Appendix D: Self-Test Solutions and Answers to Even-Numbered Exercises

Chapter 1

2. a. 10
b. 5
c. Categorical variables: Size and Fuel
Quantitiative variables: Cylinders, City MPG, and Highway MPG
d.

Variable	Measurement Scale
Size	Ordinal
Cylinders	Ratio
City MPG	Ratio
Highway MPG	Ratio
Fuel	Nominal

3. a. Average for city driving = 182/10 = 18.2 mpg
b. Average for highway driving = 261/10 = 26.1 mpg
On average, the miles per gallon for highway driving is 7.9 mpg greater than for city driving
c. 3 of 10 or 30% have four-cylinder engines
d. 6 of 10 or 60% use regular fuel

4. a. 7
b. 5
c. Categorical variables: State, Campus Setting, and NCAA Division
d. Quantitiative variables: Endowment and Applicants Admitted

6. a. Quantitative
b. Categorical
c. Categorical
d. Quantitative
e. Categorical

8. a. 1015
b. Categorical
c. Percentages
d. .10(1015) = 101.5; 101 or 102 respondents

10. a. Quantitative; ratio
b. Categorical; nominal
c. Categorical; ordinal
d. Quantitative; ratio
e. Categorical; nominal

12. a. All visitors to Hawaii
b. Yes
c. First and fourth questions provide quantitative data
Second and third questions provide categorical data

13. a. Federal spending ($ trillions)
b. Quantitative
c. Time series
d. Federal spending has increased over time

14. a. Graph with a time series line for each manufacturer
b. Toyota surpasses General Motors in 2006 to become the leading auto manufacturer
c. A bar chart would show cross-sectional data for 2007; bar heights would be GM 8.8, Ford 7.9, DC 4.6, and Toyota 9.6

18. a. 36%
b. 189
c. Categorical

20. a. 43% of managers were bullish or very bullish, and 21% of managers expected health care to be the leading industry over the next 12 months
b. The average 12-month return estimate is 11.2% for the population of investment managers
c. The sample average of 2.5 years is an estimate of how long the population of investment managers think it will take to resume sustainable growth

22. a. The population consists of all customers of the chain stores in Charlotte, North Carolina
b. Some of the ways the grocery store chain could use to collect the data are
- Customers entering or leaving the store could be surveyed
- A survey could be mailed to customers who have a shopper's club card
- Customers could be given a printed survey when they check out
- Customers could be given a coupon that asks them to complete a brief online survey; if they do, they will receive a 5% discount on their next shopping trip

24. a. Correct
b. Incorrect
c. Correct
d. Incorrect
e. Incorrect

Chapter 2

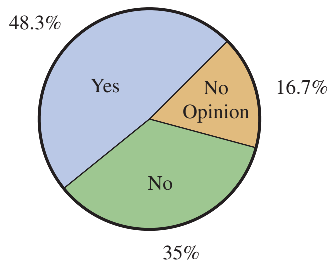
2. a. .20
b. 40
c/d.

Class	Frequency	Percent Frequency
A	44	22
B	36	18
C	80	40
D	40	20
Total	200	100

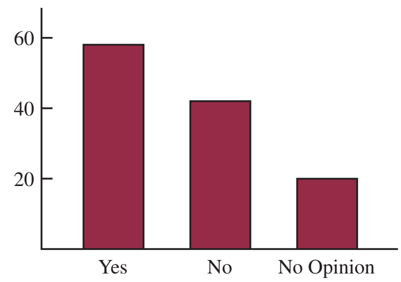
3. a. $360° \times 58/120 = 174°$

b. $360° \times 42/120 = 126°$

c.



d.



4. a. Categorical

b.

TV Show	Frequency	Percent Frequency
Law & Order	10	20%
CSL	18	36%
Without a Trace	9	18%
Desp Housewives	13	26%
Total:	50	100%

d. *CSI* had the largest viewing audience; *Desperate Housewives* was in second place

6. a.

Network	Frequency	Percent Frequency
ABC	15	30
CBS	17	34
FOX	1	2
NBC	17	34

b. CBS and NBC tied for first; ABC is close with 15

7.

Rating	Frequency	Relative Frequency
Outstanding	19	.38
Very good	13	.26
Good	10	.20
Average	6	.12
Poor	2	.04

Management should be pleased with these results: 64% of the ratings are very good to outstanding, and 84% of the ratings are good or better; comparing these ratings to previous results will show whether the restaurant is making improvements in its customers' ratings of food quality

8. a.

Position	Frequency	Relative Frequency
P	17	.309
H	4	.073
1	5	.091
2	4	.073
3	2	.036
S	5	.091
L	6	.109
C	5	.091
R	7	.127
Totals	55	1.000

b. Pitcher

c. 3rd base

d. Right field

e. Infielders 16 to outfielders 18

10. a/b.

Rating	Frequency	Percent Frequency
Excellent	20	2
Good	101	10
Fair	528	52
Bad	244	24
Terrible	122	12
Total	1015	100

c.

d. 36% a bad or a terrible job
12% a good or excellent job

e. 50% a bad or a terrible job
4% a good or excellent job
More pessimism in Spain

12.

Class	Cumulative Frequency	Cumulative Relative Frequency
≤19	10	.20
≤29	24	.48
≤39	41	.82
≤49	48	.96
≤59	50	1.00

14. b/c.

Class	Frequency	Percent Frequency
6.0–7.9	4	20
8.0–9.9	2	10
10.0–11.9	8	40
12.0–13.9	3	15
14.0–15.9	3	15
Totals	20	100

15. a/b.

Waiting Time	Frequency	Relative Frequency
0–4	4	.20
5–9	8	.40
10–14	5	.25
15–19	2	.10
20–24	1	.05
Totals	20	1.00

c/d.

Waiting Time	Cumulative Frequency	Cumulative Relative Frequency
≤4	4	.20
≤9	12	.60
≤14	17	.85
≤19	19	.95
≤24	20	1.00

e. 12/20 = .60

16. a.

Salary	Frequency
150–159	1
160–169	3
170–179	7
180–189	5
190–199	1
200–209	2
210–219	1
Total	20

b.

Salary	Percent Frequency
150–159	5
160–169	15
170–179	35
180–189	25
190–199	5
200–209	10
210–219	5
Total	100

c.

Salary	Cumulative Percent Frequency
Less than or equal to 159	5
Less than or equal to 169	20
Less than or equal to 179	55
Less than or equal to 189	80
Less than or equal to 199	85
Less than or equal to 209	95
Less than or equal to 219	100
Total	100

e. There is skewness to the right

f. 15%

18. a. Lowest $180; highest $2050

b.

Spending	Frequency	Percent Frequency
$0–249	3	12
250–499	6	24
500–749	5	20
750–999	5	20
1000–1249	3	12
1250–1499	1	4
1500–1749	0	0
1750–1999	1	4
2000–2249	1	4
Total	25	100

c. The distribution shows a positive skewness

d. Majority (64%) of consumers spend between $250 and $1000; the middle value is about $750; and two high spenders are above $1750

20. a.

Off-Course Income ($1000s)	Frequency	Percent Frequency
0–4,999	30	60
5,000–9,999	9	18
10,000–14,999	4	8
15,000–19,999	0	0
20,000–24,999	3	6
25,000–29,999	2	4
30,000–34,999	0	0
35,000–39,999	0	0
40,000–44,999	1	2
45,000–49,999	0	0
Over 50,000	1	2
Total	50	100

c. Off-course income is skewed to the right; only Tiger Woods earns over $50 million

d. The majority (60%) earn less that $5 million; 78% earn less than $10 million; five golfers (10%) earn between $20 million and $30 million; only Tiger Woods and Phil Mickelson earn more than $40 million

22.

5	7 8
6	4 5 8
7	0 2 2 5 5 6 8
8	0 2 3 5

23. Leaf unit = .1

6	3
7	5 5 7
8	1 3 4 8
9	3 6
10	0 4 5
11	3

24. Leaf unit = 10

11	6
12	0 2
13	0 6 7
14	2 2 7
15	5
16	0 2 8
17	0 2 3

25.

9	8 9
10	2 4 6 6
11	4 5 7 8 8 9
12	2 4 5 7
13	1 2
14	4
15	1

26. a.

1	0 3 7 7
2	4 5 5
3	0 0 5 5 9
4	0 0 0 5 5 8
5	0 0 0 4 5 5

b.

0	5 7
1	0 1 1 3 4
1	5 5 5 8
2	0 0 0 0 0 0
2	5 5
3	0 0 0
3	6
4	
4	
5	
5	
6	3

28. a.

2	14
2	67
3	011123
3	5677
4	003333344
4	6679
5	00022
5	5679
6	14
6	6
7	2

b. 40–44 with 9

c. 43 with 5

d. 10%; relatively small participation in the race

29. a.

		y: 1	2	Total
x	A	5	0	5
	B	11	2	13
	C	2	10	12
	Total	18	12	30

b.

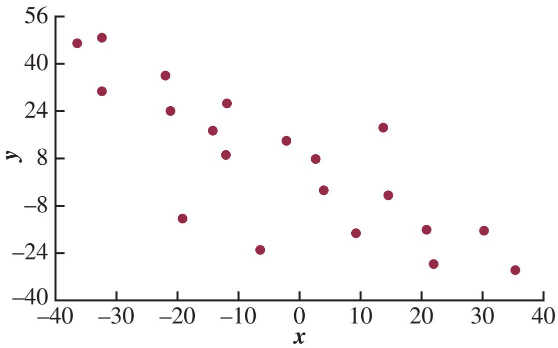
		y: 1	2	Total
x	A	100.0	0.0	100.0
	B	84.6	15.4	100.0
	C	16.7	83.3	100.0

c.

		y: 1	2
x	A	27.8	0.0
	B	61.1	16.7
	C	11.1	83.3
	Total	100.0	100.0

d. *A* values are always in $y = 1$
B values are most often in $y = 1$
C values are most often in $y = 2$

30. a.



b. A negative relationship between x and y; y decreases as x increases

32. a.

Education Level	Household Income ($1000s) Under 25	25.0–49.9	50.0–74.9	75.0–99.9	100 or more	Total
Not H.S. Graduate	32.10	18.71	9.13	5.26	2.20	13.51
H.S. Graduate	37.52	37.05	33.04	25.73	16.00	29.97
Some College	21.42	28.44	30.74	31.71	24.43	27.21
Bachelor's Degree	6.75	11.33	18.72	25.19	32.26	18.70
Beyond Bach. Deg.	2.21	4.48	8.37	12.11	25.11	10.61
Total	100.00	100.00	100.00	100.00	100.00	100.00

13.51% of the heads of households did not graduate from high school

b. 25.11%, 53.54%

c. Positive relationship between income and education level

34. a.

Fund Type	5-Year Average Return 0–9.99	10–19.99	20–29.99	30–39.99	40–49.99	50–59.99	Total
DE	1	25	1	0	0	0	27
FI	9	1	0	0	0	0	10
IE	0	2	3	2	0	1	8
Total	10	28	4	2	0	1	45

b.

5-Year Average Return	Frequency
0–9.99	10
10–19.99	28
20–29.99	4
30–39.99	2
40–49.99	0
50–59.99	1
Total	45

c.

Fund Type	Frequency
DE	27
FI	10
IE	8
Total	45

d. The margin of the crosstabulation shows these frequency distributions

e. Higher returns—International Equity funds
Lower returns—Fixed Income funds

36. b. Higher 5-year returns are associated with higher net asset values.

38. a.

Displace	Highway MPG 15–19	20–24	25–29	30–34	35–39	Total
1.0–2.9	0	6	72	46	4	128
3.0–4.9	3	56	86	0	0	145
5.0–6.9	23	14	1	0	0	38
Total	26	76	159	46	4	311

b. Higher fuel efficiencies are associated with smaller displacement engines
Lower fuel efficiencies are associated with larger displacement engines

d. Lower fuel efficiencies are associated with larger displacement engines

e. Scatter diagram

40. a.

Division	Frequency	Percent
Buick	10	5
Cadillac	10	5
Chevrolet	122	61
GMC	24	12
Hummer	2	1
Pontiac	18	9
Saab	2	1
Saturn	12	6
Total	200	100

b. Chevrolet, 61%

c. Hummer and Saab, both only 1%
Maintain Chevrolet and GMC

42. a.

SAT Score	Frequency
800–999	1
1000–1199	3
1200–1399	6
1400–1599	10
1600–1799	7
1800–1999	2
2000–2199	1
Total	30

b. Nearly symmetrical

c. 33% of the scores fall between 1400 and 1599
A score below 800 or above 2200 is unusual
The average is near or slightly above 1500

44. a.

Population	Frequency	Percent Frequency
0.0–2.4	17	34
2.5–4.9	12	24
5.0–7.4	9	18
7.5–9.9	4	8
10.0–12.4	3	6
12.5–14.9	1	2
15.0–17.4	1	2
17.5–19.9	1	2
20.0–22.4	0	0
22.5–24.9	1	2
25.0–27.4	0	0
27.5–29.9	0	0
30.0–32.4	0	0
32.5–34.9	0	0
35.0–37.4	1	2
Total	50	100

c. High positive skewness

d. 17 (34%) with population less than 2.5 million
29 (58%) with population less than 5 million
8 (16%) with population greater than 10 million
Largest 35.9 million (California)
Smallest .5 million (Wyoming)

46. a. High Temperatures

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

b. Low Temperatures

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

c. The most frequent range for high is in 60s (9 of 20) with only one low temperature above 54
High temperatures range mostly from 41 to 68, while low temperatures range mostly from 21 to 47
Low was 11; high was 84

d.

High Temp	Frequency	Low Temp	Frequency
10–19	0	10–19	1
20–29	0	20–29	5
30–39	1	30–39	5
40–49	4	40–49	5
50–59	3	50–59	3
60–69	9	60–69	1
70–79	2	70–79	0
80–89	1	80–89	0
Total	20	Total	20

48. a.

Level of Support	Percent Frequency
Strongly favor	30.10
Favor more than oppose	34.83
Oppose more than favor	21.13
Strongly oppose	13.94
Total	100.00

Overall favor higher tax $= 30.10\% + 34.83\%$
$= 64.93\%$

b. 20.2, 19.5, 20.6, 20.7, 19.0
Roughly 20% per country

c. The crosstabulation with column percentages:

	Country				
Support	Great Britain	Italy	Spain	Germany	United States
Strongly favor	31.00	31.96	45.99	19.98	20.98
Favor more than oppose	34.04	39.04	32.01	36.99	32.06
Oppose more than favor	23.00	17.99	13.98	24.03	26.96
Strongly oppose	11.96	11.01	8.03	18.99	20.00
Total	100.00	100.00	100.00	100.00	100.00

Considering the percentage of respondents who favor the higher tax by either saying "strongly favor" or "favor more than oppose," 65.04%, 71.00%, 78.00%, 56.97%, and 53.04% for the five countries; all show more than 50% support, but all European countries show more support for the tax than the United States; Italy and Spain show the highest level of support.

50. a. Row totals: 247; 54; 82; 121
Column totals: 149; 317; 17; 7; 14

b.

Year	Freq.	Fuel	Freq.
1973 or before	247	Elect.	149
1974–79	54	Nat. Gas	317
1980–86	82	Oil	17
1987–91	121	Propane	7
Total	504	Other	14
		Total	504

c. Crosstabulation of column percentages

Year Constructed	Fuel Type				
	Elect.	Nat. Gas	Oil	Propane	Other
1973 or before	26.9	57.7	70.5	71.4	50.0
1974–1979	16.1	8.2	11.8	28.6	0.0
1980–1986	24.8	12.0	5.9	0.0	42.9
1987–1991	32.2	22.1	11.8	0.0	7.1
Total	100.0	100.0	100.0	100.0	100.0

d. Crosstabulation of row percentages

Year Constructed	Fuel Type					
	Elect.	Nat. Gas	Oil	Propane	Other	Total
1973 or before	16.2	74.1	4.9	2.0	2.8	100.0
1974–1979	44.5	48.1	3.7	3.7	0.0	100.0
1980–1986	45.1	46.4	1.2	0.0	7.3	100.0
1987–1991	39.7	57.8	1.7	0.0	0.8	100.0

52. a. Crosstabulation of market value and profit

Market Value ($1000s)	Profit ($1000s) 0–300	300–600	600–900	900–1200	Total
0–8000	23	4			27
8000–16,000	4	4	2	2	12
16,000–24,000		2	1	1	4
24,000–32,000		1	2	1	4
32,000–40,000		2	1		3
Total	27	13	6	4	50

b. Crosstabulation of row percentages

Market Value ($1000s)	Profit ($1000s) 0–300	300–600	600–900	900–1200	Total
0–8000	85.19	14.81	0.00	0.00	100
8000–16,000	33.33	33.33	16.67	16.67	100
16,000–24,000	0.00	50.00	25.00	25.00	100
24,000–32,000	0.00	25.00	50.00	25.00	100
32,000–40,000	0.00	66.67	33.33	0.00	100

c. A positive relationship is indicated between profit and market value; as profit goes up, market value goes up

54. b. A positive relationship is demonstrated between market value and stockholders' equity

Chapter 3

2. 16, 16.5

3. Arrange data in order: 15, 20, 25, 25, 27, 28, 30, 34

$i = \frac{20}{100}(8) = 1.6$; round up to position 2

20th percentile = 20

$i = \frac{25}{100}(8) = 2$; use positions 2 and 3

25th percentile $= \frac{20 + 25}{2} = 22.5$

$i = \frac{65}{100}(8) = 5.2$; round up to position 6

65th percentile = 28

$i = \frac{75}{100}(8) = 6$; use positions 6 and 7

75th percentile $= \frac{28 + 30}{2} = 29$

4. 59.73, 57, 53

6. a. 18.42
b. 6.32
c. 34.3%
d. Reductions of only .65 shots and .9% made shots per game
Yes, agree but not dramatically

8. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{3200}{20} = 160$

Order the data from low 100 to high 360

Median: $i = \left(\frac{50}{100}\right)20 = 10$; use 10th and 11th positions

Median $= \left(\frac{130 + 140}{2}\right) = 135$

Mode = 120 (occur 3 times)

b. $i = \left(\frac{25}{100}\right)20 = 5$; use 5th and 6th positions

$Q_1 = \left(\frac{115 + 115}{2}\right) = 115$

$i = \left(\frac{75}{100}\right)20 = 15$; use 15th and 16th positions

$Q_3 = \left(\frac{180 + 195}{2}\right) = 187.5$

c. $i = \left(\frac{90}{100}\right)20 = 18$; use 18th and 19th positions

90th percentile $= \left(\frac{235 + 255}{2}\right) = 245$

90% of the tax returns cost $245 or less

10. a. .4%, 3.5%
b. 2.3%, 2.5%, 2.7%
c. 2.0%, 2.8%
d. optimistic

12. Disney: 3321, 255.5, 253, 169, 325
Pixar: 3231, 538.5, 505, 363, 631
Pixar films generate approximately twice as much box office revenue per film

14. 16, 4

15. Range = 34 − 15 = 19
Arrange data in order: 15, 20, 25, 25, 27, 28, 30, 34

$i = \frac{25}{100}(8) = 2$; $Q_1 = \frac{20 + 25}{2} = 22.5$

$i = \frac{75}{100}(8) = 6; Q_3 = \frac{28 + 30}{2} = 29$

$\text{IQR} = Q_3 - Q_1 = 29 - 22.5 = 6.5$

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{204}{8} = 25.5$

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
27	1.5	2.25
25	−.5	.25
20	−5.5	30.25
15	−10.5	110.25
30	4.5	20.25
34	8.5	72.25
28	2.5	6.25
25	−.5	.25
		242.00

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1} = \frac{242}{8 - 1} = 34.57$$

$$s = \sqrt{34.57} = 5.88$$

16. a. Range = 190 − 168 = 22

b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{1068}{6} = 178$

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

$$= \frac{4^2 + (-10)^2 + 6^2 + 12^2 + (-8)^2 + (-4)^2}{6 - 1}$$

$$= \frac{376}{5} = 75.2$$

c. $s = \sqrt{75.2} = 8.67$

d. $\frac{s}{\bar{x}}(100) = \frac{8.67}{178}(100\%) = 4.87\%$

18. a. 38, 97, 9.85
b. Eastern shows more variation

20. *Dawson:* range = 2, $s = .67$
Clark: range = 8, $s = 2.58$

22. a. 1285, 433
Freshmen spend more
b. 1720, 352
c. 404, 131.5
d. 367.04, 96.96
e. Freshmen have more variability

24. *Quarter-milers:* $s = .0564$, Coef. of Var. = 5.8%
Milers: $s = .1295$, Coef. of Var. = 2.9%

26. .20, 1.50, 0, −.50, −2.20

27. Chebyshev's theorem: *at least* $(1 - 1/z^2)$

a. $z = \frac{40 - 30}{5} = 2; 1 - \frac{1}{(2)^2} = .75$

b. $z = \frac{45 - 30}{5} = 3; 1 - \frac{1}{(3)^2} = .89$

c. $z = \frac{38 - 30}{5} = 1.6; 1 - \frac{1}{(1.6)^2} = .61$

d. $z = \frac{42 - 30}{5} = 2.4; 1 - \frac{1}{(2.4)^2} = .83$

e. $z = \frac{48 - 30}{5} = 3.6; 1 - \frac{1}{(3.6)^2} = .92$

28. a. 95%
b. Almost all
c. 68%

29. a. $z = 2$ standard deviations

$1 - \frac{1}{z^2} = 1 - \frac{1}{2^2} = \frac{3}{4}$; at least 75%

b. $z = 2.5$ standard deviations

$1 - \frac{1}{z^2} = 1 - \frac{1}{2.5^2} = .84$; at least 84%

c. $z = 2$ standard deviations
Empirical rule: 95%

30. a. 68%
b. 81.5%
c. 2.5%

32. a. −.67
b. 1.50
c. Neither an outlier
d. Yes; $z = 8.25$

34. a. 76.5, 7
b. 16%, 2.5%
c. 12.2, 7.89; no

36. 15, 22.5, 26, 29, 34

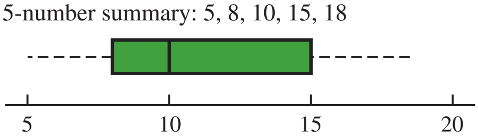
38. Arrange data in order: 5, 6, 8, 10, 10, 12, 15, 16, 18

$i = \frac{25}{100}(9) = 2.25$; round up to position 3

$Q_1 = 8$
Median (5th position) = 10

$i = \frac{75}{100}(9) = 6.75$; round up to position 7

$Q_3 = 15$
5-number summary: 5, 8, 10, 15, 18



40. a. Men's 1st place 43.73 minutes faster
b. Medians: 109.64, 131.67
Men's median time 22.03 minutes faster
c. 65.30, 87.18, 109.64, 128.40, 148.70
109.03, 122.08, 131.67, 147.18, 189.28
d. Men's Limits: 25.35 to 190.23; no outliers
Women's Limits: 84.43 to 184.83; 2 outliers
e. Women runners show less variation

41. a. Arrange data in order low to high

$i = \frac{25}{100}(21) = 5.25$; round up to 6th position

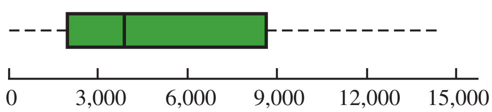
$Q_1 = 1872$
Median (11th position) = 4019
$i = \frac{75}{100}(21) = 15.75$; round up to 16th position
$Q_3 = 8305$
5-number summary: 608, 1872, 4019, 8305, 14,138

b. IQR $= Q_3 - Q_1 = 8305 - 1872 = 6433$
Lower limit: $1872 - 1.5(6433) = -7777.5$
Upper limit: $8305 + 1.5(6433) = 17,955$

c. No; data are within limits

d. 41,138 > 27,604; 41,138 would be an outlier; data value would be reviewed and corrected

e.



42. a. 73.5
b. 68, 71.5, 73.5, 74.5, 77
c. Limits: 67 and 79; no outliers
d. 66, 68, 71, 73, 75; 60.5 and 80.5
63, 65, 66, 67.6, 69; 61.25 and 71.25
75, 77, 78.5, 79.5, 81; 73.25 and 83.25
No outliers for any of the services
e. Verizon is highest rated
Sprint is lowest rated

44. a. 18.2, 15.35
b. 11.7, 23.5
c. 3.4, 11.7, 15.35, 23.5, 41.3
d. Yes; Alger Small Cap 41.3

45. b. There appears to be a negative linear relationship between x and y

c.

