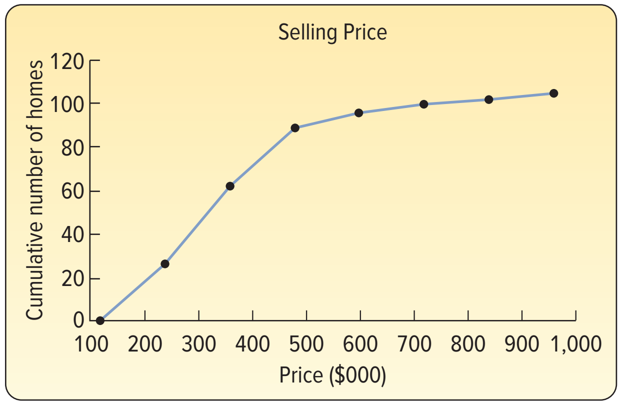
47. **a.**



b. 35.5%, found by (99.3 + 60.5)/450.7

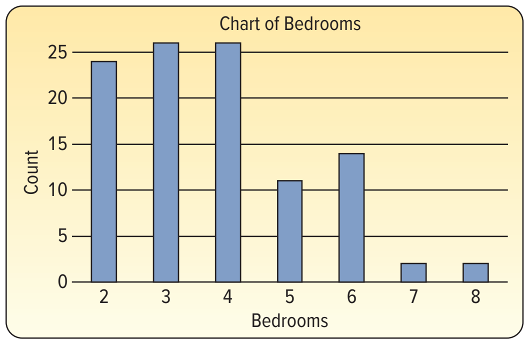
c. 70.4%, found by (99.3 + 60.5)/(99.3 + 60.5 + 34.5 + 18.3 + 14.3)

c.



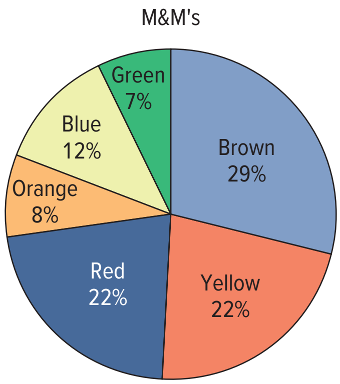
Fifty percent (about 52) of the homes sold for about $320,000 or less.
The top 10% (about 90) of homes sold for at least $520,000.
About 41% (about 41) of the homes sold for less than $300,000.

d.



Two-, three-, and four-bedroom houses are most common with about 25 houses each. Seven- and eight-bedroom houses are rather rare.

49.



Brown, yellow, and red make up almost 75% of the candies. The other 25% is composed of blue, orange, and green.

51. There are many choices and possibilities here. For example you could choose to start the first class at 160,000 rather than 120,000. The choice is yours!
$i >= (919{,}480 - 167{,}962)/7 = 107{,}360$. Use intervals of 120,000.

Selling Price (000)	Frequency	Cumulative Frequency
$120 up to $240	26	26
240 up to 360	36	62
360 up to 480	27	89
480 up to 600	7	96
600 up to 720	4	100
720 up to 840	2	102
840 up to 960	1	105

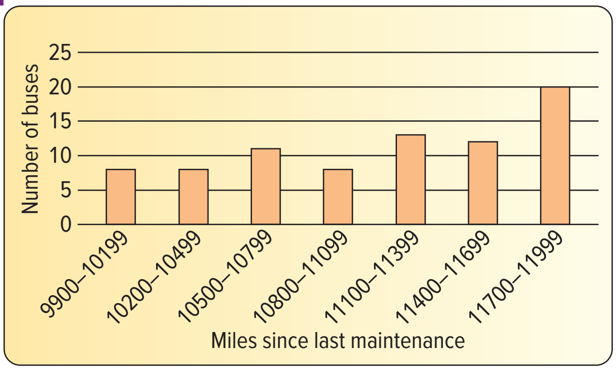
a. Most homes (60%) sell between $240,000 and $480,000.

b. The typical price in the first class is $180,000 and in the last class it is $900,000.

53. Since $2^6 = 64 < 80 < 128 = 2^7$, use seven classes. The interval should be at least (11,973 − 10,000)/7 = 281 miles. Use 300. The resulting frequency distribution is:

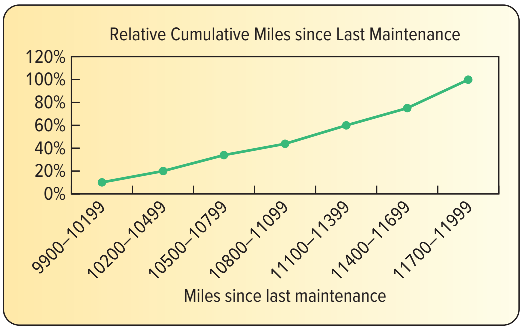
Class	*f*
9,900 up to 10,200	8
10,200 up to 10,500	8
10,500 up to 10,800	11
10,800 up to 11,100	8
11,110 up to 11,400	13
11,400 up to 11,700	12
11,700 up to 12,000	20

a. The typical amount driven, or the middle of the distribution is about 11,100 miles. Based on the frequency distribution, the range is from 9,900 up to 12,000 miles.



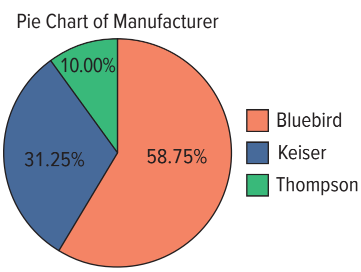
b. The distribution is somewhat "skewed" with a longer "tail" to the left and no outliers.

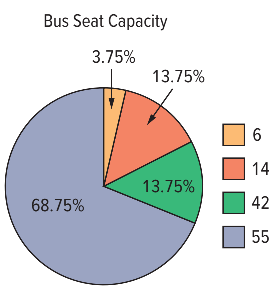
c.



Forty percent of the buses were driven fewer than about 10,800 miles. About 30% of the 80 buses (about 24) were driven less than 10,500 miles.

d. The first diagram shows that Bluebird makes about 59% of the buses, Keiser about 31%, and Thompson only about 10%. The second chart shows that nearly 69% of the buses have 55 seats.





CHAPTER 3

1. $\mu = 5.4$, found by 27/5

3. **a.** $\bar{x} = 7.0$, found by 28/4
b. $(5 - 7) + (9 - 7) + (4 - 7) + (10 - 7) = 0$

5. $\bar{x} = 14.58$, found by 43.74/3

7. **a.** 15.4, found by 154/10
b. Population parameter, since it includes all the salespeople at Midtown Ford

9. **a.** $54.55, found by $1,091/20
b. A sample statistic—assuming that the power company serves more than 20 customers

11. $\bar{x} = \dfrac{\Sigma x}{n}$ so

$\Sigma x = \bar{x} \cdot n = (\$5{,}430)(30) = \$162{,}900$

13. **a.** No mode
b. The given value would be the mode.
c. 3 and 4 bimodal

15. **a.** Mean = 3.583
b. Median = 5
c. Mode = 5

17. **a.** Median = 2.9
b. Mode = 2.9

19. $\bar{x} = \dfrac{647}{11} = 58.82$

Median = 58, Mode = 58
Any of the three measures would be satisfactory.

21. **a.** $\bar{x} = \dfrac{85.9}{12} = 7.16$
b. Median = 7.2. There are several modes: 6.6, 7.2, and 7.3.
c. $\bar{x} = \dfrac{30.7}{4} = 7.675$,
Median = 7.85
About 0.5 percentage point higher in winter

23. $46.09, found by $\dfrac{300(\$53) + 400(\$42) + 400(\$45)}{300 + 400 + 400}$

25. $22.50, found by [50($12) + 50($20) + 100($29)]/200

27. **a.** 7, found by 10 – 3
b. 6, found by 30/5
c. 6.8, found by 34/5
d. The difference between the highest number sold (10) and the smallest number sold (3) is 7. The typical squared deviation from 6 is 6.8.

29. **a.** 30, found by 54 – 24
b. 38, found by 380/10
c. 74.4, found by 744/10
d. The difference between 54 and 24 is 30. The average of the squared deviations from 38 is 74.4.

31.

State	Mean	Median	Range
California	33.10	34.0	32
Iowa	24.50	25.0	19

The mean and median ratings were higher, but there was also more variation in California.

33. **a.** 5
b. 4.4, found by

$$\frac{(8 - 5)^2 + (3 - 5)^2 + (7 - 5)^2 + (3 - 5)^2 + (4 - 5)^2}{5}$$

35. **a.** $2.77
b. 1.26, found by

$$\frac{(2.68 - 2.77)^2 + (1.03 - 2.77)^2 + (2.26 - 2.77)^2 + (4.30 - 2.77)^2 + (3.58 - 2.77)^2}{5}$$

37. **a.** Range: 7.3, found by 11.6 – 4.3. Arithmetic mean: 6.94, found by 34.7/5. Variance: 6.5944, found by 32.972/5. Standard deviation: 2.568, found by $\sqrt{6.5944}$.
b. Dennis has a higher mean return (11.76 > 6.94). However, Dennis has greater spread in its returns on equity (16.89 > 6.59).

39. **a.** $\bar{x} = 4$

$$s^2 = \frac{(7 - 4)^2 + \cdots + (3 - 4)^2}{5 - 1} = \frac{22}{5 - 1} = 5.5$$

b. $s = 2.3452$

41. **a.** $\bar{x} = 38$

$$s^2 = \frac{(28 - 38)^2 + \cdots + (42 - 38)^2}{10 - 1} = \frac{744}{10 - 1} = 82.667$$

b. $s = 9.0921$

43. a. $\bar{x} = \dfrac{951}{10} = 95.1$

$$s^2 = \frac{(101 - 95.1)^2 + \cdots + (88 - 95.1)^2}{10 - 1}$$

$$= \frac{1{,}112.9}{9} = 123.66$$

b. $s = \sqrt{123.66} = 11.12$

45. About 69%, found by $1 - 1/(1.8)^2$

47. a. About 95%
b. 47.5%, 2.5%

49. a. Mean = 5, found by (6 + 4 + 3 + 7 + 5)/5. Median is 5, found by rearranging the values and selecting the middle value.
b. Population, because all partners were included
c. $\Sigma(x - \mu) = (6 - 5) + (4 - 5) + (3 - 5) + (7 - 5) + (5 - 5) = 0$

51. $\bar{x} = \dfrac{545}{16} = 34.06$

Median = 37.50

53. The mean is 35.675, found by 1,427/40. The median is 36, found by sorting the data and averaging the 20th and 21st observations.

55. $\bar{x}_w = \dfrac{\$5.00(270) + \$6.50(300) + \$8.00(100)}{270 + 300 + 100} = \$6.12$

57. $\bar{x}_w = \dfrac{15{,}300(4.5) + 10{,}400(3.0) + 150{,}600(10.2)}{176{,}300} = 9.28$

59. a. 55, found by 72 – 17
b. 17.6245, found by the square root of 2795.6/9

61. a. This is a population because it includes all the public universities in Ohio.
b. The mean is 25,097.
c. The median is 20,491 (University of Toledo).
d. There is no mode for this data.
e. I would select the median because the mean is biased by a few schools (Ohio State, Cincinnati, Kent State, and Ohio University) that have extremely high enrollments compared to the other schools.
f. The range is (67,524 – 1,748) = 65,776.
g. The standard deviation is 17,307.39.

63. a. The mean is \$717.20, found by \$17,930/25. The median is \$717.00 and there are two modes, \$710 and \$722.
b. The range is \$90, found by \$771 – \$681, and the standard deviation is \$24.87, found by the square root of 14,850/24.
c. From \$667.46 up to \$766.94, found by \$717.20 ± 2(\$24.87)

65. a. $\bar{x} = \dfrac{273}{30} = 9.1$, Median = 9
b. Range = 18 – 4 = 14

$$s = \sqrt{\frac{368.7}{30 - 1}} = 3.57$$

67. a. The mean team salary is \$139,174,000 and the median is \$141,715,000. Since the distribution is skewed, the median value of \$141,715,000 is more typical.
b. The range is \$158,590,000; found by \$227,400,000 – \$68,810,000. The standard deviation is \$41,101,000. At least 95% of the team salaries are between \$56,971,326 and \$221,376,000; found by \$139,174,000 plus or minus 2(\$41,101,000).

CHAPTER 4

1. In a histogram, observations are grouped so their individual identity is lost. With a dot plot, the identity of each observation is maintained.

3. a. Dot plot
b. 15
c. 1, 7
d. 2 and 3

5. Median = 53, found by $(11 + 1)(\frac{1}{2})$ ∴ 6th value in from lowest

$Q_1 = 49$, found by $(11 + 1)(\frac{1}{4})$ ∴ 3rd value in from lowest

$Q_3 = 55$, found by $(11 + 1)(\frac{3}{4})$ ∴ 9th value in from lowest

7. a. $Q_1 = 33.25$, $Q_3 = 50.25$
b. $D_2 = 27.8$, $D_8 = 52.6$
c. $P_{67} = 47$

9. a. 800
b. $Q_1 = 500$, $Q_3 = 1{,}200$
c. 700, found by 1,200 – 500
d. Less than 200 or more than 1,800
e. There are no outliers.
f. The distribution is positively skewed.

11.

```
                   +--------------------------+
         ---------I           +               I------------------------
                   +--------------------------+
 -----+----------+----------+----------+----------+----------+---
     14.0       21.0       28.0       35.0       42.0       49.0
```

The distribution is somewhat positively skewed. Note that the dashed line above 35 is longer than below 18.

13. a. The mean is 30.8, found by 154/5. The median is 31.0, and the standard deviation is 3.96, found by

$$s = \sqrt{\frac{62.8}{4}} = 3.96$$

b. –0.15, found by $\dfrac{3(30.8 - 31.0)}{3.96}$

c.

Salary	$\left(\dfrac{x - \bar{x}}{s}\right)$	$\left(\dfrac{x - \bar{x}}{s}\right)^3$
36	1.313131	2.264250504
26	−1.212121	−1.780894343
33	0.555556	0.171467764
28	−0.707071	−0.353499282
31	0.050505	0.000128826
		0.301453469

0.125, found by [5/(4 × 3)] × 0.301

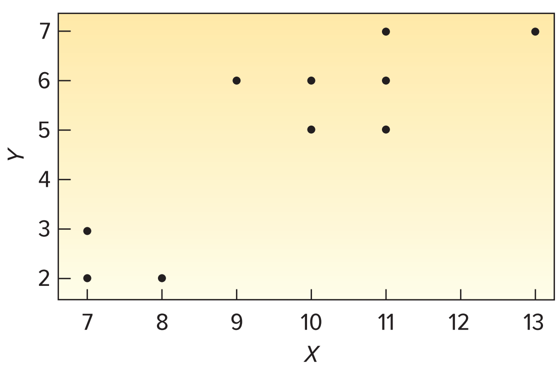
15. a. The mean is 21.93, found by 328.9/15. The median is 15.8, and the standard deviation is 21.18, found by

$$s = \sqrt{\frac{6{,}283}{14}} = 21.18$$

b. 0.868, found by [3(21.93 – 15.8)]/21.18
c. 2.444, found by [15/(14 × 13)] × 29.658

17. The correlation coefficient is 0.86. Larger values of x are associated with larger values of y. The relationship is fairly strong.

Scatter Diagram of Y versus X



There is a positive relationship between the variables.

19. **a.** Both variables are nominal scale.
b. Contingency table
c. Yes, 58.5%, or more than half of the customers order dessert.
No, only 32% of lunch customers order dessert.
Yes, 85% of dinner customers order dessert.

21. **a.** Dot plot **b.** 15 **c.** 5

23. **a.** $L_{50} = (20 + 1)\frac{50}{100} = 10.50$

$\text{Median} = \dfrac{83.7 + 85.6}{2} = 84.65$

$L_{25} = (21)(.25) = 5.25$

$Q_1 = 66.6 + .25(72.9 - 66.6) = 68.175$

$L_{75} = 21(.75) = 15.75$

$Q_3 = 87.1 + .75(90.2 - 87.1) = 89.425$

b. $L_{26} = 21(.26) = 5.46$
$P_{26} = 66.6 + .46(72.9 - 66.6) = 69.498$
$L_{83} = 21(.83) = 17.43$
$P_{83} = 93.3 + .43(98.6 - 93.3)$
$= 95.579$

c.

```
                 -------------------------
-------------I                   +       I---------------
                 -------------------------

--------+---------+---------+---------+---------+--------C20
      64.0      72.0      80.0      88.0      96.0
```

25. **a.** $Q_1 = 26.25$, $Q_3 = 35.75$, Median = 31.50

```
             -----------------------
-------------I          +          I-------------------
             -----------------------

------+---------+---------+---------+---------+---------+
    24.5      28.0      31.5      35.0      38.5      42.0
```

b. $Q_1 = 33.25$, $Q_3 = 38.75$, Median = 37.50

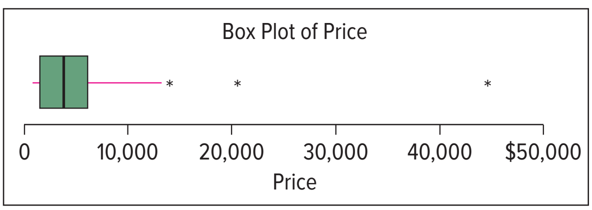
```
      ---------------------
------I          +        I----------------------
      ---------------------

------+---------+---------+---------+---------+---------+
    32.5      35.0      37.5      40.0      42.5      45.0
```

c. The median time for public transportation is about 6 minutes less. There is more variation in public transportation. The difference between Q_1 and Q_3 is 9.5 minutes for public transportation and 5.5 minutes for private transportation.

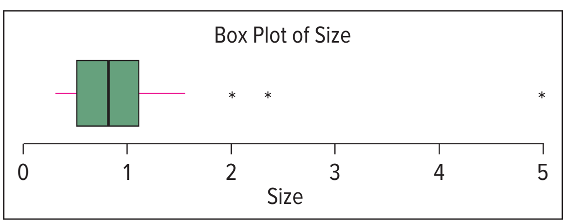
27. The distribution is positively skewed. The first quartile is about \$20 and the third quartile is about \$90. There is one outlier located at \$255. The median is about \$50.

29. **a.**



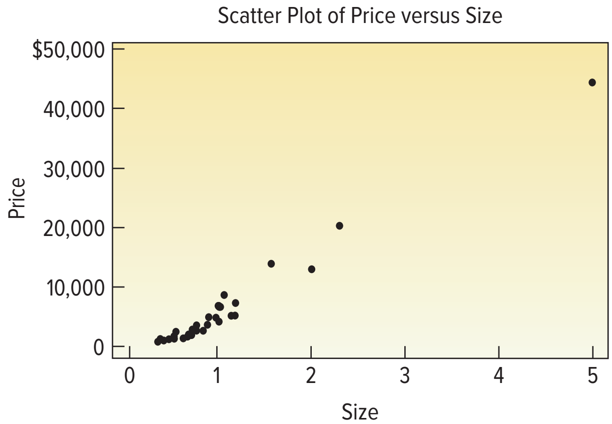
Median is 3,733. First quartile is 1,478. Third quartile is 6,141. So prices over 13,135.5, found by 6,141 + 1.5 × (6,141 − 1,478), are outliers. There are three (13,925; 20,413; and 44,312).

b.



Median is 0.84. First quartile is 0.515. Third quartile is 1.12. So sizes over 2.0275, found by 1.12 + 1.5 (1.12 − 0.515), are outliers. There are three (2.03; 2.35; and 5.03).

c.



There is a direct association between them. The first observation is larger on both scales.

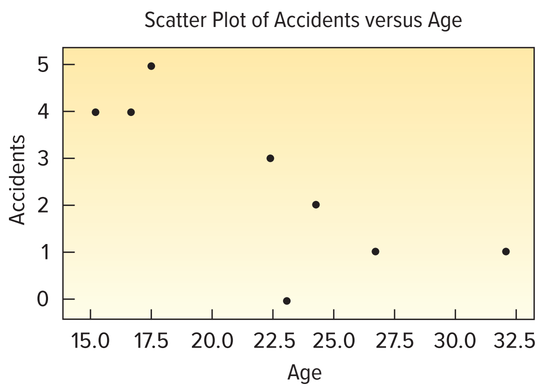
d.

Shape/ Cut	Average	Good	Ideal	Premium	Ultra Ideal	All
Emerald	0	0	1	0	0	1
Marquise	0	2	0	1	0	3
Oval	0	0	0	1	0	1
Princess	1	0	2	2	0	5
Round	1	3	3	13	3	23
Total	2	5	6	17	3	33

The majority of the diamonds are round (23). Premium cut is most common (17). The Round Premium combination occurs most often (13).

31. $sk = 0.065$ or $sk = \dfrac{3(7.7143 - 8.0)}{3.9036} = -0.22$

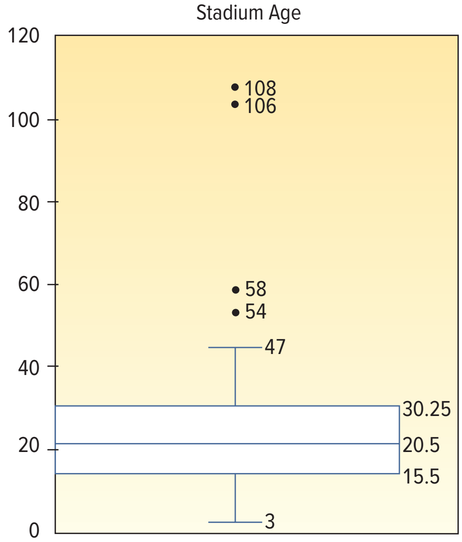
33.



As age increases, the number of accidents decreases.

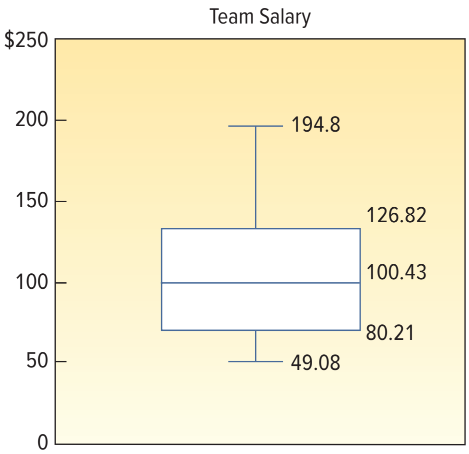
35. **a.** 139,340,000
b. 5.4% unemployed, found by (7,523/139,340)100
c. Men = 5.64%
Women = 5.12%

37. **a.** Box plot of age assuming the current year is 2020.



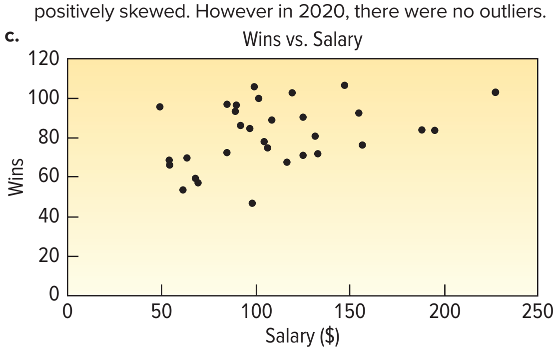
Distribution of stadium is highly positively skewed to the right. Any stadium older than 52.375 years (Q3 + 1.5) (Q3 − Q1) = 30.25 + 1.5(30.25 − 15.5) is an outlier. Boston Red Sox, Chicago Cubs, LA Dodgers, Oakland Athletics, and LA Angels.

b.



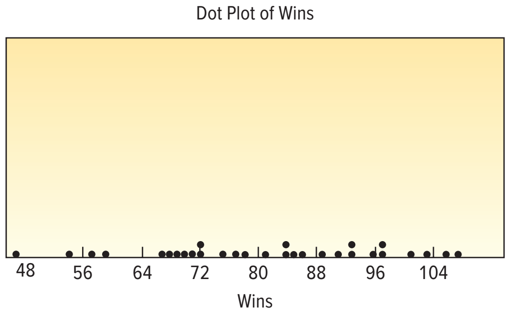
The first quartile is \$80.21 million and the third is \$126.82 million. Outliers are greater than (Q3 + 1.5(Q3 − Q1) or 126.82 + 1.5*(126.82 − 80.21) = \$196.74 million. The distribution is positively skewed. However in 2020, there were no outliers.

c.



The correlation coefficient is 0.29. The relationship is generally positive but the relationship is weak. Higher salaries are not strongly associated with more wins.

d.



The dot plot shows a range of wins from the 40s to the 100s. Most teams appear to win between 65 and 90 games in a season. 7 teams won 96 or more games. 9 teams won less than 72 games. 16 teams won 81 or more games out of a total possible 162 games.

CHAPTER 5

1.

	Person	
Outcome	**1**	**2**
1	A	A
2	A	F
3	F	A
4	F	F

3. **a.** .176, found by $\frac{6}{34}$ **b.** Empirical

5. **a.** Empirical
b. Classical
c. Classical
d. Empirical, based on seismological data

7. **a.** The survey of 40 people about environmental issues
b. 26 or more respond yes, for example.
c. 10/40 = .25
d. Empirical
e. The events are not equally likely, but they are mutually exclusive.

