APPENDIX C: ANSWERS TO ODD-NUMBERED CHAPTER EXERCISES & REVIEW 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 MP3 players f. SAT scores i. Number of home computers	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 MP3 players i. Number of home computers	a. Salary

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

17. **a.**

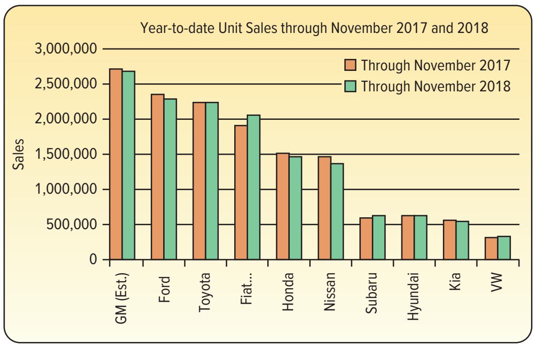
Manufacturer	Difference (units)
Fiat Chrysler	151,254
Tesla (Est.)	65,730
Subaru	30,980
Volvo	17,609
Land Rover	14,767
Mitsubishi	13,903
VW	12,622
Mazda	11,878
BMW	5,225

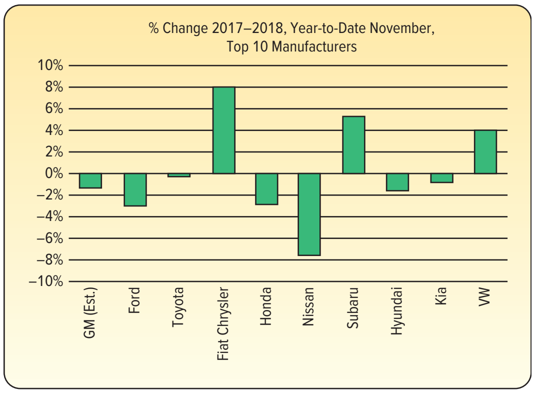
Manufacturer	Difference (units)
Porsche	1,609
Audi	1,024
MINI	(1,607)
Others	(1,650)
smart	(1,751)
Kia	(4,384)
Toyota	(5,771)
Jaguar	(9,159)
Hyundai	(9,736)
Mercedes (includes Sprinter)	(14,978)
General Motors (Est.)	(36,925)
Honda	(42,399)
Ford	(68,700)
Nissan	(110,081)

b. Percentage differences with top five and bottom five.

Manufacturer	% Change from 2017
Tesla (Est.)	163.0%
Volvo	24.5%
Land Rover	22.1%
Mitsubishi	14.6%
Fiat Chrysler	8.0%
Subaru	5.3%
Mazda	4.5%
VW	4.1%
Porsche	3.1%
BMW	1.9%
Audi	0.5%
Toyota	−0.3%
Kia	−0.8%
GM (Est.)	−1.4%
Hyundai	−1.6%
Honda	−2.8%
Ford	−2.9%
MINI	−3.8%
Mercedes (includes Sprinter)	−4.5%
Nissan	−7.6%
Others	−8.7%
Jaguar	−25.3%
smart	−60.3%

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.

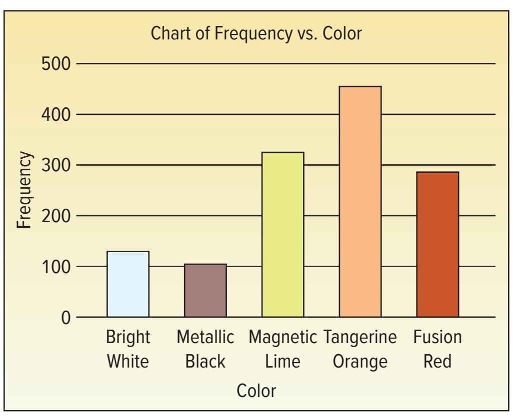
3.

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

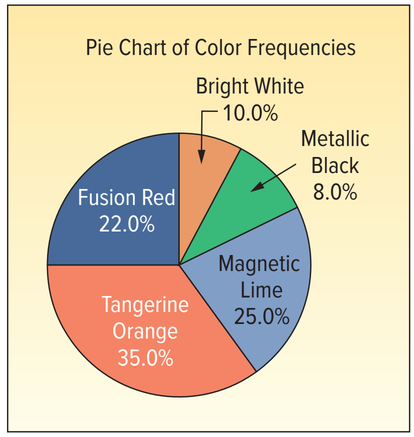
5. **a.** A frequency table.

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

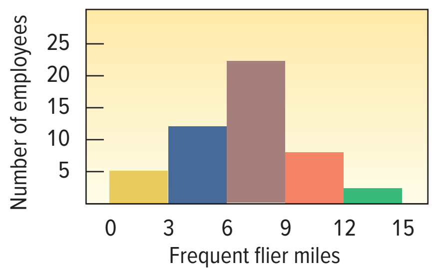
b. The largest group of shoppers (21) shop at the BiLo Supermarket 3, 4, or 5 times during a month period. Some customers visit the store only 1 time during the month, but others shop as many as 15 times.

c.

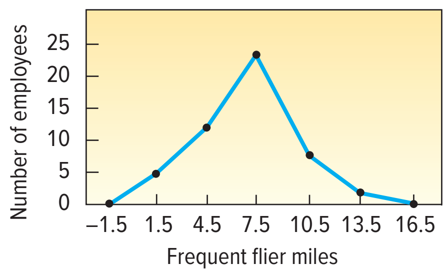
Number of Visits	Percent of Total
0 up to 3	17.65
3 up to 6	41.18
6 up to 9	25.49
9 up to 12	7.84
12 up to 15	5.88
15 up to 18	1.96
Total	100.00

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.



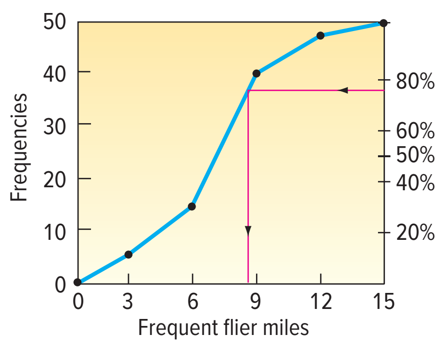
f. For the 50 employees, about half traveled between 6,000 and 9,000 miles. Five employees traveled less than 3,000 miles, and 2 traveled more than 12,000 miles.

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

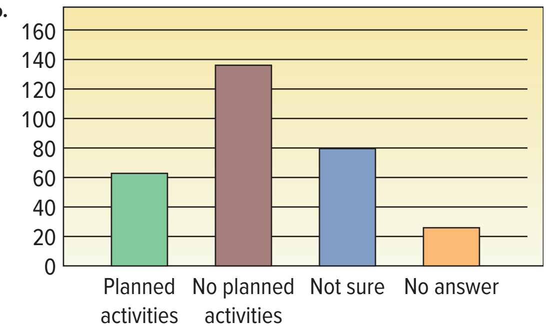
c.



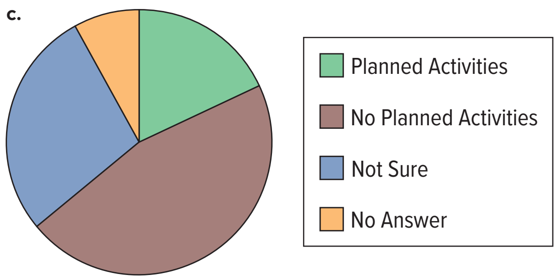
d. About 8.7 thousand miles

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. $i \geq \frac{48 - 16}{5} = 6.4$ Use interval of 7.
c. 15
d.

Class		Frequency
15 up to 22	III	3
22 up to 29	~~IIII~~ III	8
29 up to 36	~~IIII~~ II	7
36 up to 43	~~IIII~~	5
43 up to 50	II	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. $i = \frac{10 - 1}{6} = 1.5$ 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 Calls	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.

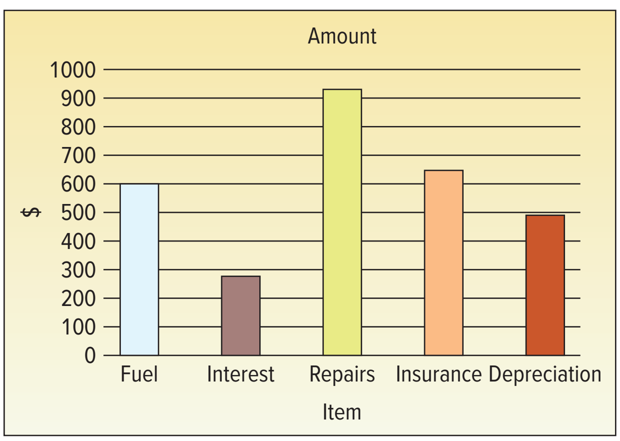
35. **a.** 56
b. 10 (found by 60 – 50)
c. 55
d. 17

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

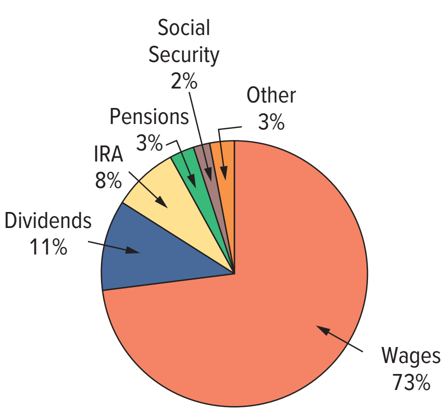
\$ 70 up to \$105	4
105 up to 140	17
140 up to 175	14
175 up to 210	2
210 up to 245	6
245 up to 280	1

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.



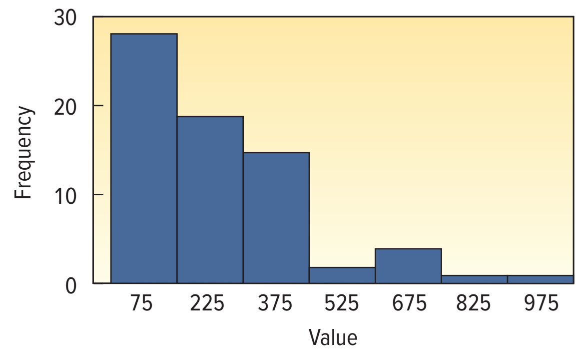
41.



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 has 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.

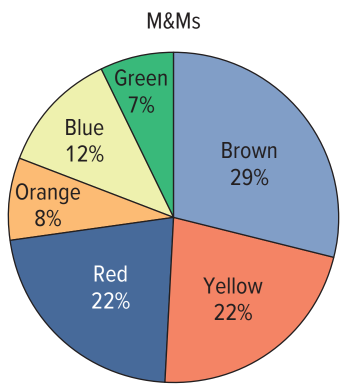


45. **a.** Pie chart
b. 700, found by 0.7(1,000)
c. Yes, 0.70 + 0.20 = 0.90

47. **a.**

b. 34.9%, found by (84.6 + 62.3)/420.9
c. 69.3% found by (84.6 + 62.3)/(84.6 + 62.3 + 32.4 + 18.6 + 14.1))

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!

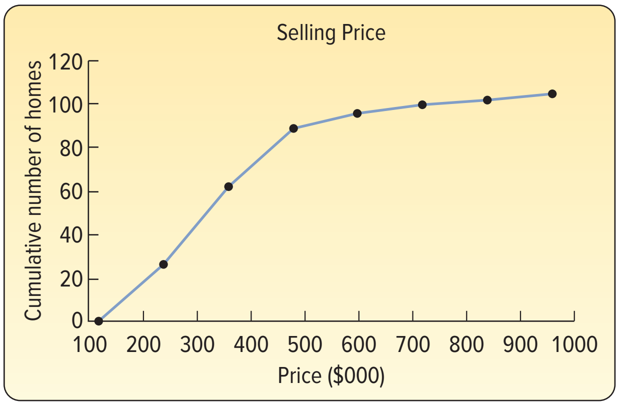
i > = (919,480 – 167,962)/7 = 107,360. Use intervals of 120,000.

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.

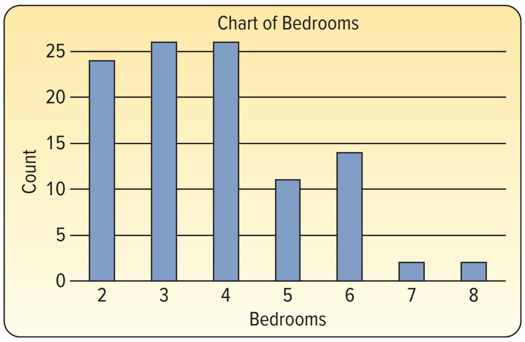
b. The typical price in the first class is $180,000 and in the last class it is $900,000

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.

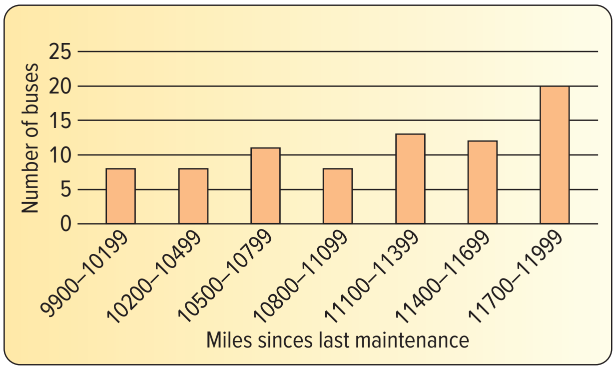


Two-, three-, and four-bedroom houses are most common with about 25 houses each. Seven- and eight-bedroom houses are rather rare.

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:

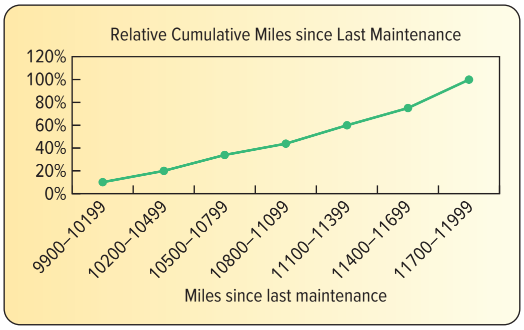
Class	*f*
9,900 up to 10,200	8
10,200 up to 10,500	8
10,500 up to 10,800	11
10,800 up to 11,100	8
11,110 up to 11,400	13
11,400 up to 11,700	12
11,700 up to 12,000	20

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.



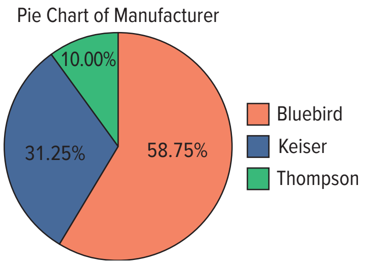
b. The distribution is somewhat "skewed" with a longer "tail" to the left and no outliers.

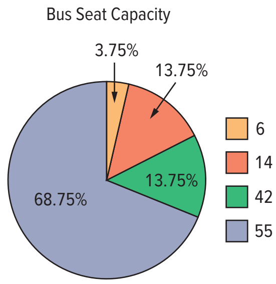
c.



Forty percent of the buses were driven fewer than about 10800 miles. About 30% of the 80 buses (about 24) were driven less than 10500 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. 12.8%, found by $\sqrt[5]{(1.08)(1.12)(1.14)(1.26)(1.05)} = 1.128$

29. 12.28% increase, found by
$\sqrt[5]{(1.094)(1.138)(1.117)(1.119)(1.147)} = 1.1228$

31. 1.60%, found by $\sqrt[7]{\dfrac{239.051}{213.967}} - 1$

33. In 2017, 2.28% found by $\sqrt[6]{\dfrac{265.9}{232.2}} - 1$
In 2020, 1.34% found by $\sqrt[3]{\dfrac{276.7}{265.9}} - 1$
The annual percent increase of subscribers is forecast to increase over the next 3 years.

35. **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.

37. **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.

39.

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.

41. **a.** 5
b. 4.4, found by
$$\frac{(8-5)^2 + (3-5)^2 + (7-5)^2 + (3-5)^2 + (4-5)^2}{5}$$

43. **a.** $2.77
b. 1.26, found by
$$\frac{(2.68-2.77)^2 + (1.03-2.77)^2 + (2.26-2.77)^2 + (4.30-2.77)^2 + (3.58-2.77)^2}{5}$$

45. **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).

47. **a.** $\bar{x} = 4$
$$s^2 = \frac{(7-4)^2 + \cdots + (3-4)^2}{5-1} = \frac{22}{5-1} = 5.5$$
b. $s = 2.3452$

49. **a.** $\bar{x} = 38$
$$s^2 = \frac{(28-38)^2 + \cdots + (42-38)^2}{10-1}$$
$$= \frac{744}{10-1} = 82.667$$
b. $s = 9.0921$

51. **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. $s = \sqrt{123.66} = 11.12$

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

55. **a.** About 95%
b. 47.5%, 2.5%

57. Because the exact values in a frequency distribution are not known, the midpoint is used for every member of that class.

59.

Class	f	M	fM	$(M - \bar{x})$	$f(M - \bar{x})^2$
20 up to 30	7	25	175	−22.29	3,477.909
30 up to 40	12	35	420	−12.29	1,812.529
40 up to 50	21	45	945	−2.29	110.126
50 up to 60	18	55	990	7.71	1,069.994
60 up to 70	12	65	780	17.71	3,763.729
	70		3,310		10,234.287

$$\bar{x} = \frac{3{,}310}{70} = 47.29$$
$$s = \sqrt{\frac{10{,}234.287}{70-1}} = 12.18$$

61.

Number of Clients	f	M	fM	$(M - \bar{x})$	$f(M - \bar{x})^2$
20 up to 30	1	25	25	−19.8	392.04
30 up to 40	15	35	525	−9.8	1,440.60
40 up to 50	22	45	990	0.2	0.88
50 up to 60	8	55	440	10.2	832.32
60 up to 70	4	65	260	20.2	1,632.16
	50		2,240		4,298.00

$\bar{x} = \frac{2,240}{50} = 44.8$ $\qquad s = \sqrt{\frac{4,298}{50 - 1}} = 9.37$

63. **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. $\Sigma(x - \mu) = (6 - 5) + (4 - 5) + (3 - 5) + (7 - 5) + (5 - 5) = 0$

65. $\bar{x} = \frac{545}{16} = 34.06$
Median = 37.50

67. 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.

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

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

73. $GM = \sqrt[52]{\frac{5000000}{42000}} - 1$, so about 9.63%

75. **a.** 55, found by 72 − 17
b. 17.6245, found by the square root of 2795.6/9

77. **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.

79. **a.** There were 13 flights, so all items are considered.
b. $\mu = \frac{2,259}{13} = 173.77$
c. Range = 301 − 7 = 294
$s = \sqrt{\frac{133,846}{13}} = 101.47$

81. **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)

83. **a.** $\bar{x} = \frac{273}{30} = 9.1$, Median = 9
b. Range = 18 − 4 = 14
$s = \sqrt{\frac{368.7}{30 - 1}} = 3.57$
c. $2^5 = 32$, so suggest 5 classes $i = \frac{18 - 4}{5} = 2.8$ use $i = 3$

Class	M	f	fM	$M - \bar{x}$	$(M - \bar{x})^2$	$f(M - \bar{x})^2$
3.5 up to 6.5	5	10	50	−4	16	160
6.5 up to 9.5	8	6	48	−1	1	6
9.5 up to 12.5	11	9	99	2	4	36
12.5 up to 15.5	14	4	56	5	25	100
15.5 up to 18.5	17	1	17	8	64	64
			270			366

d. $\bar{x} = \frac{270}{30} = 9.0$
$s = \sqrt{\frac{366}{30 - 1}} = 3.552$
The mean and standard deviation from grouped data are estimates of the mean and standard deviations of the actual values.

85. $\bar{x} = 13 = \frac{910}{70}$
$s = 5.228 = \sqrt{1807.5/69}$

87. **a.** 1. 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.
2. 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 $; found by $139,174,000 plus or minus 2($41,101,000).
b. 4.10% per year, found by $\sqrt[18]{\frac{4,100,000}{1,990,000}} - 1 = 1,04097 = 4.10\%$

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 $\frac{3(30.8 - 31.0)}{3.96}$
c.

Salary	$\left(\frac{x - \bar{x}}{s}\right)$	$\left(\frac{x - \bar{x}}{s}\right)^3$
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

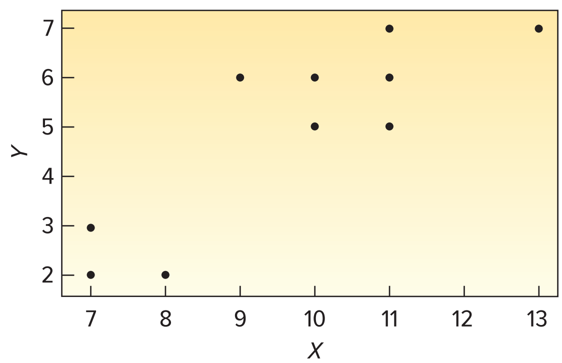
15. **a.** The mean is 21.93, found by 328.9/15. The median is 15.8, and the standard deviation is 21.18, found by

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

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

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

$L_{25} = (21)(.25) = 5.25$
$Q_1 = 66.6 + .25(72.9 - 66.6) = 68.175$
$L_{75} = 21(.75) = 15.75$
$Q_3 = 87.1 + .75(90.2 - 87.1) = 89.425$

