CHAPTER 1

WHAT IS STATISTICS?

1. a. Interval
   b. Ratio
   c. Ratio
   d. Nominal
   e. Ordinal
   f. Ratio
   g. Nominal
   h. Ordinal
   i. Nominal
   j. Ratio **(LO5)**

2. a. Ratio
   b. Ratio
   c. Ratio
   d. Ratio
   e. Ratio
   f. Ratio **(LO5)**

3. Answers will vary **(LO5)**

4. a. Sample
   b. Population
   c. Population
   d. Sample **(LO2)**

5. Qualitative data are not numerical, whereas quantitative data are numerical. Examples will vary by student. **(LO3)**

6. A population is the entire group which you are studying. A sample is a subset taken from a population. **(LO2)**

7. Nominal, ordinal, interval, and ratio. Examples will vary. **(LO4)**

8. a. A sample is used because it is difficult to locate every student.
   b. A population is employed because the information is easy to find.
   c. A population is used because the information is easy to find.
   d. A sample works because it is difficult to locate every musical. **(LO2)**

9. a. continuous, quantitative, ratio
   b. discrete, qualitative, nominal
   c. discrete, quantitative, ratio
   d. discrete, qualitative, nominal
   e. continuous, quantitative, interval
   f. continuous, quantitative, interval
   g. discrete, qualitative, ordinal
   h. discrete, qualitative, ordinal
   i. discrete, quantitative, ratio **(LO3,4&5)**

10. The cell phone provider is nominal level data. The minutes used are ratio level. Satisfaction is ordinal level. **(LO5)**

11. If you were using the Ridgedale Mall store as typical of all Gap stores then it would be sample data. However, if you were considering it as the only store of interest, then the data would be population data. **(LO2)**

12. Various answers. **(LO5)**

13. Based on these findings, we can infer that 270/300 or 90 percent of the executives would move. **(LO2)**

14. The clear majority of customers tested (400/500, or 80%) believe this take-out service is excellent. Based on these findings, we can expect a similar proportion of all customers to feel the same way. **(LO2)**

15. a 2010 total sales = 1 000 772; 2011 total sales = 942 973; total sales declined about 6% from 2010 to 2011.
    b. Hockey (Men's Finals) and Hockey – Women's experienced losses of 17 and 19 percent, respectively. Meanwhile, Hockey (Women's Finals) gained 9.5 percent and Figure Skating – Juniors about 9 percent. So it would appear that there has been a significant shift within the market from 2010 to 2011. **(LO5)**

16. a. Qualitative: type
    Quantitative: list price, number of bedrooms, full baths and half baths and total square feet.
    b. Type is nominal; the rest are ratio. **(LO3&5)**

17. a. Type is a qualitative variable; dollars is quantitative.
    b. Type is a nominal level variable; dollars is ratio level. **(LO3&5)**

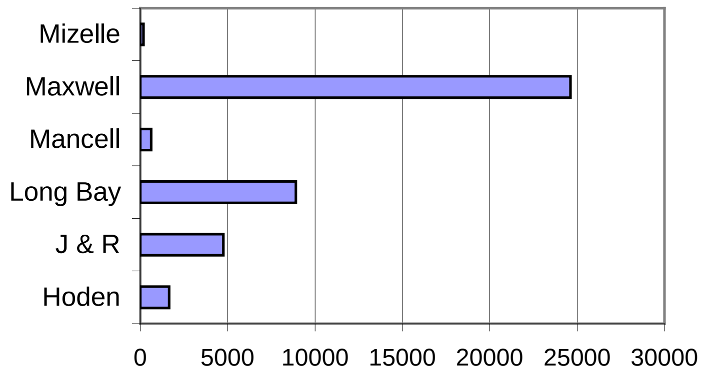
18. a. Women and men are qualitative variables; the others are quantitative.
    b. Year is interval; women and men are nominal; earnings ratio is ratio. **(LO3&5)**

19. a. Region is a qualitative variable; average house price is quantitative.
    b. Region is nominal; average house price is ratio. (LO3&5)

CHAPTER 2

DESCRIBING DATA: FREQUENCY DISTRIBUTIONS AND GRAPHIC PRESENTATION

1. Maxwell Heating & Air Conditioning far exceeds the other corporations in sales. Mancell Electric & Plumbing and Mizelle Roofing & Sheet Metal are the two corporations with the least amount of fourth quarter sales.

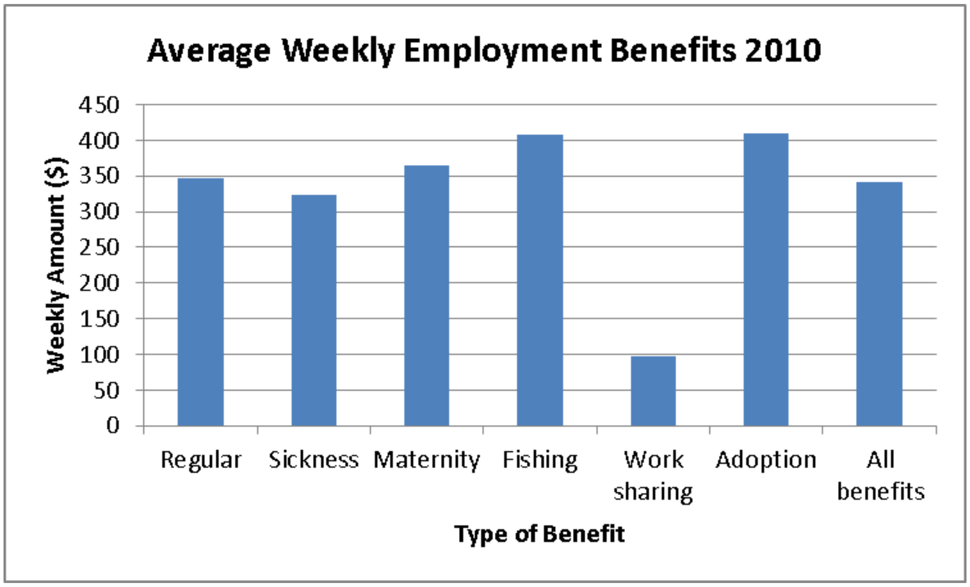


Maxwell has the highest sales, and Mizelle the lowest.

2. Three classes are needed, one for each player. **(LO1)**

3. There are four classes: winter, spring, summer, and fall.
The relative frequencies are 0.1, 0.3, 0.4, and 0.2, respectively. **(LO1)**

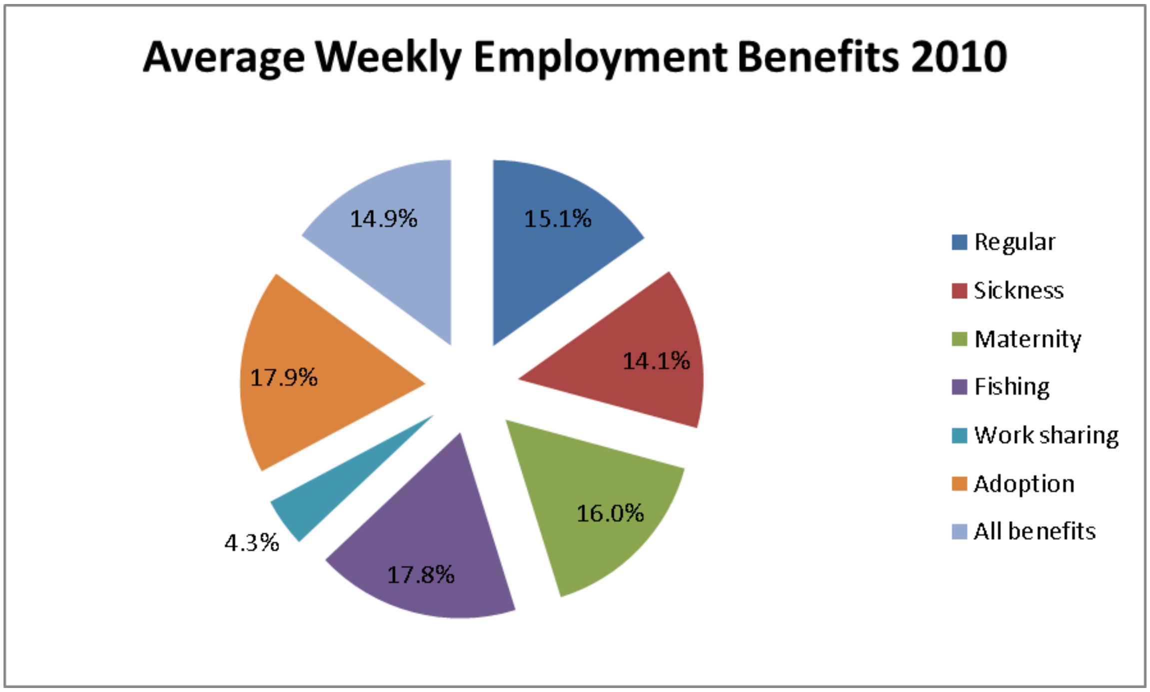
4. a.



b.

Regular	346.94	0.151
Sickness	323.02	0.141
Maternity	365.57	0.160
Fishing	407.34	0.178
Work sharing	97.49	0.043
Adoption	410.03	0.179
All Benefits	341.22	0.149

c.



Average Weekly Employment Benefits 2010
15.1%
14.1%
16.0%
17.8%
4.3%
17.9%
14.9%
Regular
Sickness
Maternity
Fishing
Work sharing
Adoption
All benefits

(LO2)

5. a.

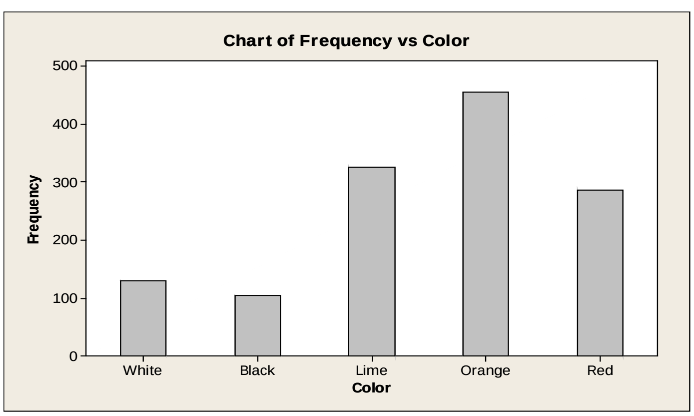
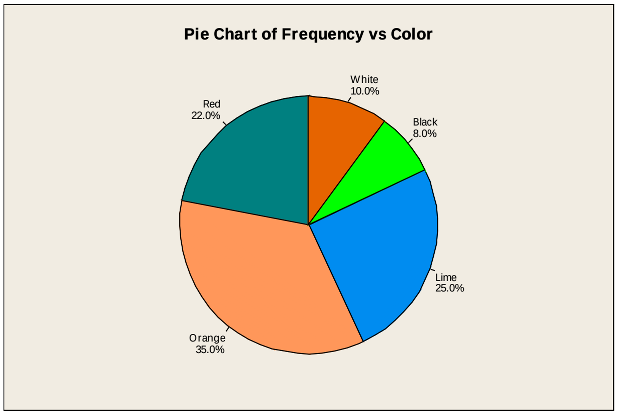


Chart of Frequency vs Color
Frequency
0
100
200
300
400
500
White
Black
Lime
Orange
Red
Color

b.

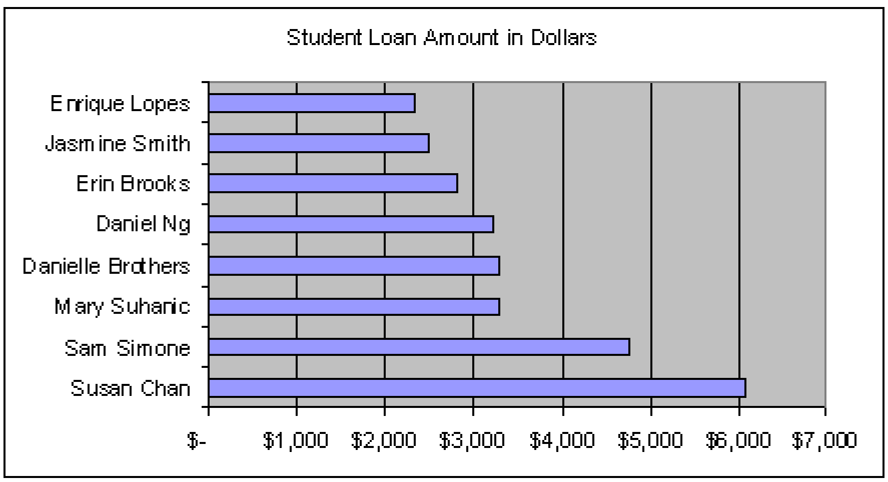
Type	Number	Relative Frequencies
Bright white	130	0.10
Metallic black	104	0.08
Magnetic lime	325	0.25
Tangerine orange	455	0.35
Fusion red	286	0.22
Total	1300	1.00

c.



(LO2)

6. Student loan amounts for 2010. Susan Chan's loan is the largest.



(LO2)

7. $2^5 = 32,\ 2^6 = 64$; therefore 6 classes **(LO3)**

8. $2^5 = 32,\ 2^6 = 64$ suggests 6 classes. $i \geq \frac{\$29 - \$0}{6} = 4.47$ Use interval of 5. **(LO3)**

9. $2^7 = 128,\ 2^8 = 256$ suggests 8 classes $i \geq \frac{567 - 235}{8} = 41.5$ Use interval of 45. **(LO3)**

10. a. $2^5 = 32,\ 2^6 = 64$ suggests 6 classes.
    b. $i \geq \frac{129 - 42}{6} = 14$ Use interval of 15 and start first class at 40. **(LO3)**

11. a. $2^4 = 16$ suggests 5 classes
    b. 6/5 = 1.2 Use interval of 1.5
    c. 24

d.

Units	*f*	**Relative frequency**
24 to under 25.5	2	0.125
25.5 to under 27	4	0.250
27 to under28.5	8	0.500
28.5 to under 30	0	0.000
30 to under 31.5	2	0.125
Total	16	1.000

e. The largest concentration is in the 27 up to 28.5 class (8). **(LO3)**

12. a. $2^4 = 16$, $2^5 = 32$, suggest 5 classes

b. 47/5 = 9.4 Use interval of 10.

c. 50

d.

	f	Relative frequency
50 to under 60	4	0.20
60 to under 70	5	0.25
70 to under 80	6	0.30
80 to under 90	2	0.10
90 to under 100	3	0.15
Total	20	1.00

e. The fewest number is about 50, the highest about 100. The greatest concentration is in classes 60 up to 70 and 70 up to 80. **(LO3)**

13. a.

Shoppers	*f*
0 to under 3	9
3 to under 6	21
6 to under 9	13
9 to under 12	4
12 to under 15	3
15 to under 18	1
Total	51

b. The largest group of shopper's (21) shop at Food Queen 3, 4 or 5 times during a month. Some customers visit the store only 1 time during the month, but others shop as many as 15 times.

c.

Number of Shoppers	*Percent of Total*
0 to under 3	17.65
3 to under 6	41.18
6 to under 9	25.49
9 to under 12	7.84
12 to under 15	5.88
15 to under 18	1.96
Total	100.00

(LO3)

14. a. An interval of 10 is more convenient to work with. The distribution using 10 is:

	f
15 to under 25	1
25 to under 35	2
35 to under 45	5
45 to under 55	10
55 to under 65	15
65 to under 75	4
75 to under 85	3
Total	40

b. Data tends to cluster in classes 45 up to 55 and 55 up to 65.

c. Based on the distribution, the least spent is $15 years (actually $18 from the raw data). The most spent was less than $85. The largest concentration of spending is between $45 up to $65.

d.

$ Spent	*Percent of Total*
15 to under 25	2.5
25 to under 35	5.0
35 to under 45	12.5
45 to under 55	25.0
55 to under 65	37.5
65 to under 75	10.0
75 to under 85	7.5
Total	100.0

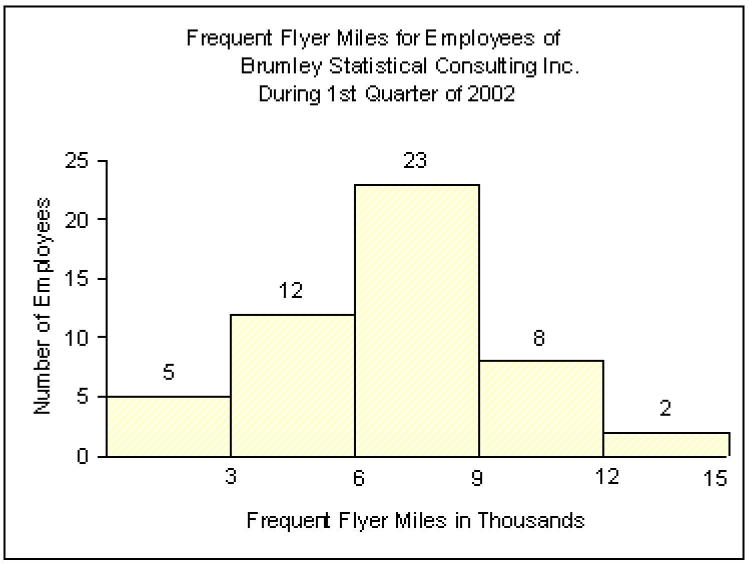
(LO3)

15. a. Histogram
b. 100
c. 5
d. 28
e. 0.28

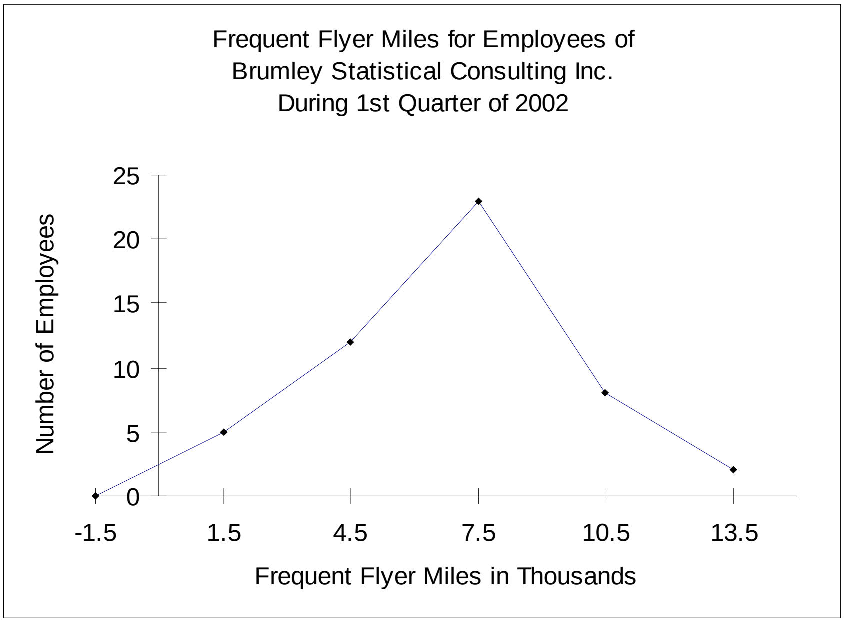
f. 12.5
g. 13 **(LO4)**

16. a. 3
b. approximately 276
c. 76
d. 2
e. frequency polygon **(LO4)**

17. a. 50
b. 1.5 thousands or 1500 miles
c. Using lower limits on the X-axis

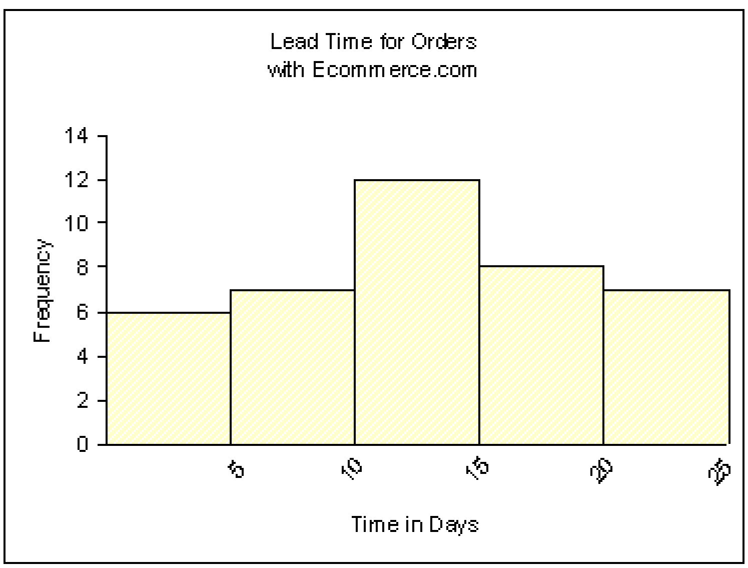


d. 1.5, 5
e.

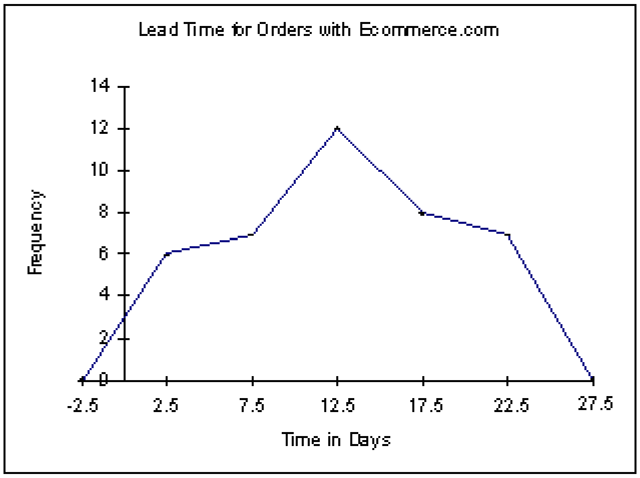


f. Most between 6000-9000, even spread on both sides **(LO4)**

18. a. 40
b. 2.5
c. 2.5, 6 (always draw a frequency polygon using the midpoints)
d.



e.



f. Most orders take around 10-15 days. **(LO4)**

19. a. 40
    b. 5
    c. 11 or 12
    d. about $18 per hour
    e. about $9 per hour
    f. about 75% **(LO4)**

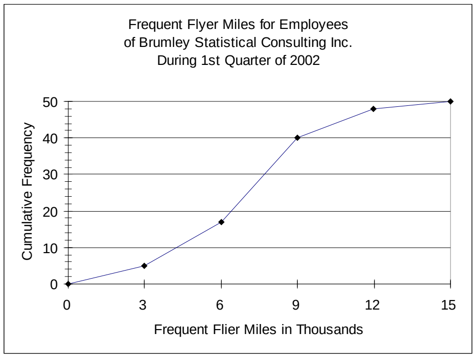
20. a. 200
    b. 50 or $50,000
    c. approximately $175,000
    d. about $240,000
    e. about 60 homes
    f. about 130 homes **(LO4)**

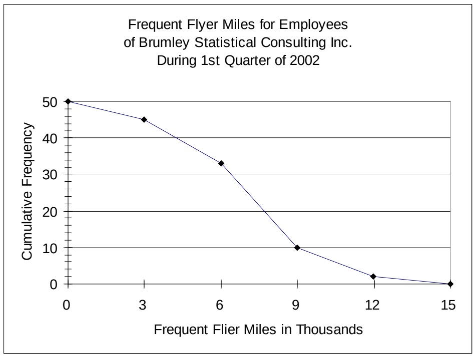
21. a. 5

b.

Miles	f	Cumulative Frequency Less than	Cumulative Frequency More than
0 to under 3	5	5	50
3 to under 6	12	17	45
6 to under 9	23	40	33
9 to under 12	8	48	10
12 to under 15	2	50	2

c.





d. about 8500 miles

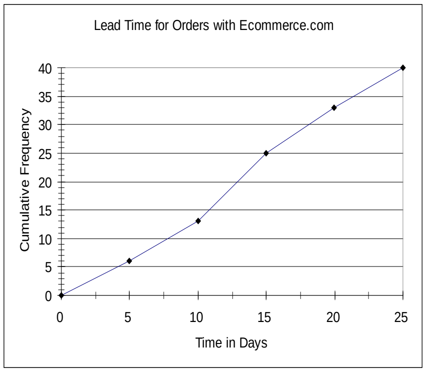
e. about 7500 miles **(LO4)**

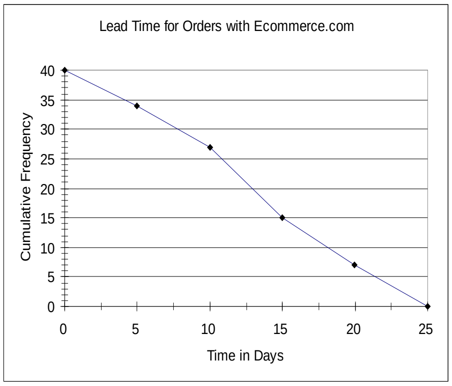
22. a. 13, 25

b.