9. **a.** Answers will vary. Here are some possibilities: 1236, 5124, 6125, 9999.
b. $(1/10)^4$
c. Classical

11. $P(A \text{ or } B) = P(A) + P(B) = .30 + .20 = .50$
$P(\text{neither}) = 1 - .50 = .50.$

13. **a.** 102/200 = .51
b. .49, found by 61/200 + 37/200 = .305 + .185. Special rule of addition.

15. $P(\text{above } C) = .25 + .50 = .75$

17. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = .20 + .30 - .15 = .35$

19. When two events are mutually exclusive, it means that if one occurs, the other event cannot occur. Therefore, the probability of their joint occurrence is zero.

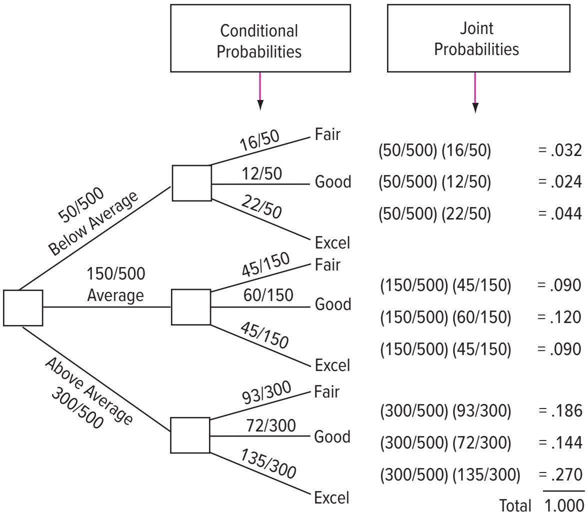
21. Let A denote the event the fish is green and B be the event the fish is male.
a. $P(A) = 80/140 = 0.5714$
b. $P(B) = 60/140 = 0.4286$
c. $P(A \text{ and } B) = 36/140 = 0.2571$
d. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 80/140 + 60/140 - 36/140 = 104/140 = 0.7429$

23. $P(A \text{ and } B) = P(A) \times P(B|A) = .40 \times .30 = .12$

25. .90, found by (.80 + .60) − .5.
.10, found by (1 − .90).

27. **a.** $P(A_1) = 3/10 = .30$
b. $P(B_1|A_2) = 1/3 = .33$
c. $P(B_2 \text{ and } A_3) = 1/10 = .10$

29. **a.** A contingency table
b. .27, found by 300/500 × 135/300
c. The tree diagram would appear as:



31. **a.** Out of all 545 students, 171 prefer skiing. So the probability is 171/545, or 0.3138.
b. Out of all 545 students, 155 are in junior college. Thus, the probability is 155/545, or 0.2844.
c. Out of 210 four-year students, 70 prefer ice skating. So the probability is 70/210, or 0.3333.
d. Out of 211 students who prefer snowboarding, 68 are in junior college. So the probability is 68/211, or 0.3223.
e. Out of 180 graduate students, 74 prefer skiing and 47 prefer ice skating. So the probability is (74 + 47)/180 = 121/180, or 0.6722.

33. **a.** 78,960,960
b. 840, found by (7)(6)(5)(4). That is 7!/3!
c. 10, found by 5!/3!2!

35. 210, found by (10)(9)(8)(7)/(4)(3)(2)

37. 120, found by 5!

39. (4)(8)(3) = 96 combinations

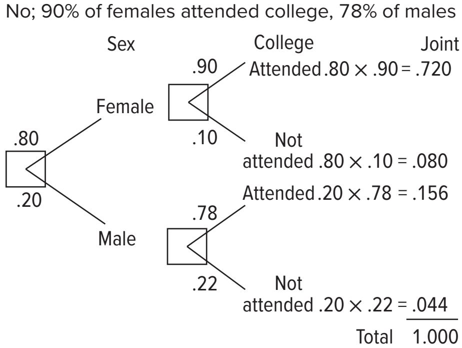
41. **a.** Asking teenagers to compare their reactions to a newly developed soft drink.
b. Answers will vary. One possibility is more than half of the respondents like it.

43. Subjective

45. **a.** 4/9, found by (2/3) · (2/3)
b. 3/4, because (3/4) · (2/3) = 0.5

47. **a.** .8145, found by $(.95)^4$
b. Special rule of multiplication
c. $P(A \text{ and } B \text{ and } C \text{ and } D) = P(A) \times P(B) \times P(C) \times P(D)$

49. **a.** .08, found by .80 × .10
b. No; 90% of females attended college, 78% of males
c.



d. Yes, because all the possible outcomes are shown on the tree diagram.

51. **a.** 0.57, found by 57/100
b. 0.97, found by (57/100) + (40/100)
c. Yes, because an employee cannot be both.
d. 0.03, found by 1 − 0.97

53. **a.** 1/2, found by (2/3)(3/4)
b. 1/12, found by (1/3)(1/4)
c. 11/12, found by 1 − 1/12

55. **a.** 0.9039, found by $(0.98)^5$
b. 0.0961, found by 1 − 0.9039

57. **a.** 0.0333, found by (4/10)(3/9)(2/8)
b. 0.1667, found by (6/10)(5/9)(4/8)
c. 0.8333, found by 1 − 0.1667
d. Dependent

59. **a.** 0.3818, found by (9/12)(8/11)(7/10)
b. 0.6182, found by 1 − 0.3818

61. **a.** $P(S) \cdot P(R|S) = .60(.85) = 0.51$
b. $P(S) \cdot P(PR|S) = .60(1 - .85) = 0.09$

63. **a.** $P(\text{not perfect}) = P(\text{bad sector}) + P(\text{defective})$

$$= \frac{112}{1{,}000} + \frac{31}{1{,}000} = .143$$

b. $P(\text{defective} \mid \text{not perfect}) = \dfrac{.031}{.143} = .217$

65. **a.** 0.1 + 0.02 = 0.12
b. 1 − 0.12 = 0.88
c. $(0.88)^3 = 0.6815$
d. 1 − .6815 = 0.3185

67. Yes, 256 is found by 2^8.

69. .9744, found by $1 - (.40)^4$

71. **a.** 0.193, found by .15 + .05 − .0075 = .193
b. .0075, found by (.15)(.05)

73. **a.** $P(F \text{ and } >60) = .25$, found by solving with the general rule of multiplication: $P(F) \cdot P(>60|F) = (.5)(.5)$
b. 0
c. .3333, found by 1/3

75. $26^4 = 456{,}976$

77. 1/(3,628,800) ≈ 0.00000028

79. 0.512, found by $(0.8)^3$

81. .525, found by $1 - (.78)^3$

83. **a.**

Wins	# Teams
40–49	1
50–59	3
60–69	4
70–79	6
80–89	6
90–99	6
100–109	4
Grand Total	**30**

1. 10/30 = 0.33
2. 10/10 = 1.00
3. In 2019, winning 90 or more games in a season assured a place in the end-of-season playoffs.

b.

Frequency (# Teams) by League			
Home Runs	**American**	**National**	**Grand Total**
140–169	2	3	5
170–199	1	0	1
200–229	5	6	11
230–259	4	5	9
260–289	1	1	2
270–319	2	0	2
Grand Total	**15**	**15**	**30**

	Relative Frequency		
Home Runs	**American**	**National**	**Both Leagues**
140–169	13.33%	20.00%	16.67%
170–199	6.67%	0.00%	3.33%
200–229	33.33%	40.00%	36.67%
230–259	26.67%	33.33%	30.00%
260–289	6.67%	6.67%	6.67%
270–319	13.33%	0.00%	6.67%
Grand Total	**100.00**%	**100.00**%	**100.00**%

1. In the American League, the probability that a team hits 200 or more home runs is 0.80.
2. In the National League, the probability that a team hits 200 or more home runs is 0.80.
3. For the 2019 season, the distributions of home runs for each league appear to be very similar. Using the designated hitter in the American League did not provide an advantage.

CHAPTER 6

1. Mean = 1.3, variance = .81, found by:

$\mu = 0(.20) + 1(.40) + 2(.30) + 3(.10) = 1.3$

$\sigma^2 = (0 - 1.3)^2(.2) + (1 - 1.3)^2(.4)$
$\quad + (2 - 1.3)^2(.3) + (3 - 1.3)^2(.1)$
$\quad = .81$

3. Mean = 14.5, variance = 27.25, found by:

$\mu = 5(.1) + 10(.3) + 15(.2) + 20(.4) = 14.5$

$\sigma^2 = (5 - 14.5)^2(.1) + (10 - 14.5)^2(.3)$
$\quad + (15 - 14.5)^2(.2) + (20 - 14.5)^2(.4)$
$\quad = 27.25$

5. a.

Calls, x	Frequency	$P(x)$	$xP(x)$	$(x - \mu)^2 P(x)$
0	8	.16	0	.4624
1	10	.20	.20	.0980
2	22	.44	.88	.0396
3	9	.18	.54	.3042
4	1	.02	.08	.1058
	50		1.70	1.0100

b. Discrete distribution, because only certain outcomes are possible.
c. 0.20 found by $P(x = 3) + P(x = 4) = 0.18 + 0.02 = 0.20$
d. $\mu = \Sigma x \cdot P(x) = 1.70$
e. $\sigma = \sqrt{1.01} = 1.005$

7.

Amount	$P(x)$	$xP(x)$	$(x - \mu)^2 P(x)$
10	.50	5	60.50
25	.40	10	6.40
50	.08	4	67.28
100	.02	2	124.82
		21	259.00

a. 0.10 found by $P(x = 50) + P(x = 100) = 0.08 + 0.02 = 0.10$
b. $\mu = \Sigma\, xP(x) = 21$
c. $\sigma^2 = \Sigma(x - \mu)^2 P(x) = 259$

$\sigma = \sqrt{259} = 16.093$

9. Using the binomial table, Excel, or the binomial formula:

x	$P(x)$
0	0.4096
1	0.4096
2	0.1536
3	0.0256
4	0.0016

Using the binomial formula with $x = 2$ as an example:

$P(2) = \frac{4!}{2!(4-2)!}(.2)^2(.8)^{4-2} = 0.1536$

11. a.

x	$P(x)$
0	.064
1	.288
2	.432
3	.216

b. $\mu = 1.8$
$\sigma^2 = 0.72$
$\sigma = \sqrt{0.72} = .8485$

13. a. .2668, found by $P(2) = \frac{9!}{(9-2)!2!}(.3)^2(.7)^7$
b. .1715, found by $P(4) = \frac{9!}{(9-4)!4!}(.3)^4(.7)^5$
c. .0404, found by $P(0) = \frac{9!}{(9-0)!0!}(.3)^0(.7)^9$

15. a. .2824, found by $P(0) = \frac{12!}{(12-0)!0!}(.1)^0(.9)^{12}$
b. .3766, found by $P(1) = \frac{12!}{(12-1)!1!}(.1)^1(.9)^{11}$
c. .2301, found by $P(2) = \frac{12!}{(12-2)!2!}(.1)^2(.9)^{10}$
d. $\mu = 1.2$, found by 12(.1)
$\sigma = 1.0392$, found by $\sqrt{1.08}$

17. a. The random variable is the count of the 15 accountants who have a CPA. The random variable follows a binomial probability distribution. The random variable meets all 4 criteria for a binomial distributor: (1) Fixed number of trials (15), (2) each trial results in a success or failure (the accountant has a CPA or not), (3) known probability of success (0.52), and (4) each trial is independent of any other selection.
b. Using the binomial table, Excel, or the binomial formula, the probability distribution follows. P(5 of the 15 accountants with a CPA) = 0.0741.

x	$P(x)$	x	$P(x)$
0	0.0000	8	Q.2020
1	0.0003	9	0.1702
2	0.0020	10	0.1106
3	0.0096	11	0,0545
4	0.0311	12	0.0197
5	0.0741	13	0.0049
6	0.1338	14	0.0008
7	0.1864	15	0.0001

c. 0.3884, found by $P(x = 7) + P(x = 8)$
d. Mean $= n\pi = (15)(.52) = 7.8$ accountants
e. Variance $= n\pi(1 - \pi) = (15)(.52)(.48) = 3.744$

19. a. 0.296, found by using Appendix B.1 with n of 8, π of 0.30, and x of 2
b. $P(x \le 2) = 0.058 + 0.198 + 0.296 = 0.552$
c. 0.448, found by $P(x \ge 3) = 1 - P(x \le 2) = 1 - 0.552$

21. a. 0.387, found from Appendix B.1 with n of 9, π of 0.90, and x of 9
b. $P(x < 5) = 0.001$
c. 0.992, found by 1 – 0.008
d. 0.947, found by 1 – 0.053

23. a. $\mu = 10.5$, found by 15(0.7) and $\sigma = \sqrt{15(0.7)(0.3)} = 1.7748$
b. 0.2061, found by $\frac{15!}{10!5!}(0.7)^{10}(0.3)^5$
c. 0.4247, found by 0.2061 + 0.2186
d. 0.5154, found by 0.2186 + 0.1700 + 0.0916 + 0.0305 + 0.0047

25. **a.** .6703
b. .3297

27. **a.** .0613
b. .0803

29. $\mu = 6$
$P(x \geq 5) = 1 - (.0025 + .0149 + .0446 + .0892 + .1339) = .7149$

31. A random variable is an outcome that results from a chance experiment. A probability distribution also includes the likelihood of each possible outcome.

33. $\mu = 1{,}000(.25) + \$2{,}000(.60) + \$5{,}000(.15) = \$2{,}200$
$$\sigma^2 = (1{,}000 - 2{,}200)^2 .25 + (2{,}000 - 2{,}200)^2 .60 + (5{,}000 - 2{,}200)^2 .15 = 1{,}560{,}000$$

35. $\mu = 12(.25) + \cdots + 15(.1) = 13.2$
$\sigma^2 = (12 - 13.2)^2 .25 + \cdots + (15 - 13.2)^2 .10 = 0.86$
$\sigma = \sqrt{0.86} = .927$

37. **a.** $10(.35) = 3.5$
b. $P(x = 4) = {}_{10}C_4 (.35)^4 (.65)^6 = 210(.0150)(.0754) = .2375$
c. $P(x \geq 4) = {}_{10}C_x (.35)^x (.65)^{10-x} = .2375 + .1536 + \cdots + .0000 = .4862$

39. **a.** 6, found by 0.4×15
b. 0.0245, found by $\frac{15!}{10!5!}(0.4)^{10}(0.6)^5$
c. 0.0338, found by $0.0245 + 0.0074 + 0.0016 + 0.0003 + 0.0000$
d. 0.0093, found by $0.0338 - 0.0245$

41. **a.** $\mu = 20(0.075) = 1.5$
$\sigma = \sqrt{20(0.075)(0.925)} = 1.1779$
b. 0.2103, found by $\frac{20!}{0!20!}(0.075)^0(0.925)^{20}$
c. 0.7897, found by $1 - 0.2103$

43. **a.** 0.2285, found by $\frac{16!}{3!13!}(0.15)^3(0.85)^{13}$
b. 2.4, found by (0.15)(16)
c. 0.79, found by .0743 + .2097 + .2775 + .2285

45. 0.2784, found by 0.1472 + 0.0811 + 0.0348 + 0.0116 + 0.0030 + 0.0006 + 0.0001 + 0.0000

47. **a.**

0	0.0002	7	0.2075
1	0.0019	8	0.1405
2	0.0116	9	0.0676
3	0.0418	10	0.0220
4	0.1020	11	0.0043
5	0.1768	12	0.0004
6	0.2234		

b. $\mu = 12(0.52) = 6.24$ $\quad \sigma = \sqrt{12(0.52)(0.48)} = 1.7307$
c. 0.1768
d. 0.3343, found by $0.0002 + 0.0019 + 0.0116 + 0.0418 + 0.1020 + 0.1768$

49. **a.** .0183 **b.** .1954
c. .6289 **d.** .5665

51. **a.** 0.1733, found by $\frac{(3.1)^4 e^{-3.1}}{4!}$
b. 0.0450, found by $\frac{(3.1)^0 e^{-3.1}}{0!}$
c. 0.9550, found by $1 - 0.0450$

53. $\mu = n\pi = 23\left(\frac{2}{113}\right) = .407$
$$P(2) = \frac{(.407)^2 e^{-.407}}{2!} = 0.0551$$
$$P(0) = \frac{(.407)^0 e^{-.407}}{0!} = 0.6656$$

55. **a.** Let $\mu = n\pi = 12(2/3) = 8$
b. 0.2384
c. 0.8223

57. **a.** Using the entire binomial probability distribution, with a probability of success equal to 30% and number of trials equal to 40, there is an 80% chance of leasing 10 or more cars. Note that the expected value or number of cars sold with probability of success equal to 30% and trials equal to 40 is: $n\pi = (40)(0.30) = 12$.
b. Of the 40 vehicles that Zook Motors sold only 10, or 25%, were leased. So Zook's probability of success (leasing a car) is 25%. Using .25 as the probability of success, Zook's probability of leasing 10 or more vehicles in 40 trials is only 56%. The data indicate that Zoot leases vehicles at a lower rate than the national average.

59. The mean number of home runs per game is 2.7885. The average season home runs per team is 225.87. Then 225.87/162*2 = 2.7885.
a. $P(x = 0) = \frac{2.7885^0 e^{-2.7885}}{0!} = .0615$
b. $P(x = 2) = \frac{2.7885^2 e^{-2.7885}}{2!} = .2392$
c. $P(X \geq 4) = 0.3055$, found by $1 - P(X < 4) = (.0615 + .1715 + .2392 + .2223) = .6945$

CHAPTER 7

1. **a.** $b = 10, a = 6$ **b.** $\mu = \frac{6 + 10}{2} = 8$
c. $\sigma = \sqrt{\frac{(10 - 6)^2}{12}} = 1.1547$
d. $\text{Area} = \frac{1}{(10 - 6)} \cdot \frac{(10 - 6)}{1} = 1$
e. $P(x > 7) = \frac{1}{(10 - 6)} \cdot \frac{10 - 7}{1} = \frac{3}{4} = .75$
f. $P(7 \leq x \leq 9) = \frac{1}{(10 - 6)} \cdot \frac{(9 - 7)}{1} = \frac{2}{4} = .50$
g. $P(x = 7.91) = 0$
For a continuous probability distribution, the area for a point value is zero.

3. **a.** 0.30, found by (30 − 27)/(30 − 20)
b. 0.40, found by (24 − 20)/(30 − 20)

5. **a.** $a = 0.5, b = 3.00$
b. $\mu = \frac{0.5 + 3.00}{2} = 1.75$
$\sigma = \sqrt{\frac{(3.00 - .50)^2}{12}} = .72$
c. $P(x < 1) = \frac{1}{(3.0 - 0.5)} \cdot \frac{1 - .5}{1} = \frac{.5}{2.5} = 0.2$
d. 0, found by $\frac{1}{(3.0 - 0.5)} \cdot \frac{(1.0 - 1.0)}{1}$
e. $P(x > 1.5) = \frac{1}{(3.0 - 0.5)} \cdot \frac{3.0 - 1.5}{1} = \frac{1.5}{2.5} = 0.6$

7. The actual shape of a normal distribution depends on its mean and standard deviation. Thus, there is a normal distribution, and an accompanying normal curve, for a mean of 7 and a standard deviation of 2. There is another normal curve for a mean of \$25,000 and a standard deviation of \$1,742, and so on.

9. **a.** 490 and 510, found by $500 \pm 1(10)$
b. 480 and 520, found by $500 \pm 2(10)$
c. 470 and 530, found by $500 \pm 3(10)$

11. $$z_{Rob} = \frac{\$70{,}000 - \$80{,}000}{\$5{,}000} = -2$$
$$z_{Rachel} = \frac{\$70{,}000 - \$55{,}000}{\$8{,}000} = 1.875$$

Adjusting for their industries, Rob is well below average and Rachel well above.

13. **a.** 1.25, found by $z = \dfrac{25 - 20}{4.0} = 1.25$

b. 0.3944, found in Appendix B.3

c. 0.3085, found by $z = \dfrac{18 - 20}{2.5} = -0.5$

Find 0.1915 in Appendix B.3 for $z = -0.5$, then 0.5000 − 0.1915 = 0.3085.

15. **a.** 0.2131, found by $z = \dfrac{35.00 - 29.81}{9.31} = 0.56$

Then find 0.2131 in Appendix B.3 for a $z = 0.56$.

b. 0.2869, found by 0.5000 − 0.2131 = 0.2869

c. 0.1469, found by $z = \dfrac{20.00 - 29.81}{9.31} = -1.05$

For a $z = -1.05$, find 0.3531 in Appendix B.3, then 0.5000 − 0.3531 = 0.1469.

17. **a.** 0.8276: First find $z = -1.5$, found by (44 − 50)/4 and $z = 1.25 = (55 - 50)/4$. The area between −1.5 and 0 is 0.4332 and the area between 0 and 1.25 is 0.3944, both from Appendix B.3. Then adding the two areas we find that 0.4332 + 0.3944 = 0.8276.

b. 0.1056, found by 0.5000 − .3944, where $z = 1.25$

c. 0.2029: Recall that the area for $z = 1.25$ is 0.3944, and the area for $z = 0.5$, found by (52 − 50)/4, is 0.1915. Then subtract 0.3944 − 0.1915 and find 0.2029.

19. **a.** 0.1151: Begin by using formula (7–5) to find the z-value for $3,500, which is (3,500 − 2,878)/520, or 1.20. Then see Appendix B.3 to find the area between 0 and 1.20, which is 0.3849.

Finally, since the area of interest is beyond 1.20, subtract that probability from 0.5000. The result is 0.5000 − 0.3849, or 0.1151.

b. 0.0997: Use formula (7–5) to find the z-value for $4,000, which is (4,000 − 2,878)/520, or 2.16. Then see Appendix B.3 for the area under the standard normal curve. That probability is 0.4846. Since the two points (1.20 and 2.16) are on the same side of the mean, subtract the smaller probability from the larger. The result is 0.4846 − 0.3849 = 0.0997.

c. 0.8058: Use formula (7–5) to find the z-value for $2,400, which is −0.92, found by (2,400 − 2,878)/520. The corresponding area is 0.3212. Since −0.92 and 2.16 are on different sides of the mean, add the corresponding probabilities. Thus, we find 0.3212 + 0.4846 = 0.8058.

21. **a.** 0.0764, found by $z = (20 - 15)/3.5 = 1.43$, then 0.5000 − 0.4236 = 0.0764

b. 0.9236, found by 0.5000 + 0.4236, where $z = 1.43$

c. 0.1185, found by $z = (12 - 15)/3.5 = -0.86$.

The area under the curve is 0.3051, then $z = (10 - 15)/3.5 = -1.43$. The area is 0.4236. Finally, 0.4236 − 0.3051 = 0.1185.

23. $x = 56.60$, found by adding 0.5000 (the area left of the mean) and then finding a z-value that forces 45% of the data to fall inside the curve. Solving for x: $1.65 = (x - 50)/4$, so $x = 56.60$.

25. $1,630, found by $2,100 − 1.88($250)

27. **a.** 214.8 hours: Find a z-value where 0.4900 of area is between 0 and z. That value is $z = 2.33$. Then solve for x: $2.33 = (x - 195)/8.5$, so $x = 214.8$ hours.

b. 270.2 hours: Find a z-value where 0.4900 of area is between 0 and ($-z$). That value is $z = -2.33$. Then solve for x: $-2.33 = (x - 290)/8.5$, so $x = 270.2$ hours.

29. 41.7%, found by 12 + 1.65(18)

31. **a.** 0. For a continuous probability distribution, there is no area for a point value.

b. 0. For a continuous probability distribution, there is no area for a point value.

33. **a.** $\mu = \dfrac{11.96 + 12.05}{2} = 12.005$

b. $\sigma = \sqrt{\dfrac{(12.05 - 11.96)^2}{12}} = .0260$

c. $P(x < 12) = \dfrac{1}{(12.05 - 11.96)} \dfrac{12.00 - 11.96}{1} = \dfrac{.04}{.09} = .44$

d. $P(x > 11.98) = \dfrac{1}{(12.05 - 11.96)}\left(\dfrac{12.05 - 11.98}{1}\right)$

$= \dfrac{.07}{.09} = .78$

e. All cans have more than 11.00 ounces, so the probability is 100%.

35. **a.** $\mu = \dfrac{4 + 10}{2} = 7$

b. $\sigma = \sqrt{\dfrac{(10 - 4)^2}{12}} = 1.732$

c. $P(x < 6) = \dfrac{1}{(10 - 4)} \times \left(\dfrac{6 - 4}{1}\right) = \dfrac{2}{6} = .33$

d. $P(x > 5) = \dfrac{1}{(10 - 4)} \times \left(\dfrac{10 - 5}{1}\right) = \dfrac{5}{6} = .83$

37. Based on the friend's information, the probability that the wait time is any value more than 30 minutes is zero. Given the data (wait time was 35 minutes), the friend's information should be rejected. It was false.

39. **a.** 0.4015, z for 900 is: $\dfrac{900 - 1{,}054.5}{120} = -1.29$. Using the z-table, probability is .4015.

b. 0.0985, found by 0.5000 − 0.4015 [0.4015 found in part (a)]

c. 0.7884; z for 900 is: $\dfrac{900 - 1{,}054.5}{120} = -1.29$, z for 1,200 is: $\dfrac{1{,}200 - 1{,}054.5}{120} = 1.21$. Adding the two corresponding probabilities, 0.4015 + 0.3869 = .7884.

d. 0.2279; z for 900 is: $\dfrac{900 - 1{,}054.5}{120} = -1.29$, z for 1,000 is: $\dfrac{1{,}000 - 1{,}054.5}{120} = -0.45$. Subtracting the two corresponding probabilities, 0.4015 − 0.1736 = .2279.