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

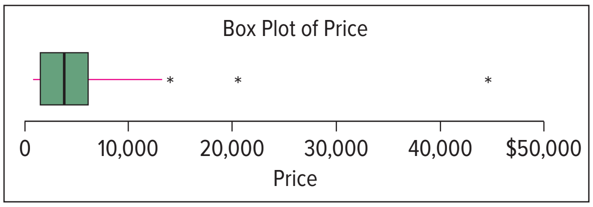
```
         ---------------------
------I              +       I-----------------------
         ---------------------

------+---------+---------+---------+---------+---------+
     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.

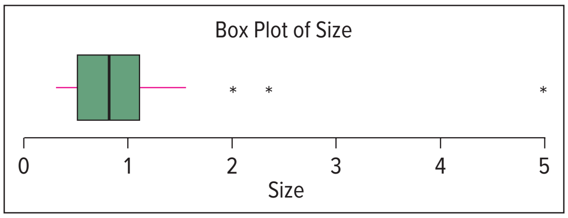
27. The distribution is positively skewed. The first quartile is about \$20 and the third quartile is about \$90. There is one outlier located at \$255. The median is about \$50.

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).

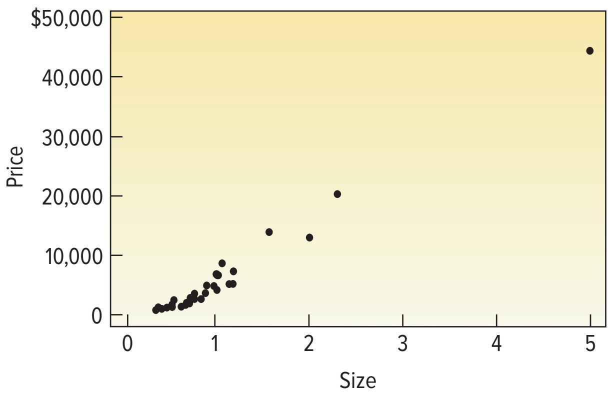
b.



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.

Scatter Plot of Price versus Size



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

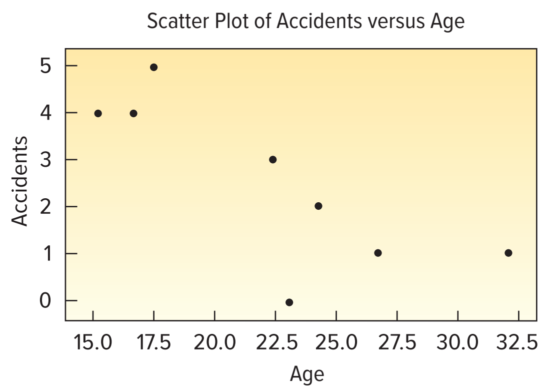
d.

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 $sk = \frac{3(7.7143 - 8.0)}{3.9036} = -0.22$

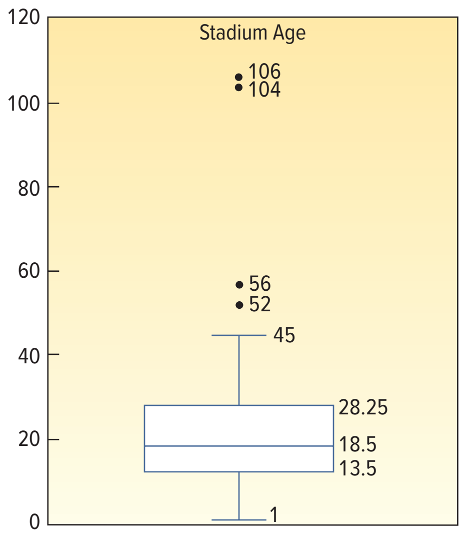
33.



As age increases, the number of accidents decreases.

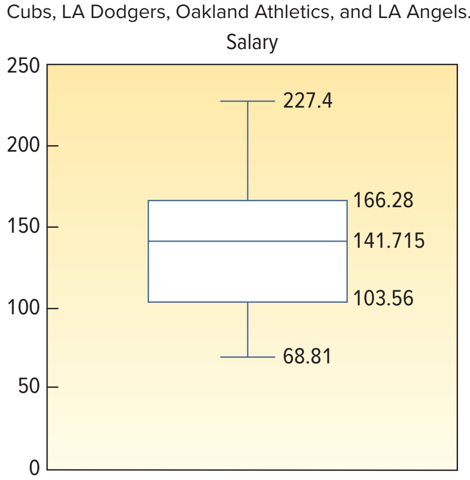
35. **a.** 139,340,000
b. 5.4% unemployed, found by (7,523/139,340)100
c. Men = 5.64%
Women = 5.12%

37. **a.** Box plot of age assuming the current year is 2018.



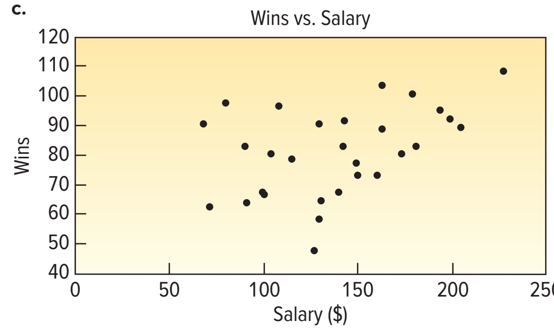
Distribution of stadium is highly positively skewed to the right. Any stadium older than 50.375 years (Q3 + 1.5(Q3–Q1) = 28.25 + 1.5(28.25–13.5) is an outlier. Boston, Chicago Cubs, LA Dodgers, Oakland Athletics, and LA Angels.

b.



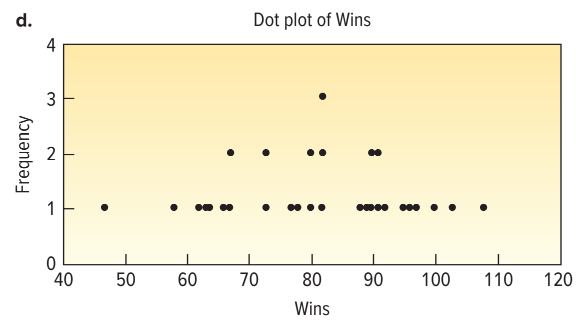
The first quartile is $103.56 million and the third is $166.28 million. Outliers are greater than Q3 + 1.5(Q3–Q1) or 166.28 + 1.5*(166.28–103.56) = $260.36 million. The distribution is positively skewed. However in 2018, there were no outliers.

c.



The correlation coefficient is 0.43. The relationship is generally positive but the relationship is generally weak. Higher salaries are not strongly associated with more wins.

d.



The dot plot shows a range of wins from the high 40s to the 100s. Most teams appear to win between 65 and 90 games in a season. 13 teams won 90 or more games. 9 teams won less than 70 games. That leaves 16 teams that won between 70 and 90 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.

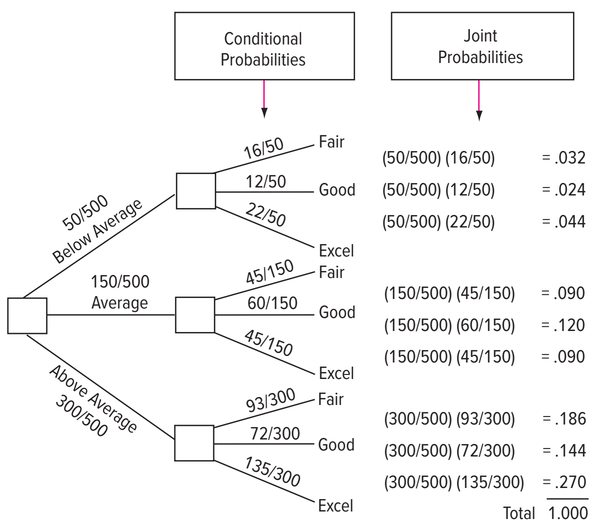
21. Let A denote the event the fish is green and B be the event the fish is male.
a. $P(A) = 80/140 = 0.5714$
b. $P(B) = 60/140 = 0.4286$
c. $P(A \text{ and } B) = 36/140 = 0.2571$
d. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 80/140 + 60/140 - 36/140 = 104/140 = 0.7429$

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.
$$P(A_1 \mid B_1) = \frac{P(A_1) \times P(B_1 \mid A_1)}{P(A_1) \times P(B_1 \mid A_1) + P(A_2) \times P(B_1 \mid A_2)}$$
$$= \frac{.60 \times .05}{(.60 \times .05) + (.40 \times .10)} = .4286$$

35.
$$P(\text{night} \mid \text{win}) = \frac{P(\text{night})P(\text{win} \mid \text{night})}{P(\text{night})P(\text{win} \mid \text{night}) + P(\text{day})P(\text{win} \mid \text{day})}$$
$$= \frac{(.70)(.50)}{[(.70)(.50)] + [(.30)(.90)]} = .5645$$

37. $P(\text{cash} \mid > \$50)$
$$= \frac{P(\text{cash})\, P(> \$50 \mid \text{cash})}{[P(\text{cash})\, P(> \$50 \mid \text{cash}) + P(\text{credit})\, P(> \$50 \mid \text{credit}) + P(\text{debit})\, P(> \$50 \mid \text{debit})]}$$
$$= \frac{(.30)(.20)}{(.30)(.20) + (.30)(.90) + (.40)(.60)} = .1053$$

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

41. 210, found by (10)(9)(8)(7)/(4)(3)(2)

43. 120, found by 5!

45. (4)(8)(3) = 96 combinations

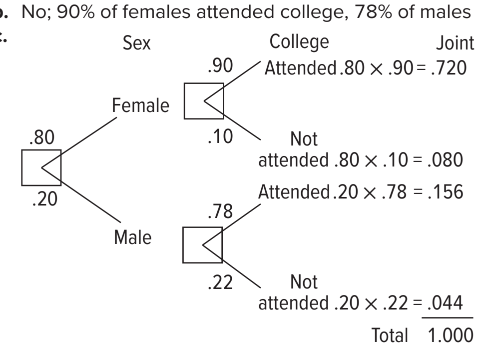
47. **a.** Asking teenagers to compare their reactions to a newly developed soft drink.
b. Answers will vary. One possibility is more than half of the respondents like it.

49. Subjective

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

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

55. **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.

57. **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

59. **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

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

63. **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

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

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

69. **a.** $P(\text{not perfect}) = P(\text{bad sector}) + P(\text{defective})$
$$= \frac{112}{1{,}000} + \frac{31}{1{,}000} = .143$$
b. $P(\text{defective} \mid \text{not perfect}) = \dfrac{.031}{.143} = .217$

71.
$$P(\text{poor} \mid \text{profit}) = \frac{.10(.20)}{.10(.20) + .60(.80) + .30(.60)} = .0294$$

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

75. Yes, 256 is found by 2^8.
77. .9744, found by $1 - (.40)^4$
79. a. 0.193, found by .15 + .05 − .0075 = .193
 b. .0075, found by (.15)(.05)
81. a. P(F 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
83. $26^4 = 456{,}976$
85. 1/3, 628,800
87. a. $P(D) = .20(.03) + .30(.04) + .25(.07) + .25(.065) = .05175$
 b. $$P(\text{Tyson} \mid \text{defective}) = \frac{.20(.03)}{[.20(.03) + .30(.04) + .25(.07) + .25(.065)]} = .1159$$

Supplier	Joint	Revised
Tyson	.00600	.1159
Fuji	.01200	.2319
Kirkpatricks	.01750	.3382
Parts	.01625	.3140
	.05175	1.0000

89. 0.512, found by $(0.8)^3$
91. .525, found by $1 - (.78)^3$
93. a.

Wins	# Teams
40–49	1
50–59	1
60–69	6
70–79	4
80–89	7
90–99	8
100–109	3
Grand Total	**30**

 1. 11/30 = 0.37
 2. 10/11 = 0.91
 3. Winning 90 or more games in a season does not guarantee a place in the end-ofseason playoffs.

 b.

	Frequency (# teams) by League		
Home Runs	**American**	**National**	**Grand Total**
120–149	1	2	3
150–179	4	7	11
180–209	5	3	8
210–239	4	3	7
240–269	1		1
Grand Total	**15**	**15**	**30**

Home Runs	Relative Frequency		
	American	**National**	**Grand Total**
120–149	6.67%	13.33%	10.00%
150–179	26.67%	46.67%	36.67%
180–209	33.33%	20.00%	26.67%
210–239	26.67%	20.00%	23.33%
240–269	6.67%	0.00%	3.33%
Grand Total	**100.00%**	**100.00%**	**100.00%**

 1. In the American League, the probability that a team hits 180 or more homeruns if 0.67.
 2. In the National League, the probability that a team hits 180 or more homeruns is 0.40.
 3. There is clear difference in the distribution of homeruns between the American and National Leagues. There are many potential reasons for the difference. One of the reasons may be the use of a designated hitter.

CHAPTER 6

1. Mean = 1.3, variance = .81, found by:
 $\mu = 0(.20) + 1(.40) + 2(.30) + 3(.10) = 1.3$
 $\sigma^2 = (0 - 1.3)^2(.2) + (1 - 1.3)^2(.4) + (2 - 1.3)^2(.3) + (3 - 1.3)^2(.1) = .81$
3. Mean = 14.5, variance = 27.25, found by:
 $\mu = 5(.1) + 10(.3) + 15(.2) + 20(.4) = 14.5$
 $\sigma^2 = (5 - 14.5)^2(.1) + (10 - 14.5)^2(.3) + (15 - 14.5)^2(.2) + (20 - 14.5)^2(.4) = 27.25$
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	0.2020
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. $P(x \leq 2) = 0.058 + 0.198 + 0.296 = 0.552$

c. 0.448, found by $P(x \geq 3) = 1 - P(x \leq 2) = 1 - 0.552$

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. Given N = 12, 7 boys and 5 girls.

$P(\text{3 boys on a team of 5}) = \frac{({}_7C_3)({}_5C_2)}{({}_{12}C_5)} = .4419$

$P(\text{2 girls on a team of 5}) = \frac{({}_5C_2)({}_7C_3)}{({}_{12}C_5)} = .4419$

Using the multiplication rule, the probability is (.4419) (.4419) = .1953

b. $P(\text{5 boys on a team of 5}) = \frac{({}_7C_5)({}_5C_0)}{({}_{12}C_5)} = 0.027$

c. Using the complement rule: P(1 or more girls) = 1 – P(0 girls on a team of 5)

$= 1 - \frac{({}_5C_0)({}_7C_5)}{({}_{12}C_5)} = 1 - 0.027 = 0.973$

27. N is 10, the number of loans in the population; S is 3, the number of underwater loans in the population; x is 0, the number of selected underwater loans in the sample; and n is 2, the size of the sample. Use formula (6–6) to find

$$P(0) = \frac{({}_7C_2)({}_3C_0)}{{}_{10}C_2} = \frac{21(1)}{45} = 0.4667$$

29. $P(2) = \frac{[{}_9C_3][{}_6C_2]}{[{}_{15}C_5]} = \frac{84(15)}{3003} = .4196$

31. a. .6703
b. .3297

33. a. .0613
b. .0803

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

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

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

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

43. a. 5 10(.35) = 3.5

b. $P(x = 4) = {}_{10}C_4 (.35)^4 (.65)^6 = 210(.0150)(.0754) = .2375$

c. $P(x \geq 4) = {}_{10}C_x (.35)^x (.65)^{10-x}$
$= 2375 + .1536 + \cdots + .0000 = .4862$

45. 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

47. 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

49. 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

51. 0.2784, found by 0.1472 + 0.0811 + 0.0348 + 0.0116 + 0.0030 + 0.0006 + 0.0001 + 0.0000

53. 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. $\mu = 12(0.52) = 6.24$ $\sigma = \sqrt{12(0.52)(0.48)} = 1.7307$

c. 0.1768

d. 0.3343, found by 0.0002 + 0.0019 + 0.0116 + 0.0418 + 0.1020 + 0.1768

55. a. $P(1) = \frac{[{}_7C_2][{}_3C_1]}{[{}_{10}C_3]} = \frac{(21)(3)}{120} = .5250$

b. $P(0) = \frac{[{}_7C_3][{}_3C_0]}{[{}_{10}C_3]} = \frac{(35)(1)}{120} = .2917$
$P(x \geq 1) = 1 - P(0) = 1 - .2917 = .7083$

57. $P(x = 0) = \frac{[{}_8C_4][{}_4C_0]}{[{}_{12}C_4]} = \frac{70}{495} = .141$

59. a. .0498 **b.** .7746, found by $(1 - .0498)^5$

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

63. 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

65. $\mu = n\pi = 23\left(\frac{2}{113}\right) = .407$

$P(2) = \frac{(.407)^2 e^{-.407}}{2!} = 0.0551$

$P(0) = \frac{(.407)^0 e^{-.407}}{0!} = 0.6656$

67. Let $\mu = n\pi = 155(1/3{,}709) = 0.042$

$$P(4) = \frac{0.042^4 e^{-0.042}}{4!} = 0.00000012 \quad \text{Very unlikely!}$$

69. **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 indicates that Zoot leases vehicles at a lower rate than the national average.

71. The mean number of home runs per game is 2.2984. The average season home runs per team is 186.167. Then 186.167/162 × 2 = 2.2984.

a. $P(x = 0) = \frac{\mu^0 e^{-2.2984}}{0!} = .1004$

b. $P(x = 2) = \frac{\mu^2 e^{-2.2984}}{2!} = .2652$

c. $P(X \geq 4) = 0.2004$, found by
$1 - P(X < 4) = (.1004 + .2308 + .2652 + .2032) = .7996$

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 ± 1(10)

b. 480 and 520, found by 500 ± 2(10)

c. 470 and 530, found by 500 ± 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 $z = \frac{25 - 20}{4.0} = 1.25$

b. 0.3944, found in Appendix B.3

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

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

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

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 $z = \frac{20.00 - 29.81}{9.31} = -1.05$

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.** 0.3935, found by $1 - e^{[(-1/60)(30)]}$

b. 0.1353, found by $e^{[(-1/60)(120)]}$

c. 0.1859, found by $e^{[(-1/60)(45)]} - e^{[(-1/60)(75)]}$

d. 41.59 seconds, found by −60 ln(0.5)

33. **a.** 0.5654, found by $1 - e^{[(-1/18)(15)]}$,
and 0.2212, found by $1 - e^{[(-1/60)(15)]}$

b. 0.0013, found by $e^{[(-1/18)(120)]}$, and 0.1353, found by $e^{[(-1/60)(120)]}$

c. 0.1821, found by $e^{[(-1/18)(30)]} - e^{[(-1/18)(90)]}$,
and 0.3834, found by $e^{[(-1/60)(30)]} - e^{[(-1/60)(90)]}$

d. 4 minutes, found by −18 ln(0.8), and 13.4 minutes, found by −60 ln(0.8)

35. **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.

37. **a.** $\mu = \frac{11.96 + 12.05}{2} = 12.005$

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

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

d. $P(x > 11.98) = \frac{1}{(12.05 - 11.96)} \left(\frac{12.05 - 11.98}{1}\right)$
$= \frac{.07}{.09} = .78$

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

39. **a.** $\mu = \frac{4 + 10}{2} = 7$

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

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

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

41. 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.

43. **a.** 0.4015, z for 900 is: $\frac{900 - 1{,}054.5}{120} = -1.29$. 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: $\frac{900 - 1{,}054.5}{120} = -1.29$, z for 1200 is: $\frac{1{,}200 - 1{,}054.5}{120} = 1.21$. Adding the two corresponding probabilities, 0.4015 + 0.3869 = .7884.
d. 0.2279; z for 900 is: $\frac{900 - 1{,}054.5}{120} = -1.29$, z for 1000 is: $\frac{1{,}000 - 1{,}054.5}{120} = -0.45$. Subtracting the two corresponding probabilities, 0.4015 − 0.1736 = .2279.

45. **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)

47. **a.** 0.0968, z for 300 is: $\frac{300 - 270}{23} = 1.30$. Using the z-table, probability is .4032. Subtracting from 0.5, 0.5000 − 0.4032 = 0.0968.
b. 0.9850, z for 220 is: $\frac{220 - 270}{23} = -2.17$. 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.

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

51. **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 virutally 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.

53. **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.

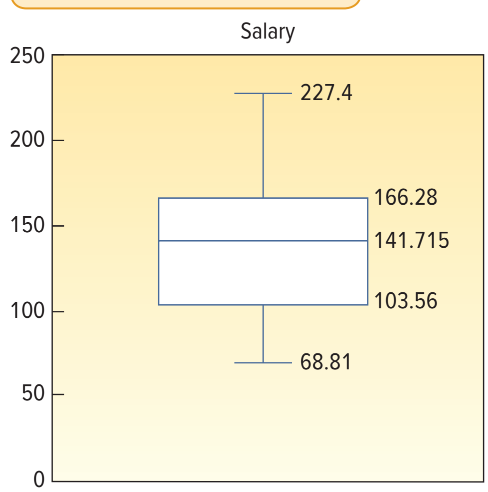
55. The z-score associated with $50,000 is 8.25: (50,000 − 33,500)/2000. That is, $50,000 is 8.25 standard deviations above the mean salary. Conclusion: The probability that someone in the same business has a salary of $50,000 is zero. This salary would be exceptionally unusual.

57. **a.** 0.4262, found by $1 - e^{[(-1/27)(15)]}$
b. 0.1084, found by $e^{[(-1/27)(60)]}$
c. 0.1403, found by $e^{[(-1/27)(30)]} - e^{[(-1/27)(45)]}$
d. 2.84 secs, found by −27 ln(0.9)

59. **a.** 0.2835, found by $1 - e^{[(-1/300{,}000)(100{,}000)]}$
b. 0.1889, found by $e^{[(-1/300{,}000)(500{,}000)]}$
c. 0.2020, found by $e^{[(-1/300{,}000)(200{,}000)]} - e^{[(-1/300{,}000)(350{,}000)]}$
d. Both the mean and standard deviation are 300,000 hours.

61. **a.**

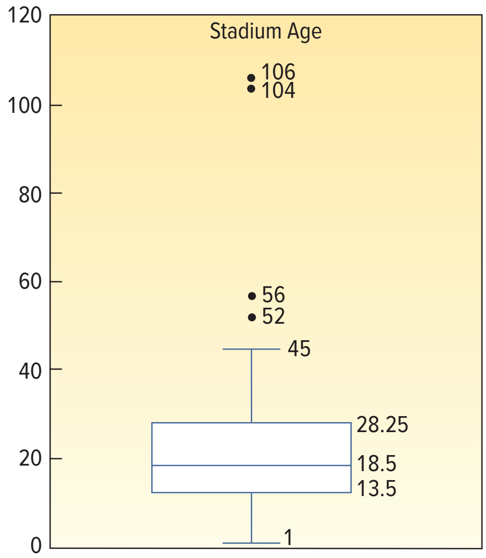
Salary ($ mil)	
Mean	139.17
Median	141.72
Population standard deviation	40.41
Skewness	0.17
Range	158.59
Minimum	68.81
Maximum	227.40



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

b.

Stadium Age	
Mean	27.4
Median	18.5
Population standard deviation	24.7
Skewness	2.2
Range	105.0
Minimum	1.0
Maximum	106.0



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 Mercury
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.**

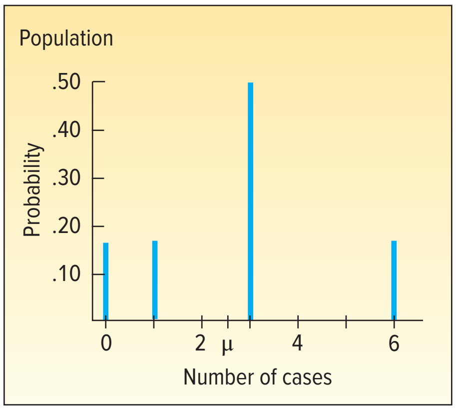
Sample	Values	Sum	Mean
1	12, 12, 14	38	12.66
2	12, 12, 15	39	13.00
3	12, 12, 20	44	14.66
4	14, 15, 20	49	16.33
5	12, 14, 15	41	13.66
6	12, 14, 15	41	13.66
7	12, 15, 20	47	15.66
8	12, 15, 20	47	15.66
9	12, 14, 20	46	15.33
10	12, 14, 20	46	15.33

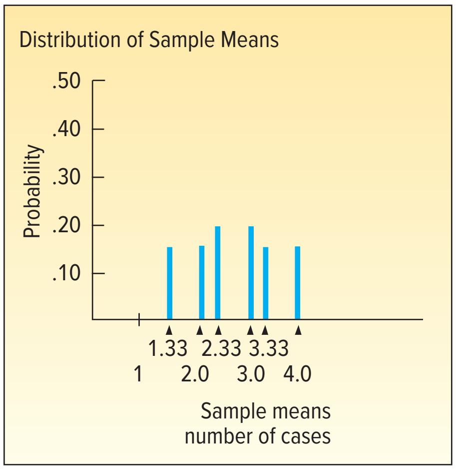
b. $\mu_{\bar{x}} = \frac{(12.66 + \cdots + 15.33 + 15.33)}{10} = 14.6$
$\mu = (12 + 12 + 14 + 15 + 20)/5 = 14.6$
c. The dispersion of the population is greater than that of the sample means. The sample means vary from 12.66 to 16.33, whereas the population varies from 12 to 20.

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 $\frac{53.33}{20}$
$\mu = 2.67$, found by $(3 + 6 + 3 + 3 + 0 + 1)/6$.
They are equal.
d.

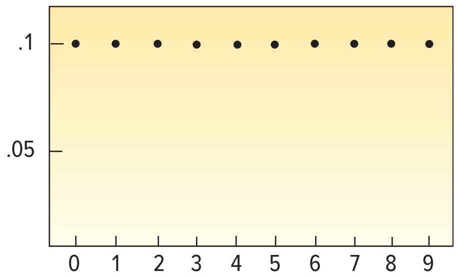




Sample Mean	Number of Means	Probability
1.33	3	.1500
2.00	3	.1500
2.33	4	.2000
3.00	4	.2000
3.33	3	.1500
4.00	3	.1500
	20	1.0000

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.

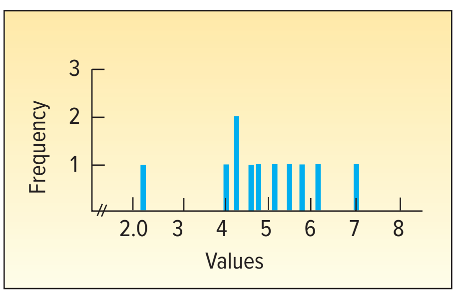
11. **a.**



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

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.

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.

15. **a.** $z = \frac{63 - 60}{12/\sqrt{9}} = 0.75$. So probability is 0.2266, found by 0.5000 – 0.2734.

b. $z = \frac{56 - 60}{12/\sqrt{9}} = -1$. So the probability is 0.1587, found by 0.5000 – 0.3413

c. 0.6147, found by 0.3413 + 0.2734

17. $z = \frac{1{,}950 - 2{,}200}{250/\sqrt{50}} = -7.07$ $p = 1$, 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 + ⋯ + 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 + ⋯ + 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 sores however, this student has a 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. $\sigma_{\bar{x}} = \dfrac{5.5}{\sqrt{25}} = 1.1$

c. $z = \dfrac{36 - 35}{5.5/\sqrt{25}} = 0.91$

$p = 0.1814$, found by $0.5000 - 0.3186$

d. $z = \dfrac{34.5 - 35}{5.5/\sqrt{25}} = -0.45$

$p = 0.6736$, 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. $z = \dfrac{34 - 35}{5.5/\sqrt{25}} = -0.91$; $z = \dfrac{36 - 35}{5.5/\sqrt{25}} = 0.91$. Subtracting: $0.5 - .3186 = .1814$ in each tail. Multiplying by 2, the final probability is .3628.

33. $z = \dfrac{\$335 - \$350}{\$45/\sqrt{40}} = -2.11$

$p = 0.9826$, found by $0.5000 + 0.4826$

35. $z = \dfrac{29.3 - 29}{2.5/\sqrt{60}} = 0.93$

$p = 0.8238$, found by $0.5000 + 0.3238$

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

39. $z = \dfrac{900 - 947}{205/\sqrt{60}} = -1.78$

$p = 0.0375$, 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.** $z = \dfrac{600 - 510}{14.28/\sqrt{10}} = 19.9$, $P = 0.00$, or virtually never

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

$p = 0.4864 + 0.5000 = 0.9864$

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

$p = 0.5000 - 0.4864 = 0.0136$

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

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

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

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

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

d. .0162, found by $.5000 - .4838$

47. Mean 2018 attendance was 2.322 million. Likelihood of a sample mean this large or larger is .1611, found by $0.5000 - .3389$, where $z = \dfrac{2.322 - 2.45}{\dfrac{0.71}{\sqrt{30}}} = -0.99$.