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
4	50	−4	4	−16
6	50	−2	4	−8
11	40	3	−6	−18
3	60	−5	14	−70
16	30	8	−16	−128
40	230	0	0	−240

$\bar{x} = 8$; $\bar{y} = 46$

$$s_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n - 1} = \frac{-240}{4} = -60$$

The sample covariance indicates a negative linear association between x and y

d. $r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{-60}{(5.43)(11.40)} = -.969$

The sample correlation coefficient of −.969 is indicative of a strong negative linear relationship

46. b. There appears to be a positive linear relationship between x and y
c. $s_{xy} = 26.5$
d. $r_{xy} = .693$

48. −.91; negative relationship

50. b. .910
c. Strong positive linear relationship; no

52. a. 3.69
b. 3.175

53. a.

f_i	M_i	$f_i M_i$
4	5	20
7	10	70
9	15	135
5	20	100
25		325

$$\bar{x} = \frac{\Sigma f_i M_i}{n} = \frac{325}{25} = 13$$

b.

f_i	M_i	$(M_i - \bar{x})$	$(M_i - \bar{x})^2$	$f_i(M_i - \bar{x})^2$
4	5	−8	64	256
7	10	−3	9	63
9	15	2	4	36
5	20	7	49	245
25				600

$$s^2 = \frac{\Sigma f_i (M_i - \bar{x})^2}{n - 1} = \frac{600}{25 - 1} = 25$$

$$s = \sqrt{25} = 5$$

54. a.

Grade x_i	Weight w_i
4 (A)	9
3 (B)	15
2 (C)	33
1 (D)	3
0 (F)	0
	60 credit hours

$$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{9(4) + 15(3) + 33(2) + 3(1)}{9 + 15 + 33 + 3}$$

$$= \frac{150}{60} = 2.5$$

b. Yes

56. 3.8, 3.7

58. a. 1800, 1351
b. 387, 1710
c. 7280, 1323
d. 3,675,303, 1917
e. High positive skewness
f. Using a box plot: 4135 and 7450 are outliers

60. a. 2.3, 1.85
b. 1.90, 1.38
c. Altria Group 5%
d. −.51, below mean
e. 1.02, above mean
f. No

62. a. \$670
b. \$456
c. $z = 3$; yes
d. Save time and prevent a penalty cost

64. a. 215.9
b. 55%
c. 175.0, 628.3
d. 48.8, 175.0, 215.9, 628.3, 2325.0
e. Yes, any price over 1308.25
f. 482.1; prefer median

66. a. 364 rooms
b. \$457
c. −.293; slight negative correlation
Higher cost per night tends to be associated with smaller hotels

68. a. .268, low or weak positive correlation
b. Very poor predictor; spring training is practice and does not count toward standings or playoffs

70. a. 60.68
b. $s^2 = 31.23$; $s = 5.59$

Chapter 4

2. $\binom{6}{3} = \frac{6!}{3!3!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(3 \cdot 2 \cdot 1)(3 \cdot 2 \cdot 1)} = 20$

ABC	ACE	BCD	BEF
ABD	ACF	BCE	CDE
ABE	ADE	BCF	CDF
ABF	ADF	BDE	CEF
ACD	AEF	BDF	DEF

4. b. (H,H,H), (H,H,T), (H,T,H), (H,T,T), (T,H,H), (T,H,T), (T,T,H), (T,T,T)
c. ⅛

6. $P(E_1) = .40$, $P(E_2) = .26$, $P(E_3) = .34$
The relative frequency method was used

8. a. 4: Commission Positive—Council Approves
Commission Positive—Council Disapproves
Commission Negative—Council Approves
Commission Negative—Council Disapproves

9. $\binom{50}{4} = \frac{50!}{4!46!} = \frac{50 \cdot 49 \cdot 48 \cdot 47}{4 \cdot 3 \cdot 2 \cdot 1} = 230{,}300$

10. a. Using the table, P(Debt) = .94
b. Five of the 8 institutions, P(over 60%) = 5/8 = .625
c. Two of the 8 institutions, P(more than \$30,000) = 2/8 = .25
d. P(No debt) = 1 − P(debt) = 1 − .72 = .28
e. A weighted average with 72% having average debt of \$32,980 and 28% having no debt

$$\text{Average debt per graduate} = \frac{.72(\$32{,}980) + .28(\$0)}{.72 + .28} = \$23{,}746$$

12. a. 3,478,761
b. 1/3,478,761
c. 1/146,107,962

14. a. ¼
b. ½
c. ¾

15. a. S = {ace of clubs, ace of diamonds, ace of hearts, ace of spades}
b. S = {2 of clubs, 3 of clubs, . . . , 10 of clubs, J of clubs, Q of clubs, K of clubs, A of clubs}
c. There are 12; jack, queen, or king in each of the four suits
d. For (a): 4/52 = 1/13 = .08
For (b): 13/52 = 1/4 = .25
For (c): 12/52 = .23

16. a. 36
c. ⅙
d. $\frac{5}{18}$
e. No; P(odd) = P(even) = $\frac{1}{2}$
f. Classical

17. a. (4, 6), (4, 7), (4, 8)
b. .05 + .10 + .15 = .30
c. (2, 8), (3, 8), (4, 8)
d. .05 + .05 + .15 = .25
e. .15

18. a. .0222
b. .8226
c. .1048

20. a. .108
b. .096
c. .434

22. a. .40, .40, .60
b. .80, yes
c. $A^c = \{E_3, E_4, E_5\}$; $C^c = \{E_1, E_4\}$;
$P(A^c) = .60$; $P(C^c) = .40$
d. (E_1, E_2, E_5); .60
e. .80

23. a. $P(A) = P(E_1) + P(E_4) + P(E_6)$
$= .05 + .25 + .10 = .40$
$P(B) = P(E_2) + P(E_4) + P(E_7)$
$= .20 + .25 + .05 = .50$
$P(C) = P(E_2) + P(E_3) + P(E_5) + P(E_7)$
$= .20 + .20 + .15 + .05 = .60$
b. $A \cup B = \{E_1, E_2, E_4, E_6, E_7\}$;
$P(A \cup B) = P(E_1) + P(E_2) + P(E_4) + P(E_6) + P(E_7)$
$= .05 + .20 + .25 + .10 + .05$
$= .65$
c. $A \cap B = \{E_4\}$; $P(A \cap B) = P(E_4) = .25$
d. Yes, they are mutually exclusive
e. $B^c = \{E_1, E_3, E_5, E_6\}$;
$P(B^c) = P(E_1) + P(E_3) + P(E_5) + P(E_6)$
$= .05 + .20 + .15 + .10$
$= .50$

24. a. .05
b. .70

26. a. .64
b. .48
c. .36
d. .76

28. Let B = rented a car for business reasons
P = rented a car for personal reasons
a. $P(B \cup P) = P(B) + P(P) - P(B \cap P)$
$= .540 + .458 - .300$
$= .698$
b. $P(\text{Neither}) = 1 - .698 = .302$

30. a. $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{.40}{.60} = .6667$
b. $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{.40}{.50} = .80$
c. No, because $P(A \mid B) \neq P(A)$

32. a.

	Car	Light Truck	Total
U.S.	.1330	.2939	.4269
Non-U.S.	.3478	.2253	.5731
Total	.4808	.5192	1.0000

b. .4269, .5731 Non-U.S. higher
.4808, .5192 Light Truck slightly higher
c. .3115, .6885 Light Truck higher
d. .6909, .3931 Car higher
e. .5661, U.S. higher for Light Trucks

33. a.

	Reason for Applying			
	Quality	Cost/ Convenience	Other	Total
Full-time	.218	.204	.039	.461
Part-time	.208	.307	.024	.539
Total	.426	.511	.063	1.000

b. A student is most likely to cite cost or convenience as the first reason (probability = .511); school quality is the reason cited by the second largest number of students (probability = .426)
c. $P(\text{quality} \mid \text{full-time}) = .218/.461 = .473$
d. $P(\text{quality} \mid \text{part-time}) = .208/.539 = .386$
e. For independence, we must have $P(A)P(B) = P(A \cap B)$; from the table
$P(A \cap B) = .218$, $P(A) = .461$, $P(B) = .426$
$P(A)P(B) = (.461)(.426) = .196$
Because $P(A)P(B) \neq P(A \cap B)$, the events are not independent

34. a.

	On Time	Late	Total
Southwest	.3336	.0664	.40
US Airways	.2629	.0871	.35
JetBlue	.1753	.0747	.25
Total	.7718	.2282	1.00

b. Southwest (.40)
c. .7718
d. US Airways (.3817); Southwest (.2910)

36. a. .7921
b. .9879
c. .0121
d. .3364, .8236, .1764
Don't foul Jerry Stackhouse

38. a. .70
b. .30
c. .67, .33
d. .20, .10
e. .40
f. .20
g. No; $P(S \mid M) \neq P(S)$

39. a. Yes, because $P(A_1 \cap A_2) = 0$
b. $P(A_1 \cap B) = P(A_1)P(B \mid A_1) = .40(.20) = .08$
$P(A_2 \cap B) = P(A_2)P(B \mid A_2) = .60(.05) = .03$
c. $P(B) = P(A_1 \cap B) + P(A_2 \cap B) = .08 + .03 = .11$
d. $P(A_1 \mid B) = \dfrac{.08}{.11} = .7273$
$P(A_2 \mid B) = \dfrac{.03}{.11} = .2727$

40. a. .10, .20, .09
b. .51
c. .26, .51, .23

42. M = missed payment
D_1 = customer defaults
D_2 = customer does not default
$P(D_1) = .05$, $P(D_2) = .95$, $P(M \mid D_2) = .2$, $P(M \mid D_1) = 1$
a. $P(D_1 \mid M) = \dfrac{P(D_1)P(M \mid D_1)}{P(D_1)P(M \mid D_1) + P(D_2)P(M \mid D_2)}$
$= \dfrac{(.05)(1)}{(.05)(1) + (.95)(.2)}$
$= \dfrac{.05}{.24} = .21$
b. Yes, the probability of default is greater than .20

44. a. .47, .53, .50, .45
b. .4963
c. .4463
d. 47%, 53%

46. a. .60
b. .26
c. .40
d. .74

48. a. 315
b. .29
c. No
d. Republicans

50. a. .76
b. .24

52. b. .2022
c. .4618
d. .4005

54. a. .49
b. .44
c. .54
d. No
e. Yes

56. a. .25
b. .125
c. .0125
d. .10
e. No

58. a.

	Young Adult	Older Adult	Total
Blogger	.0432	.0368	.08
Nonblogger	.2208	.6992	.92
Total	.2640	.7360	1.00

b. .2640
c. .0432
d. .1636

60. a. .40
b. .67

Chapter 5

1. a. Head, Head (H, H)
Head, Tail (H, T)
Tail, Head (T, H)
Tail, Tail (T, T)
b. x = number of heads on two coin tosses
c.

Outcome	Values of x
(H, H)	2
(H, T)	1
(T, H)	1
(T, T)	0

d. Discrete; 0, 1, and 2

2. a. x = time in minutes to assemble product
b. Any positive value: $x > 0$
c. Continuous

3. Let Y = position is offered
N = position is not offered
a. $S = \{(Y, Y, Y), (Y, Y, N), (Y, N, Y), (Y, N, N), (N, Y, Y), (N, Y, N), (N, N, Y), (N, N, N)\}$
b. Let N = number of offers made; N is a discrete random variable
c.

Experimental Outcome	(Y, Y, Y)	(Y, Y, N)	(Y, N, Y)	(Y, N, N)	(N, Y, Y)	(N, Y, N)	(N, N, Y)	(N, N, N)
Value of N	3	2	2	1	2	1	1	0

4. $x = 0, 1, 2, \ldots, 9$

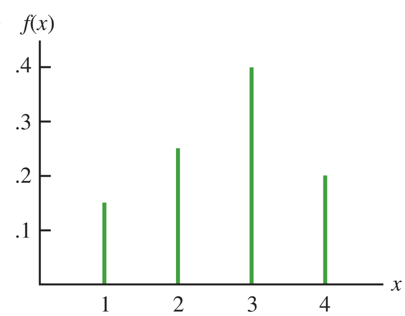
6. a. 0, 1, 2, . . . , 20; discrete
b. 0, 1, 2, . . . ; discrete
c. 0, 1, 2, . . . , 50; discrete
d. $0 \le x \le 8$; continuous
e. $x > 0$; continuous

7. a. $f(x) \ge 0$ for all values of x
$\Sigma f(x) = 1$; therefore, it is a valid probability distribution
b. Probability $x = 30$ is $f(30) = .25$
c. Probability $x \le 25$ is $f(20) + f(25) = .20 + .15 = .35$
d. Probability $x > 30$ is $f(35) = .40$

8. a.

x	$f(x)$
1	3/20 = .15
2	5/20 = .25
3	8/20 = .40
4	4/20 = .20
	Total 1.00

b.



c. $f(x) \ge 0$ for $x = 1, 2, 3, 4$
$\Sigma f(x) = 1$

10. a.

x	1	2	3	4	5
$f(x)$	.05	.09	.03	.42	.41

b.

x	1	2	3	4	5
$f(x)$	.04	.10	.12	.46	.28

c. .83
d. .28
e. Senior executives are more satisfied

12. a. Yes
b. .15
c. .10

14. a. .05
b. .70
c. .40

16. a.

y	$f(y)$	$yf(y)$
2	.20	.4
4	.30	1.2
7	.40	2.8
8	.10	.8
Totals	1.00	5.2

$E(y) = \mu = 5.2$

b.

y	$y - \mu$	$(y - \mu)^2$	$f(y)$	$(y - \mu)^2 f(y)$
2	−3.20	10.24	.20	2.048
4	−1.20	1.44	.30	.432
7	1.80	3.24	.40	1.296
8	2.80	7.84	.10	.784
			Total	4.560

$\text{Var}(y) = 4.56$

$\sigma = \sqrt{4.56} = 2.14$

18. a/b.

x	$f(x)$	$xf(x)$	$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 f(x)$
0	0.04	0.00	−1.84	3.39	0.12
1	0.34	0.34	−0.84	0.71	0.24
2	0.41	0.82	0.16	0.02	0.01
3	0.18	0.53	1.16	1.34	0.24
4	0.04	0.15	2.16	4.66	0.17
Total	1.00	1.84			0.79
		↑			↑
		$E(x)$			$\text{Var}(x)$

c/d.

y	$f(y)$	$yf(y)$	$y - \mu$	$(y - \mu)^2$	$y - \mu^2 f(y)$
0	0.00	0.00	−2.93	8.58	0.01
1	0.03	0.03	−1.93	3.72	0.12
2	0.23	0.45	−0.93	0.86	0.20
3	0.52	1.55	0.07	0.01	0.00
4	0.22	0.90	1.07	1.15	0.26
Total	1.00	2.93			0.59
		↑			↑
		$E(y)$			$\text{Var}(y)$

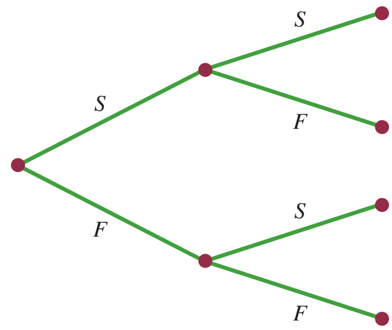
e. The number of bedrooms in owner-occupied houses is greater than in renter-occupied houses; the expected number of bedrooms is 2.93 − 1.84 = 1.09 greater, and the variability in the number of bedrooms is less for the owner-occupied houses

20. a. 430
b. −90; concern is to protect against the expense of a large loss

22. a. 445
b. $1250 loss

24. a. Medium: 145; large: 140
b. Medium: 2725; large: 12,400

25. a.



b. $f(1) = \binom{2}{1}(.4)^1(.6)^1 = \frac{2!}{1!1!}(.4)(.6) = .48$

c. $f(0) = \binom{2}{0}(.4)^0(.6)^2 = \frac{2!}{0!2!}(1)(.36) = .36$

d. $f(2) = \binom{2}{2}(.4)^2(.6)^0 = \frac{2!}{2!0!}(.16)(.1) = .16$

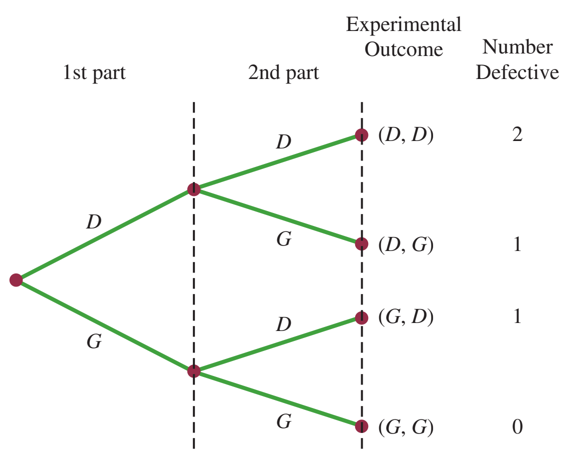
e. $P(x \geq 1) = f(1) + f(2) = .48 + .16 = .64$

f. $E(x) = np = 2(.4) = .8$
$\text{Var}(x) = np(1 - p) = 2(.4)(.6) = .48$
$\sigma = \sqrt{.48} = .6928$

26. a. .3487
b. .1937
c. .9298
d. .6513
e. 1
f. .9, .95

28. a. .2789
b. .4181
c. .0733

30. a. Probability of a defective part being produced must be .03 for each part selected; parts must be selected independently
b. Let D = defective
G = not defective



c. Two outcomes result in exactly one defect

d. $P(\text{no defects}) = (.97)(.97) = .9409$
$P(1 \text{ defect}) = 2(.03)(.97) = .0582$
$P(2 \text{ defects}) = (.03)(.03) = .0009$

32. a. .90
b. .99
c. .999
d. Yes

34. a. .2262
b. .8355

36. a. .1897
b. .9757
c. $f(12) = .0008$; yes
d. 5

38. a. $f(x) = \dfrac{3^x e^{-3}}{x!}$
b. .2241
c. .1494
d. .8008

39. a. $f(x) = \dfrac{2^x e^{-2}}{x!}$
b. $\mu = 6$ for 3 time periods
c. $f(x) = \dfrac{6^x e^{-6}}{x!}$
d. $f(2) = \dfrac{2^2 e^{-2}}{2!} = \dfrac{4(.1353)}{2} = .2706$
e. $f(6) = \dfrac{6^6 e^{-6}}{6!} = .1606$
f. $f(5) = \dfrac{4^5 e^{-4}}{5!} = .1563$

40. a. .1952
b. .1048
c. .0183
d. .0907

42. a. $f(0) = \dfrac{7^0 e^{-7}}{0!} = e^{-7} = .0009$
b. probability $= 1 - [f(0) + f(1)]$
$f(1) = \dfrac{7^1 e^{-7}}{1!} = 7e^{-7} = .0064$
probability $= 1 - [.0009 + .0064] = .9927$
c. $\mu = 3.5$
$f(0) = \dfrac{3.5^0 e^{-3.5}}{0!} = e^{-3.5} = .0302$
probability $= 1 - f(0) = 1 - .0302 = .9698$
d.

$$\begin{aligned}\text{probability} &= 1 - [f(0) + f(1) + f(2) + f(3) + f(4)]\\ &= 1 - [.0009 + .0064 + .0223 + .0521 + .0912]\\ &= .8271\end{aligned}$$

44. a. $\mu = 1.25$
b. .2865
c. .3581
d. .3554

46. a. $f(1) = \dfrac{\dbinom{3}{1}\dbinom{10-3}{4-1}}{\dbinom{10}{4}} = \dfrac{\left(\dfrac{3!}{1!2!}\right)\left(\dfrac{7!}{3!4!}\right)}{\dfrac{10!}{4!6!}}$
$= \dfrac{(3)(35)}{210} = .50$
b. $f(2) = \dfrac{\dbinom{3}{2}\dbinom{10-3}{2-2}}{\dbinom{10}{2}} = \dfrac{(3)(1)}{45} = .067$
c. $f(0) = \dfrac{\dbinom{3}{0}\dbinom{10-3}{2-0}}{\dbinom{10}{2}} = \dfrac{(1)(21)}{45} = .4667$
d. $f(2) = \dfrac{\dbinom{3}{2}\dbinom{10-3}{4-2}}{\dbinom{10}{4}} = \dfrac{(3)(21)}{210} = .30$
e. $x = 4$ is *greater than* $r = 3$; thus, $f(4) = 0$

48. a. .5250
b. .8167

50. $N = 60$, $n = 10$
a. $r = 20$, $x = 0$

$$\begin{aligned}f(0) &= \frac{\dbinom{20}{0}\dbinom{40}{10}}{\dbinom{60}{10}} = \frac{(1)\left(\dfrac{40!}{10!30!}\right)}{\dfrac{60!}{10!50!}}\\ &= \left(\frac{40!}{10!30!}\right)\left(\frac{10!50!}{60!}\right)\\ &= \frac{40 \cdot 39 \cdot 38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 \cdot 33 \cdot 32 \cdot 31}{60 \cdot 59 \cdot 58 \cdot 57 \cdot 56 \cdot 55 \cdot 54 \cdot 53 \cdot 52 \cdot 51}\\ &= .0112\end{aligned}$$

b. $r = 20$, $x = 1$

$$\begin{aligned}f(1) &= \frac{\dbinom{20}{1}\dbinom{40}{9}}{\dbinom{60}{10}} = 20\left(\frac{40!}{9!31!}\right)\left(\frac{10!50!}{60!}\right)\\ &= .0725\end{aligned}$$

c. $1 - f(0) - f(1) = 1 - .0112 - .0725 = .9163$
d. Same as the probability one will be from Hawaii; .0725

52. a. .2917
b. .0083
c. .5250, .1750; 1 bank
d. .7083
e. .90, .49, .70

54. a.

x	1	2	3	4	5
$f(x)$	.24	.21	.10	.21	.24

b. 3.00, 2.34
c. Bonds: $E(x) = 1.36$, $\text{Var}(x) = .23$
Stocks: $E(x) = 4$, $\text{Var}(x) = 1$

56. a. .0596
b. .3585

c. 100
d. 95, 9.75

58. a. .9510
b. .0480
c. .0490

60. a. 240
b. 12.96
c. 12.96

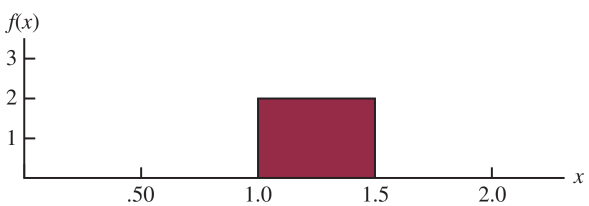
62. .1912

64. a. .2240
b. .5767

66. a. .4667
b. .4667
c. .0667

Chapter 6

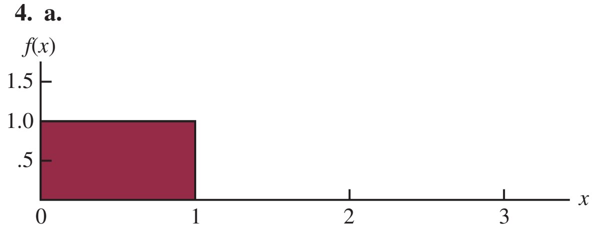
1. a.



b. $P(x = 1.25) = 0$; the probability of any single point is zero because the area under the curve above any single point is zero
c. $P(1.0 \le x \le 1.25) = 2(.25) = .50$
d. $P(1.20 < x < 1.5) = 2(.30) = .60$

2. b. .50
c. .60
d. 15
e. 8.33

4. a.



b. $P(.25 < x < .75) = 1(.50) = .50$
c. $P(x \le .30) = 1(.30) = .30$
d. $P(x > .60) = 1(.40) = .40$

6. a. .125
b. .50
c. .25

10. a. .9332
b. .8413
c. .0919
d. .4938

12. a. .2967
b. .4418
c. .3300
d. .5910
e. .8849
f. .2389

13. a. $P(-1.98 \le z \le .49) = P(z \le .49) - P(z < -1.98) = .6879 - .0239 = .6640$
b. $P(.52 \le z \le 1.22) = P(z \le 1.22) - P(z < .52) = .8888 - .6985 = .1903$
c. $P(-1.75 \le z \le -1.04) = P(z \le -1.04) - P(z < -1.75) = .1492 - .0401 = .1091$

14. a. $z = 1.96$
b. $z = 1.96$
c. $z = .61$
d. $z = 1.12$
e. $z = .44$
f. $z = .44$

15. a. The z value corresponding to a cumulative probability of .2119 is $z = -.80$
b. Compute .9030/2 = .4515; the cumulative probability of .5000 + .4515 = .9515 corresponds to $z = 1.66$
c. Compute .2052/2 = .1026; z corresponds to a cumulative probability of .5000 + .1026 = .6026, so $z = .26$
d. The z value corresponding to a cumulative probability of .9948 is $z = 2.56$
e. The area to the left of z is 1 − .6915 = .3085, so $z = -.50$