41. **a.** 0.3015, found by 0.5000 − 0.1985

b. 0.2579, found by 0.4564 − 0.1985

c. 0.0011, found by 0.5000 − 0.4989

d. $1,818, found by 1,280 + 1.28(420)

43. **a.** 0.0968, z for 300 is: $\dfrac{300 - 270}{23} = 1.30$. Using the z-table, probability is .4032. Subtracting from 0.5, 0.5000 − 0.4032 = 0.0968.

b. 0.9850, z for 220 is: $\dfrac{220 - 270}{23} = -2.17$. Using the z-table, probability is .4850. Adding 0.5, 0.5000 + 0.4850 = 0.9850.

c. 0.8882; Using the results from parts (a) and (b), the z for 220 is −2.17 with a probability of .4850; the z for 300 is 1.30 with a probability of 0.4032. Adding the two probabilities, (0.4032 + 0.4850) = 0.8882.

d. 307.7; The z-score for the upper 15% of the distribution is 1.64. So the time associated with the upper 15% is 1.64 standard deviations added to the mean, or 270 + 1.64(23) = 307.7 minutes.

45. About 4,099 units, found by solving for x. $1.65 = (x - 4{,}000)/60$

47. **a.** 15.39%, found by (8 − 10.3)/2.25 = −1.02, then 0.5000 − 0.3461 = 0.1539.

b. 17.31%, found by:

$z = (12 - 10.3)/2.25 = 0.76$. Area is 0.2764.

$z = (14 - 10.3)/2.25 = 1.64$. Area is 0.4495.

The area between 12 and 14 is 0.1731, found by 0.4495 − 0.2764.

c. The probability is virtually zero. Applying the Empirical Rule, for 99.73% of the days, returns are between 3.55 and 17.05, found by 10.3 ± 3(2.25). Thus, the chance of less than 3.55 returns is rather remote.

49. **a.** 21.19%, found by $z = (9.00 - 9.20)/0.25 = -0.80$, so $0.5000 - 0.2881 = 0.2119$.

b. Increase the mean. $z = (9.00 - 9.25)/0.25 = -1.00$, $P = 0.5000 - 0.3413 = 0.1587$.

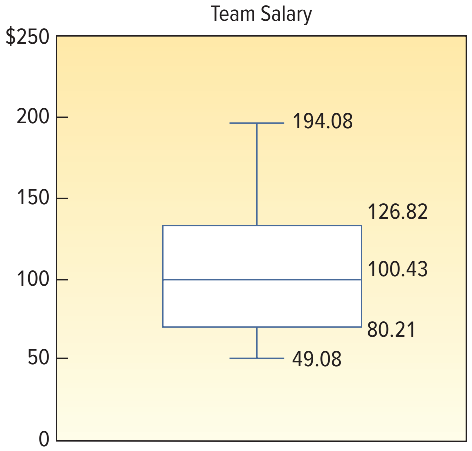
Reduce the standard deviation. $\sigma = (9.00 - 9.20)/0.15 = -1.33$; $P = 0.5000 - 0.4082 = 0.0918$.

Reducing the standard deviation is better because a smaller percent of the hams will be below the limit.

51. The z-score associated with $50,000 is 8.25: (50,000 − 33,500)/2,000. That is, $50,000 is 8.25 standard deviations above the mean salary. Conclusion: The probability that someone in the same business has a salary of $50,000 is zero. This salary would be exceptionally unusual.

53. **a.**

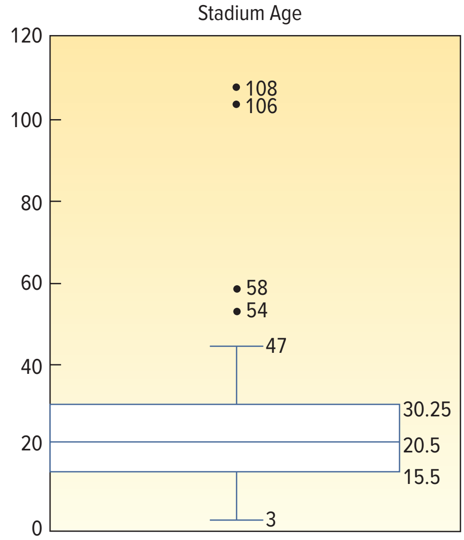
Team Salary	
Mean	105.41
Median	100.43
Standard Deviation	37.36
Sample Variance	1396.04
Skewness	0.64
Range	145.00
Minimum	49.08
Maximum	194.08
Count	30



The distribution of salary is approximately normal. The mean and median are about the same, and skewness is slightly greater than zero. These statistics indicate a slightly positively skewed distribution. The box plot also supports a conclusion that the distribution of salary is positively skewed.

b.

Stadium Age	
Mean	29.37
Median	20.50
Mode	20.00
Standard Deviation	25.12
Sample Variance	631.00
Skewness	2.16
Range	105.00
Minimum	3.00
Maximum	108.00
Count	30



Based on the descriptive statistics and the box plot, stadium age is not normally distributed. The distribution is highly skewed toward the oldest stadiums. See the coefficient of skewness. Also see that the mean and median are very different. The difference is because the mean is affected by the two oldest stadium ages.

CHAPTER 8

1. **a.** 303 Louisiana, 5155 S. Main, 3501 Monroe, 2652 W. Central

b. Answers will vary.

c. 630 Dixie Hwy, 835 S. McCord Rd, 4624 Woodville Rd

d. Answers will vary.

3. **a.** Bob Schmidt Chevrolet
Great Lakes Ford Nissan
Grogan Towne Chrysler
Southside Lincoln Mercury
Rouen Chrysler Jeep Eagle

b. Answers will vary.

c. York Automotive
Thayer Chevrolet/Toyota
Franklin Park Lincoln
Mathews Ford Oregon Inc.
Valiton Chrysler

5. **a.**

Sample	Values	Sum	Mean
1	12, 12	24	12
2	12, 14	26	13
3	12, 16	28	14
4	12, 14	26	13
5	12, 16	28	14
6	14, 16	30	15

b. $\mu_{\bar{x}} = (12 + 13 + 14 + 13 + 14 + 15)/6 = 13.5$

$\mu = (12 + 12 + 14 + 16)/4 = 13.5$

c. More dispersion with population data compared to the sample means. The sample means vary from 12 to 15, whereas the population varies from 12 to 16.

7. a.

Sample	Values	Sum	Mean
1	12, 12, 14	38	12.66
2	12, 12, 15	39	13.00
3	12, 12, 20	44	14.66
4	14, 15, 20	49	16.33
5	12, 14, 15	41	13.66
6	12, 14, 15	41	13.66
7	12, 15, 20	47	15.66
8	12, 15, 20	47	15.66
9	12, 14, 20	46	15.33
10	12, 14, 20	46	15.33

b. $\mu_{\bar{x}} = \dfrac{(12.66 + \cdots + 15.33 + 15.33)}{10} = 14.6$

$\mu = (12 + 12 + 14 + 15 + 20)/5 = 14.6$

c. The dispersion of the population is greater than that of the sample means. The sample means vary from 12.66 to 16.33, whereas the population varies from 12 to 20.

9. a. 20, found by ${}_6C_3$

b.

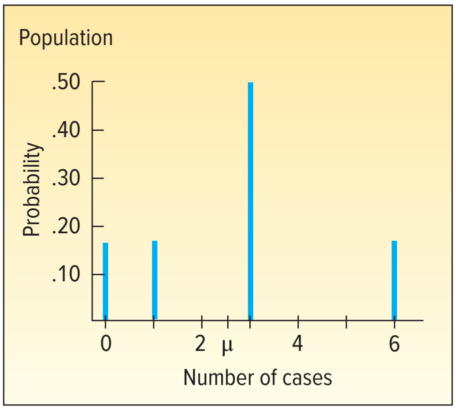
Sample	Cases	Sum	Mean
Ruud, Wu, Sass	3, 6, 3	12	4.00
Ruud, Sass, Flores	3, 3, 3	9	3.00
⋮	⋮	⋮	⋮
Sass, Flores, Schueller	3, 3, 1	7	2.33

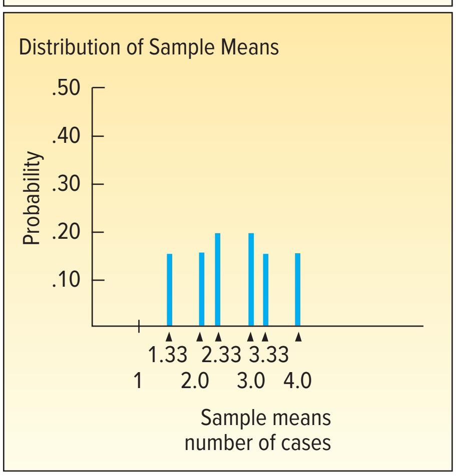
c. $\mu_{\bar{x}} = 2.67$, found by $\dfrac{53.33}{20}$

$\mu = 2.67$, found by $(3 + 6 + 3 + 3 + 0 + 1)/6$.

They are equal.

d.

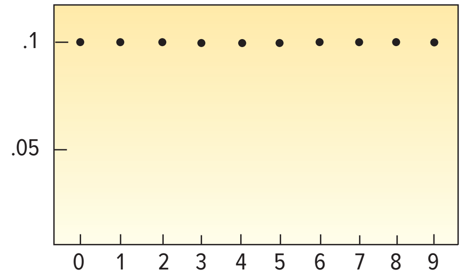




Sample Mean	Number of Means	Probability
1.33	3	.1500
2.00	3	.1500
2.33	4	.2000
3.00	4	.2000
3.33	3	.1500
4.00	3	.1500
	20	1.0000

The population has more dispersion than the sample means. The sample means vary from 1.33 to 4.0. The population varies from 0 to 6.

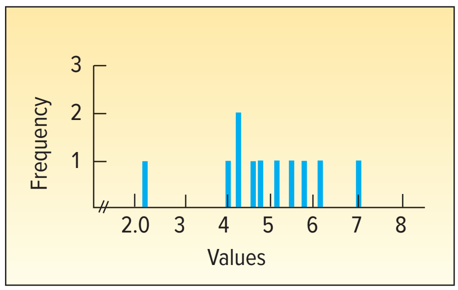
11. a.



$\mu = \dfrac{0 + 1 + \cdots + 9}{10} = 4.5$

b.

Sample	Sum	$\bar{x}$	Sample	Sum	$\bar{x}$
1	11	2.2	6	20	4.0
2	31	6.2	7	23	4.6
3	21	4.2	8	29	5.8
4	24	4.8	9	35	7.0
5	21	4.2	10	27	5.4



The mean of the 10 sample means is 4.84, which is close to the population mean of 4.5. The sample means range from 2.2 to 7.0, whereas the population values range from 0 to 9. From the above graph, the sample means tend to cluster between 4 and 5.

13. a.–c. Answers will vary depending on the coins in your possession.

15. a. The sampling distribution of the sample mean will be normally distributed with the mean equal to the population mean, 60, and the standard error $= (12/\sqrt{9}) = 4$. Applying the central limit theorem requires the population distribution to be normal.

b. 4; Standard Error $= (12/\sqrt{9}) = 4$

c. .2266; $z = \dfrac{63 - 60}{12/\sqrt{9}} = .75$. The probability is $0.5 - .2734 = .2266$

d. .1587; $z = \frac{56 - 60}{12/\sqrt{9}} = -1.00$. The probability is $0.5 - .3413 =$ .1587

e. .6147; Using the information from parts c and d, add the two probabilities,
$.2734 + .3413 = .6147$

f. .0244; $z = \frac{51 - 60}{12/\sqrt{9}} = -2.25$, $z = \frac{69 - 60}{12/\sqrt{9}} = 2.25$; For each tail, $.5000 - .4878 = .0122$. Then multiply by 2. The probability of a sampling error of 9 or more is .0244

17. $z = \frac{1{,}950 - 2{,}200}{250/\sqrt{50}} = -7.07$ $\quad p = 1$, or virtually certain

19. a. Kiehl's, Banana Republic, Cariloha, Nike, and Windsor.
b. Answers may vary.
c. Tilly's, Fabletics, Banana Republic, Madewell, Nike, Guess, Ragstock, Soma

21. a.

Samples	Mean	Deviation from Mean	Square of Deviation
1, 1	1.0	−1.0	1.0
1, 2	1.5	−0.5	0.25
1, 3	2.0	0.0	0.0
2, 1	1.5	−0.5	0.25
2, 2	2.0	0.0	0.0
2, 3	2.5	0.5	0.25
3, 1	2.0	0.0	0.0
3, 2	2.5	0.5	0.25
3, 3	3.0	1.0	1.0

b. Mean of sample means is $(1.0 + 1.5 + 2.0 + \cdots + 3.0)/9 = 18/9 = 2.0$. The population mean is $(1 + 2 + 3)/3 = 6/3 = 2$. They are the same value.

c. Variance of sample means is $(1.0 + 0.25 + 0.0 + \cdots + 3.0)/9 = 3/9 = 1/3$. Variance of the population values is $(1 + 0 + 1)/3 = 2/3$. The variance of the population is twice as large as that of the sample means.

d. Sample means follow a triangular shape peaking at 2. The population is uniform between 1 and 3.

23. Larger samples provide narrower estimates of a population mean. So the company with 200 sampled customers can provide more precise estimates. In addition, they selected consumers who are familiar with laptop computers and may be better able to evaluate the new computer.

25. a. We selected 60, 104, 75, 72, and 48. Answers will vary.
b. We selected the third observation. So the sample consists of 75, 72, 68, 82, 48. Answers will vary.
c. Number the first 20 motels from 00 to 19. Randomly select three numbers. Then number the last five numbers 20 to 24. Randomly select two numbers from that group.

27. a. $(79 + 64 + 84 + 82 + 92 + 77)/6 = 79.67\%$
b. 15, found by ${}_6C_2$
c.

Sample	Value	Sum	Mean
1	79, 64	143	71.5
2	79, 84	163	81.5
⋮	⋮	⋮	⋮
15	92, 77	169	84.5
			1,195.0

d. $\mu_{\bar{x}} = 79.67$, found by 1,195/15.
$\mu = 79.67$, found by 478/6.
They are equal.

e. Answers will vary. Not likely as the student is not graded on all available information. Based on these test scores however, this student has an 8/15 chance of receiving a higher grade with this method than the average and a 7/15 chance of receiving a lower grade.

29. a. 10, found by ${}_5C_2$
b.

Number of Shutdowns	Mean	Number of Shutdowns	Mean
4, 3	3.5	3, 3	3.0
4, 5	4.5	3, 2	2.5
4, 3	3.5	5, 3	4.0
4, 2	3.0	5, 2	3.5
3, 5	4.0	3, 2	2.5

Sample Mean	Frequency	Probability
2.5	2	.20
3.0	2	.20
3.5	3	.30
4.0	2	.20
4.5	1	.10
	10	1.00

c. $\mu_{\bar{x}} = (3.5 + 4.5 + \cdots + 2.5)/10 = 3.4$
$\mu = (4 + 3 + 5 + 3 + 2)/5 = 3.4$
The two means are equal.

d. The population values are relatively uniform in shape. The distribution of sample means tends toward normality.

31. a. The distribution will be normal.
b. $\sigma_{\bar{x}} = \frac{5.5}{\sqrt{25}} = 1.1$
c. $z = \frac{36 - 35}{5.5/\sqrt{25}} = 0.91$
$p = 0.1814$, found by $0.5000 - 0.3186$
d. $z = \frac{34.5 - 35}{5.5/\sqrt{25}} = -0.45$
$p = 0.6736$, found by $0.5000 + 0.1736$
e. 0.4922, found by $0.3186 + 0.1736$
f. Sampling error of more than 1 hour corresponds to times of less than 34 or more than 36 hours. $z = \frac{34 - 35}{5.5/\sqrt{25}} = -0.91$; $z = \frac{36 - 35}{5.5/\sqrt{25}} = 0.91$. Subtracting: $0.5 - .3186 = .1814$ in each tail. Multiplying by 2, the final probability is .3628.

33. $z = \frac{\$335 - \$350}{\$45/\sqrt{40}} = -2.11$
$p = 0.9826$, found by $0.5000 + 0.4826$

35. $z = \frac{29.3 - 29}{2.5/\sqrt{60}} = 0.93$
$p = 0.8238$, found by $0.5000 + 0.3238$

37. Between 5,954 and 6,046, found by $6{,}000 \pm 1.96(150/\sqrt{40})$

39. $z = \frac{900 - 947}{205/\sqrt{60}} = -1.78$
$p = 0.0375$, found by $0.5000 - 0.4625$

41. a. Alaska, Connecticut, Georgia, Kansas, Nebraska, South Carolina, Virginia, Utah
b. Arizona, Florida, Iowa, Massachusetts, Nebraska, North Carolina, Rhode Island, Vermont

43. **a.** $z = \frac{600 - 510}{14.28/\sqrt{10}} = 19.9$, $P = 0.00$, or virtually never

b. $z = \frac{500 - 510}{14.28/\sqrt{10}} = -2.21$,

$p = 0.4864 + 0.5000 = 0.9864$

c. $z = \frac{500 - 510}{14.28/\sqrt{10}} = -2.21$,

$p = 0.5000 - 0.4864 = 0.0136$

45. **a.** $\sigma_{\bar{x}} = \frac{2.1}{\sqrt{81}} = 0.23$

b. $z = \frac{7.0 - 6.5}{2.1/\sqrt{81}} = 2.14$, $z = \frac{6.0 - 6.5}{2.1/\sqrt{81}} = -2.14$,

$p = .4838 + .4838 = .9676$

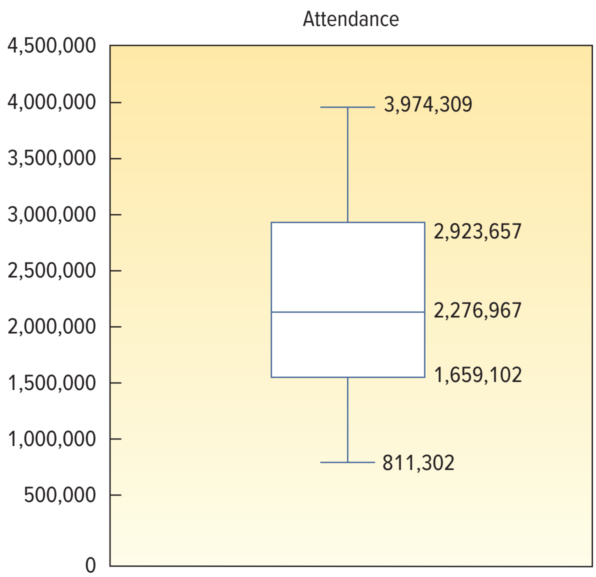
c. $z = \frac{6.75 - 6.5}{2.1/\sqrt{81}} = 1.07$, $z = \frac{6.25 - 6.5}{2.1/\sqrt{81}} = -1.07$,

$p = .3577 + .3577 = .7154$

d. .0162, found by .5000 − .4838

47. Mean 2019 attendance was 2.283 million. Likelihood of a sample mean this large or larger is .1151, found by 0.5000 − .3349, where $z = \frac{2.283 - 2.45}{.762/\sqrt{30}} = -1.20$.

Attendance	
Mean	2,283,163
Median	2,276,967
Standard Deviation	762,164
Skewness	0
Range	3,163,007
Minimum	811,302
Maximum	3,974,309
Count	30.0



CHAPTER 9

1. 51.314 and 58.686, found by $55 \pm 2.58(10/\sqrt{49})$
3. **a.** 1.581, found by $\sigma_{\bar{x}} = 25/\sqrt{250}$
   b. The population is normally distributed and the population variance is known. In addition, the central limit theorem says that the sampling distribution of sample means will be normally distributed.
   c. 16.901 and 23.099, found by 20 ± 3.099
5. **a.** \$20. It is our best estimate of the population mean.
   b. \$18.60 and \$21.40, found by $\$20 \pm 1.96(\$5/\sqrt{49})$. About 95% of the intervals similarly constructed will include the population mean.
7. **a.** 8.60 gallons
   b. 7.83 and 9.37, found by $8.60 \pm 2.58(2.30/\sqrt{60})$
   c. If 100 such intervals were determined, the population mean would be included in about 99 intervals.
9. **a.** 2.201
   b. 1.729
   c. 3.499
11. **a.** The population mean is unknown, but the best estimate is 20, the sample mean.
    b. Use the *t*-distribution since the standard deviation is unknown. However, assume the population is normally distributed.
    c. 2.093
    d. Margin of error $= 2.093(2/\sqrt{20}) = 0.94$
    e. Between 19.06 and 20.94, found by $20 \pm 2.093(2/\sqrt{20})$
    f. Neither value is reasonable because they are not inside the interval.
13. Between 95.39 and 101.81, found by $98.6 \pm 1.833(5.54/\sqrt{10})$
15. **a.** 0.8, found by 80/100
    b. Between 0.72 and 0.88, found by

    $$0.8 \pm 1.96\left(\sqrt{\frac{0.8(1 - 0.8)}{100}}\right)$$

    c. We are reasonably sure the population proportion is between 72% and 88%.
17. **a.** 0.625, found by 250/400
    b. Between 0.563 and 0.687, found by

    $$0.625 \pm 2.58\left(\sqrt{\frac{0.625(1 - 0.625)}{400}}\right)$$

    c. We are reasonably sure the population proportion is between 56% and 69%. Because the estimated population proportion is more than 50%, the results indicate that Netflix should schedule the new romcom series.
19. 97, found by $n = \left(\frac{1.96 \times 10}{2}\right)^2 = 96.04$
21. 196, found by $n = 0.15(0.85)\left(\frac{1.96}{0.05}\right)^2 = 195.9216$
23. 554, found by $n = \left(\frac{1.96 \times 3}{0.25}\right)^2 = 553.19$
25. **a.** 577, found by $n = 0.60(0.40)\left(\frac{1.96}{0.04}\right)^2 = 576.24$
    b. 601, found by $n = 0.50(0.50)\left(\frac{1.96}{0.04}\right)^2 = 600.25$
27. 6.13 years to 6.87 years, found by $6.5 \pm 1.989(1.7/\sqrt{85})$
29. **a.** The sample mean, \$1,147, is the point estimate of the population mean.
    b. The sample standard deviation, \$50, is the point estimate of the population standard deviation.
    c. Margin of error $= 2.426\left(\frac{50}{\sqrt{40}}\right) = 19.18$
    d. Between \$1,127.82 and 1,166.18, found by

    $1{,}147 \pm 2.426\left(\frac{50}{\sqrt{40}}\right)$. \$1,250 is not reasonable because it is outside of the confidence interval.

31. **a.** The population mean is unknown. The point estimate of the population mean is the sample mean, 8.32 years.
b. Between 7.50 and 9.14, found by $8.32 \pm 1.685(3.07/\sqrt{40})$
c. 10 is not reasonable because it is outside the confidence interval.

33. **a.** 65.49 up to 71.71 hours, found by $68.6 \pm 2.680(8.2/\sqrt{50})$
b. The value suggested by the NCAA is included in the confidence interval. Therefore, it is reasonable.
c. Changing the confidence interval to 95 would reduce the width of the interval. The value of 2.680 would change to 2.010.

35. 61.47, rounded to 62. Found by solving for n in the equation: $1.96(16/\sqrt{n}) = 4$

37. **a.** Between 52,461.11 up to 57,640.77 found by

$$55{,}051 \pm 1.711\left(\frac{7{,}568}{\sqrt{25}}\right)$$

b. $55,000 is reasonable because it is inside of the confidence interval.

39. **a.** 82.58, found by 991/12
b. 3.94 is the sample standard deviation.
c. Margin of error $= 1.796\left(\frac{3.94}{\sqrt{12}}\right) = 2.04$
d. Between 80.54 and $84.62, found by

$$82.58 \pm 1.796\left(\frac{3.94}{\sqrt{12}}\right)$$

e. 80 is not reasonable because it is outside of the confidence interval.

41. **a.** 89.467, found by 1,342/15, is the point estimate of the population mean.
b. Between 84.992 and 93.942, found by

$$89.4667 \pm 2.145\left(\frac{8.08}{\sqrt{15}}\right)$$

c. No, the stress level is higher because even the lower limit of the confidence interval is above 80.

43. **a.** 14/400 = .035, or 3.5%, is the point estimate of the population proportion.
b. Margin of error $= 2.576\left(\sqrt{\frac{(0.035)(1-0.035)}{400}}\right) = .024$
c. The confidence interval is between 0.011 and 0.059;

$$0.035 \pm 2.576\left(\sqrt{\frac{(0.035)(1-0.035)}{400}}\right).$$

d. It would be reasonable to conclude that 5% of the employees are failing the test because 0.05, or 5%, is inside the confidence interval.