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 Fox TV should schedule the new comedy show.

19. 97, found by $n = \left(\dfrac{1.96 \times 10}{2}\right)^2 = 96.04$

21. 196, found by $n = 0.15(0.85)\left(\dfrac{1.96}{0.05}\right)^2 = 195.9216$

23. 554, found by $n = \left(\dfrac{1.96 \times 3}{0.25}\right)^2 = 553.19$

25. **a.** 577, found by $n = 0.60(0.40)\left(\dfrac{1.96}{0.04}\right)^2 = 576.24$

b. 601, found by $n = 0.50(0.50)\left(\dfrac{1.96}{0.04}\right)^2 = 600.25$

27. 33.41 and 36.59, found by

$$35 \pm 2.030\left(\frac{5}{\sqrt{36}}\right)\sqrt{\frac{300-36}{300-1}}$$

29. 1.683 and 2.037, found by

$$1.86 \pm 2.680\left(\frac{0.5}{\sqrt{50}}\right)\sqrt{\frac{400-50}{400-1}}$$

31. 6.13 years to 6.87 years, found by $6.5 \pm 1.989(1.7/\sqrt{85})$

33. **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 $= 2.426\left(\dfrac{50}{\sqrt{40}}\right) = 19.18$

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.

35. 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.

37. 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.

39. 61.47, rounded to 62. Found by solving for n in the equation: $1.96(16/\sqrt{n}) = 4$

41. 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.

43. 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.

45. a. 89.467, found by 1342/15, 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.

47. 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.

49. a. Between 0.648 and 0.752, found by $.7 \pm 2.58\left(\sqrt{\frac{0.7(1-0.7)}{500}}\right)\left(\sqrt{\frac{20{,}000-500}{20{,}000-1}}\right)$
b. Based on this sample we would confirm Ms. Miller will receive a majority of the votes as the lower limit of the confidence interval is above 0.500.

51. a. Margin of error $= 2.032\left(\frac{4.50}{\sqrt{35}}\right)\sqrt{\frac{(500-35)}{500-1}} = \$1.49$
b. $52.51 and $55.49, found by $\$54.00 \pm 2.032\frac{\$4.50}{\sqrt{35}}\sqrt{\frac{(500-35)}{500-1}}$

53. 369, found by $n = 0.60(1-0.60)(1.96/0.05)^2$

55. 97, found by $[(1.96 \times 500)/100]^2$

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

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

61. 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 $n = \{(1.96 \times 24)/10\}^2 = 22.13$

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

65. a. Between 0.156 and 0.184, found by $0.17 \pm 1.96\sqrt{\frac{(0.17)(1-0.17)}{2700}}$
b. Yes, because 18% are inside the confidence interval.
c. 21,682; found by $0.17(1-0.17)[1.96/0.005]^2$

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

69. 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.

71. 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. $H_0: \mu \le 40$ $H_1: \mu > 40$
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. $H_0: \mu \le 50{,}000$ $H_1: \mu > 50{,}000$
Reject H_0 if $t > 1.833$.

$$t = \frac{(60000 - 50000)}{(10000/\sqrt{10})} = 3.16$$

Reject H_0 and conclude that the mean income in Wilmington is greater than \$50,000.

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

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

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

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. $H_0: \mu \le 1.4$ $H_1: \mu > 1.4$
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. $H_0: \mu \le 67$ $H_1: \mu > 67$
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. 1.05, found by $z = (9{,}992 - 9{,}880)/(400/\sqrt{100})$. Then 0.5000 − 0.3531 = 0.1469, which is the probability of a Type II error.

23. $H_0: \mu \ge 60$ $H_1: \mu < 60$
Reject H_0 if $z < -1.282$; the critical value is 59.29.

$$z = \frac{58 - 60}{(2.7/\sqrt{24})} = -3.629$$

Reject H_0. The mean assembly time is less than 60 minutes.
Using the sample mean, 58, as μ_1, the *z*-score for 59.29 is 2.34. So the probability for values between 58 and 59.29 is .4904. The Type II error is the area to the right of 59.29 or .5000 − .4904 = .0096.

25. $H_0: \mu = \$45{,}000$ $H_1: \mu \ne \$45{,}000$
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.

27. $H_0: \mu \ge 10$ $H_1: \mu < 10$
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.

29. $H_0: \mu \ge 7.0$ $H_1: \mu < 7.0$
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.

31. $H_0: \mu \ge 3.13$ $H_1: \mu < 3.13$
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.

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

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

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

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 rhe mean interest paid is greater than \$6,658.

35. $H_0: \mu = 3.1$ $H_1: \mu \ne 3.1$ 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.

37. $H_0: \mu \ge 6.5$ $H_1: \mu < 6.5$ 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.

39. $H_0: \mu = 0$ $H_1: \mu \ne 0$
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.

41. $H_0: \mu \le 100$ $H_1: \mu > 100$ 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.

43. 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.

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

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

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

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.

47. **a.** $9.00 \pm 1.645(1/\sqrt{36}) = 9.00 \pm 0.274$.
So the limits are 8.726 and 9.274.

b. $z = \frac{8.726 - 8.6}{1/\sqrt{36}} = 0.756$.

$P(z < 0.756) = 0.5000 + 0.2764 = .7764$

c. $z = \frac{9.274 - 9.6}{1/\sqrt{36}} = -1.956$.

$P(z > -1.96) = 0.4750 + 0.5000 = .9750$

49. $50 + 2.33\frac{10}{\sqrt{n}} = 55 - .525\frac{10}{\sqrt{n}} \qquad n = (5.71)^2 = 32.6$

Let $n = 33$

51. 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.

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

$$t = \frac{139.17 - 100}{41.1/\sqrt{30}} = 5.22$$

Using a p-value calculator or statistical software, the p-value is .000014.
Reject the null. The mean salary is probably not $100.0 million.

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

$$t = \frac{2.3224 - 2.0}{.7420/\sqrt{30}} = 2.38$$

Using a p-value calculator or statistical software, p-value is .0121.
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: $\mu_{\text{married}} \neq \mu_{\text{unmarried}}$
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: $\mu_{\text{Pitchers}} \neq \mu_{\text{Position Players}}$
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 $df = 20 + 16 - 2 = 34$ Reject H_0 if t is not between −2.728 and 2.728.

$$s_p^2 = \frac{(20 - 1)(8.218)^2 + (16 - 1)(6.002)^2}{20 + 16 + 2} = 53.633$$

$$t = \frac{4.953 - 4.306}{\sqrt{53.633\left(\frac{1}{20} + \frac{1}{16}\right)}} = .2634$$

Using a p-value calculator or statistical software, the p-value is .7938.
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.** $$df = \frac{\left(\frac{697{,}225}{16} + \frac{2{,}387{,}025}{18}\right)^2}{\frac{\left(\frac{697{,}225}{16}\right)^2}{16 - 1} + \frac{\left(\frac{2{,}387{,}025}{18}\right)^2}{18 - 1}} = 26.7 \rightarrow 26df$$

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

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. Reject H_0 if $t > 2.353$.

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

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

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

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:** $\bar{d} = -25.917$

$$s_d = 40.791 \qquad t = \frac{-25.917}{40.791/\sqrt{12}} = -2.201$$

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: $\mu_{\text{Men}} \neq \mu_{\text{Women}}$
Reject H_0 if $t < -2.645$ or $t > 2.645$.

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

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

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

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 $$\frac{\left(\frac{9200^2}{40} + \frac{7100^2}{30}\right)^2}{\frac{\left(\frac{9200^2}{40}\right)^2}{39} + \frac{\left(\frac{7100^2}{30}\right)^2}{29}} = 67.9$$

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

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 $$\frac{\left(\frac{0.56^2}{12} + \frac{0.3^2}{12}\right)^2}{\frac{\left(\frac{0.56^2}{12}\right)^2}{11} + \frac{\left(\frac{0.3^2}{12}\right)^2}{11}} = 16.8$$

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

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 $$\frac{\left(\frac{10.5^2}{10} + \frac{14.25^2}{12}\right)^2}{\frac{\left(\frac{10.5^2}{10}\right)^2}{9} + \frac{\left(\frac{14.25^2}{12}\right)^2}{11}} = 19.8$$

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

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 $$\frac{\left(\frac{2.33^2}{10} + \frac{2.55^2}{14}\right)^2}{\frac{\left(\frac{2.33^2}{10}\right)^2}{9} + \frac{\left(\frac{2.55^2}{14}\right)^2}{13}} = 20.6$$

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

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

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

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

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: $\mu_{\text{Reduced}} \le \mu_{\text{Regular}}$ H_1: $\mu_{\text{Reduced}} > \mu_{\text{Regular}}$
Reject H_0 if $t > 2.650$.

b. $\bar{X}_1 = 125.125$ $s_1 = 15.094$ $\bar{X}_2 = 117.714$ $s_2 = 19.914$

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

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

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 does not provide evidence that the reduced price increased sales.

37. a. H_0: $\mu_{\text{Before}} - \mu_{\text{After}} = \mu_d \le 0$ H_1: $\mu_d > 0$
Reject H_0 if $t > 1.895$.

b. $\bar{d} = 1.75$ $s_d = 2.9155$ $t = \frac{1.75}{2.9155/\sqrt{8}} = 1.698$

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: $\mu_1 = \mu_2$ H_1: $\mu_1 \ne \mu_2$
Reject H_0 if $t < -2.024$ or $t > 2.024$.

b. $$s_p^2 = \frac{(15-1)(40000)^2 + (25-1)(30000)^2}{15+25-2} = 1{,}157{,}894{,}737$$

$$t = \frac{150000 - 180000}{\sqrt{1{,}157{,}894{,}737\left(\frac{1}{15} + \frac{1}{25}\right)}} = -2.699$$

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 indicates that the population means are different.

41. a. H_0: $\mu_{\text{Before}} - \mu_{\text{After}} = \mu_d \le 0$ H_1: $\mu_d > 0$
Reject H_0 if $t > 1.895$.

b. $\bar{d} = 3.113$ $s_d = 2.911$ $t = \frac{3.113}{2.911/\sqrt{8}} = 3.025$

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: $\mu_{\text{Ocean Drive}} = \mu_{\text{Rio Rancho}}$; 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: $\mu_{\text{US 17}} - \mu_{\text{SC 707}} = \mu_d \le 0$ H_1: $\mu_d > 0$
Reject H_0 if $t > 1.711$.

b. $\bar{d} = 2.8$ $s_d = 6.589$ $t = \frac{2.8}{6.589/\sqrt{25}} = 2.125$

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. a. 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: $\sigma_1^2 = \sigma_2^2$ 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 $F = \frac{(12)^2}{(10)^2} = 1.44$
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: $\mu_{Southwyck} = \mu_{Franklin} = \mu_{Old\ Orchard}$ 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 tracks of land.

11. **a.** *Ho:* $\mu_1 = \mu_2 = \mu_3$ 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. **a.** H_0: $\mu_1 = \mu_2$ H_1: Not all treatment means are equal.
b. Reject H_0 if $F > 18.5$.
c. H_0: $\mu_A = \mu_B = \mu_C$ H_1: Not all block means are equal
Reject H_0 if $F > 19.0$
d. $SSTotal = (46.0 - 36.5)^2 + \cdots (35.0 - 36.5)^2 = 289.5$

$$SST = 3\left(\left(42.33 - {}^{219}/_{6}\right)^2\right) + 3\left(\left(30.67 - {}^{219}/_{6}\right)^2\right) = 204.167$$

$$SSB = 2\left(\left(38.5 - {}^{219}/_{6}\right)^2\right) + 2\left(\left(31.5 - {}^{219}/_{6}\right)^2\right) + 2\left(\left(39.5 - {}^{219}/_{6}\right)^2\right) = 76.00$$

$SSE = SSTotal - SST - SSB = 289.5 - 204.1667 - 76 = 9.333$

e.

Source	SS	df	MS	F	p-value
Treatment	204.167	1	204.167	43.75	0.0221
Blocks	76.000	2	38.000	8.14	0.1094
Error	9.333	2	4.667		
Total	289.5000	5			

f. The *F*-statistic is significant: $43.75 > 18.5$; *p*-value is less then .05. so reject H_0. There is a difference in the treatment means: $8.14 < 19.0$. For the blocks, $8.14 < 19.0$; *p*-value is more than .05, so fail to reject H_0 for blocks. There is no difference between blocks.

17. **a.** For treatment
H_0: $\mu_{Day} = \mu_{Afternoon} = \mu_{Night}$
H_1: Not all means equal
b. Reject if $F > 4.46$.
c. For blocks: H_0: $\mu_S = \mu_L = \mu_C = \mu_T = \mu_M$ H_1: Not all means are equal. Reject if $F > 3.84$.
d. $SSTotal = (31 - 433/15)^2 + \cdots + (27 - 433/15)^2 = 139.73$

$$SST = 5\left(\left(30 - {}^{433}/_{15}\right)^2\right) + 5\left(\left(26 - {}^{433}/_{15}\right)^2\right) + 5\left(\left(30.6 - {}^{433}/_{15}\right)^2\right) = 62.53$$

$$SSB = 3\left(\left(30.33 - {}^{433}/_{15}\right)^2\right) + 3\left(\left(30.67 - {}^{433}/_{15}\right)^2\right) + 3\left(\left(27.3 - {}^{433}/_{15}\right)^2\right) + 3\left(\left(29 - {}^{433}/_{15}\right)^2\right) + 3\left(\left(27 - {}^{433}/_{15}\right)^2\right) = 33.73$$

$SSE = (SSTotal - SST - SSB) = 139.73 - 62.53 - 33.73 = 43.47$

e. Here is the ANOVA table:

Source	SS	df	MS	F	p-value
Treatment	62.53	2	31.2667	5.75	.0283
Blocks	33.73	4	8.4333	1.55	.2767
Error	43.47	8	5.4333		
Total	139.73	14			

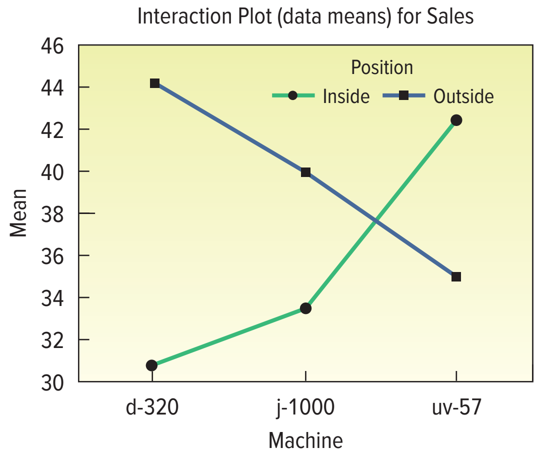
f. As $5.75 > 4.46$ the null for treatments is rejected, but the null for blocks is not rejected as $1.55 < 3.84$. There is a difference in means by shifts, but not by employee.

19.

Source	SS	df	MS	F	P
Size	156.333	2	78.1667	1.98	0.180
Weight	98.000	1	98.000	2.48	0.141
Interaction	36.333	2	18.1667	0.46	0.642
Error	473.333	12	39.444		
Total	764.000	17			

a. Since the *p*-value (0.18) is greater than 0.05, there is no difference in the Size means.
b. The *p*-value for Weight (0.141) is also greater than 0.05. Thus, there is no difference in those means.
c. There is no significant interaction because the *p*-value (0.642) is greater than 0.05.

21. a.



Yes, there appears to be an interaction effect. Sales are different based on machine position, either in the inside or outside position.

b.

Two-way ANOVA: Sales versus Position, Machine					
Source	***df***	**SS**	**MS**	***F***	***P***
Position	1	104.167	104.167	9.12	0.007
Machine	2	16.333	8.167	0.72	0.502
Interaction	2	457.333	228.667	20.03	0.000
Error	18	205.500	11.417		
Total	23	783.333			

The position and the interaction of position and machine effects are significant. The effect of machine on sales is not significant.

c.

One-way ANOVA: D-320 Sales versus Position					
Source	***df***	**SS**	**MS**	***F***	***P***
Position	1	364.50	364.50	40.88	0.001
Error	6	53.50	8.92		
Total	7	418.00			

One-way ANOVA: J-1000 Sales versus Position					
Source	***df***	**SS**	**MS**	***F***	***P***
Position	1	84.5	84.5	5.83	0.052
Error	6	87.0	14.5		
Total	7	171.5			

One-way ANOVA: UV-57 Sales versus Position					
Source	***df***	**SS**	**MS**	***F***	***P***
Position	1	112.5	112.5	10.38	0.018
Error	6	65.0	10.8		
Total	7	177.5			

Recommendations using the statistical results and mean sales plotted in part (a): Position the D-320 machine outside. Statistically, the position of the J-1000 does not matter. Position the UV-57 machine inside.

23. 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 oceanfront homes.

25. 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.

27. 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.

29. a. H_0: $\mu_{Discount} = \mu_{Variety} = \mu_{Department}$ 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.

31. a. H_0: $\mu_{Rec\ Center} = \mu_{Key\ Street} = \mu_{Monclova} = \mu_{Whitehouse}$ 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.

33. a. H_0: $\mu_{Lecture} = \mu_{Distance}$ H_1: $\mu_{Lecture} \neq \mu_{Distance}$
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. $t = \dfrac{37 - 45}{\sqrt{9.5\left(\dfrac{1}{6} + \dfrac{1}{8}\right)}} - 4.806$

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.

35. **a.** H_0: $\mu_{Compact} = \mu_{Midsize} = \mu_{Large}$ 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.

37. 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.

39. For color, the critical value of *F* is 4.76; for size, it is 5.14.

Source	SS	df	MS	F
Treatment	25.0	3	8.3333	5.88
Blocks	21.5	2	10.75	7.59
Error	8.5	6	1.4167	
Total	55.0	11		

H_0s for both treatment and blocks (color and size) are rejected. At least one mean differs for color and at least one mean differs for size.

41. **a.** Critical value of *F* is 3.49. Computed *F* is 0.668. Do not reject H_0.

b. Critical value of *F* is 3.26. Computed *F* value is 100.204. Reject H_0 for block means.

There is a difference in homes but not assessors.

43. For gasoline:

H_0: $\mu_1 = \mu_2 = \mu_3$; H_1: Mean mileage is not the same.

Reject H_0 if $F > 3.89$.

For automobile:

H_0: $\mu_1 = \mu_2 = \ldots = \mu_7$; H_1: Mean mileage is not the same.

Reject H_0 if $F > 3.00$.

ANOVA Table

Source	SS	df	MS	F
Gasoline	44.095	2	22.048	26.71
Autos	77.238	6	12.873	15.60
Error	9.905	12	0.825	
Total	131.238	20		

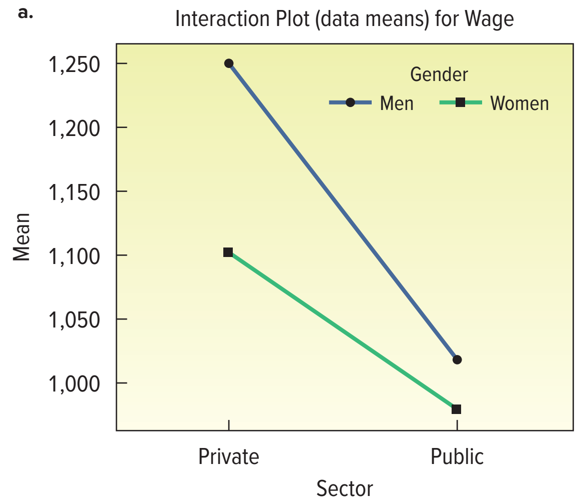
There is a difference in both autos and gasoline.

45. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$; H_1: The treatment means are not equal. Reject H_0 if $F > 2.37$.

Source	SS	df	MS	F
Treatment	0.03478	5	0.00696	3.86
Error	0.10439	58	0.0018	
Total	0.13917	63		

H_0 is rejected. There is a difference in the mean weight of the colors.

47. **a.**



b. Two-way ANOVA: Wage versus Gender, Sector