Lead Time	f	Cumulative Frequency Less than	Cumulative Frequency More than
0 to under 5	6	6	40
5 to under 10	7	13	34
10 to under 15	12	25	27
15 to under 20	8	33	15
20 to under 25	7	40	7

c.





d. About 14 days
e. 27; 15
f. About 18 days **(LO4)**

23. a. 621 to 629
b. 5
c. 621, 623, 623, 627, 629 **(LO5)**

24. a. 210 - 219
b. 6
c. 210, 211, 213, 215, 217, 219 **(LO5)**

25. a. 25
b. 1
c. 38, 106
d. 60, 61, 63, 63, 65, 65, 69
e. No values
f. 9
g. 9
h. 76
i. 16 **(LO5)**

26. a. 50
b. one
c. 126, 270
d. 155, 158, 159
e. No values
f. 13
g. 12
h. 193.5
i. 19 **(LO5)**

27.

Stem	Leaves
0	5
1	28
2	
3	0024789
4	12366
5	2

There were a total of 16 calls studied. The number of calls ranged from 5 to 52 received. Typical was 30-39 calls, smallest was 5, largest was 52 **(LO5)**

28.

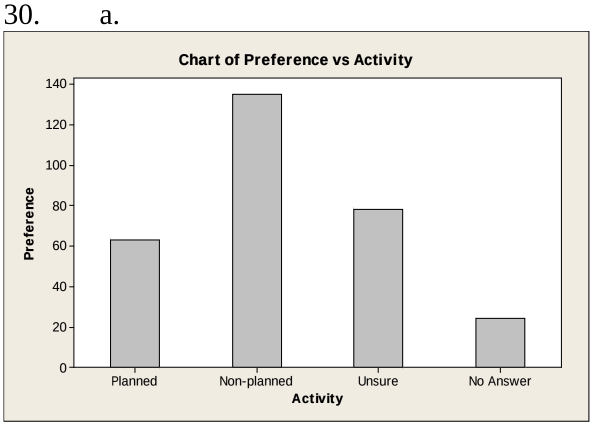
Stem	Leaves
3	6
4	7
5	22499
6	0113458
7	035678
8	0344447
9	055

The daily usage ranged from 36 to 95. In a typical day the ATM is typically used between 52-87 times, smallest was 36, largest was 95; clustered between 52-87 times **(LO5)**

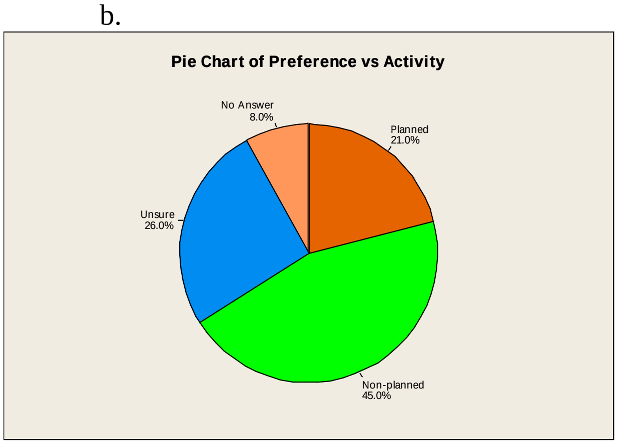
29. a. Qualitative variables are ordinarily nominal level of measurement, but some are ordinal. Quantitative variables are commonly of interval or ratio level of measurement.

b. Yes, both types depict samples and populations. **(LO1&3)**

30. a.



b.



c. Both are readable, but the pie chart may be easier to comprehend. **(LO2)**

31. $2^6 = 64$ and $2^7 = 128$ suggest 7 classes **(LO3)**

32. $2^7 = 128$, $2^8 = 256$ suggests 8 classes. $i \geq \frac{490 - 56}{8} = 54.25$ Use interval of 55. **(LO3)**

33.
a. 5 because $2^4 = 16 < 25$ and $2^5 = 32 > 25$
b. $i \geq \frac{48 - 16}{5} = 6.4$ use interval of 7.
c. 15
d.

Class	*Frequency*
15 to under 22	3
22 to under 29	8
29 to under 36	7
36 to under 43	5
43 to under 50	2
	25

d. The values are clustered between 22 and 36. **(LO3)**

34.
a. 6 because $2^5 = 32 < 45$ and $2^6 = 64 > 45$
b. 90, found by $\frac{570 - 41}{6} = 88.17$
c. 40
d.

Class	*Frequency*
40 to under 130	6
130 to under 220	10
220 to under 310	17
310 to under 400	8
400 to under 490	3
490 to under 580	1
	45

(LO3)

35.
a. 70
b. 1
c. 0, 145
d. 30, 30, 32, 39
e. 24
f. 21
g. 77.5
h. 25 **(LO5)**

36.
a. 55
b. two
c. 91, 237
d. 141, 143, 145
e. 8
f. 12
g. three
h. 180 **(LO5)**

37. a. 56
b. 10 (found by 60 – 50)
c. 55
d. 17 **(LO4)**

38. a. ogive
b. 250
c. 50 (found by 100 – 50)
d. approx $240,000
e. approx $230,000 **(LO4)**

39. a. $36.60, (found by 265 – 82)/5
b. approx $40
c.

$80 to under $120	8
120 to under 160	19
160 to under 200	10
200 to under 240	6
240 to under 280	1
Total	44

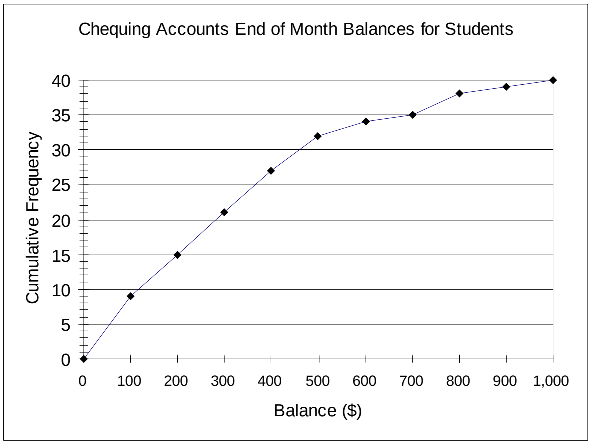
d. The purchases ranged from a low of about $80 to a high of about $280. The concentration is in the $120 to under $160 class. **(LO3)**

40. a. ***Student Chequing Accounts End of Month Balances***

Balance	*f*	*CF*
0 to under 100	9	9
100 to under 200	6	15
200 to under 300	6	21
300 to under 400	6	27
400 to under 500	5	32
500 to under 600	2	34
600 to under 700	1	35
700 to under 800	3	38
800 to under 900	1	39
900 to under 1000	1	40
Total	40	

Probably a class interval of $200 would be better.

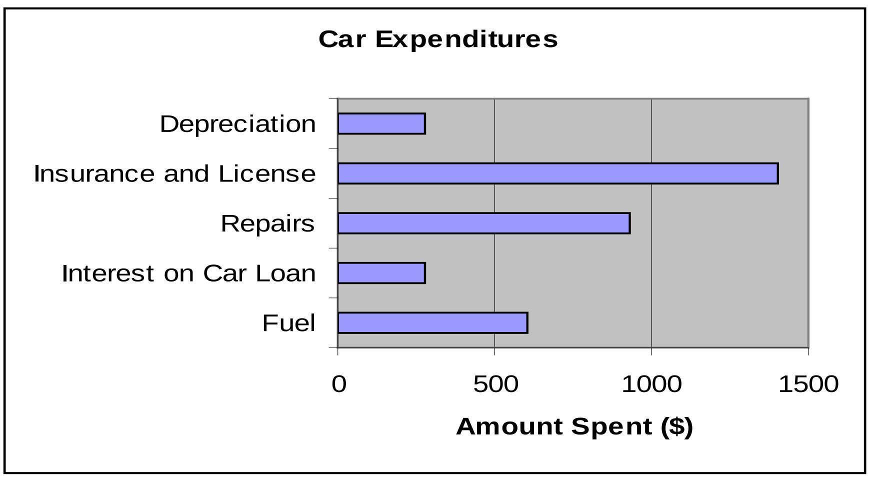
b.



c. About 67% have less than a $400 balance. Therefore, about 33% would be considered "preferred."

d. Approx $50 would be a convenient cutoff point. **(LO4)**

41.



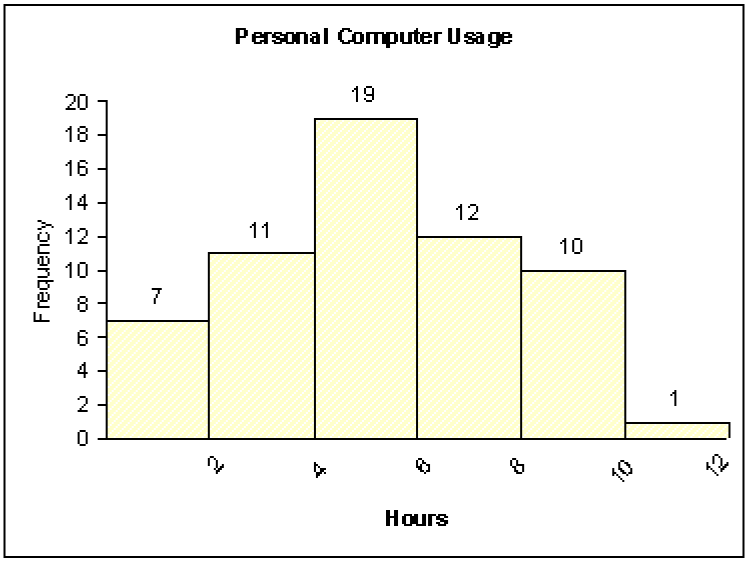
A pie chart is also acceptable. From the graph we can see that insurance and license fees are the highest expense at close to $1500 per year. **(LO2)**

42. a. Since $2^5 = 32 < 60 < 64 = 2^6$, 6 classes are recommended. The interval should be at least (10.1 - 0.4)/6 = 1.6. So we will use 2 as a convenient value.

Personal Computer Usage (Hours)							*cumulative*	
lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*Frequency*	*percent*
0	<	2	1	2	7	11.7	7	11.7
2	<	4	3	2	11	18.3	18	30.0
4	<	6	5	2	19	31.7	37	61.7
6	<	8	7	2	12	20.0	49	81.7
8	<	10	9	2	10	16.7	59	98.3

10	<	12	11	2	1	1.7	60	100.0
					60	100.0		

b.

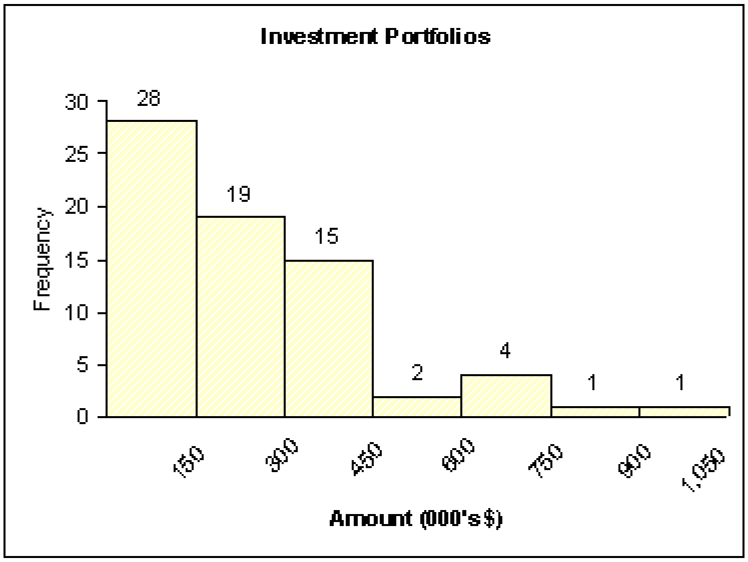


Interpretations will vary. The "typical" person used the computer about 5 hours per week and everyone is within about five hours of that amount. **(LO3&4)**

43. a. Since $2^6 = 64 < 70 < 128 = 2^7$, 7 classes are recommended. The interval should be at least (1002.2 - 3.3)/7 = 142.7; use 150 as a convenient value.

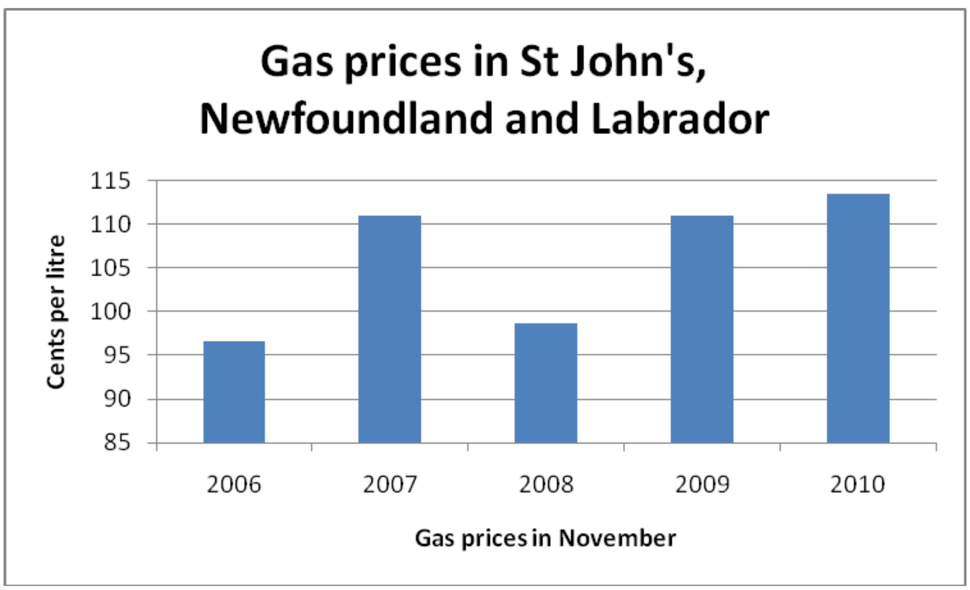
							cumulative	
Investment Portfolios								
lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*frequency*	*percent*
0	<	150	75	150	28	40.0	28	40.0
150	<	300	225	150	19	27.1	47	67.1
300	<	450	375	150	15	21.4	62	88.6
450	<	600	525	150	2	2.9	64	91.4
600	<	750	675	150	4	5.7	68	97.1
750	<	900	825	150	1	1.4	69	98.6
900	<	1,050	975	150	1	1.4	70	100.0
					70	100.0		

b.



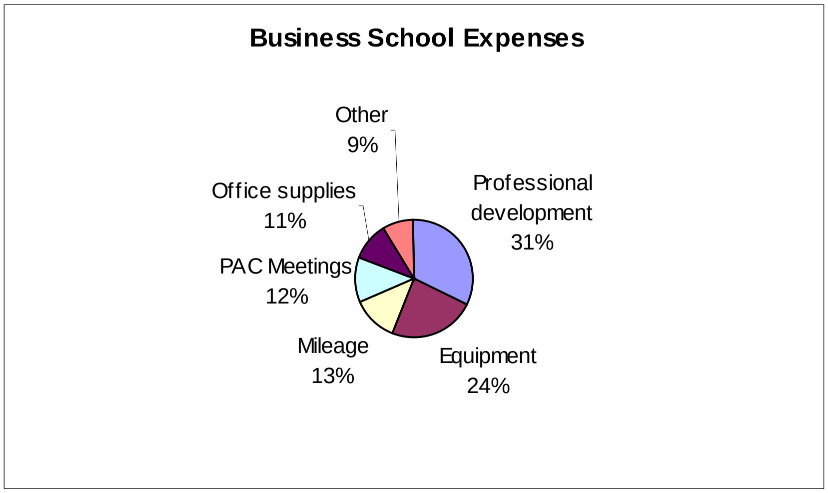
There will be many answers for the interpretation. **(LO3&4)**

44.



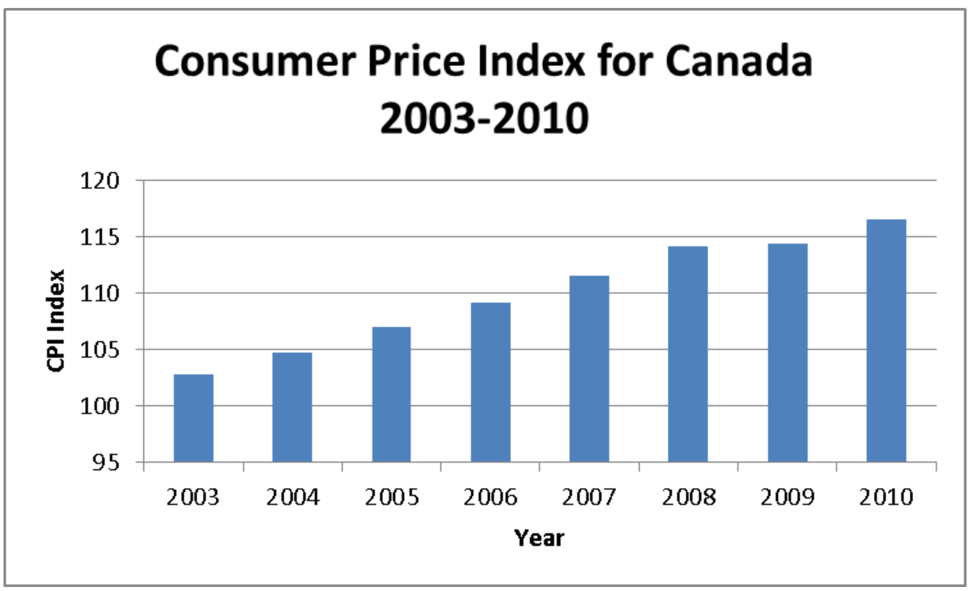
(LO2)

45.



Professional development is the largest expense. **(LO2)**

46.



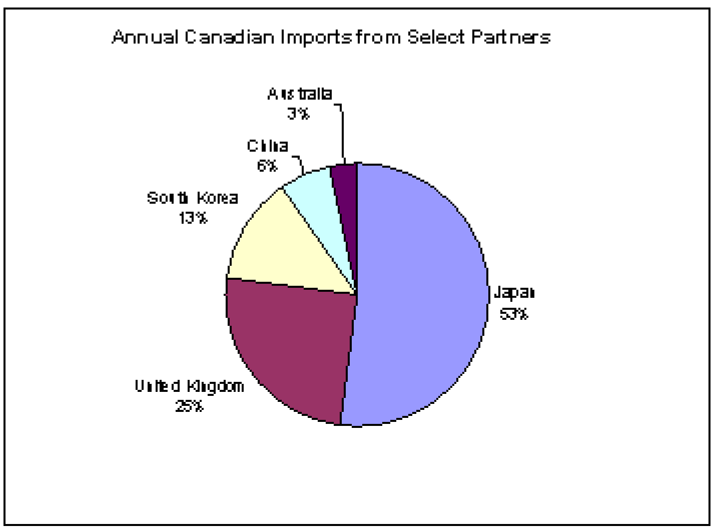
(LO2)

47.

(LO2)

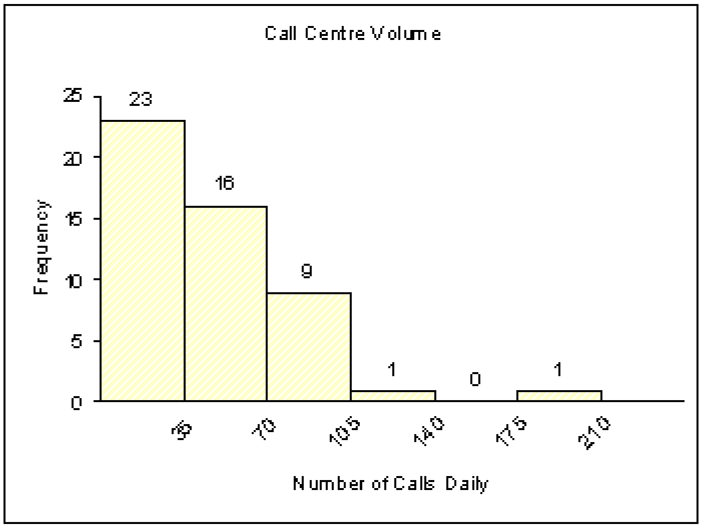
.

48. Japan is the dominant trading partner with over 50% of all imports.



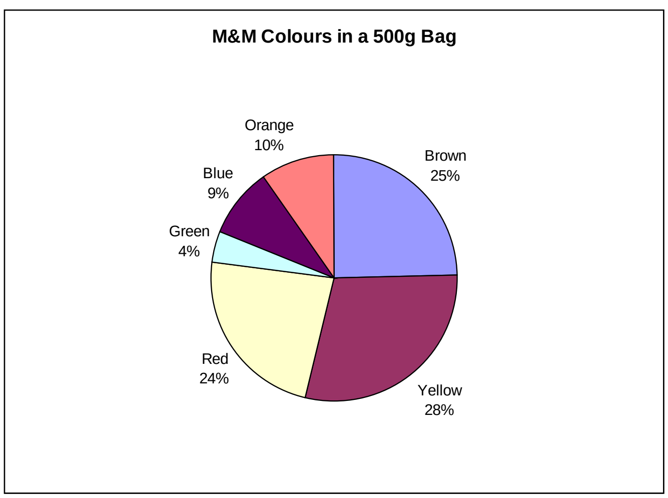
(LO2)

49. There are 50 observations so the recommended number of classes is 6.



Twenty-nine of the 50 days, or 58 percent, have fewer than 40 calls waiting. There are three days that have more than 100 calls waiting. **(LO3&4)**

50. There will be many answers. The following pie chart shows the breakdown of the six colours. About 75% of the candies are either brown, yellow or red. Each of these colours represents about 25% of the total. The percent of orange and blue is less than 10 percent each. About 4 percent of the candies are green.



(LO2)

51. a. 2^5 = 32 <36 < 64 = 2^6. Thus 6 classes are recommended.
b. The interval width should be at least 2, found by (15-3) /6. Use 2.2 for convenience and to ensure there are only 6 classes
c. 2.2
d.

Class	Frequency
2.2 up to 4.6	2
4.6 up to 6.8	7
6.8 up to 9	11
9 up to 11.2	12
11.2 up to 13.4	2
13.4 up to 15.6	2

e. The distribution is slightly right-skewed with the largest concentration in the middle two classes of 6.8 up to 11.2. **(LO3)**

52. **(LO3,4&5)**

Answers will vary depending on how the data is organized. One possible frequency distribution created using MegaStat is:

a. n = 50 use 6 classes interval = (1100000-136900)/6 = 160512

using 200,000 as the class width will give 6 classes

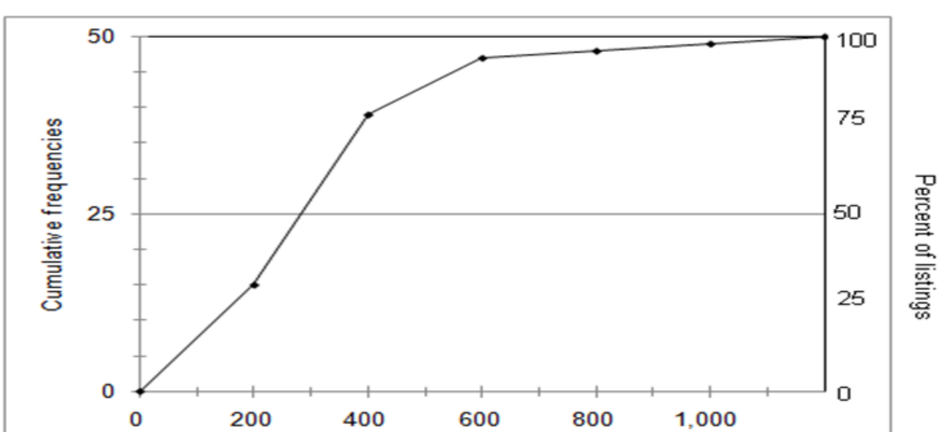
List Price			**frequency**
0	<	200,000	15
200,000	<	400,000	24
400,000	<	600,000	8
600,000	<	800,000	1
800,000	<	1,000,000	1
1,000,000	<	1,200,000	1
		Total	50

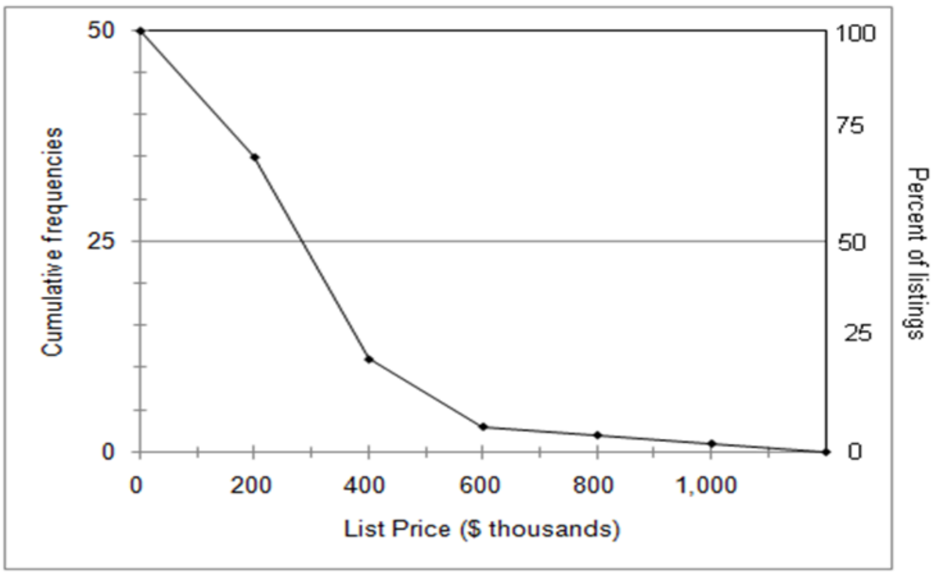
1. the prices are clustered between $200,000 to under $400,000.