45. 369, found by $n = 0.60(1 - 0.60)(1.96/0.05)^2$

47. 97, found by $[(1.96 \times 500)/100]^2$

49. **a.** Between 7,849 and 8,151, found by $8{,}000 \pm 2.756(300/\sqrt{30})$
b. 554, found by $n = \left(\frac{(1.96)(300)}{25}\right)^2$

51. **a.** Between 75.44 and 80.56, found by $78 \pm 2.010(9/\sqrt{50})$
b. 220, found by $n = \left(\frac{(1.645)(9)}{1.0}\right)^2$

53. **a.** The point estimate of the population mean is the sample mean, $650.
b. The point estimate of the population standard deviation is the sample standard deviation, $24.
c. 4, found by $24/\sqrt{36}$
d. Between $641.88 and $658.12, found by

$$650 \pm 2.030\left(\frac{24}{\sqrt{36}}\right)$$

e. 23, found by $n = \{(1.96 \times 24)/10\}^2 = 22.13$

55. **a.** 708.13, rounded up to 709, found by $0.21(1 - 0.21)(1.96/0.03)^2$
b. 1,068, found by $0.50(0.50)(1.96/0.03)^2$

57. **a.** Between 0.156 and 0.184, found by

$$0.17 \pm 1.96\sqrt{\frac{(0.17)(1-0.17)}{2{,}700}}$$

b. Yes, because 18% are inside the confidence interval.
c. 21,682; found by $0.17(1 - 0.17)[1.96/0.005]^2$

59. Between 12.69 and 14.11, found by $13.4 \pm 1.96(6.8/\sqrt{352})$

61. **a.** Answers will vary.
b. Answers will vary.
c. Answers will vary.
d. Answers may vary.
e. Select a different sample of 20 homes and compute a confidence interval using the new sample. There is a 5% probability that a sample mean will be more than 1.96 standard errors from the mean. If this happens, the confidence interval will not include the population mean.

63. **a.** Between $4,033.1476 and $5,070.6274, found by $4{,}551.8875 \pm 518.7399$.
b. Between 71,040.0894 and 84,877.1106, found by $77{,}958.6000 \pm 6{,}918.5106$.
c. In general, the confidence intervals indicate that the average maintenance cost and the average odometer reading suggest an aging bus fleet.

CHAPTER 10

1. **a.** Two-tailed
b. Reject H_0 when z does not fall in the region between −1.96 and 1.96.
c. −1.2, found by $z = (49 - 50)/(5/\sqrt{36}) = -1.2$
d. Fail to reject H_0.
e. Using the z-table, the p-value is .2302, found by 2(.5000 − .3849). A 23.02% chance of finding a z-value this large when H_0 is true.

3. **a.** One-tailed
b. Reject H_0 when $z > 1.65$.
c. 1.2, found by $z = (21 - 20)/(5/\sqrt{36})$
d. Fail to reject H_0 at the .05 significance level.
e. Using the z-table, the p-value is .1151, found by .5000 − .3849. An 11.51% chance of finding a z-value this large or larger.

5. **a.** H_0: $\mu = 60{,}000$ H_1: $\mu \neq 60{,}000$
b. Reject H_0 if $z < -1.96$ or $z > 1.96$.
c. −0.69, found by:

$$z = \frac{59{,}500 - 60{,}000}{(5{,}000/\sqrt{48})}$$

d. Do not reject H_0.
e. Using the z-table, the p-value is .4902, found by 2(.5000 − .2549). Crosset's experience is not different from that claimed by the manufacturer. If H_0 is true, the probability of finding a value more extreme than this is .4902.

7. **a.** H_0: $\mu \geq 6.8$ H_1: $\mu < 6.8$
b. Reject H_0 if $z < -1.65$.
c. $z = \frac{6.2 - 6.8}{1.8/\sqrt{36}} = -2.0$
d. H_0 is rejected.
e. Using the z-table, the p-value is 0.0228. The mean number of DVDs watched is less than 6.8 per month. If H_0 is true, you will get a statistic this small less than one time out of 40 tests.

9. **a.** Reject H_0 when $t < 1.833$.
b. $t = \frac{12 - 10}{(3/\sqrt{10})} = 2.108$
c. Reject H_0. The mean is greater than 10.

11. H_0: $\mu \leq 40$ H_1: $\mu > 40$
Reject H_0 if $t > 1.703$.

$$t = \frac{42 - 40}{(2.1/\sqrt{28})} = 5.040$$

Reject H_0 and conclude that the mean number of calls is greater than 40 per week.

13. H_0: $\mu \leq 60{,}000$ H_1: $\mu > 60{,}000$
Reject H_0 if $t > 1.833$.

$$t = \frac{(70{,}000 - 60{,}000)}{(10{,}000/\sqrt{10})} = 3.16$$

Reject H_0 and conclude that the mean income in Wilmington is greater than $60,000.

15. **a.** Reject H_0 if $t < -3.747$.

b. $\bar{x} = 17$ and $s = \sqrt{\dfrac{50}{5-1}} = 3.536$

$$t = \frac{17 - 20}{(3.536/\sqrt{5})} = -1.90$$

c. Do not reject H_0. We cannot conclude the population mean is less than 20.

d. Using a *p*-value calculator or statistical software, the *p*-value is .0653.

17. H_0: $\mu \leq 1.4$ H_1: $\mu > 1.4$
Reject H_0 if $t > 2.821$.

$$t = \frac{1.6 - 1.4}{0.216/\sqrt{10}} = 2.93$$

Reject H_0 and conclude that the water consumption has increased. Using a *p*-value calculator or statistical software, the *p*-value is .0084. There is a slight probability that the sampling error, .2 liters, could occur by chance.

19. H_0: $\mu \leq 67$ H_1: $\mu > 67$
Reject H_0 if $t > 1.796$.

$$t = \frac{(82.5 - 67)}{(59.5/\sqrt{12})} = 0.902$$

Fail to reject H_0 and conclude that the mean number of text messages is not greater than 67. Using a *p*-value calculator or statistical software, the *p*-value is .1932. There is a good probability (about 19%) this could happen by chance.

21. H_0: $\mu = \$45{,}000$ H_1: $\mu \neq \$45{,}000$
Reject H_0 if $z < -1.65$ or $z > 1.65$.

$$z = \frac{\$45{,}500 - \$45{,}000}{\$3000/\sqrt{120}} = 1.83$$

Using the *z*-table, the *p*-value is 0.0672, found by 2(0.5000 − 0.4664).
Reject H_0. We can conclude that the mean salary is not $45,000.

23. H_0: $\mu \geq 10$ H_1: $\mu < 10$
Reject H_0 if $z < -1.65$.

$$z = \frac{9.0 - 10.0}{2.8/\sqrt{50}} = -2.53$$

Using the *z*-table, *p*-value = 0.5000 − 0.4943 = 0.0057.
Reject H_0. The mean weight loss is less than 10 pounds.

25. H_0: $\mu \geq 7.0$ H_1: $\mu < 7.0$
Assuming a 5% significance level, reject H_0 if $t < -1.677$.

$$t = \frac{6.8 - 7.0}{0.9/\sqrt{50}} = -1.57$$

Using a *p*-value calculator or statistical software, the *p*-value is 0.0614.
Do not reject H_0. West Virginia students are not sleeping less than 6 hours.

27. H_0: $\mu \geq 3.13$ H_1: $\mu < 3.13$
Reject H_0 if $t < -1.711$.

$$t = \frac{2.86 - 3.13}{1.20/\sqrt{25}} = -1.13$$

We fail to reject H_0 and conclude that the mean number of residents is not necessarily less than 3.13.

29. H_0: $\mu \leq \$6{,}658$ H_1: $\mu > \$6{,}658$
Reject H_0 if $t > 1.796$.

$$\bar{x} = \frac{85{,}963}{12} = 7{,}163.58 \quad s = \sqrt{\frac{9{,}768{,}674.92}{12-1}} = 942.37$$

$$t = \frac{7163.58 - 6{,}658}{942.37/\sqrt{12}} = 1.858$$

Reject H_0. First, the test statistic (1.858) is more than the critical value, 1.796. Second, using a *p*-value calculator or statistical software, the *p*-value is .0451 and less than the significance level, .05. We conclude that the mean interest paid is greater than $6,658.

31. H_0: $\mu = 3.1$ H_1: $\mu \neq 3.1$ Assume a normal population.
Reject H_0 if $t < -2.201$ or $t > 2.201$.

$$\bar{x} = \frac{41.1}{12} = 3.425$$

$$s = \sqrt{\frac{4.0625}{12-1}} = .6077$$

$$t = \frac{3.425 - 3.1}{.6077/\sqrt{12}} = 1.853$$

Using a *p*-value calculator or statistical software, the *p*-value is .0910.
Do not reject H_0. Cannot show a difference between senior citizens and the national average.

33. H_0: $\mu \geq 6.5$ H_1: $\mu < 6.5$ Assume a normal population.
Reject H_0 if $t < -2.718$.
$\bar{x} = 5.1667$ $s = 3.1575$

$$t = \frac{5.1667 - 6.5}{3.1575/\sqrt{12}} = -1.463$$

Using a *p*-value calculator or statistical software, the *p*-value is .0861.
Do not reject H_0.

35. H_0: $\mu = 0$ H_1: $\mu \neq 0$
Reject H_0 if $t < -2.110$ or $t > 2.110$.
$\bar{x} = -0.2322$ $s = 0.3120$

$$t = \frac{-0.2322 - 0}{0.3120/\sqrt{18}} = -3.158$$

Using a *p*-value calculator or statistical software, the *p*-value is .0057.
Reject H_0. The mean gain or loss does not equal 0.

37. H_0: $\mu \leq 100$ H_1: $\mu > 100$ Assume a normal population.
Reject H_0 if $t > 1.761$.

$$\bar{x} = \frac{1{,}641}{15} = 109.4$$

$$s = \sqrt{\frac{1{,}389.6}{15-1}} = 9.9628$$

$$t = \frac{109.4 - 100}{9.9628/\sqrt{15}} = 3.654$$

Using a *p*-value calculator or statistical software, the *p*-value is .0013.
Reject H_0. The mean number with the scanner is greater than 100.

39. H_0: $\mu = 1.5$ H_1: $\mu \neq 1.5$
Reject H_0 if $t > 3.250$ or $t < -3.250$.

$$t = \frac{1.3 - 1.5}{0.9/\sqrt{10}} = -0.703$$

Using a *p*-value calculator or statistical software, the *p*-value is .4998.
Fail to reject H_0.

41. H_0: $\mu \geq 30$ H_1: $\mu < 30$
Reject H_0 if $t < -1.895$.

$$\bar{x} = \frac{238.3}{8} = 29.7875 \quad s = \sqrt{\frac{5.889}{8-1}} = 0.9172$$

$$t = \frac{29.7875 - 30}{0.9172/\sqrt{8}} = -0.655$$

Using a p-value calculator or statistical software, the p-value is .2667.
Do not reject H_0. The cost is not less than \$30,000.

43. H_0: $\mu \geq 8$ H_1: $\mu < 8$
Reject H_0 if $t < -1.714$.

$$t = \frac{7.5 - 8}{3.2/\sqrt{24}} = -0.77$$

Using a p-value calculator or statistical software, the p-value is .2246.
Do not reject the null hypothesis. The time is not less.

45. **a.** H_0: $\mu = 100$ H_1: $\mu \neq 100$
Reject H_0 if t is not between -2.045 and 2.045.

$$t = \frac{105.41 - 100}{37.4/\sqrt{30}} = 0.792$$

Using a p-value calculator or statistical software, the p-value is .4348.
Fail to reject the null. The data do not provide enough evidence to reject the null hypothesis.

b. H_0: $\mu \leq 2{,}000{,}000$ H_1: $\mu > 2{,}000{,}000$
Reject H_0 if t is > 1.699.

$$t = \frac{2.283 - 2.0}{.762/\sqrt{30}} = 2.03$$

Using a p-value calculator or statistical software, p-value is 0.0258.
Reject the null. The mean attendance was more than 2,000,000.

CHAPTER 11

1. **a.** Two-tailed test
b. Reject H_0 if $z < -2.05$ or $z > 2.05$.
c. $$z = \frac{102 - 99}{\sqrt{\frac{5^2}{40} + \frac{6^2}{50}}} = 2.59$$
d. Reject H_0.
e. Using the z-table, the p-value is = .0096, found by 2(.5000 − .4952).

3. **Step 1** H_0: $\mu_1 \geq \mu_2$ H_1: $\mu_1 < \mu_2$
Step 2 The .05 significance level was chosen.
Step 3 Reject H_0 if $z < -1.65$.
Step 4 −0.94, found by:

$$z = \frac{7.6 - 8.1}{\sqrt{\frac{(2.3)^2}{40} + \frac{(2.9)^2}{55}}} = -0.94$$

Step 5 Fail to reject H_0.
Step 6 Babies using the Gibbs brand did not gain less weight. Using the z-table, the p-value is = .1736, found by .5000 − .3264.

5. **Step 1** H_0: $\mu_{\text{married}} = \mu_{\text{unmarried}}$ H_1: $\mu_{\text{married}} = \mu_{\text{unmarried}}$
Step 2 The 0.05 significance level was chosen.
Step 3 Use a z-statistic as both population standard deviations are known.
Step 4 If $z < -1.960$ or $z > 1.960$, reject H_0.
Step 5 $$z = \frac{4.0 - 4.4}{\sqrt{\frac{(1.2)^2}{45} + \frac{(1.1)^2}{39}}} = -1.59$$

Fail to reject the null.
Step 6 It is reasonable to conclude that the time that married and unmarried women spend each week is not significantly different. Using the z-table, the p-value is .1142. The difference of 0.4 hour per week could be explained by sampling error.

7. **a.** Reject H_0 if $t > 2.120$ or $t < -2.120$. $df = 10 + 8 - 2 = 16$.
b. $$s_p^2 = \frac{(10-1)(4)^2 + (8-1)(5)^2}{10 + 8 - 2} = 19.9375$$
c. $$t = \frac{23 - 26}{\sqrt{19.9375\left(\frac{1}{10} + \frac{1}{8}\right)}} = -1.416$$
d. Do not reject H_0.
e. Using a p-value calculator or statistical software, the p-value is .1759. From the t-table we estimate the p-value is greater than 0.10 and less than 0.20.

9. **Step 1** H_0: $\mu_{\text{Pitchers}} = \mu_{\text{Position Players}}$
H_1: $\mu_{\text{Pitchers}} = \mu_{\text{Position Players}}$
Step 2 The 0.01 significance level was chosen.
Step 3 Use a t-statistic assuming a pooled variance with the standard deviation unknown.
Step 4 $df = 15 + 14 - 2 = 27$
Reject H_0 if t is not between −2.771 and 2.770.

$$s_p^2 = \frac{(15-1)(10.089)^2 + (14-1)(6.727)^2}{15 + 14 - 2} = 74.570$$

$$t = \frac{7.859 - 6.175}{\sqrt{74.570\left(\frac{1}{15} + \frac{1}{14}\right)}} = .5250$$

Using a p-value calculator or statistical software, the p-value is 0.6039.
Step 5 Do not reject H_0.
Step 6 There is no difference in the mean salaries of pitchers and position players.

11. **Step 1** H_0: $\mu_s \leq \mu_a$ H_1: $\mu_s > \mu_a$
Step 2 The .10 significance level was chosen.
Step 3 $df = 6 + 7 - 2 = 11$
Reject H_0 if $t > 1.363$.
Step 4 $$s_p^2 = \frac{(6-1)(12.2)^2 + (7-1)(15.8)^2}{6 + 7 - 2} = 203.82$$

$$t = \frac{142.5 - 130.3}{\sqrt{203.82\left(\frac{1}{6} + \frac{1}{7}\right)}} = 1.536$$

Step 5 Using a p-value calculator or statistical software, the p-value is 0.0763. Reject H_0.
Step 6 The mean daily expenses are greater for the sales staff.

13. **a.** $$df = \frac{\left(\frac{25}{15} + \frac{225}{12}\right)^2}{\frac{\left(\frac{25}{15}\right)^2}{15-1} + \frac{\left(\frac{225}{12}\right)^2}{12-1}} = \frac{416.84}{0.1984 + 31.9602}$$
$= 12.96 \rightarrow 12df$
b. H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$
Reject H_0 if $t > 2.179$ or $t < -2.179$.
c. $$t = \frac{50 - 46}{\sqrt{\frac{25}{15} + \frac{225}{12}}} = 0.8852$$
d. Fail to reject the null hypothesis.

15. a. $df = \dfrac{\left(\dfrac{697{,}225}{16} + \dfrac{2{,}387{,}025}{18}\right)^2}{\dfrac{\left(\dfrac{697{,}225}{16}\right)^2}{16-1} + \dfrac{\left(\dfrac{2{,}387{,}025}{18}\right)^2}{18-1}} = 26.7 \rightarrow 26df$

b. H_0: $\mu_{\text{Private}} \leq \mu_{\text{Public}}$ H_1: $\mu_{\text{Private}} > \mu_{\text{Public}}$
Reject H_0 if $t > 1.706$.

c. $t = \dfrac{12{,}840 - 11{,}045}{\sqrt{\dfrac{2{,}387{,}025}{18} + \dfrac{697{,}225}{16}}} = 4.276$

d. Reject the null hypothesis. The mean adoption cost from a private agency is greater than the mean adoption cost from a public agency.

17. a. $\bar{d} = \dfrac{12}{4} = 3.00$

$s_d = \sqrt{\dfrac{(2-3)^2 + (3-3)^2 + (3-3)^2 + (4-3)^2}{4-1}} = 0.816$

$t = \dfrac{3}{0.816/\sqrt{4}} = 7.353$

Using a *p*-value calculator or statistical software, the *p*-value is .0026.

b. Reject the H_0. The test statistic is greater than the critical value. The *p*-value is less than .05.

c. There are more defective parts produced on the day shift.

19. a. **Step 1:** H_0: $\mu_d \geq 0$ H_1: $\mu_d < 0$
Step 2: The 0.05 significance level was chosen.
Step 3: Use a *t*-statistic with the standard deviation unknown-for a paired sample.
Step 4: Reject H_0 if $t < -1.796$.

b. **Step 5:** $\bar{d} = -25.917$

$s_d = 40.791 \quad t = \dfrac{-25.917}{40.791/\sqrt{12}} = -2.201$

Using a *p*-value calculator or statistical software, the *p*-value is .0250.

c. Reject H_0. The test statistic is greater than the critical value. The *p*-value is less than .05.

d. **Step 6:** The incentive plan resulted in an increase in daily income.

21. a. $H_0 = \mu_{\text{Men}} = \mu_{\text{Women}}$ H_1: $\mu_{\text{Men}} = \mu_{\text{Women}}$
Reject H_0 if $t < -2.645$ or $t > 2.645$.

b. $s_p^2 = \dfrac{(35-1)(4.48)^2 + (40-1)(3.86)^2}{35+40-2} = 17.31$

$t = \dfrac{24.51 - 22.69}{\sqrt{17.31\left(\dfrac{1}{35} + \dfrac{1}{40}\right)}} = 1.890$

c. Using a *p*-value calculator or statistical software, the *p*-value is .0627.

d. Do not reject the null hypothesis. The test statistic is less than the critical value. The *p*-value is more than .01.

e. There is no difference in the means.

23. a. H_0: $\mu_{\text{Clark}} = \mu_{\text{Murnen}}$; H_1: $\mu_{\text{Clark}} \neq \mu_{\text{Murnen}}$
Reject H_0 if $z < -1.96$ or $z > 1.96$.

b. $z = \dfrac{4.77 - 5.02}{\sqrt{\dfrac{(1.05)^2}{40} + \dfrac{(1.23)^2}{50}}} = -1.04$

c. Using a *z*-table or a *p*-value calculator or statistical software, the *p*-value is .2983.

d. H_0 is not rejected. The test statistic is less than the critical value. The *p*-value is more than .05.

e. There is no difference in the mean number of calls.

25. a. H_0: $\mu_A \geq \mu_B$ H_1: $\mu_A < \mu_B$
Reject H_0 if $t < -1.668$.

b. $df = 67$, found by $\dfrac{\left({}^{9200^2}/_{40} + {}^{7100^2}/_{30}\right)^2}{\dfrac{\left({}^{9200^2}/_{40}\right)^2}{39} + \dfrac{\left({}^{7100^2}/_{30}\right)^2}{29}} = 67.9$

$t = \dfrac{57000 - 61000}{\sqrt{\dfrac{9200^2}{40} + \dfrac{7100^2}{30}}} = -2.053$

c. Using a *p*-value calculator or statistical software, the *p*-value is .0220.

d. Reject H_0. The test statistic is less than the critical value. Reject H_0 if $t < -1.668$. The *p*-value is less than .05.

e. The mean income of those selecting Plan B is larger.

27. a. H_0: $\mu_{\text{Apple}} = \mu_{\text{Spotify}}$ H_1: $\mu_{\text{Apple}} \neq \mu_{\text{Spotify}}$;
Reject H_0 if $t < -2.120$ or $t > 2.120$.

b. $df = 16$, found by $\dfrac{\left({}^{0.56^2}/_{12} + {}^{0.3^2}/_{12}\right)^2}{\dfrac{\left({}^{0.56^2}/_{12}\right)^2}{11} + \dfrac{\left({}^{0.3^2}/_{12}\right)^2}{11}} = 16.8$

$t = \dfrac{1.65 - 2.2}{\sqrt{\dfrac{0.56^2}{12} + \dfrac{0.3^2}{12}}} = -2.999$

c. Using a *p*-value calculator or statistical software, the *p*-value is .0085.

d. Reject H_0. The test statistic is outside the interval. The *p*-value is less than .05.

e. The number of average monthly households using Apple Music and Spotify differ.

29. a. H_0: $\mu_n = \mu_s$ H_1: $\mu_n \neq \mu_s$;
Reject H_0 if $t < -2.093$ or $t > 2.093$.

b. $df = 19$, found by $\dfrac{\left({}^{10.5^2}/_{10} + {}^{14.25^2}/_{12}\right)^2}{\dfrac{\left({}^{10.5^2}/_{10}\right)^2}{9} + \dfrac{\left({}^{14.25^2}/_{12}\right)^2}{11}} = 19.8$

$t = \dfrac{83.55 - 78.8}{\sqrt{\dfrac{10.5^2}{10} + \dfrac{14.25^2}{12}}} = 0.899$

c. Using a *p*-value calculator or statistical software, the *p*-value is .3799.

d. Do not reject H_0. The test statistic is inside the interval. The *p*-value is large and greater than .05.

e. There is no difference in the mean number of hamburgers sold at the two locations.

31. a. H_0: $\mu_{\text{Peach}} = \mu_{\text{Plum}}$ H_1: $\mu_{\text{Peach}} \neq \mu_{\text{Plum}}$
Reject H_0 if $t < -2.845$ or $t > 2.845$.

b. $df = 20$, found by $\dfrac{\left({}^{2.33^2}/_{10} + {}^{2.55^2}/_{14}\right)^2}{\dfrac{\left({}^{2.33^2}/_{10}\right)^2}{9} + \dfrac{\left({}^{2.55^2}/_{14}\right)^2}{13}} = 20.6$

$t = \dfrac{15.87 - 18.29}{\sqrt{\dfrac{2.33^2}{10} + \dfrac{2.55^2}{14}}} = -2.411$

c. Using a *p*-value calculator or statistical software, the *p*-value is .0256.

d. Do not reject H_0. The test statistic is inside the interval. The *p*-value is more than .01.

e. There is no difference in the mean amount purchased at the 1% level of significance.

33. a. H_0: $\mu_{\text{Under 25}} \leq \mu_{\text{Over 65}}$ H_1: $\mu_{\text{Under 25}} > \mu_{\text{Over 65}}$;
Reject H_0 if $t > 2.602$.

b. $df = 15$, found by $\dfrac{\left({}^{2.264^2}/_{8} + {}^{2.461^2}/_{11}\right)^2}{\dfrac{\left({}^{2.264^2}/_{8}\right)^2}{7} + \dfrac{\left({}^{2.461^2}/_{11}\right)^2}{10}} = 15.953$

$t = \dfrac{10.375 - 5.636}{\sqrt{\dfrac{2.264^2}{8} + \dfrac{2.461^2}{11}}} = 4.342$

c. Using a p-value calculator or statistical software, the p-value is .0003.
d. Reject H_0. The test statistic is greater than the critical value. The p-value is less than .01.
e. Customers who are under 25 years of age use ATMs more than customers who are over 60 years of age.

35. **a.** H_0: $\mu_{\text{Reduced}} \le \mu_{\text{Regular}}$ $\quad H_1$: $\mu_{\text{Reduced}} > \mu_{\text{Regular}}$
Reject H_0 if $t > 2.650$.
b. $\bar{X}_1 = 125.125 \quad s_1 = 15.094 \quad \bar{X}_2 = 117.714 \quad s_2 = 19.914$

$$s_p^2 = \frac{(8-1)(15.094)^2 + (7-1)(19.914)^2}{8+7-2} = 305.708$$

$$t = \frac{125.125 - 117.714}{\sqrt{305.708\left(\frac{1}{8}+\frac{1}{7}\right)}} = 0.819$$

c. Using a p-value calculator or statistical software, the p-value is .2133.
d. Do not reject H_0. The test statistic is inside the interval. The p-value is more than .01.
e. The sample data do not provide evidence that the reduced price increased sales.

37. **a.** H_0: $\mu_{\text{Before}} - \mu_{\text{After}} = \mu_d \le 0$ $\quad H_1$: $\mu_d > 0$
Reject H_0 if $t > 1.895$.
b. $\bar{d} = 1.75 \quad s_d = 2.9155 \quad t = \dfrac{1.75}{2.9155/\sqrt{8}} = 1.698$
c. Using a p-value calculator or statistical software, the p-value is .0667.
d. Do not reject H_0. The test statistic is less than the critical value. The p-value is greater than .05.
e. We fail to find evidence the change reduced absences.