```
Source        DF      SS       MS      F      P
Gender         1   44086    44086  11.44  0.004
Sector         1  156468   156468  40.61  0.000
Interaction    1   14851    14851   3.85  0.067
Error         16   61640     3853
Total         19  277046
```

There is no interaction effect of gender and sector on wages. However, there are significant differences in mean wages based on gender and significant differences in mean wages based on sector.

c. One-way ANOVA: Wage versus Sector

```
Source    DF      SS       MS      F       P
Sector     1  156468   156468  23.36   0.000
Error     18  120578     6699
Total     19  277046

s = 81.85   R-Sq = 56.48%   R-Sq(adj) = 54.06%
```

```
One-way ANOVA: Wage versus Gender
Source    DF      SS      MS     F      P
Gender     1   44086   44086  3.41  0.081
Error     18  232960   12942
Total     19  277046

s = 113.8  R-Sq = 15.91%  R-Sq(adj) = 11.24%
```

d. The statistical results show that only sector, private or public, has a significant effect on the wages of accountants.

49. **a.** H_0: $\sigma_p^2 = \sigma_{np}^2$ 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: $\sigma_g^2 = \sigma_{ng}^2$ 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: $\mu_c = \mu_l = \mu_m = \mu_p = \mu_r$ 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).

51. a. H_0: $\mu_B = \mu_K = \mu_T$ 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 is not different.

b. H_0: $\mu_B = \mu_K = \mu_T$ 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 is not different. P-value = 0.4828.

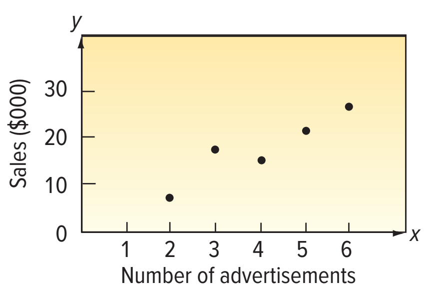
CHAPTER 13

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.



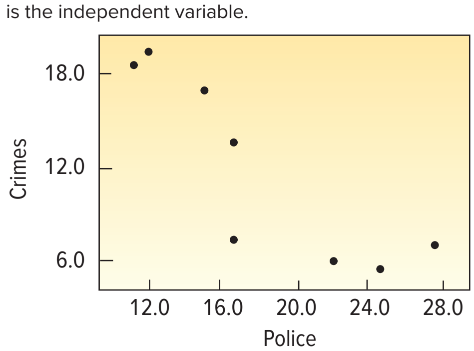
c. $\Sigma(x - \bar{x})(y - \bar{y}) = 36$, $n = 5$, $s_x = 1.5811$, $s_y = 6.1237$

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

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: $\rho \le 0$; 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: $\rho \le 0$ 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. $\Sigma(x - \bar{x})(y - \bar{y}) = 44.6$, $s_x = 2.726$, $s_y = 2.011$

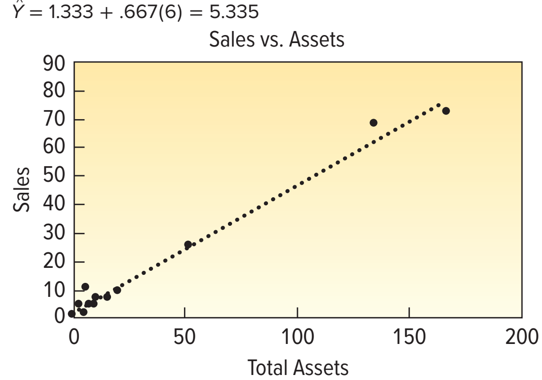
$$r = \frac{44.6}{(10-1)(2.726)(2.011)} = .904$$

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

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

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

17. a.



b. Computing correlation in Excel, $r = .9916$

c.

	Total Assets	12-Month Sales
Mean	36.1038	17.8088
Standard deviation	55.6121	25.2208
Count	12	12

$$b = .9916\frac{25.2208}{55.6121} = .4497;\ a = 17.8088 - .4497(36.1038)$$
$= 1.5729$

d. $\hat{Y} = 1.5729 + .4497(100.0) = 451.2729$ (\$ 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. $H_0: \beta \geq 0 \quad H_1: \beta < 0 \quad df = n - 2 = 8 - 2 = 6$
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. $H_0: \beta = 0 \quad H_1: \beta \neq 0 \quad df = n - 2 = 12 - 2 = 10$
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.

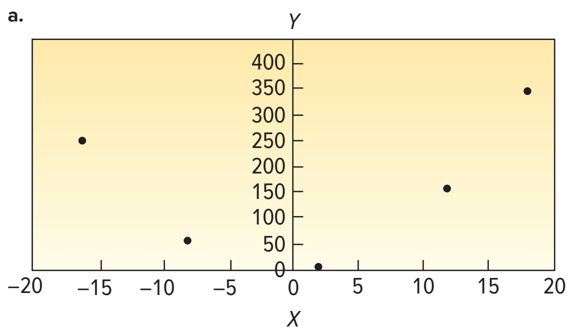
27. The standard error of estimate is 0.913, found by $\sqrt{\frac{6.667}{10-2}}$.
The coefficient of determination is 0.82, found by 29.733/36.4. Eighty-two percent of the variation in kilowatt hours can be explained by the variation in the number of rooms.

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. $6.308 \pm (3.182)(.993)\sqrt{1 + 1/5 + .0671}$
$= [2.751, 9.865]$

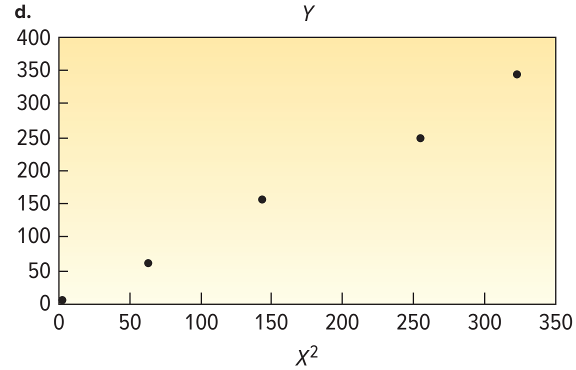
33. **a.** 4.2939, 6.3721
b. 2.9854, 7.6806

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 in 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.

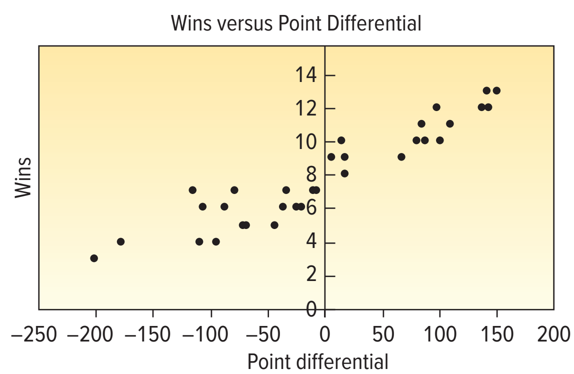
39. $H_0: \rho \leq 0$; $H_1: \rho > 0$. Reject H_0 if $t > 2.764$.

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

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 −0.0937. The linear relationship between points allowed and points scored is very, very weak.
b. $H_0: \rho \geq 0 \quad H_1: \rho < 0 \quad$ Reject H_0 if $t < -1.697$. df = 30
$t = \frac{-0.0937\sqrt{32-2}}{\sqrt{1-(-0.0937)^2}} = -1.680$. Using an online calculator, p-value = .6224
Fail to reject H_0. The evidence suggests no significant inverse relationship between points scored and points allowed.

43. **a.** There is a positive relationship between wins and point differential. Also, all teams with a "losing" season record (winning 7 or less games) recorded a negative point differential.



b. $r = .9367$. There is a strong, positive relationship between wins and point differential.
c. The R^2 = 87.78%. Point differential accounts for 87.78% of the variance of wins.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.9367
R Square	0.8775
Adjusted *R* Square	0.8734
Standard Error	1.0302
Observations	32

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	228.0365	228.0365	214.8686	0.0000
Residual	30	31.8385	1.0613		
Total	31	259.8750			

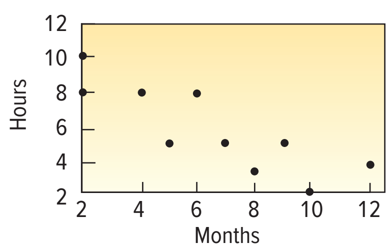
	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	7.9375	0.1821	43.5856	0.0000	7.5656	8.3094	7.5656	8.3094
Point differential	0.0282	0.0019	14.6584	0.0000	0.0242	0.0321	0.0242	0.0321

d. Wins = 7.9375 + .0282 (point differential)

e. Setting wins = 8, solve 8 = 7.9375 + .0282 (point differential) for point differential. The point differential is +2.2163 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 = 35.46 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 35 points; winning 10 games would require a point differential of about 70 points, etc.

45. a.



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

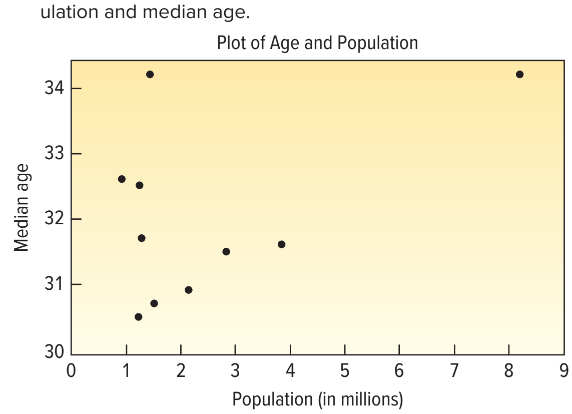
b. $r = -0.827$ The correlation coefficent indicates a strong, inverse linear relationship between months owned and hours exercised.

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. The 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

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

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

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

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 that 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			

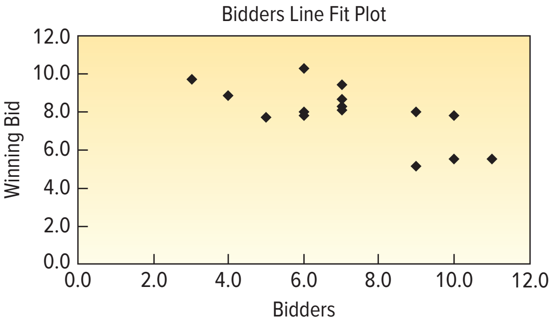
	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	31.3672	0.6158	50.9348	0.0000	29.9471	32.7873	29.9471	32.7873
Population	0.2722	0.1901	1.4319	0.1901	−0.1662	0.7106	−0.1662	0.7106

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).

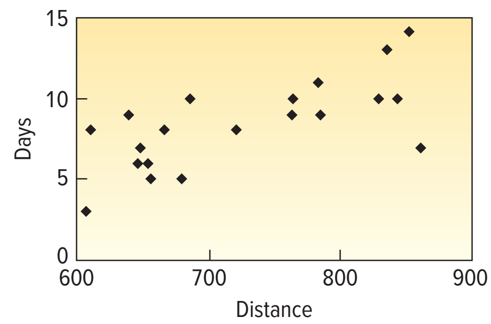
e. This indicates there is a negative relationship between the number of bids (X) and the winning bid (Y) and that for each additional bidder the winning bid decreases by 0.4667 million. The slope is significantly different from zero because its p-value, .0032, is less than .05.

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

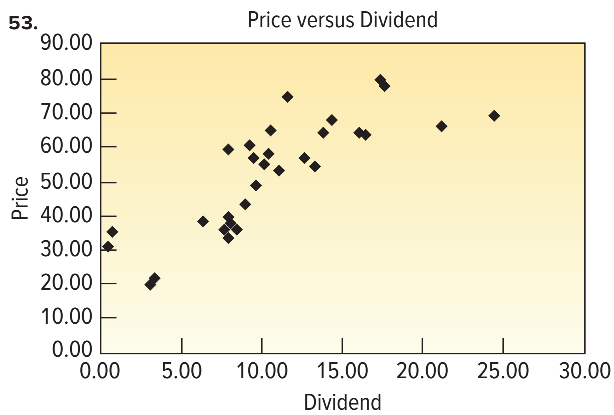
$$t = \frac{0.6921\sqrt{20 - 2}}{1 - (0.6921)^2} = 3.4562;$$ the one-sided p-value (.0007/2) is .0004. H_0 is rejected. There is a positive association between shipping distance and shipping time.

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

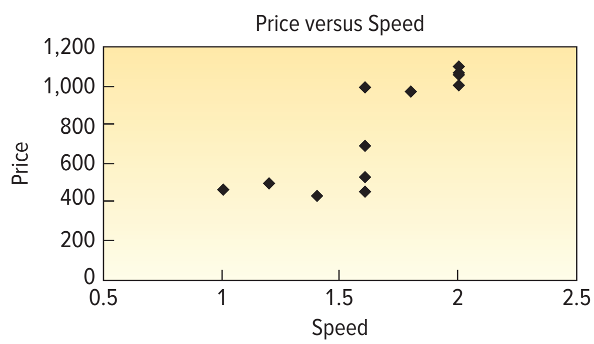
d. $r = \sqrt{0.932} = 0.966$

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

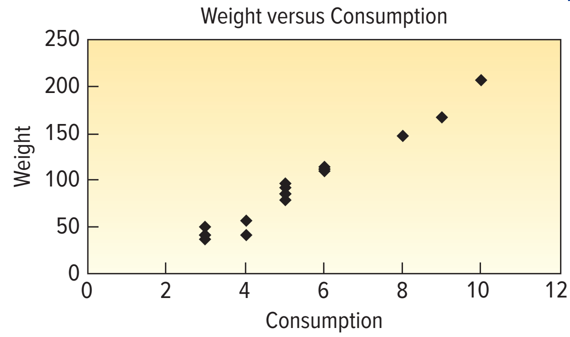
a. The correlation of Speed and Price is 0.8346.
H_0: $\rho \leq 0$ H_1: $\rho > 0$ Reject H_0 if $t > 1.8125$.

$t = \dfrac{0.8346\sqrt{12-2}}{\sqrt{1-(0.8346)^2}} = 4.7911$ 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

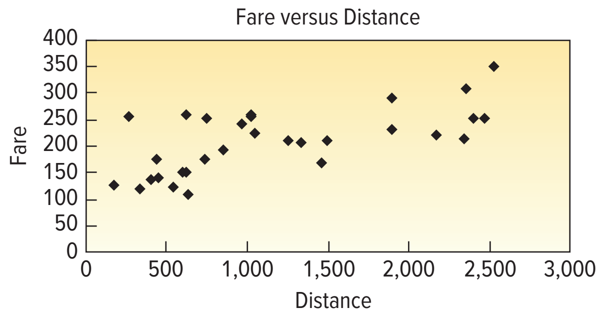
a. The correlation of Weight and Consumption is 0.9872.
H_0: $\rho \leq 0$ H_1: $\rho > 0$ Reject H_0 if $t > 1.746$.

$t = \dfrac{0.9872\sqrt{18-2}}{1-(0.9872)^2} = 24.7627$. 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 that 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.

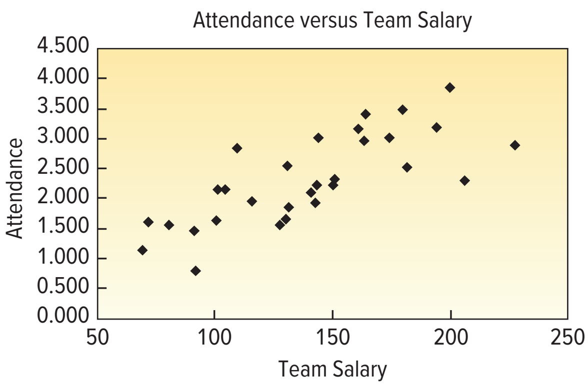
Reject H_0. There is a significant positive correlation between fares and distances.

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 1500-mile flight would cost $226.1312, found by $147.0812 + 0.0527(1500).

e. A flight of 4218 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*	0.7516
R Square	0.5649
Adjusted *R* Square	0.5494
Standard Error	0.4981
Observations	30

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	9.0205	9.0205	36.3547	0.0000
Residual	28	6.9475	0.2481		
Total	29	15.9680			

	Coefficients	Standard Error	*t*-Stat	*p*-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.4339	0.3261	1.3303	0.1942	−0.2342	1.1019	−0.2342	1.1019
Team Salary	0.0136	0.0023	6.0295	0.0000	0.0090	0.0182	0.0090	0.0182

The regression equation is: Attendance = .4339 + .0136(Team Salary). Expected Attendance with a salary of $100 million is 1.7939 million, found by .4339 + 0.0136 (100)

c. Increasing the salary by 30 million will increase attendance by 0.408 million on average, found by 0.0136 (30).

d. H_0: $\beta \le 0$ H_1: $\beta > 0$ df = n − 2 = 30 − 2 = 28
Reject H_0 if $t > 1.701$
$t = 0.0136/0.0023 = 6.0295$, Using a *p*-value calculator or statistical software, the *p*-value is less than .00001.
Reject H_0 and conclude the slope is positive.

e. 0.5649 or 56.49% of the variation in attendance is explained by variation in salary.

f.

Correlation Matrix

	Attendance	ERA	BA
Attendance	1		
ERA	−0.5612	1	
BA	0.2184	−0.4793	1

The correlation between attendance and batting average is 0.2184.
H_0: $\rho \le 0$ H_1: $\rho > 0$ At the 5% level, reject H_0 if $t > 1.701$.

$$t = \frac{0.2184\sqrt{30-2}}{\sqrt{1-(0.2184)^2}} = 1.1842$$

Using a *p*-value calculator or statistical software, the *p*-value is .1231. Fail to reject H_0.
The batting average and attendance are not positively correlated.
The correlation between attendance and ERA is −0.5612. The correlation between attendance and ERA is stronger than the correlation between attendance and batting average.
H_0: $\rho \ge 0$ H_1: $\rho < 0$ At the 5% level, reject H_0 if $t < -1.701$

$$t = \frac{-0.5612\sqrt{30-2}}{\sqrt{1-(-0.5612)^2}} = -3.5883$$

Using a *p*-value calculator or statistical software, the *p*-value is .0006. 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.
$$R^2_{adj} = 1 - \frac{\frac{\text{SSE}}{n-(k+1)}}{\frac{\text{SS total}}{n-1}} = 1 - \frac{\frac{583.693}{65-(2+1)}}{\frac{661.6}{65-1}} = 1 - \frac{9.414}{10.3375} = 1 - .911 = .089$$

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: $\beta_1 \ne 0$ H_1: $\beta_2 \ne 0$
$t = 1.99$ $t = -2.38$
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: $\beta_2 \ne 0$
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.

d. When you consider the interaction variable, the regression equation is Performance = 38.7 + 3.80 Aptitude − 0.1 Union + 3.61 x_1x_2.

```
Predictor     Coef   SE Coef      T       P
Constant     38.69     15.62   2.48   0.031
Aptitude     3.802     2.179   1.74   0.109
Union        -0.10     23.14  -0.00   0.997
X1X2         3.610     3.473   1.04   0.321
```

The *t*-value corresponding to the interaction term is 1.04. The *p*-value is .321 This is not significant. So we conclude there is no interaction between aptitude and union membership when predicting job performance.

11. a. The regression equation is
Price = 3,080 − 54.2 Bidders + 16.3 Age

```
Predictor     Coef   SE Coef      T       P
Constant    3080.1     343.9   8.96   0.000
Bidders     -54.19     12.28  -4.41   0.000
Age         16.289     3.784   4.30   0.000
```

The price decreases $54.2 as each additional bidder participates. Meanwhile the price increases $16.3 as the painting gets older. While one would expect older paintings to be worth more, it is unexpected that the price goes down as more bidders participate!

b. The regression equation is
Price = 3,972 − 185 Bidders + 6.35 Age + 1.46 x_1x_2

Predictor	Coef	SE Coef	T	P
Constant	3971.7	850.2	4.67	0.000
Bidders	−185.0	114.9	−1.61	0.122
Age	6.353	9.455	0.67	0.509
X_1X_2	1.462	1.277	1.15	0.265

The *t*-value corresponding to the interaction term is 1.15. This is not significant. So we conclude there is no interaction.

c. In the stepwise procedure, the number of bidders enters the equation first. Then the interaction term enters. The variable age would not be included as it is not significant. Response is Price on 3 predictors, with $N = 25$.

Step	1	2
Constant	4,507	4,540
Bidders	−57	−256
T-Value	−3.53	−5.59
P-Value	0.002	0.000
X_1X_2		2.25
T-Value		4.49
P-Value		0.000
S	295	218
R-Sq	35.11	66.14
R-Sq(adj)	32.29	63.06

Commentary: The stepwise method is misleading. In this problem, the first step is to run the "full" model with interaction. The result is that none of the independent variables are different from zero. So, remove the interaction term and rerun. Now we get the result in part (a). This is the model that should be used to predict price.

13. 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. $s_{y\cdot1234} = \sqrt{500/35} = 3.7796$

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.

15. 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.

17. a. $28,000

b. $R^2 = \dfrac{SSR}{SS\ total} = \dfrac{3{,}050}{5{,}250} = .5809$

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.

19. 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.

21. 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 TV's are, on average, $187.10 higher in price than Samsung TVs.

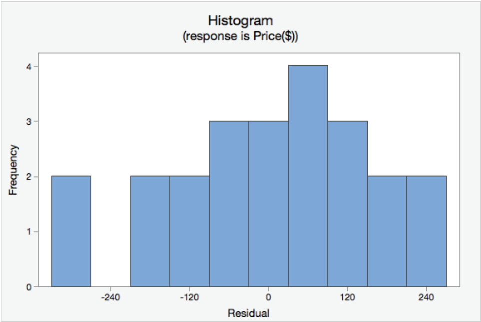
e. Here is some of the output.

Coefficients

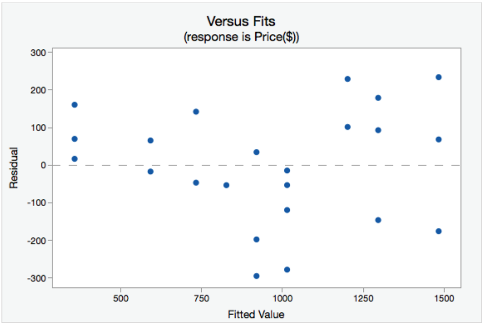
Term	Coef	SE Coef	95% CI	*t*-Value	*p*-Value
Constant	−1145.6	220.7	(−1606.1, −685.2)	−5.19	<0.0001
Screen	46.955	5.149	(36.215, 57.695)	9.12	<0.0001
Sony					
1	187.10	71.84	(37.24, 336.96)	2.60	0.0170

Based on the *p*-values, screen size and manufacturer are both significant in predicting price.

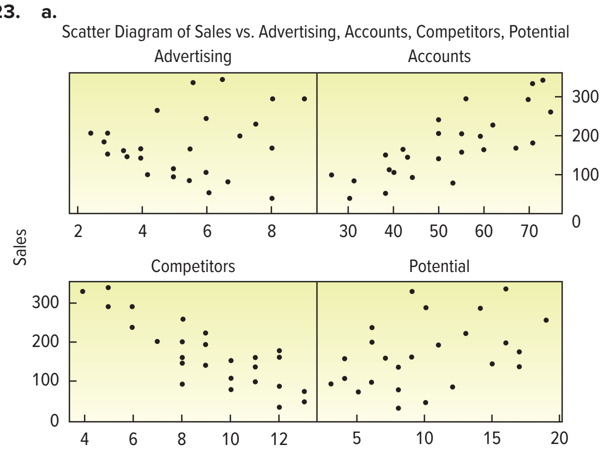
f. A histogram of the residuals indicates they follow a normal distribution.



g. There is no apparent relationship in the residuals, but the residual variation may be increasing with larger fitted values.



23. **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.

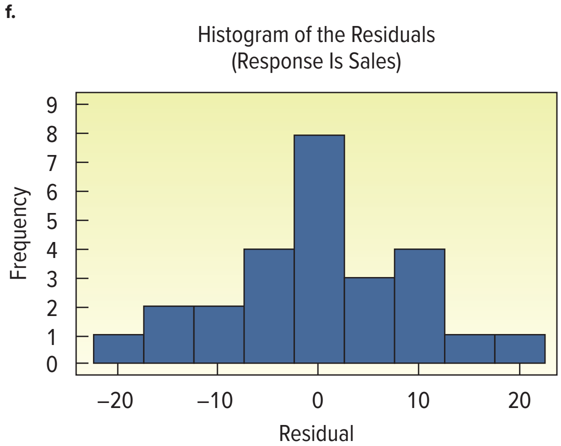
e. If you omit potential, the regression equation is:
Sales = 180 + 1.68 Advertising + 3.37 Accounts − 21.2 Competitors

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.

25. 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.

27. **a.** The correlation between the independent variables, yield and EPS, is small, .16195. Multicollinearity should not be a 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

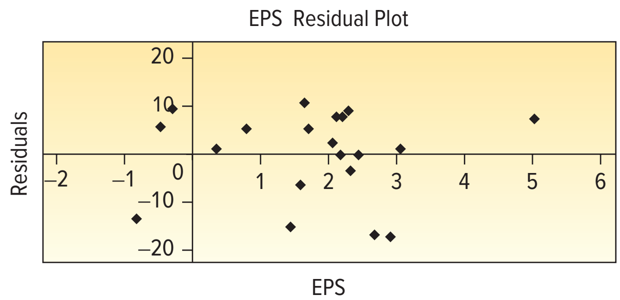
The regression equation is P/E = 29.913 − 5.324 EPS + 1.4 49 Yield.

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.

29. **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

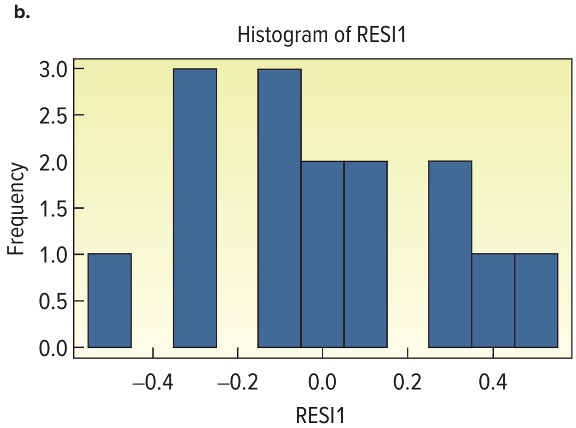
Regression Statistics	
Multiple *R*	0.6023
R Square	0.3628
Adjusted *R* Square	0.3274
Standard Error	9.4562
Observations	20

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.

31. a. The regression equation is

Auction price = −118,929 + 1.63 Loan + 2.1 Monthly payment + 50 Payments made

```
Analysis of Variance
Source       DF          SS          MS      F      P
Regression    3  5966725061  1988908354  39.83  0.000
Residual
  Error      16   798944439    49934027
Total        19  6765669500
```

The computed F is 39.83. It is much larger than the critical value 3.24. The p-value is also quite small. Thus, the null hypothesis that all the regression coefficients are zero can be rejected. At least one of the multiple regression coefficients is different from zero.

b.