2. the smallest list price is between $0 and $200,000 and the largest is more than $1,000,000 but under $1,200,000.

b.

List Price			frequency	less-than	more-than
0	<	200,000	15	15	50
200,000	<	400,000	24	39	35
400,000	<	600,000	8	47	11
600,000	<	800,000	1	48	3
800,000	<	1,000,000	1	49	2
1,000,000	<	1,200,000	1	50	1
		Total	50		





1. about 20
2. about 45%
3. about 78%

c.

min = 567
max = 5150 interval = (5150-567)/6 = 764

using 750 as the class width will give 7 classes and using 1000 will give 6 classes

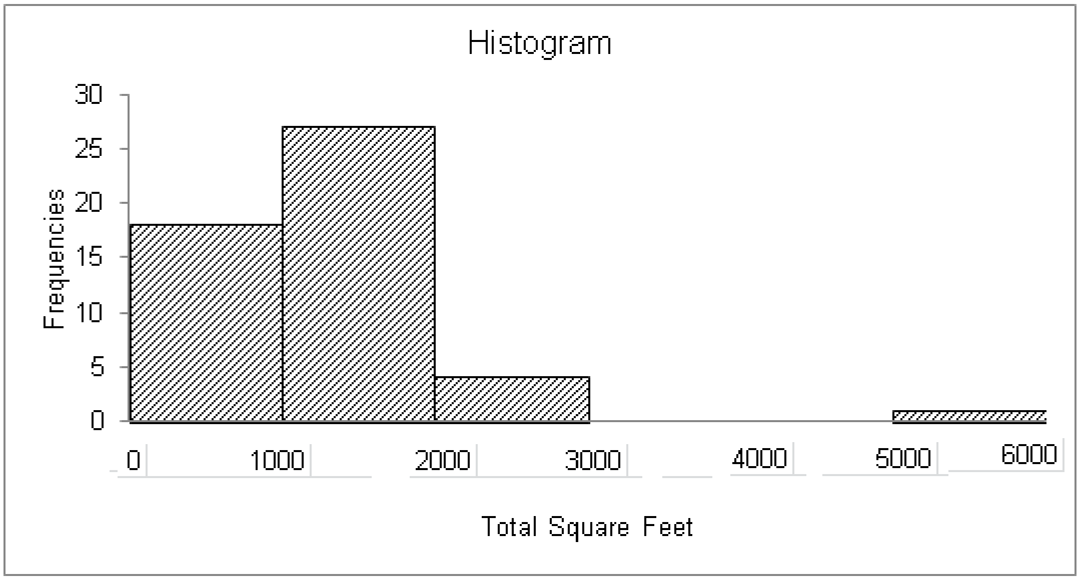
Square Feet			frequency
0	<	1,000	18
1,000	<	2,000	27
2,000	<	3,000	4
3,000	<	4,000	0
4,000	<	5,000	0
5,000	<	6,000	1

Total 50

1. A typical size is from 1000 to under 2000 feet.

2. The homes are clustered from 0 to 2000 square feet with 1

d.



1. A typical size is from 1000 to under 2000 feet.

2. The homes are clustered from 0 to 2000 square feet with 1 home between 5000 to under 6000 square feet.

e.

Stem and Leaf plot for List Price
stem unit = 100000
leaf unit = 10000

Frequency	Stem	Leaf
15	1	3 3 4 4 5 5 5 5 5 6 6 6 6 8 9
10	2	1 1 1 2 2 4 5 7 8 9
14	3	0 0 1 2 2 2 3 3 4 4 6 6 7 9
2	4	1 6
6	5	0 0 4 7 9 9
1	6	6
0	7	
1	8	9
0	9	
0	10	
1	11	0

50

1. The list prices are clustered between 130000 and 390000.
2. There is one value at 1 100 000 which is much higher than the rest of the values. This value may be an outlier.
3. Answers will vary but should contain the above information.

f.

Stem and Leaf plot for Total Square Feet
stem unit = 1000
leaf unit = 100

Frequency	Stem	Leaf
18	0	5 5 5 7 7 7 8 8 8 8 8 8 8 9 9 9 9 9
27	1	0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 5 6 7 9
4	2	3 3 3 6
0	3	
0	4	
1	5	1
50		

1. The square footage is clustered between 500 and 2600.
2. There is one large value at 5100. This may be an outlier.
3. Answers will vary but should contain the above information.

53. **(LO3,4&5)**

a.

min = 500
max = 5200 interval = (5200-500)/7 = 671

use 750 as the interval

Total Square Feet

lower		*upper*	frequency
0	<	750	11
750	<	1,500	48
1,500	<	2,250	18
2,250	<	3,000	8
3,000	<	3,750	6
3,750	<	4,500	2
4,500	<	5,250	5

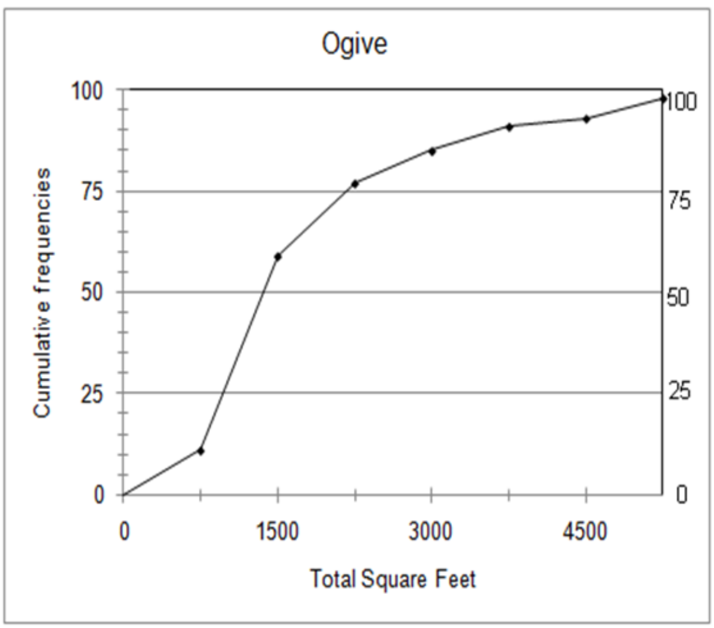
1. A typical size is from 750 to 1500. The range of the data is from about 0 to under 5250.

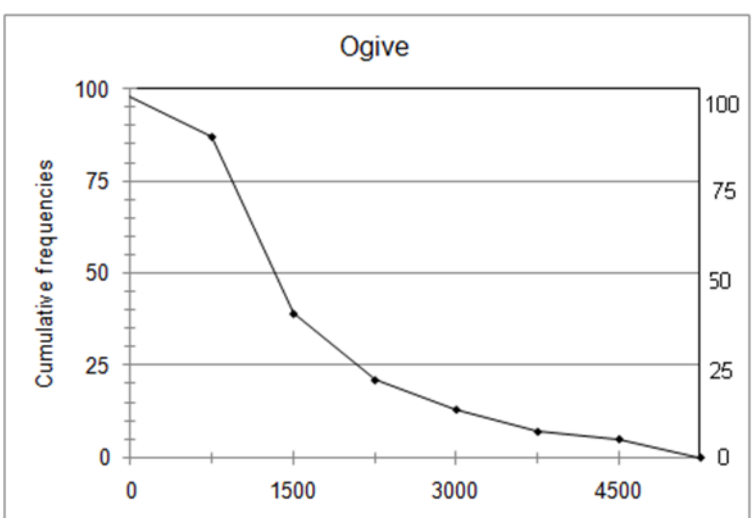
2. There are 7 values between 3750 and 5250 square feet. These values are much larger than

the typical number of square feet.

b.

Total Square Feet					
lower		*upper*	frequency	less-than	more-than
0	<	750	11	11	98
750	<	1,500	48	59	87
1,500	<	2,250	18	77	39
2,250	<	3,000	8	85	21
3,000	<	3,750	6	91	13
3,750	<	4,500	2	93	7
4,500	<	5,250	5	98	5





1. about 1200 square feet

2. about 75

3. about 75%

c.

Stem and Leaf plot for	List Price
stem unit =	1000000
leaf unit =	100000

Frequency	Stem	Leaf
88	0	1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4
9	0	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 6 6 6 6 6 7 7 7 7 7 7 8 9 9 9 9 9
<u>1</u>	1	1 3 3 3 4 5 5 5 8
98	2	4

1. The values are clustered between \$100 000 and \$900 000.

2. There are 10 homes listed for more than \$900 000 and one home listed for more than \$2 400 000. Is is possible that one or more of these may be outliers.

3. Answers will vary but should contain the above information.

d.

Stem and Leaf plot for	Total Square Feet
stem unit =	1000
leaf unit =	100

Frequency	Stem	Leaf
19	0	5 5 5 5 6 6 7 7 7 7 7 7 7 8 8 8 8 9 9
53	1	0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 2 2 2 2 2
13	1	3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 5 5 6 6 6 6 6 6 6 7 8 9 9
6	2	1 1 1 1 1 2 2 2 3 4 4 5 8
4	3	0 2 2 3 5 6
<u>3</u>	4	0 3 7 8
98	5	0 2 2

1. The values are clustered between 1100 and 1900.

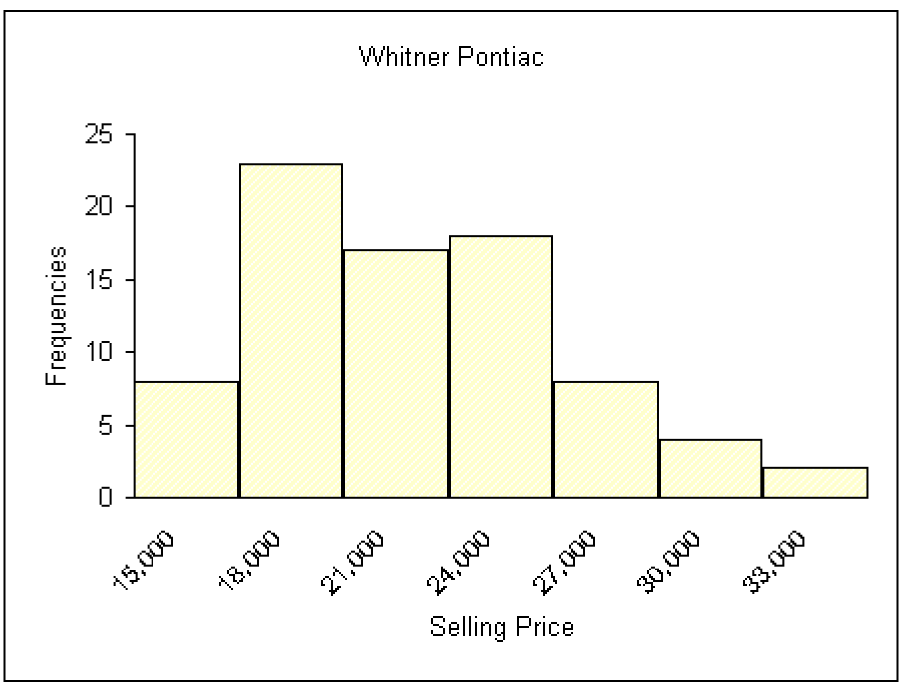
2. There are 7 homes with square footage more than 4000 sq ft and 3 of them with more than 5000 sq ft. It possible that one or more of these may be outliers.

3. Answers will vary but should contain the above information.

CASE (LO3&4)

Answers may vary.

Price							*cumulative*	
lower		*upper*	*Midpoint*	*width*	*frequency*	*percent*	*frequency*	*percent*
15,000	<	18,000	16,500	3,000	8	10.0	8	10.0
18,000	<	21,000	19,500	3,000	23	28.8	31	38.8
21,000	<	24,000	22,500	3,000	17	21.3	48	60.0
24,000	<	27,000	25,500	3,000	18	22.5	66	82.5
27,000	<	30,000	28,500	3,000	8	10.0	74	92.5
30,000	<	33,000	31,500	3,000	4	5.0	78	97.5
33,000	<	36,000	34,500	3,000	2	2.5	80	100.0
					80	100.0		



The selling prices range from about $15 000 to about $36 000. The selling prices are concentrated between $18 000 and $27 000. A total of 58, or 72,5%, of the vehicles sold within this range. The highest frequency is in the $18 000 to under $21 000 class. So we say that a typical selling price is $19 500. Six vehicles sold for less than $18 000, and two sold for more than $33 000.

CHAPTER 3

DESCRIBING DATA: NUMERICAL MEASURES

1. $\mu = 5.4$ found by 27/5 **(LO1)**

2. $\mu = 5.5$ found by 33/6 **(LO1)**

3. a. Mean = 7.0, found by 28/4
   b. $(5-7)+(9-7)+(4-7)+(10-7)=0$ **(LO1)**

4. a. 4.2 found by 21/5
   b. $(1.3-4.2)+(7.0-4.2)+(3.6-4.2)+(4.1-4.2)+(5.0-4.2)=0$ **(LO1)**

5. 14.58, found by 43.74/3 **(LO1)**

6. $20.95, found by $125.68/6 **(LO1)**

7. a. 15.4, found by 154/10 **(LO1)**
   b. Population parameter since it includes all the salespersons at Midtown Ford.

8. a. 23.9, found by 167/7 **(LO1)**
   b. Population parameter since it includes all the calls during a seven-day period.

9. a. $54.55, found by $1091/20 **(LO1)**
   b. A sample statistic, assuming that the power company serves more than 20 customers.

10. a. 10.73, found by 161/15 **(LO1)**
    b. Statistic

11. Yes, $162,900 found by 30($5430) **(LO1)**

12. Veteran sales people likely are above average and new recruits are below average. So, the sales goal is impractical and leads to more exits. **(LO1)**

13. $22.91, found by $\dfrac{300(\$20)+400(\$25)+400(\$23)}{300+400+400}=\dfrac{\$25,200}{1100}$ **(LO1)**

14. $1.50 found by ($40 + $35)/50 **(LO1)**

15. $23.00, found by ($800 + $1000 + $2800)/200 **(LO1)**

16. $143.75, found by ($1000 +$750 + $4000)/40 **(LO1)**

17. a. no mode
    b. The given value would be the mode
    c. 3 and 4; bimodal **(LO1)**

18. Median = 33, Mode = 15 **(LO1)**

19. Median = 5 Mode = 5 **(LO1)**

20. Median = 10.5 Mode = 8 **(LO1)**

21. a. Median = 63.8%
b. Mode = no mode **(LO1)**

22. Median = 9.2 Modes are 8.2, 8.5, and 10.3 **(LO1)**

23. Mean = 58.82; Median = 58.00; Mode = 58.00. All three measures are nearly identical. **(LO1)**

24. Mean = 94.5; Median = 99.5; Applicants are not better than the general population. **(LO1)**

25. a. 6.72, found by 80.6/12
b. 6.6 is both the median and the mode
c. positively skewed; the mean is the highest value **(LO1)**

26. Treat Wind Direction as nominal, Temperature as interval, and Pavement as ordinal data. That would lead to a mode (Southwest) for the first column, a mean (91) for the second column, and a median (Trace) for a third. **(LO1)**

27. 12.8% increase found by $\sqrt[5]{(1.08)(1.12)(1.14)(1.26)(1.05)} - 1$ **(LO1)**

28. 6% increase found by $\sqrt[8]{(1.02)(1.08)(1.06)(1.04)(1.10)(1.06)(1.08)(1.04)} - 1$ **(LO1)**

29. 12.28% increase found by $\sqrt[5]{(1.094)(1.138)(1.117)(1.119)(1.147)} - 1$ **(LO1)**

30. 2.10% increase found by $\sqrt[10]{\frac{114.4}{92.9}} - 1$ **(LO1)**

31. 1.14% increase found by $\sqrt[39]{\frac{34108800}{21961999}} - 1$ **(LO1)**

32. 3.06% increase found by $\sqrt[20]{\frac{117.1}{64.1}} - 1$ **(LO1)**

33. 10.76% found by $\sqrt[5]{\frac{70}{42}} - 1$ **(LO1)**

34. 5.85% increase found by $\sqrt[11]{\frac{14701}{7863}} - 1$ **(LO1)**

35. a. 7, found by 10 – 3 **(LO4)**
b. 6, found by 30/5
c. 2.4, found by 12/5

d. The difference between the highest number sold (10) and the smallest number sold (3) is 7. On the average the number of service reps on duty deviates by 2.4 from the mean of 6.

36. a. 24, found by 52 – 28 **(LO4)**
b. 38
c. 6.25, found by 50/8
d. The difference between 28 and 52 is 24. On the average the number of students enrolled deviates 6.25 from the mean of 38.

37. a. 30, found by 54 – 24 **(LO4)**
b. 38, found by 380/10
c. 7.2, found by 72/10
d. The difference between 54 and 24 is 30. On the average the number of minutes required to install a door deviates 7.2 minutes from the mean of 38 minutes.

38. a. 7.6%, found by 18.2 – 10.6 **(LO4)**
b. 13.85% found by 110.8/8
c. 2%, found by 16/8
d. The difference between 18.2 and 10.6 is 7.6%. On the average the return on investment deviates two percent from the mean of 13.85%.

39. British Columbia: Median = 34, Mean = 33.1 Mode = 34 and Range = 32
Manitoba: Median = 25, Mean = 24.5, Mode = 25, and Range = 19
In BC, there was a greater average preference for the pizza than in Manitoba, however BC also had a greater dispersion in preference. **(LO4)**

40. Sales employees: Mean = 3.5, Median = 2.5, Mode = 2 and Range = 10
Warehouse employees: Mean = 1.125, Median = 0.5, Mode = 0 and Range = 5
Sales employees are more likely on average to experience days off due to illness than warehouse employees, however, there is greater dispersion amongst days lost with sales employees. **(LO4)**

41. a. 5
b. 4.4 found by $\dfrac{(8-5)^2+(3-5)^2+(7-5)^2+(3-5)^2+(4-5)^2}{5}$ **(LO1&4)**

42. a. 8 **(LO1&4)**
b. 9.67 found by $\dfrac{(13-8)^2+(3-8)^2+(8-8)^2+(10-8)^2+(8-8)^2+(6-8)^2}{6}$

43. a. \$2.77 **(LO1&4)**
b. $s^2=\dfrac{(2.68-2.77)^2+\ldots+(4.30-2.77)^2+(3.58-2.77)^2}{5}=1.26$

44. a. 11.76%, found by 58.8/5
b. 16.89, found by 84.452/5 **(LO1&4)**

45. a. Range = 7.3, found by 11.6 – 4.3 **(LO1&4)**
Arithmetic mean = 6.94, found by 34.7/5
Variance = 6.5944, found by 32.972/5 Standard Deviation = 2.568

b. Dennis has a higher mean return (11.76 > 6.94). However, Dennis has greater spread in their returns on equity (16.89 > 6.59).

46. a. \$18,000, found by \$140 000 – 122 000 **(LO1&4)**
b. \$129,600, found by \$648,000/5
c. Variance = 40,240,000, found by 201,200,000/5 Standard Deviation = \$6343.50
d. Means about the same, but less dispersion in salary for TMV vice presidents.

47. a. $\overline{X} = 4$ $s^2 = \frac{(7-4)^2 + \ldots + (3-4)^2}{5-1} = 5.5$

b. $s^2 = \frac{102 - \frac{(20)^2}{5}}{5-1} = 5.50$

c. $s = 2.3452$ **(LO4)**

48. a. $\overline{X} = 8$ $s^2 = \frac{(11-8)^2 + \ldots + (7-8)^2}{5-1} = 5.5$

b. $s^2 = \frac{342 - \frac{(40)^2}{5}}{5-1} = 5.50$

c. $s = 2.3452$ **(LO4)**

49. a. $\overline{X} = 38$ $s^2 = \frac{(28-38)^2 + \ldots + (42-38)^2}{10-1} = 82.6667$

b. $s^2 = \frac{15{,}184 - \frac{(380)^2}{10}}{10-1} = 82.6667$

c. $s = 9.0921$ **(LO4)**

50. a. $\overline{X} = 13.85$ $s^2 = \frac{(10.6-13.85)^2 + \ldots + (15.6-13.85)^2}{8-1} = 6.0086$

b. $s^2 = \frac{1576.64 - \frac{(110.8)^2}{8}}{8-1} = 6.0086$

c. $s = 2.4512$ **(LO4)**

51. a. $\overline{X} = 95.1$ $s^2 = \frac{(101-95.1)^2 + (97-95.1)^2 + \cdots + (88-95.1)^2}{10-1} = 123.66$

b. $s^2 = \sqrt{\frac{91553 - \frac{(951)^2}{10}}{10-1}} = 123.66$

c. 11.12 **(LO4)**

52. a. $\overline{X} = 104.1$ $s^2 = \frac{(110-104.1)^2 + (126-104.1)^2 + \cdots + (100-104.1)^2}{10-1} = 120.78$

b. $s^2 = \sqrt{\dfrac{109455 - \dfrac{(1041)^2}{10}}{10 - 1}} = 120.77$

c. 10.99 **(LO4)**

53. $1 - 1/(1.8)^2 = .69135 = 69.1\%$ **(LO6)**

54. $1 - 1/(2.5)^2 = .84 = 84\%$ **(LO6)**

55. a. About 95 %
b. 47.5%, 2.5% **(LO6)**

56. a. 85, halfway between the endpoints of 140 and 30
b. About 18, found by (140 – 30)/6
c. 103 and 67, found by $85 \pm (1)18$
d. 121 and 49, found by $85 \pm (2)18$ **(LO6)**

57. 8.06%, found by (0.25/3.10)(100) **(LO7)**

58. Domestic 23.81%, found by (5/21)(100). Overseas 20%, found by (7/35)(100). There is slightly more relative dispersion in the weights of luggage for domestic passengers. **(LO7)**

59. a. Because the two series are in different units of measurement.
b. P.E. ratio = 36.73%; ROI 52%, less spread in the P.E. ratios **(LO7)**

60. a. The data are in the same units but the means, relatively speaking, are far apart.
b. The relative dispersion in stocks under $10 is 28.95%. For stocks over $60, 5.71%. Less relative dispersion in stocks over $60. **(LO7)**

61. a. The mean is 30.8, found by 154/5. The median is 31.0 and the standard deviation is 3.96, found as $\sqrt{\dfrac{4806 - \dfrac{154^2}{5}}{4}}$.
b. - 0.15, found by $\dfrac{3(30.8 - 31.0)}{3.96}$. **(LO1, 4&7)**

62. a. The mean is 542, found by 8130/15. The median is 546 and the standard deviation is 25.08, found as $\sqrt{\dfrac{4{,}415{,}268 - \dfrac{8130^2}{15}}{14}}$. **(LO1, 4&7)**
b. –0.478, found by $\dfrac{3(542 - 546)}{25.08}$

63. a. The mean is 21.93, found by 328.9 / 15. The median is 15.8 and the standard deviation is 21.18, found as $\sqrt{\dfrac{13{,}494.676 - \dfrac{328.9^2}{15}}{14}}$.