39. **a.** H_0: $\mu_1 = \mu_2$ $\quad H_1$: $\mu_1 \ne \mu_2$
Reject H_0 if $t < -2.024$ or $t > 2.024$.
b.

$$s_p^2 = \frac{(15-1)(40{,}000)^2 + (25-1)(30{,}000)^2}{15+25-2} = 1{,}157{,}894{,}737$$

$$t = \frac{150{,}000 - 180{,}000}{\sqrt{1{,}157{,}894{,}737\left(\frac{1}{15}+\frac{1}{25}\right)}} = -2.699$$

c. Using a p-value calculator or statistical software, the p-value is .0103.
d. Reject H_0. The test statistic is outside the interval. The p-value is less than .05.
e. The data indicate that the population means are different.

41. **a.** H_0: $\mu_{\text{Before}} - \mu_{\text{After}} = \mu_d \le 0$ $\quad H_1$: $\mu_d > 0$
Reject H_0 if $t > 1.895$.
b. $\bar{d} = 3.113 \quad s_d = 2.911 \quad t = \dfrac{3.113}{2.911/\sqrt{8}} = 3.025$
c. Using a p-value calculator or statistical software, the p-value is .0096.
d. Reject H_0. The test statistic is outside the interval. The p-value is less than .05.
e. We find evidence the average contamination is lower after the new soap is used.

43. **a.** H_0: $\mu_{\text{Ocean Drive}} = \mu_{\text{Rio Rancho}}$; $\quad H_1$: $\mu_{\text{Ocean Drive}} \ne \mu_{\text{Rio Rancho}}$
Reject H_0 if $t < -2.008$ or $t > 2.008$.
b.

$$s_p^2 = \frac{(25-1)(23.43)^2 + (28-1)(24.12)^2}{25+28-2} = 566$$

$$t = \frac{86.2 - 92.0}{\sqrt{566\left(\frac{1}{25}+\frac{1}{28}\right)}} = -0.886$$

c. Using a p-value calculator or statistical software, the p-value is .3798.
d. Do not reject H_0. The test statistic is inside the interval. The p-value is more than .05.
e. It is reasonable to conclude there is no difference in the mean number of cars in the two lots.

45. **a.** H_0: $\mu_{\text{US 17}} - \mu_{\text{SC 707}} = \mu_d \le 0$ $\quad H_1$: $\mu_d > 0$
Reject H_0 if $t > 1.711$.
b. $\bar{d} = 2.8 \quad s_d = 6.589 \quad t = \dfrac{2.8}{6.589/\sqrt{25}} = 2.125$
c. Using a p-value calculator or statistical software, the p-value is .0220.
d. Reject H_0. The test statistic is greater than the test statistic. The p-value is less than .05.
e. On average, there are more cars in the US 17 lot.

47. **a.** Using statistical software, the result is that we fail to reject the null hypothesis that the mean prices of homes with and without pools are equal. Assuming equal population variances, the p-value is 0.4908.
b. Using statistical software, the result is that we reject the null hypothesis that the mean prices of homes with and without garages are equal. There is a large difference in mean prices between homes with and without garages. Assuming equal population variances, the p-value is less than 0.0001.
c. Using statistical software, the result is that we fail to reject the null hypothesis that the mean prices of homes are equal with mortgages in default and not in default. Assuming equal population variances, the p-value is 0.6980.

49. Using statistical software, the result is that we reject the null hypothesis that the mean maintenance cost of buses powered by diesel and gasoline engines is the same. Assuming equal population variances, the p-value is less than 0.0001.

CHAPTER 12

1. 9.01, from Appendix B.6

3. Reject H_0 if $F > 10.5$, where degrees of freedom in the numerator are 7 and 5 in the denominator. Computed $F = 2.04$, found by:

$$F = \frac{s_1^2}{s_2^2} = \frac{(10)^2}{(7)^2} = 2.04$$

Do not reject H_0. There is no difference in the variations of the two populations.

5. **a.** $H_0{:}\sigma_1^2 = \sigma_2^2$ $\quad H_1{:}\sigma_1^2 \ne \sigma_2^2$
b. df in numerator are 11 and 9 in the denominator.
Reject H_0 where $F > 3.10$ (3.10 is about halfway between 3.14 and 3.07).
c. $F = 1.44$, found by $F = \dfrac{(12)^2}{(10)^2} = 1.44$
d. Using a p-value calculator or statistical software, the p-value is .2964.
e. Do not reject H_0.
f. It is reasonable to conclude variations of the two populations could be the same.

7. **a.** H_0: $\mu_1 = \mu_2 = \mu_3$; H_1: Treatment means are not all the same.
b. Reject H_0 if $F > 4.26$.
c & d.

Source	SS	*df*	MS	*F*
Treatment	62.17	2	31.08	21.94
Error	12.75	9	1.42	
Total	74.92	11		

e. Reject H_0. The treatment means are not all the same.

9. **a.** H_0: $\mu_{\text{Southwyck}} = \mu_{\text{Franklin}} = \mu_{\text{Old Orchard}}$ $\quad H_1$: Treatment means are not all the same.
b. Reject H_0 if $F > 4.26$.
c.

Source	SS	*df*	MS	*F*
Treatment	276.50	2	138.25	14.18
Error	87.75	9	9.75	

d. Using a p-value calculator or statistical software, the p-value is .0017.
e. Reject H_0. The test statistic is greater than the critical value. The p-value is less than .05.
f. The mean incomes are not all the same for the three tracts of land.

11. **a.** $H_0: \mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same.
b. Reject H_0 if $F > 4.26$.
c. SST = 107.20 SSE = 9.47 SS total = 116.67
d. Using Excel,

ANOVA						
Source of Variation	SS	*df*	MS	*F*	*p*-Value	*F* crit
Treatment	107.2000	2	53.6000	50.9577	0.0000	4.2565
Error	9.4667	9	1.0519			
Total	116.6667	11				

e. Since 50.96 > 4.26, H_0 is rejected. At least one of the means differ.
f. $(\bar{X}_1 - \bar{X}_2) \pm t\sqrt{MSE(1/n_1 + 1/n_2)}$
$(9.667 - 2.20) \pm 2.262\sqrt{1.052(1/3 + 1/5)}$
7.467 ± 1.69
[5.777, 9.157] Yes, we can conclude that treatments 1 and 2 have different means.

13. **a.** $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: Treatment means are not all equal. Reject H_0 if $F > 3.71$.
b. The *F*-test statistic is 2.36.
c. The *p*-value is .133.
d. H_0 is not rejected. The test statistic, 2.36, is less than the critical value, 3.71. The *p*-value is more than .05.
e. There is no difference in the mean number of weeks.

15. $H_0: \sigma_1^2 \le \sigma_2^2; H_1: \sigma_1^2 > \sigma_2^2$. $df_1 = 21 - 1 = 20$; $df_2 = 18 - 1 = 17$. H_0 is rejected if $F > 3.16$.
$$F = \frac{(45{,}600)^2}{(21{,}330)^2} = 4.57$$
Reject H_0. There is more variation in the selling price of ocean-front homes.

17. Sharkey: $n = 7$ $s_s = 14.79$
White: $n = 8$ $s_w = 22.95$
$H_0: \sigma_w^2 \le \sigma_s^2$; $H_1: \sigma_w^2 > \sigma_s^2$. $df_s = 7 - 1 = 6$;
$df_w = 8 - 1 = 7$. Reject H_0 if $F > 8.26$.
$$F = \frac{(22.95)^2}{(14.79)^2} = 2.41$$
Cannot reject H_0. There is no difference in the variation of the monthly sales.

19. **a.** $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$
H_1: Treatment means are not all equal.
b. $\alpha = .05$ Reject H_0 if $F > 3.10$.
c.

Source	SS	*df*	MS	*F*
Treatment	50	4 − 1 = 3	50/3	1.67
Error	200	24 − 4 = 20	10	
Total	250	24 − 1 = 23		

d. Do not reject H_0.

21. **a.** $H_0: \mu_{\text{Discount}} = \mu_{\text{Variety}} = \mu_{\text{Department}}$ H_1: Not all means are equal. H_0 is rejected if $F > 3.89$.
b. From Excel, single-factor ANOVA,

ANOVA						
Source of Variation	SS	*df*	MS	*F*	*p*-Value	*F* crit
Treatment	63.3333	2	31.6667	13.3803	0.0009	3.8853
Error	28.4000	12	2.3667			
Total	91.7333	14				

c. The *F*-test statistic is 13.3803.
d. *p*-value = .0009.
e. H_0 is rejected. The *F*-statistic exceeds the critical value; the *p*-value is less than .05.
f. There is a difference in the treatment means.

23. **a.** $H_0: \mu_{\text{Rec Center}} = \mu_{\text{Key Street}} = \mu_{\text{Monclova}} = \mu_{\text{Whitehouse}}$ H_1: Not all means are equal. H_0 is rejected if $F > 3.10$.
b. From Excel, single-factor ANOVA,

ANOVA						
Source of Variation	SS	*df*	MS	*F*	*p*-Value	*F* crit
Treatment	87.7917	3	29.2639	9.1212	0.0005	3.0984
Error	64.1667	20	3.2083			
Total	151.9583	23				

c. The *F*-test statistic is 9.1212.
d. *p*-value = .0005.
e. Since computed *F* of 9.1212 > 3.10, and the *p*-value is less than .05, the null hypothesis of no difference is rejected.
f. There is evidence the number of crimes differs by district.

25. **a.** $H_0: \mu_{\text{Lecture}} = \mu_{\text{Distance}}$ $H_1: \mu_{\text{Lecture}} \ne \mu_{\text{Distance}}$
Critical value of $F = 4.75$. Reject H_0 if the F-stat > 4.75.

ANOVA						
Source of Variation	SS	*df*	MS	*F*	*p*-Value	*F* crit
Treatment	219.4286	1	219.4286	23.0977	0.0004	4.7472
Error	114.0000	12	9.5000			
Total	333.4286	13				

Reject H_0 in favor of the alternative.
b. $$t = \frac{37 - 45}{\sqrt{9.5\left(\frac{1}{6} + \frac{1}{8}\right)}} - 4.806$$
Since $t^2 = F$. That is $(-4.806)^2 = 23.098$. The *p*-value for this statistic is 0.0004 as well. Reject H_0 in favor of the alternative.
c. There is a difference in the mean scores between lecture and distance-based formats.

27. **a.** $H_0: \mu_{\text{Compact}} = \mu_{\text{Midsize}} = \mu_{\text{Large}}$ H_1: Not all means are equal. H_0 is rejected if $F > 3.10$.
b. The *F*-test statistic is 8.258752.
c. *p*-value is .0019.
d. The null hypothesis of equal means is rejected because the *F*-statistic (8.258752) is greater than the critical value (3.10). The *p*-value (0.0019) is also less than the significance level (0.05).
e. The mean miles per gallon for the three car types are different.

29. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. H_1: At least one mean is different. Reject H_0 if $F > 2.7395$. Since 13.74 > 2.74, reject H_0. You can also see this from the *p*-value of 0.0001 < 0.05. Priority mail express is faster than all three of the other classes, and priority mail is faster than either first-class or standard. However, first-class and standard mail may be the same.

31. **a.** $H_0: \mu_R = \mu_O = \mu_Y = \mu_B = \mu_B = \mu_G$ H_1: The treatment means are not equal.
Reject H_0 if $F > 2.37$.
b. From Excel, Single Factor ANOVA,

ANOVA						
Source of Variation	SS	*df*	MS	*F*	*p*-Value	*F* crit
Treatment	0.0348	5	0.0070	3.8645	0.0043	2.3738
Error	0.1044	58	0.0018			
Total	0.1392	63				

For treatments, $F = 3.8645$.

c. For treatments, P-value = .0043.
d. Reject H_0 for color.
e. The mean weight of M&M's differ by color.

33. a. $H_0: \sigma_p^2 = \sigma_{np}^2 \quad H_1: \sigma_p^2 \neq \sigma_{np}^2$
Reject H_0. The p-value is less than 0.05. There is a difference in the variance of average selling prices between houses with pools and houses without pools.
b. $H_0: \sigma_g^2 = \sigma_{ng}^2 \quad H_1: \sigma_g^2 \neq \sigma_{ng}^2$
Reject H_0. There is a difference in the variance of average selling prices between house with garages and houses without garages. The p-value is < 0.0001.
c. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$; H_1: Not all treatment means are equal.
Fail to reject H_0. The p-value is much larger than 0.05. There is no statistical evidence of differences in the mean selling price between the five townships.
d. H_0: $\mu_c = \mu_i = \mu_m = \mu_p = \mu_r$ H_1: Not all treatment means are equal.
Fail to reject H_0. The p-value is much larger than 0.05. There is no statistical evidence of differences in the mean selling price between the five agents. Is fairness of assignment based on the overall mean price, or based on the comparison of the means of the prices assigned to the agents?
While the p-value is not less than 0.05, it may indicate that the pairwise differences should be reviewed. These indicate that Marty's comparisons to the other agents are significantly different.
e. The results show that the mortgage type is a significant effect on the mean years of occupancy (p = 0.0227). The interaction effect is also significant (p = 0.0026).

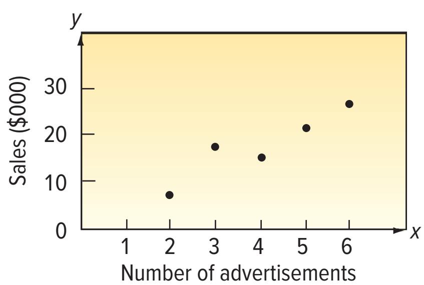
35. a. H_0: $\mu_B = \mu_K = \mu_T$ $\quad H_1$: Not all treatment (manufacturer) mean maintenance costs are equal.
Do not reject H_0. (p = 0.7664). The mean maintenance costs by the bus manufacturer are not different.
b. H_0: $\mu_B = \mu_K = \mu_T$ $\quad H_1$: Not all treatments have equal mean miles since the last maintenance.
Do not reject H_0. The mean miles since the last maintenance by the bus manufacturer are not different. p-Value = 0.4828.

CHAPTER 13

1. $\Sigma(x - \bar{x})(y - \bar{y}) = 10.6$, $s_x = 2.7$, $s_y = 1.3$

$$r = \frac{10.6}{(5-1)(2.709)(1.38)} = 0.75$$

3. a. Sales
b.

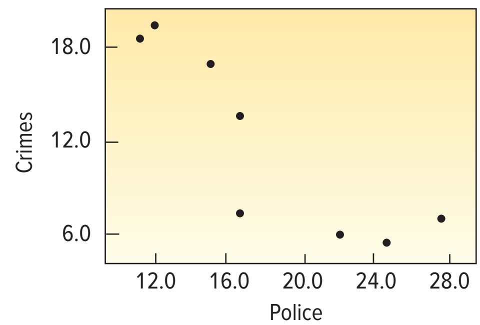


c. $\Sigma(x - \bar{x})(y - \bar{y}) = 36$, $n = 5$, $s_x = 1.5811$, $s_y = 6.1237$

$$r = \frac{36}{(5-1)(1.5811)(6.1237)} = 0.9295$$

d. There is a strong positive association between the variables.

5. a. Either variable could be independent. In the scatter plot, police is the independent variable.
b.



c. $n = 8$, $\Sigma(x - \bar{x})(y - \bar{y}) = -231.75$
$s_x = 5.8737$, $s_y = 6.4462$

$$r = \frac{-231.75}{(8-1)(5.8737)(6.4462)} = -0.8744$$

d. Strong inverse relationship. As the number of police increases, the crime decreases or, as crime increases the number of police decrease.

7. Reject H_0 if $t > 1.812$.

$$t = \frac{.32\sqrt{12-2}}{\sqrt{1-(.32)^2}} = 1.068$$

Do not reject H_0.

9. H_0: $\rho \leq 0$; H_1: $\rho > 0$. Reject H_0 if $t > 2.552$. $df = 18$.

$$t = \frac{.78\sqrt{20-2}}{\sqrt{1-(.78)^2}} = 5.288$$

Reject H_0. There is a positive correlation between gallons sold and the pump price.

11. H_0: $\rho \leq 0 \quad H_1$: $\rho > 0$
Reject H_0 if $t > 2.650$ with $df = 13$.

$$t = \frac{0.667\sqrt{15-2}}{\sqrt{1-0.667^2}} = 3.228$$

Reject H_0. There is a positive correlation between the number of passengers and plane weight.

13. a. $\hat{y} = 3.7671 + 0.3630x$

$$b = 0.7522\left(\frac{1.3038}{2.7019}\right) = 0.3630$$

$$a = 5.8 - 0.3630(5.6) = 3.7671$$

b. 6.3081, found by $\hat{y} = 3.7671 + 0.3630(7)$

15. a. $\Sigma(x - \bar{x})(y - \bar{y}) = 44.6$, $s_x = 2.726$, $s_y = 2.011$

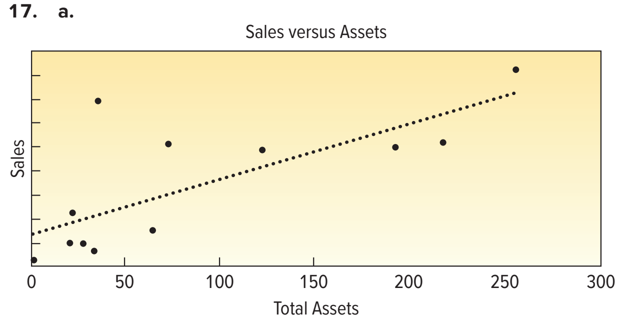
$$r = \frac{44.6}{(10-1)(2.726)(2.011)} = .904$$

$$b = .904\left(\frac{2.011}{2.726}\right) = 0.667$$

$$a = 7.4 - .677(9.1) = 1.333$$

b. $\hat{Y} = 1.333 + .667(6) = 5.335$

17. a.



b. Computing correlation in Excel, $r = .7340$

c.

	Total Assets	12-Month Sales
Mean	89.9342	60.32
Standard deviation	87.9057	49.8566
Count	12	12

$b = .7340 \frac{49.8566}{87.9057} = .4163; \quad a = 60.32 - .4163(89.9342)$

$= 22.8804$

d. $\hat{y} = 22.8804 + .4163(100.0) = 64.5104$ ($ billion)

19. a. $b = -.8744\left(\frac{6.4462}{5.8737}\right) = -0.9596$

$a = \frac{95}{8} - (-0.9596)\left(\frac{146}{8}\right) = 29.3877$

b. 10.1957, found by $29.3877 - 0.9596(20)$

c. For each police officer added, crime goes down by almost one.

21. H_0: $\beta \geq 0 \qquad H_1$: $\beta < 0 \qquad df = n - 2 = 8 - 2 = 6$

Reject H_0 if $t < -1.943$.

$$t = -0.96/0.22 = -4.364$$

Reject H_0 and conclude the slope is less than zero.

23. H_0: $\beta = 0 \qquad H_1$: $\beta = 0 \qquad df = n - 2 = 12 - 2 = 10$

Reject H_0 if t not between -2.228 and 2.228.

$$t = 0.08/0.03 = 2.667$$

Reject H_0 and conclude the slope is different from zero.

25. The standard error of estimate is 3.378, found by $\sqrt{\frac{68.4814}{8-2}}$.

The coefficient of determination is 0.76, found by $(-0.874)^2$. Seventy-six percent of the variation in crimes can be explained by the variation in police.

27. The standard error of estimate is 0.913, found by $\sqrt{\frac{6.667}{10-2}}$.

The coefficient of determination is 0.82, found by 29.733/36.4. Eighty-two percent of the variation in kilowatt=hours can be explained by the variation in the number of rooms.

29. a. $r^2 = \frac{1{,}000}{1{,}500} = .6667$

b. $r = \sqrt{.6667} = .8165$

c. $s_{y \cdot x} = \sqrt{\frac{500}{13}} = 6.2017$

31. a. $6.308 \pm (3.182)(.993)\sqrt{.2 + \frac{(7-5.6)^2}{29.2}}$

$= 6.308 \pm 1.633$

$= [4.675, 7.941]$

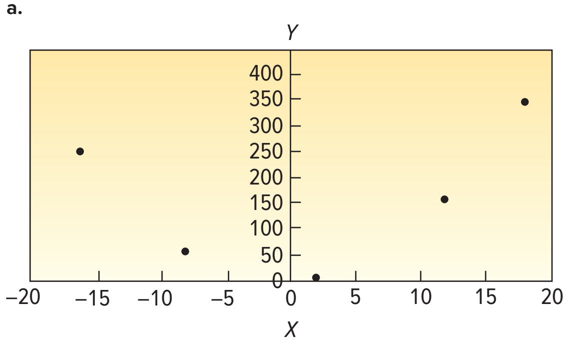
b. $6.308 \pm (3.182)(.993)\sqrt{1 + 1/5 + .0671}$

$= [2.751, 9.865]$

33. a. 4.2939, 6.3721

b. 2.9854, 7.6806

35. a.

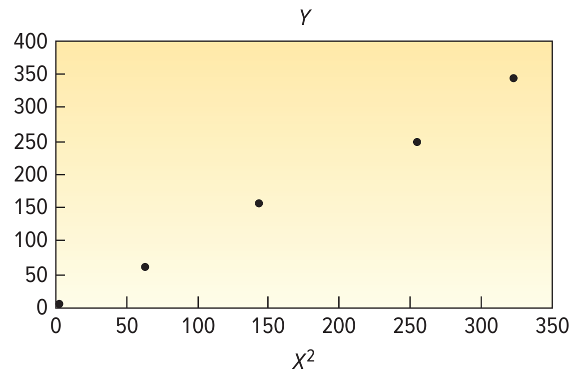


The correlation of X and Y is 0.2975. The scatter plot reveals the variables do not appear to be linearly related. In fact, the pattern is U-shaped.

b. The correlation coefficient is .2975.

c. Perform the task.

d.



e. The correlation between Y and $X^2 = .9975$.

f. The relationship between Y and X is nonlinear. The relationship between Y and the transformed X^2 is nearly perfectly linear.

g. Linear regression analysis can be used to estimate the linear relationship: $Y = a + b\,(X)^2$.

37. H_0: $\rho \leq 0$; H_1: $\rho > 0$. Reject H_0 if $t > 1.714$.

$$t = \frac{.94\sqrt{25-2}}{\sqrt{1-(.94)^2}} = 13.213$$

Reject H_0. There is a positive correlation between passengers and weight of luggage.

39. H_0: $\rho \leq 0$; H_1: $\rho > 0$. Reject H_0 if $t > 2.764$.

$$t = \frac{.47\sqrt{12-2}}{\sqrt{1-(.47)^2}} = 1.684$$

Do not reject H_0. Using an online p-value calculator or statistical software, the p-value is 0.0615.

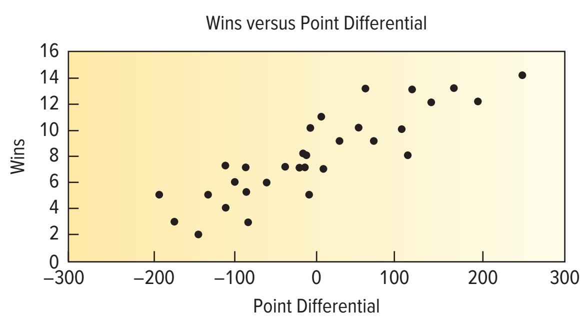
41. a. The correlation is -0.3156. The linear relationship between points allowed and points scored is weak.

b. H_0: $\rho \geq 0 \qquad H_1$: $\rho < 0 \qquad$ Reject H_0 if $t < -1.697$. $df = 30$

$t = \frac{-0.3156\sqrt{32-2}}{\sqrt{1-(-0.3156)^2}} = -1.8217$. Using an online calculator, p-value $= .03924$

Reject H_0. The evidence suggests a statistically significant inverse relationship between points scored and points allowed.

43. a. There is a positive relationship between wins and point differential.



b. $r = .8744$. There is a strong, positive relationship between wins and point differential.

c. The $R^2 = 76.45\%$. Point differential accounts for 76.45% of the variance of wins.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.8744
R Square	0.7645
Adjusted *R* Square	0.7567
Standard Error	1.5822
Observations	32

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	243.8647	243.8647	97.4108	0.0000
Residual	30	75.1040	2.5035		
Total	31	318.9688			

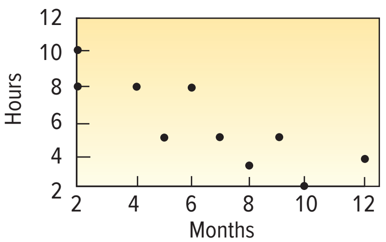
	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	7.9688	0.2797	28.4901	0.0000	7.3975	8.5400	7.3975	8.5400
Point differential	0.0260	0.0026	9.8697	0.0000	0.0207	0.0314	0.0207	0.0314

d. Wins = 7.9688 + .0260 (point differential)

e. Setting wins = 8, solve 8 = 7.9375 + .0282 (point differential) for point differential. The point differential is +1.1999 points; points scored and points allowed would be nearly equal.

f. The slope indicates that for every positive single point increase in point differential, wins increase .0282. Slope equals: (change in wins)/(for a unit change in point differential). Setting (change in Wins to 1), solve (Change in point differential) = 1/.0282 = 38.46 increase in the point differential. So, given that a team can win 8 of 16 games with about a zero point differential, we can predict that winning 9 games would require a point differential of about 38 points; winning 10 games would require a point differential of about 76 points, etc.