```
Predictor        Coef    SE Coef       T        P
Constant      -118929      19734   -6.03    0.000
Loan           1.6268     0.1809    8.99    0.000
Monthly
  Payment        2.06      14.95    0.14    0.892
Payments
  Made           50.3      134.9    0.37    0.714
```

The null hypothesis is that the coefficient is zero in the individual test. It would be rejected if t is less than −2.120 or more than 2.120. In this case, the t value for the loan variable is larger than the critical value. Thus, it should not be removed. However, the monthly payment and payments made variables would likely be removed.

c. The revised regression equation is: Auction price = −119,893 + 1.67 Loan

33. 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-the-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

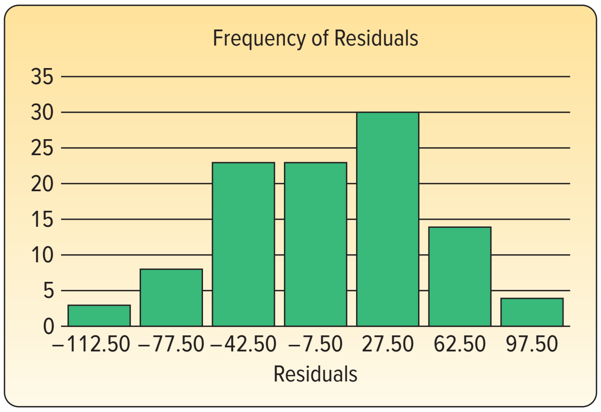
Regression Statistics	
Multiple *R*	0.955
R Square	0.913
Adjusted *R* Square	0.911
Standard Error	47914.856
Observations	105

ANOVA

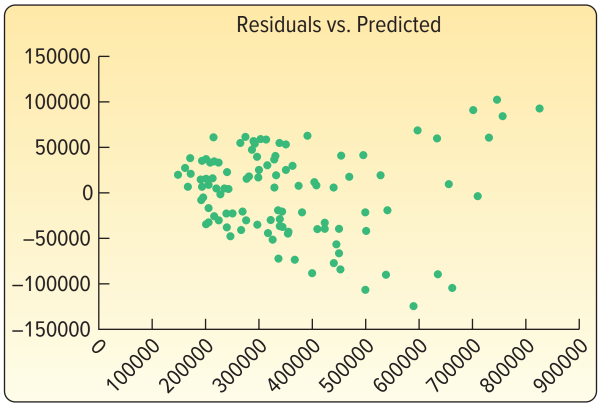
	df	SS	MS	*F*	Significance *F*
Regression	2.00	2451577033207.43	1225788516603.72	533.92	0.00
Residual	102.00	234175013207.24	2295833462.82		
Total	104.00	2685752046414.68			

	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.



35. 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

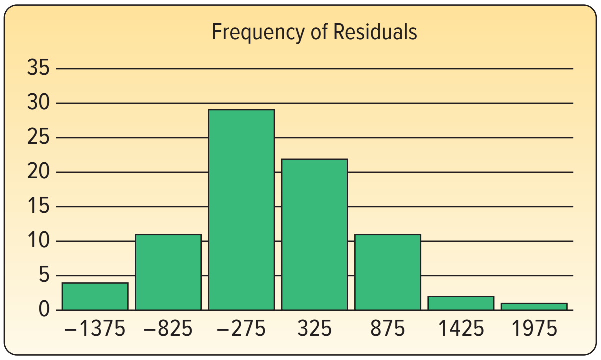
ANOVA

	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			

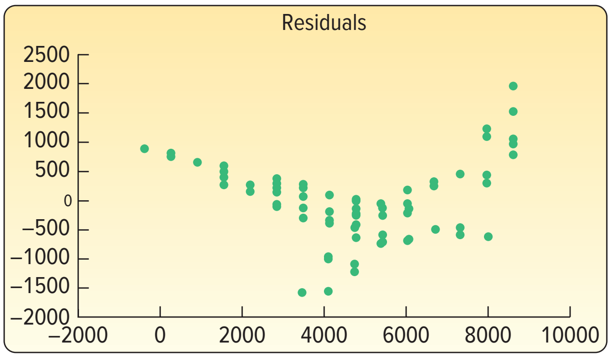
	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	−1028.539	213.761	−4.812	0.000
Age (years)	644.528	27.157	23.733	0.000
Engine Type (0=diesel)	3190.481	156.100	20.439	0.000

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.

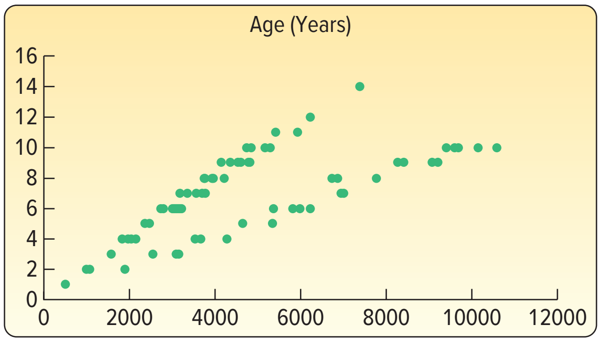
c. The normality conjecture appears realistic.



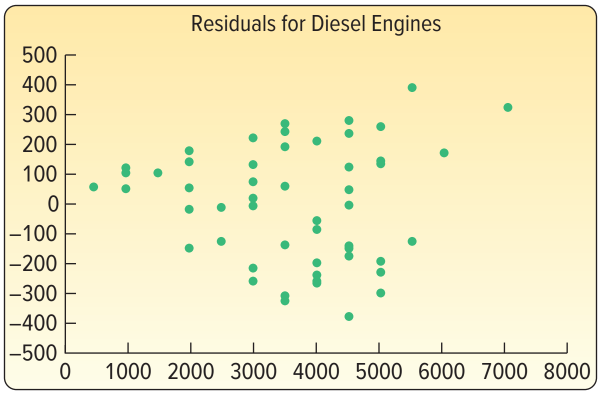
d. The plot of residuals versus predicted values shows the following. There are clearly patterns in the graph that indicate that the residuals do not follow the assumptions required for the tests of hypotheses.

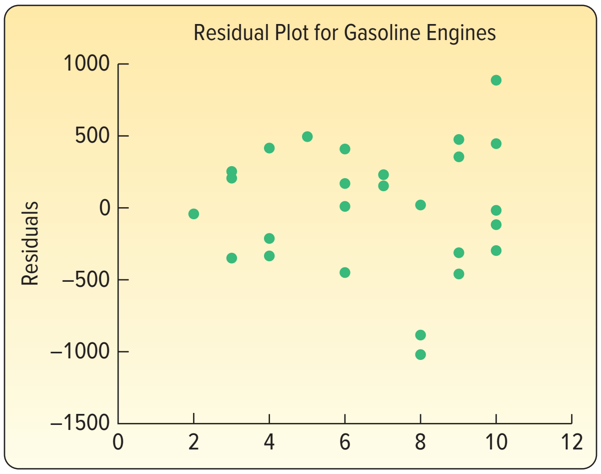


Let's remember the scatter plot of costs versus age. The graph clearly shows the effect of engine type on costs. So there are essentially two regression equations depending on the type of engine.



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:** $H_0: \pi = 0.10$ $H_1: \pi \neq 0.10$
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.** $H_0: \pi \geq 0.90$ $H_1: \pi < 0.90$
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. **a.** $H_0: \pi_1 = \pi_2 \qquad H_1: \pi_1 \neq \pi_2$

b. H_0 is rejected if $z < -1.96$ or $z > 1.96$.

c. $p_c = \dfrac{24 + 40}{400 + 400} = 0.08$

d. −2.09, found by

$$z = \frac{0.06 - 0.10}{\sqrt{[(0.08 \times 0.92)/400] + [(0.08 \times 0.92)/400]}}$$

e. H_0 is rejected. The proportion infested is not the same in the two fields.

11. $H_0: \pi_d \leq \pi_r \qquad H_1: \pi_d > \pi_r$

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.

Number of Clients	z-Values	Area	Found by	f_e
Under 30	Under −1.58	0.0571	0.5000 − 0.4429	2.855
30 up to 40	−1.58 up to −0.51	0.2479	0.4429 − 0.1950	12.395
40 up to 50	−0.51 up to 0.55	0.4038	0.1950 + 0.2088	20.19
50 up to 60	0.55 up to 1.62	0.2386	0.4474 − 0.2088	11.93
60 or more	1.62 or more	0.0526	0.5000 − 0.4474	2.63

The first and last class both have expected frequencies smaller than 5. They are combined with adjacent classes.
H_0: The population of clients follows a normal distribution.
H_1: The population of clients does not follow a normal distribution.
Reject the null if $\chi^2 > 5.991$.

Number of Clients	Area	f_e	f_o	$f_e - f_o$	$(f_o - f_e)^2$	$[(f_o - f_e)^2]/f_e$
Under 40	0.3050	15.25	16	−0.75	0.5625	0.0369
40 up to 50	0.4038	20.19	22	−1.81	3.2761	0.1623
50 or more	0.2912	14.56	12	2.56	6.5536	0.4501
Total	1.0000	50.00	50	0		0.6493

Since 0.6493 is not greater than 5.991, we fail to reject the null hypothesis. These data could be from a normal distribution.

27. 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.

29. 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.

31. **a.** $H_0: \pi = 0.50 \qquad H_1: \pi \neq 0.50$

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. $z = \dfrac{\frac{36}{53} - 0.5}{\sqrt{0.5(1 - 0.5)/53}} = 2.61$

We reject the null hypothesis.

e. Using a p-value calculator (rounding to three decimal places) or a z-table, the p-value is 0.009, found by 2(0.5000 − 0.4955). The data indicates that the National Football Conference is luckier than the American Conference in calling the flip of a coin.

33. $H_0: \pi \leq 0.60 \qquad H_1: \pi > 0.60$

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.

35. $H_0: \pi \leq 0.44 \qquad H_1: \pi > 0.44$

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.

37. $H_0: \pi \leq 0.20 \qquad H_1: \pi > 0.20$

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.

39. $H_0: \pi \geq 0.0008 \qquad H_1: \pi < 0.0008$

H_0 is rejected if $z < -1.645$.

$$z = \frac{0.0006 - 0.0008}{\sqrt{\dfrac{0.0008\,(0.9992)}{10{,}000}}} = -0.707 \qquad H_0 \text{ is not rejected.}$$

These data do not prove there is a reduced fatality rate.

41. 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.

43. 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.

45. 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.

47. 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.

49. 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.

51. 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.

53. 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.

55.

Hourly Wage	f	M	fM	$M - x$	$(M - x)^2$	$f(M - x)^2$
\$5.50 up to 6.50	20	6	120	−2.222	4.938	98.8
6.50 up to 7.50	24	7	168	−1.222	1.494	35.9
7.50 up to 8.50	130	8	1040	−0.222	0.049	6.4
8.50 up to 9.50	68	9	612	0.778	0.605	41.1
9.50 up to 10.50	28	10	280	1.778	3.161	88.5
Total	270		2220			270.7

The sample mean is 8.222, found by 2,220/270.
The sample standard deviation is 1.003, found as the square root of 270.7/269.
H_0: The population of wages follows a normal distribution.
H_1: The population of hourly wages does not follow a normal distribution.
Reject the null if $\chi^2 > 4.605$.

Wage	z-values	Area	Found by	f_e	f_o	$f_e - f_o$	$(f_o - f_e)^2$	$[(f_o - f_e)^2]/f_e$
Under \$6.50	Under −1.72	0.0427	0.5000 − 0.4573	11.529	20	−8.471	71.7578	6.2241
6.50 up to 7.50	−1.72 up to −0.72	0.1931	0.4573 − 0.2642	52.137	24	28.137	791.6908	15.1848
7.50 up to 8.50	−0.72 up to 0.28	0.3745	0.2642 + 0.1103	101.115	130	−28.885	834.3432	8.2514
8.50 up to 9.50	0.28 up to 1.27	0.2877	0.3980 − 0.1103	77.679	68	9.679	93.6830	1.2060
9.50 or more	1.27 or more	0.1020	0.5000 − 0.3980	27.54	28	−0.46	0.2116	0.0077
Total		1.0000		270	270	0		30.874

Since 30.874 is greater than 4.605, we reject the null hypothesis not from a normal distribution.

57. 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.

59. 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.

61. 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.

63. 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	10	4	14
Yes	5	11	16
Total	15	15	

$$\chi^2 = \frac{(10 - 7)^2}{7} + \frac{(4 - 7)^2}{7} + \frac{(5 - 8)^2}{8} + \frac{(11 - 8)^2}{8} = 4.82$$

Reject H_0. Conclude that salary and winning are related.

CHAPTER 16

1. **a.** If the number of pluses (successes) in the sample is 9 or more, reject H_0.
b. Reject H_0 because the cumulative probability associated with nine or more successes (.073) does not exceed the significance level (.10).

3. **a.** H_0: $\pi \le .50$; H_1: $\pi > .50$; $n = 10$
b. H_0 is rejected if there are nine or more plus signs. A "+" represents a loss.
c. Reject H_0. It is an effective program because there were nine people who lost weight.

5. **a.** H_0: median \$81,500 H_1: median > \$81,500
b. Reject H_0 if 12 or more earned than \$81,500.
c. 13 of the 18 chiropractors earned more than \$81,500 so reject H_0. The results indicate the starting salary for chiropractors is more than \$81,500.

7.

Couple	Difference	Rank
1	550	7
2	190	5
3	250	6
4	−120	3
5	−70	1
6	130	4
7	90	2

Sums: −4, +24. So $T = 4$ (the smaller of the two sums). From Appendix B.8, .05 level, one-tailed test, $n = 7$, the critical value is 3. Since the T of 4 > 3, do not reject H_0 (one-tailed test). There is no difference in square footage. Professional couples do not live in larger homes.

9. **a.** H_0: The production is the same for the two systems.
H_1: Production using the new procedure is greater.
b. H_0 is rejected if $T \le 21$, $n = 13$.
c. The calculations for the first three employees are:

Employee	Old	New	d	Rank	R^1	R^2
A	60	64	4	6	6	
B	40	52	12	12.5	12.5	
C	59	58	−1	2		2

The sum of the negative ranks is 6.5. Since 6.5 is less than 21, H_0 is rejected. Production using the new procedure is greater.

11. H_0: The distributions are the same. H_1: The distributions are not the same. Reject H_0 if z, 21.96 or $z > 1.96$.

A		B	
Score	Rank	Score	Rank
38	4	26	1
45	6	31	2
56	9	35	3
57	10.5	42	5
61	12	51	7
69	14	52	8
70	15	57	10.5
79	16	62	13
	86.5		49.5

$$z = \frac{86.5 - \frac{8(8 + 8 + 1)}{2}}{\sqrt{\frac{8(8)(8 + 8 + 1)}{12}}} = 1.943$$

H_0 is not rejected. There is no difference in the two populations.

13. H_0: The distributions are the same. H_1: The distribution of Campus is to the right. Reject H_0 if $z > 1.65$.

Campus		Online	
Age	Rank	Age	Rank
26	6	28	8
42	16.5	16	1
65	22	42	16.5
38	13	29	9.5
29	9.5	31	11
32	12	22	3
59	21	50	20
42	16.5	42	16.5
27	7	23	4
41	14	25	5
46	19		94.5
18	2		
	158.5		

$$z = \frac{158.5 - \frac{12(12 + 10 + 1)}{2}}{\sqrt{\frac{12(10)(12 + 10 + 1)}{12}}} = 1.35$$

H_0 is not rejected. There is no difference in the distributions.

15. ANOVA requires that we have two or more populations, the data are interval- or ratio-level, the populations are normally distributed, and the population standard deviations are equal. Kruskal-Wallis requires only ordinal-level data, and no assumptions are made regarding the shape of the populations.

17. **a.** H_0: The three population distributions are equal.
H_1: Not all of the distributions are the same.
b. Reject H_0 if $H > 5.991$.

c.

Sample 1 Rank	Sample 2 Rank	Sample 3 Rank
8	5	1
11	6.5	2
14.5	6.5	3
14.5	10	4
16	12	9
64	13	19
	53	

$$H = \frac{12}{16(16+1)}\left[\frac{(64)^2}{5} + \frac{(53)^2}{6} + \frac{(19)^2}{5}\right] - 3(16+1)$$

$= 59.98 - 51 = 8.98$

d. Reject H_0 because $8.98 > 5.991$. The three distributions are not equal.

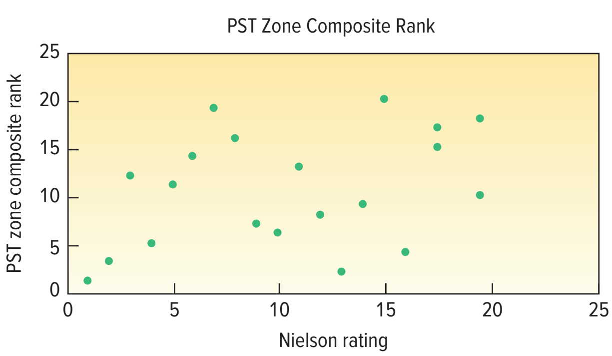
19. H_0: The distributions of the lengths of life are the same.
H_1: The distributions of the lengths of life are not the same.
H_0 is rejected if $H > 9.210$.

Salt		Fresh		Others	
Hours	Rank	Hours	Rank	Hours	Rank
167.3	3	160.6	1	182.7	13
189.6	15	177.6	11	165.4	2
177.2	10	185.3	14	172.9	7
169.4	6	168.6	4	169.2	5
180.3	12	176.6	9	174.7	8
	46		39		35

$$H = \frac{12}{15(16)}\left[\frac{(46)^2}{5} + \frac{(39)^2}{5} + \frac{(35)^2}{5}\right] - 3(16) = 0.62$$

H_0 is not rejected. There is no difference in the three distributions.

21. a.



b. $r_s = 1 - \dfrac{6\Sigma d^2}{n(n^2-1)} = 1 - \dfrac{6(845)}{20(20^2-1)} = .635$

c. H_0: No correlation among the ranks
H_1: A positive correlation among the ranks
Reject H_0 if $t > 1.734$.

$$t = .635\sqrt{\frac{20-2}{1-.635^2}} = 2.489$$

H_0 is rejected. We conclude the correlation in population among the ranks is positive. The Nielson rankings and the PST zone composite rank are significantly, positively related.

23.

Representative	Sales	Rank	Training Rank	d	d^2
1	319	8	8	0	0
2	150	1	2	1	1
3	175	2	5	3	9
4	460	10	10	0	0
5	348	9	7	−2	4
6	300	6.5	1	5.5	30.25
7	280	5	6	1	1
8	200	4	9	5	25
9	190	3	4	1	1
10	300	6.5	3	−3.5	12.25
					83.50

a. $r_s = 1 - \dfrac{6(83.5)}{10(10^2-1)} = 0.494$

A moderate positive correlation

b. H_0: No correlation among the ranks. H_1: A positive correlation among the ranks. Reject H_0 if $t > 1.860$.

$$t = 0.494\sqrt{\frac{10-2}{1-(0.494)^2}} = 1.607$$

H_0 is not rejected. The correlation in population among the ranks could be 0.

25. H_0: $\pi = .50$. H_1: $\pi = .50$. Use a software package to develop the binomial probability distribution for $n = 19$ and $\pi = .50$. H_0 is rejected if there are either 5 or fewer "+" signs, or 14 or more. The total of 12 "+" signs falls in the acceptance region. H_0 is not rejected. There is no preference between the two shows.

27. H_0: $\pi = .50$ H_1: $\pi = .50$
H_0 is rejected if there are 12 or more or 3 or fewer plus signs. Because there are only 8 plus signs, H_0 is not rejected. There is no preference with respect to the two brands of components.

29. a. H_0: $= 0.50$ H_1: 0.50 $n = 22$; 2 were indifferent, so $n = 20$. 5 preferred pulp; 15 preferred no pulp.

b. As a two-tailed test, Reject if 5 or less preferred pulp, or 14 or more preferred pulp.

c. Reject H_0. There is a difference in the preference for the two types of orange juice.

31. H_0: Rates are the same; H_1: The rates are not the same. H_0 is rejected if $H > 5.991$. $H = .082$. Do not reject H_0.

33. H_0: The populations are the same. H_1: The populations differ. Reject H_0 if $H > 7.815$. $H = 14.30$. Reject H_0.

35. $r_s = 1 - \dfrac{6(78)}{12(12^2-1)} = 0.727$

H_0: There is no correlation between the rankings of the coaches and of the sportswriters.
H_1: There is a positive correlation between the rankings of the coaches and of the sportswriters. Reject H_0 if $t > 1.812$.

$$t = 0.727\sqrt{\frac{12-2}{1-(.727)^2}} = 3.348$$

H_0 is rejected. There is a positive correlation between the sportswriters and the coaches.

37. a. H_0: There is no difference in the distributions of the selling prices in the five townships.
H_1: There is a difference in the distributions of the selling prices of the five townships.
H_0 is rejected if H is greater than 9.488. The computed value of H is 2.70, so the null hypothesis is not rejected. The sample data does not suggest a difference in the distributions of selling prices.

b. H_0: There is no difference in the distributions of the selling prices depending on the number of bedrooms.
H_1: There is a difference in the distributions of the selling prices depending on the number of bedrooms.

H_0 is rejected if H is greater than 9.488. The computed value of H is 75.71, so the null hypothesis is rejected. The sample data indicates there is a difference in the distributions of selling prices based on the number of bedrooms.

c. H_0: There is no difference in the distributions of FICO scores depending on the type of mortgage the occupant has on the home.
H_1: There is a difference in the distributions of FICO scores depending on the type of mortgage the occupant has on the home.
H_0 is rejected if H is greater than 3.841. The computed value of H is 41.04, so the null hypothesis is rejected. The sample data suggests a difference in the distributions of the FICO scores. The data shows that home occupants with lower FICO scores tended to use adjustable rate mortgages.

39. **a.** H_0: The distributions of the maintenance costs are the same for all capacities.
H_1: The distributions of the costs are not the same.
H_0 is rejected if $H > 7.815$, from χ^2 with 3 degrees of freedom.

$$H = \frac{12}{80(81)}\left[\frac{(132)^2}{3} + \frac{(501)^2}{11} + \frac{(349)^2}{11} + \frac{(2258)^2}{55}\right] - 3(81) = 2.186$$