b. 0.868, found by $\dfrac{3(21.93 - 15.8)}{21.18}$. **(LO1, 4&7)**

64. a. The mean is 5961.82, found by 166931/28. The median is 1375 and the standard deviation is 7148.75, found as $\sqrt{\dfrac{2375037105 - \dfrac{166931^2}{28}}{27}}$.

b. 1.925, found by $\dfrac{3(5961.82 - 1375)}{7148.75}$ **(LO1, 4&7)**

65. Median = 53 found by (11 + 1)(1/2); therefore 6th value in from lowest.
Q_1 = 49 found by (11 + 1)(1/4); therefore 3rd value in from lowest
Q_3 = 55 found by (11 + 1)(3/4); therefore 9th value in from lowest **(LO8)**

66. Median = 9.53, found by (9.45 + 9.61)/2
Q_1 = 7.69 found by 7.59 + (7.99 – 7.59) ¼
Q_3 = 12.59 found by 12.22 + (12.71 – 12.22)3/4 **(LO8)**

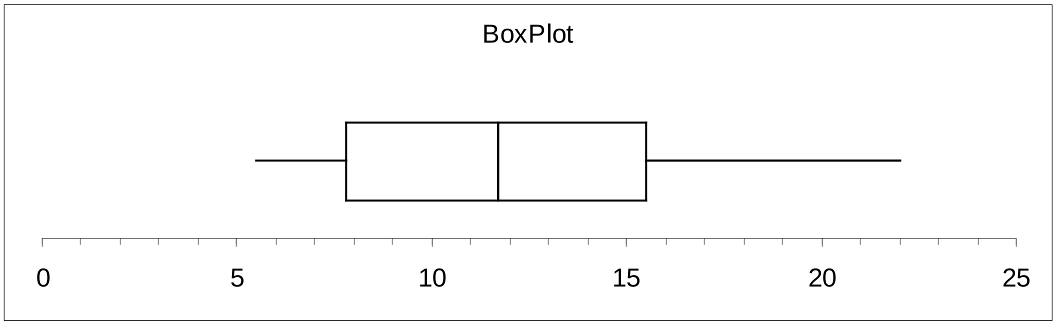
67. a. Q_1 = 33.25 Q_3 = 50.25
b. D_2 = 27.8 D_8 = 52.6
c. P_{67} = 47 **(LO8)**

68. a. Median = 58
b. Q_1 = 51.25 Q_3 = 66.0 (Megastat shows Q_1 as 51.75)
c. D_1 = 45.3 D_9 = 76.4
d. P_{33} = 53.53 **(LO8)**

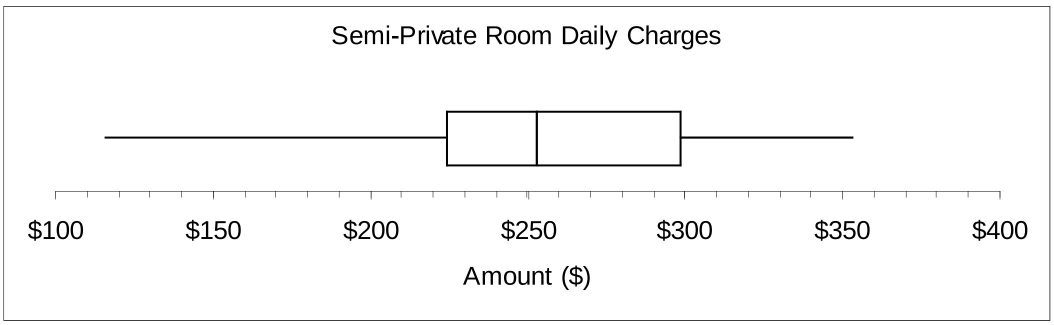
69. a. 350
b. Q_1 = 175 Q_3 = 930
c. 755, found by 930 – 175
d. Less than zero, or more than about 2060
e. There are no outliers
f. The distribution is positively skewed **(LO3&8)**

70. a. 450
b. Q_1 = 300 Q_3 = 700
c. 400, found by 700 – 300
d. Less than zero or more than 1300
e. One outlier at about 1500
f. Distribution is positively skewed **(LO3&8)**

71. The distribution is somewhat positively skewed. Note that the line above 15.5 is longer than below 7.8. **(LO3&9)**



72. The median is \$253, the smallest value is \$116 and the largest is \$353. About 25% of the semi-private rooms are less than \$214 and 25% above \$304.25. The distribution is negatively skewed.



(LO3&9)

73. Because the exact values in a frequency distribution are not known, the midpoint of the class is used for every member of that class. **(LO1)**

74.

Class	*f*	*M*	*fM*	*fM*
0 to under 5	2	2.5	5.00	12.50
5 to under 10	7	7.5	52.50	393.75
10 to under 15	12	12.5	150.00	1875.00
15 to under 20	6	17.5	105.00	1837.50
20 to under 25	3	22.5	67.50	1518.75
Total	30		380.00	5637.50

$\bar{X} = \frac{380}{30} = 12.67$ $\quad$ $s = \sqrt{\frac{5637.5 - (380)^2/30}{29}} = 5.33$

Median = $10 + \frac{\frac{30}{2} - 9}{12}(5) = 12.5$ **(LO1)**

75.

Class	*f*	*X*	*M*	*fM*^{2}
20 to under 30	7	25	175	4375
30 to under 40	12	35	420	14,700
40 to under 50	21	45	945	42,525
50 to under 60	18	55	990	54,450
60 to under 70	12	65	780	50,700
Total	70		3310	166,750

$$\bar{X} = \frac{3310}{70} = 47.2857 \qquad s = \sqrt{\frac{166,750 - (3310)^2/70}{69}} = 12.179$$

$$\text{Median} = 40 + \frac{\frac{70}{2} - 19}{21}(10) = 47.6$$

(LO1)

76.

Age	*f*	*M*	*fM*	*f*
10 to under 20	3	15	45	675
20 to under 30	7	25	175	4375
30 to under 40	18	35	630	22,050
40 to under 50	20	45	900	40,500
50 to under 60	12	55	660	36,300
Total	60		2410	103,900

$$\bar{X} = \frac{2410}{60} = 40.17 \qquad s = \sqrt{\frac{103,900 - (2410)^2/60}{60 - 1}} = 10.97$$

$$\text{Median} = 40 + \frac{\frac{60}{2} - 28}{20}(10) = 41.0$$

(LO1)

77.

Amount	*f*	*X*	*M*	*fM*^{2}
20 to under 30	1	25	25	625
30 to under 40	15	35	525	18,375
40 to under 50	22	45	990	44,550
50 to under 60	8	55	440	24,200
60 to under 70	4	65	260	16,900
Total	50		2240	104,650

$$\bar{X} = \frac{2240}{50} = 44.8 \qquad s = \sqrt{\frac{104,650 - (2240)^2/50}{50 - 1}} = 9.37$$

$$\text{Median} = 40 + \frac{\frac{50}{2} - 16}{22}(10) = 44.1$$

(LO1)

78.

Expenditure	*f*	*X*	*M*	*fM*^{2}
25 to under 35	5	30	150	4500
35 to under 45	10	40	400	16,000
45 to under 55	21	50	1050	52,500
55 to under 65	16	60	960	57,600
65 to under 75	8	70	560	39,200
Total	60		3120	169,800

$\bar{X} = \frac{3120}{60} = 52$ $\qquad s = \sqrt{\frac{169{,}800 - (3120)^2/60}{60-1}} = 11.32$

Median = $45 + \frac{\frac{60}{2} - 15}{21}(10) = 52.14$ **(LO1)**

79. a. Mean = 5, found by (6 + 4 + 3 + 7 + 5)/5
Median is 5, found by ordering the values and selecting the middle value.
b. Population because all partners were included.
c. $\sum(X - \mu) = (6-5)+(4-5)+(3-5)+(7-5)+(5-5) = 0$ **(LO1)**

80. a. Mean = 21.71, Median = 22.00 **(LO1)**
b. (23-21.7)+(19-21.7)+(26-21.7)+(17-21.7)+(21-21.7)+(24-21.7)+(22-21.7)=0

81. $\bar{X} = \frac{545}{16} = 34.06$ Median = 37.50 **(LO1)**

82. $\bar{X} = \frac{2116}{30} = 70.5333$ **(LO1)**

83. The Communications industry has older workers than the Retail Trade. Production workers have the most age difference. **(LO1)**

84. a. 4.84, found by 121/25
b. Median = 4.0
c. On half the days she made at least 4 appointments. The arithmetic mean number of appointments per day is 4.84. **(LO1)**

85. $\bar{X}_w = \frac{\$5.00(270) + \$6.50(300) + \$8.00(100)}{270 + 300 + 100} = \$6.12$ **(LO1)**

86. $\bar{X}_w = \frac{3(4) + 3(4) + 5(3) + 2(3) + 1(4)}{3 + 3 + 5 + 2 + 1} = 3.50$; weighted mean **(LO1)**

87. $\bar{X}_w = \frac{[15{,}300(4.5) + 10{,}400(3.0) + 150{,}600(10.2)]}{176{,}300} = 9.28$ **(LO1)**

88. a. arithmetic mean = 6.49 km
b. median = 6.3 km
c. mode = 5.3 and 4.6 km (bimodal) **(LO1)**

89.

Wage (X)	Freq (f)	fX	fX^2
13.00	20	260	3380
15.50	12	186	2883
18.00	8	144	2592

Totals: 40 590 8855

Mean = 590/40 = 14.75 Variance = $\dfrac{8855 - \dfrac{(590)^2}{40}}{40} = 3.8125$

Standard deviation = 1.95 **(LO1&4)**

90. GM = $\sqrt[10]{\dfrac{33,598}{25,000}} - 1 = 1.0300 - 1 = 0.03$ or 3.0 percent **(LO1)**

GM = $\sqrt[10]{\dfrac{44,771}{25,000}} - 1 = 1.0599995 - 1 = 0.06$ or 6.0 percent **(LO1)**

91. a. 55, found by 72 – 17
b. 14.4, found by 144/10 where $\overline{X}$ = 43.2
c. 17.6245
d. $1 - 1/k^2$ = 1 – ¼ = .75 = 75%
e. 43.2 ± 2(17.6245) = 7.95 to 78.45 **(LO1,4&6)**

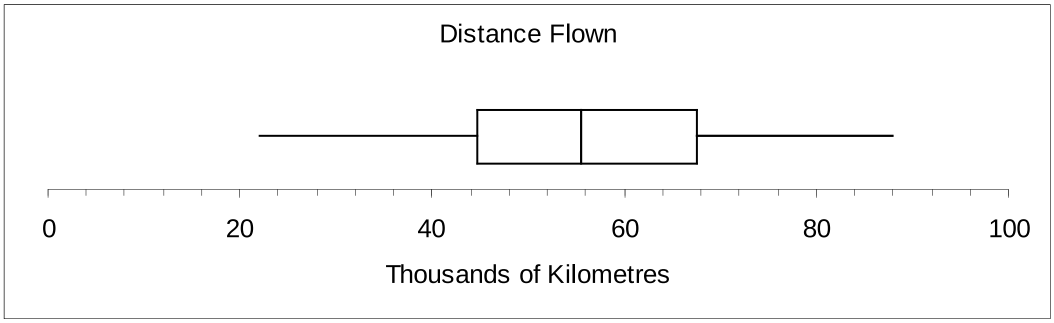
92. a. 9, found by 12 – 3
b. 2.72, found by 13.6/5 where mean = 7.6
c. 3.5071
d. $1 - 1/k^2$ = 1 – 1/6.25 = .84 = 84%; yes, all are
e. 7.6 ± 3.5071 = 4.09 to 11.1; 60% are **(LO1,4&6)**

93. a. population
b. 183.47
c. 94.92%; a lot of variability compared to the mean **(LO4&7)**

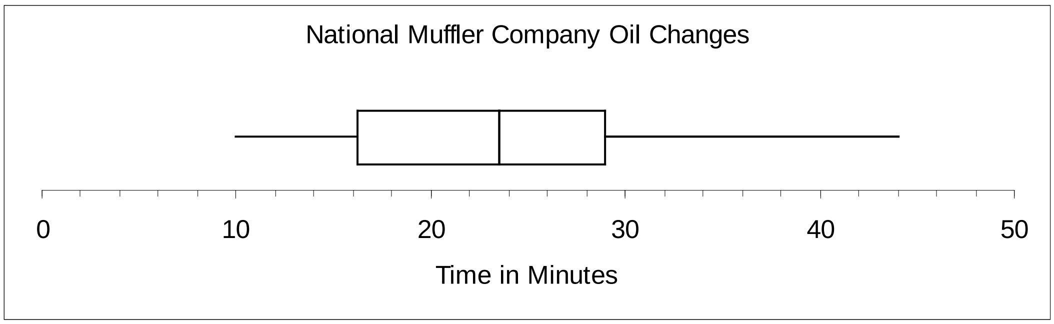
94. a. 30 found by 30 – 0
b. 6.09 found by $\sqrt{\dfrac{34,758 - \dfrac{(2094)^2}{150}}{149}}$ **(LO4)**

95. Comments may vary. The following results were found using statistical software. **(LO8&9)**

Q_1 = 44.25 Q_3 = 68.50 Median = 55.50
The distribution is approximately symmetric. The box plot is as follows.



96. a.



b. No outliers

c. Summaries may vary. The distribution is positively skewed. The median time to change a muffler is 23.50 minutes. The first quartile is 15.75 minutes and the third quartile is 29.25 minutes. The range of time is 10 minutes to 44 minutes. **(LO3,8&9)**

97. The distribution is positively skewed. The first quartile is approximately $20 and the third quartile is approximately $90. There is one outlier located at approx $255. The median is about 50. **(LO9)**

98. The distribution is positively skewed. The first quartile is equal to 10 and the third quartile is equal to 40. There are four outliers located at approx 85, 90, 95 and 100. The median is about 25. **(LO9)**

99. a. $\overline{X} = \frac{857.90}{50} = 17.158$ Median = 16.35 **(LO1,4,6,8&9)**

b. $s = \sqrt{\frac{20{,}206.73 - \frac{(857.90)^2}{50}}{50 - 1}} = 10.58$

c. $17.158 \pm (2)(10.58) = -4.002$ and 38.318

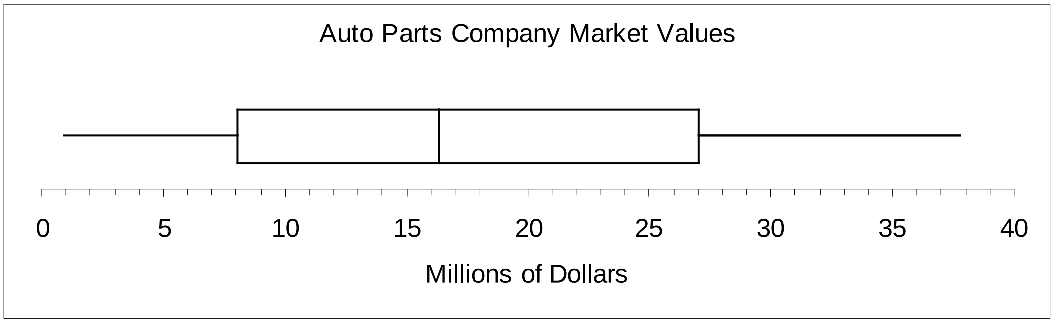
d. $CV = \frac{10.58}{17.158}(100) = 61.66\%$

e. $sk = \frac{3(17.158 - 16.35)}{10.58} = 0.23$

f. $L_p = (50+1)\frac{25}{100} = 12.75$ $L_p = (50+1)\frac{75}{100} = 38.25$

Excel will give you 8.075 & 27.025

$Q_1 = 7.825$ $Q_3 = 27.400$



g. The distribution is nearly symmetrical. The mean is 17.158, the median is 16.35 and the standard deviation is 10.58. About 75 percent of the companies have a value less than 27.4 and 25 percent have a value less than 7.825.

100. a. The times are a population because all tables for that night are included.
b. The mean is 40.84, found by 1021/25. The median is 39.
c. The range is 44, found by 67 – 23. The standard deviation is 14.55, found by the square root of 5291.4/25. **(LO1&4)**

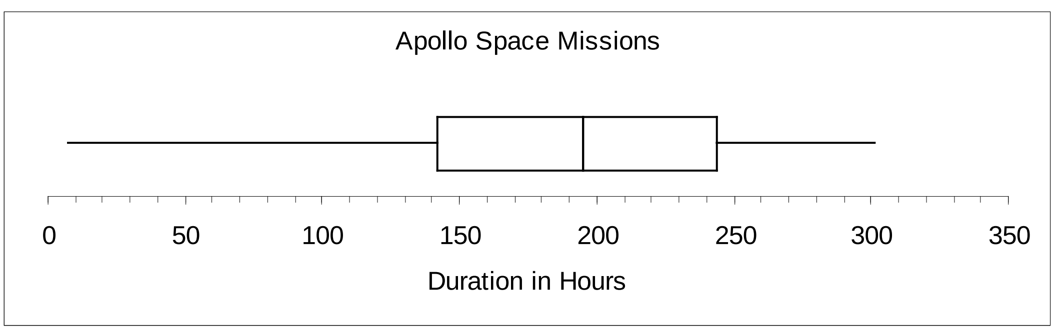
101. a. The mean is 173.77 hours, found by 2259/13. The median is 195 hours.

s = 101.47 hours, found by $\sqrt{\dfrac{526{,}391 - \dfrac{2259^2}{13}}{13}}$.

b. CV = 58.4%, found by (101.47/173.77) × 100
Coefficient of skewness is –0.697; slight negative skewness

c. $L_{45} = (14)(.45) = 6.3$. So the 45th percentile is 192 + 0.3(195- 192) = 192.9.
$L_{82} = (14)(.82) = 11.48$. So the 82nd percentile is 260 + 0.48(295- 260) = 276.8.

d.



There is a slight negative skewness visible, but no outliers.
(LO1,4,6,7,8&9)

102. $\overline{X} = 96.55$ found by 10,620/110

$$s = \sqrt{\frac{1{,}029{,}937.5 - \frac{(10{,}620)^2}{110}}{109}} = 6.514$$

$$\text{Median} = 95 + \frac{\frac{110}{2} - 47}{35}(5) = 96.14$$

(LO1&4)

103. Mean is 13, found by 910/70

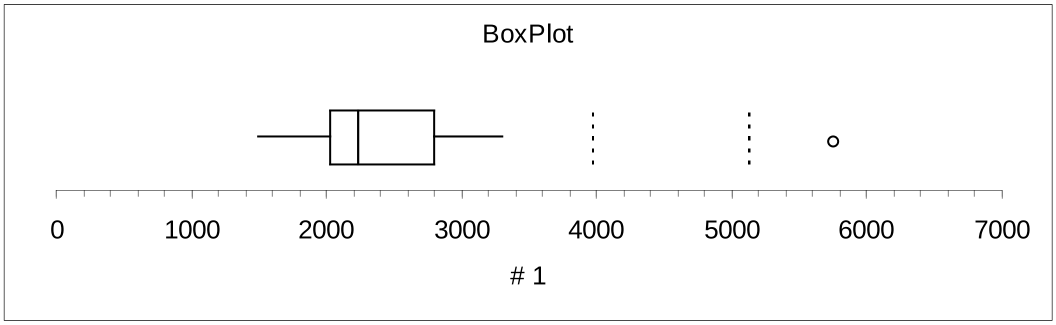
$$s = \sqrt{\frac{13{,}637.50 - \frac{(910)^2}{70}}{69}} = 5.118$$

$$\text{Median} = 10 + \frac{\frac{70}{2} - 19}{27}(5) = 12.96 \text{ kms.}$$

(LO1&4)

104. a. mean = $2706.67; median = $2235 **(LO1&9)**

b.



There is one extreme outlier. This observation will distort the value of the mean.

c. The median will be more representative of the data than the mean due to the high extreme outlier.

105.

Grade	Number of Students	fx
3	4	12
4	4	16
5	6	30
7	20	140
8	6	48
10	1	10
	41	256

Mean = 256/41 =6.24
Mode = 7

$$s = \sqrt{\frac{1714 - \frac{256^2}{41}}{40}} = 1.70;$$

IQR = 7-5 = 2 **(LO1,4&8)**

106. Mean = 76.83; mode = 80 ; standard deviation = 9.07; IQR = 80-70 = 10 **(LO1,4&8)**

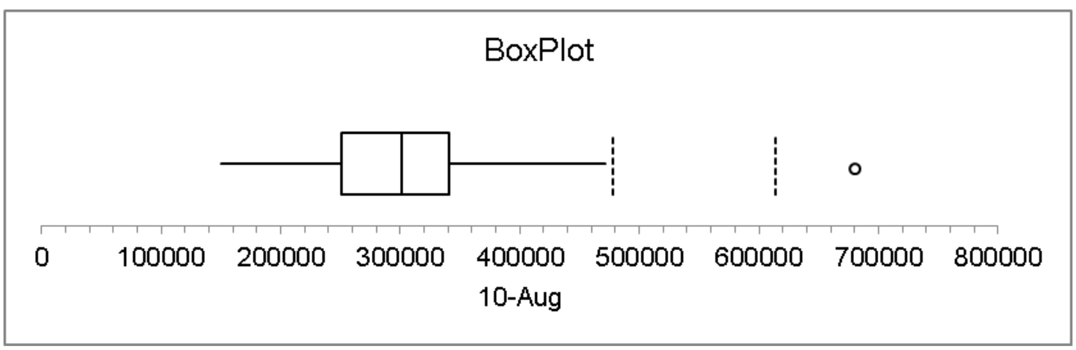
107. **(LO1,4,7,8&9)**

a.

1. Mean 318,811.44
   Median 301,759.50
   sample standard deviation 127,004.01

2. skewness 1.57
   there is moderate positive skewness

3. There is one extreme high outlier.



1st quartile	250,377.00
3rd quartile	341,340.50

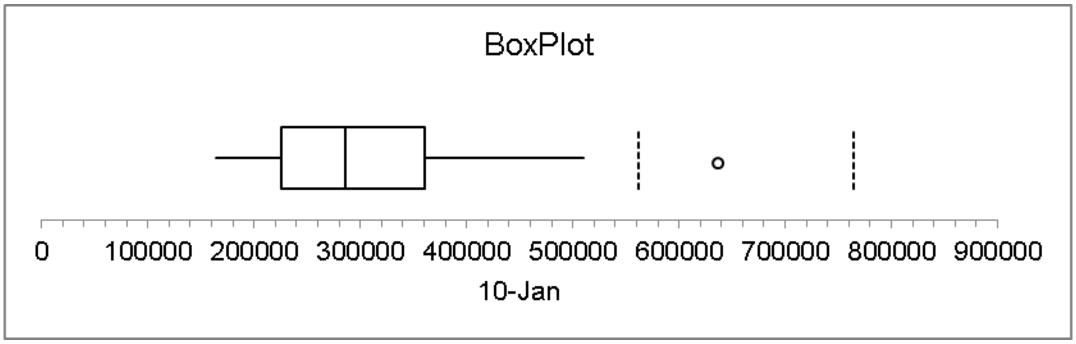
4. Answers will vary but should include the above and that the median is a more representative measure of location as the extreme high outlier affects the mean.

b.

1. Mean 314,036.94
   Median 286,390.50
   sample standard deviation 124,987.94

2. skewness 1.31
   there is moderate positive skewness

3. There is one high outlier.



1st quartile	236,227.50
3rd quartile	360,349.00

4. Answers will vary but should include the above and that the median may be a more representative
measure of location as the high outlier affects the mean.

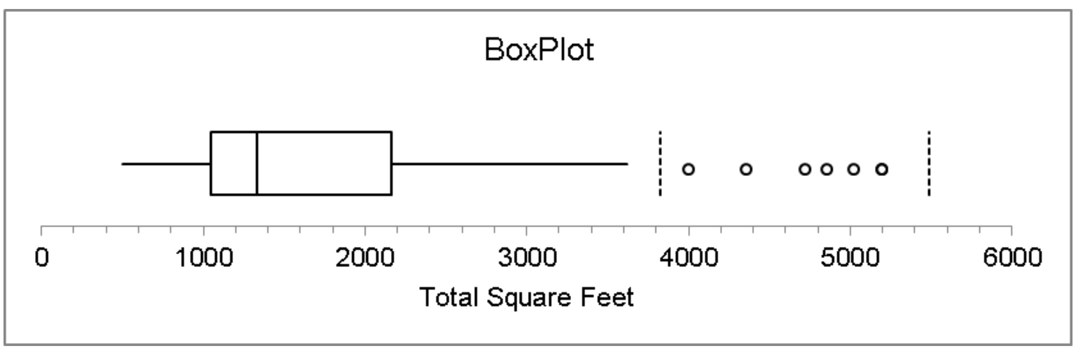
108. **(LO1,4,7,8&9)**

a.