16. a. $z = 2.33$
b. $z = 1.96$
c. $z = 1.645$
d. $z = 1.28$

18. $\mu = 30$ and $\sigma = 8.2$
a. At $x = 40$, $z = \dfrac{40 - 30}{8.2} = 1.22$
$P(z \le 1.22) = .8888$
$P(x \ge 40) = 1.000 - .8888 = .1112$
b. At $x = 20$, $z = \dfrac{20 - 30}{8.2} = -1.22$
$P(z \le -1.22) = .1112$
$P(x \le 20) = .1112$
c. A z value of 1.28 cuts off an area of approximately 10% in the upper tail
$x = 30 + 8.2(1.28)$
$= 40.50$
A stock price of \$40.50 or higher will put a company in the top 10%

20. a. .0885
b. 12.51%
c. 93.8 hours or more

22. a. .7193
b. \$35.59
c. .0233

24. a. 200, 26.04
b. .2206

c. .1251
d. 242.84 million

26. a. $\mu = np = 100(.20) = 20$
$\sigma^2 = np(1 - p) = 100(.20)(.80) = 16$
$\sigma = \sqrt{16} = 4$
b. Yes, because $np = 20$ and $n(1 - p) = 80$
c. $P(23.5 \le x \le 24.5)$
$z = \dfrac{24.5 - 20}{4} = 1.13 \qquad P(z \le 1.13) = .8708$
$z = \dfrac{23.5 - 20}{4} = .88 \qquad P(z \le .88) = .8106$
$P(23.5 \le x \le 24.5) = P(.88 \le z \le 1.13)$
$= .8708 - .8106 = .0602$
d. $P(17.5 \le x \le 22.5)$
$z = \dfrac{22.5 - 20}{4} = .63 \qquad P(z \le .63) = .7357$
$z = \dfrac{17.5 - 20}{4} = -.63 \qquad P(z \le -.63) = .2643$
$P(17.5 \le x \le 22.5) = P(-.63 \le z \le .63)$
$= .7357 - .2643 = .4714$
e. $P(x \le 15.5)$
$z = \dfrac{15.5 - 20}{4} = -1.13 \qquad P(z \le -1.13) = .1292$
$P(x \le 15.5) = P(z \le -1.13) = .1292$

28. a. $\mu = np = 250(.20) = 50$
b. $\sigma^2 = np(1 - p) = 250(.20)(1 - 20) = 40$
$\sigma = \sqrt{40} = 6.3246$
$P(x < 40) = P(x \le 39.5)$
$z = \dfrac{x - \mu}{\sigma} = \dfrac{39.5 - 50}{6.3246} = -1.66 \quad \text{Area} = .0485$
$P(x \le 39.5) = .0485$
c. $P(55 \le x \le 60) = P(54.5 \le x \le 60.5)$
$z = \dfrac{x - \mu}{\sigma} = \dfrac{54.5 - 50}{6.3246} = .71 \quad \text{Area} = .7611$
$z = \dfrac{x - \mu}{\sigma} = \dfrac{60.5 - 50}{6.3246} = 1.66 \quad \text{Area} = .9515$
$P(54.5 \le x \le 60.5) = .9515 - .7611 = .1904$
d. $P(x \ge 70) = P(x \ge 69.5)$
$z = \dfrac{x - \mu}{\sigma} = \dfrac{69.5 - 50}{6.3246} = 3.08 \quad \text{Area} = .9990$
$P(x \ge 69.5) = 1 - .9990 = .0010$

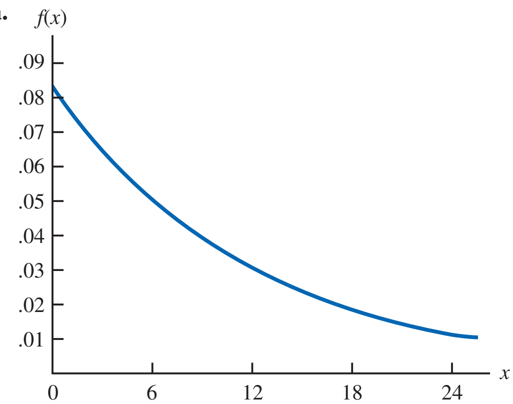
30. a. 220
b. .0392
c. .8962

32. a. .5276
b. .3935
c. .4724
d. .1341

33. a. $P(x \le x_0) = 1 - e^{-x_0/3}$
b. $P(x \le 2) = 1 - e^{-2/3} = 1 - .5134 = .4866$
c. $P(x \ge 3) = 1 - P(x \le 3) = 1 - (1 - e^{-3/3})$
$= e^{-1} = .3679$
d. $P(x \le 5) = 1 - e^{-5/3} = 1 - .1889 = .8111$
e. $P(2 \le x \le 5) = P(x \le 5) - P(x \le 2)$
$= .8111 - .4866 = .3245$

34. a. .5624
b. .1915
c. .2461
d. .2259

35. a.



b. $P(x \le 12) = 1 - e^{-12/12} = 1 - .3679 = .6321$
c. $P(x \le 6) \ = 1 - e^{-6/12} = 1 - .6065 = .3935$
d. $P(x \ge 30) = 1 - P(x < 30)$
$= 1 - (1 - e^{-30/12})$
$= .0821$

36. a. .3935
b. .2386
c. .1353

38. a. $f(x) = 5.5e^{-5.5x}$
b. .2528
c. .6002

40. a. $3780 or less
b. 19.22%
c. $8167.50

42. a. 3229
b. .2244
c. $12,383 or more

44. a. .0228
b. $50

46. a. 38.3%
b. 3.59% better, 96.41% worse
c. 38.21%

48. $\mu = 19.23$ ounces

50. a. Lose $240
b. .1788
c. .3557
d. .0594

52. a. $\frac{1}{7}$ minute
b. $7e^{-7x}$
c. .0009
d. .2466

54. **a.** 2 minutes
b. .2212
c. .3935
d. .0821

Chapter 7

1. **a.** AB, AC, AD, AE, BC, BD, BE, CD, CE, DE
b. With 10 samples, each has a $\frac{1}{10}$ probability
c. E and C because 8 and 0 do not apply; 5 identifies E; 7 does not apply; 5 is skipped because E is already in the sample; 3 identifies C; 2 is not needed because the sample of size 2 is complete

2. 22, 147, 229, 289

3. 459, 147, 385, 113, 340, 401, 215, 2, 33, 348

4. **a.** Bell South, LSI Logic, General Electric
b. 120

6. 2782, 493, 825, 1807, 289

8. ExxonMobil, Chevron, Travelers, Microsoft, Pfizer, and Intel

10. **a.** finite; **b.** infinite; **c.** infinite; **d.** finite; **e.** infinite

11. **a.** $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{54}{6} = 9$

b. $s = \sqrt{\dfrac{\Sigma(x_i - \bar{x})^2}{n-1}}$

$\Sigma(x_i - \bar{x})^2 = (-4)^2 + (-1)^2 + 1^2 + (-2)^2 + 1^2 + 5^2 = 48$

$s = \sqrt{\dfrac{48}{6-1}} = 3.1$

12. **a.** .50
b. .3667

13. **a.** $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{465}{5} = 93$

b.

	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
	94	+1	1
	100	+7	49
	85	−8	64
	94	+1	1
	92	−1	1
Totals	465	0	116

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{116}{4}} = 5.39$$

14. **a.** .45
b. .15
c. .45

16. **a.** .10
b. 20
c. .72

18. **a.** 200
b. 5
c. Normal with $E(\bar{x}) = 200$ and $\sigma_{\bar{x}} = 5$
d. The probability distribution of $\bar{x}$

19. **a.** The sampling distribution is normal with
$E(\bar{x}) = \mu = 200$
$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 50/\sqrt{100} = 5$
For ± 5, $195 \le \bar{x} \le 205$
Using the standard normal probability table:

At $\bar{x} = 205$, $z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \dfrac{5}{5} = 1$

$P(z \le 1) = .8413$

At $\bar{x} = 195$, $z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \dfrac{-5}{5} = -1$

$P(z < -1) = .1587$
$P(195 \le \bar{x} \le 205) = .8413 - .1587 = .6826$

b. For ± 10, $190 \le \bar{x} \le 210$
Using the standard normal probability table:

At $\bar{x} = 210$, $z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \dfrac{10}{5} = 2$

$P(z \le 2) = .9772$

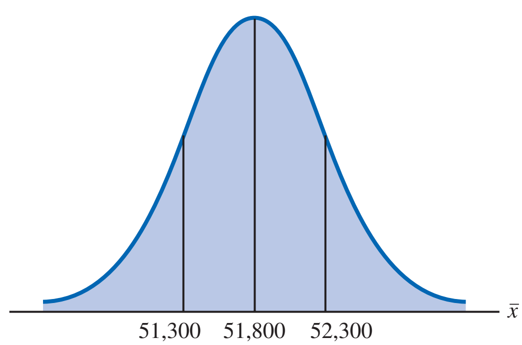
At $\bar{x} = 190$, $z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \dfrac{-10}{5} = -2$

$P(z < -2) = .0228$
$P(190 \le \bar{x} \le 210) = .9722 - .0228 = .9544$

20. 3.54, 2.50, 2.04, 1.77
$\sigma_{\bar{x}}$ decreases as n increases

22. **a.** Normal with $E(\bar{x}) = 51{,}800$ and $\sigma_{\bar{x}} = 516.40$
b. $\sigma_{\bar{x}}$ decreases to 365.15
c. $\sigma_{\bar{x}}$ decreases as n increases

23. **a.**



$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4000}{\sqrt{60}} = 516.40$$

At $\bar{x} = 52{,}300$, $z = \dfrac{52{,}300 - 51{,}800}{516.40} = .97$

$$P(\bar{x} \le 52{,}300) = P(z \le .97) = .8340$$

$$\text{At } \bar{x} = 51{,}300,\ z = \frac{51{,}300 - 51{,}800}{516.40} = -.97$$

$$P(\bar{x} < 51{,}300) = P(z < -.97) = .1660$$

$$P(51{,}300 \le \bar{x} \le 52{,}300) = .8340 - .1660 = .6680$$

b. $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{4000}{\sqrt{120}} = 365.15$

$$\text{At } \bar{x} = 52{,}300,\ z = \frac{52{,}300 - 51{,}800}{365.15} = 1.37$$

$$P(\bar{x} \le 52{,}300) = P(z \le 1.37) = .9147$$

$$\text{At } \bar{x} = 51{,}300,\ z = \frac{51{,}300 - 51{,}800}{365.15} = -1.37$$

$$P(\bar{x} < 51{,}300) = P(z < -1.37) = .0853$$

$$P(51{,}300 \le \bar{x} \le 52{,}300) = .9147 - .0853 = .8294$$

24. a. Normal with $E(\bar{x}) = 17.5$ and $\sigma_{\bar{x}} = .57$
b. .9198
c. .6212

26. a. .4246, .5284, .6922, .9586
b. Higher probability the sample mean will be close to population mean

28. a. Normal with $E(\bar{x}) = 95$ and $\sigma_{\bar{x}} = 2.56$
b. .7580
c. .8502
d. Part (c), larger sample size

30. a. $n/N = .01$; no
b. 1.29, 1.30; little difference
c. .8764

32. a. $E(\bar{p}) = .40$

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(.40)(.60)}{200}} = .0346$$

Within $\pm.03$ means $.37 \le \bar{p} \le .43$

$$z = \frac{\bar{p} - p}{\sigma_{\bar{p}}} = \frac{.03}{.0346} = .87$$

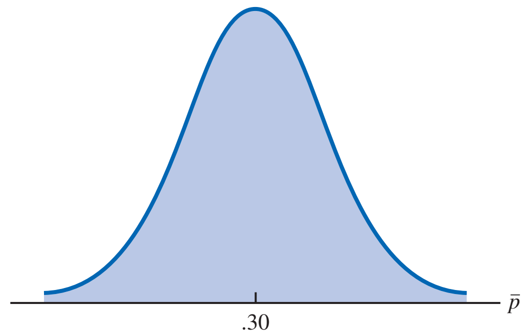
$$\begin{aligned} P(.37 \le \bar{p} \le .43) &= P(-.87 \le z \le .87) \\ &= .8078 - .1922 \\ &= .6156 \end{aligned}$$

b. $z = \dfrac{\bar{p} - p}{\sigma_{\bar{p}}} = \dfrac{.05}{.0346} = 1.44$

$$\begin{aligned} P(.35 \le \bar{p} \le .45) &= P(-1.44 \le z \le 1.44) \\ &= .9251 - .0749 \\ &= .8502 \end{aligned}$$

34. a. .6156
b. .7814
c. .9488
d. .9942
e. Higher probability with larger n

35. a.



$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.30(.70)}{100}} = .0458$$

The normal distribution is appropriate because $np = 100(.30) = 30$ and $n(1-p) = 100(.70) = 70$ are both greater than 5

b. $P(.20 \le \bar{p} \le .40) = ?$

$$z = \frac{.40 - .30}{.0458} = 2.18$$

$$\begin{aligned} P(.20 \le \bar{p} \le .40) &= P(-2.18 \le z \le 2.18) \\ &= .9854 - .0146 \\ &= .9708 \end{aligned}$$

c. $P(.25 \le \bar{p} \le .35) = ?$

$$z = \frac{.35 - .30}{.0458} = 1.09$$

$$\begin{aligned} P(.25 \le \bar{p} \le .35) &= P(-1.09 \le z \le 1.09) \\ &= .8621 - .1379 \\ &= .7242 \end{aligned}$$

36. a. Normal with $E(\bar{p}) = .66$ and $\sigma_{\bar{p}} = .0273$
b. .8584
c. .9606
d. Yes, standard error is smaller in part (c)
e. .9616, the probability is larger because the increased sample size reduces the standard error

38. a. Normal with $E(\bar{p}) = .56$ and $\sigma_{\bar{p}} = .0248$
b. .5820
c. .8926

40. a. Normal with $E(\bar{p}) = .76$ and $\sigma_{\bar{p}} = .0214$
b. .8384
c. .9452

42. 122, 99, 25, 55, 115, 102, 61

44. a. Normal with $E(\bar{x}) = 115.50$ and $\sigma_{\bar{x}} = 5.53$
b. .9298
c. $z = -2.80$, .0026

46. a. 955
b. .50
c. .7062
d. .8230

48. a. 625
b. .7888

50. a. Normal with $E(\bar{p}) = .28$ and $\sigma_{\bar{p}} = .0290$
b. .8324
c. .5098

52. a. .8882
b. .0233

54. a. 48
b. Normal, $E(\bar{p}) = .25$, $\sigma_{\bar{p}} = .0625$
c. .2119

Chapter 8

2. Use $\bar{x} \pm z_{\alpha/2}(\sigma/\sqrt{n})$
a. $32 \pm 1.645(6/\sqrt{50})$
32 ± 1.4; 30.6 to 33.4
b. $32 \pm 1.96(6/\sqrt{50})$
32 ± 1.66; 30.34 to 33.66
c. $32 \pm 2.576(6/\sqrt{50})$
32 ± 2.19; 29.81 to 34.19

4. 54

5. a. $1.96\sigma/\sqrt{n} = 1.96(5/\sqrt{49}) = 1.40$
b. 24.80 ± 1.40; 23.40 to 26.20

6. 8.1 to 8.9

8. a. Population is at least approximately normal
b. 3.1
c. 4.1

10. a. \$113,638 to \$124,672
b. \$112,581 to \$125,729
c. \$110,515 to \$127,795
d. Width increases as confidence level increases

12. a. 2.179
b. −1.676
c. 2.457
d. −1.708 and 1.708
e. −2.014 and 2.014

13. a. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{80}{8} = 10$
b. $s = \sqrt{\dfrac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\dfrac{84}{7}} = 3.464$
c. $t_{.025}\left(\dfrac{s}{\sqrt{n}}\right) = 2.365\left(\dfrac{3.46}{\sqrt{8}}\right) = 2.9$
d. $\bar{x} \pm t_{.025}\left(\dfrac{s}{\sqrt{n}}\right)$
10 ± 2.9 (7.1 to 12.9)

14. a. 21.5 to 23.5
b. 21.3 to 23.7
c. 20.9 to 24.1
d. A larger margin of error and a wider interval

15. $\bar{x} \pm t_{\alpha/2}(s/\sqrt{n})$
90% confidence: $df = 64$ and $t_{.05} = 1.669$
$19.5 \pm 1.669\left(\dfrac{5.2}{\sqrt{65}}\right)$
19.5 ± 1.08 or (18.42 to 20.58)
95% confidence: $df = 64$ and $t_{.025} = 1.998$
$19.5 \pm 1.998\left(\dfrac{5.2}{\sqrt{65}}\right)$
19.5 ± 1.29 or (18.21 to 20.79)

16. a. 1.69
b. 47.31 to 50.69
c. Fewer hours and higher cost for United

18. a. 22 weeks
b. 3.8020
c. 18.20 to 25.80
d. Larger n next time

20. $\bar{x} = 22$; 21.48 to 22.52

22. a. \$9,269 to \$12,541
b. 1523
c. 4,748,714, \$34 million

24. a. Planning value of $\sigma = \dfrac{\text{Range}}{4} = \dfrac{36}{4} = 9$
b. $n = \dfrac{z_{.025}^2\sigma^2}{E^2} = \dfrac{(1.96)^2(9)^2}{(3)^2} = 34.57$; use $n = 35$
c. $n = \dfrac{(1.96)^2(9)^2}{(2)^2} = 77.79$; use $n = 78$

25. a. Use $n = \dfrac{z_{\alpha/2}^2\sigma^2}{E^2}$
$n = \dfrac{(1.96)^2(6.84)^2}{(1.5)^2} = 79.88$; use $n = 80$
b. $n = \dfrac{(1.645)^2(6.84)^2}{(2)^2} = 31.65$; use $n = 32$

26. a. 18
b. 35
c. 97

28. a. 328
b. 465
c. 803
d. n gets larger; no to 99% confidence

30. 81

31. a. $\bar{p} = \dfrac{100}{400} = .25$
b. $\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}} = \sqrt{\dfrac{.25(.75)}{400}} = .0217$
c. $\bar{p} \pm z_{.025}\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$
$.25 \pm 1.96(.0217)$
$.25 \pm .0424$; .2076 to .2924

32. a. .6733 to .7267
b. .6682 to .7318

34. 1068

35. a. $\bar{p} = \dfrac{1760}{2000} = .88$
b. Margin of error
$z_{.05}\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}} = 1.645\sqrt{\dfrac{.88(1-.88)}{2000}} = .0120$

c. Confidence interval
$.88 \pm .0120$ or .868 to .892

d. Margin of error

$$z_{.05} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 1.96\sqrt{\frac{.88(1-.88)}{2000}} = .0142$$

95% confidence interval
$.88 \pm .0142$ or .8658 to .8942

36. a. .23
b. .1716 to .2884

38. a. .1790
b. .0738, .5682 to .7158
c. 354

39. a. $n = \dfrac{z_{.025}^2 p^*(1-p^*)}{E^2} = \dfrac{(1.96)^2(.156)(1-.156)}{(.03)^2}$
$= 562$

b. $n = \dfrac{z_{.005}^2 p^*(1-p^*)}{E^2} = \dfrac{(2.576)^2(.156)(1-.156)}{(.03)^2}$
$= 970.77$; use 971

40. .0346 (.4854 to .5546)

42. a. .0442
b. 601, 1068, 2401, 9604

44. a. 4.00
b. \$29.77 to \$37.77

46. a. 122
b. \$1751 to \$1995
c. \$172, 316 million
d. Less than \$1873

48. a. 14 minutes
b. 13.38 to 14.62
c. 32 per day
d. Staff reduction

50. 37

52. 176

54. a. .5420
b. .0508
c. .4912 to .5928

56. a. .8273
b. .7957 to .8589

58. a. 1267
b. 1509

60. a. .3101
b. .2898 to .3304
c. 8219; no, this sample size is unnecessarily large

Chapter 9

2. a. H_0: $\mu \le 14$
H_a: $\mu > 14$
b. No evidence that the new plan increases sales
c. The research hypothesis $\mu > 14$ is supported; the new plan increases sales

4. a. H_0: $\mu \ge 220$
H_a: $\mu < 220$

5. a. Rejecting H_0: $\mu \le 56.2$ when it is true
b. Accepting H_0: $\mu \le 56.2$ when it is false

6. a. H_0: $\mu \le 1$
H_a: $\mu > 1$
b. Claiming $\mu > 1$ when it is not true
c. Claiming $\mu \le 1$ when it is not true

8. a. H_0: $\mu \ge 220$
H_a: $\mu < 220$
b. Claiming $\mu < 220$ when it is not true
c. Claiming $\mu \ge 220$ when it is not true

10. a. $z = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{26.4-25}{6/\sqrt{40}} = 1.48$
b. Using normal table with $z = 1.48$: p-value $= 1.0000 - .9306 = .0694$
c. p-value $> .01$, do not reject H_0
d. Reject H_0 if $z \ge 2.33$
$1.48 < 2.33$, do not reject H_0

11. a. $z = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{14.15-15}{3/\sqrt{50}} = -2.00$
b. p-value $= 2(.0228) = .0456$
c. p-value $\le .05$, reject H_0
d. Reject H_0 if $z \le -1.96$ or $z \ge 1.96$
$-2.00 \le -1.96$, reject H_0

12. a. .1056; do not reject H_0
b. .0062; reject H_0
c. ≈ 0; reject H_0
d. .7967; do not reject H_0

14. a. .3844; do not reject H_0
b. .0074; reject H_0
c. .0836; do not reject H_0

15. a. H_0: $\mu \ge 1056$
H_a: $\mu < 1056$
b. $z = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{910-1056}{1600/\sqrt{400}} = -1.83$
p-value $= .0336$
c. p-value $\le .05$, reject H_0; the mean refund of "last-minute" filers is less than \$1056
d. Reject H_0 if $z \le -1.645$
$-1.83 \le -1.645$; reject H_0

16. a. H_0: $\mu \le 3173$
H_a: $\mu > 3173$
b. .0207
c. Reject H_0, conclude mean credit card balance for undergraduate student has increased

18. a. H_0: $\mu = 4.1$
H_a: $\mu \ne 4.1$
b. -2.21, .0272
c. Reject H_0; return for Mid-Cap Growth Funds differs from that for U.S. Diversified Funds

20. a. H_0: $\mu \ge 32.79$
H_a: $\mu < 32.79$

b. -2.73
c. .0032
d. Reject H_0; conclude the mean monthly Internet bill is less in the southern state

22. a. $H_0: \mu = 8$
$H_a: \mu \neq 8$
b. .1706
c. Do not reject H_0; we cannot conclude the mean waiting time differs from 8 minutes
d. 7.83 to 8.97; yes

24. a. $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{17 - 18}{4.5/\sqrt{48}} = -1.54$
b. Degrees of freedom $= n - 1 = 47$
Area in lower tail is between .05 and .10
p-value (two-tail) is between .10 and .20
Exact p-value = .1303
c. p-value $> .05$; do not reject H_0
d. With $df = 47$, $t_{.025} = 2.012$
Reject H_0 if $t \leq -2.012$ or $t \geq 2.012$
$t = -1.54$; do not reject H_0

26. a. Between .02 and .05; exact p-value = .0397; reject H_0
b. Between .01 and .02; exact p-value = .0125; reject H_0
c. Between .10 and .20; exact p-value = .1285; do not reject H_0

27. a. $H_0: \mu \geq 238$
$H_a: \mu < 238$
b. $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{231 - 238}{80/\sqrt{100}} = -.88$
Degrees of freedom $= n - 1 = 99$
p-value is between .10 and .20
Exact p-value = .1905
c. p-value $> .05$; do not reject H_0
Cannot conclude mean weekly benefit in Virginia is less than the national mean
d. $df = 99$, $t_{.05} = -1.66$
Reject H_0 if $t \leq -1.66$
$-.88 > -1.66$; do not reject H_0

28. a. $H_0: \mu \geq 9$
$H_a: \mu < 9$
b. Between .005 and .01
Exact p-value = .0072
c. Reject H_0; mean tenure of a CEO is less than 9 years

30. a. $H_0: \mu = 600$
$H_a: \mu \neq 600$
b. Between .20 and .40
Exact p-value = .2491
c. Do not reject H_0; cannot conclude there has been a change in mean CNN viewing audience
d. A larger sample size

32. a. $H_0: \mu = 10{,}192$
$H_a: \mu \neq 10{,}192$
b. Between .02 and .05
Exact p-value = .0304
c. Reject H_0; mean price at dealership differs from national mean price

34. a. $H_0: \mu = 2$
$H_a: \mu \neq 2$
b. 2.2
c. .52
d. Between .20 and .40
Exact p-value = .2535
e. Do not reject H_0; no reason to change from 2 hours for cost estimating

36. a. $z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1 - p_0)}{n}}} = \dfrac{.68 - .75}{\sqrt{\dfrac{.75(1 - .75)}{300}}} = -2.80$
p-value = .0026
p-value $\leq .05$; reject H_0
b. $z = \dfrac{.72 - .75}{\sqrt{\dfrac{.75(1 - .75)}{300}}} = -1.20$
p-value = .1151
p-value $> .05$; do not reject H_0
c. $z = \dfrac{.70 - .75}{\sqrt{\dfrac{.75(1 - .75)}{300}}} = -2.00$
p-value = .0228
p-value $\leq .05$; reject H_0
d. $z = \dfrac{.77 - .75}{\sqrt{\dfrac{.75(1 - .75)}{300}}} = .80$
p-value = .7881
p-value $> .05$; do not reject H_0