Fail to reject H_0. There is no difference in the maintenance cost for the four bus capacities.

b. H_0: The distributions of maintenance costs by fuel type are the same.
H_1: The distributions are different.
Reject H_0 if $z < -1.96$ or $z > 1.96$.

$$z = \frac{1693 - \dfrac{53(53+27+1)}{2}}{\sqrt{\dfrac{(53)(27)(53+27+1)}{12}}} = -4.614$$

We reject reject H_0 and conclude that maintenance costs are different for diesel and gasoline fueled buses.

c. H_0: The distributions of the maintenance costs are the same for the three bus manufacturers.
H_1: The distributions of the costs are not the same.
H_o is rejected if $H > 5.991$, from χ^2 with 2 degrees of freedom.

$$H = \frac{12}{80(81)}\left[\frac{(414)^2}{8} + \frac{(1005)^2}{25} + \frac{(1821)^2}{47}\right] - 3(81) = 2.147$$

H_o is not rejected. There may be no difference in the maintenance cost for the three different manufacturers. The distributions could be the same.

CHAPTER 17

1.

Year	Loans ($ millions)	Index (base = 2010)
2010	55,177	100.0
2011	65,694	119.1
2012	83,040	150.5
2013	88,378	160.2
2014	97,420	176.6
2015	98,608	178.7
2016	101,364	183.7
2017	110,527	200.3
2018	116,364	210.9

3. The mean sales for the earliest 3 years is $(486.6 + 506.8 + 522.2)/3 or $505.2.
2017: 90.4, found by (456.6/505.2) (100)
2018: 85.8, found by (433.3/505.2) (100)
Net sales decreased by 9.6% and 14.2% from the 2009–2010 period to 2017 and 2018, respectively.

5. **a.** $P_t = \frac{3.35}{2.49}(100) = 134.54$ $P_s = \frac{4.49}{3.29}(100) = 136.47$

$P_c = \frac{4.19}{1.59}(100) = 263.52$ $P_a = \frac{2.49}{1.79}(100) = 139.11$

b. $P = \frac{14.52}{9.16}(100) = 158.52$

c. $P = \frac{\$3.35(6) + 4.49(4) + 4.19(2) + 2.49(3)}{\$2.49(6) + 3.29(4) + 1.59(2) + 1.79(3)}(100) = 147.1$

d. $P = \frac{\$3.35(6) + 4.49(5) + 4.19(3) + 2.49(4)}{\$2.49(6) + 3.29(5) + 1.59(3) + 1.79(4)}(100) = 150.2$

e. $I = \sqrt{(147.1)(150.2)} = 148.64$

7. **a.** $P_W = \frac{0.10}{0.07}(100) = 142.9$ $P_C = \frac{0.03}{0.04}(100) = 75.0$

$P_S = \frac{0.15}{0.15}(100) = 100$ $P_H = \frac{0.10}{0.08}(100) = 125.0$

b. $P = \frac{0.38}{0.34}(100) = 111.8$

c.

$$P = \frac{0.10(17{,}000) + 0.03(125{,}000) + 0.15(40{,}000) + 0.10(62{,}000)}{0.07(17{,}000) + 0.04(125{,}000) + 0.15(40{,}000) + 0.08(62{,}000)} \times (100) = 102.92$$

d.

$$P = \frac{0.10(20{,}000) + 0.03(130{,}000) + 0.15(42{,}000) + 0.10(65{,}000)}{0.07(20{,}000) + 0.04(130{,}000) + 0.15(42{,}000) + 0.08(65{,}000)} \times (100) = 103.32$$

e. $I = \sqrt{102.92(103.32)} = 103.12$

9.

Grain	2015 Price/ Bushel	2015 Production (1000 MT)	$p_0 q_0$	2018 Price/ Bushel	2018 Production (1000 of MT)	$p_t q_t$
Oats	2.488	1,298	3,229	2.571	815	2,095
Wheat	5.094	26,117	133,040	4.976	51,287	255,204
Corn	3.783	345,506	1,307,049	3.704	366,287	1,356,727
Barley	5.084	4,750	24,149	4.976	3,333	16,585
			Sum = 1,467,467			Sum = 1,630,611
		Value Index	111.117			

11. **a.** $I = \frac{6.8}{5.3}(0.20) + \frac{362.26}{265.88}(0.40) + \frac{125.0}{109.6}(0.25) + \frac{622{,}864}{529{,}917}(0.15) = 1.263.$
Index is 126.3.

b. Business activity increased 26.3% from 2000 to 2018.

13. The real income is X = ($86,829)/2.51107 = $34,578.
"Real" salary increased $34,578 − $19,800 = $17,778.

15.

Year	Tinora	Tinora Index	National Index
2000	$28,650	100.0	100
2010	$33,972	118.6	122.5
2018	$37,382	130.5	136.9

The Tinora teachers received smaller increases than the national average.

17.

Year	Domestic Sales (base = 2010)
2010	100.0
2011	43.8
2012	101.3
2013	108.4
2014	118.2
2015	121.2
2016	128.4
2017	135.4
2018	142.3

Compared to 2010, domestic sates are 42.3% higher.

19.

Year	International Sales (base = 2010)
2010	100.0
2011	99.9
2012	116.4
2013	122.7
2014	123.1
2015	107.0
2016	106.1
2017	113.9
2018	123.6

Compared to 2010, international sales are 23.6% higher.

21.

Year	Employees (base = 2010)
2010	100.0
2011	103.4
2012	111.9
2013	112.4
2014	111.0
2015	111.5
2016	110.9
2017	117.5
2018	118.5

Compared to 2010, the number of employees is 18.5% higher.

23.

Year	Revenue (millions $)	Simple Index, Revenue (base = 2013)
2013	113,245	100.0
2014	117,184	103.5
2015	117,386	103.7
2016	123,693	109.2
2017	122,092	107.8
2018	125,615	110.9

Compared to 2013, revenue increased 10.9%.

25.

Year	Employees (thousands)	Simple Index, Employees (base = 2013)
2013	307	100.0
2014	305	99.3
2015	333	108.5
2016	295	96.1
2017	313	102.0
2018	283	92.2

Compared to 2013, employees decreased 7.8%.

27. $P_{ma} = \frac{2.00}{0.81}(100) = 246.91$ $\quad P_{sh} = \frac{1.88}{0.84}(100) = 223.81$

$P_{mi} = \frac{2.89}{1.44}(100) = 200.69$ $\quad P_{po} = \frac{3.99}{2.91}(100) = 137.11$

29. $P = \frac{\$2.00(18) + 1.88(5) + 2.89(70) + 3.99(27)}{\$0.81(18) + 0.84(5) + 1.44(70) + 2.91(27)}(100) = 179.37$

31. $I = \sqrt{179.37(178.23)} = 178.80$

33. $P_R = \frac{0.60}{0.50}(100) = 120$ $\quad P_S = \frac{0.90}{1.20}(100) = 75.0$

$P_W = \frac{1.00}{0.85}(100) = 117.65$

35. $P = \frac{0.60(320) + 0.90(110) + 1.00(230)}{0.50(320) + 1.20(110) + 0.85(230)}(100) = 106.87$

37. $P = \sqrt{(106.87)(106.04)} = 106.45$

39. $P_C = \frac{0.05}{0.06}(100) = 83.33$ $\quad P_C = \frac{0.12}{0.10}(100) = 120$

$P_P = \frac{0.18}{0.20}(100) = 90$ $\quad P_E = \frac{.015}{0.15}(100) = 100$

41. $P = \frac{0.05(2{,}000) + 0.12(200) + 0.18(400) + 0.15(100)}{0.06(2{,}000) + 0.10(200) + 0.20(400) + 0.15(100)}(100)$

$= 89.79$

43. $I = \sqrt{(89.79)(91.25)} = 90.52$

45. $P_A = \frac{.86}{.82}(100) = 104.9$ $\quad P_N = \frac{2.99}{4.37}(100) = 68.4$

$P_{PET} = \frac{58.15}{71.21}(100) = 81.7$ $\quad P_{PL} = \frac{1292.53}{1743.6}(100) = 74.1$

47. $P_{Laspeyres} =$

$\frac{.86(1000) + 2.99(5000) + 58.15(60000) + 1292.53(500)}{.82(1000) + 4.37(5000) + 71.21(60000) + 1743.60(500)}(100)$

$= 80.34$

49. $I = \sqrt{(80.34)(80.14)} = 80.24$

51. $I = 100\left[\frac{1971.0}{1159.0}(0.20) + \frac{91}{87}(0.10) + \frac{114.7}{110.6}(0.40) + \frac{1501000}{1214000}(0.30)\right] = 123.05$

The economy is up 23.05 percent from 1996 to 2016.

53. February: $I = 100\left[\frac{6.8}{8.0}(0.40) + \frac{23}{20}(0.35) + \frac{303}{300}(0.25)\right]$

$= 99.50$

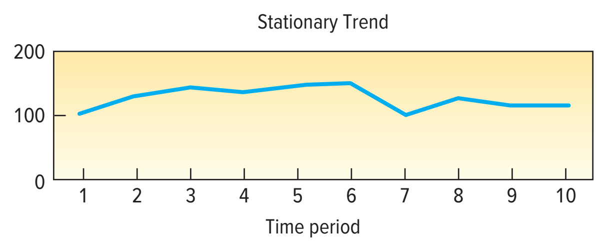
March: $I = 100\left[\frac{6.4}{8.0}(0.40) + \frac{21}{20}(0.35) + \frac{297}{300}(0.25)\right]$

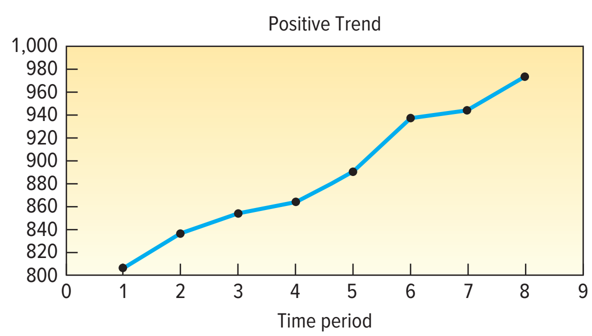
$= 93.50$

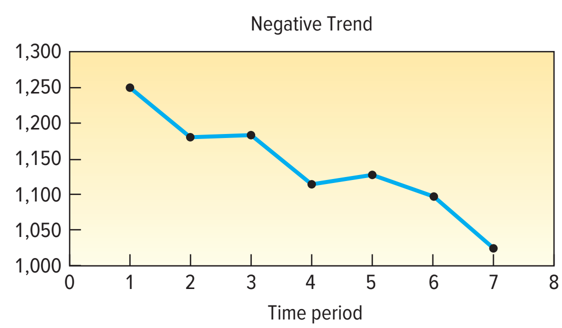
55. For 2006: $1,495,327, found by $2,400,000/1.605
For 2018: $1,715,686, found by $3,500,000/2.040

CHAPTER 18

1. Any graphs similar to the following:

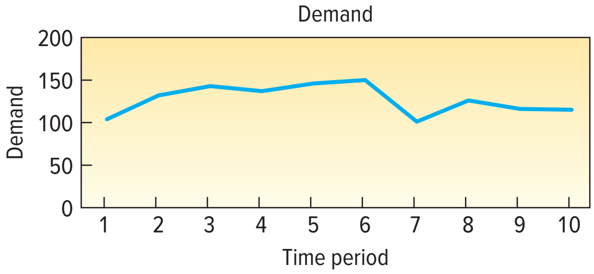






3. The irregular component is the randomness in a time series that cannot be described by any trend, seasonal, or cyclical pattern. The irregular component is used to estimate forecast error.

5. **a.** The graph shows a stationary pattern.



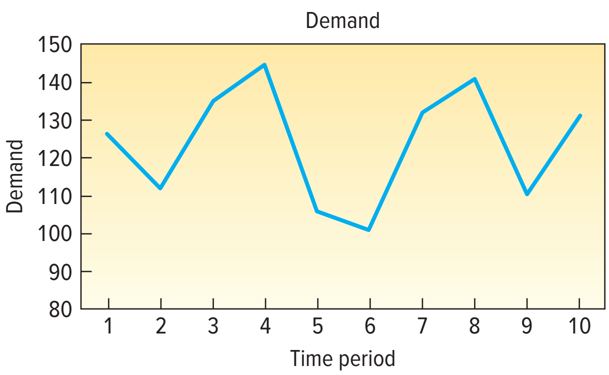
b. & c.

Period	Demand	3-Month SMA	Absolute Errors
1	104		
2	132		
3	143		
4	137	126.33	10.67
5	146	137.33	8.67
6	150	142	8
7	101	144.33	43.33
8	126	132.33	6.33
9	116	125.67	9.67
10	115	114.33	0.67
11		119	
			MAD = 12.48

d. The forecast demand for period 11 is 119.

e. The MAD of 12.48 is the reported measure of error. So, the forecast is 119 ± 12.48.

7. **a.** The graph shows a stationary pattern.

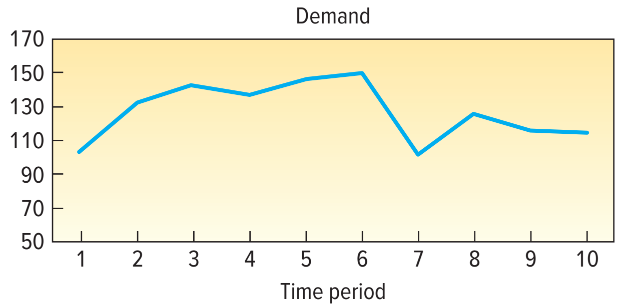


b.–d.

Period	Demand	4-Month SMA	6-Month SMA	4-Month Absolute Error	6-Month Absolute Error
1	126				
2	112				
3	135				
4	145				
5	106	129.50		23.50	
6	101	124.50		23.50	
7	132	121.75	120.80	10.25	11.20
8	141	121.00	121.80	20.00	19.20
9	110	120.00	126.70	10.00	16.70
10	131	121.00	122.50	10.00	8.50
11		128.50	120.20		
				4-month MAD	16.21
				6-month MAD	13.90

e. The 6-month moving average MAD of 13.90, is less than the 4-month MAD of 16.21. This is one reason to prefer using the 6-month rather than the 4-month moving average.

9. **a.** The graph shows a stationary pattern.



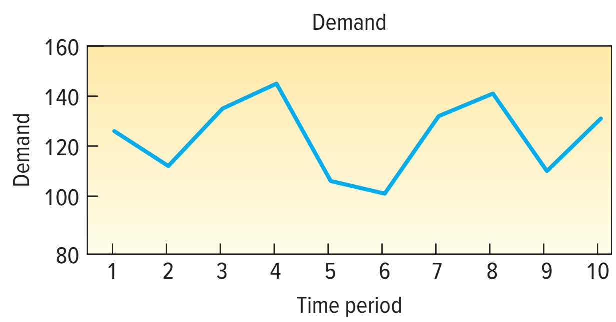
b. & c.

Period	Alpha = Demand	0.3 Exp Smooth	Error	Absolute Error
1	104			
2	132	104	28	28
3	143	112.4	30.6	30.6
4	137	121.6	15.4	15.4
5	146	126.2	19.8	19.8
6	150	132.1	17.9	17.9
7	101	137.5	−36.5	36.5
8	126	126.6	−0.6	0.6
9	116	126.4	−10.4	10.4
10	115	123.3	−8.3	8.3
11		120.8		18.61

d. Period 11 forecast = 120.80

e. The MAD of 18.61 is the reported measure of error. So, the forecast is 120.8 ± 18.61.

11. **a.** The graph shows a stationary pattern.



b.

Period	Alpha Demand	0.35 Exp Smooth	Error	Absolute Error
1	126			
2	112	126.00	−14.00	14.00
3	135	121.10	13.90	13.90
4	145	125.97	19.04	19.04
5	106	132.63	−26.63	26.63
6	101	123.31	−22.31	22.31
7	132	115.50	16.50	16.50
8	141	121.28	19.73	19.73
9	110	128.18	−18.18	18.18
10	131	121.82	9.18	9.18
11		125.03		
				MAD = 17.72

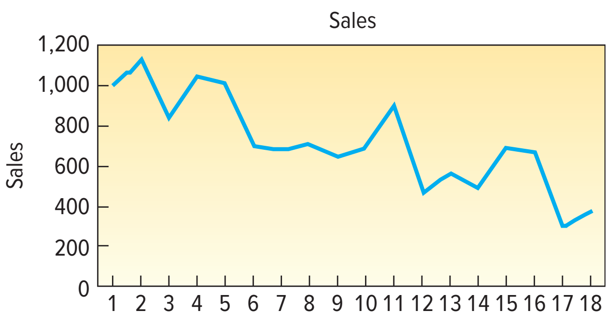
c.

Period	Alpha Demand	0.85 Exp Smooth	Error	Absolute Error
1	126			
2	112	126.00	−14.00	14.00
3	135	114.10	20.90	20.90
4	145	131.87	13.14	13.14
5	106	143.03	−37.03	37.03
6	101	111.55	−10.55	10.55
7	132	102.58	29.42	29.42
8	141	127.59	13.41	13.41
9	110	138.99	−28.99	28.99
10	131	114.35	16.65	16.65
11		128.50		
				MAD = 20.45

d. For an alpha of .35, MAD = 17.72.
For an alpha of .85, MAD = 20.45.

e. Because it has the lower measure of error (MAD), choose exponential smoothing with alpha = .35.

13. a.



b. The graph shows a negative time series trend in sales.

c. A trend model is appropriate because we want to estimate the decreasing change in sales per time period.

d. Based on the regression analysis, the trend equation is Sales = 1062.86 – 36.24 (time period); sales are declining at a historical rate of 36.24 for each increment of one time period. Based on the MAD, the average forecast error for this model is 111.15.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.83
R Square	0.69
Adjusted R Square	0.67
Standard Error	134.97
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*p*-Value
Regression	1	636183.06	636183.06	34.92	0.00
Residual	16	291455.22	18215.95		
Total	17	927638.28			

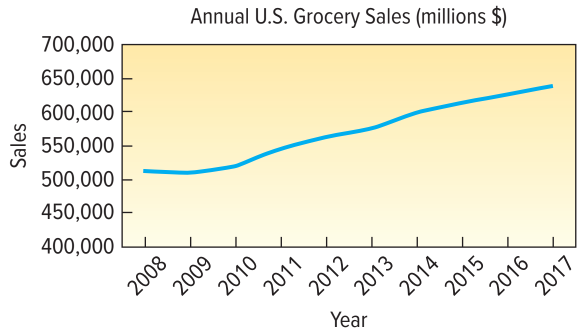
	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	1062.86	66.37	16.01	0.00
Period	−36.24	6.13	−5.91	0.00

Period	Sales	Predicted Sales	Absolute Error
1	1001	1026.62	25.62
2	1129	990.38	138.62
3	841	954.15	113.15
4	1044	917.91	126.09
5	1012	881.67	130.33
6	703	845.44	142.44
7	682	809.20	127.20
8	712	772.97	60.97
9	646	736.73	90.73
10	686	700.49	14.49
11	909	664.26	244.74
12	469	628.02	159.02
13	566	591.78	25.78
14	488	555.55	67.55
15	688	519.31	168.69
16	675	483.07	191.93
17	303	446.84	143.84
18	381	410.60	29.60
			MAD = 111.15

e. Sales are declining at a historical rate of 36.24 for each increment of one time period.

f. Sales (19) = 1062.86 – 36.24 (19) = 374.37
Sales (20) = 1062.86 – 36.24 (20) = 338.13
Sales (21) = 1062.86 – 36.24 (21) = 301.89
The MAD or error associated with each forecast is 111.15.

15. a.



b. The graph shows a positive time series trend in grocery sales.

c. A trend model is appropriate because we want to estimate the increasing change in sales per time period.

d. Predicted annual U.S. grocery sales = −30,990,548.25 + 15,682.503 (year). The forecast error computed with the MAD is 4,373.15.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.992
R Square	0.984
Adjusted R Square	0.982
Standard Error	6443.504
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*p*-Value
Regression	1	2.03E+10	2.03E+10	4.89E+02	1.85E− 08
Residual	8	3.32E+08	4.15E+07		
Total	9	2.06E+10			

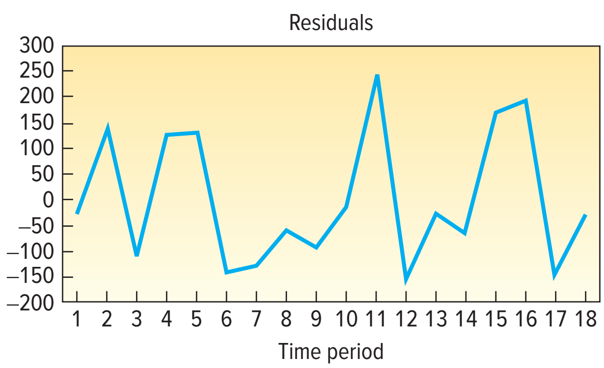
	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	−30990548.248	1427681.970	−21.707	0.000
period	15682.503	709.406	22.107	0.000

Period	Sales	Predicted Sales	Absolute Error
2008	511,222	499,917.84	11,304.16
2009	510,033	515,600.34	5,567.34
2010	520,750	531,282.84	10,532.84
2011	547,476	546,965.35	510.65
2012	563,645	562,647.85	997.15
2013	574,547	578,330.35	3,783.35
2014	599,603	594,012.85	5,590.15
2015	613,159	609,695.36	3,463.64
2016	625,295	625,377.86	82.86
2017	639,161	641,060.36	1,899.36
			MAD = 4,373.15

e. Sales are increasing at a historical rate of $15,682.503 million per year.

f. Annual U.S. grocery sales (2018) = −30,990,548.25 + 15,682.503 (2018) = $656,742.87 (millions).
Annual U.S. grocery sales (2019) = −30,990,548.25 + 15,682.503 (2019) = $672,425.37 (millions).
Annual U.S. grocery sales (2020) = −30,990,548.25 + 15,682.503 (2020) = $688,107.87 (millions).
The forecast error associated with each forecast is 4,373.15.

17. a. The graph of residuals does not show evidence of a pattern or autocorrelation.

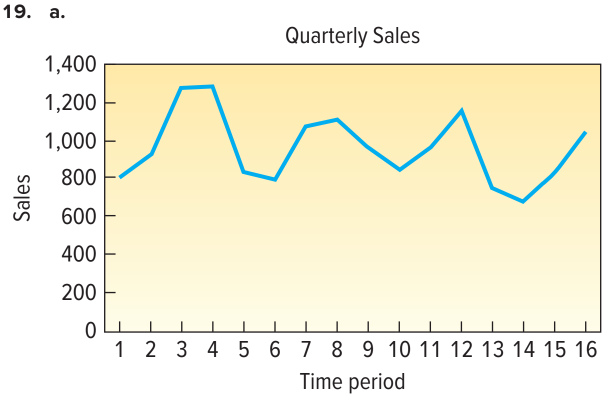


b.

Period	Sales	Forecast	Residuals	Lagged Residuals	Squared Difference	Squared Residuals
1	1001	1026.620	−25.620			656.38
2	1129	990.384	138.616	−25.620	26973.57	19214.52
3	841	954.147	−113.147	138.616	63384.95	12802.30
4	1044	917.911	126.089	−113.147	57234.02	15898.46
5	1012	881.675	130.325	126.089	17.95	16984.72
6	703	845.438	−142.438	130.325	74400.02	20288.66
7	682	809.202	−127.202	−142.438	232.15	16180.33
8	712	772.966	−60.966	−127.202	4387.25	3716.80
9	646	736.729	−90.729	−60.966	885.88	8231.80
10	686	700.493	−14.493	−90.729	5811.98	210.05
11	909	664.257	244.743	−14.493	67203.47	59899.32
12	469	628.020	−159.020	244.743	163025.10	25287.45
13	566	591.784	−25.784	−159.020	17751.92	664.81
14	488	555.548	−67.548	−25.784	1744.20	4562.68
15	688	519.311	168.689	−67.548	55807.60	28455.87
16	675	483.075	191.925	168.689	539.93	36835.21
17	303	446.839	−143.839	191.925	112737.24	20689.56
18	381	410.602	−29.602	−143.839	13049.94	876.30
				Column sums:	665187.17	291455.22
					d = 2.282	
			No autocorrelation		*d* table values: 1.16,1.39	

c. The d-test statistic is 2.282. It is larger than the upper d-critical value of 1.39. Therefore, fail to reject the null hypothesis of "no autocorrelation" and conclude that there is no autocorrelation in the data. We can use the results of the hypothesis tests associated with the regression analysis.

19. a.



b. The quarterly time series shows two patterns, negative trend and seasonality. The seasonality is indicated with quarters 3 and 4 always with high sales, and quarters 1 and 2 always with low sales.

c. A trend model is appropriate because we want to estimate the decreasing change in sales per quarter. A model with seasonal indexes is appropriate because we want to quantify the seasonal effects for each quarter.

d. e.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.262
R Square	0.069
Adjusted *R* Square	0.002
Standard Error	182.671
Observations	16

ANOVA

	df	*SS*	*MS*	*F*	*p*-Value
Regression	1	34552.224	34552.224	1.035	0.326
Residual	14	467162.213	33368.730		
Total	15	501714.438			

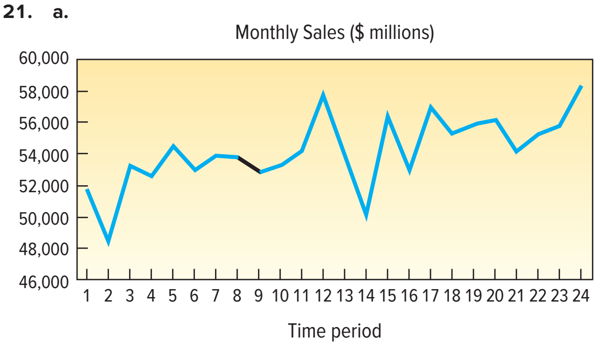
	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	1041.875	95.794	10.876	0.000
period	−10.081	9.907	−1.018	0.326

Period	Quarter	Sales	Trend	Index	Forecast	Absolute Error
1	1	812	1031.794	0.787	893.851	81.851
2	2	920	1021.713	0.900	856.087	63.913
3	3	1268	1011.632	1.253	1092.042	175.958
4	4	1280	1001.551	1.278	1218.004	61.996
5	1	832	991.471	0.839	858.918	26.918
6	2	791	981.390	0.806	822.300	31.300
7	3	1071	971.309	1.103	1048.514	22.486
8	4	1109	961.228	1.154	1168.966	59.966
9	1	965	951.147	1.015	823.986	141.014
10	2	844	941.066	0.897	788.513	55.487