1. Mean 1,710.01
Median 1,330.00
sample standard deviation 1,103.16

2. Skewness 1.69
there is moderate positive skewness

3. There are 7 high outliers.



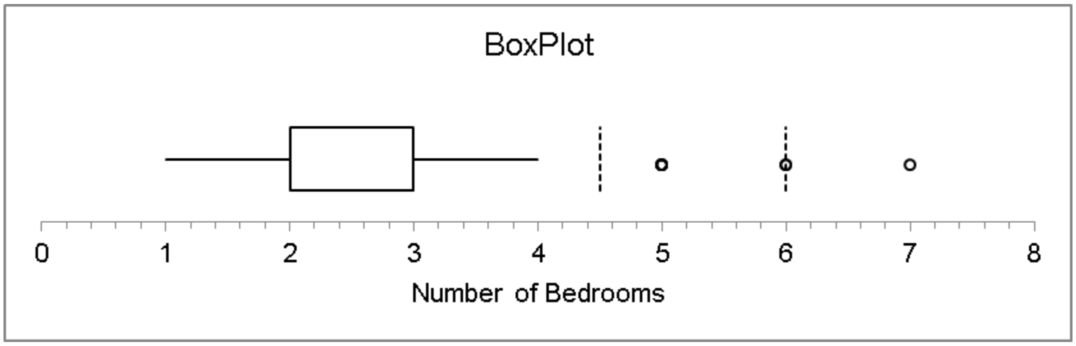
1st quartile	1,050.00
3rd quartile	2,160.00

4. Answers will vary but should include the above and that the median is a more representative
measure of location as the 7 high outliers affect the mean.

b.

1.	Mean	2.49
	Median	2.00
	sample standard deviation	1.29
2.	Skewness	1.14
	there is moderate positive skewness	

3. There are 9 high outliers, one of which is an extreme high outlier.



1st quartile	2.00
3rd quartile	3.00

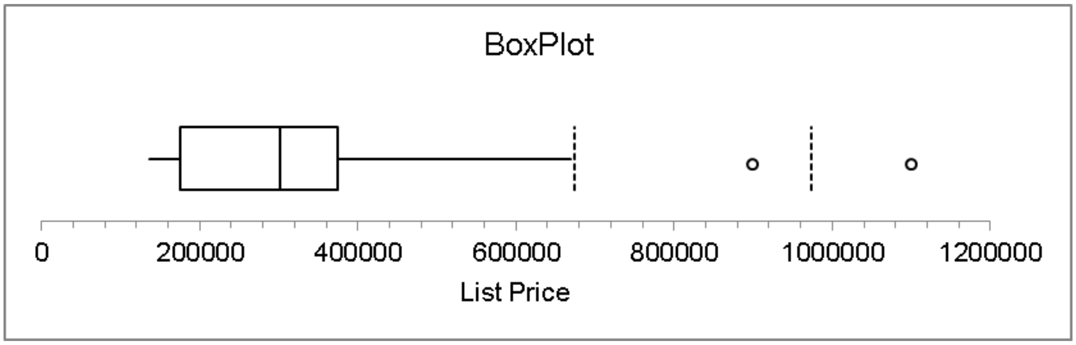
4. Answers will vary but should include the above and that the median is a more representative
measure of location as the 9 high outliers affect the mean.

109. **(LO1,4,7,8&9)**

a.

1.	Mean	331,233.98
	Median	301,200.00
	sample standard deviation	197,121.89
2.	Skewness	1.86
	there is moderate positive skewness	

3 There are 2 high outliers, one of which is an extreme high outlier.



1st quartile	174,900.00
3rd quartile	374,650.00

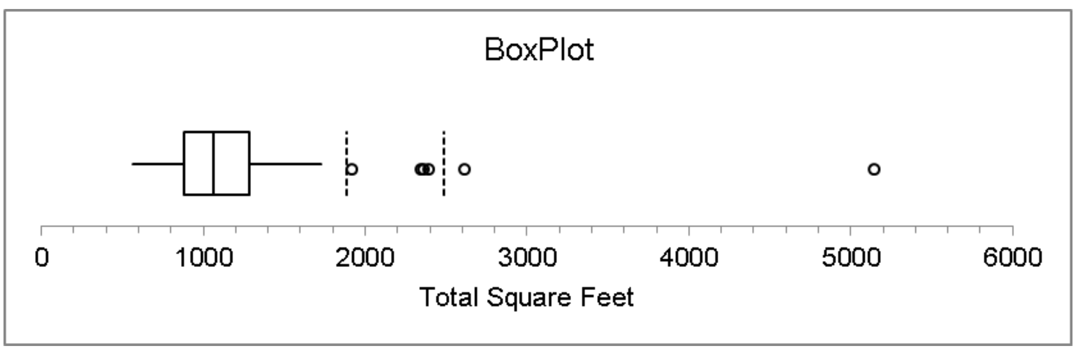
4. Answers will vary but should include the above and that the median is a more representative
measure of location as the 2 high outliers affect the mean.

b.

1. Mean 1,250.78
Median 1,066.00
sample standard deviation 730.87

2. Skewness 3.51
extreme positive skewness

3. There are 6 high outliers, two of which are extreme high outliers.



1st quartile	882.00
3rd quartile	1,283.75

4. Answers will vary but should include the above and that the median is a more representative
measure of location as the 6 high outliers affect the mean.

c.

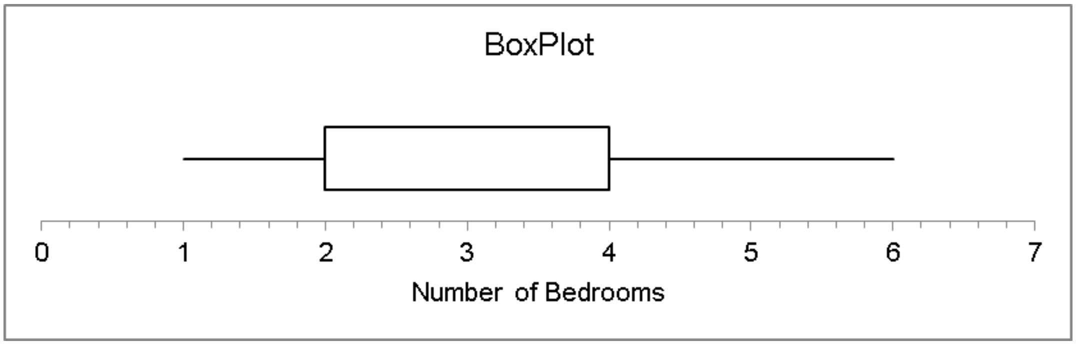
1.
mean 2.62

median	2.00
sample standard deviation	1.24

2.

skewness	0.44

slight positive skewness



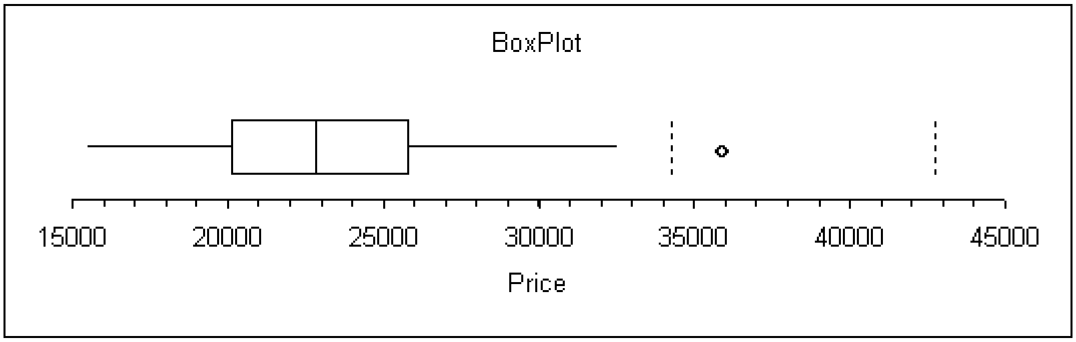
1st quartile	2.00
3rd quartile	4.00

4.
Answers will vary but should include the above information.

Case Study   Answers will vary

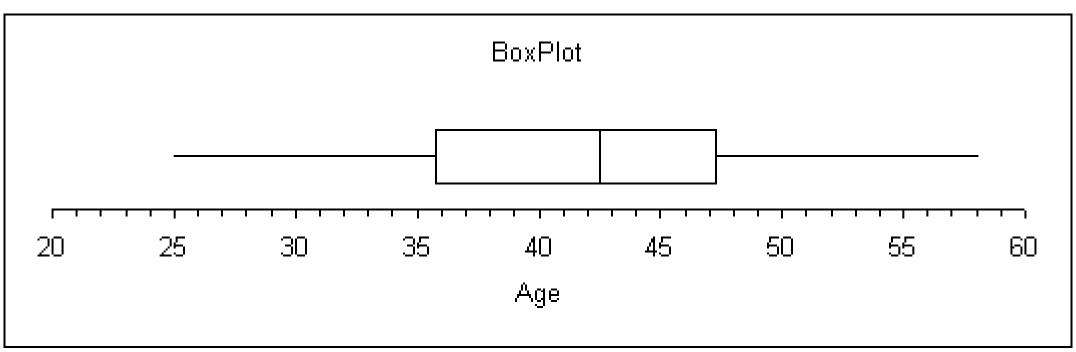
Descriptive statistics

	Price
Count	80
Mean	23,218.16
Sample variance	18,961,128.64
Sample standard deviation	4,354.44
1st quartile	20,128.00
Median	22,831.00
3rd quartile	25,787.00
interquartile range	5,659.00
Mode	20,642.00
low extremes	0
low outliers	0
high outliers	2
high extremes	0



Descriptive statistics

	Age
Count	80
Mean	42.15
Sample variance	67.52
Sample standard deviation	8.22
1st quartile	35.75
Median	42.50
3rd quartile	47.25
interquartile range	11.50
Mode	46.00
low extremes	0
low outliers	0
high outliers	0
high extremes	0



Summary Statistics	Price	Age
mean	23218.16	42.15
median	22831.00	42.5
mode	20642.00	46

standard deviation	4354.44	8.22

The mean selling price is $23 218 and the median is $22 831. These two values are less than $400 apart, so either measure is reasonable. The typical value is then about $23 000.
The mean age of the customer is about 42 years and the median age is 42.5 years. So a typical customer is about 42 years old.

CHAPTER 4

A SURVEY OF PROBABILITY CONCEPTS

1.

Outcome	*1*	*2*	
1	A	A	
2	A	F	
3	F	A	
4	F	F	**(LO 2)**

2.

Outcome	*1*	*2*	
1	A	A	
2	A	R	
3	A	S	
4	R	A	
5	R	R	
6	R	S	
7	S	A	
8	S	R	
9	S	S	**(LO 2)**

3. a. 0.176 found by 6/34
   b. Empirical **(LO 2&3)**

4. a. 0.40, found by 2/5
   b. Classical **(LO 2&3)**

5. a. Empirical
   b. Classical
   c. Classical
   d. Subjective **(LO 3)**

6. a.

Outcome	*1st*	*2nd*
1	M	M
2	M	F
3	F	F
4	F	M

   b. Classical **(LO 3)**

7. a. The survey of 40 people about environmental issues
   b. 26 or more, respond yes for example
   c. 0.25 found by 10/40
   d. Empirical
   e. The events are probably not equally likely (we don't know for sure) but they are mutually exclusive. **(LO 2&3)**

8. a. Recording the number of violations
   b. at least one violation, for example
   c. 0.009, found by 18/2000
   d. Empirical **(LO 2&3)**

9. a. Answers will vary, here are some possibilities: 123, 124, 125, 999

b. $\left(\frac{1}{10}\right)^4$

c. Classical **(LO 2&3)**

10. a. Answers will vary, value goes up 1/16, goes up 1/8, goes down 1

b. 0.01, 0.02, 0.01

c. Subjective **(LO 2&3)**

11. a. 78,960,960

b. 840, found by (7)(6)(5)(4). That is 7!/3!

c. 10, found by 5!/3!2! **(LO 2)**

12. a. 6,840

b. 504

c. 21 **(LO 2)**

13. 210, found by (10)(9)(8)(7)/(4)(3)(2) **(LO 2)**

14. 10,000, found by $(10)^4$ **(LO 2)**

15. 120, found by 5! **(LO 2)**

16. 3003, found by ${}_{15}C_{10} = \frac{15 \times 14 \times 13 \times 12 \times 11}{5 \times 4 \times 3 \times 2}$ **(LO 2)**

17. 10,897,286,400 found by ${}_{15}P_{10} = 15 \times 14 \times 13 \times 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6$ **(LO 2)**

18. 210 **(LO 2)**

19 P(A or B) = P(A) + P(B) **(LO 5)**

$= 0.30 + 0.20$

$= 0.50$

P(neither) $= 1 - 0.50 = 0.50$

20. P(X or Y) = P(X) + P(Y) **(LO 5)**

$= 0.05 + 0.02$

$= 0.07$

P(neither) $= 1 - 0.07 = 0.93$

21. a. 0.51 found by 102/200 **(LO 5)**

b. 0.49 found by (1 - .51) or by 61/200 + 37/200 = 0.305 + 0.185 Special rule of addition

22. a. 80%, found by 50% + 30%

b. 80%, found by 100% – 20% **(LO 5)**

23. 0.75, found by 0.25 + 0.50 **(LO 5)**

24. yes, no **(LO 4)**

25. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ **(LO 5)**

$$= 0.30 + 0.20 - 0.15$$
$$= 0.35$$

26. $P(X \text{ or } Y) = P(X) + P(Y) - P(X \text{ and } Y)$ **(LO 5)**

$$= 0.55 + 0.35 - 0.20$$
$$= 0.70$$

27. When two events are mutually exclusive it means that if one occurs the other event cannot occur. Therefore, the probability of their joint occurrence is zero **(LO 4)**.

28. $P(H \text{ or } M) = P(H) + P(M) - P(H \text{ and } M)$ **(LO 5)**

$$= 0.60 + 0.70 - 0.50$$
$$= 0.80$$

29.
a. 0.20
b. 0.30
c. No, because a store could have both.
d. Joint probability
e. 0.90, found by 1.0 – 0.10 **(LO 5)**

30.
a. 0.55, found by 0.50 + 0.40 – 0.35
b. Joint probability
c. No, a vacationer can visit both attractions **(LO 5)**

31. $P(A \text{ and } B) = P(A) \times P(B|A)$ **(LO 4&5)**

$$= 0.40 \times 0.30.$$
$$= 0.12$$

32. $P(X_1 \text{ and } Y_2) = P(X_1) \times P(Y_2|X_1)$ **(LO 4&5)**

$$= 0.75 \times 0.40$$
$$= 0.30$$

33. 0.90, found by (0.80 + 0.60) – 0.50
0.10, found by (1 – 0.90) **(LO 5)**

34. 5%, found by (1 – 0.95) **(LO 5)**

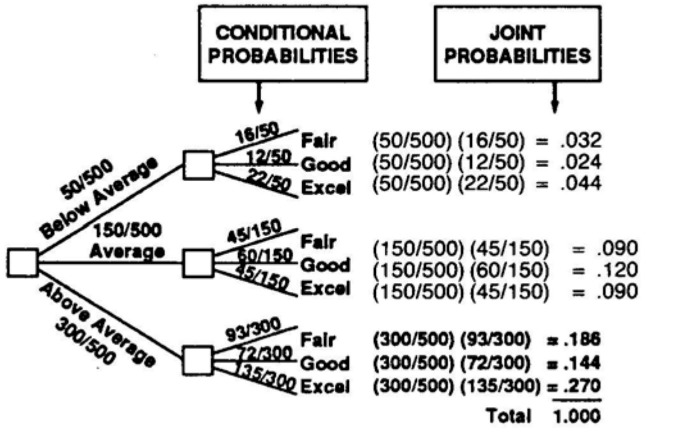
35.
a. $P(A_1) = 3/10 = 0.30$
b. $P(B_1|A_2) = 1/3 = 0.33$
c. $P(B_2 \text{ and } A_3) = 1/10 = 0.10$ **(LO 5)**

36.
a. 6/380 or 0.01579, found by 3/20 x 2/19
b. 272/380 or 0.7158, found by 17/20 x 16/19 **(LO 5)**

37. a. A contingency table

b. 0.27, found by 300/500 × 135 ÷ 300

c. A tree diagram would appear as: **(LO 5&6)**



38.

Outcome	*A*	*B*	*C*	*Outcome*	*A*	*B*	*C*	*Outcome*	*A*	*B*	*C*
1	U	U	U✓	10	D	U	U✓	19	S	U	U✓
2	U	U	D✓	11	D	U	D	20	S	U	D
3	U	U	S✓	12	D	U	S	21	S	U	S
4	U	D	U✓	13	D	D	U	22	S	D	U
5	U	D	D	14	D	D	D	23	S	D	D
6	U	D	S	15	D	D	S	24	S	D	S
7	U	S	U✓	16	D	S	U	25	S	S	U
8	U	S	D	17	D	S	D	26	S	S	D
9	U	S	S	18	D	S	S	27	S	S	S

U = Stock is up D = Stock is down S = Stock is same

✓ indicates the criteria that two stocks went up is met.

2 of the stocks went up on seven of the 27 outcomes P(A) = 7/27 = 0.259. **(LO 5)**

39. Probability the 1st presentation wins 3/5 = 0.60

Probability the 2nd presentation wins 2/5(3/4) = 0.30

Probability the 3rd presentation wins (2/5)(1/4)(3/3) = 0.10 **(LO 5)**

40. a. $\left(\frac{1}{7}\right)^3 = 0.0029$

b. $\left(\frac{7}{7}\right)\left(\frac{6}{7}\right)\left(\frac{5}{7}\right) = 0.612$

c. $\left(\frac{6}{7}\right)^3 = 0.630$ **(LO 5)**

41. a. nominal **(LO 5)**

b. 32/200 = 16%

c. 85/200 = 42.5%

d. Yes, as 32% of men ordered dessert compared to 15% of women

42. a. 19/50 = 38%

b. 8/50 = 16%
c. 15/50 + 26/50 – 8/50 = 33/50 = 66%
d. 6/26 = 23.1% **(LO 5)**

43. a. 106/659 = 16.1%
b. 143/659 = 21.7%
c. 12/659 = 1.8%
d. 233/659 + 233/659 – 87/659 = 57.5%
e. 40/161 = 24.8% **(LO 5)**

44. a. 55/195 = 28.2%
b. 25/195 = 12.8%
c. 60/90 = 66.7% **(LO 5)**

45. a. Asking teenagers their reactions to a newly developed soft drink.
b. Answers will vary. One possibility is more than half of the respondents like it. **(LO 2)**

46 Empirical **(LO 3)**

47 Subjective **(LO 3)**

48. a. 0.10, found by 50/500
b. Yes, mutually exclusive, because a given tip cannot fall in more than one category.
c. 1.00
d. 0.60, found by 300/500
e. 0.90, found by 450/500 or 1 – (50/500) **(LO 5)**

49. a. The likelihood an event will occur, assuming that another event has already occurred.
b. The collection of one or more outcomes of an experiment.
c. A measure of the likelihood that two or more events will happen concurrently.**(LO 2&4)**

50. $2^6 = 64$ **(LO 2)**

51. 26 x 10 x 26 x 10 x 26 x10 = 17 576 000 ways **(LO 2)**

52. P(4,4) = 24 ways **(LO 2)**

53. C(52,7) = 133 784 560 ways **(LO 2)**

54. 6 x 6 x 6 = 216 ways (numbers repeated) **(LO 2)**

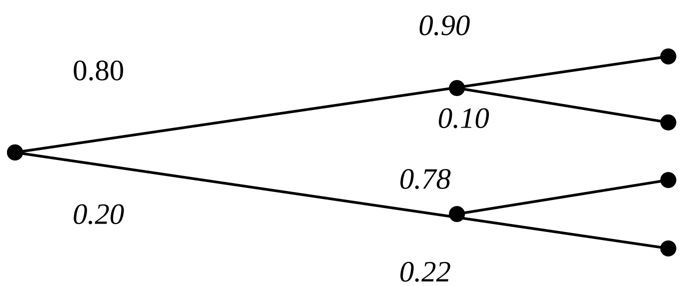
55. P(15,6) = 3 603 600 ways **(LO 2)**

56. a. 4/52, or 0.077
    b. 3/51, or 0.059
    c. 0.0045, found by (4/52)(3/51) **(LO 5)**

57. a. 0.8145, found by $(0.95)^4$
    b. Special rule of multiplication
    c. P(A and B and C and D) = P(A) x P(B) x P(C) x P(D) **(LO 5)**

58. a. Venn diagram
    b. Complement rule
    c. 1 **(LO 2)**

59 a. 0.08, found by 0.80 x 0.10
    b. the text answer is labeled



0.80 x 0.90 = 0.72

0.80 x 0.10 = 0.080

0.20 x 0.78 = 0.156

0.20 x 0.22 = 0.044

    c. Yes, because all the possible outcomes are shown on the tree diagram. **(LO 6)**

60. 0.6561 found by (0.9)(0.9)(0.9)(0.9)
    b. 0.0001 found by (0.1)(0.1)(0.1)(0.1)
    c. 0.3439 found by 1 - 0.6561 **(LO 5)**

61. a. 0.57 found by 57/100
    b. 0.97 found by (57/100) + (40/100)
    c. Yes, because an employee cannot be both.
    d. 0.03 found by 1 - 0.97 **(LO 5)**

62. a. empirical
    b. 0.046 found by $(0.359)^3$
    c. 0.263 found by $(1 - 0.359)^3$
    d. 0.737 found by 1 - 0.263 **(LO 3&5)**

63. a. 0.5 found by (2/3)(3/4)
    b. 1/12 found by (1/3)(1/4)
    c. 11/12 found by 1 – 1/12 **(LO 5)**

64. a. 1/9 found by (1/3)(1/3)
    b. 1/27 found by (1/3)(1/3)(1/3)
    c. 2/9 found by (2/3)(1/3) **(LO 5)**

65. a. 0.9039 found by $(0.98)^5$

b. 0.0961 found by 1 – 0.9039 **(LO 5)**

66. a. 0.064 found by $(0.4)^3$
b. 0.216 found by $(0.6)^3$
c. 0.784 found by 1 – 0.216
d. independent **(LO 4&5)**

67. a. 0.0333 found by (4/10)(3/9)(2/8)
b. 0.1667 found by (6/10)(5/9)(4/8)
c. 0.8333 found by 1 – 0.1667
d. dependent **(LO 4&5)**

68. 0.0889 found by (2/10)(4/9) **(LO 5)**

69. a. 0.3818, found by (9/12)(8/11)(7/10)
b. 0.6182, found by 1 – 0.3818 **(LO 5)**

70. C(7,2)C(50,3) = (21)(19 600) = 411 600 ways **(LO 2)**

71. C(20,4)C(15,3) = (4845)(455) = 2 204 475 ways **(LO 2)**

72. C(10,3)C(12,3) + C(10,2)C(12,4) + C(10,1)C(12,5) + C(10,0)C(12,6) = 57 519 ways **(LO 2)**

73. C(30,4)C(20,4) + C(30,5)C(20,3) + C(30,6)C(20,2) + C(30,7)C(20,1) + C(30,8)C(20,0)
= 454 620 240 ways **(LO 2)**

74. a. 0.40, found by 200/500
b. 0.60, found by 100/500 + 200/500
c. 0.60, found by (200/500) + (200/500) – (100/500) general rule of addition
d. 0.33, found by 100/300
e. 0.1595, found by (200/500)(199/499) **(LO 5)**

75. a. 0.30, found by 6/20
b. 0.45, found by (6 + 7 - 4)/20
c. 0.5714, found by 4/7
d. 0.0789, found by (6/20)*(5/19) **(LO 5)**

76. a. 0.42
b. 0.70(0.40) = 0.28
c. 0.88 **(LO 5)**

77. a. P(P or D) =(1/50) + (1/10) = 0.10 + 0.02 = 0.12
b. P(No) = (49/50)(9/10) = 0.882
c. P(No on 3) = $(0.882)^3$ = 0.686
d. P(at least one prize) = 1 – 0.686 = 0.314 **(LO 5)**

78. a. (10)(9)(8) = 720
b. 0.00139, found by 1/720
c. 0.99583, found by 1 – 3/720 **(LO 2&5)**

79. Yes, 256 is found by 2^8 **(LO 2)**

80. 0.70, found by $= P(A) + P(B) - P(A \text{ and } B)$ **(LO 5)**

$$= 0.60 + 0.40 - 0.30$$
$$= 0.70$$

81. 0.9744, found by $1 – (0.40)^4$ **(LO 2)**

82. 15, found by 5 x 3 **(LO 2)**

83. a. 0.185, found by (0.15)(0.95) + (0.05)(0.85)
b. 0.0075, found by (0.15)(0.05) **(LO 5)**

84. a. 17,576,000 found by (26)(26)(26)(10)(10)(10)
b. $(26)^4(10)^3$ = 456 976 000 **(LO 2)**

85. a. P(F and > 60) = 0.25, found by solving with the general rule of multiplication:
1 = P(F) ×P(>60|F) = (0.50)(0.50)
b. 0
c. 0.3333, found by 1/3 **(LO 5)**

86. a. 0.333, found by (6/10)(5/9)
b. 0.9286, found by 1 – [(6/10)(5/9)(4/8)(3/7)]
c. dependent

87. 456,976 found by 26^4 **(LO 2)**

88. a. 2024, found by $_{24}C_3 = \dfrac{24!}{3!(24-3)!}$
b. 0.125, found by 1 – [(23/24)(22/23)(21/22)] **(LO 5)**

89. 3, 628,800 matches are possible. So the probability is 1/3, 628,800. **(LO 2)**

90. For the system to operate both components in the series must work. The probability they both work is 0.81, found by (0.90)(0.90) **(LO 5)**

91. 0.512, found by $(0.8)^3$ **(LO 2)**

92.

a. 58/98 = .592

b. = 1-14/98 = .857

c. = 41/98 = .418

d. = 37/98 = 0.378

e. = 78/98 = .796

f. = 39/98 = .398

g. = 33/98 = .337

h. = 34/98 = .347

I. = 7/98 = .071

j. = 28/98 + 34/98 = 62/98 = .633

k. = 36/98 = .367

l.