45. a.



There is an inverse relationship between the variables. As the months owned increase, the number of hours exercised decreases.

b. $r = -0.827$ The correlation coefficent indicates a strong, inverse linear relationship between months owned and hours exercised.

c. H_0: $\rho \geq 0$; H_1: $\rho < 0$. Reject H_0 if $t < -2.896$.

$$t = \frac{-0.827\sqrt{10-2}}{\sqrt{1-(-0.827)^2}} = -4.16$$

Reject H_0. There is a negative association between months owned and hours exercised.

47. a. There appears to be a weak positive relationship between population and median age.

b. Compute by hand or use Excel to compute the correlation coefficient.

Population (millions) X	Median Age Y	$(X-\bar{X})$	$(X-\bar{X})^2$	$(Y-\bar{Y})$	$(Y-\bar{Y})^2$	$(X-\bar{X})(Y-\bar{Y})$
2.833	31.5	0.3612	0.130465	−0.54	0.2916	−0.19505
1.233	30.5	−1.2388	1.534625	−1.54	2.3716	1.907752
2.144	30.9	−0.3278	0.107453	−1.14	1.2996	0.373692
3.849	31.6	1.3772	1.89668	−0.44	0.1936	−0.60597
8.214	34.2	5.7422	32.97286	2.16	4.6656	12.40315
1.448	34.2	−1.0238	1.048166	2.16	4.6656	−2.21141
1.513	30.7	−0.9588	0.919297	−1.34	1.7956	1.284792
1.297	31.7	−1.1748	1.380155	−0.34	0.1156	0.399432
1.257	32.5	−1.2148	1.475739	0.46	0.2116	−0.55881
0.93	32.6	−1.5418	2.377147	0.56	0.3136	−0.86341
24.718	320.4		43.84259		15.924	11.93418

$\bar{X} = \frac{24.718}{10} = 2.4718 \quad \bar{Y} = \frac{320.4}{10} = 32.04 \quad s_x = \sqrt{\frac{43.84259}{9}}$
$= 2.207$

$s_y = \sqrt{\frac{15.924}{9}} = 1.330$

$r = \frac{11.93418}{(10 - 1)(2.207)(1.330)} = 0.452$

The correlation coefficient indicates a weak positive relationship between population and median age.

c. The slope of 0.272 indicates that for each increase of 1 million in the population the median age increases on average by 0.272 year.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.4517
R Square	0.2040
Adjusted *R* Square	0.1045
Standard Error	1.2587
Observations	10

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	3.2485	3.2485	2.0503	0.1901
Residual	8	12.6755	1.5844		
Total	9	15.9240			

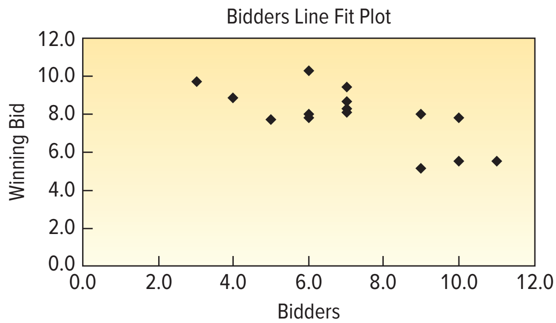
	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	31.3672	0.6158	50.9348	0.0000	29.9471	32.7873	29.9471	32.7873
Population	0.2722	0.1901	1.4319	0.1901	−0.1662	0.7106	−0.1662	0.7106

d. Median age = 31.3672 + .2722 (population). For a city with 2.5 million people, the predicted median age is 32.08 years, found by 31.4 + 0.272 (2.5).

e. The *p*-value (0.190) for the population variable is greater than, say 0.05. A test for significance of that coefficient would fail to be rejected. In other words, it is possible the population coefficient is zero.

f. The results indicate no significant linear relationship between a city's median age and its population.

49. **a.** The scatter plot indicates an inverse relationship between the winning bid and the number of bidders.



b. Using the following Excel software output, the correlation coefficient is −.7064. It indicates a moderate inverse relationship between winning bid and number of bidders.

c.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.7064
R Square	0.4990
Adjusted *R* Square	0.4604
Standard Error	1.1138
Observations	15

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	16.0616	16.0616	12.9467	0.0032
Residual	13	16.1277	1.2406		
Total	14	32.1893			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	11.2360	0.9689	11.5961	0.0000	9.1427	13.3293	9.1427	13.3293
Bidders	−0.4667	0.1297	−3.5982	0.0032	−0.7470	−0.1865	−0.7470	−0.1865

The R^2 = 49.90%.; the "number of bidders" accounts for 49.90% of the variance of the "winning bid cost".

d. The regression equation is Winning bid = 11.236 − 0.4667 (number of bidders).

e. This indicates there is a negative relationship between the number of bids (X) and the winning bid (Y) and that for each additional bidder the winning bid decreases by 0.4667 million. The slope is significantly different from zero because its p-value, .0032, is less than .05.

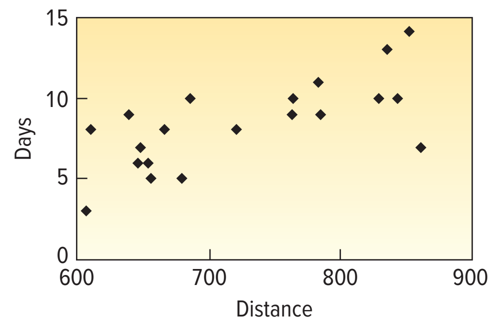
f. "Winning bid cost" = 11.235986 − 0.466727(7.0) = 7.968897 million

g. $7.9689 \pm (2.160)(1.114)\sqrt{1 + \frac{1}{15} + \frac{(7 - 7.1333)^2}{837 - \frac{(107)^2}{15}}}$

7.9689 ± 2.4854

[5.4835,10.4543]

51. a. There appears to be a relationship between the two variables. As the distance increases, so does the shipping time.



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.6921
R Square	0.4790
Adjusted R Square	0.4501
Standard Error	2.0044
Observations	20

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	66.4864	66.4864	16.5495	0.0007
Residual	18	72.3136	4.0174		
Total	19	138.8000			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−7.1264	3.8428	−1.8545	0.0801	−15.1999	0.9471	−15.1999	0.9471
Miles	0.0214	0.0053	4.0681	0.0007	0.0103	0.0324	0.0103	0.0324

b. From the regression output, $r = .6921$

$H_0: \rho \leq 0$ $H_0: \rho > 0$ Reject H_0 if $t > 1.734$.

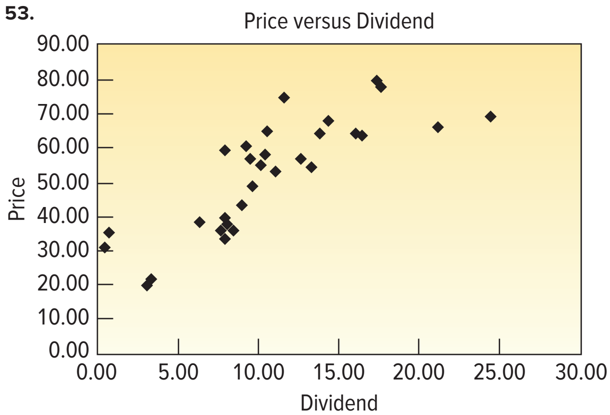
$t = \frac{0.6921\sqrt{20 - 2}}{1 - (0.6921)^2} = 3.4562$; the one-sided p-value (.0007/2) is .0004. H_0 is rejected. There is a positive association between shipping distance and shipping time.

c. $R^2 = (0.6921)^2 = 0.4790$, nearly half of the variation in shipping time is explained by shipping distance.

d. The standard error of estimate is $2.0044 = \sqrt{72.3136/18}$.

e. Predicting days based on miles will not be very accurate. The standard error of the estimate indicates that the prediction of days may be off by nearly 2 days. The regression equation only accounts for about half of the variation in shipping time with distance.

53.



SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.8114
R Square	0.6583
Adjusted *R* Square	0.6461
Standard Error	9.6828
Observations	30

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	5057.5543	5057.5543	53.9438	0.0000
Residual	28	2625.1662	93.7559		
Total	29	7682.7205			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	26.8054	3.9220	6.8346	0.0000	18.7715	34.8393	18.7715	34.8393
Dividend	2.4082	0.3279	7.3446	0.0000	1.7365	3.0798	1.7365	3.0798

a. The regression equation is: Price = 26.8054 + 2.4082 dividend. For each additional dollar paid out in a dividend, the per share price increases by \$2.4082 on average.

b. $H_0: \beta = 0 \quad H_1: \beta \neq 0$ At the 5% level, reject H_0 if t is not between −2.048 and 2.048. $t = 2.4082/0.3279 = 7.3446$ Reject H_0 and conclude slope is not zero.

c. $R^2 = \frac{Reg\ SS}{Total\ SS} = \frac{5057.5543}{7682.7205} = .6583$. 65.83% of the variation in price is explained by the dividend.

d. $r = \sqrt{.6583} = .8114$; 28 *df*; $H_0: \rho \leq 0$ $H_1: \rho > 0$
At the 5% level, reject H_0 when $t > 1.701$.
$t = \frac{0.8114\sqrt{30-2}}{\sqrt{1-(0.8114)^2}} = 7.3457$; using a *p*-value calculator, *p*-value is less than .00001.
Thus H_0 is rejected. The population correlation is positive.

e. Price = 26.8054 + 2.4082 (\$10) = \$50.8874

f. $\$50.8874 \pm 2.048(9.6828)\sqrt{1+\frac{1}{30}+\frac{(10-10.6777)^2}{872.1023}}$

The interval is (\$30.7241, \$71.0507).

55. **a.** 35

b. $s_{y \cdot x} = \sqrt{29{,}778{,}406} = 5{,}456.96$

c. $r^2 = \frac{13{,}548{,}662{,}082}{14{,}531{,}349{,}474} = 0.932$

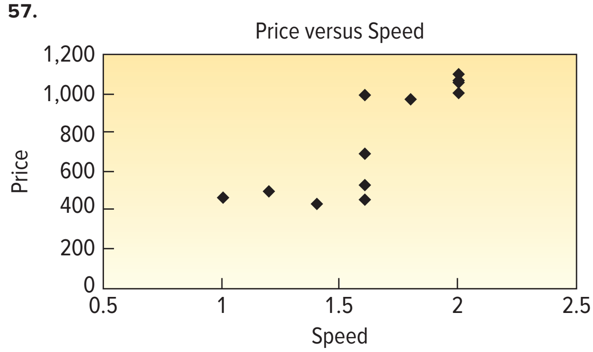
d. $r = \sqrt{0.932} = 0.966$

e. $H_0: \rho \leq 0$, $H_1: \rho > 0$; reject H_0 if $t > 1.692$.

$$t = \frac{.966\sqrt{35-2}}{\sqrt{1-(.966)^2}} = 21.46$$

Reject H_0. There is a direct relationship between size of the house and its market value.

57.



SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.8346
R Square	0.6966
Adjusted *R* Square	0.6662
Standard Error	161.6244
Observations	12

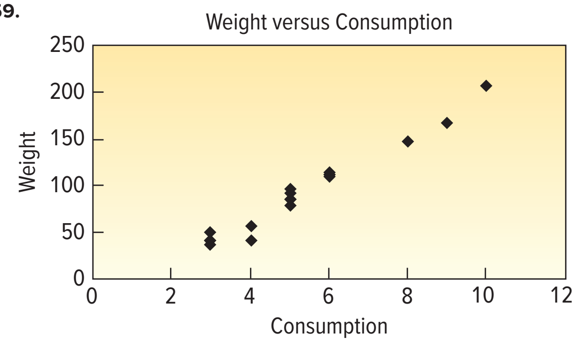
ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	599639.0413	599639.0413	22.9549	0.0007
Residual	10	261224.4587	26122.4459		
Total	11	860863.5000			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−386.5455	246.8853	−1.5657	0.1485	−936.6403	163.5494	−936.6403	163.5494
Speed	703.9669	146.9313	4.7911	0.0007	376.5837	1031.3502	376.5837	1031.3502

a. The correlation of Speed and Price is 0.8346.
$H_0: \rho \leq 0 \qquad H_1: \rho > 0 \qquad$ Reject H_0 if $t > 1.81252$.
$t = \dfrac{0.8346\sqrt{12-2}}{\sqrt{1-(0.8346)^2}} = 4.7911$ Using a p-value calculator or statistical software, the p-value is 0.0004.
Reject H_0. It is reasonable to say the population correlation is positive.
b. The regression equation is Price = −386.5455 + 703.9669 Speed.
c. The standard error of the estimate is 161.6244. Any prediction with a residual more than the standard error would be unusual. The computers 2, 3, and 10 have errors in excess of $200.00.

59.



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9872
R Square	0.9746
Adjusted R Square	0.9730
Standard Error	7.7485
Observations	18

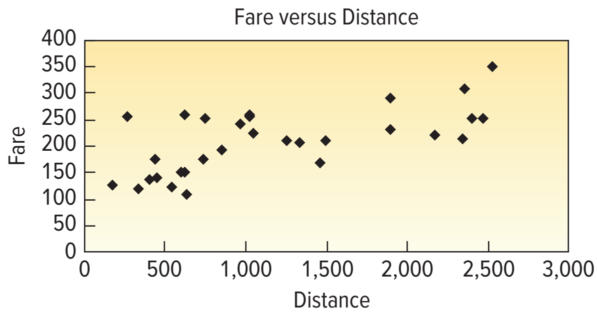
ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	36815.6444	36815.6444	613.1895	0.0000
Residual	16	960.6333	60.0396		
Total	17	37776.2778			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	−29.7000	5.2662	−5.6398	0.0000	−40.8638	−18.5362	−40.8638	−18.5362
Consumption	22.9333	0.9261	24.7627	0.0000	20.9700	24.8966	20.9700	24.8966

a. The correlation of Weight and Consumption is 0.9872.
$H_0: \rho \leq 0 \qquad H_1: \rho > 0 \qquad$ Reject H_0 if $t > 1.746$.
$t = \dfrac{0.9872\sqrt{18-2}}{1-(0.9872)^2} = 24.7627$. Using a p-value calculator or statistical software, the p-value is less than .00001.
Reject H_0. It is quite reasonable to say the population correlation is positive!
b. The regression equation is Weight = −29.7000 + 22.9333(Consumption). Each additional cup increases the estimated weight by 22.9333 pounds.
c. The fourth dog has the largest residual weighing 21 pounds less than the regression equation would estimate. The 16th dog's residual of 10.03 also exceeds the standard error of the estimate; it weights 10.03 pounds more than the predicted weight.

61. a. The relationship is direct. Fares increase for longer flights.



b. The correlation between Distance and Fare is 0.6556.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.6556
R Square	0.4298
Adjusted *R* Square	0.4094
Standard Error	46.3194
Observations	30

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	45279.0688	45279.0688	21.1043	0.0001
Residual	28	60073.5978	2145.4856		
Total	29	105352.6667			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	147.0812	15.8503	9.2794	0.0000	114.6133	179.5490	114.6133	179.5490
Distance	0.0527	0.0115	4.5939	0.0001	0.0292	0.0761	0.0292	0.0761

H_0: $\rho \le 0$; H_1: $\rho > 0$; Reject H_0 if $t > 1.701$. $df = 28$

$t = \dfrac{0.6556\sqrt{30-2}}{\sqrt{1-(0.6556)^2}} = 4.5939$ Using a *p*-value calculator or statistical software, the *p*-value is .000042.

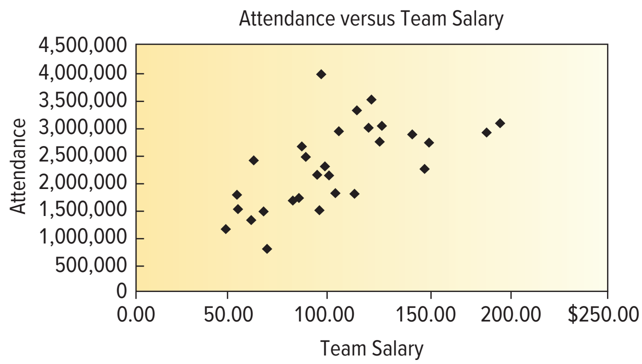
Reject H_0. There is a significant positive correlation between fares and distances.

c. 42.98 percent, found by $(0.6556)^2$, of the variation in fares is explained by the variation in distance.

d. The regression equation is Fare = 147.0812 + 0.0527(Distance). Each additional mile adds \$0.0527 to the fare. A 1,500-mile flight would cost \$226.1312, found by \$147.0812 + 0.0527(1500).

e. A flight of 4,218 miles is outside the range of the sampled data. So the regression equation may not be useful.

63. a. There does seem to be a direct relationship between the variables.



b. The regression analysis of attendance versus team salary follows:

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.6353
R Square	0.4037
Adjusted *R* Square	0.3824
Standard Error	598985.7062
Observations	30

ANOVA

	df	SS	MS	*F*	*p-Value*
Regression	1	6799983921035.39	6799983921035.39	18.95	0.00
Residual	28	10045948534908.80	358783876246.74		
Total	29	16845932455944.20			

	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	917096.4020	332297.3286	2.7599	0.0101
Team Salary	12960.0475	2976.9333	4.3535	0.0002

The regression equation is: Attendance = 917,096.4020 + 12,960.0475(Team Salary).

Expected Attendance with a salary of \$100 million is 2,213,101.152, found by 917,096.4020 + 12,960.0475 (100).

c. Increasing the salary by \$30 million will increase attendance by 338,801.425 on average, found by 12,960.0475 (30).

d. H_0: $\beta \le 0$ H_1: $\beta > 0$ $df = n - 2 = 30 - 2 = 28$ Reject H_0 if $t > 1.701$ $t = 12{,}960.0475/2{,}976.9333 = 4.3535$.

Using a *p*-value calculator or statistical software, the *p*-value is less than .0002.

Reject H_0 and conclude the slope is positive.

e. 0.4037, or 40.37%, of the variation in attendance is explained by variation in salary.

f.

Correlation Matrix

	Attendance	ERA	BA
Attendance	1.0000		
ERA	−0.3596	1.0000	
BA	0.3099	−0.2116	1.0000

The correlation between attendance and batting average is 0.3099.

H_0: $\rho \leq 0$ H_1: $\rho > 0$ At the 5% level, reject H_0 if $t > 1.701$

$$t = \frac{0.3099\sqrt{30 - 2}}{\sqrt{1 - (0.3099)^2}} = 1.7247$$

Using a *p*-value calculator or statistical software, the *p*-value is .0478.

Reject H_0.

The batting average and attendance are positively correlated.

The correlation between attendance and ERA is −0.3596. The correlation between attendance and ERA is stronger than the correlation between attendance and batting average.

H_0: $\rho \geq 0$ H_1: $\rho < 0$ At the 5% level, reject H_0 if $t < -1.701$

$$t = \frac{-0.3596\sqrt{30 - 2}}{\sqrt{1 - (.0.3596)^2}} = -2.0392$$

Using a *p*-value calculator or statistical software, the *p*-value is .0255.

Reject H_0.

The ERA and attendance are negatively correlated. Attendance increases when ERA decreases.

CHAPTER 14

1. **a.** It is called multiple regression analysis because the analysis is based on more than one independent variable.

b. +9.6 is the coefficient of the independent variable, per capita income. It means that for a 1-unit increase in per capita income, sales will increase \$9.60.

c. −11,600 is the coefficient of the independent variable, regional unemployment rate. Note that this coefficient is negative. It means that for a 1-unit increase in regional unemployment rate, sales will decrease \$11,600.

d. \$374,748 found by = 64,100 + 0.394(796,000) + 9.6(6940) 11,600(6.0)

3. **a.** 497.736, found by

$$\hat{y} = 16.24 + 0.017(18) + 0.0028(26{,}500) + 42(3)$$
$$+0.0012(156{,}000) + 0.19(141) + 26.8(2.5)$$

b. Two more social activities. Income added only 28 to the index; social activities added 53.6.

5. **a.** $$s_{Y\cdot 12} = \sqrt{\frac{SSE}{n - (k + 1)}} = \sqrt{\frac{583.693}{65 - (2 + 1)}}$$
$$= \sqrt{9.414} = 3.068$$

Based on the empirical rule, about 95% of the residuals will be between ± 6.136, found by 2(3.068).

b. $$R^2 = \frac{\text{SSR}}{\text{SS total}} = \frac{77.907}{661.6} = .118$$

The independent variables explain 11.8% of the variation.

c. $$R^2_{adj} = 1 - \frac{\dfrac{\text{SSE}}{n - (k + 1)}}{\dfrac{\text{SS total}}{n - 1}} = 1 - \frac{\dfrac{583.693}{65 - (2 + 1)}}{\dfrac{661.6}{65 - 1}}$$
$$= 1 - \frac{9.414}{10.3375} = 1 - .911 = .089$$

7. **a.** $\hat{y} = 84.998 + 2.391x_1 - 0.4086x_2$

b. 90.0674, found by $\hat{y} = 84.998 + 2.391(4) - 0.4086(11)$

c. $n = 65$ and $k = 2$

d. H_0: $\beta_1 = \beta_2 = 0$ H_1: Not all βs are 0

Reject H_0 if $F > 3.15$.

$F = 4.14$, reject H_0. Not all net regression coefficients equal zero.

e.

For x_1	For x_2
H_0: $\beta_1 = 0$	H_0: $\beta_2 = 0$
H_1: $\beta_1 \neq 0$	H_1: $\beta_2 \neq 0$
$t = 1.99$	$t = -2.38$

Reject H_0 if $t > 2.0$ or $t < -2.0$.

Delete variable 1 and keep 2.

f. The regression analysis should be repeated with only x_2 as the independent variable.

9. **a.** The regression equation is: Performance = 29.3 + 5.22 Aptitude + 22.1 Union

```
Predictor     Coef    SE Coef      T        P
Constant     29.28      12.77    2.29    0.041
Aptitude     5.222      1.702    3.07    0.010
Union       22.135      8.852    2.50    0.028

S = 16.9166 R-Sq = 53.3% R-Sq (adj) = 45.5%

Analysis of Variance
Source            DF       SS       MS       F        P
Regression         2   3919.3   1959.6    6.85    0.010
Residual Error    12   3434.0    286.2
Total             14   7353.3
```

b. These variables are both statistically significant in predicting performance. They explain 45.5% of the variation in performance. In particular union membership increases the typical performance by 22.1. A 1-unit increase in aptitude predicts a 5.222 increase in performance score.

c. H_0: $\beta_2 = 0$ H_1: $\beta_2 \neq 0$

Reject H_0 if $t < -2.179$ or $t > 2.179$. Since 2.50 is greater than 2.179, we reject the null hypothesis and conclude that union membership is significant and should be included. The corresponding *p*-value is .028.

11. **a.** $n = 40$

b. 4

c. $R^2 = \dfrac{750}{1{,}250} = .60$ Note total SS is the sum of regression SS and error SS.

d. $s_{y\cdot 1234} = \sqrt{500/35} = 3.7796$

e. H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_1: Not all the βs equal zero.

H_0 is rejected if $F > 2.65$.

$$F = \frac{750/4}{500/35} = 13.125$$

H_0 is rejected. At least one β_i does not equal zero.

13. **a.** $n = 26$

b. $R^2 = 100/140 = .7143$

c. 1.4142, found by $\sqrt{2}$

d. H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

H_1: Not all the βs are 0.

H_0 is rejected if $F > 2.71$.

Computed $F = 10.0$. Reject H_0. At least one regression coefficient is not zero.

e. H_0 is rejected in each case if $t < -2.086$ or $t > 2.086$.

x_1 and x_5 should be dropped.

15. **a.** \$28,000

b. $R^2 = \dfrac{\text{SSR}}{\text{SS total}} = \dfrac{3{,}050}{5{,}250} = .5809$

c. 9.199, found by $\sqrt{84.62}$

d. H_0 is rejected if $F > 2.97$ (approximately).

$$\text{Computed } F = \frac{1{,}016.67}{84.62} = 12.01$$

H_0 is rejected. At least one regression coefficient is not zero.

e. If computed t is to the left of -2.056 or to the right of 2.056, the null hypothesis in each of these cases is rejected. Computed t for x_2 and x_3 exceed the critical value. Thus, "population" and "advertising expenses" should be retained and "number of competitors," x_1, dropped.

17. a. The strongest correlation is between High School GPA and Paralegal GPA. No problem with multicollinearity.

b. $R^2 = \dfrac{4.3595}{5.0631} = .8610$

c. H_0 is rejected if $F > 5.41$.

$$F = \frac{1.4532}{0.1407} = 10.328$$

At least one coefficient is not zero.

d. Any H_0 is rejected if $t < -2.571$ or $t > 2.571$. It appears that only High School GPA is significant. Verbal and math could be eliminated.

e. $R^2 = \dfrac{4.2061}{5.0631} = .8307$

R^2 has only been reduced .0303.

f. The residuals appear slightly skewed (positive) but acceptable.

g. There does not seem to be a problem with the plot.