11	3	961	930.985	1.032	1004.985	43.985
12	4	1160	920.904	1.260	1119.928	40.072
13	1	751	910.824	0.825	789.053	38.053
14	2	674	900.743	0.748	754.727	80.727
15	3	828	890.662	0.930	961.456	133.456
16	4	1033	880.581	1.173	1070.890	37.890
						MAD = 68.442

Quarter	Index
1	0.866
2	0.838
3	1.079
4	1.216

f. Sales = [1041.875 – 10.081 (Time period)] [Quarterly index for the time period]
Period 17 sales = [1041.875 – 10.081 (17)][.866]
= [870.498][.866] = 753.851
Period 18 sales = [1041.875 – 10.081 (18)][.838]
= [860.417][.838] = 721.029
Period 19 sales = [1041.875 – 10.081 (19)][1.079]
= [850.336][1.079] = 917.513
Period 20 sales = [1041.875 – 10.081 (20)][1.216]
= [840.255][1.216] = 1021.750

21. a.



b. The monthly time series shows two patterns, positive trend and seasonality. The seasonality is indicated with month 2 (February), the lowest of the months 1 through 12, and month 12 (December), the highest sales among the 12 months.

c. A trend model is appropriate because we want to estimate the increasing change in sales per month. A model with seasonal indexes is appropriate because we want to quantify the seasonal effects for each month.

d. & e.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.636
R Square	0.405
Adjusted *R* Square	0.378
Standard Error	1818.644
Observations	24

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	49456604.543	49456604.543	14.953	0.001
Residual	22	72764221.957	3307464.634		
Total	23	122220826.500			

	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	51576.022	766.286	67.306	0.000
period	207.378	53.629	3.867	0.001

Period	Month-Year	Month	Sales ($ millions)	Trend	Index	Forecast	Absolute Error
1	Jan-2017	1	51756	51783.400	0.999	51561.725	194.275
2	Feb-2017	2	48335	51990.778	0.930	48047.961	287.039
3	Mar-2017	3	53311	52198.157	1.021	53598.497	287.497
4	Apr-2017	4	52512	52405.535	1.002	51476.380	1035.620
5	May-2017	5	54479	52612.913	1.035	54469.068	9.932
6	Jun-2017	6	52941	52820.291	1.002	52852.076	88.924
7	Jul-2017	7	53859	53027.670	1.016	53613.256	245.744
8	Aug-2017	8	53769	53235.048	1.010	53716.695	52.305
9	Sep-2017	9	52865	53442.426	0.989	52263.675	601.325
10	Oct-2017	10	53296	53649.804	0.993	53038.863	257.137
11	Nov-2017	11	54191	53857.183	1.006	53766.627	424.373
12	Dec-2017	12	57847	54064.561	1.070	56775.978	1071.022
13	Jan-2018	1	53836	54271.939	0.992	54039.611	203.611
14	Feb-2018	2	50047	54479.317	0.919	50347.778	300.778
15	Mar-2018	3	56455	54686.696	1.032	56153.797	301.203
16	Apr-2018	4	52836	54894.074	0.963	53920.797	1084.797
17	May-2018	5	57035	55101.452	1.035	57045.402	10.402
18	Jun-2018	6	55249	55308.830	0.999	55342.113	93.113
19	Jul-2018	7	55872	55516.209	1.006	56129.276	257.276
20	Aug-2018	8	56173	55723.587	1.008	56227.750	54.750
21	Sep-2018	9	54068	55930.965	0.967	54697.326	629.326
22	Oct-2018	10	55230	56138.343	0.984	55499.064	269.064
23	Nov-2018	11	55807	56345.722	0.990	56250.982	443.982
24	Dec-2018	12	58269	56553.100	1.030	59389.321	1120.321
							MAD = 388.492

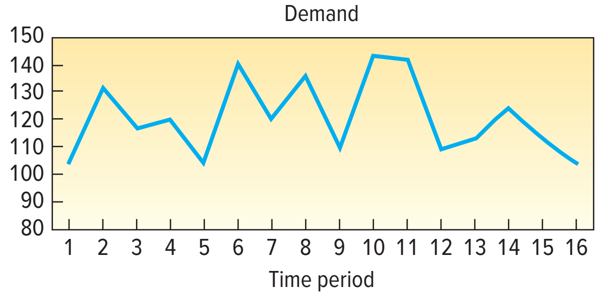
Month	Index	Month	Index
1	0.996	7	1.011
2	0.924	8	1.009
3	1.027	9	0.978
4	0.982	10	0.989
5	1.035	11	0.998
6	1.001	12	1.050

f. Sales = [51,576.022 + 207.378 (Time period)] [Quarterly index for the time period]
Period 17 sales = [51,576.022 + 207.378 (25)][0.996]
= [56,760.478][0.996] = 56,517.497
Period 18 sales = [51,576.022 + 207.378 (26)][0.924]
= [56,967.857][0.924] = 52,647.594
Period 19 sales = [51,576.022 + 207.378 (27)][1.027]
= [57,175.235][1.027] = 58,709.097
Period 20 sales = [51,576.022 + 207.378 (28)][0.982]
= [57,382.613][0.982] = 56,365.215

23. Both techniques are used when a time series has no trend or seasonality. The pattern only shows random variation. Simple moving average selects a fixed number of data points from the past and uses the average as a forecast. All past data is equally weighted. Simple exponential smoothing uses all past available data and adjusts the weights of the past information based on the forecaster's choice of a smoothing constant.

25. The time series has no seasonality.

27. **a.** The graph does not show any trend or seasonality. The graph shows a stationary pattern. Simple moving average models would be a good choice to compute forecasts.

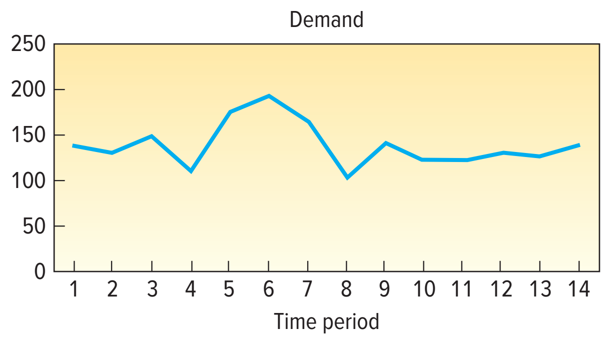


b.

Period	Demand	5-Month Moving Average	Absolute Error
1	104		
2	132		
3	117		
4	120		
5	104		
6	141	115.4	25.6
7	120	122.8	2.8
8	136	120.4	15.6
9	109	124.2	15.2
10	143	122.0	21.0
11	142	129.8	12.2
12	109	130.0	21.0
13	113	127.8	14.8
14	124	123.2	0.8
15	113	126.2	13.2
16	104	120.2	16.2
17		112.6	
			MAD = 14.4

d. The forecast demand for period 17 is 112.6 units.
e. The forecasting error is estimated with the MAD, which is 14.4. Applying the error, the forecast demand is most likely between 98.2, or 112.6 – 14.4, and 127.0, or 112.6 + 14.4.

29. **a.** The graph does not show any trend or seasonality. The graph shows a stationary pattern. Simple exponential smoothing models would be a good choice to compute forecasts.



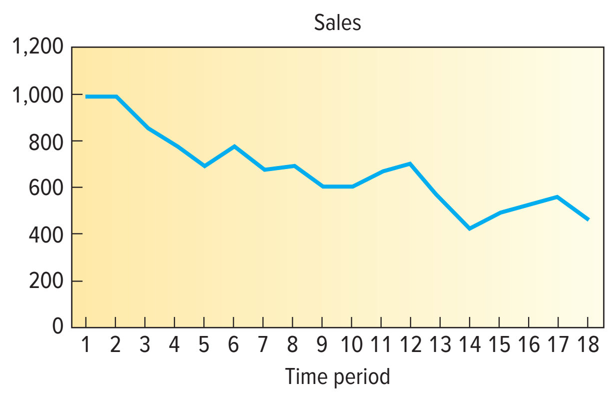
b.–d.

Period	Alpha = Demand	0.4 Exp Smooth	Error	Absolute Error
1	138			
2	131	138.000	–7.000	7.000
3	149	135.200	13.800	13.800
4	110	140.720	–30.720	30.720
5	175	128.432	46.568	46.568
6	194	147.059	46.941	46.941
7	166	165.836	0.164	0.164
8	103	165.901	–62.901	62.901
9	142	140.741	1.259	1.259
10	122	141.244	–19.244	19.244
11	121	133.547	–12.547	12.547
12	130	128.528	1.472	1.472
13	126	129.117	–3.117	3.117
14	140	127.870	12.130	12.130
15		132.722		
			MAD = 19.836	

Period	Alpha = Demand	0.9 Exp Smooth	Error	Absolute Error
1	138			
2	131	138.000	–7.000	7.000
3	149	131.700	17.300	17.300
4	110	147.270	–37.270	37.270
5	175	113.727	61.273	61.273
6	194	168.873	25.127	25.127
7	166	191.487	–25.487	25.487
8	103	168.549	–65.549	65.549
9	142	109.555	32.445	32.445
10	122	138.755	–16.755	16.755
11	121	123.676	–2.676	2.676
12	130	121.268	8.732	8.732
13	126	129.127	–3.127	3.127
14	140	126.313	13.687	13.687
15		138.631		
			MAD = 24.341	

e. Comparing the MAD's the simple exponential smoothing model with $\alpha = 0.4$ forecasts with less error.

31. **a.**



b. The graph shows a downward, negative trend in sales.
c. Forecasting with a trend model will reveal the average, per period, change in sales.
d. The Trend forecast model is Sales = 930.954 – 27.457 (time period).
Based on the MAD, the forecasting error is ± 58.525.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.900
R Square	0.811
Adjusted *R* Square	0.799
Standard Error	72.991
Observations	18

ANOVA

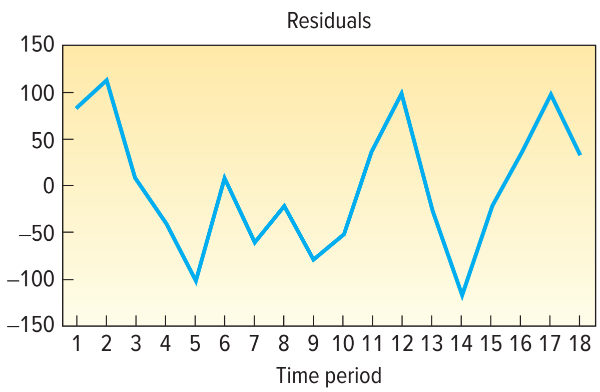
	df	SS	MS	*F*	*p*-Value
Regression	1	365262.764	365262.764	68.559	0.000
Residual	16	85243.014	5327.688		
Total	17	450505.778			

	Coefficient	Standard Error	*t*-Stat	*p*-Value
Intercept	930.954	35.894	25.936	0.000
Time period	– 27.457	3.316	– 8.280	0.000

Period	Sales	Trend Forecast	Absolute Error
1	988	903.497	84.503
2	990	876.040	113.960
3	859	848.583	10.417
4	781	821.126	40.126
5	691	793.668	102.668
6	776	766.211	9.789
7	677	738.754	61.754
8	690	711.297	21.297
9	605	683.840	78.840
10	604	656.383	52.383
11	670	628.925	41.075
12	703	601.468	101.532
13	550	574.011	24.011
14	427	546.554	119.554
15	493	519.097	26.097
16	524	491.639	32.361
17	563	464.182	98.818
18	471	436.725	34.275
			MAD = 58.525

e. Based on the slope from the regression analysis, sales are decreasing at a rate of 27.457 units per period.

33. a.

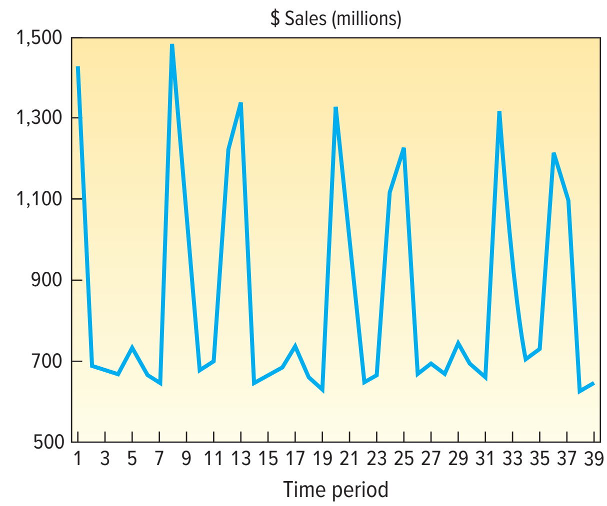


b.

Period	Sales	Predicted	Residuals	Lagged Residual	Squared Differences	Squared Residuals
1	988	903.4971	84.5029			7140.7442
2	990	876.0399	113.9601	84.5029	867.7250	12986.9036
3	859	848.5827	10.4173	113.9601	10721.1172	108.5195
4	781	821.1256	−40.1256	10.4173	2554.5774	1610.0605
5	691	793.6684	−102.6684	−40.1256	3911.6053	10540.7976
6	776	766.2112	9.7888	−102.6684	12646.6156	95.8203
7	677	738.7540	−61.7540	9.7888	5118.3762	3813.5617
8	690	711.2969	−21.2969	−61.7540	1636.7828	453.5567
9	605	683.8397	−78.8397	−21.2969	3311.1770	6215.6979
10	604	656.3825	−52.3825	−78.8397	699.9820	2743.9289
11	670	628.9254	41.0746	−52.3825	8734.2431	1687.1267
12	703	601.4682	101.5318	41.0746	3655.0697	10308.7104
13	550	574.0110	−24.0110	101.5318	15761.0016	576.5285
14	427	546.5538	−119.5538	−24.0110	9128.4319	14293.1196
15	493	519.0967	−26.0967	−119.5538	8734.2431	681.0358
16	524	491.6395	32.3605	−26.0967	3417.2410	1047.2026
17	563	464.1823	98.8177	32.3605	4416.5558	9764.9342
18	471	436.7251	34.2749	98.8177	4165.7766	1174.7656
				Sums:	99480.5211	85243.0141
						d = 1.1670

c. Based on d = 1.1670, the result of the hypothesis test is inconclusive. We cannot make any determination regarding the presence of autocorrelation in the data.

35. a.



b. The time series has definite seasonality with peaks occurring in December and January, followed by a regular peak in August. There may be a slight negative trend over the time span.

c. The choice of this forecasting model is appropriate because of the seasonality and the hint of a small negative trend.

d.–f.

SUMMARY OUTPUT

Regression Statistics	
Multiple *R*	0.0460644
R Square	0.0021219
Adjusted *R* Square	– 0.024848
Standard Error	276.24029
Observations	39

ANOVA

	df	SS	MS	*F*	*p*-Value
Regression	1	6003.82915	6003.829	0.078678	0.7806592
Residual	37	2823421.761	76308.7		
Total	38	2829425.59			

	Coefficients	Standard Error	*t*-Stat	*p*-Value
Intercept	877.94602	90.19686822	9.733664	9.53E-12
Period	– 1.102429	3.930280461	– 0.2805	0.780659

Month Number	Sales ($ millions)	Period	Trend	Index	Forecast	Absolute Error
1	1428	1	876.8436	1.6286	1305.0642	122.9358
2	687	2	875.7412	0.7845	673.3858	13.6142
3	679	3	874.6387	0.7763	686.8877	7.8877
4	669	4	873.5363	0.7659	686.5074	17.5074
5	738	5	872.4339	0.8459	753.2469	15.2469
6	673	6	871.3314	0.7724	685.2503	12.2503
7	647	7	870.2290	0.7435	655.8122	8.8122
8	1484	8	869.1266	1.7075	1401.3967	82.6033
9	1024	9	868.0242	1.1797	990.5175	33.4825
10	675	10	866.9217	0.7786	685.7084	10.7084
11	702	11	865.8193	0.8108	708.7451	6.7451
12	1216	12	864.7169	1.4062	1202.6100	13.3900
1	1346	13	863.6144	1.5586	1285.3744	60.6256
2	651	14	862.5120	0.7548	663.2134	12.2134
3	667	15	861.4096	0.7743	676.4983	9.4983
4	689	16	860.3072	0.8009	676.1107	12.8893
5	741	17	859.2047	0.8624	741.8250	0.8250
6	664	18	858.1023	0.7738	674.8464	10.8464
7	629	19	856.9999	0.7340	645.8426	16.8426
8	1334	20	855.8974	1.5586	1380.0658	46.0658
9	957	21	854.7950	1.1196	975.4215	18.4215
10	649	22	853.6926	0.7602	675.2445	26.2445
11	663	23	852.5901	0.7776	697.9160	34.9160
12	1117	24	851.4877	1.3118	1184.2115	67.2115
1	1231	25	850.3853	1.4476	1265.6846	34.6846
2	669	26	849.2829	0.7877	653.0411	15.9589
3	694	27	848.1804	0.8182	666.1090	27.8910
4	670	28	847.0780	0.7910	665.7140	4.2860
5	746	29	845.9756	0.8818	730.4032	15.5968
6	687	30	844.8731	0.8131	664.4425	22.5575
7	661	31	843.7707	0.7834	635.8730	25.1270
8	1324	32	842.6683	1.5712	1358.7349	34.7349
9	946	33	841.5659	1.1241	960.3255	14.3255
10	701	34	840.4634	0.8341	664.7807	36.2193
11	728	35	839.3610	0.8673	687.0868	40.9132
12	1219	36	838.2586	1.4542	1165.8130	53.1870
1	1104	37	837.1561	1.3188	1245.9948	141.9948
2	626	38	836.0537	0.7488	642.8688	16.8688
3	645	39	834.9513	0.7725	655.7196	10.7196
						MAD = 29.6628

Month	Index	Month	Index
1	1.4884	7	0.7536
2	0.7689	8	1.6124
3	0.7853	9	1.1411
4	0.7859	10	0.7910
5	0.8634	11	0.8186
6	0.7864	12	1.3908

e. The regression analysis shows a small decline of –$1.1024 million per month.

f. Book sales are highest in December (39.08% higher than average) and January (48.84% higher than average). There is also a "back-to-school" effect in August and September. Sales are the highest in August, 61.24% higher than average. Book sales are lowest from February through July.

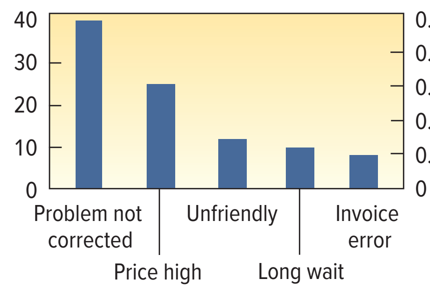
g.

Year	Month	Forecast Sales
2019	April	655.3173
2019	May	718.9813
2019	June	654.0385
2019	July	625.9034
2019	August	1337.4039
2019	September	945.2295
2019	October	654.3168
2019	November	676.2576
2019	December	1147.4145

h. Given the MAD's estimate of error in the forecasting model, the forecasts could be off by ± $29.6628 million. For July, this is 29.6628/625.9034, which is a 4.7% error. The percentage errors for the other months would be even less. Because this time series approach to forecasting replicates historical patterns in sales, the disclaimer is that the forecasts assume that the future sales will be similar to sales over the previous 39 months.

CHAPTER 19

1.



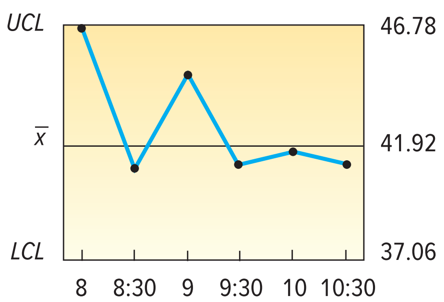
Count	38	23	12	10	8
Percent	42	25	13	11	9
Cum %	42	67	80	91	100

About 67% of the complaints concern the problem not being corrected and the price being too high.

3. Chance variation is random in nature; because the cause is a variety of factors, it cannot be entirely eliminated. Assignable variation is not random; it is usually due to a specific cause and can be eliminated.

5. **a.** The A_2 factor is 0.729.
b. The value for D_3 is 0, and for D_4 it is 2.282.

7. a.



Time	$\bar{x}$, Arithmetic Means	*R*, Range
8:00 a.m.	46	16
8:30 a.m.	40.5	6
9:00 a.m.	44	6
9:30 a.m.	40	2
10:00 a.m.	41.5	9
10:30 a.m.	39.5	1
	251.5	40

$\bar{\bar{x}} = \frac{251.5}{6} = 41.92 \qquad \bar{R} = \frac{40}{6} = 6.67$

$UCL = 41.92 + 0.729(6.67) = 46.78$

$LCL = 41.92 - 0.729(6.67) = 37.06$

b. Interpreting, the mean reading was 341.92 degrees Fahrenheit. If the oven continues operating as evidenced by the first six hourly readings, about 99.7% of the mean readings will lie between 337.06 degrees and 346.78 degrees.

9. a. The fraction defective is 0.0507. The upper control limit is 0.0801 and the lower control limit is 0.0213.

b. Yes, the seventh and ninth samples indicate the process is out of control.

c. The process appears to stay the same.

11. $\bar{c} = \frac{37}{14} = 2.64$

$2.64 \pm 3\sqrt{2.64}$

The control limits are 0 and 7.5. The process is out of control on the seventh day.

13. $\bar{c} = \frac{6}{11} = 0.545$

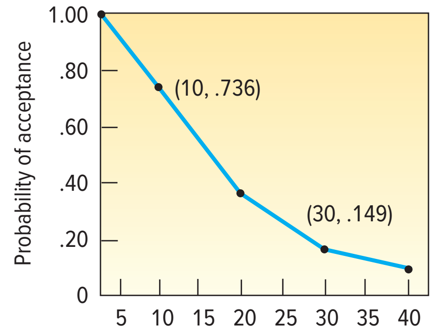
$0.545 \pm 2\sqrt{0.545} = 0.545 \pm 2.215$

The control limits are from 0 to 2.760, so there are no receipts out of control.

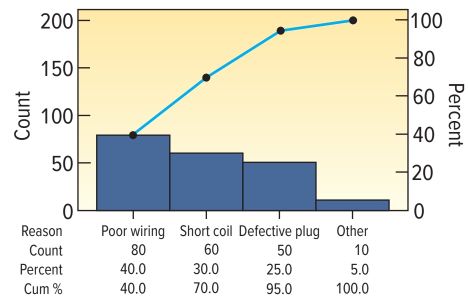
15.

Percent Defective	Probability of Accepting Lot
10	.889
20	.558
30	.253
40	.083

17. $P(x \leq 1 \mid n = 10, \pi = .10) = .736$
$P(x \leq 1 \mid n = 10, \pi = .20) = .375$
$P(x \leq 1 \mid n = 10, \pi = .30) = .149$
$P(x \leq 1 \mid n = 10, \pi = .40) = .046$



19.



21. a. Mean: $UCL = 10.0 + 0.577(0.25) = 10.0 + 0.14425 = 10.14425$

$LCL = 10.0 - 0.577(0.25) = 10.0 - 0.14425 = 9.85575$

Range: $UCL = 2.115(0.25) = 0.52875$

$LCL = 0(0.25) = 0$

b. The mean is 10.16, which is above the upper control limit and is out of control. There is too much cola in the soft drinks. The process is in control for variation; an adjustment is needed.

23. a. $\bar{\bar{x}} = \frac{611.3333}{20} = 30.57$

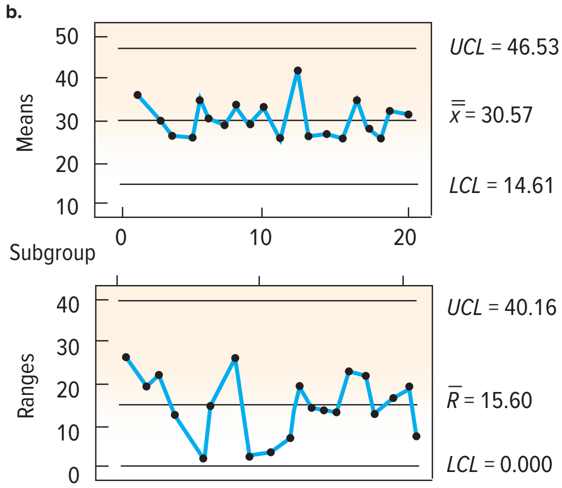
$\bar{R} = \frac{312}{20} = 15.6$

Mean: $UCL = 30.5665 + (1.023)(15.6) = 46.53$

$LCL = 30.5665 - (1.023)(15.6) = 14.61$

Range: $UCL = 2.575(15.6) = 40.17$

b.



c. The points all seem to be within the control limits. No adjustments are necessary.

25. $\bar{\bar{X}} = \frac{-0.5}{18} = -0.0278 \qquad \bar{R} = \frac{27}{18} = 1.5$

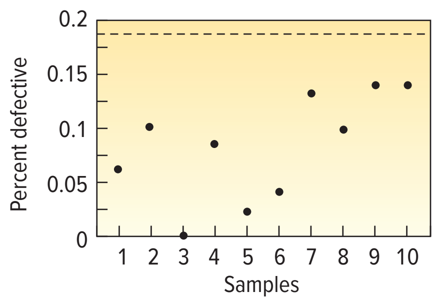