	Square Feet								
Type	500>1000	1000>1500	1500>2000	2000>2500	2500>3000	3000>3500	3500>4000	4000+	Totals
Apartment	18	32	6	2	0	0	0	0	58
House	0	4	4	4	2	3	2	7	26
Townhouse	1	4	3	5	0	1	0	0	14
Totals	19	40	13	11	2	4	2	7	98

1. 58/98 = .592

2. = 8/26 = .308

3. = 4/98 = .041

4. =18/58 = .310

m.

		List price ($000's)				
# Bedrooms	0>500	500>1000	1000>1500	1500>2000	2000>2500	Totals
1	20	0	0	0	0	20
2	32	7	2	0	0	41
3	10	8	0	1	0	19
4	0	6	1	2	0	9
5+	2	3	2	1	1	9
	64	24	5	4	1	98

1. = 19/98 = .94

2. = 64/98 = .653

3. = 7/98 = .071

4. = 1/10 = .010

n.

		List price ($000's)				
Square Feet	0>500	500>1000	1000>1500	1500>2000	2000>2500	Totals
0 > 1000	19	0	0	0	0	19

1000 > 2000	38	13	1	1	0	53
2000 > 3000	6	6	1	0	0	13
3000+	1	5	3	3	1	13
Totals	64	24	5	4	1	98

1. =1 - 19/98 = 79/98 = .806
2. = 38/64 = .594
3. = 1/98 = .010
4. = 53/98 + 5/98 - 1/98 = 57/98 = .582

93.

a. = 24/50 = .48

b. =26/50 = .52

c. =20/50 = .40

d. =3/50 = .06

e. =18/50 = .36

f. =23/50 = .46

g. =12/50 = .24

h. =9/50 = .18

i. =15/50 = .30

j. = 25/50 + 9/50 = 34/50 = .68

k. = 1 - 34/50 = 16/50 = .32

l.

	List price ($000's)			
Type	0 > 500	500>1000	1000>1500	Totals
Apartment	23	1	0	24
House	16	7	1	24
Townhouse	2	0	0	2
Totals	41	8	1	50

1. = 24/50 = .48
2. = 7/24 = .292

3. = 1/50 = .02

4. = 1/24 = .042

m.

# Bedrooms	List price ($000's) 0>500	500>1000	1000>1500	Totals
1	9	1	0	10
2	17	0	0	17
3	6	2	0	8
4+	9	5	1	15
	41	8	1	50

1. = 8/50 = .16

2. = 41/50 = .82

3. = 0

4. = 2/8 = .25

n.

Square Feet	List price ($000's) 0 > 500	500>1000	1000>1500	Totals
0 > 1000	18	0	0	18
1000 > 2000	23	4	0	27
2000+	0	4	1	5
Totals	41	8	1	50

1. = 32/50 = .64

2. =23 /50 = .460

3. =23 /41 = .561

4. =18/50 + 41/50 – 18/50 = .82

CHAPTER 5

DISCRETE PROBABILITY DISTRIBUTIONS

1. Mean = 1.3 Variance = 0.81 found by: **(LO3)**

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

$$\sigma^2 = \sum (X - \mu)^2 P(X)$$
$$= (0 - 1.3)^2(0.2) + (1 - 1.3)^2(0.4) + (2 - 1.3)^2(0.3) + (3 - 1.3)^2(0.1)$$
$$= 0.81$$

2. Mean = 5.4 Variance = 12.04 found by: **(LO3)**

$\mu = \sum XP(X) = 2(.5) + 8(.3) + 10(.2) = 5.4$

$$\sigma^2 = \sum (X - \mu)^2 P(X)$$
$$= (2 - 5.4)^2(.5) + (8 - 5.4)^2(.3) + (10 - 5.4)^2(.2)$$
$$= 12.04$$

3. a. the middle one
   b. (1) 0.3 = 30%
      (2) 0.3 = 30%
      (3) 0.9 = 90%
   c. $\mu = 5(0.1) + 10(0.2) + 15(0.3) + 20(0.4) = 15$
      $\sigma^2 = (5 – 15)^2 (0.1) + (10 – 15)^2 (0.2) + (15 – 15)^2 (0.2) + (20 – 15)^2 (0.4)$
      $= 25$
      $\sigma = 5$ **(LO3)**

4. a. discrete
   b. continuous
   c. discrete
   d. continuous
   e. continuous **(LO2)**

5. a. **(LO2&3)**

Number of Calls	Probability
0	0.16
1	0.20
2	0.44
3	0.18
4	0.02

   b. discrete
   c. 1.7, found by 0.16(0) + 0.20(1) + 0.44(2) + 0.18(3) + 0.02(4)
   d. 1.005, found by

$$\sqrt{0.16(0 - 1.7)^2 + 0.20(1 - 1.7)^2 + 0.44(2 - 1.7)^2 + 0.18(3 - 1.7)^2 + 0.02(4 - 1.7)^2}$$

6. $\mu = 1000(.6) + 1200(.3) + 1500(.1) = 1110$

$\sigma^2 = (1000 - 1110)^2(.6) + (1200 - 1110)^2(.3) + (1500 - 1110)^2(.1) = 24{,}900$

$\sigma = 157.8$ **(LO3)**

7. a. 0.20
b. 0.55
c. 0.95
d. $\mu = 0(.45) + 10(.3) + 100(.2) + 500(0.05) = 48$

$\sigma^2 = (0 - 48)^2(.45) + (10 - 48)^2(.3) + (100 - 48)^2(.2) + (500 - 48)^2(.05) = 12{,}226$

$\sigma = 110.57$ (LO3)

8. a. 0.5000
b. 0.666
c. 0
d. $\mu = 0(1/3) + 1(1/2) + 2(0) + 3(1/6) = 1$ **(LO3)**

9. a. 21, found by 0.50(10) + 0.40(25) + 0.08(50) + 0.02(100) **(LO3)**
b. 16.09, found by

$$\sqrt{0.50(10 - 21)^2 + 0.40(25 - 21)^2 + 0.08(50 - 21)^2 + 0.02(100 - 21)^2}$$

10. a. **(LO2&3)**

Hours	Probability
1	0.080
2	0.152
3	0.212
4	0.180
5	0.160
6	0.052
7	0.020
8	0.144

This is a discrete probability distribution.

b. The mean is 4.144, found by
0.08(1) + 0.152(2) + 0.212(3) + 0.180(4) + 0.160(5) +0.052(6) + 0.020(7) +0.144(8)
The standard deviation is 2.0908, found by the square root of
$0.08(1-4.144)^2 + 0.152(2-4.144)^2 + 0.212(3-4.144)^2 + 0.180(4-4.144)^2 + 0.160(5-4.144)^2 + 0.052(6-4.144)^2 + 0.020(7-4.144)^2 + 0.144(8-4.144)^2$
The “typical” customer is parked for 4.144 hours.

c. The mean is 11.53, found by
0.08(3) + 0.152(6) + 0.212(9) + 0.180(12) + 0.160(14) +0.052(16) + 0.020(18) +0.144(20)
The standard deviation is 5.01, found by the square root of
$0.08(3-11.53)^2 + 0.152(6-11.53)^2 + 0.212(9-11.53)^2 + 0.180(12-11.53)^2 + 0.160(14-11.53)^2 + 0.052(16-11.53)^2 + 0.020(18-11.53)^2 + 0.144(20-11.53)^2$

11. a. $P(2) = \dfrac{4!}{2!(4-2)!}(.25)^2(.75)^{4-2} = 0.2109$

b. $P(3)=\frac{4!}{3!(4-3)!}(.25)^3(.75)^{4-3}=0.0469$

c. P(2) + P(3) + P(4) = 0.2109 + 0.0469 + 0.0039 = 0.2617

d. P(0) + P(1) + P(2) = 0.3164 + 0.4219 + 0.2109 = 0.9492 **(LO4)**

12. a. $P(1)=\frac{5!}{1!(5-1)!}(.4)^1(.6)^{5-1}=0.2592$

b. $P(2)=\frac{5!}{2!(5-2)!}(.4)^2(.6)^{5-2}=0.3456$

c. P(3) + P(4) + P(5) = 0.2304 + 0.0768 + 0.0102 = 0.3174

d. 1 – P(5) = 1 – 0.0102 = 0.9898 **(LO4)**

13. a.

x	*P(X)*
0	0.064
1	0.288
2	0.432
3	0.216

b. $\mu=0(0.064)+\ldots 3(0.216)=1.8$ $\sigma^2=(0-1.8)^2 0.064+\ldots+(3-1.8)^2 0.216=0.72$

$\sigma=\sqrt{0.72}=0.8485$ **(LO4)**

14. a.

x	*P(x)*
0	0.168
1	0.360
2	0.309
3	0.132
4	0.028
5	0.002

b. $\mu=0(0.168)+\ldots+5(0.002)=1.5$ $\sigma^2=(0-1.5)^2 0.168+\ldots+(5-1.5)^2 0.002=1.05$

$\sigma=\sqrt{1.05}=1.0247$ **(LO4)**

15. a. 0.2668, found by $P(2)=\frac{9!}{2!(9-2)!}(.3)^2(.7)^7$

b. 0.1715, found by $P(4)=\frac{9!}{4!(9-4)!}(.3)^4(.7)^5$

c. 0.0404, found by $P(0)=\frac{9!}{0!(9-0)!}(.3)^0(.7)^9$ **(LO4)**

16. a. 0.7351, found by $P(6)=\frac{6!}{6!(6-6)!}(.95)^6(.05)^0$

b. 0.2321, found by $P(5)=\frac{6!}{5!(6-5)!}(.95)^5(.05)^1$

c. μ= 5.7, found by 6(.95)

d. $\sigma^2=0.285$, found by 6(.95)(.05), $\sigma^2=\sqrt{0.285}=0.5339$ **(LO4)**

17. a. 0.2824, found by $P(0) = \frac{12!}{0!(12-0)!}(0.1)^0(0.9)^{12}$

b. 0.3765, found by $P(1) = \frac{12!}{1!(12-1)!}(0.1)^1(0.9)^{11}$

c. 0.2301, found by $P(2) = \frac{12!}{2!(12-2)!}(0.1)^2(0.9)^{10}$

d. P(0) + P(1) + P(2) = 0.2824 + 0.3765 + 0.2301 = 0.8890

e. μ = 1.2, found by 12(0.1), $\sigma^2 = 1.08$, $\sigma = 1.04$ **(LO4)**

18. a. 0.2311, found by ${}_{12}C_4(0.3)^4(0.7)^8$

b. 0.0138, found by ${}_{12}C_0(0.3)^0(0.7)^{12}$

c. 0.1678, found by ${}_{12}C_2(0.3)^2(0.7)^{10}$

d. 3.6, found by 12(0.3) **(LO4)**

19. a. 0.1858, found by $\frac{15!}{2!13!}(0.23)^2(0.77)^{13}$

b. 0.1416, found by $\frac{15!}{5!10!}(0.23)^5(0.77)^{10}$

c. 3.45 found by (0.23)(15) **(LO4)**

20. a. 0.2479, found by $\frac{8!}{0!8!}(0.16)^0(0.84)^8$

b. 0.7521, found by 1 – 0.2479

c. 0.0038, found by 0.0035 + 0.0003 + 0.0000, where P(5) = 0.0035 found by $\frac{8!}{5!3!}(0.16)^5(0.84)^3$, $P(6)$ = 0.0003 found by $\frac{8!}{6!2!}(0.16)^6(0.84)^2$, P(7) = 0.0000 found by $\frac{8!}{7!1!}(0.16)^7(0.84)^1$ and so forth. **(LO4)**

21. a. 0.296, found by using Appendix A.3 with *n* of 8, *p* of 0.30 and *x* of 2. **(LO4)**
b. $P(x \le 2) = 0.058 + 0.198 + 0.296 = 0.552$
c. 0.448, found by $P(x \ge 3) = 1 - P(x \le 2) = 1 - 0.552$

22. a. 0.101, found from Appendix A.3, *n* = 12, *p* = 0.60, *x* = 5 **(LO4)**
b. $P(x \le 5) = 0.002 + 0.012 + 0.042 + 0.101 = 0.157$
c. $P(x \ge 6) = 1 - P(x \le 5) = 1 - 0.157 = 0.843$

23. a. 0.387, found from Appendix A.3 with *n* of 9, *p* of 0.90, and an *x* of 9 **(LO4)**
b. P(*x* < 5) = 0.001
c. 0.992, found by 1 – 0.008
d. 0.947, found by 1 – 0.053

24. a. 0.358, found from Appendix A.3 with *n* of 20, *p* of 0.05, and an *x* of 0 **(LO4)**
b. 0.642, found by 1 – 0.358
c. 0.076, found by 1 – [0.358 + 0.377 + 0.189]

25. a. μ = 10.5, found by 15(0.7) and $\sigma = \sqrt{15(0.7)(0.3)} = 1.7748$
b. 0.2061, found by $\frac{15!}{10!5!}(0.7)^{10}(0.3)^5$
c. 0.4247, found by 0.2061 + 0.2186
d. 0.5154, found by 0.2186 + 0.1700 + 0.0916 + 0.0305 + 0.0047 **(LO4)**

26. a. 6, found by (0.3)(20)
b. 0.1789, found by $\frac{20!}{5!15!}(0.3)^5(0.7)^{15}$
c. 0.0479, found by 0.0308 + 0.0120 + 0.0039 + 0.0010 + 0.0002 + 0.0000
d. Yes, probability is 0.9992, found by 1 - 0.0008 **(LO4)**

27. $$P(2) = \frac{({}_6C_2)({}_4C_1)}{{}_{10}C_3} = \frac{15(4)}{120} = 0.50$$ **(LO5)**

28. $$P(3) = \frac{({}_{10}C_3)({}_5C_1)}{{}_{15}C_4} = \frac{120(5)}{1365} = 0.4396$$ **(LO5)**

29. $$P(0) = \frac{({}_7C_2)({}_3C_0)}{{}_{10}C_2} = \frac{21(1)}{45} = 0.4667$$ **(LO5)**

30. a. $$P(3) = \frac{({}_6C_3)({}_2C_0)}{{}_8C_3} = \frac{20(1)}{56} = 0.3571$$
b. $P(X \le 2) = 1 - P(X = 3) = 1 - 0.3571 = 0.6429$ **(LO5)**

31. $$P(2) = \frac{({}_9C_3)({}_6C_2)}{{}_{15}C_5} = \frac{84(15)}{3003} = 0.4196$$ **(LO5)**

32. $P(0) = \frac{({}_{11}C_5)({}_4C_0)}{{}_{15}C_5} = \frac{462(1)}{3003} = 0.1538$ $\quad P(X \geq 1) = 1 - P(0) = 1 - 0.1538 = 0.8462$ **(LO5)**

33. a. 0.6703
b. 0.3297 **(LO6)**

34. a. 0.1465
b. 0.2381
c. 0.7619 **(LO6)**

35. a. 0.0613
b. 0.0803 **(LO6)**

36. a. 0.1353
b. 0.8647 **(LO6)**

37. $\mu = 6,\ P(X \geq 5) = 0.7149 = 1 - (.0025 + .0149 + .0446 + .0892 + .1339)$ **(LO6)**

38. 0.8088 **(LO6)**

39. A random variable is a quantitative or qualitative outcome, which results from a chance experiment. A probability distribution also includes the likelihood of each possible outcome. **(LO1)**

40. A discrete probability distribution can take on only values that are clearly separated from each other. On the other hand, a continuous probability distribution can take on any value in a range.
a. continuous
b. discrete
c. discrete
d. discrete
e. continuous
f. discrete
g. continuous **(LO2)**

41. The binomial distribution is a discrete probability distribution where the there are only two possible outcomes. A second important part is that data collected is a result of counts. Additionally, one trial is independent from the next and the chance for success remains the same from one trial to the next. **(LO4)**

42. When *n* is large and p is small the Poisson and the binomial distribution will yield approximately the same results. **(LO6)**

43. $\mu = 0(.1) + 1(.2) + 2(.3) + 3(.4) = 2.00$ $\quad \sigma^2 = (0 - 2)^2(.1) + ... + (3 - 2)^2(.40) = 1.0$ $\quad \sigma = 1$
(LO4)

44. $\mu = 0.25(\$1000) + 0.60(\$2000) + 0.15(\$5000) = \$2200$

$\sigma^2 = (1000 - 2200)^2(.25) + \ldots + (5000 - 2200)^2(.15) = 1{,}560{,}000$ **(LO4)**

45. $\mu = 0(.4) + 1(.2) + 2(.2) + 3(.1) + 4(.1) = 1.3$ $\sigma^2 = (0 - 1.30)^2(.4) + \ldots + (4 - 1.30)^2(.10) = 1.81$
$\sigma = 1.3454$ **(LO4)**

46. ¼ (.75) + ½ (.50) + ¼ (-1.00) = .1875 **(LO4)**

47. μ = 13.2, found by 12(.25) + 13(.4) + 14(.25) + 15(.1) = 3.0 + 5.2 + 3.5 + 1.5
σ^2 = 0.86, found by 0.36 + 0.016 + 0.16 + 0.324
$\sigma = \sqrt{0.86} = 0.9274$ **(LO4)**

48. a. 3.5. found by (0.35)(10)
b. 0.2377, found by 210(0.01501)(0.07542)
c. 0.4862, found by 0.2377 + 0.1536 + 0.0689 + 0.0212 + 0.0043 + 0.0005 **(LO4)**

49. a. discrete
b. continuous
c. discrete
d. discrete
e. continuous **(LO2)**

50. a. the 3rd table
b. 1. .10 = 10%
2. .80 = 80%
3. .50 = 50%
c. μ = 25(.5) + 50(.3) + 75(.1) + 100(.1) = 45
$\sigma^2 = (25 – 45)^2 (0.5) + (50 – 45)^2 (0.3) + (75 – 45)^2 (0.1) + (100 – 45)^2 (0.1)$
= 600
σ = 24.5 **(LO4)**

51. a. 6, found by (0.4)(15)
b. 0.0245, found by $\frac{15!}{10!5!}(0.4)^{10}(0.6)^5$
c. 0.0338, found by 0.0245 + 0.0074 + 0.0016 + 0.0003 + 0.0000
d. 0.4032, found by 0.0005 + 0.0047 + 0.0219 + 0.0634 + 0.1268 + 0.1859

52. 0.1294, found by $\frac{18!}{14!4!}\left(\frac{2}{3}\right)^{14}\left(\frac{1}{3}\right)^{4}$ **(LO4)**

53. a. μ = 20(0.075) = 1.5 and $\sigma = \sqrt{20(0.075)(0.925)} = 1.1779$
b. 0.2103, found by $\frac{20!}{0!20!}(0.075)^0(0.925)^{20}$
c. 0.7897, found by $1 - 0.2103$ **(LO4)**

54. a. μ=12(0.07) = 0.84 and $\sigma = \sqrt{12(0.07)(0.93)} = 0.8839$

b. 0.4186, found by $\frac{12!}{0!12!}(0.07)^0(0.93)^{12}$

c. 0.5814, found by 1 - 0.4186 **(LO4)**

55. a. 0.1311, found by $\frac{16!}{4!12!}(0.15)^4(0.85)^{12}$

b. 2.4, found by (0.15)(16)

c. 0.2100, found by 1 - 0.0743 - 0.2097 - 0.2775 - 0.2285 **(LO4)**

56. a.

0	0.0025
1	0.0207
2	0.0763
3	0.1665
4	0.2384
5	0.2340
6	0.1596
7	0.0746
8	0.0229
9	0.0042
10	0.0003

b. $\mu = 10(0.45) = 4.5$ and $\sigma = \sqrt{10(0.45)(0.55)} = 1.5732$

c. 0.2384

d. 0.5044, found by 0.0025 + 0.0207 + 0.0763 + 0.1665 + 0.2384 **(LO4)**

57 a.

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

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

c. 0.1768

d. 0.3343, found by 0.0002 + 0.0019 + 0.0116 + 0.0418 + 0.1020 + 0.1768 **(LO4)**

58. a. .002(1000) = 2

b. 0.1353, found by $\frac{2^0 e^{-2}}{0!}$

c. 0.8647, found by 1 – 0.1353 **(LO6)**

59. a. 0.0498

b. 0.7746, found by $(1 – 0.0498)^5$

60. a. $\mu = 3$, probability = 0.0498

b. 0.5768, found by 1 – (0.0498 + 0.1494 + 0.2240) **(LO6)**

61. $\mu = 4.0$; from Appendix A.4

a. 0.0183

b. 0.1954

c. 0.6289

d. 0.5665 **(LO6)**

62. a. 0.2707, found by $\frac{2^1 e^{-2}}{1!}$

b. 0.0527, found by 0.0361 + 0.0120 + 0.0034 + 0.0009 + 0.0002 + 0.0000

c. 0.1353, found by $\frac{2^0 e^{-2}}{0!}$ **(LO6)**

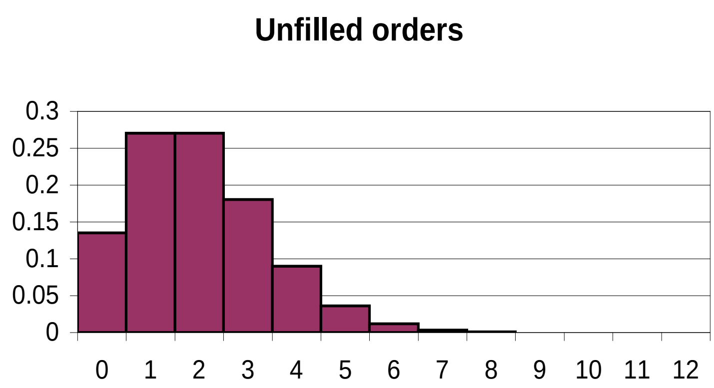
63. a. 0.00005, found by $\frac{(18.4)e^{-18.4}}{4!}$

b. almost 0, found by $\frac{(18.4)e^{-18.4}}{0!}$

c. 0.38489, found by 1 - 0.61511 **(LO6)**

64. a. $P(X < 5 \mid \mu = 2) = 0.94735$ so they are very close, but not quite over the goal. **(LO6)**

b.



65. P(0) = 0.68588 and P(2) = 0.04741

66. Let n = 34, and p = 29/34 = 0.8529 $P(5) = {}_{34}C_5(0.8529)^{29}(0.1471)^5 = 0.1899$

67. Let μ=155(1/3709) = 0.042 $P(5) = \frac{0.042^5 e^{-0.042}}{5!} = 0.000000001$ Very Unlikely!

68. a.

X	Days	P(X Sold)
0	4	0.1333
1	15	0.5000

2	5	0.1667
3	3	0.1000
4	2	0.0667
5	1	0.0333
	30	1.00

b. μ = 1.5667 per day; this represents the average number of big screen TV's sold

c. 1.5667($3575)(.15) = $840.14 per day, so for a 30-day month, multiply by 30.