19. a. The correlation of Screen and Price is 0.893. So there does appear to be a linear relationship between the two.

b. Price is the "dependent" variable.

c. The regression equation is Price = −1242.1 + 50.671 (screen size). For each inch increase in screen size, the price increases $50.671 on average.

d. Using a "dummy" variable for Sony, the regression equation is Price = 11145.6 + 46.955 (Screen) + 187.10 (Sony). If we set "Sony" = 0, then the manufacturer is Samsung and the price is predicted only by screen size. If we set "Sony" = 1, then the manufacturer is Sony. Therefore, Sony TVs are, on average, $187.10 higher in price than Samsung TVs.

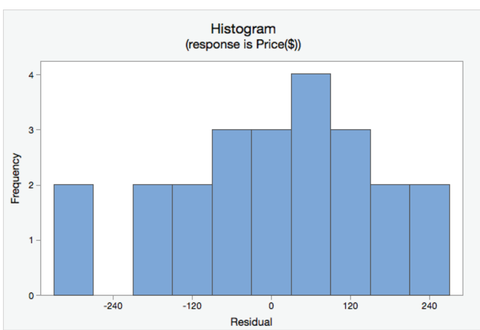
e. Here is some of the output.

Coefficients

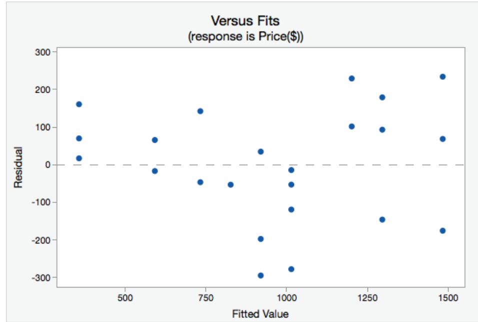
Term	Coef	SE Coef	95% CI	t-Value	p-Value
Constant	−1145.6	220.7	(−1606.1, −685.2)	−5.19	<0.0001
Screen	46.955	5.149	(36.215, 57.695)	9.12	<0.0001
Sony 1	187.10	71.84	(37.24, 336.96)	2.60	0.0170

Based on the p-values, screen size and manufacturer are both significant in predicting price.

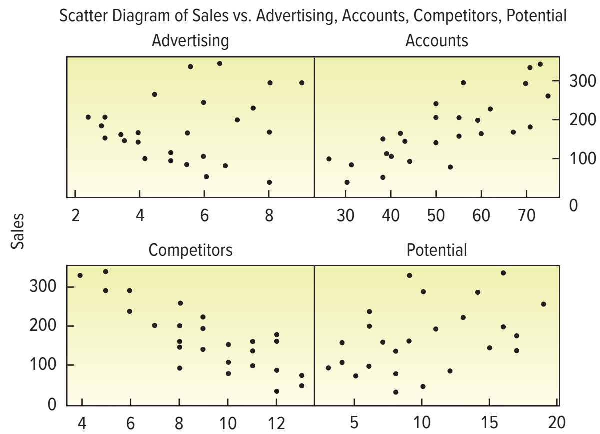
f. A histogram of the residuals indicates they follow a normal distribution.



g. There is no apparent relationship in the residuals, but the residual variation may be increasing with larger fitted values.



21. a.



Sales seem to fall with the number of competitors and rise with the number of accounts and potential.

b. Pearson correlations

	Sales	Advertising	Accounts	Competitors
Advertising	0.159			
Accounts	0.783	0.173		
Competitors	−0.833	−0.038	−0.324	
Potential	0.407	−0.071	0.468	−0.202

The number of accounts and the market potential are moderately correlated.

c. The regression equation is:

Sales = 178 + 1.81 Advertising + 3.32 Accounts −21.2 Competitors + 0.325 Potential

Predictor	Coef	SE Coef	T	P
Constant	178.32	12.96	13.76	0.000
Advertising	1.807	1.081	1.67	0.109
Accounts	3.3178	0.1629	20.37	0.000
Competitors	−21.1850	0.7879	−26.89	0.000
Potential	0.3245	0.4678	0.69	0.495

S = 9.60441   R-Sq = 98.9%   R-Sq(adj) = 98.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	176777	44194	479.10	0.000
Residual Error	21	1937	92		
Total	25	178714			

The computed F-value is quite large. So we can reject the null hypothesis that all of the regression coefficients are zero. We conclude that some of the independent variables are effective in explaining sales.

d. Market potential and advertising have large *p*-values (0.495 and 0.109, respectively). You would probably drop them.

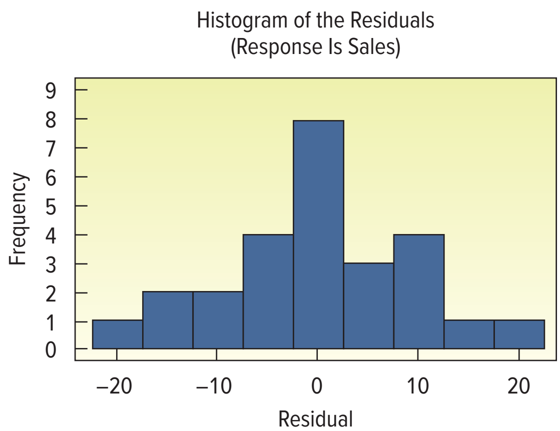
e. If you omit potential, the regression equation is:
Sales = 180 + 1.68 Advertising + 3.37 Accounts − 21.2 Competitors

Predictor	Coef	SE Coef	T	P
Constant	179.84	12.62	14.25	0.000
Advertising	1.677	1.052	1.59	0.125
Accounts	3.3694	0.1432	23.52	0.000
Competitors	−21.2165	0.7773	−27.30	0.000

Now advertising is not significant. That would also lead you to cut out the advertising variable and report that the polished regression equation is: Sales = 187 + 3.41 Accounts −21.2 Competitors

Predictor	Coef	SE Coef	T	P
Constant	186.69	12.26	15.23	0.000
Accounts	3.4081	0.1458	23.37	0.000
Competitors	−21.1930	0.8028	−26.40	0.000

f.



The histogram looks to be normal. There are no problems shown in this plot.

g. The variance inflation factor for both variables is 1.1. They are less than 10. There are no troubles as this value indicates the independent variables are not strongly correlated with each other.

23. The computer output is:

Predictor	*Coef*	*StDev*	*t-ratio*	*p*
Constant	651.9	345.3	1.89	0.071
Service	13.422	5.125	2.62	0.015
Age	−6.710	6.349	−1.06	0.301
Gender	205.65	90.27	2.28	0.032
Job	−33.45	89.55	−0.37	0.712

Analysis of Variance

SOURCE	*DF*	*SS*	*MS*	*F*	*p*
Regression	4	1066830	266708	4.77	0.005
Error	25	1398651	55946		
Total	29	2465481			

a. $\hat{y} = 651.9 + 13.422x_1 - 6.710x_2 + 205.65x_3 - 33.45x_4$

b. $R^2 = .433$, which is somewhat low for this type of study.

c. H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$; H_1: Not all βs equal zero.
Reject H_0 if $F > 2.76$.

$$F = \frac{1{,}066{,}830/4}{1{,}398{,}651/25} = 4.77$$

H_0 is rejected. Not all the β_is equal 0.

d. Using the .05 significance level, reject the hypothesis that the regression coefficient is 0 if $t < -2.060$ or $t > 2.060$. Service and gender should remain in the analyses; age and job should be dropped.

e. Following is the computer output using the independent variables service and gender.

Predictor	*Coef*	*StDev*	*t-ratio*	*p*
Constant	784.2	316.8	2.48	0.020
Service	9.021	3.106	2.90	0.007
Gender	224.41	87.35	2.57	0.016

Analysis of Variance

SOURCE	*DF*	*SS*	*MS*	*F*	*p*
Regression	2	998779	499389	9.19	0.001
Error	27	1466703	54322		
Total	29	2465481			

A man earns $224 more per month than a woman. The difference between management and engineering positions is not significant.

25. **a.** The correlation between the independent variables, yield and EPS, is small, .16195. Multicollinearity should not be an issue.

Correlation Matrix

	P/E	EPS	Yield
P/E	1		
EPS	−0.60229	1	
Yield	0.05363	0.16195	1

b. Here is part of the software output:

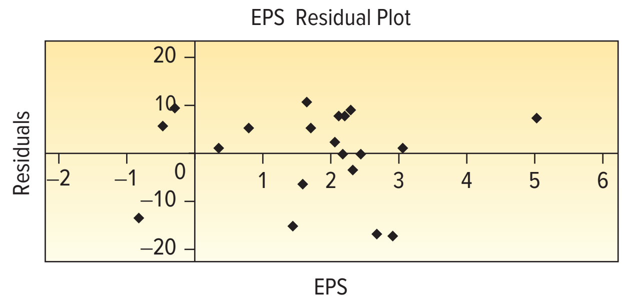
Predictor	Coef	SE Coef	*t*	*p*-Value
Constant	29.913	5.767	5.19	0.000
EPS	−5.324	1.634	−3.26	0.005
Yield	1.449	1.798	0.81	0.431

The regression equation is P/E = 29.913 − 5.324 EPS + 1.449 Yield.

c. Thus EPS has a significant relationship with P/E but not with Yield.
The regression equation is P/E = 33.5668 – 5.1107 EPS.

d. If EPS increases by one, P/E decreases by 5.1107

e. Yes, the residuals are evenly distributed above and below the horizontal line (residual = 0).



f. No. the adjusted R^2 indicates that the regression equation only accounts for 32.78% of the variation in P/E. The predictions will not be accurate.

27. **a.** The regression equation is
Sales (000) = 1.02 + 0.0829 Infomercials.

Predictor	Coef	SE Coef	T	P
Constant	1.0188	0.3105	3.28	0.006
Infomercials	0.08291	0.01680	4.94	0.000

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2.3214	2.3214	24.36	0.000
Residual Error	13	1.2386	0.0953		
Total	14	3.5600			

The global test demonstrates there is a relationship between sales and the number of infomercials.

SUMMARY OUTPUT

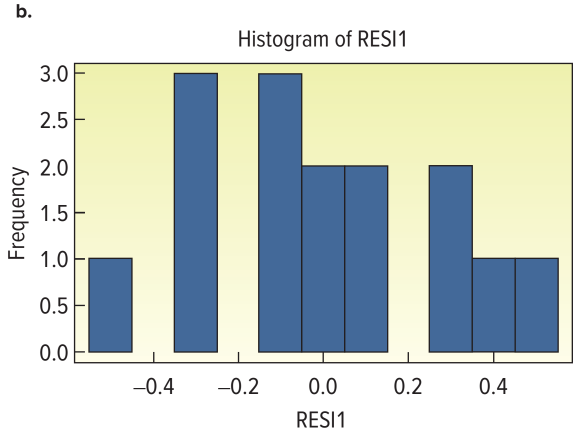
Regression Statistics	
Multiple *R*	0.6023
R Square	0.3628
Adjusted *R* Square	0.3274
Standard Error	9.4562
Observations	20

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	916.2448	916.2448	10.2466	0.0050
Residual	18	1609.5483	89.4193		
Total	19	2525.7931			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	33.5688	3.5282	9.5145	0.0000	26.1564	40.9812	26.1564	40.9812
EPS	−5.1107	1.5966	−3.2010	0.0050	−8.4650	−1.7564	−8.4650	−1.7564

b.



The residuals appear to follow the normal distribution.

29. a. The correlation matrix is as follows:

	Price	Bedrooms	Size (square feet)	Baths	Days on Market
Price	1.000				
Bedrooms	0.844	1.000			
Size (square feet)	0.952	0.877	1.000		
Baths	0.825	0.985	0.851	1.000	
Days on market	0.185	0.002	0.159	−0.002	1

The correlations for strong, positive relationships between "Price" and the independent variables "Bedrooms," "Size," and "Baths." There appears to be no relationship between "Price" and "Days on Market". The correlations among the independent variables are very strong. So, there would be a high degree of multicollinearity in a multiple regression equation if all the variables were included. We will need to be careful in selecting the best independent variable to predict price.

b.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.952
R Square	0.905
Adjusted *R* Square	0.905
Standard Error	49655.822
Observations	105.000

ANOVA

	df	*SS*	*MS*	*F*	Significance *F*
Regression	1	2.432E+12	2.432E+12	9.862E+02	1.46136E-54
Residual	103	2.540E+11	2.466E+09		
Total	104	2.686E+12			

	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	−15775.955	12821.967	−1.230	0.221
Size (square feet)	108.364	3.451	31.405	0.000

The regression analysis shows a significant relationship between price and house size. The *p*-value of the *F*-statistic is 0.00, so the null hypothesis of "no relationship" is rejected. Also, the *p*-value associated with the regression coefficient of "size" is 0.000. Therefore, this coefficient is clearly different from zero.

The regression equation is: Price = −15775.995 + 108.364 Size.

In terms of pricing, the regression equation suggests that houses are priced at about $108 per square foot.

c. The regression analyses of price and size with the qualitative variables pool and garage follow. The results show that the variable "pool" is statistically significant in the equation. The regression coefficient indicates that if a house has a pool, it adds about $28,575 to the price. The analysis of including "garage" to the analysis indicates that it does not affect the pricing of the house.

Adding pool to the regression equation increases the *R*-square by about 1%.

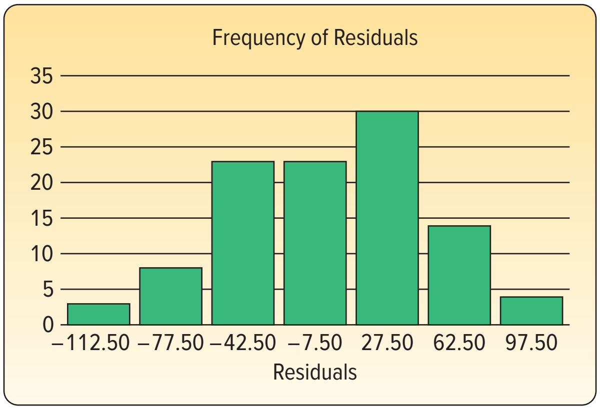
SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.955
R Square	0.913
Adjusted R Square	0.911
Standard Error	47914.856
Observations	105

ANOVA

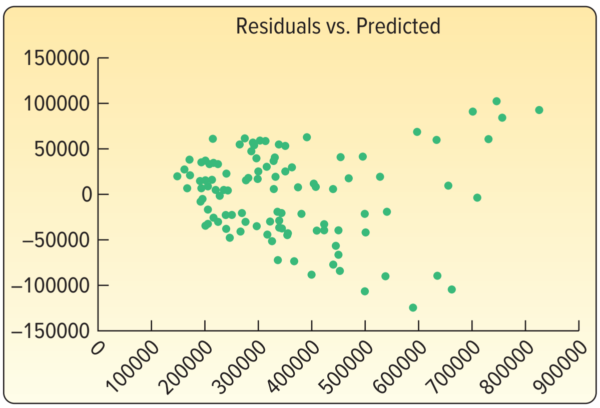
	df	SS	MS	F	Significance F
Regression	2.00	2451577033207.43	1225788516603.72	533.92	0.00
Residual	102.00	234175013207.24	2295833462.82		
Total	104.00	2685752046414.68			

	Coefficients	Standard Error	t-Stat	p-Value
Intercept	−34640.573	13941.203	−2.485	0.015
Size (square feet)	108.547	3.330	32.595	0.000
Pool (yes is 1)	28575.145	9732.223	2.936	0.004

d. The following histogram was developed using the residuals from part (c). The normality assumption is reasonable.



e. The following scatter diagram is based on the residuals in part (c) with the predicted dependent variable on the horizontal axis and residuals on the vertical axis. There does appear that the variance of the residuals increases with higher values of the predicted price. You can experiment with transformations such as the Log of Price or the square root of price and observe the changes in the graphs of residuals. Note that the transformations will make the interpretation of the regression equation more difficult.



31. a.

	Maintenance Cost ($)	Age (years)	Odometer Miles	Miles since Last Maintenance
Maintenance cost ($)	1			
Age (years)	0.710194278	1		
Odometer miles	0.700439797	0.990675674	1	
Miles since last maint.	−0.160275988	−0.140196856	−0.118982823	1

The correlation analysis shows that age and odometer miles are positively correlated with cost and that “miles since last maintenance” shows that costs increase with fewer miles between maintenance. The analysis also shows a strong correlation between age and odometer miles. This indicates the strong possibility of multicollinearity if age and odometer miles are included in a regression equation.

b. There are a number of analyses to do. First, using Age or Odometer Miles as an independent variable. When you review these analyses, both result in significant relationships. However, Age has a slightly higher R^2. So I would select age as the first independent variable. The interpretation of the coefficient using age is bit more useful for practical use. That is, we can expect about an average of $600 increase in maintenance costs for each additional year a bus ages. The results are:

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.708
R Square	0.501
Adjusted R Square	0.494
Standard Error	1658.097
Observations	80

ANOVA

	df	SS	MS	F	Significance F
Regression	1	215003471.845	215003471.845	78.203	0.000
Residual	78	214444212.142	2749284.771		
Total	79	429447683.988			

	Coefficients	Standard Error	t-Stat	p-Value
Intercept	337.297	511.372	0.660	0.511
Age (years)	603.161	68.206	8.843	0.000

We can also explore including the variable “miles since last maintenance” with Age. Your analysis will show that “miles since last maintenance” is not significantly related to costs.

Last, it is possible that maintenance costs are different for diesel versus gasoline engines. So, adding this variable to the analysis shows:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.960
R Square	0.922
Adjusted R Square	0.920
Standard Error	658.369
Observations	80

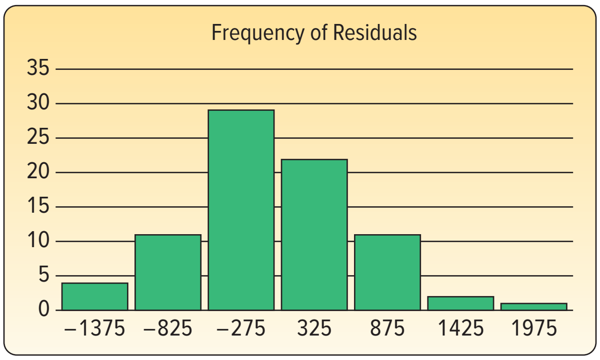
ANOVA

	df	SS	MS	*F*	Significance *F*
Regression	2	396072093.763	198036046.881	456.884	0.000
Residual	77	33375590.225	433449.224		
Total	79	429447683.988			

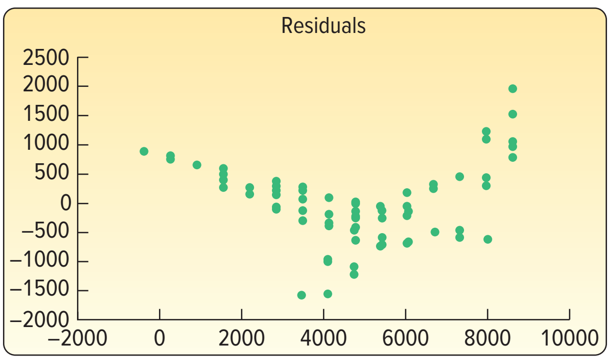
	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	–1028.539	213.761	–4.812	0.000
Age (years)	644.528	27.157	23.733	0.000
Engine Type (0=diesel)	3190.481	156.100	20.439	0.000

The results show that the engine type is statistically significant and increases the R^2 to 92.2%. Now the practical interpretation of the analysis is that, on average, buses with gasoline engines cost about $3,190 more to maintain. Also, the maintenance costs increase with bus age at an average of $644 per year of bus age.

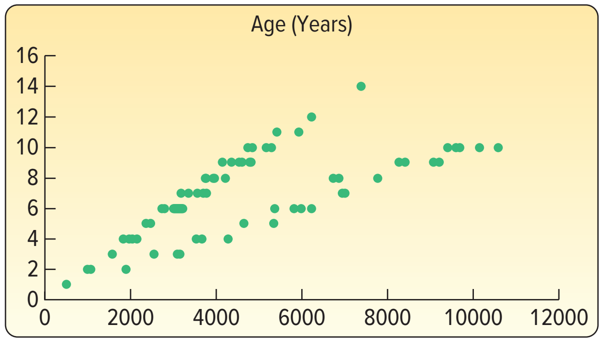
c. The normality conjecture appears realistic.



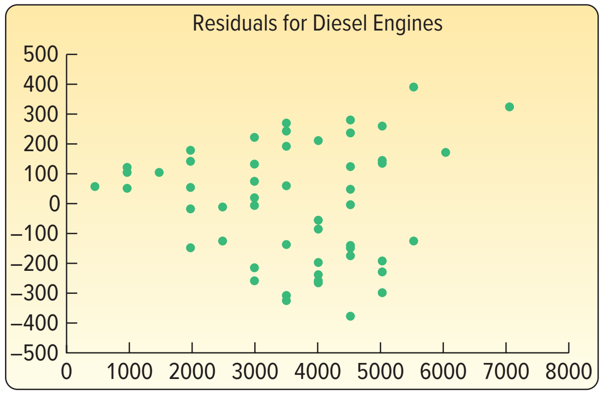
d. The plot of residuals versus predicted values shows the following. There are clearly patterns in the graph that indicate that the residuals do not follow the assumptions required for the tests of hypotheses.

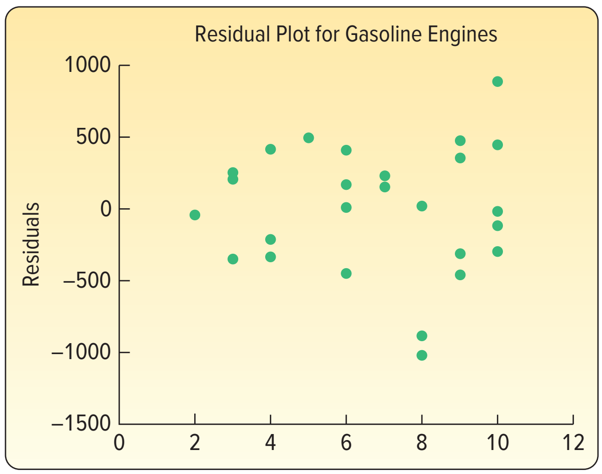


Let's remember the scatter plot of costs versus age. The graph clearly shows the effect of engine type on costs. So there are essentially two regression equations depending on the type of engine.



So based on our knowledge of the data, let's create a residual plot of costs for each engine type.





The graphs show a much better distribution of residuals.

CHAPTER 15

1. **a.** H_0 is rejected if $z > 1.65$.

b. 1.09, found by $z = (0.75 - 0.70)/\sqrt{(0.70 \times 0.30)/100}$

c. H_0 is not rejected.

3. **Step 1:** H_0: $\pi = 0.10$ $\quad H_1$: $\pi \neq 0.10$
Step 2: The 0.01 significance level was chosen.
Step 3: Use the z-statistic as the binomial distribution can be approximated by the normal distribution as $n\pi = 30 > 5$ and $n(1-\pi) = 270 > 5$.
Step 4: Reject H_0 if $z > 2.326$.
Step 5:

$$z = \frac{\{(^{63}/_{300}) - 0.10\}}{\sqrt{\{^{0.10(0.90)}/_{300}\}}} = 6.35$$

Reject H_0.
Step 6: We conclude that the proportion of carpooling cars on the Turnpike is not 10%.

5. **a.** H_0: $\pi \geq 0.90$ $\quad H_1$: $\pi < 0.90$
b. H_0 is rejected if $z < -1.28$.
c. -2.67, found by $z = (0.82 - 0.90)/\sqrt{(0.90 \times 0.10)/100}$
d. H_0 is rejected. Fewer than 90% of the customers receive their orders in less than 10 minutes.

7. **a.** H_0 is rejected if $z > 1.65$.
b. 0.64, found by $p_c = \dfrac{70 + 90}{100 + 150}$
c. 1.61, found by

$$z = \frac{0.70 - 0.60}{\sqrt{[(0.64 \times 0.36)/100] + [(0.64 \times 0.36)/150]}}$$

d. H_0 is not rejected.

9. H_0: $\pi_1 = \pi_2$ $\quad H_1$: $\pi_1 \neq \pi_2$
H_0 is rejected if $z < -1.96$ or $z > 1.96$.

$$p_c = \frac{24 + 40}{400 + 400} = 0.08$$

-2.09, found by

$$z = \frac{0.06 - 0.10}{\sqrt{[(0.08 \times 0.92)/400] + [(0.08 \times 0.92)/400]}}$$

H_0 is rejected. The proportion infested is not the same in the two fields.

11. H_0: $\pi_d \leq \pi_r$ $\quad H_1$: $\pi_d > \pi_r$
H_0 is rejected if $z > 2.05$.

$$p_c = \frac{168 + 200}{800 + 1{,}000} = 0.2044$$

$$z = \frac{0.21 - 0.20}{\sqrt{\dfrac{(0.2044)(0.7956)}{800} + \dfrac{(0.2044)(0.7956)}{1{,}000}}} = 0.52$$

H_0 is not rejected. We cannot conclude that a larger proportion of Democrats favor lowering the standards. p-value = .3015.

13. **a.** 3
b. 7.815

15. **a.** Reject H_0 if $\chi^2 > 5.991$.
b. $\chi^2 = \dfrac{(10-20)^2}{20} + \dfrac{(20-20)^2}{20} + \dfrac{(30-20)^2}{20} = 10.0$
c. Reject H_0. The proportions are not equal.