38. a. $H_0: p = .64$
$H_a: p \neq .64$
b. $\bar{p} = 52/100 = .52$
$z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1 - p_0)}{n}}} = \dfrac{.52 - .64}{\sqrt{\dfrac{.64(1 - .64)}{100}}} = -2.50$
p-value = 2(.0062) = .0124
c. p-value $\leq .05$; reject H_0
Proportion differs from the reported .64
d. Yes, because $\bar{p} = .52$ indicates that fewer believe the supermarket brand is as good as the name brand

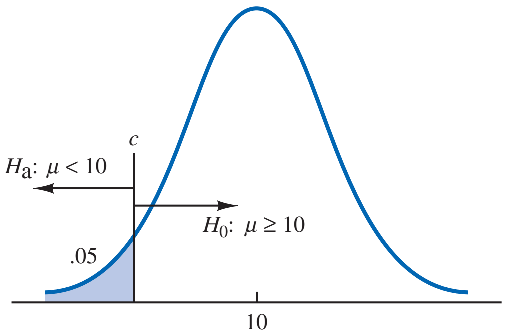
40. a. .2702
b. $H_0: p \leq .22$
$H_a: p > .22$
p-value ≈ 0; reject H_0; there is a significant increase after viewing commercials
c. Helps evaluate the effectiveness of commercials

42. a. $\bar{p} = .15$
b. .0718 to .2282
c. The return rate for the Houston store is different than the national average

44. a. $H_0: p \leq .51$
$H_a: p > .51$

b. $\bar{p} = .58$, p-value $= .0026$
c. Reject H_0; people working the night shift get drowsy more often

46.



$$c = 10 - 1.645(5/\sqrt{120}) = 9.25$$
Reject H_0 if $\bar{x} \leq 9.25$

a. When $\mu = 9$,
$$z = \frac{9.25 - 9}{5/\sqrt{120}} = .55$$
$$P(\text{Reject } H_0) = (1.0000 - .7088) = .2912$$
b. Type II error
c. When $\mu = 8$,
$$z = \frac{9.25 - 8}{5/\sqrt{120}} = 2.74$$
$$\beta = (1.0000 - .9969) = .0031$$

48. a. Concluding $\mu \leq 15$ when it is not true
b. .2676
c. .0179

49. a. H_0: $\mu \geq 25$
H_a: $\mu < 25$
Reject H_0 if $z \leq -2.05$
$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 25}{3/\sqrt{30}} = -2.05$$
Solve for $\bar{x} = 23.88$
Decision Rule: Accept H_0 if $\bar{x} > 23.88$
Reject H_0 if $\bar{x} \leq 23.88$
b. For $\mu = 23$,
$$z = \frac{23.88 - 23}{3/\sqrt{30}} = 1.61$$
$$\beta = 1.0000 - .9463 = .0537$$
c. For $\mu = 24$,
$$z = \frac{23.88 - 24}{3/\sqrt{30}} = -.22$$
$$\beta = 1.0000 - .4129 = .5871$$
d. The Type II error cannot be made in this case; note that when $\mu = 25.5$, H_0 is true; the Type II error can only be made when H_0 is false

50. a. Concluding $\mu = 28$ when it is not true
b. .0853, .6179, .6179, .0853
c. .9147

52. .1151, .0015
Increasing n reduces β

54. $n = \dfrac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu_0 - \mu_a)^2} = \dfrac{(1.645 + 1.28)^2(5)^2}{(10 - 9)^2} = 214$

56. 109

57. At $\mu_0 = 400$, $\alpha = .02$; $z_{.02} = 2.05$
At $\mu_a = 385$, $\beta = .10$; $z_{.10} = 1.28$
With $\sigma = 30$,
$$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu_0 - \mu_a)^2} = \frac{(2.05 + 1.28)^2(30)^2}{(400 - 385)^2} = 44.4 \text{ or } 45$$

58. 324

60. a. H_0: $\mu = 16$
H_a: $\mu \neq 16$
b. .0286; reject H_0
Readjust line
c. .2186; do not reject H_0
Continue operation
d. $z = 2.19$; reject H_0
$z = -1.23$; do not reject H_0
Yes, same conclusion

62. a. H_0: $\mu \leq 119{,}155$
H_a: $\mu > 119{,}155$
b. .0047
c. Reject H_0; mean annual income for theatergoers in Bay Area is higher

64. $t = -1.05$
p-value between .20 and .40
Exact p-value $= .2999$
Do not reject H_0; there is no evidence to conclude that the age at which women had their first child has changed

66. $t = 2.26$
p-value between .01 and .025
Exact p-value $= .0155$
Reject H_0; mean cost is greater than \$125,000

68. a. H_0: $p \leq .50$
H_a: $p > .50$
b. .64
c. .0026; reject H_0; college graduates have a greater stop smoking success rate

70. a. H_0: $p \leq .80$
H_a: $p > .80$
b. .84
c. .0418
d. Reject H_0; more than 80% of customers are satisfied with service of home agents

72. H_0: $p \geq .90$
H_a: $p < .90$
p-value $= .0808$
Do not reject H_0; claim of at least 90% cannot be rejected

74. a. H_0: $\mu \leq 72$
H_a: $\mu > 72$
b. .2912
c. .7939
d. 0, because H_0 is true

76. a. 45
b. .0192, .2358, .7291, .7291, .2358, .0192

Chapter 10

1. a. $\bar{x}_1 - \bar{x}_2 = 13.6 - 11.6 = 2$

b. $z_{\alpha/2} = z_{.05} = 1.645$

$$\bar{x}_1 - \bar{x}_2 \pm 1.645\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$2 \pm 1.645\sqrt{\frac{(2.2)^2}{50} + \frac{(3)^2}{35}}$$

$2 \pm .98$ (1.02 to 2.98)

c. $z_{\alpha/2} = z_{.05} = 1.96$

$$2 \pm 1.96\sqrt{\frac{(2.2)^2}{50} + \frac{(3)^2}{35}}$$

2 ± 1.17 (.83 to 3.17)

2. a. $$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(25.2 - 22.8) - 0}{\sqrt{\frac{(5.2)^2}{40} + \frac{(6)^2}{50}}} = 2.03$$

b. p-value $= 1.0000 - .9788 = .0212$

c. p-value $\le .05$; reject H_0

4. a. $\bar{x}_1 - \bar{x}_2 = 85.36 - 81.40 = 3.96$

b. $$z_{.025}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 1.96\sqrt{\frac{(4.55)^2}{37} + \frac{(3.97)^2}{44}} = 1.88$$

c. 3.96 ± 1.88 (2.08 to 5.84)

6. p-value $= .0351$
Reject H_0; mean price in Atlanta lower than mean price in Houston

8. a. Reject H_0; customer service has improved for Rite Aid

b. Do not reject H_0; the difference is not statistically significant

c. p-value $= .0336$; reject H_0; customer service has improved for Expedia

d. 1.80

e. The increase for J.C. Penney is not statistically significant

9. a. $\bar{x}_1 - \bar{x}_2 = 22.5 - 20.1 = 2.4$

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

$$= \frac{\left(\frac{2.5^2}{20} + \frac{4.8^2}{30}\right)^2}{\frac{1}{19}\left(\frac{2.5^2}{20}\right)^2 + \frac{1}{29}\left(\frac{4.8^2}{30}\right)^2} = 45.8$$

c. $df = 45$, $t_{.025} = 2.014$

$$t_{.025}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.014\sqrt{\frac{2.5^2}{20} + \frac{4.8^2}{30}} = 2.1$$

d. 2.4 ± 2.1 (.3 to 4.5)

10. a. $$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(13.6 - 10.1) - 0}{\sqrt{\frac{5.2^2}{35} + \frac{8.5^2}{40}}} = 2.18$$

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

$$= \frac{\left(\frac{5.2^2}{35} + \frac{8.5^2}{40}\right)^2}{\frac{1}{34}\left(\frac{5.2^2}{35}\right)^2 + \frac{1}{39}\left(\frac{8.5^2}{40}\right)^2} = 65.7$$

Use $df = 65$

c. $df = 65$, area in tail is between .01 and .025; two-tailed p-value is between .02 and .05
Exact p-value $= .0329$

d. p-value $\le .05$; reject H_0

12. a. $\bar{x}_1 - \bar{x}_2 = 22.5 - 18.6 = 3.9$ miles

b. $$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

$$= \frac{\left(\frac{8.4^2}{50} + \frac{7.4^2}{40}\right)^2}{\frac{1}{49}\left(\frac{8.4^2}{50}\right)^2 + \frac{1}{39}\left(\frac{7.4^2}{40}\right)^2} = 87.1$$

Use $df = 87$, $t_{.025} = 1.988$

$$3.9 \pm 1.988\sqrt{\frac{8.4^2}{50} + \frac{7.4^2}{40}}$$

3.9 ± 3.3 (.6 to 7.2)

14. a. $H_0\colon \mu_1 - \mu_2 \ge 0$
$H_a\colon \mu_1 - \mu_2 < 0$

b. -2.41

c. Using t table, p-value is between .005 and .01
Exact p-value $= .009$

d. Reject H_0; nursing salaries are lower in Tampa

16. a. $H_0\colon \mu_1 - \mu_2 \le 0$
$H_a\colon \mu_1 - \mu_2 > 0$

b. 38

c. $t = 1.80$, $df = 25$
Using t table, p-value is between .025 and .05
Exact p-value $= .0420$

d. Reject H_0; conclude higher mean score if college grad

18. a. $H_0\colon \mu_1 - \mu_2 \ge 120$
$H_a\colon \mu_1 - \mu_2 < 120$

b. -2.10
Using t table, p-value is between .01 and .025
Exact p-value $= .0195$

c. 32 to 118

d. Larger sample size

19. a. 1, 2, 0, 0, 2

b. $\bar{d} = \Sigma d_i/n = 5/5 = 1$

c. $$s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n - 1}} = \sqrt{\frac{4}{5 - 1}} = 1$$

d. $t = \dfrac{\bar{d} - \mu}{s_d/\sqrt{n}} = \dfrac{1 - 0}{1/\sqrt{5}} = 2.24$

$df = n - 1 = 4$

Using t table, p-value is between .025 and .05

Exact p-value = .0443

p-value $\leq$ .05; reject H_0

20. a. 3, −1, 3, 5, 3, 0, 1

b. 2

c. 2.08

d. 2

e. .07 to 3.93

21. $H_0: \mu_d \leq 0$

$H_a: \mu_d > 0$

$\bar{d} = .625$

$s_d = 1.30$

$t = \dfrac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \dfrac{.625 - 0}{1.30/\sqrt{8}} = 1.36$

$df = n - 1 = 7$

Using t table, p-value is between .10 and .20

Exact p-value = .1080

p-value $>$ .05; do not reject H_0; cannot conclude commercial improves mean potential to purchase

22. \$.10 to \$.32; earnings have increased

24. $t = 1.32$

Using t table, p-value is greater than .10

Exact p-value = .1142

Do not reject H_0; cannot conclude airfares from Dayton are higher

26. a. $t = -1.42$

Using t table, p-value is between .10 and .20

Exact p-value = .1718

Do not reject H_0; no difference in mean scores

b. −1.05

c. 1.28; yes

28. a. $\bar{p}_1 - \bar{p}_2 = .48 - .36 = .12$

b. $\bar{p}_1 - \bar{p}_2 \pm z_{.05}\sqrt{\dfrac{\bar{p}_1(1 - \bar{p}_1)}{n_1} + \dfrac{\bar{p}_2(1 - \bar{p}_2)}{n_2}}$

$.12 \pm 1.645\sqrt{\dfrac{.48(1 - .48)}{400} + \dfrac{.36(1 - .36)}{300}}$

$.12 \pm .0614$ (.0586 to .1814)

c. $.12 \pm 1.96\sqrt{\dfrac{.48(1 - .48)}{400} + \dfrac{.36(1 - .36)}{300}}$

$.12 \pm .0731$ (.0469 to .1931)

29. a. $\bar{p} = \dfrac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \dfrac{200(.22) + 300(.16)}{200 + 300} = .1840$

$z = \dfrac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}(1 - \bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

$= \dfrac{.22 - .16}{\sqrt{.1840(1 - .1840)\left(\dfrac{1}{200} + \dfrac{1}{300}\right)}} = 1.70$

p-value = 1.0000 − .9554 = .0446

b. p-value $\leq$ .05; reject H_0

30. $\bar{p}_1 = .55, \qquad \bar{p}_2 = .48$

.07 ± .0691

32. a. $H_0: p_w \leq p_m$

$H_a: p_w > p_m$

b. $\bar{p}_w = .3699$

c. $\bar{p}_m = .3400$

d. p-value = .1093

Do not reject H_0; cannot conclude women are more likely to ask directions

34. a. .64

b. .45

c. .19 ± .0813 (.1087 to .2713)

36. a. $H_0: p_1 - p_2 = 0$

$H_a: p_1 - p_2 \neq 0$

b. .13

c. p-value = .0404

d. Reject H_0; there is a significant difference between the younger and older age groups

38. a. $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 \neq 0$

$z = 2.79$

p-value = .0052

Reject H_0; a significant difference between systems exists

40. a. $H_0: \mu_1 - \mu_2 \leq 0$

$H_a: \mu_1 - \mu_2 > 0$

b. $t = .60$, $df = 57$

Using t table, p-value is greater than .20

Exact p-value = .2754

Do not reject H_0; cannot conclude that funds with loads have a higher mean rate of return

42. a. A decline of \$2.45

b. 2.45 ± 2.15 (.30 to 4.60)

c. 8% decrease

d. \$23.93

44. a. p-value ≈ 0, reject H_0

b. .0468 to .1332

46. a. .35 and .47

b. .12 ± .1037 (.0163 to .2237)

c. Yes, we would expect occupancy rates to be higher

Chapter 11

2. $s^2 = 25$

a. With 19 degrees of freedom, $\chi^2_{.05} = 30.144$ and $\chi^2_{.95} = 10.117$

$$\frac{19(25)}{30.144} \leq \sigma^2 \leq \frac{19(25)}{10.117}$$

$$15.76 \leq \sigma^2 \leq 46.95$$

b. With 19 degrees of freedom, $\chi^2_{.025} = 32.852$ and $\chi^2_{.975} = 8.907$

$$\frac{19(25)}{32.852} \le \sigma^2 \le \frac{19(25)}{8.907}$$
$$14.46 \le \sigma^2 \le 53.33$$

c. $3.8 \le \sigma \le 7.3$

4. a. .22 to .71
b. .47 to .84

6. a. .2205, 47.95, 6.92
b. 5.27 to 10.11

8. a. .4748
b. .6891
c. .2383 to 1.3687
.4882 to 1.1699

9. H_0: $\sigma^2 \le .0004$
H_a: $\sigma^2 > .0004$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(30-1)(.0005)}{.0004} = 36.25$$

From table with 29 degrees of freedom, p-value is greater than .10
p-value $> .05$; do not reject H_0
The product specification does not appear to be violated

10. H_0: $\sigma^2 \le 331.24$
H_a: $\sigma^2 > 331.24$
$\chi^2 = 52.07$, $df = 35$
p-value between .025 and .05
Reject H_0; standard deviation for Vanguard is greater

12. a. .8106
b. $\chi^2 = 9.49$
p-value greater than .20
Do not reject H_0; cannot conclude the variance for the other magazine is different

14. a. $F = 2.4$
p-value between .025 and .05
Reject H_0
b. $F_{.05} = 2.2$; reject H_0

15. a. Larger sample variance is s_1^2

$$F = \frac{s_1^2}{s_2^2} = \frac{8.2}{4} = 2.05$$

Degrees of freedom: 20, 25
From table, area in tail is between .025 and .05
p-value for two-tailed test is between .05 and .10
p-value $> .05$; do not reject H_0
b. For a two-tailed test:
$F_{\alpha/2} = F_{.025} = 2.30$
Reject H_0 if $F \ge 2.30$
$2.05 < 2.30$; do not reject H_0

16. $F = 1.59$
p-value less than .05
Reject H_0; the Fidelity Fund has greater variance

17. a. Population 1 is 4-year-old automobiles
H_0: $\sigma_1^2 \le \sigma_2^2$
H_a: $\sigma_1^2 > \sigma_2^2$

b. $F = \dfrac{s_1^2}{s_2^2} = \dfrac{170^2}{100^2} = 2.89$

Degrees of freedom: 25, 24
From tables, p-value is less than .01
p-value $\le .01$; reject H_0
Conclude that 4-year-old automobiles have a larger variance in annual repair costs compared to 2-year-old automobiles, which is expected because older automobiles are more likely to have more expensive repairs that lead to greater variance in the annual repair costs

18. $F = 1.44$
p-value greater than .20
Do not reject H_0; the difference between the variances is not statistically significant

20. $F = 5.29$
p-value ≈ 0
Reject H_0; population variances are not equal for seniors and managers

22. a. $F = 4$
p-value less than .01
Reject H_0; greater variability in stopping distance on wet pavement

24. 10.72 to 24.68

26. a. $\chi^2 = 27.44$
p-value between .01 and .025
Reject H_0; variance exceeds maximum requirements
b. .00012 to .00042

28. $\chi^2 = 31.50$
p-value between .05 and .10
Reject H_0; conclude that population variance is greater than 1

30. a. $n = 15$
b. 6.25 to 11.13

32. $F = 1.39$
Do not reject H_0; cannot conclude the variances of grade point averages are different

34. $F = 2.08$
p-value between .05 and .10
Reject H_0; conclude the population variances are not equal

Chapter 12

1. a. Expected frequencies: $e_1 = 200(.40) = 80$
$e_2 = 200(.40) = 80$
$e_3 = 200(.20) = 40$
Actual frequencies: $f_1 = 60, f_2 = 120, f_3 = 20$

$$\chi^2 = \frac{(60-80)^2}{80} + \frac{(120-80)^2}{80} + \frac{(20-40)^2}{40}$$
$$= \frac{400}{80} + \frac{1600}{80} + \frac{400}{40}$$
$$= 5 + 20 + 10 = 35$$

Degrees of freedom: $k - 1 = 2$
$\chi^2 = 35$ shows p-value is less than .005
p-value $\leq .01$; reject H_0; the proportions are not .40, .40, and .20

b. Reject H_0 if $\chi^2 \geq 9.210$
$\chi^2 = 35$; reject H_0

2. $\chi^2 = 15.33$, $df = 3$
p-value less than .005
Reject H_0; the proportions are not all .25

3. H_0: $p_{ABC} = .29$, $p_{CBS} = .28$, $p_{NBC} = .25$, $p_{IND} = .18$
H_a: The proportions are not
$p_{ABC} = .29$, $p_{CBS} = .28$, $p_{NBC} = .25$, $p_{IND} = .18$
Expected frequencies: $300(.29) = 87$, $300(.28) = 84$
$300(.25) = 75$, $300(.18) = 54$
$e_1 = 87$, $e_2 = 84$, $e_3 = 75$, $e_4 = 54$
Actual frequencies: $f_1 = 95$, $f_2 = 70$, $f_3 = 89$, $f_4 = 46$

$$\chi^2 = \frac{(95-87)^2}{87} + \frac{(70-84)^2}{84} + \frac{(89-75)^2}{75} + \frac{(46-54)^2}{54} = 6.87$$

Degrees of freedom: $k - 1 = 3$
$\chi^2 = 6.87$, p-value between .05 and .10
Do not reject H_0; cannot conclude that the audience proportions have changed

4. $\chi^2 = 29.51$, $df = 5$
p-value is less than .005
Reject H_0; the percentages differ from those reported by the company

6. a. $\chi^2 = 12.21$, $df = 3$
p-value is between .005 and .01
Conclude difference for 2003
b. 21%, 30%, 15%, 34%
Increased use of debit card
c. 51%

8. $\chi^2 = 16.31$, $df = 3$
p-value less than .005
Reject H_0; ratings differ, with telephone service slightly better

9. H_0: The column variable is independent of the row variable
H_a: The column variable is not independent of the row variable
Expected frequencies:

	A	B	C
P	28.5	39.9	45.6
Q	21.5	30.1	34.4

$$\chi^2 = \frac{(20-28.5)^2}{28.5} + \frac{(44-39.9)^2}{39.9} + \frac{(50-45.6)^2}{45.6} + \frac{(30-21.5)^2}{21.5} + \frac{(26-30.1)^2}{30.1} + \frac{(30-34.4)^2}{34.4}$$
$$= 7.86$$

Degrees of freedom: $(2-1)(3-1) = 2$
$\chi^2 = 7.86$, p-value between .01 and .025
Reject H_0; column variable and row variable are not independent

10. $\chi^2 = 19.77$, $df = 4$
p-value less than .005
Reject H_0; column variable and row variable are not independent

11. H_0: Type of ticket purchased is independent of the type of flight

H_a: Type of ticket purchased is not independent of the type of flight

Expected frequencies:

$e_{11} = 35.59$ $e_{12} = 15.41$
$e_{21} = 150.73$ $e_{22} = 65.27$
$e_{31} = 455.68$ $e_{32} = 197.32$

Ticket	Flight	Observed Frequency (f_i)	Expected Frequency (e_i)	$(f_i - e_i)^2/e_i$
First	Domestic	29	35.59	1.22
First	International	22	15.41	2.82
Business	Domestic	95	150.73	20.61
Business	International	121	65.27	47.59
Full-fare	Domestic	518	455.68	8.52
Full-fare	International	135	197.32	19.68
Totals		920		$\chi^2 = 100.43$

Degrees of freedom: $(3-1)(2-1) = 2$
$\chi^2 = 100.43$, p-value is less than .005
Reject H_0; type of ticket is not independent of type of flight

12. a. $\chi^2 = 7.95$, $df = 3$
p-value is between .025 and .05
Reject H_0; method of payment is not independent of age group
b. 18 to 24 use most

14. a. $\chi^2 = 8.47$; p-value is between .025 and .05
Reject H_0; intent to purchase again is not independent of the automobile
b. Accord 77, Camry 71, Taurus 62, Impala 57
c. Impala and Taurus below, Accord and Camry above; Accord and Camry have greater owner satisfaction, which may help future market share

16. a. 6446
b. $\chi^2 = 425.4$; p-value $= 0$
Reject H_0; attitude toward nuclear power is not independent of country
c. Italy (58%), Spain (32%)

18. $\chi^2 = 3.01$, $df = 2$
p-value is greater than .10
Do not reject H_0; couples working is independent of location; 63.3%

20. First estimate μ from the sample data (sample size = 120)

$$\mu = \frac{0(39) + 1(30) + 2(30) + 3(18) + 4(3)}{120} = \frac{156}{120} = 1.3$$

Therefore, we use Poisson probabilities with $\mu = 1.3$ to compute expected frequencies

x	Observed Frequency	Poisson Probability	Expected Frequency	Difference $(f_i - e_i)$
0	39	.2725	32.70	6.30
1	30	.3543	42.51	−12.51
2	30	.2303	27.63	2.37
3	18	.0998	11.98	6.02
4 or more	3	.0431	5.16	−2.17

$$\chi^2 = \frac{(6.30)^2}{32.70} + \frac{(-12.51)^2}{42.51} + \frac{(2.37)^2}{27.63} + \frac{(6.02)^2}{11.98} + \frac{(-2.17)^2}{5.16} = 9.04$$

Degrees of freedom: $5 - 1 - 1 = 3$
$\chi^2 = 9.04$, p-value is between .025 and .05
Reject H_0; not a Poisson distribution

21. With $n = 30$ we will use six classes with .1667 of the probability associated with each class

$$\bar{x} = 22.8, \; s = 6.27$$

The z values that create 6 intervals, each with probability .1667 are −.98, −.43, 0, .43, .98

z	Cutoff Value of x
−.98	$22.8 - .98(6.27) = 16.66$
−.43	$22.8 - .43(6.27) = 20.11$
0	$22.8 + .00(6.27) = 22.80$
.43	$22.8 + .43(6.27) = 25.49$
.98	$22.8 + .98(6.27) = 28.94$

Interval	Observed Frequency	Expected Frequency	Difference
less than 16.66	3	5	−2
16.66–20.11	7	5	2
20.11–22.80	5	5	0
22.80–25.49	7	5	2
25.49–28.94	3	5	−2
28.94 and up	5	5	0

$$\chi^2 = \frac{(-2)^2}{5} + \frac{(2)^2}{5} + \frac{(0)^2}{5} + \frac{(2)^2}{5} + \frac{(-2)^2}{5} + \frac{(0)^2}{5} = \frac{16}{5} = 3.20$$