$UCL = -.0278 + (0.729)(1.5) = 1.065$
$LCL = -.0278 - (0.729)(1.5) = -1.121$
$UCL = 2.282(1.5) = 3.423$

The X-bar chart indicates that the "process" was in control. However, the R-bar chart indicates that the performance on hole 12 was outside the limits.

27. **a.** $p = \frac{40}{10(50)} = 0.08 \qquad 3\sqrt{\frac{0.08(0.92)}{50}} = 0.115$

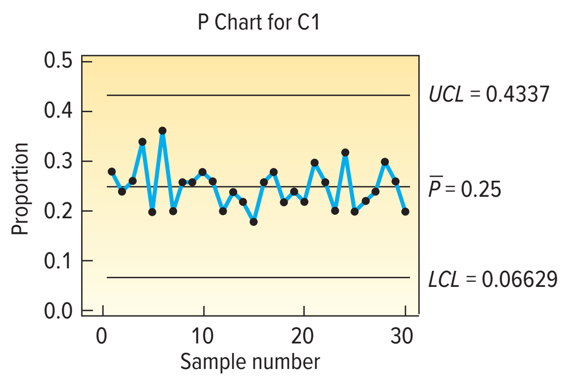
$UCL = 0.08 + 0.115 = 0.195$
$LCL = 0.08 - 0.115 = 0$

b.



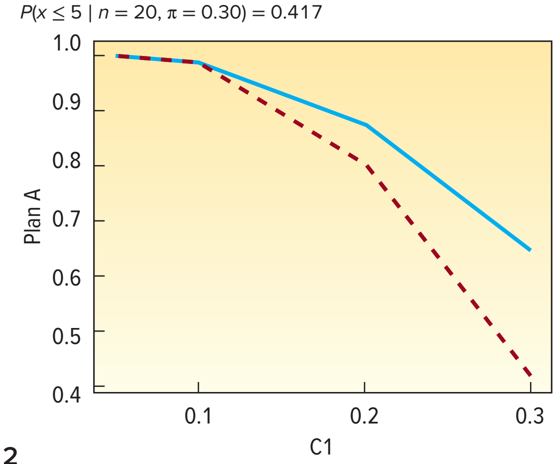
c. There are no points that exceed the limits.

29.



These sample results indicate that the odds are much less than 50-50 for an increase. The percent of stocks that increase is "in control" around 0.25, or 25%. The control limits are 0.06629 and 0.4337.

31. $P(x \le 3 \mid n = 10, \pi = 0.05) = 0.999$
$P(x \le 3 \mid n = 10, \pi = 0.10) = 0.987$
$P(x \le 3 \mid n = 10, \pi = 0.20) = 0.878$
$P(x \le 3 \mid n = 10, \pi = 0.30) = 0.649$
$P(x \le 5 \mid n = 20, \pi = 0.05) = 0.999$
$P(x \le 5 \mid n = 20, \pi = 0.10) = 0.989$
$P(x \le 5 \mid n = 20, \pi = 0.20) = 0.805$
$P(x \le 5 \mid n = 20, \pi = 0.30) = 0.417$

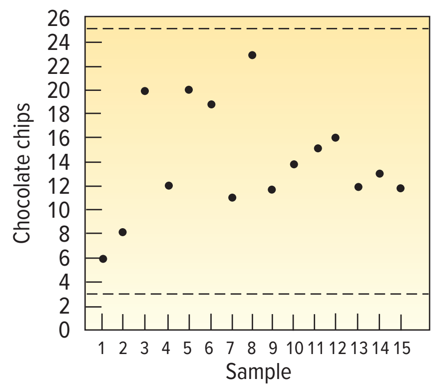


The solid line is the operating characteristic curve for the first plan, and the dashed line, the second. The supplier would prefer the first because the probability of acceptance is higher (above). However, if he is really sure of his quality, the second plan seems higher at the very low range of defect percentages and might be preferred.

33. **a.** $\bar{c} = \frac{213}{15} = 14.2; \; 3\sqrt{14.2} = 11.30$

$UCL = 14.2 + 11.3 = 25.5$
$LCL = 14.2 - 11.3 = 2.9$

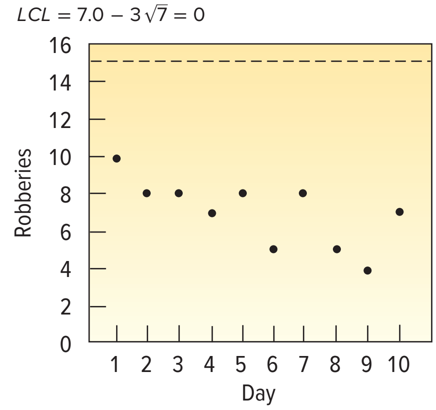
b.



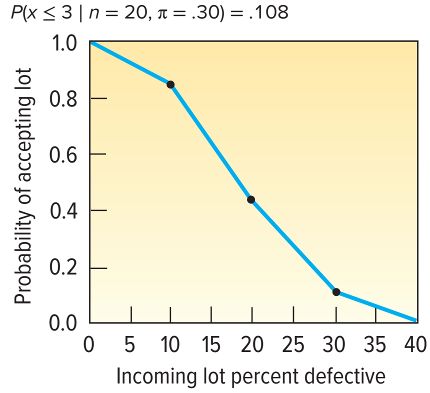
c. All the points are in control.

35. $\bar{c} = \frac{70}{10} = 7.0$

$UCL = 7.0 + 3\sqrt{7} = 14.9$
$LCL = 7.0 - 3\sqrt{7} = 0$



37. $P(x \le 3 \mid n = 20, \pi = .10) = .867$
$P(x \le 3 \mid n = 20, \pi = .20) = .412$
$P(x \le 3 \mid n = 20, \pi = .30) = .108$

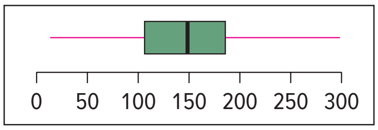


APPENDIX C: ANSWERS

Answers to Odd-Numbered Review Exercises

REVIEW OF CHAPTERS 1–4
PROBLEMS

1. **a.** Mean is 147.9. Median is 148.5. Standard deviation is 69.24.
b. The first quartile is 106. The third quartile is 186.25.
c.



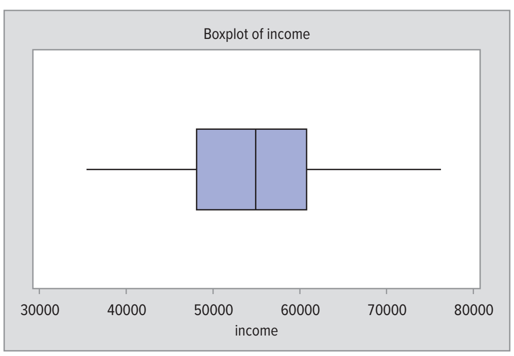
There are no outliers. The distribution is symmetric. The whiskers and the boxes are about equal on the two sides.

d. $2^6 = 64$, use six classes; $i = \frac{299 - 14}{6} = 47.5$, use $i = 50$.

Amount	Frequency
$ 0 up to $ 50	3
50 up to 100	8
100 up to 150	15
150 up to 200	13
200 up to 250	7
250 up to 300	7
Total	50

e. Answers will vary but include all of the above information.

3. **a.** Mean is $55,224. Median is $54,916. Standard deviation is $9,208.
b. The first quartile is $48,060. The third quartile is 60,730.
c.



The distribution is symmetric with no outliers.

d.

Amounts	Frequency
35000–42999	5
43000–50999	12
51000–58999	18
59000–66999	9
67000–74999	6
75000–82999	1
Total	**51**

e. Answers will vary but include all of the above information.

5. **a.** Box plot.
b. Median is 48, the first quartile is 24, and the third quartile is 84.
c. Positively skewed with the long tail to the right.
d. You cannot determine the number of observations.

REVIEW OF CHAPTERS 5–7
PROBLEMS

1. **a.** .035
b. .018
c. .648

3. **a.** .0401
b. .6147
c. 7,440

5. **a.** $\mu = 1.10$
$\sigma = 1.18$
b. About 550
c. $\mu = 1.833$

REVIEW OF CHAPTERS 8 AND 9
PROBLEMS

1. $z = \frac{8.8 - 8.6}{2.0/\sqrt{35}} = 0.59$, $.5000 - .2224 = .2776$

3. $160 \pm 2.426 \frac{20}{\sqrt{40}}$, 152.33 up to 167.67

5. $985.5 \pm 2.571 \frac{115.5}{\sqrt{6}}$, 864.27 up to 1,106.73

7. $240 \pm 2.131 \frac{35}{\sqrt{16}}$, 221.35 up to 258.65

Because 250 is in the interval, the evidence does *not* indicate an increase in production.

9. $n = \left[\frac{1.96(25)}{4}\right]^2 = 150$

11. $n = .08(.92)\left(\frac{2.33}{0.22}\right)^2 = 999$

13. $n = .4(.6)\left(\frac{2.33}{0.03}\right)^2 = 1{,}448$

REVIEW OF CHAPTERS 10–12
PROBLEMS

1. H_0: $\mu \geq 36$; H_1: $\mu < 36$. Reject H_0 if $t < -1.683$.

$$t = \frac{35.5 - 36.0}{0.9/\sqrt{42}} = -3.60$$

Reject H_0. The mean height is less than 36 inches.

3. H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$
Reject H_0 if $t < -2.845$ or $t > 2.845$.

$$s_p^2 = \frac{(12 - 1)(5)^2 + (10 - 1)(8)^2}{12 + 10 + 2} = 42.55$$

$$t = \frac{250 - 252}{\sqrt{42.55\left(\frac{1}{12} + \frac{1}{10}\right)}} = -0.716$$

H_0 is not rejected. There is no difference in the mean strength of the two glues.

5. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: The means are not all the same.
H_0 rejected if $F > 3.29$.

Source	SS	*df*	MS	*F*
Treatments	20.736	3	6.91	1.04
Error	100.00	15	6.67	
Total	120.736	18		

H_0 is not rejected. There is no difference in the mean sales.

7. a. From the graph, marketing salaries may be acting differently.
   b. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$
      H_1: At least one mean is different (for four majors).
      H_0: $\mu_1 = \mu_2 = \mu_3$
      H_1: At least one mean is different (for 3 years).
      H_0: There is no interaction.
      H_1: There is interaction.
   c. The *p*-value (.482) is high. Do not reject the hypothesis of no interaction.
   d. The *p*-value for majors is small (.034 < .05), so there is a difference among mean salaries by major. There is no difference from one year to the next in mean salaries (.894 > .05).

REVIEW OF CHAPTERS 13 AND 14
PROBLEMS

1. a. Profit
   b. $\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$
   c. \$163,200
   d. About 86% of the variation in net profit is explained by the four variables.
   e. About 68% of the net profits would be within \$3,000 of the estimates; about 95% would be within 2(\$3,000), or \$6,000, of the estimates; and virtually all would be within 3(\$3,000), or \$9,000, of the estimates.
3. a. 0.9261
   b. 2.0469, found by $\sqrt{83.8/20}$
   c. H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
      H_1: Not all coefficients are zero.
      Reject if $F > 2.87$; computed $F = 62.697$, found by 162.70/4.19.
   d. Could delete x_2 because *t*-ratio (1.29) is less than the critical *t*-value of 2.086. Otherwise, reject H_0 for x_1, x_3, and x_4 because all of those *t*-ratios are greater than 2.086.

REVIEW OF CHAPTERS 15 AND 16
PROBLEMS

1. H_0: Median $\leq$ 60
   H_1: Median > 60
   $\mu = 20(.5) = 10$
   $\sigma = \sqrt{20(.5)(.5)} = 2.2361$
   H_0 is rejected if $z > 1.65$. There are 16 observations greater than 60.

$$z = \frac{15.5 - 10.0}{2.2361} = 2.46$$

   Reject H_0. The median sales per day are greater than 60.
3. H_0: The population lengths are the same.
   H_1: The population lengths are not the same.
   H_0 is rejected if H is > 5.991.

$$H = \frac{12}{24(24+1)}\left[\frac{(104.5)^2}{7} + \frac{(125.5)^2}{9} + \frac{(70)^2}{8}\right] - 3(24+1)$$
$$= 78.451 - 75 = 3.451$$

   Do not reject H_0. The population lengths are the same.

REVIEW OF CHAPTERS 17 AND 18
PROBLEMS

1. a. 156.6, found by (16,915/10,799)100
   b. 153.0, found by (16,615/11,056.7)100. Note: 11,056.7 is the average for the period 2008 to 2010.
   c. 9,535 + 854.4*t* and 18,079, found by 9,535 + 854.4 (10)
3. 55.44, found by 1.20[3.5 + (0.7)(61)], and 44.73, found by 0.90[3.5 + (0.7)(66)]

APPENDIX C: ANSWERS

Solutions to Practice Tests

PRACTICE TEST (AFTER CHAPTER 4)
PART 1

1. statistics
2. descriptive statistics
3. population
4. quantitative and qualitative
5. discrete
6. nominal
7. nominal
8. zero
9. seven
10. 50
11. variance
12. never
13. median

PART 2

1. $\sqrt[3]{(1.18)(1.04)(1.02)} = 1.0777$, *or* 7.77%
2. **a.** \$30,000
   b. 105
   c. 52
   d. 0.19, found by 20/105
   e. 165
   f. 120 and 330
3. **a.** 70
   b. 71.5
   c. 67.8
   d. 28
   e. 9.34
4. \$44.20, found by [(200)\$36 + (300)\$40 + (500)\$50]/1,000
5. **a.** pie chart
   b. 11.1
   c. three times
   d. 65%

PRACTICE TEST (AFTER CHAPTER 7)
PART 1

1. never
2. experiment
3. event
4. joint
5. **a.** permutation
   b. combination
6. one
7. three or more outcomes
8. infinite
9. one
10. 0.2764
11. 0.0475
12. independent
13. mutually exclusive
14. only two outcomes
15. bell-shaped

PART 2

1. **a.** 0.0526, found by (5/20)(4/19)
   b. 0.4474, found by 1 – (15/20)(14/19)
2. **a.** 0.2097, found by $16(.15)(.85)^{15}$
   b. 0.9257, found by $1 - (.85)^{16}$
3. 720, found by $6 \times 5 \times 4 \times 3 \times 2$
4. **a.** 2.2, found by .2(1) + .5(2) + .2(3) + .1(4)
   b. 0.76, found by .2(1.44) + .5(0.04) + .2(0.64) + .1(3.24)
5. **a.** 0.1808. The *z*-value for \$2,000 is 0.47, found by (2,000 – 1,600)/850.
   b. 0.4747, found by 0.2939 + 0.1808
   c. 0.0301, found by 0.5000 – 0.4699
6. **a.** contingency table
   b. 0.625, found by 50/80
   c. 0.75, found by 60/80
   d. 0.40, found by 20/50
   e. 0.125, found by 10/80
7. **a.** 0.0498, found by $\frac{3^0 e^{-3}}{0!}$
   b. 0.2240, found by $\frac{3^3 e^{-3}}{3!}$
   c. 0.1847, found by 1 – [0.0498 + 0.1494 + 0.2240 + 0.2240 + 0.1680]
   d. .0025

PRACTICE TEST (AFTER CHAPTER 9)
PART 1

1. random sample
2. sampling error
3. standard error
4. become smaller
5. point estimate
6. confidence interval
7. population size
8. proportion
9. positively skewed
10. 0.5

PART 2

1. 0.0351, found by 0.5000 – 0.4649. The corresponding
   $z = \frac{11 - 12.2}{2.3/\sqrt{12}} = -1.81$
2. **a.** The population mean is unknown.
   b. 9.3 years, which is the sample mean
   c. 0.3922, found by $2/\sqrt{26}$
   d. The confidence interval is from 8.63 up to 9.97, found by
   $9.3 \pm 1.708\left(\frac{2}{\sqrt{26}}\right)$
3. 2,675, found by $.27(1 - .27)\left(\frac{2.33}{.02}\right)^2$
4. The confidence interval is from 0.5459 up to 0.7341, found by
   $.64 \pm 1.96\sqrt{\frac{.64(1 - .64)}{100}}$

PRACTICE TEST (AFTER CHAPTER 12)
PART 1

1. null hypothesis
2. significance level
3. *p*-value
4. standard deviation
5. normality
6. test statistic
7. split evenly between the two tails
8. range from negative infinity to positive infinity

9. independent
10. three and 20

PART 2

1. $H_0: \mu \le 90$ $H_1: \mu > 90$ If $t > 2.567$, reject H_0.

$$t = \frac{96 - 90}{12/\sqrt{18}} = 2.12$$

Do not reject the null. The mean time in the park could be 90 minutes.

2. $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \ne \mu_2$
$df = 14 + 12 - 2 = 24$
If $t < -2.064$ or $t > 2.064$, then reject H_0.

$$s_p^2 = \frac{(14-1)(30)^2 + (12-1)(40)^2}{14 + 12 - 2} = 1{,}220.83$$

$$t = \frac{837 - 797}{\sqrt{1{,}220.83\left(\frac{1}{14} + \frac{1}{12}\right)}} = \frac{40.0}{13.7455} = 2.910$$

Reject the null hypothesis. There is a difference in the mean miles traveled.

3. a. three, because there are 2 *df* between groups.
   b. 21, found by the total degrees of freedom plus 1.
   c. If the significance level is .05, the critical value is 3.55.
   d. $H_0: \mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same.
   e. At a 5% significance level, the null hypothesis is rejected.
   f. At a 5% significance level, we can conclude the treatment means differ.

PRACTICE TEST (AFTER CHAPTER 14) PART 1

1. vertical
2. interval
3. zero
4. −0.77
5. never
6. 7
7. decrease of .5
8. −0.9
9. zero
10. unlimited
11. linear
12. residual
13. two
14. correlation matrix
15. normal distribution

PART 2

1. a. 30
   b. The regression equation is $\hat{y} = 90.619X - 0.9401$. If X is zero, the line crosses the vertical axis at −0.9401. As the independent variable increases by one unit, the dependent variable increases by 90.619 units.
   c. 905.2499
   d. 0.3412, found by 129.7275/380.1667. Thirty-four percent of the variation in the dependent variable is explained by the independent variable.
   e. 0.5842, found by $\sqrt{0.3412}$ $H_0: p \ge 0$ $H_1: p < 0$
   Using a significance level of .01, reject H_0 if $t > 2.467$.

$$t = \frac{0.5842\sqrt{30-2}}{\sqrt{1-(0.5842)^2}} = 3.81$$

Reject H_0. There is a negative correlation between the variables.

2. a. 30
   b. 4
   c. 0.5974, found by 227.0928/380.1667
   d. $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ H_1: Not all βs are 0. Reject H_0 if $F > 4.18$ (using a 1% level of significance). Since the computed value of F is 9.27, reject H_0. Not all of the regression coefficients are zero.
   e. Reject H_0 if $t > 2.787$ or $t < -2.787$ (using a 1% level of significance). Drop variable 2 initially and then rerun. Perhaps you will delete variable(s) 1 and/or 4 also.

PRACTICE TEST (AFTER CHAPTER 16) PART 1

1. nominal
2. at least 30 observations
3. two
4. 6
5. number of categories
6. dependent
7. binomial
8. comparing two or more independent samples
9. never
10. normal populations, equal standard deviations

PART 2

1. H_0: The proportions are as stated.
H_1: The proportions are not as stated.
Using a significance level of .05, reject H_0 if $\chi^2 > 7.815$.

$$x^2 = \frac{(120-130)^2}{130} + \frac{(40-40)^2}{40} + \frac{(30-20)^2}{20} + \frac{(10-10)^2}{10} = 5.769$$

Do not reject H_0. Proportions could be as declared.

2. H_0: No relationship between gender and book type.
H_1: There is a relationship between gender and book type.
Using a significance level of .01, reject H_0 if $\chi^2 > 9.21$.

$$x^2 = \frac{(250-197.3)^2}{197.3} + \cdots + \frac{(200-187.5)^2}{187.5} = 54.84$$

Reject H_0. There is a relationship between gender and book type.

3. H_0: The distributions are the same.
H_1: The distributions are not the same.
H_0 is rejected if $H > 5.99$.

8:00 a.m.	Ranks	10:00 a.m.	Ranks	1:30 p.m.	Ranks
68	6	59	1.5	67	5
84	20	59	1.5	69	7
75	10.5	63	4	75	10.5
78	15.5	62	3	76	12.5
70	8	78	15.5	79	17
77	14	76	12.5	83	19
88	24	80	18	86	21.5
71	9			86	21.5
				87	23
Sums	107		56		137
Count	8		7		9

$$H = \frac{12}{24(25)}\left[\frac{107^2}{8} + \frac{56^2}{7} + \frac{137^2}{9}\right] - 3(25) = 4.29$$

H_0 is not rejected. There is no difference in the three distributions.

4. $H_0: \pi \le 1/3$ $H_1: \pi > 1/3$
At the .01 significance level, the decision rule is to reject H_0 if $z > 2.326$.

$$z = \frac{\left[\frac{210}{500} - \frac{1}{3}\right]}{\sqrt{\frac{(\frac{1}{3})(1-\frac{1}{3})}{500}}} = \frac{0.08667}{0.02108} = 4.11$$

Reject the null hypothesis.
The actual proportion of Louisiana children who were obese or overweight is more than one out of three.

PRACTICE TEST (AFTER CHAPTER 18) PART 1

1. denominator
2. index
3. quantity
4. base period
5. 1982–1984
6. trend
7. moving average
8. autocorrelation
9. residual
10. same

PART 2

1. a. 111.54, found by (145,000/130,000) × 100 for 2013 92.31, found by
   (120,000/130,000) × 100 for 2014 130.77, found by
   (170,000/130,000) × 100 for 2015 146.15, found by
   (190,000/130,000) × 100 for 2016
   b. 87.27, found by (120,000/137,500) × 100 for 2014 126.64, found by
   (170,000/137,500) × 100 for 2015 138.18, found by
   (190,000/137,500) × 100 for 2016
2. a. 108.91, found by (1,100/1,010) × 100
   b. 111.18, found by (4,525/4,070) × 100
   c. 110.20, found by (5,400/4,900) × 100
   d. 110.69, found by the square root of (111.18) × (110.20)
3. For January of the fifth year, the seasonally adjusted forecast is 70.0875, found by 1.05 × [5.50 + 1.25(49)].
   For February of the fifth year, the seasonally adjusted forecast is 66.844, found by 0.983 × [5.50 + 1.25(50)].