17. H_0: The outcomes are the same; H_1: The outcomes are not the same. Reject H_0 if $\chi^2 > 9.236$.

$$\chi^2 = \frac{(3-5)^2}{5} + \cdots + \frac{(7-5)^2}{5} = 7.60$$

Do not reject H_0. Cannot reject H_0 that outcomes are the same.

19. H_0: There is no difference in the proportions.
H_1: There is a difference in the proportions.
Reject H_0 if $\chi^2 > 15.086$.

$$\chi^2 = \frac{(47-40)^2}{40} + \cdots + \frac{(34-40)^2}{40} = 3.400$$

Do not reject H_0. There is no difference in the proportions.

21. **a.** Reject H_0 if $\chi^2 > 9.210$.
b. $\chi^2 = \dfrac{(30-24)^2}{24} + \dfrac{(20-24)^2}{24} + \dfrac{(10-12)^2}{12} = 2.50$
c. Do not reject H_0.

23. H_0: Proportions are as stated; H_1: Proportions are not as stated. Reject H_0 if $\chi^2 > 11.345$.

$$\chi^2 = \frac{(50-25)^2}{25} + \cdots + \frac{(160-275)^2}{275} = 115.22$$

Reject H_0. The proportions are not as stated.

25. H_0: There is no relationship between community size and section read. H_1: There is a relationship.
Reject H_0 if $\chi^2 > 9.488$.

$$\chi^2 = \frac{(170-157.50)^2}{157.50} + \cdots + \frac{(88-83.62)^2}{83.62} = 7.340$$

Do not reject H_0. There is no relationship between community size and section read.

27. H_0: No relationship between error rates and item type.
H_1: There is a relationship between error rates and item type.
Reject H_0 if $\pi^2 > 9.21$.

$$\chi^2 = \frac{(20-14.1)^2}{14.1} + \cdots + \frac{(225-225.25)^2}{225.25} = 8.033$$

Do not reject H_0. There is not a relationship between error rates and item type.

29. **a.** H_0: $\pi = 0.50$ $\quad H_1$: $\pi \neq 0.50$
b. Yes. Both $n\pi$ and $n(1-\pi)$ are equal to 25 and exceed 5.
c. Reject H_0 if z is not between -2.576 and 2.576.
d. $z = \dfrac{\frac{37}{54} - 0.5}{\sqrt{0.5(1-0.5)/54}} = 2.72$

We reject the null hypothesis.
e. Using a p-value calculator (rounding to four decimal places) or a z-table, the p-value is 0.066, found by $2(0.5000 - 0.4967)$. The data indicate that the National Football Conference is luckier than the American Conference in calling the flip of a coin.

31. H_0: $\pi \leq 0.60$ $\quad H_1$: $\pi > 0.60$
H_0 is rejected if $z > 2.33$.

$$z = \frac{.70 - .60}{\sqrt{\dfrac{.60(.40)}{200}}} = 2.89$$

H_0 is rejected. Ms. Dennis is correct. More than 60% of the accounts are more than 3 months old.

33. H_0: $\pi \leq 0.44$ $\quad H_1$: $\pi > 0.44$
H_0 is rejected if $z > 1.65$.

$$z = \frac{0.480 - 0.44}{\sqrt{(0.44 \times 0.56)/1.000}} = 2.55$$

H_0 is rejected. We conclude that there has been an increase in the proportion of people wanting to go to Europe.

35. H_0: $\pi \leq 0.20$ $\quad H_1$: $\pi > 0.20$
H_0 is rejected if $z > 2.33$.

$$z = \frac{(56/200) - 0.20}{\sqrt{(0.20 \times 0.80)/200}} = 2.83$$

H_0 is rejected. More than 20% of the owners move during a particular year. p-value = $0.5000 - 0.4977 = 0.0023$.

37. H_0: $\pi \geq 0.0008$ $\quad H_1$: $\pi < 0.0008$
H_0 is rejected if $z < -1.645$.

$$z = \frac{0.0006 - 0.0008}{\sqrt{\dfrac{0.0008(0.9992)}{10{,}000}}} = -0.707 \quad H_0 \text{ is not rejected.}$$

These data do not prove there is a reduced fatality rate.

39. $H_0: \pi_1 \le \pi_2$ $H_1: \pi_1 > \pi_2$
If $z > 2.33$, reject H_0.

$$p_c = \frac{990 + 970}{1{,}500 + 1{,}600} = 0.63$$

$$z = \frac{.6600 - .60625}{\sqrt{\frac{.63(.37)}{1{,}500} + \frac{.63(.37)}{1{,}600}}} = 3.10$$

Reject the null hypothesis. We can conclude the proportion of men who believe the division is fair is greater.

41. $H_0: \pi_1 \le \pi_2$ $H_1: \pi_1 > \pi_2$ H_0 is rejected if $z > 1.65$.

$$p_c = \frac{.091 + .085}{2} = .088$$

$$z = \frac{0.091 - 0.085}{\sqrt{\frac{(0.088)(0.912)}{5{,}000} + \frac{(0.088)(0.912)}{5{,}000}}} = 1.059$$

H_0 is not rejected. There has not been an increase in the proportion calling conditions "good." The p-value is .1446, found by .5000 − .3554. The increase in the percentages will happen by chance in one out of every seven cases.

43. $H_0: \pi_1 = \pi_2$ $H_1: \pi_1 \ne \pi_2$
H_0 is rejected if z is not between −1.96 and 1.96.

$$p_c = \frac{100 + 36}{300 + 200} = .272$$

$$z = \frac{\frac{100}{300} - \frac{36}{200}}{\sqrt{\frac{(0.272)(0.728)}{300} + \frac{(0.272)(0.728)}{200}}} = 3.775$$

H_0 is rejected. There is a difference in the replies of the sexes.

45. $H_0: \pi_s = 0.50, \pi_r = \pi_e = 0.25$
H_1: Distribution is not as given above.
$df = 2$. Reject H_0 if $\chi^2 > 4.605$.

Turn	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
Straight	112	100	12	1.44
Right	48	50	−2	0.08
Left	40	50	−10	2.00
Total	200	200		3.52

H_0 is not rejected. The proportions are as given in the null hypothesis.

47. H_0: There is no preference with respect to TV stations.
H_1: There is a preference with respect to TV stations.
$df = 3 - 1 = 2$. H_0 is rejected if $\chi^2 > 5.991$

TV Station	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
WNAE	53	50	3	9	0.18
WRRN	64	50	14	196	3.92
WSPD	33	50	−17	289	5.78
	150	150	0		9.88

H_0 is rejected. There is a preference for TV stations.

49. $H_0: \pi_n = 0.21, \pi_m = 0.24, \pi_s = 0.35, \pi_w = 0.20$
H_1: The distribution is not as given.
Reject H_0 if $\chi^2 > 11.345$.

Region	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
Northeast	68	84	−16	3.0476
Midwest	104	96	8	0.6667
South	155	140	15	1.6071
West	73	80	−7	0.6125
Total	400	400	0	5.9339

H_0 is not rejected. The distribution of order destinations reflects the population.

51. H_0: The proportions are the same.
H_1: The proportions are not the same.
Reject H_0 if $\chi^2 > 16.919$.

f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
44	28	16	256	9.143
32	28	4	16	0.571
23	28	−5	25	0.893
27	28	−1	1	0.036
23	28	−5	25	0.893
24	28	−4	16	0.571
31	28	3	9	0.321
27	28	−1	1	0.036
28	28	0	0	0.000
21	28	−7	49	1.750
				14.214

Do not reject H_0. The digits are evenly distributed.

53. H_0: Gender and attitude toward the deficit are not related.
H_1: Gender and attitude toward the deficit are related.
Reject H_0 if $\chi^2 > 5.991$.

$$\chi^2 = \frac{(244 - 292.41)^2}{292.41} + \frac{(194 - 164.05)^2}{164.05} + \frac{(68 - 49.53)^2}{49.53} + \frac{(305 - 256.59)^2}{256.59} + \frac{(114 - 143.95)^2}{143.95} + \frac{(25 - 43.47)^2}{43.47} = 43.578$$

Since 43.578 > 5.991, you reject H_0. A person's position on the deficit is influenced by his or her gender.

55. H_0: Whether a claim is filed and age are not related.
H_1: Whether a claim is filed and age are related.
Reject H_0 if $\chi^2 > 7.815$.

$$\chi^2 = \frac{(170 - 203.33)^2}{203.33} + \cdots + \frac{(24 - 35.67)^2}{35.67} = 53.639$$

Reject H_0. Age is related to whether a claim is filed.

57. $H_0: \pi_{BL} = \pi_O = .23, \pi_Y = \pi_G = .15, \pi_{BR} = \pi_R = .12$.
H_1: The proportions are not as given. Reject H_0 if $\chi^2 > 15.086$.

Color	f_o	f_e	$(f_o - f_e)^2/f_e$
Blue	12	16.56	1.256
Brown	14	8.64	3.325
Yellow	13	10.80	0.448
Red	14	8.64	3.325
Orange	7	16.56	5.519
Green	12	10.80	0.133
Total	72		14.006

Do not reject H_0. The color distribution agrees with the manufacturer's information.

59. H_0: Salary and winning are not related.
H_1: Salary and winning are related.
Reject H_0 if $\chi^2 > 3.841$ with 1 degree of freedom.

	Salary		
Winning	Lower half	Top half	Total
No	8	6	14
Yes	7	9	16
Total	15	15	

$$\chi^2 = \frac{(8 - 7)^2}{7} + \frac{(6 - 7)^2}{7} + \frac{(7 - 8)^2}{8} + \frac{(9 - 8)^2}{8} = .536$$

Fail to reject H_0. Cannot conclude that salary and winning are related.

APPENDIX C:

Solutions to Practice Tests

PRACTICE TEST—CHAPTER 1

Part I

1. Statistics
2. Descriptive statistics
3. Statistical inference
4. Sample
5. Population
6. Nominal
7. Ratio
8. Ordinal
9. Interval
10. Discrete
11. Nominal
12. Nominal

Part II

1. **a.** 11.1
   b. About 3 to 1
   c. 65%
2. **a.** Ordinal
   b. 67.7%

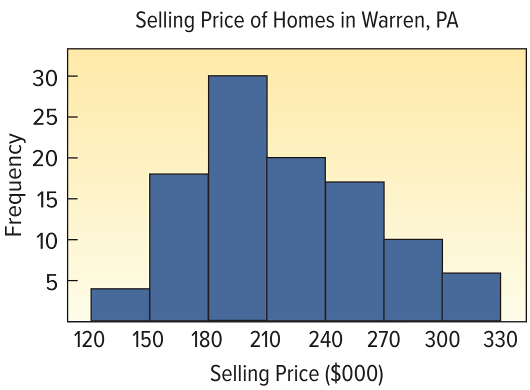
PRACTICE TEST—CHAPTER 2

Part I

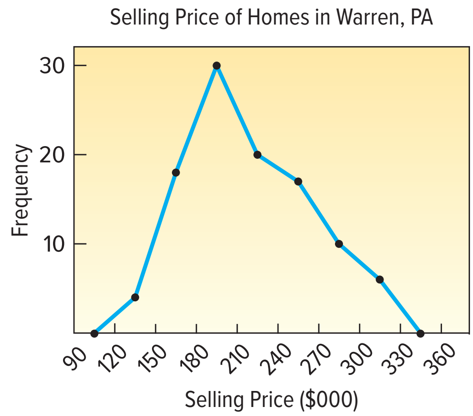
1. Frequency table
2. Frequency distribution
3. Bar chart
4. Pie chart
5. Histogram or frequency polygon
6. 7
7. Class interval
8. Midpoint
9. Total number of observations
10. Upper class limits

Part II

1. **a.** $30
   b. 105
   c. 52
   d. .19
   e. $165
   f. $120, $330
   g.



   h.



PRACTICE TEST—CHAPTER 3

Part I

1. Parameter
2. Statistic
3. Zero
4. Median
5. 50%
6. Mode
7. Range
8. Variance
9. Variance
10. Never
11. Median
12. Normal rule or empirical rule

Part II

1. **a.** $\bar{X} = \dfrac{560}{8} = 70$
   b. Median = 71.5
   c. Range = 80 − 52 = 28
   d. $s = \sqrt{\dfrac{610.0}{8-1}} = 9.335$
2. $\bar{X}_w = \dfrac{200(\$36) + 300(\$40) + 500(\$50)}{200 + 300 + 500} = \$44.20$
3. $-0.88 \pm 2(1.41)$
   -0.88 ± 2.82
   $-3.70, 1.94$

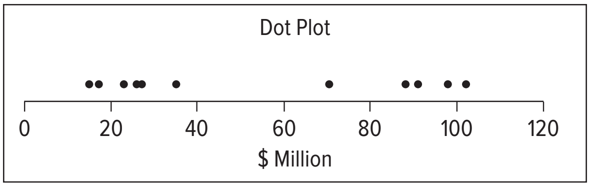
PRACTICE TEST—CHAPTER 4

Part I

1. Dot plot
2. Box plot
3. Scatter diagram
4. Contingency table
5. Quartile
6. Percentile
7. Skewness
8. First quartile
9. Interquartile range

Part II

1. **a.**



b. $L_{50} = (11 + 1)\dfrac{50}{100} = 6$

median = 35

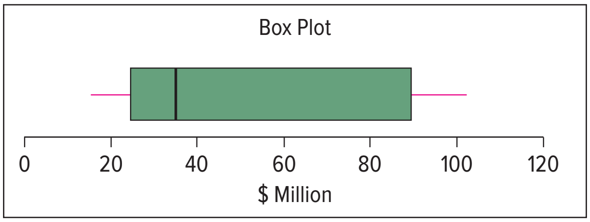
c. $L_{25} = (11 + 1)\dfrac{25}{100} = 3$

$Q_1 = 23$

d. $L_{75} = (11 + 1)\dfrac{75}{100} = 9$

$Q_3 = 91$

e.



2. **a.** $P(H) = \dfrac{144}{449} = 0.32$

b. $P(H| < 30) = \dfrac{21}{89} = 0.24$

c. $P(H| > 60) = \dfrac{75}{203} = 0.37$. Age is related to high blood pressure, because $P(H| > 60)$ is greater than $P(H| < 30)$.

PRACTICE TEST—CHAPTER 5

Part I

1. Probability
2. Experiment
3. Event
4. Relative frequency
5. Subjective
6. Classical
7. Mutually exclusive
8. Exhaustive
9. Mutually exclusive
10. Complement rule
11. Joint probability
12. Independent

Part II

1. **a.** $P(\text{Both}) = P(B_1) \cdot P(B_2|B_1)$

$$= \left(\frac{5}{20}\right)\left(\frac{4}{19}\right) = .0526$$

b. $P(\text{at least1}) = 1 - P(\text{neither})$

$$= 1 - \left(\frac{15}{20}\right)\left(\frac{14}{19}\right) = 1 - .5526 = .4474$$

2. $P(\text{at least1}) = P(\text{Jogs}) + P(\text{Bike}) - P(\text{Both})$

$= .30 + .20 - 12 = .38$

3. $X = 5! = 120$

PRACTICE TEST—CHAPTER 6

Part I

1. Probability distribution
2. Probability
3. One
4. Mean
5. Two
6. Never
7. Equal
8. π
9. .075
10. .183

Part II

1. **a.** Binomial

b. $P(x = 1) = {}_{16}C_1(.15)^1(.85)^{15} = (16)(.15)(.0874) = .210$

c. $P(x \geq 1) = 1 - P(x = 0) = 1 - {}_{16}C_0(.15)^0(.85)^{16} = .9257$

2. **a.** tPoisson

b. $P(x = 3) = \dfrac{3^3 e^{-3}}{3!} = \dfrac{27}{(6)(20.0855)} = .224$

c. $P(x = 0) = \dfrac{3^0 e^{-3}}{0!} = .050$

d. $P(x \geq 1) = 1 - P(x = 0) = 1 - .050 = .950$

3.

Exemptions x	Probability $P(x)$	$X \cdot P(x)$	$(x - 2.2)^2 \cdot P(x)$
1	0.2	0.2	0.288
2	0.5	1	0.02
3	0.2	0.6	0.128
4	0.1	0.4	0.324
		2.2	0.76

a. $\mu = 1(.2) + 2(.5) + 3(.2) + 4(.1) = 2.2$

b. $\sigma^2 = (1 - 2.2)^2(.2) + \cdots + (4 - 2.2)^2(.1) = 0.76$

PRACTICE TEST—CHAPTER 7

Part I

1. One
2. Infinite
3. Discrete
4. Always equal
5. Infinite
6. One
7. Any of these values
8. .2764
9. .9396
10. .0450

Part II

1. **a.** $z = \dfrac{2{,}000 - 1{,}600}{850} = .47$

$P(0 \leq z < .47) = .1808$

b. $z = \dfrac{900 - 1{,}600}{850} = -0.82$

$P(-0.82 \leq z \leq .47) = .2939 + .1808 = .4747$

c. $z = \dfrac{1{,}800 - 1{,}600}{850} = 0.24$

$P(0.24 \leq z \leq .47) = .1808 - .0948 = -.0860$

d. $1.65 = \dfrac{X - 1{,}600}{850}$

$X = 1{,}600 + 1.65(850) = \$3{,}002.50$

PRACTICE TEST—CHAPTER 8

Part I

1. Random sample
2. No size restriction
3. Strata
4. Sampling error
5. Sampling distribution of sample means
6. 120
7. Standard error of the mean
8. Always equal to
9. Decrease
10. Normal distribution

Part II

1. $z = \dfrac{11 - 12.2}{2.3/\sqrt{12}} = -1.81$

   $P(z < -1.81) = .5000 - .4649 = .0351$

PRACTICE TEST—CHAPTER 9

Part I

1. Point estimate
2. Confidence interval
3. Narrower than
4. Proportion
5. 95
6. Standard deviation
7. Binomial
8. Approaches the z-distribution
9. Population median
10. Population mean

Part II

1. **a.** Unknown
   b. 9.3 years
   c. $s_{\bar{x}} = \dfrac{2.0}{\sqrt{26}} = 0.392$
   d. $9.3 \pm (1.708)\dfrac{2.0}{\sqrt{26}}$

   9.3 ± 0.67
   $(8.63, 9.97)$
2. $n = (.27)(.73)\left(\dfrac{2.326}{.02}\right)^2 = 2.666$
3. $.64 \pm 1.96\sqrt{\dfrac{.64(.36)}{100}}$

   $.64 \pm .094$
   $[.546, .734]$

PRACTICE TEST—CHAPTER 10

Part I

1. Null hypothesis
2. Accept
3. Significant level
4. Test statistic
5. Critical
6. Two
7. Standard deviation (or variance)
8. p-value
9. Binomial
10. Five

Part II

1. $H_0: \mu \leq 90$, $H_1: \mu > 90$

   $df = 18 - 1 = 17$
   Reject H_0 if $t > 2.567$.

   $t = \dfrac{96 - 90}{12/\sqrt{18}} = 2.121$

   Do not reject H_0. We cannot conclude that the mean time in the park is more than 90 minutes.
2. $H_0: \mu \leq 9.75$ $H_1: \mu > 9.75$
   Reject H_0 if $z > 1.645$.
   Note σ is known, so z is used and we assume a .05 significance level.

   $z = \dfrac{9.85 - 9.75}{0.27/\sqrt{25}} = 1.852$

   Reject H_0. The mean weight is more than 9.75 ounces.
3. $H_0: \pi \geq 0.67$, $H_1: \pi < 0.67$
   Reject H_0 if $z < -1.645$.

   $$z = \frac{\dfrac{180}{300} - 0.67}{\sqrt{\dfrac{0.67(1 - 0.67)}{300}}} = -2.578$$

   Reject H_0. Less than .67 of the couples seek their mate's approval.

PRACTICE TEST—CHAPTER 11

Part I

1. Zero
2. z
3. Proportions
4. Population standard deviation
5. Differences
6. t-distribution
7. $n - 2$
8. Paired
9. Independent
10. Dependent sample

Part II

1. $H_0: \mu_y = \mu_h$; $H_1: \mu_y \neq \mu_h$
   $df = 14 + 12 - 2 = 24$
   Reject H_0 if $t < -2.064$ or $t > 2.064$.

   $$s_p^2 = \frac{(14 - 1)30^2 + (12 - 1)(40)^2}{14 + 12 - 2} = 1220.83$$

   $$t = \frac{837 - 797}{\sqrt{1220.83\left(\dfrac{1}{14} + \dfrac{1}{12}\right)}} = \frac{40.0}{13.7455} = 2.910$$

   Reject H_0. There is a difference in the mean miles traveled.
2. $H_0: \pi_E = \pi_T$ $H_1: \pi_E \neq \pi_T$
   Reject H_0 if $z < -1.96$ or $z > 1.96$.

   $$P_c = \frac{128 + 149}{300 + 400} = \frac{277}{700} = .396$$

   $$z = \frac{\dfrac{128}{300} - \dfrac{149}{400}}{\sqrt{\dfrac{.396(1 - .396)}{300} + \dfrac{.396(1 - .396)}{400}}} = \frac{.054}{.037} = 1.459$$

   Do not reject H_0. There is no difference on the proportion that liked the soap in the two cities.

PRACTICE TEST—CHAPTER 12

Part I

1. F-distribution
2. Positively skewed
3. Variances
4. Means
5. Population standard deviations
6. Error or residual
7. Equal
8. Degrees of freedom
9. Variances
10. Independent

Part II

1. $H_0: \sigma_h^2 = \sigma_y^2; H_1: \sigma_h^2 \neq \sigma_y^2$
   $df_y = 12 - 1 = 11 \quad df_h = 14 - 1 = 13$
   Reject H_0 if $F > 2.635$.
   $$F = \frac{(40)^2}{(30)^2} = 1.78$$
   Do not reject H_0. Cannot conclude there is a difference in the variation of the miles traveled.
2. **a.** 3
   b. 21
   c. 3.55
   d. $H_0: \mu_1 = \mu_2 = \mu_3$
   H_1: not all treatment means are the same.
   e. Reject H_0.
   f. The treatment means are not the same.

PRACTICE TEST—CHAPTER 13

Part I

1. Scatter diagram
2. −1 and 1
3. Less than zero
4. Coefficient of determination
5. *t*
6. Predicted or fitted
7. Sign
8. Large
9. Error
10. Independent

Part II

1. **a.** 25
   b. Shares of stock
   c. $\hat{Y} = 197.9229 + 24.9145X$
   d. Direct
   e. $r = \sqrt{\dfrac{152{,}399.0211}{208{,}333.1400}} = 0.855$
   f. $\hat{Y} = 197.9229 + 24.9145(10) = 447.0679$, or 447
   g. Increase almost 25
   h. $H_0: \beta \leq 0$
   $H_1: \beta > 0$
   Reject H_0 if $t > 1.71$.
   $$t = \frac{24.9145}{3.1473} = 7.916$$
   Reject H_0. There is a positive relationship between years and shares.

PRACTICE TEST—CHAPTER 14

Part I

1. Independent variables
2. Least squares
3. Mean square error
4. Independent variables
5. Independent variables
6. Different from zero
7. *F*-distribution
8. *t*-distribution
9. Linearity
10. Correlated
11. Multicollinearity
12. Dummy

Part II

1. **a.** Four
   b. $\hat{Y} = 70.06 + 0.42x_1 + 0.27x_2 + 0.75x_3 + 0.42x_4$
   c. $R^2 = \dfrac{1050.8}{1134.6} = 0.926$
   d. $s_{y.1234} = \sqrt{4.19} = 2.05$
   e. $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
   H_1: not all $\beta_i = 0$
   Reject H_0 if $F > 2.87$.
   $$F = \frac{262.70}{4.19} = 62.70$$
   Reject H_0. Not all the regression coefficients equal zero.
   f. $H_0: \beta_i = 0,\ H_1: \beta_i \neq 0$
   Reject H_0 if $t < -2.086$ or $t > 2.086$.

$\beta_1 = 0$	$\beta_2 = 0$	$\beta_3 = 0$	$\beta_4 = 0$
$\beta_1 \neq 0$	$\beta_2 \neq 0$	$\beta_3 \neq 0$	$\beta_4 \neq 0$
$t = 2.47$	$t = 1.29$	$t = 2.50$	$t = 6.00$
Reject H_0	Do not reject H_0	Reject H_0	Reject H_0

Conclusion. Drop variable 2 and retain the others.

PRACTICE TEST—CHAPTER 15

Part I

1. Nominal
2. No assumption
3. Can have negative values
4. 2
5. 6
6. Independent
7. 4
8. Are the same
9. 9.488
10. Degrees of freedom

Part II

1. H_0: There is no difference between the school district and census data.
   H_1: There is a difference between the school district and census data.
   Reject H_0 if $\chi^2 > 7.815$.
   $$\chi^2 = \frac{(120-130)^2}{130} + \frac{(40-40)^2}{40} + \frac{(30-20)^2}{20} + \frac{(10-10)^2}{10} = 5.77$$
   Do not reject H_0. There is no difference between the census and school district data.
2. H_0: Gender and book type are independent.
   H_1: Gender and book type are related.
   Reject H_0 if $\chi^2 > 5.991$.
   $$\chi^2 = \frac{(250-197.31)^2}{197.31} + \cdots + \frac{(200-187.5)^2}{187.5} = 54.842$$
   Reject H_0. Men and women read different types of books.