69. a.

Number of Bedrooms (X)	Count	p(x)	Xp(x)	(x-μ)	$(x-μ)^2$	$(x-μ)^2p(x)$
1	20	0.20408	0.2041	-1.4898	2.2195	0.4530
2	41	0.41837	0.8367	-0.4898	0.2399	0.1004
3	19	0.19388	0.5816	0.5102	0.2603	0.0505
4	9	0.09184	0.3673	1.5102	2.2807	0.2095
5	6	0.06122	0.3061	2.5102	6.3011	0.3858
6	2	0.02041	0.1224	3.5102	12.3215	0.2515
7	1	0.01020	0.0714	4.5102	20.3419	0.2076
Total	98	1.00000	2.4898			1.6581

mean =	2.4898	
variance =		1.6581
standard deviation =		1.2877

b.

Full Baths (x)	Count	p(x)	Xp(x)	(x-μ)	$(x-μ)^2$	$(x-μ)^2p(x)$
1	37	0.37755	0.3776	-0.7653	0.5857	0.2211
2	51	0.52041	1.0408	0.2347	0.0551	0.0287
3	6	0.06122	0.1837	1.2347	1.5245	0.0933
4	4	0.04082	0.1633	2.2347	4.9939	0.2038
Total	98	1.00000	1.7653			0.5470

mean =	1.7653	
variance =		0.5470
standard deviation =		0.7396

70. a.

Full Baths (x)	Count	p(x)	Xp(x)	(x-μ)	$(x-μ)^2$	$(x-μ)^2p(x)$
1	27	0.540	0.5400	-0.5600	0.3136	0.1693
2	20	0.400	0.8000	0.4400	0.1936	0.0774
3	2	0.040	0.1200	1.4400	2.0736	0.0829
4	0	0.000	0.0000	2.4400	5.9536	0.0000

5	1	0.020	0.1000	3.4400	11.8336	0.2367
Total	50	1.000	1.5600			0.5664

mean = 1.5600
variance = 0.5664
standard deviation = 0.7526

b.

No Beds (x)	Count	p(x)	Xp(x)	(x-μ)	$(x-\mu)^2$	$(x-\mu)^2 p(x)$
1	10	0.200	0.2000	-1.6200	2.6244	0.5249
2	17	0.340	0.6800	-0.6200	0.3844	0.1307
3	8	0.160	0.4800	0.3800	0.1444	0.0231
4	13	0.260	1.0400	1.3800	1.9044	0.4951
5	1	0.020	0.1000	2.3800	5.6644	0.1133
6	1	0.020	0.1200	3.3800	11.4244	0.2285
Total	50	1.000	2.6200			1.5156

mean = 2.6200
variance = 1.5156
standard deviation = 1.2311

CHAPTER 6

CONTINUOUS PROBABILITY DISTRIBUTIONS

1. a. $a = 6 \quad b = 10$
   b. 8, found by (6 + 10)/2
   c. 1.1547 found by $\sqrt{\frac{(10-6)^2}{12}}$
   d. [1/(10 – 6)](10 – 6) = 1
   e. 0.75, found by [1/(10 – 6)](10 – 7)
   f. 0.5, found by [1/(10 – 6)](9 – 7) **(LO2&3)**

2. a. $a = 2 \quad b = 5$
   b. 3.5, found by (2 + 5)/2
   c. 0.8660 found by $\sqrt{\frac{(5-2)^2}{12}}$
   d. [1/(5 – 2)](5 – 2) = 1
   e. 0.8, found by [1/(5 – 2)](5 – 2.6)
   f. 0.2667, found by [1/(5 – 2)](3.7 – 2.9) **(LO2&3)**

3. a. 0.3 found by [1/(30 – 20)](30 – 27)
   b. 0.4 found by [1/(30 – 20)](24 – 20) **(LO2&3)**

4. a. Mean is 2100 found by (400 + 3800) / 2
   b. 981.50 found by $\sqrt{\frac{(3800-400)^2}{12}}$
   c. 0.4706 found by [1 / (3800 - 400)] * (2000 – 400)
   d. 0.2353 found by [1 / (3800 - 400)] * (3800 – 3000) **(LO2&3)**

5. a. $a = 0.5, \; b = 8.0$
   b. Mean is 4.25, found by (0.5 + 8.0)/2

      Standard deviation is 2.16, found by $\sqrt{\frac{(8-0.5)^2}{12}}$
   c. 0.2, found by [1/(8.0 – 0.5](2.0 – 0.5)
   d. 0.0, found by [1/(8.0 – 0.5](3.0 – 3.0)
   e. 0.4, found by [1/(8.0 – 0.5](8.0 – 5.0) **(LO2&3)**

6. a. $a = 0.5, \; b = 10.0$ (using minutes as the units)
   b. Mean is 5.25, found by (0.5 + 10)/2

      Standard deviation is 2.74, found by $\sqrt{\frac{(10-0.5)^2}{12}}$
   c. 52.63%, found by [1/(10 – 0.5]*(10 – 5)

d. 2.875, found from [1/(10 – 0.5]*(x – 5) = 0.25
and 7.625 found from [1/(10 – 0.5]*(10 – x) = 0.25 **(LO2&3)**

7. The actual shape of a normal distribution depends on its mean and standard deviation. Thus, there is a normal distribution, and an accompanying normal curve, for a mean of 7 and a standard deviation of 2. There is another normal curve for a mean of $25,000 and a standard deviation of $1742, and so on. **(LO4)**

8. It is bell shaped and symmetrical about its mean. It is asymptotic. There is a family of normal curves. The mean, median, and the mode are equal. **(LO4)**

9. a. 490 and 510, found by 500 ± 1(10)
b. 480 and 520, found by 500 ± 2(10)
c. 470 and 530, found by 500 ± 3(10) **(LO6)**

10. a. 68.26 percent
b. 95.4 percent
c. 99.7 percent **(LO6)**

11. a. 68.26%
b. 95.44%
c. 99.7% **(LO6)**

12. a. 275 and 425, found by 350 ± 1(75)
b. 200 and 500, found by 350 ± 2(75)
c. 125 and 575, found by 350 ± 3(75) **(LO6)**

13. $Z_{Rob} = \frac{50,000 - 60,000}{5000} = -2.00 \quad Z_{Rachel} = \frac{50,000 - 35,000}{8000} = 1.875$

Adjusting for their industries, Rob is well below average and Rachel well above. **(LO5)**

14. $Z_1 = \frac{75 - 90}{22} = -0.68 \qquad Z_2 = \frac{100 - 90}{22} = 0.45$

The first is slightly less expensive than average and the second is slightly more. **(LO5)**

15. a. 0.8413; 0.1587
b. 0.1056; 0.8944
c. 0.9977; 0.0023
d. 0.0094; 0.9906 **(LO7)**

16. a. 0.6826
b. 0.0828
c. 0.0378
d. 0.9238 **(LO6)**

17. a. 1.25 found by $z = \frac{25 - 20}{4.0} = 1.25$
b. 0.3944 = 39.44%, found in Appendix A.1

c. 0.3085 = 30.85%, found by $z = \frac{18 - 20}{4.0} = -0.5$

Find 0.1915 in Appendix A.1 for z = – 0.5; then 0.5000 – 0.1915 = 0.3085
(LO5&7)

18. a. z = 0.84, found by $z = \frac{14.3 - 12.2}{2.5} = 0.84$

b. 0.2995 = 29.95%, found in Appendix A.1

c. 0.1894 = 18.94%, found by $z = \frac{10 - 12.5}{2.5} = -0.88$

Find 0.3106 in Appendix A.1 for z = – 0.88, then 0.5000 – 0.3106 – 0.1894 **(LO5&7)**

19. a. 0.3413, found by $z = \frac{\$24 - \$20.50}{\$3.50} = 1.00$

Then find 0.3413 in Appendix A.1 for z = 1

b. 0.1587, found by 0.5000 – 0.3413 = 0.1587

c. 0.3336, found by $z = \frac{\$19.00 - \$20.50}{\$3.50} = -0.43$

Find 0.1664 in Appendix A.1, for a z = – 0.43, then 0.5000 – 0.1664 = 0.3336 **(LO7)**

20. a. About 0.4332 from Appendix A.1, where z = 1.50

b. About 0.1915, where z = – 0.50

c. About 0.3085, found by 0.5000 – 0.1915 **(LO7)**

21. a. 0.8276, first find z = – 1.5, found by ((44 – 50)/4) and z = 1.25 = (55 – 50)/4). The area between –1.5 and 0 is 0.4332 and the area between 0 and 1.25 is 0.3944, both from Appendix A.1. Then adding the two areas, we find that 0.4332 + 0.3944 = 0.8276.

b. 0.1056, found by 0.5000 – 0.3994, where z = 1.25

c. 0.2029, recall that the area for z = 1.25 is 0.3944, and the area for z = 0.5, found by ((52 – 50)/4) is 0.1915. Then subtract 0.3944 – 0.1915 and find 0.2029. **(LO6&7)**

22. a. 0.4017, first find z = – 0.36, found by ((75 – 80)/14) and z = 0.71(((90 – 80)/14). The area between –0.36 and 0 is 0.1406 and the area between 0 and 0.71 is 0.2611, both from Appendix A.1, then adding the two area we find 0.1406 + 0.2611 = 0.4017.

b. 0.3594, found by 0.5000 – 0.1406, where z = – 0.36

c. 0.2022 found by z = (55 – 80)/14 = – 1.79, for which the area is 0.4633. The z-value for 70 is – 0.71 and the corresponding area is 0.2661. So 0.4633 – 0.2611 = 0.2022 **(LO6&7)**

23. a. 0.1525, found by subtracting 0.4938 – 0.3413, which are the areas associated with z values of 2.5 and 1, respectively.

b. 0.0062, found by 0.5000 – 0.4938

c. 0.9710, found by recalling that the area of the z value of 2.5 is 0.4938. Then find z = – 2.00 found by ((205 – 225)/10). Thus, 0.4938 + 0.4772 = 0.9710. **(LO6&7)**

24. a. 0.3085, found by z = ($80,000 – $70,000)/$20,000 = 0.50. The area is 0.1915. Then 0.5000 – 0.1915 = 0.3085

b. 0.2902, found by z = ((80,000 – 70,000)/20,000) = 0.50, the area is 0.1915
z = ((65,000 – 70,000)/20,000) = – 0.25, the area is 0.0987
Adding these values together: 0.1915 + 0.0987 = 0.2902

c. 0.5987, found by the area under the curve with a $z = -0.25$, 0.0987 + 0.5000 = 0.5987 **(LO6&7)**

25. a. 0.0764, found by $z = (20 - 15)/3.5 = 1.43$, then 0.5000 – 0.4236 = 0.0764
b. 0.9236, found by 0.5000 + 0.4236, where $z = 1.43$
c. 0.1185, found by $z = (12 - 15)/3.5 = -0.86$. The area under the curve is 0.3051, then $z = ((10 - 15/3.5)) = -1.43$. The area is 0.4236, finally, 0.4236 – 0.3051 = 0.1185 **(LO6&7)**

26. a. 0.7348, found by 0.3186 + 0.4162
b. 0.0059, found by 0.5000 - 0.4941
c. 0.0779, found by 0.4941 - 0.4162 **(LO6&7)**

27. $X = 56.58$, found by adding 0.5000(the area left of the mean) and then finding a z value that forces 45% of the data to fall inside the curve. Solving for X: $1.645 = (X - 50)/4 = 56.58$. **(LO8)**

28. $$-0.84 = \frac{X - 80}{14}$$ **(LO8)**
$$X = 80.00 - 11.76 = 68.24$$

29. 200.7; find a z value where 0.4900 of area is between 0 and z. That value is $z = 2.33$, then solve for X: $(X - 200)/0.3$ so $X = 200.7$ **(LO8)**

30. a. \$107,600, found by \$70,000 + 1.88(20,000)
b. \$44,400, found by \$70,000 - 1.28(20,000) **(LO8)**

31. 1630, found by 2100 - 1.88(250) **(LO8)**

32. 1570, found by 1200 + 1.645(225) **(LO8)**

33. 1026, found by 900 + 0.84(150) **(LO8)**

34. 10,289; found by 12,200 – 2.33(820) **(LO8)**

35. a. $\mu = np = 50(0.25) = 12.5$
$\sigma^2 = np(1 - p) = 12.5(1 - 0.25) = 9.375$ $\sigma = \sqrt{9.375} = 3.0619$
b. 0.2578, found by (14.5 – 12.5)/3.0619 = 0.65, the area is 0.2422, then 0.5000 – 0.2422 = 0.2578
c. 0.2578, found by (10.5 – 12.5)/3.0619 = – 0.65. The area is 0.2422. Then 0.5000 – 0.2422 = 0.2578. **(LO9)**

36. a. $\mu = (40)(0.55) = 22$ $\sigma^2 = 9.9$ $\sigma = 3.15$
b. 0.2148, found by (24.5 – 22)/3.15 = 0.79. The area is 0.2852. Then 0.5000 – 0.2852 = 0.2148
c. 0.0197, found by (15.5 – 22)/3.15 = – 2.06. The area is 0.4803. Then 0.5000 – 0.4803 = 0.0197
d. $z = (14.5 - 22.0)/3.15 = -2.38$ and $(25.5 - 22.0)/3.15 = 1.1$, so 0.8578, found by 0.4913 + 0.3665 = 0.8578 **(LO9)**

37. a. 0.0192, found by 0.500 - 0.4808

b. 0.0694, found by 0.500 - 0.4306

c. 0.0502, found by 0.0694 - 0.0192 **(LO9)**

38. a. 10, which is the same as μ

b. 0.1894, found by ((7.5 – 10)/2.828) = – 0.88. The area is 0.3106. Then 0.5000 – 0.3106 = 0.1894

c. 0.2981, found by ((8.5 – 10)/2.828) = – 0.53. The area is 0.2019. Then 0.5000 – 0.2019 = 0.2981

d. 0.1087, found by 0.3106 – 0.2019 **(LO9)**

39. a. Yes. (1) There are two mutually exclusive outcomes-overweight and not overweight. (2) It is the result of counting the number of successes (overweight members). (3) Each trial is independent. (4) The probability of 0.30 remains the same for each trial.

b. 0.0084, found by μ=500(0.30) = 150 σ^2 = 500(0.30)(0.70) =105

$\sigma = \sqrt{105} = 10.24695$ $z = \frac{X - \mu}{\sigma} = \frac{174.5 - 150}{10.24695} = 2.39$

The area under the curve for 2.39 is 0.4916. Then 0.5000 – 0.4916 = 0.0084

c. 0.8461, found by $z = \frac{139.5 - 150}{10.24695} = -1.02$

The area between 139.5 and 150 is 0.3461. Adding 0.3461 + 0.5000 = 0.8461 **(LO9)**

40. a. About 0.9599, found by μ=100(0.38) = 38 σ^2 = 100(0.38)(1 – 0.38) = 23.56

$\sigma = \sqrt{23.56} = 4.85$ Then (29.5 – 38)/4.85) = – 1.75. Area under the curve for – 1.75 is 0.4599. Adding 0.4599 + 0.5000 = 0.9599

b. 0.6985, found by (40.5 – 38)/4.85) = 0.52, for which the area is 0.1985. Then 0.5000 + 0.1985 = 0.6985

c. 0.6584, found by 0.4599 + 0.1985 **(LO9)**

41. a. 12.005 found by (11.96 + 12.05) / 2

b. 0.02598 found by $\sqrt{\frac{(12.05 - 11.96)^2}{12}}$

c. 0.4444 found by [1 / (12.05 – 11.96)](12 – 11.96)

d. 0.7778 found by [1 / (12.05 – 11.96)](12.05 – 11.98)

e. 1.0 found by [1 / (12.05 – 11.96)](12.05 – 11.96) **(LO&3)**

42. a. 2.1 found by (0 + 4.2) / 2

b. 1.212 found by $\sqrt{\frac{(4.2 - 0)^2}{12}}$

c. 0.7143 found by [1 / (4.2 – 0)](3 – 0)

d. 0.6429 found by [1 / (4.2 – 0)](4.2 – 1.5) **(LO2&3)**

43. a. 7 found by (4 + 10) / 2

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

c. 0.3333 found by [1 / (10 – 4)](6 – 4)

d. 0.8333 found by [1 / (10 – 4)](10 – 5) **(LO2&3)**

44. a. 1 = [1 / (3.5 - 0)](3.5 – 0)
b. 1.75 found by (0 + 3.5) / 2
c. 1.01 found by $\sqrt{\frac{(3.5-0)^2}{12}}$
d. 28.567% found by [1 / (3.5 – 0)](1 – 0)
e. 42.86% found by [1 / (3.5 – 0)](3.5 – 2) **(LO2&3)**

45. a. 0.9406 and 0.0594
b. 0.9664 and 0.0336
c. 0.2177 and 0.7823
d. 0.0071 and 0.9929 **(LO7)**

46. a. .8931, found by .4332 + .4599
b. .1059, found by .4966 - .3907
c. .1420, found by .4951 + .3531
d. .2077, found by .4929 - .2852 **(LO6,7&8)**

47. a. -0.71
b. 0.2611+ 0.4686 = .7297 = 72.97%
c. 0.2611 + 0.5 = 0.7611 = 76.11%
d. 0.4251 + 0.5 = .9251 = 92.51%
e. 0.0749 + 0.2389 = 0.3138 = 31.38%
f. 0.84 = (X – 50)/7; X = 55.88 **(LO6,7&8)**

48. a. .0475, found by .5 - .4525
b. .0150, found by .5 - .4850
c. .5934, found by 2(.2967)
d. .8664, found by 2(.4332)
e. -.84 = (X – 45000)/6000; X = 39 960
f. 1.28 = (X – 45000)/6000; X = 52 680 **(LO6,7&8)**

49. a. – 0.4 for net sales, found by (170 – 180)/25 and 2.92 for employees, found by (1850 – 1500)/120

b. Net sales are 0.4 standard deviations below the mean. Employees is 2.92 standard deviations above the mean.

c. 65.54% of the aluminum fabricators have greater net sales compared with Clarion, found by 0.1554 + 0.5000. Only 0.18% have more employees than Clarion, found by 0.5000 – 0.4982 **(LO5)**

50. a. $\frac{29-32}{2}=-1.5$ $\frac{34-32}{2}=1.0$ 0.3413
  b. 0.7745, found by 0.4332 + 0.3413
  c. 0.0495, found by 0.5000 - 0.4505
  d. 35.3, found by 32 + 1.65(2) **(LO5,6&7)**

51. a. almost 0.5000, because $z=\frac{30-490}{90}=-5.11$
  b. 0.2514, found by 0.5000 - 0.2486
  c. 0.6374, found by 0.2486 + 0.3888
  d. 0.3450, found by 0.3888 - 0.0438 **(LO6&7)**

52. a. 0.4082 = 40.82%, because $z=\frac{5-4.2}{0.6}=1.33$
  b. 0.0918 = 9.18%, found by 0.5000 - 0.4082
  c. 0.0905 = 9.05%, found by 0.4987 - 0.4082
  d. 0.6280 = 62.80%, found by 0.4987 + 0.1293
  e. 5.25, found by 4.2 + 1.75(0.6) **(LO6,7&8)**

53. a. 0.3015 = 30.15%, found by 0.5000 - 0.1985
  b. 0.2579 = 25.79%, found by 0.4564 - 0.1985
  c. 0.0011 = .11%, found by 0.5000 - 0.4989
  d. $1818, found by 1280 + 1.28(420) **(LO6,7&8)**

54. a. 0.3446, found by 0.5000 – 0.1554
  b. 0.6006, found by 0.1554 + 0.4452
  c. 0.1039, found by 0.4452 – 0.3413
  d. 44,200, found by 40,000 + 0.84(5000) **(LO6,7&8)**

55. a. 90.82%, found by z = (40 – 34)/4.5 = 1.33, then 0.5000 + 0.4082
  b. 78.23%, found by 0.5000 + 0.2823, where z = (25 – 29)/5.1 = –0.78
  c. Prob($Z > x$) = 0.01 implies Prob ($0 < Z < x$) = 0.49 and x = 2.33
  Women: 34 + 2.33(4.5) = 44.485
  Men: 29 + 2.33(5.1) = 40.883 **(LO7&8)**

56. a. 0.1314 or 13.14%, found by z = (2500 – 1994)/450 = 1.12, then 0.5000 – 0.3686
  b. 0.1189 or 11.89%, found by 0.4875 – 0.3686, when z = (3000 – 1994)/450= 2.24
  c. 0.0136 or 1.36%, found by z = (1000 – 1994)/450 = –2.21, then 0.5000 – 0.4864
  (LO6&7)

57. About 4099 units found by solving for X. 1.645 = (X – 4000)/60 **(LO8)**

58.
a. 0.0047, found by 0.5000 – 0.4953
b. 0.1241, found by 0.4292 – 0.3051
c. 0.8413, found by 0.5000 + 0.3413
d. No, because $z = \frac{70{,}000 - 60{,}000}{2000} = 5.0$ **(LO5,7&7)**

59.
a. 15.39%, found by (8 – 10.3)/2.25 = – 1.02, then 0.5000 – 0.3461 = 0.1539
b. 17.31%, found by: z = (12 – 10.3)/2.25 = 0.76. Area is 0.2764
z = (14 – 10.3)/2.25 = 1.64. Area is 0.4495
The area between 12 and 14 is 0.1731, found by 0.4495 – 0.2764.
c. Yes, but it is rather remote. Reasoning: On 99.73% of the days, returns are between 3.55 and 17.03, found by 10.3 ± 3(2.25). Thus, the chance of less than 3.55 returns is rather remote. **(LO6&7)**

60.
a. Then (4.5 – 10)/2.83 = – 1.94, for which the area is 0.4738. Then 0.5000 – 0.4738 = 0.0262
b. 0.9441, found by (5.5 – 10)/2.83 = – 1.59, for which the area is 0.4441. Then 0.5000 + 0.4441 = 0.9441
c. 0.0297, found by (4.5 – 10)/2.828 = – 1.94 and (5.5 – 10)/2.828 = – 1.59. Then 0.4738 – 0.4441 = 0.0297
d. 0.8882, found by adding the area between z = – 1.59 and z = 1.59. Then 2(0.4441) = 0.8882 **(LO8)**

61.
a. (37 – 39.5)/1.5 = -1.67. Then 0.4525 + 0.5 = 0.9525
b. (41.5 – 39.5)/1.5 = 1.33. Then 0.4082 + 0.5 = 0.9082
c. (36 – 39.5)/1.5 = -2.33. Then 0.4901 - 0.4525 = 0.0376
d. (.5 - .4901) + 0.9525 = 0.9624 **(LO6&7)**

62.
a. 0.0091, found by z = (2500 – 4200)/720 = –2.36, then 0.5000 – 0.4909
b. 0.0062, found by z = (6000 – 4200)/720 = 2.5, then 0.5000 – 0.4938
c. 0.9847, found by 0.4909 + 0.4938 **(LO7&8)**

63.
a. 0.9678, found by: m=60(0.64) = 38.4
$s^2 = 60(0.64)(0.36) = 13.824$ $s = \sqrt{13.824} = 3.72$
Then (31.5 – 38.4)/3.72 = – 1.85, for which the area is 0.4678.
Then 0.5000 + 0.4678 = 0.9678
b. 0.0853, found by (43.5 – 38.4)/3.72 = 1.37, for which the area is 0.4147. Then 0.5000 – 0.4147 = 0.0853
c. 0.8084, found by 0.4441 + 0.3643
d. 0.0348 found by 0.4495 – 0.4147 **(LO8)**

64.
a. m= 60(0.1) = 6 and $s = \sqrt{60(0.1)(0.9)} = 2.32$ **(LO8)**
b. 0.0393, found by 0.4738 - 0.4345
c. 0.9738, found by 0.5000 + 0.4738

z = (24.5 – 20)/3.46 = 1.30. The area is 0.4032. Then for 25 or more, 0.5000 – 0.4032 = 0.0968

65. 0.0968, found by: $\mu = 50(0.40) = 20$ **(LO8)**

$\sigma^2 = 50(0.40)(0.60) = 12$ $\sigma = \sqrt{12} = 3.46$

$z = (24.5 - 20)/3.46 = 1.30$. The area is 0.4032. Then for 25 or more, 0.5000 – 0.4032 = 0.0968

66. $\mu = 800(0.80) = 640$ $\sigma = \sqrt{800(0.80)(0.20)} = 11.31$ **(LO8)**

$z = \dfrac{664.5 - 640}{11.31} = 2.17$ Probability is 0.5000 – 0.4850 = 0.0150

67.
a. $1.645 = (45 - \mu)/5$ $\mu = 36.78$
b. $1.645 = (45 - \mu)/10$ $\mu = 28.55$
c. $z = (30 - 28.5)/10 = 0.15$, then 0.5000 + 0.0596 = 0.5596 **(LO7&8)**

68.
a. $\dfrac{2 - 3.1}{0.3} = -3.67$ $\dfrac{3 - 3.1}{0.3} = -0.33$ 0.3707, found by 0.5000 - 0.1293
b. almost 0
c. 0.0228, found by 0.5000 - 0.4772; leads to 228 students, found by 10,000(0.0228).
d. 3.484, found by 3.1 + 1.28(0.3) **(LO7&8)**

69.
a. 21.19 percent found by $z = (3 - 3.1)/0.125 = -0.80$; so 0.5000 – 0.2881 = 0.2119
b. Increase the mean. $z = (3 - 3.1)/0.125 = -1.2$; probability is 0.5000 – 0.3849 = 0.1151
Reduce the standard deviation. $z = (3 - 3.1)/0.1 = -1.0$; the probability = 0.5000 – 0.3413 = 0.1587
Increasing the mean is better because a smaller percent of the hams will be below the limit. **(LO7)**

70. 0.5000 – 0.3333 = 0.1667 so $z = 0.43$ $-0.43 = \dfrac{40 - 43.9}{\sigma}$ $\sigma = 9.07$

0.5000 – 0.2000 = 0.3000, so $z = 0.84$ $0.84 = \dfrac{49 - 43.9}{\sigma}$ $\sigma = 6.07$

There is about a 50 percent difference between the two standard deviations. The distribution is not normal. Something may be wrong with the report. **(LO7)**

71.
a. $z = (100 - 85)/8 = 1.88$, so 0.5000 – 0.4699 = 0.0301
b. Let $z = 0.67$, so $0.67 = (X - 85)/8$ and X = 90.36, set mileage at 90 360
c. $z = (72 - 85)/8 = -1.63$, so 0.5000 – 0.4484 = 0.0516 **(LO7&8)**

72.
a. $z = (45.00 - 42.000)/2.25 = 1.33$, $p(z > 1.33) = 0.5000 - 0.4082 = 0.0918$. It is over \$45 about 22 days, found by 240(0.0918).
b. $z = (38.00 - 42.00)/2.25 = -1.78$ and $z = (40.00 - 42.00)/2.25 = -0.89$. So 0.4625 – 0.3133 = 0.1492 or 14.92 percent of the days.
c. \$45.44 (z=1.53) **(LO6,7&8)**

73. $\dfrac{470 - \mu}{\sigma} = 0.25$ $\dfrac{500 - \mu}{\sigma} = 1.28$ $\sigma = 29.126$ and $\mu = 462.718$ **(LO7)**

74. 32.56 found by – 1.28 = (X – 36.84)/3.34 **(LO8)**

75. m=150(0.15) = 22.5 $s\sqrt{150(0.15)(0.85)} = 4.37$

z = (29.5 – 22.5)/4.37 = 1.60 P(z > 1.60) = 0.5000 – 0.4452 = 0.0548 **(LO8)**

76.