Degrees of freedom: $6 - 2 - 1 = 3$
$\chi^2 = 3.20$, p-value greater than .10
Do not reject H_0
Assumption of a normal distribution is not rejected

22. $\chi^2 = 4.30$, $df = 2$
p-value greater than .10
Do not reject H_0; assumption of Poisson distribution is not rejected

24. $\chi^2 = 2.8$, $df = 3$
p-value greater than .10
Do not reject H_0; assumption of normal distribution is not rejected

26. $\chi^2 = 8.04$, $df = 3$
p-value between .025 and .05
Reject H_0; potentials are not the same for each sales territory

28. $\chi^2 = 4.64$, $df = 2$
p-value between .05 and .10
Do not reject H_0; cannot conclude market shares have changed

30. $\chi^2 = 42.53$, $df = 4$
p-value is less than .005
Reject H_0; conclude job satisfaction differs

32. $\chi^2 = 23.37$, $df = 3$
p-value is less than .005
Reject H_0; employment status is not independent of region

34. a. 71%, 22%, slower preferred
b. $\chi^2 = 2.99$, $df = 2$
p-value greater than .10
Do not reject H_0; cannot conclude men and women differ in preference

36. $\chi^2 = 6.17$, $df = 6$
p-value is greater than .10
Do not reject H_0; assumption that county and day of week are independent cannot be rejected

38. $\chi^2 = 7.75$, $df = 3$
p-value is between .05 and .10
Do not reject H_0; cannot conclude office vacancies differ by metropolitan area

Chapter 13

1. a. $\bar{\bar{x}} = (156 + 142 + 134)/3 = 144$

$$\text{SSTR} = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2 = 6(156 - 144)^2 + 6(142 - 144)^2 + 6(134 - 144)^2 = 1488$$

b. $\text{MSTR} = \dfrac{\text{SSTR}}{k - 1} = \dfrac{1488}{2} = 744$

c. $s_1^2 = 164.4, \quad s_2^2 = 131.2, \quad s_3^2 = 110.4$

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2$$

$= 5(164.4) + 5(131.2) + 5(110.4)$
$= 2030$

d. $\text{MSE} = \dfrac{\text{SSE}}{n_T - k} = \dfrac{2030}{18 - 3} = 135.3$

e.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	1488	2	744	5.50	.0162
Error	2030	15	135.3		
Total	3518	17			

f. $F = \dfrac{\text{MSTR}}{\text{MSE}} = \dfrac{744}{135.3} = 5.50$

From the F table (2 numerator degrees of freedom and 15 denominator), p-value is between .01 and .025

Using Excel or Minitab, the p-value corresponding to $F = 5.50$ is .0162

Because p-value $\le \alpha = .05$, we reject the hypothesis that the means for the three treatments are equal

2.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	300	4	75	14.07	.0000
Error	160	30	5.33		
Total	460	34			

4.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	150	2	75	4.80	.0233
Error	250	16	15.63		
Total	400	18			

Reject H_0 because p-value $\le \alpha = .05$

6. Because p-value $= .0082$ is less than $\alpha = .05$, we reject the null hypothesis that the means of the three treatments are equal

8. $\bar{\bar{x}} = (79 + 74 + 66)/3 = 73$

$$\text{SSTR} = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2 = 6(79 - 73)^2 + 6(74 - 73)^2 + 6(66 - 73)^2 = 516$$

$$\text{MSTR} = \frac{\text{SSTR}}{k - 1} = \frac{516}{2} = 258$$

$s_1^2 = 34 \quad s_2^2 = 20 \quad s_3^2 = 32$

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 5(34) + 5(20) + 5(32) = 430$$

$$\text{MSE} = \frac{\text{SSE}}{n_T - k} = \frac{430}{18 - 3} = 28.67$$

$$F = \frac{\text{MSTR}}{\text{MSE}} = \frac{258}{28.67} = 9.00$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	516	2	258	9.00	.003
Error	430	15	28.67		
Total	946	17			

Using F table (2 numerator degrees of freedom and 15 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $F = 9.00$ is .003

Because p-value $\le \alpha = .05$, we reject the null hypothesis that the means for the three plants are equal; in other words, analysis of variance supports the conclusion that the population mean examination scores at the three NCP plants are not equal

10. p-value $= .0000$
Because p-value $\le \alpha = .05$, we reject the null hypothesis that the means for the three groups are equal

12. p-value $= .0038$
Because p-value $\le \alpha = .05$, we reject the null hypothesis that the mean meal prices are the same for the three types of restaurants

13. a. $\bar{\bar{x}} = (30 + 45 + 36)/3 = 37$

$$\text{SSTR} = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2 = 5(30 - 37)^2 + 5(45 - 37)^2 + 5(36 - 37)^2 = 570$$

$$\text{MSTR} = \frac{\text{SSTR}}{k - 1} = \frac{570}{2} = 285$$

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 4(6) + 4(4) + 4(6.5) = 66$$

$$\text{MSE} = \frac{\text{SSE}}{n_T - k} = \frac{66}{15 - 3} = 5.5$$

$$F = \frac{\text{MSTR}}{\text{MSE}} = \frac{285}{5.5} = 51.82$$

Using F table (2 numerator degrees of freedom and 12 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $F = 51.82$ is .0000

Because p-value $\le \alpha = .05$, we reject the null hypothesis that the means of the three populations are equal

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\dfrac{1}{n_i} + \dfrac{1}{n_j}\right)}$

$$= t_{.025}\sqrt{5.5\left(\frac{1}{5} + \frac{1}{5}\right)}$$

$$= 2.179\sqrt{2.2} = 3.23$$

$|\bar{x}_1 - \bar{x}_2| = |30 - 45| = 15 > \text{LSD}$; significant difference
$|\bar{x}_1 - \bar{x}_3| = |30 - 36| = 6 > \text{LSD}$; significant difference
$|\bar{x}_2 - \bar{x}_3| = |45 - 36| = 9 > \text{LSD}$; significant difference

c. $\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

$(30 - 45) \pm 2.179\sqrt{5.5\left(\frac{1}{5} + \frac{1}{5}\right)}$

$-15 \pm 3.23 = -18.23$ to -11.77

14. a. Significant; p-value $= .0106$

b. LSD $= 15.34$

1 and 2; significant

1 and 3; not significant

2 and 3; significant

15. a.

	Manufacturer 1	Manufacturer 2	Manufacturer 3
Sample Mean	23	28	21
Sample Variance	6.67	4.67	3.33

$\bar{\bar{x}} = (23 + 28 + 21)/3 = 24$

$$\text{SSTR} = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2$$
$$= 4(23 - 24)^2 + 4(28 - 24)^2 + 4(21 - 24)^2$$
$$= 104$$

$$\text{MSTR} = \frac{\text{SSTR}}{k - 1} = \frac{104}{2} = 52$$

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2$$
$$= 3(6.67) + 3(4.67) + 3(3.33) = 44.01$$

$$\text{MSE} = \frac{\text{SSE}}{n_T - k} = \frac{44.01}{12 - 3} = 4.89$$

$$F = \frac{\text{MSTR}}{\text{MSE}} = \frac{52}{4.89} = 10.63$$

Using F table (2 numerator degrees of freedom and 9 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $F = 10.63$ is .0043

Because p-value $\le \alpha = .05$, we reject the null hypothesis that the mean time needed to mix a batch of material is the same for each manufacturer.

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_1} + \frac{1}{n_3}\right)}$

$$= t_{.025}\sqrt{4.89\left(\frac{1}{4} + \frac{1}{4}\right)}$$
$$= 2.262\sqrt{2.45} = 3.54$$

Since $|\bar{x}_1 - \bar{x}_3| = |23 - 21| = 2 < 3.54$, there does not appear to be any significant difference between the means for manufacturer 1 and manufacturer 3

16. $\bar{x}_1 - \bar{x}_2 \pm \text{LSD}$

$23 - 28 \pm 3.54$

$-5 \pm 3.54 = -8.54$ to -1.46

18. a. Significant; p-value $= .0000$

b. Significant; $2.3 > \text{LSD} = 1.19$

20. a. Significant; p-value $= .011$

b. Comparing North and South

$|7702 - 5566| = 2136 > \text{LSD} = 1620.76$

significant difference

Comparing North and West

$|7702 - 8430| = 728 > \text{LSD} = 1620.76$

no significant difference

Comparing South and West

$|5566 - 8430| = 2864 > \text{LSD} = 1775.45$

significant difference

21. ***Treatment Means***

$\bar{x}_{\cdot 1} = 13.6, \quad \bar{x}_{\cdot 2} = 11.0, \quad \bar{x}_{\cdot 3} = 10.6$

Block Means

$\bar{x}_{1\cdot} = 9,\ \bar{x}_{2\cdot} = 7.67,\ \bar{x}_{3\cdot} = 15.67,\ \bar{x}_{4\cdot} = 18.67,\ \bar{x}_{5\cdot} = 7.67$

Overall Mean

$\bar{\bar{x}} = 176/15 = 11.73$

Step 1

$$\text{SST} = \sum_i \sum_j (x_{ij} - \bar{\bar{x}})^2$$
$$= (10 - 11.73)^2 + (9 - 11.73)^2 + \cdots + (8 - 11.73)^2$$
$$= 354.93$$

Step 2

$$\text{SSTR} = b\sum_j (\bar{x}_{\cdot j} - \bar{\bar{x}})^2$$
$$= 5[(13.6 - 11.73)^2 + (11.0 - 11.73)^2 + (10.6 - 11.73)^2] = 26.53$$

Step 3

$$\text{SSBL} = k\sum_j (\bar{x}_{i\cdot} - \bar{\bar{x}})^2$$
$$= 3[(9 - 11.73)^2 + (7.67 - 11.73)^2 + (15.67 - 11.73)^2 + (18.67 - 11.73)^2 + (7.67 - 11.73)^2] = 312.32$$

Step 4

$$\text{SSE} = \text{SST} - \text{SSTR} - \text{SSBL}$$
$$= 354.93 - 26.53 - 312.32 = 16.08$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	26.53	2	13.27	6.60	.0203
Blocks	312.32	4	78.08		
Error	16.08	8	2.01		
Total	354.93	14			

From the F table (2 numerator degrees of freedom and 8 denominator), p-value is between .01 and .025

Actual p-value $= .0203$

Because p-value $\le \alpha = .05$, we reject the null hypothesis that the means of the three treatments are equal

22.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	310	4	77.5	17.69	.0005
Blocks	85	2	42.5		
Error	35	8	4.38		
Total	430	14			

Significant; p-value $\leq \alpha = .05$

24. p-value $= .0453$

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the mean tune-up times are the same for both analyzers

26. a. Significant; p-value $= .0231$

b. Writing section

28. *Step 1*

$$\text{SST} = \sum_i \sum_j \sum_k (x_{ijk} - \bar{\bar{x}})^2$$
$$= (135 - 111)^2 + (165 - 111)^2 + \cdots + (136 - 111)^2 = 9028$$

Step 2

$$\text{SSA} = br \sum_i (\bar{x}_{j\cdot} - \bar{\bar{x}})^2$$
$$= 3(2)[(104 - 111)^2 + (118 - 111)^2] = 588$$

Step 3

$$\text{SSB} = ar \sum_j (\bar{x}_{\cdot j} - \bar{\bar{x}})^2$$
$$= 2(2)[(130 - 111)^2 + (97 - 111)^2 + (106 - 111)^2] = 2328$$

Step 4

$$\text{SSAB} = r \sum_i \sum_j (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{\bar{x}})^2$$
$$= 2[(150 - 104 - 130 + 111)^2 + (78 - 104 - 97 + 111)^2 + \cdots + (128 - 118 - 106 + 111)^2] = 4392$$

Step 5

$$\text{SSE} = \text{SST} - \text{SSA} - \text{SSB} - \text{SSAB} = 9028 - 588 - 2328 - 4392 = 1720$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Factor A	588	1	588	2.05	.2022
Factor B	2328	2	1164	4.06	.0767
Interaction	4392	2	2196	7.66	.0223
Error	1720	6	286.67		
Total	9028	11			

Factor A: $F = 2.05$

Using F table (1 numerator degree of freedom and 6 denominator), p-value is greater than .10

Using Excel or Minitab, the p-value corresponding to $F = 2.05$ is .2022

Because p-value $> \alpha = .05$, Factor A is not significant

Factor B: $F = 4.06$

Using F table (2 numerator degrees of freedom and 6 denominator), p-value is between .05 and .10

Using Excel or Minitab, the p-value corresponding to $F = 4.06$ is .0767

Because p-value $> \alpha = .05$, Factor B is not significant

Interaction: $F = 7.66$

Using F table (2 numerator degrees of freedom and 6 denominator), p-value is between .01 and .025

Using Excel or Minitab, the p-value corresponding to $F = 7.66$ is .0223

Because p-value $\leq \alpha = .05$, interaction is significant

30. Design: p-value $= .0104$; significant
Size: p-value $= .1340$; not significant
Interaction: p-value $= .2519$; not significant

32. Class: p-value $= .0002$; significant
Type: p-value $= .0006$; significant
Interaction: p-value $= .4229$; not significant

34. Significant; p-value $= .0134$

36. Significant; p-value $= .046$

38. Not significant; p-value $= .2455$

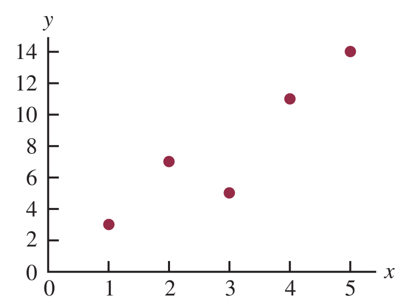
40. a. Significant; p-value $= .0175$

42. Significant; p-value $= .004$

44. Type of machine (p-value $= .0226$) is significant; type of loading system (p-value $= .7913$) and interaction (p-value $= .0671$) are not significant

Chapter 14

1. a.



b. There appears to be a positive linear relationship between x and y

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between x and y; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion

d. Summations needed to compute the slope and y-intercept:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{15}{5} = 3, \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{40}{5} = 8,$$
$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 26, \quad \Sigma(x_i - \bar{x})^2 = 10$$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{26}{10} = 2.6$$

$b_0 = \bar{y} - b_1\bar{x} = 8 - (2.6)(3) = 0.2$

$\hat{y} = 0.2 - 2.6x$

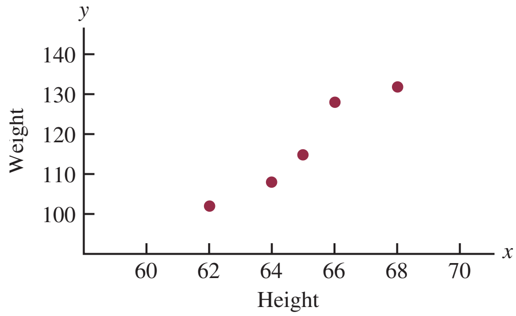
e. $\hat{y} = .2 + 2.6x = .2 + 2.6(4) = 10.6$

2. b. There appears to be a negative linear relationship between x and y

d. $\hat{y} = 68 - 3x$

e. 38

4. a.



b. There appears to be a positive linear relationship between x = height and y = weight

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between height and weight; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion

d. Summations needed to compute the slope and y-intercept:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{325}{5} = 65, \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{585}{5} = 117,$$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 110, \quad \Sigma(x_i - \bar{x})^2 = 20$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{110}{20} = 5.5$$

$b_0 = \bar{y} - b_1\bar{x} = 117 - (5.5)(65) = -240.5$

$\hat{y} = -240.5 + 5.5x$

e. $\hat{y} = -240.5 + 5.5(63) = 106$

The estimate of weight is 106 pounds

6. c. $\hat{y} = 8.9412 - .02633x$

e. 6.3 or approximately \$6300

8. c. $\hat{y} = 359.2668 - 5.2772x$

d. \$254

10. c. $\hat{y} = -6{,}745.44 + 149.29x$

d. 4003 or \$4,003,000

12. c. $\hat{y} = -8129.4439 + 22.4443x$

d. \$8704

14. c. $\hat{y} = 37.1217 + .51758x$

d. 73

15. a. $\hat{y}_i = .2 + 2.6x_i$ and $\bar{y} = 8$

x_i	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
1	3	2.8	.2	.04	−5	25
2	7	5.4	1.6	2.56	−1	1
3	5	8.0	−3.0	9.00	−3	9
4	11	10.6	.4	.16	3	9
5	14	13.2	.8	.64	6	36
				SSE = 12.40		SST = 80

SSR = SST − SSE = 80 − 12.4 = 67.6

b. $r^2 = \frac{\text{SSR}}{\text{SST}} = \frac{67.6}{80} = .845$

The least squares line provided a good fit; 84.5% of the variability in y has been explained by the least squares line

c. $r_{xy} = \sqrt{.845} = +.9192$

16. a. SSE = 230, SST = 1850, SSR = 1620

b. $r^2 = .876$

c. $r_{xy} = -.936$

18. a. The estimated regression equation and the mean for the dependent variable:

$\hat{y} = 1790.5 + 581.1x, \quad \bar{y} = 3650$

The sum of squares due to error and the total sum of squares:

$\text{SSE} = \Sigma(y_i - \hat{y}_i)^2 = 85{,}135.14$

$\text{SST} = \Sigma(y_i - \bar{y})^2 = 335{,}000$

Thus, SSR = SST − SSE

= 335,000 − 85,135.14 = 249,864.86

b. $r^2 = \frac{\text{SSR}}{\text{SST}} = \frac{249{,}864.86}{335{,}000} = .746$

The least squares line accounted for 74.6% of the total sum of squares

c. $r_{xy} = \sqrt{.746} = +.8637$

20. a. $\hat{y} = 12.0169 + .0127x$

b. $r^2 = .4503$

c. 53

22. a. .77

b. Yes

c. $r_{xy} = +.88$, strong

23. a. $s^2 = \text{MSE} = \frac{\text{SSE}}{n - 2} = \frac{12.4}{3} = 4.133$

b. $s = \sqrt{\text{MSE}} = \sqrt{4.133} = 2.033$

c. $\Sigma(x_i - \bar{x})^2 = 10$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{2.033}{\sqrt{10}} = .643$$

d. $t = \frac{b_1 - \beta_1}{s_{b_1}} = \frac{2.6 - 0}{.643} = 4.044$

From the t table (3 degrees of freedom), area in tail is between .01 and .025

p-value is between .02 and .05

Using Excel or Minitab, the p-value corresponding to $t = 4.04$ is .0272

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

e. $\text{MSR} = \dfrac{\text{SSR}}{1} = 67.6$

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{67.6}{4.133} = 16.36$$

From the F table (1 numerator degree of freedom and 3 denominator), p-value is between .025 and .05

Using Excel or Minitab, the p-value corresponding to $F = 16.36$ is .0272

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	67.6	1	67.6	16.36	.0272
Error	12.4	3	4.133		
Total	80	4			

24. **a.** 76.6667
b. 8.7560
c. .6526
d. Significant; p-value = .0193
e. Significant; p-value = .0193

26. **a.** $s^2 = \text{MSE} = \dfrac{\text{SSE}}{n-2} = \dfrac{85{,}135.14}{4} = 21{,}283.79$

$s = \sqrt{\text{MSE}} = \sqrt{21{,}283.79} = 145.89$

$\Sigma(x_i - \bar{x})^2 = .74$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{145.89}{\sqrt{.74}} = 169.59$$

$$t = \frac{b_1 - \beta_1}{s_{b_1}} = \frac{581.08 - 0}{169.59} = 3.43$$

From the t table (4 degrees of freedom), area in tail is between .01 and .025

p-value is between .02 and .05

Using Excel or Minitab, the p-value corresponding to $t = 3.43$ is .0266

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

b. $\text{MSR} = \dfrac{\text{SSR}}{1} = \dfrac{249{,}864.86}{1} = 249{,}864.86$

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{249{,}864.86}{21{,}283.79} = 11.74$$

From the F table (1 numerator degree of freedom and 4 denominator), p-value is between .025 and .05

Using Excel or Minitab, the p-value corresponding to $F = 11.74$ is .0266

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

c.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	249,864.86	1	249,864.86	11.74	.0266
Error	85,135.14	4	21,283.79		
Total	335,000	5			

28. They are related; p-value = .000

30. Significant; p-value = .0042

32. **a.** $s = 2.033$

$\bar{x} = 3$, $\Sigma(x_i - \bar{x})^2 = 10$

$$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}}$$

$$= 2.033\sqrt{\frac{1}{5} + \frac{(4-3)^2}{10}} = 1.11$$

b. $\hat{y} = .2 + 2.6x = .2 + 2.6(4) = 10.6$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

$10.6 \pm 3.182(1.11)$

10.6 ± 3.53, or 7.07 to 14.13

c. $s_{\text{ind}} = s\sqrt{1 + \dfrac{1}{n} + \dfrac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}}$

$$= 2.033\sqrt{1 + \frac{1}{5} + \frac{(4-3)^2}{10}} = 2.32$$

d. $\hat{y}_p \pm t_{\alpha/2} s_{\text{ind}}$

$10.6 \pm 3.182(2.32)$

10.6 ± 7.38, or 3.22 to 17.98

34. Confidence interval: 8.65 to 21.15
Prediction interval: −4.50 to 41.30

35. **a.** $s = 145.89$, $\bar{x} = 3.2$, $\Sigma(x_i - \bar{x})^2 = .74$

$\hat{y} = 1790.5 + 581.1x = 1790.5 + 581.1(3)$
$= 3533.8$

$$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}}$$

$$= 145.89\sqrt{\frac{1}{6} + \frac{(3-3.2)^2}{.74}} = 68.54$$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

$3533.8 \pm 2.776(68.54)$

3533.8 ± 190.27, or \$3343.53 to \$3724.07

b. $s_{\text{ind}} = s\sqrt{1 + \dfrac{1}{n} + \dfrac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}}$

$$= 145.89\sqrt{1 + \frac{1}{6} + \frac{(3-3.2)^2}{.74}} = 161.19$$

$\hat{y}_p \pm t_{\alpha/2} s_{\text{ind}}$

$3533.8 \pm 2.776(161.19)$

3533.8 ± 447.46, or \$3086.34 to \$3981.26

36. **a.** \$201
b. 167.25 to 234.65
c. 108.75 to 293.15

38. **a.** \$5046.67
b. \$3815.10 to \$6278.24
c. Not out of line

40. **a.** 9
b. $\hat{y} = 20.0 + 7.21x$
c. 1.3626
d. SSE = SST − SSR = 51,984.1 − 41,587.3 = 10,396.8

MSE = 10,396.8/7 = 1485.3

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{41{,}587.3}{1485.3} = 28.0$$

From the F table (1 numerator degree of freedom and 7 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $F = 28.0$ is .0011

Because p-value $\leq \alpha = .05$, we reject H_0: $\beta_1 = 0$

e. $\hat{y} = 20.0 + 7.21(50) = 380.5$, or \$380,500

42. a. $\hat{y} = 80.0 + 50.0x$

b. 30

c. Significant; p-value $= .000$

d. \$680,000

44. b. Yes

c. $\hat{y} = 2044.38 - 28.35$ weight

d. Significant; p-value $= .000$

e. .774; a good fit

45. a. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{70}{5} = 14,\ \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{76}{5} = 15.2,$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 200,\ \Sigma(x_i - \bar{x})^2 = 126$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{200}{126} = 1.5873$$

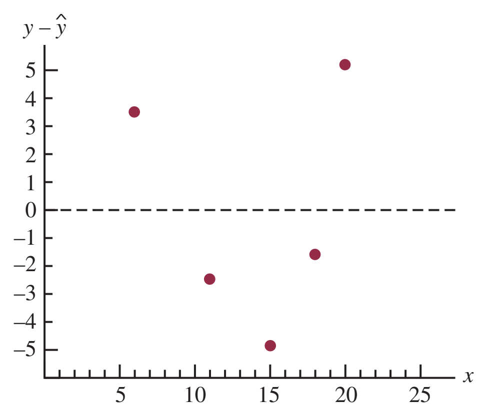
$b_0 = \bar{y} - b_1\bar{x} = 15.2 - (1.5873)(14) = -7.0222$

$\hat{y} = -7.02 + 1.59x$

b.

x_i	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$
6	6	2.52	3.48
11	8	10.47	−2.47
15	12	16.83	−4.83
18	20	21.60	−1.60
20	30	24.78	5.22

c.



With only five observations, it is difficult to determine whether the assumptions are satisfied; however, the plot does suggest curvature in the residuals, which would indicate that the error term assumptions are not satisfied; the scatter diagram for these data also indicates that the underlying relationship between x and y may be curvilinear

d. $s^2 = 23.78$

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\Sigma(x_i - \bar{x})^2} = \frac{1}{5} + \frac{(x_i - 14)^2}{126}$$

x_i	h_i	$s_{y_i - \hat{y}_i}$	$y_i - \hat{y}_i$	Standardized Residuals
6	.7079	2.64	3.48	1.32
11	.2714	4.16	−2.47	−.59
15	.2079	4.34	−4.83	−1.11
18	.3270	4.00	−1.60	−.40
20	.4857	3.50	5.22	1.49

e. The plot of the standardized residuals against $\hat{y}$ has the same shape as the original residual plot; as stated in part (c), the curvature observed indicates that the assumptions regarding the error term may not be satisfied

46. a. $\hat{y} = 2.32 + .64x$

b. No; the variance appears to increase for larger values of x

47. a. Let x = advertising expenditures and y = revenue

$\hat{y} = 29.4 + 1.55x$

b. SST = 1002, SSE = 310.28, SSR = 691.72

$$\text{MSR} = \frac{\text{SSR}}{1} = 691.72$$

$$\text{MSE} = \frac{\text{SSE}}{n-2} = \frac{310.28}{5} = 62.0554$$

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{691.72}{62.0554} = 11.15$$

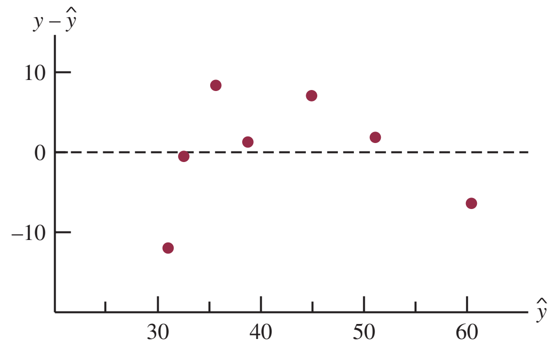
From the F table (1 numerator degree of freedom and 5 denominator), p-value is between .01 and .025

Using Excel or Minitab, p-value $= .0206$

Because p-value $\leq \alpha = .05$, we conclude that the two variables are related

c.

x_i	y_i	$\hat{y}_i = 29.40 + 1.55x_i$	$y_i - \hat{y}_i$
1	19	30.95	−11.95
2	32	32.50	−.50
4	44	35.60	8.40
6	40	38.70	1.30
10	52	44.90	7.10
14	53	51.10	1.90
20	54	60.40	−6.40



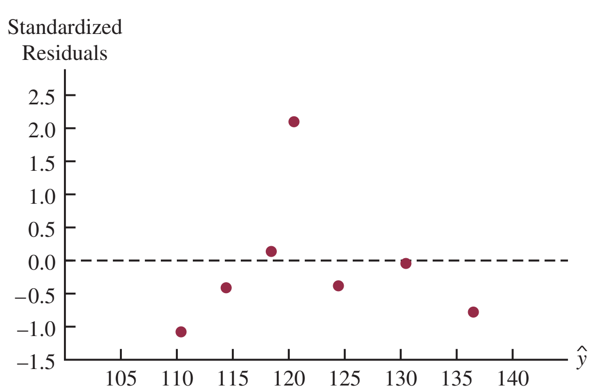
d. The residual plot leads us to question the assumption of a linear relationship between x and y; even though the relationship is significant at the $\alpha = .05$ level, it would be extremely dangerous to extrapolate beyond the range of the data

48. b. Yes

50. a. Using Minitab, we obtained the estimated regression equation $\hat{y} = 66.1 + .402x$; a portion of the Minitab output is shown in Figure D14.50; the fitted values and standardized residuals are shown:

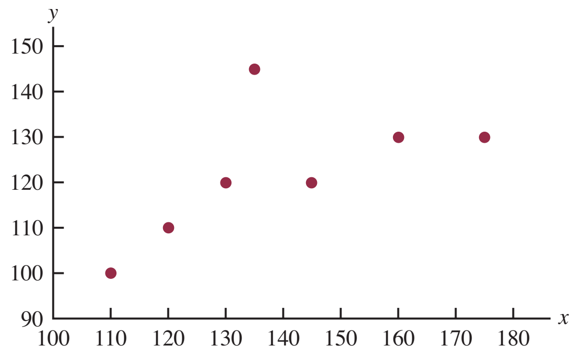
x_i	y_i	$\hat{y}_i$	Standardized Residuals
135	145	120.41	2.11
110	100	110.35	−1.08
130	120	118.40	.14
145	120	124.43	−.38
175	130	136.50	−.78
160	130	130.47	−.04
120	110	114.38	−.41

b.



The standardized residual plot indicates that the observation $x = 135$, $y = 145$ may be an outlier; note that this observation has a standardized residual of 2.11

c. The scatter diagram is shown:



The scatter diagram also indicates that the observation $x = 135$, $y = 145$ may be an outlier; the implication is that for simple linear regression outliers can be identified by looking at the scatter diagram

52. a. A portion of the Minitab output is shown in Figure D14.52
b. Minitab identifies observation 1 as having a large standardized residual; thus, we would consider observation 1 to be an outlier

54. b. Value = −252 + 5.83 Revenue
c. There are five unusual observations (9, 19, 21, 22, and 32).

58. a. $\hat{y} = 9.26 + .711x$
b. Significant; p-value = .001
c. $r^2 = .744$; good fit
d. \$13.53

60. b. GR(%) = 25.4 + .285 RR(%)
c. Significant; p-value = .000
d. No; $r^2 = .449$
e. Yes
f. Yes

62. a. $\hat{y} = 22.2 - .148x$
b. Significant relationship; p-value = .028
c. Good fit; $r^2 = .739$
d. 12.294 to 17.271

64. a. $\hat{y} = 220 + 132x$
b. Significant; p-value = .000
c. $r^2 = .873$; very good fit
d. \$559.50 to \$933.90

66. a. Market beta = .95
b. Significant; p-value = .029
c. $r^2 = .470$; not a good fit
d. Xerox has a higher risk

68. b. There appears to be a positive linear relationship between the two variables
c. $\hat{y} = 9.37 + 1.2875$ Top Five (%)
d. Significant; p-value = .000
e. $r^2 = .741$; good fit
f. $r_{xy} = .86$

FIGURE D14.50

```
The regression equation is
Y = 66.1 + 0.402 X

Predictor      Coef    SE Coef      T       p
Constant      66.10      32.06   2.06   0.094
X            0.4023     0.2276   1.77   0.137

S = 12.62    R-sq = 38.5%    R-sq(adj) = 26.1%

Analysis of Variance

SOURCE           DF       SS       MS       F       p
Regression        1    497.2    497.2    3.12   0.137
Residual Error    5    795.7    159.1
Total             6   1292.9

Unusual Observations
Obs     X        Y      Fit    SE Fit    Residual    St Resid
  1   135   145.00   120.42      4.87       24.58       2.11R

 R denotes an observation with a large standardized residual
```

FIGURE D14.52

```
The regression equation is
Shipment = 4.09 + 0.196 Media$

Predictor      Coef    SE Coef      T       p
Constant      4.089      2.168   1.89   0.096
Media$      0.19552    0.03635   5.38   0.000

S = 5.044    R-Sq = 78.3%    R-Sq(adj) = 75.6%

Analysis of Variance
Source           DF       SS       MS       F       p
Regression        1   735.84   735.84   28.93   0.000
Residual Error    8   203.51    25.44
Total             9   939.35

Unusual Observations
Obs    Media$    Shipment     Fit    SE Fit    Residual    St Resid
  1       120       36.30   27.55      3.30        8.75       2.30R

R denotes an observation with a large standardized residual
```

Chapter 15

2. a. The estimated regression equation is
$\hat{y} = 45.06 + 1.94x_1$
An estimate of y when $x_1 = 45$ is
$\hat{y} = 45.06 + 1.94(45) = 132.36$

b. The estimated regression equation is
$\hat{y} = 85.22 + 4.32x_2$
An estimate of y when $x_2 = 15$ is
$\hat{y} = 85.22 + 4.32(15) = 150.02$

c. The estimated regression equation is
$\hat{y} = -18.37 + 2.01x_1 + 4.74x_2$
An estimate of y when $x_1 = 45$ and $x_2 = 15$ is
$\hat{y} = -18.37 + 2.01(45) + 4.74(15) = 143.18$

4. a. $255,000

5. a. The Minitab output is shown in Figure D15.5a

b. The Minitab output is shown in Figure D15.5b

c. It is 1.60 in part (a) and 2.29 in part (b); in part (a) the coefficient is an estimate of the change in revenue

FIGURE D15.5a

```
The regression equation is
Revenue = 88.6 + 1.60 TVAdv

Predictor       Coef    SE Coef       T       p
Constant      88.638      1.582   56.02   0.000
TVAdv         1.6039     0.4778    3.36   0.015

S = 1.215     R-sq = 65.3%     R-sq(adj) = 59.5%

Analysis of Variance

SOURCE            DF        SS        MS       F       p
Regression         1    16.640    16.640   11.27   0.015
Residual Error     6     8.860     1.477
Total              7    25.500
```

FIGURE D15.5b

```
The regression equation is
Revenue = 83.2 + 2.29 TVAdv + 1.30 NewsAdv

Predictor       Coef    SE Coef       T       p
Constant      83.230      1.574   52.88   0.000
TVAdv         2.2902     0.3041    7.53   0.001
NewsAdv       1.3010     0.3207    4.06   0.010

S = 0.6426     R-sq = 91.9%     R-sq(adj) = 88.7%

Analysis of Variance

SOURCE            DF        SS        MS       F       p
Regression         2    23.435    11.718   28.38   0.002
Residual Error     5     2.065     0.413
Total              7    25.500
```

due to a one-unit change in television advertising expenditures; in part (b) it represents an estimate of the change in revenue due to a one-unit change in television advertising expenditures when the amount of newspaper advertising is held constant

d. Revenue = 83.2 + 2.29(3.5) + 1.30(1.8) = 93.56 or $93,560

6. a. Proportion Won = .354 + .000888 HR
b. Proportion Won = .865 − .0837 ERA
c. Proportion Won = .709 + .00140 HR − .103 ERA

8. a. $\hat{y} = 31054 + 1328.7$ Reliability
b. $\hat{y} = 21313 + 136.69$ Score $- 1446.3$ Reliability
c. $26,643

10. a. PCT = −1.22 + 3.96 FG%
b. Increase of 1% in FG% will increase PCT by .04
c. PCT = −1.23 + 4.82 FG% − 2.59 Opp 3 Pt% + .0344 Opp TO
d. Increase FG%; decrease Opp 3 Pt%; increase Opp TO
e. .638

12. a. $R^2 = \dfrac{\text{SSR}}{\text{SST}} = \dfrac{14{,}052.2}{15{,}182.9} = .926$

b. $R_a^2 = 1 - (1 - R^2)\dfrac{n-1}{n-p-1}$

$$= 1 - (1 - .926)\frac{10-1}{10-2-1} = .905$$

c. Yes; after adjusting for the number of independent variables in the model, we see that 90.5% of the variability in y has been accounted for

14. a. .75 **b.** .68

15. a. $R^2 = \frac{\text{SSR}}{\text{SST}} = \frac{23.435}{25.5} = .919$

$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}$

$= 1 - (1 - .919)\frac{8-1}{8-2-1} = .887$

b. Multiple regression analysis is preferred because both R^2 and R_a^2 show an increased percentage of the variability of y explained when both independent variables are used

16. a. No, $R^2 = .153$

b. Better fit with multiple regression

18. a. $R^2 = .564$, $R_a^2 = .511$

b. The fit is not very good

19. a. $\text{MSR} = \frac{\text{SSR}}{p} = \frac{6216.375}{2} = 3108.188$

$\text{MSE} = \frac{\text{SSE}}{n-p-1} = \frac{507.75}{10-2-1} = 72.536$

b. $F = \frac{\text{MSR}}{\text{MSE}} = \frac{3108.188}{72.536} = 42.85$

From the F table (2 numerator degrees of freedom and 7 denominator), p-value is less than .01

Using Excel or Minitab the p-value corresponding to $F = 42.85$ is .0001

Because p-value $\le \alpha$, the overall model is significant

c. $t = \frac{b_1}{s_{b_1}} = \frac{.5906}{.0813} = 7.26$

p-value = .0002

Because p-value $\le \alpha$, β_1 is significant

d. $t = \frac{b_2}{s_{b_2}} = \frac{.4980}{.0567} = 8.78$

p-value = .0001

Because p-value $\le \alpha$, β_2 is significant

20. a. Significant; p-value = .000

b. Significant; p-value = .000

c. Significant; p-value = .002

22. a. SSE = 4000, $s^2 = 571.43$, MSR = 6000

b. Significant; p-value = .008

23. a. $F = 28.38$

p-value = .002

Because p-value $\le \alpha$, there is a significant relationship

b. $t = 7.53$

p-value = .001

Because p-value $\le \alpha$, β_1 is significant and x_1 should not be dropped from the model

c. $t = 4.06$

p-value = .010

Because p-value $\le \alpha$, β_2 is significant and x_2 should not be dropped from the model

24. a. $\hat{y} = -.682 + .0498$ Revenue $+ .0147$ % Wins

b. Significant; p-value = .001

c. Revenue is significant; p-value = .001
%Wins is significant; p-value = .025

26. a. Significant; p-value = .000

b. All significant; p-values are all $< \alpha = .05$

28. a. Using Minitab, the 95% confidence interval is 132.16 to 154.16

b. Using Minitab, the 95% prediction interval is 111.13 at 175.18

29. a. See Minitab output in Figure D15.5b.

$\hat{y} = 83.23 + 2.29(3.5) + 1.30(1.8) = 93.555$ or \$93,555

b. Minitab results: 92.840 to 94.335, or \$92,840 to \$94,335

c. Minitab results: 91.774 to 95.401, or \$91,774 to \$95,401

30. a. 46.758 to 50.646

b. 44.815 to 52.589

32. a. $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

where $x_2 = \begin{cases} 0 \text{ if level 1} \\ 1 \text{ if level 2} \end{cases}$

b. $E(y) = \beta_0 + \beta_1 x_1 + \beta_2(0) = \beta_0 + \beta_1 x_1$

c. $E(y) = \beta_0 + \beta_1 x_1 + \beta_2(1) = \beta_0 + \beta_1 x_1 + \beta_2$

d. $\beta_2 = E(y \mid \text{level 2}) - E(y \mid \text{level 1})$

β_1 is the change in $E(y)$ for a 1-unit change in x_1 holding x_2 constant

34. a. \$15,300

b. $\hat{y} = 10.1 - 4.2(2) + 6.8(8) + 15.3(0) = 56.1$

Sales prediction: \$56,100

c. $\hat{y} = 10.1 - 4.2(1) + 6.8(3) + 15.3(1) = 41.6$

Sales prediction: \$41,600

36. a. $\hat{y} = 1.86 + 0.291$ Months $+ 1.10$ Type $- 0.609$ Person

b. Significant; p-value = .002

c. Person is not significant; p-value = .167

38. a. $\hat{y} = -91.8 + 1.08$ Age $+ .252$ Pressure $+ 8.74$ Smoker

b. Significant; p-value = .01

c. 95% prediction interval is 21.35 to 47.18 or a probability of .2135 to .4718; quit smoking and begin some type of treatment to reduce his blood pressure

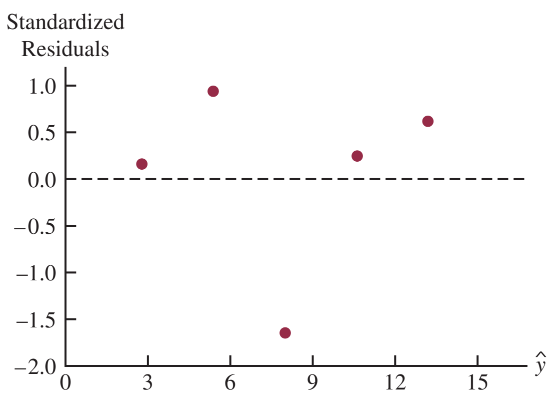
39. a. The Minitab output is shown in Figure D15.39

b. Minitab provides the following values:

x_i	y_i	$\hat{y}_i$	Standardized Residual
1	3	2.8	.16
2	7	5.4	.94
3	5	8.0	−1.65
4	11	10.6	.24
5	14	13.2	.62

FIGURE D15.39

```
The regression equation is
Y = 0.20 + 2.60 X

Predictor       Coef     SE Coef       T        p
Constant       0.200       2.132    0.09    0.931
X             2.6000      0.6429    4.04    0.027

S = 2.033     R-sq = 84.5%     R-sq(adj) = 79.3%

Analysis of Variance
SOURCE            DF         SS         MS        F        p
Regression         1     67.600     67.600    16.35    0.027
Residual Error     3     12.400      4.133
Total              4     80.000
```



The point (3,5) does not appear to follow the trend of the remaining data; however, the value of the standardized residual for this point, −1.65, is not large enough for us to conclude that (3,5) is an outlier

c. Minitab provides the following values:

x_i	y_i	Studentized Deleted Residual
1	3	.13
2	7	.91
3	5	−4.42
4	11	.19
5	14	.54

$t_{.025} = 4.303$ ($n - p - 2 = 5 - 1 - 2 = 2$ degrees of freedom)

Because the studentized deleted residual for (3,5) is $-4.42 < -4.303$, we conclude that the 3rd observation is an outlier

40. a. $\hat{y} = -53.3 + 3.11x$

b. −1.94, −.12, 1.79, .40, −1.90; no

c. .38, .28, .22, .20, .92; no

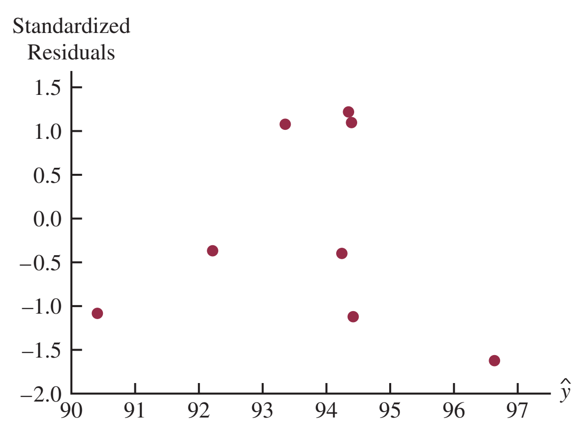
d. .60, .00, .26, .03, 11.09; yes, the fifth observation

41. a. The Minitab output appears in Figure D15.5b; the estimated regression equation is

Revenue = 83.2 + 2.29 TVAdv + 1.30 NewsAdv

b. Minitab provides the following values:

$\hat{y}_i$	Standardized Residual	$\hat{y}_i$	Standardized Residual
96.63	−1.62	94.39	1.10
90.41	−1.08	94.24	−.40
94.34	1.22	94.42	−1.12
92.21	−.37	93.35	1.08



With relatively few observations, it is difficult to determine whether any of the assumptions regarding ϵ have been violated; for instance, an argument could be made that there does not appear to be any pattern in the plot; alternatively, an argument could be made that there is a curvilinear pattern in the plot

c. The values of the standardized residuals are greater than −2 and less than +2; thus, using this test, there are no outliers

As a further check for outliers, we used Minitab to compute the following studentized deleted residuals:

Observation	Studentized Deleted Residual	Observation	Studentized Deleted Residual
1	−2.11	5	1.13
2	−1.10	6	−.36
3	1.31	7	−1.16
4	−.33	8	1.10

$t_{.025} = 2.776$ ($n - p - 2 = 8 - 2 - 2 = 4$ degrees of freedom)

Because none of the studentized deleted residuals are less than −2.776 or greater than 2.776, we conclude that there are no outliers in the data

d. Minitab provides the following values:

Observation	h_i	D_i
1	.63	1.52
2	.65	.70
3	.30	.22
4	.23	.01
5	.26	.14
6	.14	.01
7	.66	.81
8	.13	.06

The critical leverage value is

$$\frac{3(p + 1)}{n} = \frac{3(2 + 1)}{8} = 1.125$$

Because none of the values exceed 1.125, we conclude that there are no influential observations; however, using Cook's distance measure, we see that $D_1 > 1$ (rule of thumb critical value); thus, we conclude that the first observation is influential

Final conclusion: observation 1 is an influential observation

42. b. Unusual trend
c. No outliers
d. Observation 2 is an influential observation

44. a. $E(y) = \dfrac{e^{\beta_0+\beta_1 x}}{1 + e^{\beta_0+\beta_1 x}}$
b. Estimate of the probability that a customer who does not have a Simmons credit card will make a purchase
c. $\hat{g}(x) = -0.9445 + 1.0245x$
d. .28 for customers who do not have a Simmons credit card
.52 for customers who have a Simmons credit card
e. Estimated odds ratio = 2.79

46. a. $E(y) = \dfrac{e^{\beta_0+\beta_1 x}}{1 + e^{\beta_0+\beta_1 x}}$
b. $E(y) = \dfrac{e^{-2.6355+0.22018x}}{1 + e^{-2.6355+0.22018x}}$
c. Significant; p-value = .0002
d. .39
e. $1200
f. Estimated odds ratio = 1.25

48. a. $E(y) = \dfrac{e^{\beta_0+\beta_1 x}}{1 + e^{\beta_0+\beta_1 x}}$
b. $\hat{g}(x) = -2.805 + 1.1492x$
c. .86
d. Estimated odds ratio = 3.16

50. b. 67.39

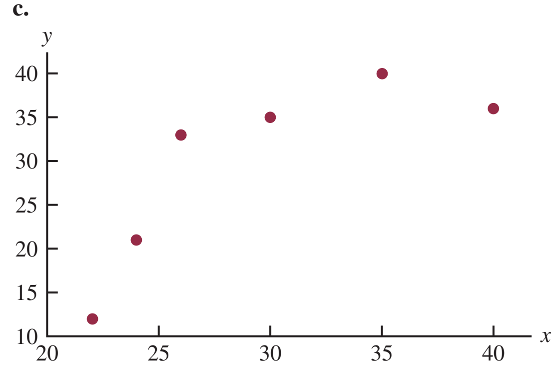
52. a. $\hat{y} = -1.41 + .0235x_1 + .00486x_2$
b. Significant; p-value = .0001
c. Both significant
d. $R^2 = .937$; $R_a^2 = 9.19$; good fit

54. a. Buy Again = −7.522 + 1.8151 Steering
b. Yes
c. Buy Again = −5.388 + .6899 Steering + .9113 Treadwear
d. Significant; p-value = .001

56. a. $\hat{y}$ = 4.9090 + 10.4658 FundDE + 21.6823 FundIE
b. $R^2 = .6144$; reasonably good fit
c. $\hat{y}$ = 1.1899 + 6.8969 FundDE + 17.6800 FundIE + 0.0265 Net Asset Value ($) + 6.4564 Expense Ratio (%)
Net Asset Value ($) is not significant and can be deleted
d. $\hat{y}$ = −4.6074 + 8.1713 FundDE + 19.5194 FundIE + 5.5197 Expense Ratio (%) + 5.9237 3StarRank + 8.2367 4StarRank + 6.6241 5StarRank
e. 15.28%

Chapter 16

1. a. The Minitab output is shown in Figure D16.1a
b. Because the p-value corresponding to $F = 6.85$ is $.059 > \alpha = .05$, the relationship is not significant
c.



The scatter diagram suggests that a curvilinear relationship may be appropriate
d. The Minitab output is shown in Figure D16.1d
e. Because the p-value corresponding to $F = 25.68$ is $.013 < \alpha = .05$, the relationship is significant
f. $\hat{y} = -168.88 + 12.187(25) - .17704(25)^2 = 25.145$

2. a. $\hat{y} = 9.32 + .424x$; p-value = .117 indicates a weak relationship between x and y

FIGURE D16.1a

```
The regression equation is
Y = - 6.8 + 1.23 X

Predictor       Coef    SE Coef        T        p
Constant       -6.77      14.17    -0.48    0.658
X             1.2296     0.4697     2.62    0.059

S = 7.269     R-sq = 63.1%     R-sq(adj) = 53.9%

Analysis of Variance

SOURCE            DF         SS       MS       F        p
Regression         1     362.13   362.13    6.85    0.059
Residual Error     4     211.37    52.84
Total              5     573.50
```

FIGURE D16.1d

```
The regression equation is
Y = - 169 + 12.2 X - 0.177 XSQ

Predictor        Coef    SE Coef        T        p
Constant      -168.88      39.79    -4.74    0.024
X              12.187      2.663     4.58    0.020
XSQ          -0.17704    0.04290    -4.13    0.026

S = 3.248     R-sq = 94.5%     R-sq(adj) = 90.8%

Analysis of Variance

SOURCE            DF         SS       MS       F        p
Regression         2     541.85   270.92   25.68    0.013
Residual Error     3      31.65    10.55
Total              5     573.50
```

b. $\hat{y} = -8.10 + 2.41x - .0480x^2$
$R_a^2 = .932$; a good fit
c. 20.965

4. a. $\hat{y} = 943 + 8.71x$
b. Significant; p-value $= .005 < \alpha = .01$

5. a. The Minitab output is shown in Figure D16.5a
b. Because the p-value corresponding to $F = 73.15$ is $.003 < \alpha = .01$, the relationship is significant; we would reject H_0: $\beta_1 = \beta_2 = 0$
c. See Figure D16.5c

6. b. No, the relationship appears to be curvilinear
c. Several possible models; e.g.,
$\hat{y} = 2.90 - .185x + .00351x^2$

8. a. It appears that a simple linear regression model is not appropriate
b. Price = 33829 − 4571 Rating + 154 RatingSq
c. logPrice = −10.2 + 10.4 logRating
d. Part (c); higher percentage of variability is explained

10. a. Significant; p-value $= .000$
b. Significant; p-value $= .000$

11. a. SSE = 1805 − 1760 = 45

$$F = \frac{\text{MSR}}{\text{MSE}} = \left(\frac{1760/4}{45/25}\right) = 244.44$$

Because p-value $= .000$ the relationship is significant
b. $\text{SSE}(x_1, x_2, x_3, x_4) = 45$
c. $\text{SSE}(x_2, x_3) = 1805 - 1705 = 100$
d. $F = \dfrac{(100 - 45)/2}{1.8} = 15.28$

Because p-value $= .000$, x_1 and x_2 are significant

FIGURE D16.5a

```
The regression equation is
Y = 433 + 37.4 X -0.383 XSQ

Predictor        Coef     SE Coef        T        p
Constant        432.6       141.2     3.06    0.055
X              37.429       7.807     4.79    0.017
XSQ           -0.3829      0.1036    -3.70    0.034

S = 15.83      R-sq = 98.0%     R-sq(adj) = 96.7%

Analysis of Variance

SOURCE              DF        SS        MS        F        p
Regression           2     36643     18322    73.15    0.003
Residual Error       3       751       250
Total                5     37395
```

FIGURE D16.5c

```
    Fit     Stdev.Fit              95% C.I.                95% P.I.
1302.01       9.93        (1270.41, 1333.61)    (1242.55, 1361.47)
```

12. a. The Minitab output is shown in Figure D16.12a
b. The Minitab output is shown in Figure D16.12b
c. $F = \dfrac{[\text{SSE(reduced)} - \text{SSE(full)}]/(\text{\# extra terms})}{\text{MSE(full)}}$
$= \dfrac{(7.2998 - 4.3240)/2}{.1663} = 8.95$
The p-value associated with $F = 8.95$ (2 numerator degrees of freedom and 26 denominator) is .001; with a p-value $< \alpha = .05$, the addition of the two independent variables is significant

14. a. $\hat{y} = -111 + 1.32$ Age $+ .296$ Pressure
b. $\hat{y} = -123 + 1.51$ Age $+ .448$ Pressure $+ 8.87$ Smoker $- .00276$ AgePress
c. Significant; p-value $= .000$

16. a. Weeks $= -8.9 + 1.51$ Age
b. Weeks $= -.07 + 1.73$ Age $- 2.7$ Manager $- 15.1$ Head $- 17.4$ Sales
c. Same as part (b)
d. Same as part (b)
e. Weeks $= 13.1 + 1.64$ Age $- 9.76$ Married $- 19.4$ Head $- 29.0$ Manager $- 19.0$ Sales

18. a. RPG $= -4.05 + 27.6$ OBP
b. A variety of models will provide a good fit; the five-variable model identified using Minitab's Stepwise Regression procedure with Alpha-to-Enter $= .10$ and Alpha-to-Remove $= .10$ follows:
RPG $= -.0909 + 32.2$ OBP $+ .109$ HR $- 21.5$ AVG $+ .244$ 3B $- .0223$ BB

20.

x_1	x_2	x_3	Treatment
0	0	0	A
1	0	0	B
0	1	0	C
0	0	1	D

$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

22. Factor A: $x_1 = 0$ if level 1 and 1 if level 2
Factor B:

x_2	x_3	Level
0	0	1
1	0	2
0	1	3

$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1 x_3$

24. a. Not significant at the .05 level of significance; p-value $= .093$
b. 139

26. Overall significant; p-value $= .029$
Individually, none of the variables are significant at the .05 level of significance; a larger sample size would be helpful

28. $d = 1.60$; test is inconclusive

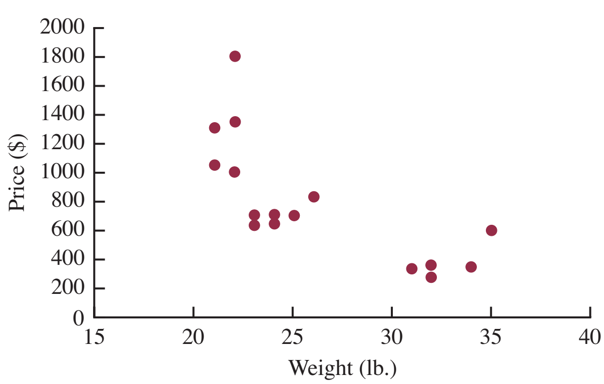
FIGURE D16.12a

```
The regression equation is
Scoring Avg. = 46.3 + 14.1 Putting Avg.

Predictor          Coef    SE Coef      T       p
Constant         46.277      6.026   7.68   0.000
Putting Avg.     14.103      3.356   4.20   0.000

S = 0.510596    R-Sq = 38.7%    R-Sq(adj) = 36.5%

Analysis of Variance

SOURCE            DF        SS        MS       F        p
Regression         1    4.6036    4.6036   17.66   0.0000
Residual Error    28    7.2998    0.2607
Total             29   11.9035
```

FIGURE D16.12b

```
The regression equation is
Scoring Avg. = 59.0 - 10.3 Greens in Reg.
  + 11.4 Putting Avg - 1.81 Sand Saves

Predictor              Coef    SE Coef        T        p
Constant             59.022      5.774    10.22    0.000
Greens in Reg.      -10.281      2.877    -3.57    0.001
Putting Avg.         11.413      2.760     4.14    0.000
Sand Saves          -1.8130     0.9210    -1.97    0.060

S = 0.407808     R-Sq = 63.7%     R-Sq(adj) = 59.5%

Analysis of Variance

Source             DF        SS        MS        F        p
Regression          3    7.5795    2.5265    15.19    0.000
Residual Error     26    4.3240    0.1663
Total              29   11.9035
```

30. a.



There appears to be a curvilinear relationship between weight and price

b. A portion of the Minitab output follows:

```
The regression equation is
Price = 11376 - 728 Weight + 12.0 WeightSq

Predictor    Coef   SE Coef       T       p
Constant    11376      2565    4.43   0.000
Weight     -728.3     193.7   -3.76   0.002
WeightSq   11.974     3.539    3.38   0.004

S = 242.804    R-Sq = 77.0%    R-Sq(adj) = 74.1%

Analysis of Variance

SOURCE           DF        SS        MS       F       p
Regression        2   3161747   1580874   26.82   0.000
Residual Error   16    943263     58954
Total            18   4105011
```

The results obtained support the conclusion that there is a curvilinear relationship between weight and price

c. A portion of the Minitab output follows:

```
The regression equation is
Price = 1284 - 572 Type_Fitness - 907 Type_Comfort

Predictor         Coef   SE Coef       T       p
Constant       1283.75     95.22   13.48   0.000
Type_Fitness    -571.8     153.5   -3.72   0.002
Type_Comfort    -907.1     145.5   -6.24   0.000

S = 269.328    R-Sq = 71.7%    R-Sq(adj) = 68.2%

Analysis of Variance

SOURCE           DF        SS        MS       F       p
Regression        2   2944410   1472205   20.30   0.000
Residual Error   16   1160601     72538
Total            18   4105011
```

Type of bike appears to be a significant factor in predicting price, but the estimated regression equation developed in part (b) appears to provide a slightly better fit

d. A portion of the Minitab output follows; in this output WxF denotes the interaction between the weight of the bike and the dummy variable Type_Fitness and WxC denotes the interaction between the weight of the bike and the dummy variable Type_Comfort

```
The regression equation is
Price = 5924 - 214 Weight - 6343 Type_Fitness - 7232
        Type_Comfort + 261 WxF + 266 WxC

Predictor         Coef   SE Coef       T       p
Constant          5924      1547    3.83   0.002
Weight         -214.56     71.42   -3.00   0.010
Type_Fitness     -6343      2596   -2.44   0.030
```

```
Type_Comfort      -7232       2518   -2.87   0.013
WxF               261.3      111.8    2.34   0.036
WxC              266.41      93.98    2.83   0.014

S = 224.438    R-Sq = 84.0%    R-Sq(adj) = 77.9%

Analysis of Variance

SOURCE           DF        SS       MS       F       p
Regression        5   3450170   690034   13.70   0.000
Residual Error   13    654841    50372
Total            18   4105011
```

By taking into account the type of bike, the weight, and the interaction between these two factors, this estimated regression equation provides an excellent fit

32. a. Delay = 63.0 + 11.1 Industry; no significant positive autocorrelation

34. Significant differences between comfort levels for the three types of browsers; p-value = .034

Chapter 17

1. a.

Item	Price Relative
A	103 = (7.75/7.50)(100)
B	238 = (1500/630)(100)

b. $I_{2009} = \dfrac{7.75 + 1500.00}{7.50 + 630.00}(100) = \dfrac{1507.75}{637.50}(100) = 237$

c. $I_{2009} = \dfrac{7.75(1500) + 1500.00(2)}{7.50(1500) + 630.00(2)}(100)$

$= \dfrac{14{,}625.00}{12{,}510.00}(100) = 117$

d. $I_{2009} = \dfrac{7.75(1800) + 1500.00(1)}{7.50(1800) + 630.00(1)}(100)$

$= \dfrac{15{,}450.00}{14{,}130.00}(100) = 109$

2. a. 32%

b. $8.14

3. a. Price relatives for A = (6.00/5.45)100 = 110
B = (5.95/5.60)100 = 106
C = (6.20/5.50)100 = 113

b. $I_{2009} = \dfrac{6.00 + 5.95 + 6.20}{5.45 + 5.60 + 5.50}(100) = 110$

c. $I_{2009} = \dfrac{6.00(150) + 5.95(200) + 6.20(120)}{5.45(150) + 5.60(200) + 5.50(120)}(100)$
$= 109$

9% increase over the two-year period

4. $I_{2009} = 114$

6.

Item	Price Relative	Base Period Price	Base Period Usage	Weight	Weighted Price Relative
A	150	22.00	20	440	66,000
B	90	5.00	50	250	22,500
C	120	14.00	40	560	67,200
			Totals	1250	155,700

$$I = \frac{155{,}700}{1250} = 125$$

7. a. Price relatives for A = (3.95/2.50)100 = 158
B = (9.90/8.75)100 = 113
C = (.95/.99)100 = 96

b.

Item	Price Relative	Base Price	Quantity	Weight $P_{i0}Q_i$	Weighted Price Relative
A	158	2.50	25	62.5	9,875
B	113	8.75	15	131.3	14,837
C	96	.99	60	59.4	5,702
			Totals	253.2	30,414

$$I = \frac{30{,}414}{253.2} = 120$$

Cost of raw materials is up 20% for the chemical

8. I = 105; portfolio is up 5%

10. a. Deflated 1996 wages: $\dfrac{\$11.86}{154.9}(100) = \$7.66$

Deflated 2009 wages: $\dfrac{\$18.55}{212.2}(100) = \$8.74$

b. $\dfrac{18.55}{11.86}(100) = 156.4$; the percentange increase in actual wages is 56.4%

c. $\dfrac{8.74}{7.66}(100) = 114.1$; the change in real wages is an increase of 14.1%

12. a. 2428, 2490, 2451
Manufacturing shipments increased slightly in constant dollars
b. 3043, 3132, 3050
c. PPI

14. $$I = \frac{300(18.00) + 400(4.90) + 850(15.00)}{350(18.00) + 220(4.90) + 730(15.00)}(100)$$
$$= \frac{20,110}{18,328}(100) = 110$$

15. $$I = \frac{95(1200) + 75(1800) + 50(2000) + 70(1500)}{120(1200) + 86(1800) + 35(2000) + 60(1500)}(100)$$
$= 99$
Quantities are down slightly

16. $I = 83$

18. a. 151, 197, 143, 178
b. $I = 170$

20. $I_{\text{Jan}} = 73.5, I_{\text{Mar}} = 70.1$

22. $I = 182.7$

24. \$36,082; \$32,528; \$27,913; \$34,387; \$40,551; \$42,651; \$46,458; \$56,324

26. $I = 143$; quantity is up 43%

Chapter 18

1. The following table shows the calculations for parts (a), (b), and (c):

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	18						
2	13	18	−5	5	25	−38.46	38.46
3	16	13	3	3	9	18.75	18.75
4	11	16	−5	5	25	−45.45	45.45
5	17	11	6	6	36	35.29	35.29
6	14	17	−3	3	9	−21.43	21.43
			Totals	22	104	−51.30	159.38

a. $\text{MAE} = \frac{22}{5} = 4.4$
b. $\text{MSE} = \frac{104}{5} = 20.8$
c. $\text{MAPE} = \frac{159.38}{5} = 31.88$
d. Forecast for week 7 is 14

2. The following table shows the calculations for parts (a), (b), and (c):

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	18						
2	13	18.00	−5.00	5.00	25.00	−38.46	38.46
3	16	15.50	0.50	0.50	0.25	3.13	3.13
4	11	15.67	−4.67	4.67	21.81	−42.45	42.45
5	17	14.50	2.50	2.50	6.25	14.71	14.71
6	14	15.00	−1.00	1.00	1.00	−7.14	7.14
			Totals	13.67	54.31	−70.21	105.86

a. $\text{MAE} = \dfrac{13.67}{5} = 2.73$

b. $\text{MSE} = \dfrac{54.31}{5} = 10.86$

c. $\text{MAPE} = \dfrac{105.89}{5} = 21.18$

d. Forecast for week 7 is

$$\frac{18 + 13 + 16 + 11 + 17 + 14}{6} = 14.83$$

4. a. $\text{MSE} = \dfrac{363}{6} = 60.5$

Forecast for month 8 is 15

b. $\text{MSE} = \dfrac{216.72}{6} = 36.12$

Forecast for month 8 is 18

c. The average of all the previous values is better because MSE is smaller

5. a. The data appear to follow a horizontal pattern

b. Three-week moving average.

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	18			
2	13			
3	16			
4	11	15.67	−4.67	21.78
5	17	13.33	3.67	13.44
6	14	14.67	−0.67	0.44
			Total	35.67

$$\text{MSE} = \frac{35.67}{3} = 11.89$$

The forecast for week 7 $= \dfrac{(11 + 17 + 14)}{3} = 14$

c. Smoothing constant = .2

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	18			
2	13	18.00	−5.00	25.00
3	16	17.00	−1.00	1.00
4	11	16.80	−5.80	33.64
5	17	15.64	1.36	1.85
6	14	15.91	−1.91	3.66
			Total	65.15

$$\text{MSE} = \frac{65.15}{5} = 13.03$$

The forecast for week 7 is $.2(14) + (1 - .2)15.91 = 15.53$

d. The three-week moving average provides a better forecast since it has a smaller MSE

e. Smoothing constant = .4

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	18			
2	13	18.00	−5.00	25.00
3	16	16.00	0.00	0.00
4	11	16.00	−5.00	25.00
5	17	14.00	3.00	9.00
6	14	15.20	−1.20	1.44
			Total	60.44

$$\text{MSE} = \frac{60.44}{5} = 12.09$$

The exponential smoothing forecast using $\alpha = .4$ provides a better forecast than the exponential smoothing forecast using $\alpha = .2$ since it has a smaller MSE

6. a. The data appear to follow a horizontal pattern

b. $\text{MSE} = \dfrac{110}{4} = 27.5$

The forecast for week 8 is 19

c. $\text{MSE} = \dfrac{252.87}{6} = 42.15$

The forecast for week 7 is 19.12

d. The three-week moving average provides a better forecast since it has a smaller MSE

e. MSE = 39.79

The exponential smoothing forecast using $\alpha = .4$ provides a better forecast than the exponential smoothing forecast using $\alpha = .2$ since it has a smaller MSE

8. a.

Week	4	5	6	7	8	9	10	11	12
Forecast	19.33	21.33	19.83	17.83	18.33	18.33	20.33	20.33	17.83

b. MSE = 11.49

Prefer the unweighted moving average here; it has a smaller MSE

c. You could always find a weighted moving average at least as good as the unweighted one; actually the unweighted moving average is a special case of the weighted ones where the weights are equal

10. b. The more recent data receive the greater weight or importance in determining the forecast; the moving averages method weights the last n data values equally in determining the forecast

12. a. The data appear to follow a horizontal pattern

b. MSE(3-Month) = .12

MSE(4-Month) = .14

Use 3-Month moving averages

c. 9.63

13. a. The data appear to follow a horizontal pattern

b.

Month	Time-Series Value	3-Month Moving Average Forecast	(Error)^2	α = .2 Forecast	(Error)^2
1	240				
2	350			240.00	12100.00
3	230			262.00	1024.00
4	260	273.33	177.69	255.60	19.36
5	280	280.00	0.00	256.48	553.19
6	320	256.67	4010.69	261.18	3459.79
7	220	286.67	4444.89	272.95	2803.70
8	310	273.33	1344.69	262.36	2269.57
9	240	283.33	1877.49	271.89	1016.97
10	310	256.67	2844.09	265.51	1979.36
11	240	286.67	2178.09	274.41	1184.05
12	230	263.33	1110.89	267.53	1408.50
		Totals	17,988.52		27,818.49

MSE (3-Month) = 17,988.52/9 = 1998.72

MSE (α = .2) = 27,818.49/11 = 2528.95

Based on the preceding MSE values, the 3-Month moving averages appear better; however, exponential smoothing was penalized by including month 2, which was difficult for any method to forecast; using only the errors for months 4 to 12, the MSE for exponential smoothing is

$$\text{MSE}(\alpha = .2) = 14{,}694.49/9 = 1632.72$$

Thus, exponential smoothing was better considering months 4 to 12

c. Using exponential smoothing,

$$F_{13} = \alpha Y_{12} + (1 - \alpha)F_{12}$$
$$= .20(230) + .80(267.53) = 260$$

14. a. The data appear to follow a horizontal pattern

b. Values for months 2–12 are as follows:

105.00	114.00	115.80	112.56	105.79	110.05
120.54	126.38	118.46	106.92	104.85	

MSE = 510.29

c. Values for months 2–12 are as follows:

105.00	120.00	120.00	112.50	101.25	110.63
127.81	133.91	116.95	98.48	99.24	

MSE = 540.55

Conclusion: A smoothing constant of .3 is better than a smoothing constant of .5 since the MSE is less for 0.3

16. a. The time series plot indicates a possible linear trend in the data; this could be due to decreasing viewer interest in watching the Masters, but closer inspection of the data indicates that the two highest ratings correspond to years 1997 and 2001, years in which Tiger Woods won the tournament; the pattern observed may be simply due to the effect Tiger Woods has on ratings and not necessarily on any long-term decrease in viewer interest

b. The methods discussed in this section are only applicable for a time series that has a horizontal pattern, so if there is a really a long-term linear trend in the data, the methods discussed in this section are not appropriate

c. The time series plot for the data for years 2002–2008 exhibits a horizontal pattern; it seems reasonable to conclude that the extreme values observed in 1997 and 2001 are more attributable to viewer interest in the performance of Tiger Woods; basing the forecast on years 2002–2008 does seem reasonable, but, because of the injury that Tiger Woods experienced in the 2008 reason, if he is able to play in the 2009 Masters, then the rating for 2009 may be significantly higher than suggested by the data for years 2002–2008

17. a. The time series plot shows a linear trend

b. $\bar{t} = \dfrac{\sum_{t=1}^{n} t}{n} = \dfrac{15}{5} = 3 \quad \bar{Y} = \dfrac{\sum_{t=1}^{n} Y_t}{n} = \dfrac{55}{5} = 11$

$$\Sigma(t - \bar{t})(Y_t - \bar{Y}) = 21 \quad \Sigma(t - \bar{t})^2 = 10$$

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{21}{10} = 2.1$$

$$b_0 = \bar{Y} - b_1\bar{t} = 11 - (2.1)(3) = 4.7$$

$$T_t = 4.7 + 2.1t$$

c. $T_6 = 4.7 + 2.1(6) = 17.3$

18. Forecast for week 6 is 21.16

20. a. The time series plot exhibits a curvilinear trend

b. $T_t = 107.857 - 28.9881t + 2.65476t^2$

c. 45.86

21. a. The time series plot shows a linear trend

b. $\bar{t} = \dfrac{\sum_{t=1}^{n} t}{n} = \dfrac{45}{9} = 5 \quad \bar{Y} = \dfrac{\sum_{t=1}^{n} Y_t}{n} = \dfrac{108}{9} = 12$

$$\Sigma(t - \bar{t})(Y_t - \bar{Y}) = 87.4 \quad \Sigma(t - \bar{t})^2 = 60$$

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{87.4}{60} = 1.4567$$

$$b_0 = \bar{Y} - b_1\bar{t} = 12 - (1.4567)(5) = 4.7165$$

$$T_t = 4.7165 + 1.4567t$$

c. $T_{10} = 4.7165 + 1.4567(10) = 19.28$

22. a. The time series plot shows a downward linear trend

b. $T_t = 13.8 - .7t$

c. 8.2

d. If SCF can continue to decrease the percentage of funds spent on administrative and fund-raising by .7% per year, the forecast of expenses for 2015 is 4.70%

24. a. The time series plot shows a linear trend
b. $T_t = 7.5623 - .07541t$
c. 6.7328
d. Given the uncertainty in global market conditions, making a prediction for December using only time is not recommended

26. a. A linear trend is not appropriate
b. $T_t = 5.702 + 2.889t - 1618t^2$
c. 17.90

28. a. The time series plot shows a horizontal pattern, but there is a seasonal pattern in the data; for instance, in each year the lowest value occurs in quarter 2 and the highest value occurs in quarter 4
b. A portion of the Minitab regression output is shown;

```
The regression equation is
Value = 77.0 - 10.0 Qtr1 - 30.0
        Qtr2 - 20.0 Qtr3
```

c. The quarterly forecasts for next year are as follows:

Quarter 1 forecast = 77.0 − 10.0(1) − 30.0(0) − 20.0(0) = 67
Quarter 2 forecast = 77.0 − 10.0(0) − 30.0(1) − 20.0(0) = 47
Quarter 3 forecast = 77.0 − 10.0(0) − 30.0(0) − 20.0(1) = 57
Quarter 4 forecast = 77.0 − 10.0(0) − 30.0(0) − 20.0(0) = 77

30. a. There appears to be a seasonal pattern in the data and perhaps a moderate upward linear trend
b. A portion of the Minitab regression output follows:

```
The regression equation is
Value = 2492 - 712 Qtr1 - 1512
        Qtr2 + 327 Qtr3
```

c. The quarterly forecasts for next year are as follows:

Quarter 1 forecast is 1780
Quarter 2 forecast is 980
Quarter 3 forecast is 2819
Quarter 4 forecast is 2492

d. A portion of the Minitab regression output follows:

```
The regression equation is
Value = 2307 - 642 Qtr1 - 1465
        Qtr2 + 350 Qtr3 + 23.1 t
```

The quarterly forecasts for next year are as follows:

Quarter 1 forecast is 2058
Quarter 2 forecast is 1258
Quarter 3 forecast is 3096
Quarter 4 forecast is 2769

32. a. The time series plot shows both a linear trend and seasonal effects
b. A portion of the Minitab regression output follows:

```
The regression equation is
Revenue = 70.0 + 10.0 Qtr1 + 105
          Qtr2 + 245 Qtr3
```

Quarter 1 forecast is 80
Quarter 2 forecast is 175
Quarter 3 forecast is 315
Quarter 4 forecast is 70

c. A portion of the Minitab regression output follows

```
The regression equation is
Revenue = -70.1 + 45.0 Qtr1 + 128
          Qtr2 + 257 Qtr3 + 11.7 Period
```

Quarter 1 forecast = is 221
Quarter 2 forecast = is 315
Quarter 3 forecast = is 456
Quarter 4 forecast = is 211

34. a. The time series plot shows seasonal and linear trend effects
b. *Note*: Jan = 1 if January, 0 otherwise; Feb = 1 if February, 0 otherwise; and so on

A portion of the Minitab regression output follows:

```
The regression equation is
Expense = 175 - 18.4 Jan - 3.72 Feb +
          12.7 Mar + 45.7 Apr + 57.1
          May + 135 Jun + 181 Jul + 105
          Aug + 47.6 Sep + 50.6 Oct +
          35.3 Nov + 1.96 Period
```

c. *Note*: The next time period in the time series is Period = 37 (January of Year 4); the forecasts for January–December are 229; 246; 264; 299; 312; 392; 440; 366; 311; 316; 302; 269

35. a. The time series plot indicates a linear trend and a seasonal pattern
b.

Year	Quarter	Time Series Value	Four-Quarter Moving Average	Centered Moving Average
1	1	4		
	2	2		
			3.50	
	3	3		3.750
			4.00	
	4	5		4.125
			4.25	
2	1	6		4.500
			4.75	

Year	Quarter	Time Series Value	Four-Quarter Moving Average	Centered Moving Average
	2	3		5.000
			5.25	
	3	5		5.375
			5.50	
	4	7		5.875
			6.25	
3	1	7		6.375
			6.50	
	2	6		6.625
			6.75	
	3	6		
	4	8		

c.

Year	Quarter	Time Series Value	Centered Moving Average	Seasonal-Irregular Component
1	1	4		
	2	2		
	3	3	3.750	0.800
	4	5	4.125	1.212
2	1	6	4.500	1.333
	2	3	5.000	0.600
	3	5	5.375	0.930
	4	7	5.875	1.191
3	1	7	6.375	1.098
	2	6	6.625	0.906
	3	6		
	4	8		

Quarter	Seasonal-Irregular Values		Seasonal Index	Adjusted Seasonal Index
1	1.333	1.098	1.216	1.205
2	0.600	0.906	0.752	0.746
3	0.800	0.930	0.865	0.857
4	1.212	1.191	1.201	1.191
		Total	4.036	

$$\text{Adjustment for seasonal index} = \frac{4.000}{4.036} = 0.991$$

36. a.

Year	Quarter	Deseasonalized Value
1	1	3.320
	2	2.681
	3	3.501
	4	4.198

Year	Quarter	Deseasonalized Value
2	1	4.979
	2	4.021
	3	5.834
	4	5.877
3	1	5.809
	2	8.043
	3	7.001
	4	6.717

b. Let Period = 1 denote the time series value in Year 1—Quarter 1; Period = 2 denote the time series value in Year 1—Quarter 2; and so on; a portion of the Minitab regression output treating Period as the independent variable and the Deseasonlized Values as the values of the dependent variable follows:

```
The regression equation is
Deseasonalized Value = 2.42 + 0.422
                       Period
```

c. The quarterly deseasonalized trend forecasts for Year 4 (Periods 13, 14, 15, and 16) are as follows:

Forecast for quarter 1 is 7.906
Forecast for quarter 2 is 8.328
Forecast for quarter 3 is 8.750
Forecast for quarter 4 is 9.172

d. Adjusting the quarterly deseasonalized trend forecasts provides the following quarterly estimates:

Forecast for quarter 1 is 9.527
Forecast for quarter 2 is 6.213
Forecast for quarter 3 is 7.499
Forecast for quarter 4 is 10.924

38. a. The time series plot shows a linear trend and seasonal effects

b. 0.71 0.78 0.83 0.97 1.02 1.30 1.50 1.23 0.98 0.99 0.93 0.79

c.

Month	Deaseasonalized Expense
1	239.44
2	230.77
3	246.99
4	237.11
5	235.29
6	242.31
7	240.00
8	235.77
9	244.90
10	242.42
11	247.31

(*Continued*)

Month	Deaseasonalized Expense
12	246.84
13	253.52
14	262.82
15	259.04
16	252.58
17	259.80
18	253.85
19	266.67
20	272.36
21	265.31
22	272.73
23	274.19
24	278.48
25	274.65
26	269.23
27	277.11
28	288.66
29	284.31
30	300.00
31	280.00
32	268.29
33	295.92
34	297.98
35	301.08
36	316.46

d. Let Period = 1 denote the time series value in January -Year 1; Period = 2 denote the time series value in February-Year 2; and so on; a portion of the Minitab regression output treating Period as the independent variable and the Deseasonlized Values as the values of the dependent variable follows:

```
The regression equation is
Deseasonalized Expense = 228 + 1.96
                         Period
```

e.

Month	Monthly Forecast
January	213.37
February	235.93
March	252.69
April	297.21
May	314.53
June	403.42
July	486.42
August	386.52
September	309.88
October	314.98
November	297.71
December	254.44

40. a. The time series plot indicates a seasonal effect; power consumption is lowest in the time period 12–4 A.M., steadily increases to the highest value in the 12–4 P.M. time period, and then decreases again. There may also be some linear trend in the data

b.

Time Period	Adjusted Seasonal Index
12–4 A.M.	0.3256
4–8 A.M.	0.4476
8–12 noon	1.3622
12–4 P.M.	1.6959
4–8 P.M.	1.4578
8–12 midnight	0.7109

c. The following Minitab output shows the results of fitting a linear trend equation to the deseasonalized time series:

```
The regression equation is
Deseasonalized Power = 63108 + 1854 t
```

Deaseasonalized Power (t = 19) = 63,108 + 1854(19) = 98,334

Forecast for 12–4 P.M. = 1.6959(98,334) = 166,764.63 or approximately 166,765 kWh

Deaseasonalized Power (t = 20) = 63,108 + 1854(20) = 100,188

Forecast for 4–8 P.M. = 1.4578(100,188) = 146,054.07 or approximately 146,054 kWh

Thus, the forecast of power consumption from noon to 8 P.M. is 166,765 + 146,054 = 312,819 kWh

42. a. The time series plot indicates a horizontal pattern

b.

$$\text{MSE}(\alpha = .2) = 1.40$$
$$\text{MSE}(\alpha = .3) = 1.27$$
$$\text{MSE}(\alpha = .4) = 1.23$$

A smoothing constant of α = .4 provides the best forecast because it has a smaller MSE

c. 31.00

44. a. There appears to be an increasing trend in the data

b. A portion the Minitab regression output follows (*Note*: t = 1 corresponds to 2001, t = 2 corresponds to 2002, and so on)

```
The regression equation is
Balance($) = 1984 + 146 t
```

The forecast for 2009 (t = 9) is Balance($) = 1984 + 146(9) = $3298

c. A portion of the Minitab regression output follows (*Note*: t = 1 corresponds to 2001, t = 2 corresponds to 2002, and so on)

```
The regression equation is
Balance($) = 2924 - 419 t + 62.7 tsq
```

The forecast for 2009 (t = 9) is Balance ($) = 2924 − 419(9) + 62.7(9)^2 = $4232

d. The quadratic trend equation provides the best forecast accuracy for the historical data
e. Linear trend equation

46. a. The forecast for July is 236.97
Forecast for August, using forecast for July as the actual sales in July, is 236.97
Exponential smoothing provides the same forecast for every period in the future; this is why it is not usually recommended for long-term forecasting
b. Using Minitab's regression procedure we obtained the linear trend equation

$$T_t = 149.72 + 18.451t$$

Forecast for July is 278.88
Forecast for August is 297.33
c. The proposed settlement is not fair since it does not account for the upward trend in sales; based upon trend projection, the settlement should be based on forecasted lost sales of $278,880 in July and $297,330 in August

48. a. The time series plot shows a linear trend
b. $T_t = -5 + 15t$
The slope of 15 indicates that the average increase in sales is 15 pianos per year
c. 85, 100

50. a.

Quarter	Adjusted Seasonal Index
1	1.2717
2	0.6120
3	0.4978
4	1.6185

Note: Adjustment for seasonal index $= \dfrac{4}{3.8985} = 1.0260$
b. The largest effect is in quarter 4; this seems reasonable since retail sales are generally higher during October, November, and December

52. a. Yes, a linear trend pattern appears to be present
b. A portion of the Minitiab regression output follows:

```
The regression equation is
Number Sold = 22.9 + 15.5 Year
```

c. Forecast in year 8 is or approximately 147 units

54. b. The centered moving average values smooth out the time series by removing seasonal effects and some of the random variability; the centered moving average time series shows the trend in the data
c.

Quarter	Adjusted Seasonal Index
1	0.899
2	1.362
3	1.118
4	0.621

d. Hudson Marine experiences the largest seasonal increase in quarter 2; since this quarter occurs prior to the peak summer boating season, this result seems reasonable, but the largest seasonal effect is the seasonal decrease in quarter 4; this is also reasonable because of decreased boating in the fall and winter

Chapter 19

1. $n = 27$ cases with a value different than 150
Normal approximation $\mu = .5n = .5(27) = 13.5$
$\sigma = \sqrt{.25\,n} = \sqrt{.25(27)} = 2.5981$
With the number of plus signs = 22 in the upper tail, use continuity correction factor as follows
$P(x \geq 21.5) = P\left(z \geq \dfrac{21.5 - 13.5}{2.5981}\right) = P(z \geq 3.08)$
p-value = (1.0000 − .9990) = .0010
p-value ≤ .01; reject H_0; conclude population median > 150

2. Dropping the no preference, the binomial probabilities for $n = 9$ and $p = .50$ are as follows

x	Probability	x	Probability
0	0.0020	5	0.2461
1	0.0176	6	0.1641
2	0.0703	7	0.0703
3	0.1641	8	0.0176
4	0.2461	9	0.0020

Number of plus signs = 7
$P(x \geq 7) = P(7) + P(8) + P(9)$
$= .0703 + .0176 + .0020$
$= .0899$
Two-tailed p-value = 2(.0899) = .1798
p-value > .05, do not reject H_0; conclude no indication that a difference exists

4. a. H_0: Median ≥ 15
H_a: Median < 15
b. $n = 9$; number of plus signs = 1
p-value = .0196
Reject H_0; bond mutual funds have lower median

6. $n = 48$; $z = 1.88$
p-value = .0301
Reject H_0; conclude median > $56.2 thousand

8. a. $n = 15$
p-value = .0768
Do not reject H_0; no significant difference for the pace
b. 25%, 68.8%; recommend larger sample

10. $n = 600$; $z = 2.41$
p-value = .0160
Reject H_0; significant difference, *American Idol* preferred

12. H_0: Median for Additive 1 − Median for Additive 2 = 0
H_a: Median for Additive 1 − Median for Additive 2 ≠ 0

Difference	Absolute Difference	Rank	Signed Ranks Negative	Signed Ranks Positive
2.07	2.07	9		9
1.79	1.79	7		7
−0.54	0.54	3	−3	
2.09	2.09	10		10
0.01	0.01	1		1
0.97	0.97	4		4
−1.04	1.04	5	−5	
3.57	3.57	12		12
1.84	1.84	8		8
3.08	3.08	11		11
0.43	0.43	2		2
1.32	1.32	6		6
	Sum of Positive Signed Ranks			$T^+ = 70$

$$\mu_{T^+} = \frac{n(n+1)}{4} = \frac{12(13)}{4} = 39$$

$$\sigma_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{12(13)(25)}{24}} = 12.7475$$

$$P(T^+ \geq 70) = P\left(z \geq \frac{69.5 - 39}{12.7475}\right) = P(z \geq 2.39)$$

p-value = 2(1.0000 − .9916) = .0168
p-value ≤ .05, reject H_0; conclude significant difference between additives

13. H_0: Median time without Relaxant − Median time with Relaxant ≤ 0
H_a: Median time without Relaxant − Median time with Relaxant > 0

Difference	Absolute Difference	Rank	Signed Ranks Negative	Signed Ranks Positive
5	5	9		9
2	2	3		3
10	10	10		10
−3	3	6.5	−6.5	
1	1	1		1
2	2	3		3
−2	2	3	−3	
3	3	6.5		6.5
3	3	6.5		6.5
3	3	6.5		6.5
	Sum of Positive Signed Ranks			$T^+ = 45.5$

$$\mu_{T^+} = \frac{n(n+1)}{4} = \frac{10(11)}{4} = 27.5$$

$$\sigma_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{10(11)(12)}{24}} = 9.8107$$

$$P(T^+ \geq 45.5) = P\left(z \geq \frac{45 - 27.5}{12.7475}\right) = P(z \geq 1.78)$$

p-value = (1.0000 − .9925) = .0375
p-value ≤ .05; reject H_0; conclude without the relaxant has a greater median time

14. $n = 11$; $T^+ = 61$; $z = 2.45$
p-value = .0142
Reject H_0; conclude significant difference; on-time % better in 2006

16. $n = 10$; $T^+ = 12.5$; $z = -1.48$
p-value = .1388
Do not reject H_0; conclude no difference between median scores

18. H_0: The two populations of additives are identical
H_a: The two populations of additives are not identical

Additive 1	Rank	Additive 2	Rank
17.3	2	18.7	8.5
18.4	6	17.8	4
19.1	10	21.3	15
16.7	1	21.0	14
18.2	5	22.1	16
18.6	7	18.7	8.5
17.5	3	19.8	11
		20.7	13
		20.2	12
$W =$	34		

$$\mu_W = \frac{1}{2}n_1(n_1 + n_2 + 1) = \frac{1}{2}7(7 + 9 + 1) = 59.5$$

$$\sigma_W = \sqrt{\frac{1}{12}n_1n_2(n_1 + n_2 + 1)} = \sqrt{\frac{1}{12}7(9)(7 + 9 + 1)}$$
$$= 9.4472$$

With $W = 34$ in lower tail, use the continuity correction

$$P(W \leq 34) = P\left(z \leq \frac{34.5 - 59.5}{9.4472}\right) = P(z \leq -2.65)$$

p-value = 2(.0040) = .0080
p-value < .05; reject H_0; conclude additives are not identical
Additive 2 tends to provide higher miles per gallon

19. a. H_0: The two populations of salaries are identical
H_a: The two populations of salaries are not identical

Public Accountant	Rank	Financial Planner	Rank
50.2	5	49.0	2
58.8	19	49.2	3
56.3	16	53.1	10
58.2	18	55.9	15
54.2	13	51.9	8.5
55.0	14	53.6	11
50.9	6	49.7	4
59.5	20	53.9	12
57.0	17	51.8	7
51.9	8.5	48.9	1
$W =$	136.5		

$$\mu_W = \frac{1}{2}n_1(n_1 + n_2 + 1) = \frac{1}{2}10(10 + 10 + 1) = 105$$

$$\sigma_W = \sqrt{\frac{1}{12}n_1n_2(n_1 + n_2 + 1)} = \sqrt{\frac{1}{12}10(10)(10 + 10 + 1)}$$
$$= 13.2288$$

With $W = 136.5$ in upper tail, use the continuity correction

$$P(W \geq 136.5) = P\left(z \geq \frac{136 - 105}{13.2288}\right) = P(z \geq 2.34)$$

p-value $= 2(1.0000 - .9904) = .0192$

p-value $\leq .05$; reject H_0; conclude populations are not identical

Public accountants tend to have higher salaries

b. Public Accountant $\dfrac{(55.0 + 56.3)}{2} = \$55.65$ thousand

Financial Planner $\dfrac{(51.8 + 51.9)}{2} = \$51.85$ thousand

20. a. \$54,900, \$40,400

b. $W = 69$; $z = 2.04$

p-value $= .0414$

Reject H_0; conclude a difference between salaries; men higher

22. $W = 157$; $z = 2.74$

p-value $= .0062$

Reject H_0; conclude a difference between ratios; Japan tends to be higher

24. $W = 116$; $z = -.22$

p-value $= .8258$

Do not reject H_0; conclude no evidence prices differ

26. H_0: All populations of product ratings are identical

H_a: Not all populations of product ratings are identical

	A	B	C
	4	11	7
	8	14	2
	10	15	1
	3	12	6
	9	13	5
Sum of Ranks	34	65	21

$$H = \left[\frac{12}{15(16)}\left(\frac{34^2}{5} + \frac{65^2}{5} + \frac{21^2}{5}\right)\right] - 3(16) = 10.22$$

χ^2 table with $df = 2$, $\chi^2 = 10.22$; the p-value is between .005 and .01

p-value $\leq .01$; reject H_0; conclude the populations of ratings are not identical

28. H_0: All populations of calories burned are identical

H_a: Not all populations calories burned are identical

	Swimming	Tennis	Cycling
	8	9	5
	4	14	1
	11	13	3
	6	10	7
	12	15	2
Sum of Ranks	41	61	18

$$H = \left[\frac{12}{15(16)}\left(\frac{41^2}{5} + \frac{61^2}{5} + \frac{18^2}{5}\right)\right] - 3(16) = 9.26$$

χ^2 table with $df = 2$, $\chi^2 = 9.26$; the p-value is between .005 and .01

p-value $\leq .05$ reject H_0; conclude that the populations of calories burned are not identical

30. $H = 8.03$ with $df = 3$

p-value is between .025 and .05

Reject H_0; conclude a difference between quality of courses

32. a. $\Sigma d_i^2 = 52$

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} = 1 - \frac{6(52)}{10(99)} = .685$$

b. $$\sigma_{r_s} = \sqrt{\frac{1}{n - 1}} = \sqrt{\frac{1}{9}} = .3333$$

$$z = \frac{r_s - 0}{\sigma_{r_s}} = \frac{.685}{.3333} = 2.05$$

p-value $= 2(1.0000 - .9798) = .0404$

p-value $\leq .05$ reject H_0; conclude significant positive rank correlation

34. $\Sigma d_i^2 = 250$

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} = 1 - \frac{6(250)}{11(120)} = -.136$$

$$\sigma_{r_s} = \sqrt{\frac{1}{n - 1}} = \sqrt{\frac{1}{10}} = .3162$$

$$z = \frac{r_s - 0}{\sigma_{r_s}} = \frac{-.136}{.3162} = -.43$$

p-value $= 2(.3336) = .6672$

p-value $> .05$ do not reject H_0; we cannot conclude that there is a significant relationship

36. $r_s = -.709$, $z = -2.13$

p-value $= .0332$

Reject H_0; conclude a significant negative rank correlation

38. Number of plus signs $= 905$, $z = -3.15$

p-value less than .0020

Reject H_0; conclude a significant difference between the preferences

40. $n = 12$; $T^+ = 6$; $z = -2.55$

p-value $= .0108$

Reject H_0; conclude significant difference between prices

42. $W = 70$; $z = -2.93$

p-value $= .0034$

Reject H_0; conclude populations of weights are not identical

44. $H = 12.61$ with $df = 2$

p-value is less than .005

Reject H_0; conclude the populations of ratings are not identical

46. $r_s = .757$, $z = 2.83$

p-value $= .0046$

Reject H_0; conclude a significant positive rank correlation

Chapter 20

2. a. 5.42
b. UCL = 6.09, LCL = 4.75

4. *R chart:*
UCL = $\bar{R}D_4$ = 1.6(1.864) = 2.98
LCL = $\bar{R}D_3$ = 1.6(.136) = .22
$\bar{x}$ chart:
UCL = $\bar{\bar{x}} + A_2\bar{R}$ = 28.5 + .373(1.6) = 29.10
LCL = $\bar{\bar{x}} - A_2\bar{R}$ = 28.5 − .373(1.6) = 27.90

6. 20.01, .082

8. a. .0470
b. UCL = .0989, LCL = −0.0049 (use LCL = 0)
c. $\bar{p}$ = .08; in control
d. UCL = 14.826, LCL = −0.726 (use LCL = 0)
Process is out of control if more than 14 defective
e. In control with 12 defective
f. *np* chart

10. $f(x) = \dfrac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$

When p = .02, the probability of accepting the lot is

$$f(0) = \frac{25!}{0!(25-0)!}(.02)^0(1-.02)^{25} = .6035$$

When p = .06, the probability of accepting the lot is

$$f(0) = \frac{25!}{0!(25-0)!}(.06)^0(1-.06)^{25} = .2129$$

12. p_0 = .02; producer's risk = .0599
p_0 = .06; producer's risk = .3396
Producer's risk decreases as the acceptance number c is increased

14. $n = 20$, $c = 3$

16. a. 95.4
b. UCL = 96.07, LCL = 94.73
c. No

18.

	***R* Chart**	**$\bar{x}$ Chart**
UCL	4.23	6.57
LCL	0	4.27

Estimate of standard deviation = .86

20.

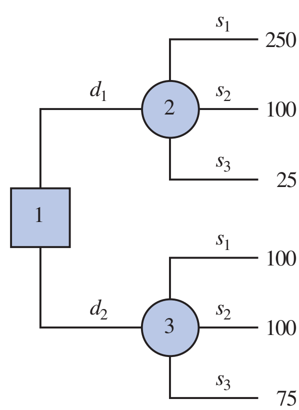
	***R* Chart**	**$\bar{x}$ Chart**
UCL	.1121	3.112
LCL	0	3.051

22. a. UCL = .0817, LCL = −.0017 (use LCL = 0)

24. a. .03
b. β = .0802

Chapter 21

1. a.



b. EV(d_1) = .65(250) + .15(100) + .20(25) = 182.5
EV(d_2) = .65(100) + .15(100) + .20(75) = 95
The optimal decision is d_1

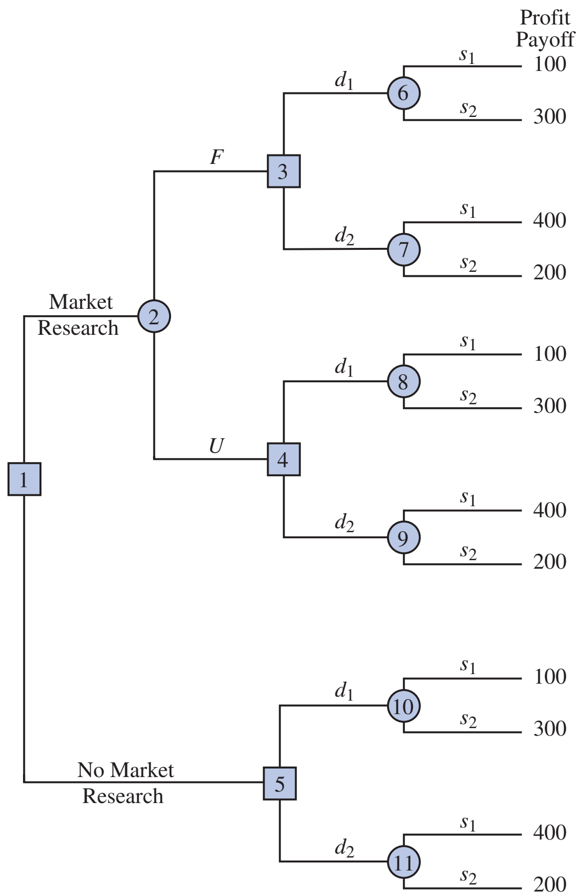
2. a. d_1; EV(d_1) = 11.3
b. d_4; EV(d_4) = 9.5

3. a. EV(own staff) = .2(650) + .5(650) + .3(600)
= 635
EV(outside vendor) = .2(900) + .5(600)
+ .3(300) = 570
EV(combination) = .2(800) + .5(650) + .3(500)
= 635
Optimal decision: hire an outside vendor with an expected cost of $570,000
b. EVwPI = .2(650) + .5(600) + .3(300)
= 520
EVPI = |520 − 570| = 50, or $50,000

4. b. Discount; EV = 565
c. Full Price; EV = 670

6. c. Chardonnay only; EV = 42.5
d. Both grapes; EV = 46.4
e. Both grapes; EV = 39.6

8. a.



b. EV (node 6) = .57(100) + .43(300) = 186
EV (node 7) = .57(400) + .43(200) = 314
EV (node 8) = .18(100) + .82(300) = 264
EV (node 9) = .18(400) + .82(200) = 236
EV (node 10) = .40(100) + .60(300) = 220
EV (node 11) = .40(400) + .60(200) = 280

EV (node 3) = Max(186,314) = 314 d_2
EV (node 4) = Max(264,236) = 264 d_1
EV (node 5) = Max(220,280) = 280 d_2

EV (node 2) = .56(314) + .44(264) = 292
EV (node 1) = Max(292,280) = 292

∴ Market Research
If Favorable, decision d_2
If Unfavorable, decision d_1

10. a. 5000 − 200 − 2000 − 150 = 2650
3000 − 200 − 2000 − 150 = 650

b. Expected values at nodes
8: 2350 5: 2350 9: 1100
6: 1150 10: 2000 7: 2000
4: 1870 3: 2000 2: 1560
1: 1560

c. Cost would have to decrease by at least $130,000

12. b. d_1, 1250
c. 1700
d. If N, d_1
If U, d_2; 1666

14.

State of Nature	$P(s_j)$	$P(I\mid s_j)$	$P(I \cap s_j)$	$P(s_j\mid I)$
s_1	.2	.10	.020	.1905
s_2	.5	.05	.025	.2381
s_3	.3	.20	.060	.5714
	1.0		$P(I)$ = .105	1.0000

16. a. .695, .215, .090
.98, .02
.79, .21
.00, 1.00
c. If C, Expressway
If O, Expressway
If R, Queen City
26.6 minutes

18. a. The Technology Sector provides the maximum expected annual return of 16.97%. Using this recommendation, the minimum annual return is −20.1% and the maximum annual return is 93.1%
b. 15.20%; 1.77%
c. Because the Technology Sector mutual fund shows the greater variation in annual return, it is considered to have more risk
d. This is a judgement recommendation and opinions may vary, but because the investor is described as being conservative, we recommend the lower risk small-cap stock mutual fund

20. a. Optimal Strategy:

Start the R&D project
If it is successful, build the facility

Expected value = $10M million
b. At node 3, payoff for sell rights would have to be $25million or more, in order to recover the $5million R&D cost, the selling price would have to be $30million or more