Descriptive statistics

	List Price
Count	98
Mean	567,496.76
sample variance	173,446,667,107.59
sample standard deviation	416,469.29

normal distribution

P(lower)	P(upper)	z	X	mean	std.dev
.4121	.5879	-0.22	475000	567496.8	416469.3

actual number of homes > $475 000 = 43

The actual number of homes listed for more than $475 000 is 43, so the probability is 43/98 = 0.438776

The values should be closer, so we will further examine the list prices to see if the distribution is skewed. We get the following results from MegaStat.

skewness 2.08

We can see that the distribution is fairly positively skewed. This is not a good approximation as a result.

77.

Descriptive statistics

	List Price
count	50
mean	331,233.98
population variance	38,079,899,213.38
population standard deviation	195,140.72

Note :

sample standard deviation 197,121.89

normal distribution

P(lower)	P(upper)	z	X	mean
.5383	.4617	0.10	350000	331234

actual number of homes > $350 000 =	15

The actual number of homes listed for more than $350 000 is 15, so the probability is 15/50 =

The values should be closer, so we will further examine the list prices to see if the distribution is skewed.
We get the following results from MegaStat.

skewness	1.86

We can see that the distribution is fairly positively skewed. This is not a good approximation as a result.

78.

a.

Descriptive statistics

	10-Aug
Count	16
Mean	318,811.44
population variance	15,121,892,767.25
population standard deviation	122,971.11

normal distribution

P(lower)	P(upper)	z	X	mean	std.dev
.7454	.2546	0.66	400000	318811.4375	122971.105

Aug-10			
$150,320			
$173,918			
$222,597			
$239,688			
$253,940			
$254,298	less than 400 000 =		
$298,588	13/16 =	0.8125	
$299,812			
$303,707	skewness =	1.57	
$305,866			
$322,281		not a good approximation	
$326,550			
$385,712			
$410,995			
$471,929			
$680,782			

b.

Descriptive statistics

	11-Jan
count	16

mean	328,706.38
population variance	19,763,677,314.98
population standard deviation	140,583.35
skewness	1.85
kurtosis	4.84
coefficient of variation (CV)	44.17%

P(lower)	P(upper)	z	X	mean	std.dev
.6940	.3060	0.51	400000	328706.375	140583.347

Jan-11		
$143,876		
$171,788		
$229,716		
$240,646	more than 400 000 =	
$252,141	3/16 =	0.1875
$260,133		
$294,436	skewness =	1.85
$300,353		
$315,483		not a good approximation
$324,598		
$325,732		
$329,640		
$394,655		
$427,159		
$486,384		
$762,562		

CHAPTER 7

SAMPLING METHODS AND THE CENTRAL LIMIT THEOREM

1. a. 303 Louisiana, 5155 S. Main, 3501 Monroe, 2652 W. Central
   b. Answers will vary
   c. 630 Dixie Hwy, 835 S. McCord Rd., 4624 Woodville Rd.
   d. Answers will vary **(LO2)**

2. Cluster **(LO2)**

3. Systematic random sampling **(LO2)**

4. Stratified **(LO2)**

5. a.

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

   b. $\mu_{\bar{X}} = (12+13+14+13+14+15)/6 = 13.5$ $\mu = (12 + 12 + 14 + 16)/4 = 13.5$
   c. More dispersion with population compared to the sample means. The sample means vary from 12 to 15 whereas the population varies from 12 to 16. **(LO3)**

6. a.

Sample	*Values*	*Sum*	*Mean*
1	2,2	4	2
2	2,4	6	3
3	2,4	6	3
4	2,8	10	5
5	2,4	6	3
6	2,4	6	3
7	2,8	10	5
8	4,4	8	4
9	4,8	12	6
10	4,8	12	6

   b. $\mu = (2 + 2 + 4 + 4 + 8)/5 = 4$ $\mu_{\bar{X}} = (2+3+3+5+3+3+5+4+6+6)/10 = 4$
   c. They are equal. The dispersion for the population is greater than that for the sample means. The population varies from 2 to 8, whereas the sample means only vary from 2 to 6. **(LO3)**

7. a.

Sample	*Values*	*Sum*	*Mean*
1	12,12,14	38	12.67
2	12,12,15	39	13.0
3	12,12,20	44	14.67
4	14,15,20	49	16.33
5	12,14,15	41	13.67
6	12,14,15	41	13.67

7	12,15,20	47	15.67
8	12,15,20	47	15.67
9	12,14,20	46	15.33
10	12,14,20	46	15.33

b. $\mu_{\bar{X}} = (12.66 + 13.0 + \ldots + 15.33 + 15.33) / 10 = 14.6$

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

c. The dispersion of the population is greater than that of the sample means. The sample means vary from 12.66 to 16.33 whereas the population varies from 12 to 20. **(LO3)**

8. a.

Sample	*Values*	*Sum*	*Mean*
1	0,0,1	1	0.33
2	0,0,3	3	1.00
3	0,0,6	6	2.00
4	0,1,3	4	1.33
5	0,3,6	9	3.00
6	0,1,3	4	1.33
7	0,3,6	9	3.00
8	1,3,6	10	3.33
9	0,1,6	7	2.33
10	0,1,6	7	2.33

b. $\mu_{\bar{X}} = (0.33 + 1.00 + \ldots + 2.33 + 2.33) / 10 = 2$ $\mu = (0 + 0 + 1 + 3 + 6)/5 = 2$

c. The dispersion of the population is greater than the sample means. The sample means vary from 0.33 to 3.33, the population varies from 0 to 6. **(LO3)**

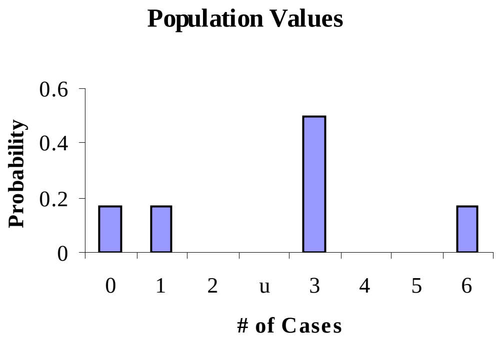
9. a. 20 found by ${}_6C_3$

b.

Sample	*Cases*	*Sum*	*Mean*
Ruud,Austin,Sass	3,6,3	12	4.0
Ruud,Sass,Palmer	3,3,3	9	3.0
Ruud,Palmer,Wilhelms	3,3,0	6	2.0
Ruud,Wilhelms,Schueller	3,0,1	4	1.33
Austin,Sass,Palmer	6,3,3	12	4.0
Austin,Palmer,Wilhelms	6,3,0	9	3.0
Austin,Wilhelms,Schueller	6,0,1	7	2.33
Sass,Palmer,Wilhelms	3,3,0	6	2.0
Sass,Wilhelms,Schueller	3,0,1	4	1.33
Palmer,Wilhelms,Schueller	3,0,1	4	1.33
Austin, Sass, Wilhelms	6,3,0	9	3.00
Ruud,Austin,Palmer	3,6,3	12	4.0
Ruud,Austin,Wilhelms	3,6,0	9	3.0
Ruud,Austin,Schueller	3,6,1	10	3.33
Ruud,Sass,Wilhelms	3,3,0	6	2.0
Ruud,Sass,Schueller	3,3,1	7	2.33
Ruud,Palmer,Schueller	3,3,1	7	2.33
Austin,Sass,Schueller	6,3,1	10	3.33
Austin,Palmer,Schueller	6,3,1	10	3.33
Sass,Palmer,Schueller	3,3,1	7	2.33

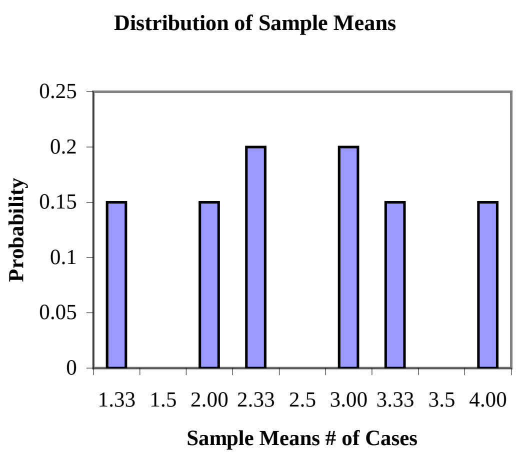
c. $\mu_{\bar{X}} = \frac{53.33}{20} = 2.667$ $\mu = (3 + 6 + 3 + 3 + 1 + 0)/6 = 2.667$ They are equal

d.



Sample Mean	*Number of Means*	*Probability*
1.33	3	0.1500
2.00	3	0.1500
2.33	4	0.2000
3.00	4	0.2000
3.33	3	0.1500
4.00	3	0.1500
	20	1.0000

More of a dispersion in population compared to sample means. The sample means vary from 1.33 to 4.0. The population varies from 0 to 6.



(LO3)

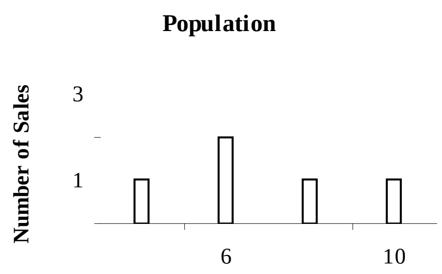
10. a. 10, found by (5!)/3!2!

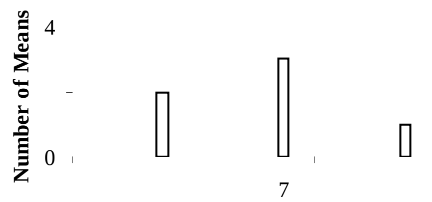
b.

Cars sold	*Sample mean*	*Cars sold*	*Sample mean*
8,6	7	6,10	8
8,4	6	6,6	6
8,10	9	4,10	7
8,6	7	4,6	5
6,4	5	10,6	8

c. 6.8 for population, 6.8 for sample means. They are identical.

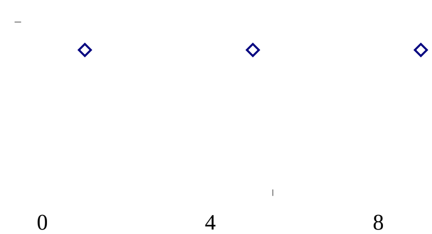
d.





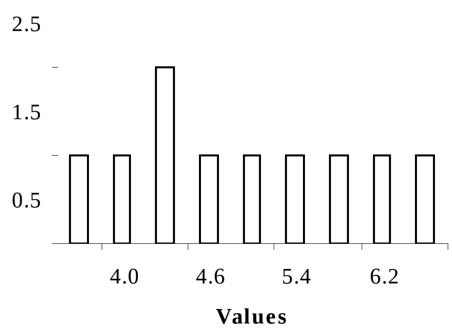
(LO3)

11. a. $\mu = \dfrac{0+1+\ldots+9}{10} = 4.5$



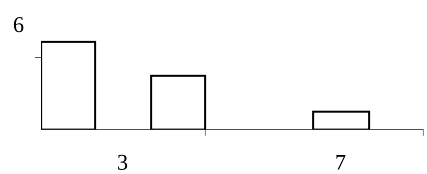
b.

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



The mean of the 10 sample means is 4.84, which is close to the population mean of 4.5. The sample means range from 2.2 to 7.0, whereas the population values range from 0 to 9. From the above graph, the sample means tend to cluster between 4 and 5. **(LO4)**

12. a.



b. $$m = \frac{2+3+\ldots+5}{20} = 3.3$$

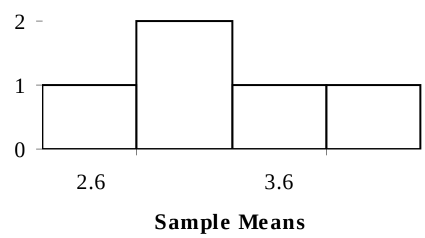
c. Answers will vary, below is one sample

Sample	*Sample Values*	*Sum*	$\overline{X}$
1	2,3,2,3,3	13	2.6
2	3,3,4,2,4	16	3.2
3	3,3,4,4,2	16	3.2
4	3,2,5,5,3	18	3.6
5	3,4,4,2,7	20	4.0

$$m_{\overline{X}} = \frac{2.6+3.2+3.2+3.6+4.0}{5} = 3.32$$

d. Sample mean is very close to the population mean. It is not to be expected that they are exact.

e.



There is less dispersion in the sample means than the population. **(LO4)**

13. a. Answers will vary
b. Answers will vary
c. The sample distribution should be more bell-shaped. **(LO4)**

14. Answers will vary. **(LO4)**

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

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

c. 0.6147, found by 0.3413 + 0.2734. **(LO5)**

16. a. z = $\frac{74-75}{5/\sqrt{40}} = -1.26$ So probability is 0.1038, found by 0.5000 - 0.3962.

b. z = $\frac{76-75}{5/\sqrt{40}} = 1.26$ So probability is 0.7924, found by 2(0.3962).

c. z = $\frac{77-75}{5/\sqrt{40}} = 2.53$ So probability is 0.0981, found by 0.4943 - 0.3962

d. 0.0057, found by 0.5000 - 0.4943 **(LO5)**

17. z = $\frac{950-1200}{250/\sqrt{50}} = -7.07$ So probability is very close to 1 or virtually certain. **(LO5)**

18. a. $s_{\bar{x}} = \frac{80}{\sqrt{40}} = 12.649$

b. z = $\frac{320-330}{80/\sqrt{40}} = -0.79$ So probability is 0.7852, found by 0.2852 + 0.5000.

c. z = $\frac{350-330}{80/\sqrt{40}} = 1.58$ So probability is 0.7281, found by 0.2852 + 0.4429.

d. 0.0571, found by 0.5000 - 0.4429.

d. No assumptions are necessary as n > 30. **(LO5)**

19. $\sqrt{\frac{.45(1-.45)}{200}}$ = 0.035178 **(LO6)**

20. $\sqrt{\frac{.09(1-.09)}{50}}$ = 0.04047 **(LO6)**

21. z = (.06 - .02)/ $\sqrt{\frac{.02(1-.02)}{50}}$ = 2.02; then 0.5 - .4783 = 0.0217 **(LO6)**

22. z = (.35 - .30)/ $\sqrt{\frac{.30(1-.30)}{650}}$ = 2.78; then 0.5 - .4973 = 0.0027 **(LO6)**

23. $z = (.80 - .75)/\sqrt{\frac{.75(1-.75)}{300}}$ = 2; then 0.5 + 0.4772 = 0.9772 **(LO6)**

24. $z = (.20 - .25)/\sqrt{\frac{.25(1-.25)}{500}}$ = -2.58; then 0.5 - 0.4951 = 0.0049 **(LO6)**

25. a. Formal Man, Summit Stationers, Bootleggers, Leather Ltd., Petries
    b. Answers will vary
    c. Gap, Fredericks, Summit, M Studios, Leather ltd, Things Remembered, County Seat, Coach House Gifts, Regis Hairstylists **(LO2)**

26. Answers will vary. One possible answer is to sample every 50th family. This would be a stratified random sample. **(LO2)**

27. The difference between a sample statistic and the population parameter. Yes, the difference could be zero. The sample mean and the population parameter are equal. **(LO4)**

28. 1. Results adequate
    2. Destructive nature of some tests
    3. Physically impossible to check all items
    4. Costly to check all items
    5. Time consuming to check all items **(LO)**

29. a. The standard error of the mean declines as the sample size grows because the sample size is in the denominator and as the denominator increases the proportion decreases.
    b. If the sample size is increased, the Central Limit theorem guarantees the distribution of the sample means becomes more normal.

c. The shape of the distribution becomes narrower since the dispersion is less and estimates of the mean are more precise. **(LO4)**

30. a.

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

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

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

d. Sample means follow a triangular shape peaking at 2. The population is uniform between 1 and 3. **(LO2)**

31. Use of either a proportional or nonproportional stratified random sample would be appropriate. For example, suppose the number of banks in the financial district were as follows:

Assets	*Number*	*Percent of Total*
$500 million and more	20	2.0
$100-499 million	324	32.4
less than $100 million	656	65.6
	1000	100

For a proportional stratified sample, if the sample size is 100, then two banks worth assets of $500 million would be selected, 32 medium-sized banks and 66 small banks. For a nonproportional sample, 10 or even all-20 large banks could be selected and fewer medium and small size banks and the sample results weighted by the appropriate percents of the total. **(LO2)**

32. A simple random sample would be appropriate, but this means each 10-foot length would have to be numbered 1, 2, 3, …, 720. A faster method would be to (1) select a pipe from the first say, 20 pipes produced, and (2) select every 20^{th} pipe produced thereafter and measure its inside diameter. Thus, the sample would include about 36 PVC pipes. **(LO2)**

33. a. We selected 60, 104, 75, 72, and 48. Answers will vary.
b. We selected the third observation. So the sample consists of 75, 72, 68, 82, 48. Answers will vary.
c. Use a stratified random sample. **(LO2)**

34. a. answers will vary
b. answers will vary
c. answers will vary
d. stratified random by gender for example **(LO2)**

35. a. 15 found by ${}_6C_2$

b.

Sample	Value	Sum	Mean
1	79,64	143	71.5
2	79,84	163	81.5
3	79,82	161	80.5
4	79,92	171	85.5
5	79,77	156	78.0
6	64,84	148	74.0
7	64,82	146	73.0
8	64,92	156	78.0
9	64,77	141	70.5
10	84,82	166	83.0
11	84,92	176	88.0
12	84,77	161	80.5
13	82,92	174	87.0
14	82,77	159	79.5
15	92,77	169	84.5
			1195.0

c. $\mu_{\bar{X}} = \frac{1195}{15} = 79.67$ μ=478/6 = 79.67 They are equal

d. No, the student is not graded on all available information. He/she is as likely to get a lower grade based on the sample as a higher grade. By dropping a lower grade, the average is 82.8. This is preferable. **(LO3)**

36. a. 10, found by ${}_5C_2$

b.

Sample	Value	Sum	Mean
1	2,3	5	2.5
2	2,5	7	3.5
3	2,3	5	2.5
4	2,5	7	3.5
5	3,5	8	4.0
6	3,3	6	3.0
7	3,5	8	4.0
8	5,3	8	4.0
9	5,5	10	5.0
10	3,5	8	4.0
			36.0

c. $\mu_{\bar{X}} = 36/10 = 3.6$ μ=18/5 = 3.6 They are equal **(LO3)**

37. a. 10, found by ${}_5C_2$

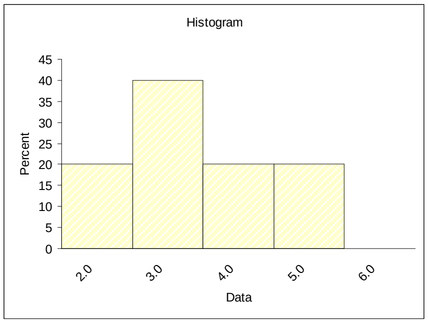
b.

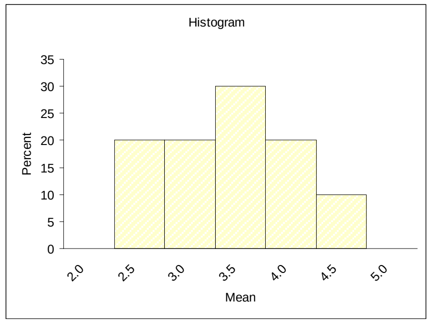
Shutdowns	Mean	Shutdowns	Mean
4,3	3.5	3,3	3.0
4,5	4.5	3,2	2.5
4,3	3.5	5,3	4.0
4,2	3.0	5,2	3.5
3,5	4.0	3,2	2.5

c. $m_{\bar{X}} = (3.5 + 4.5 + ... + 2.5) / 10 = 3.4$ $m = (4 + 3 + 5 + 3 + 2)/5 = 3.4$

The two means are equal.

d. The population values are uniform in shape. The distribution of the sample means tends toward normality. See the following charts. **(LO3)**





38. a. 15, found by ${}_6C_2$

b.

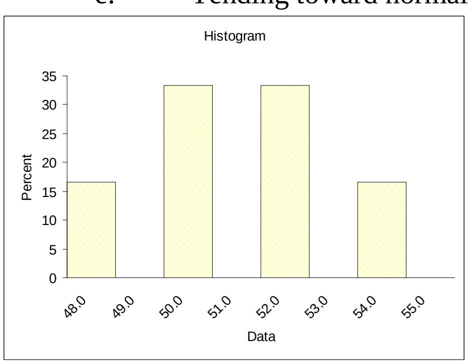
# Sold	*Mean*	*# Sold*	*Mean*
54,50	52	50,52	51
54,52	53	52,48	50
54,48	51	52,50	51
54,50	52	52,52	52
54,52	53	48,50	49
50,52	51	48,52	50
50,48	49	50,52	51
50,50	50		

c.

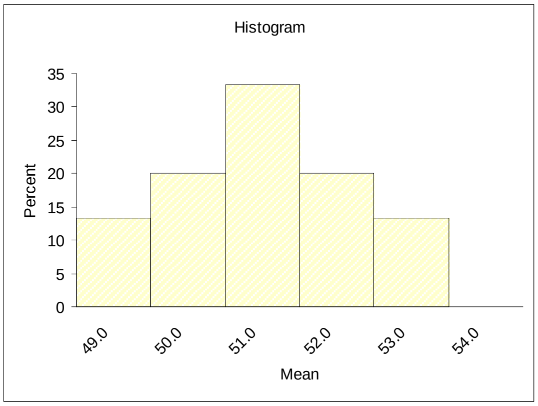
Sample Means	*Frequency*	*Probability*
49	2	0.13
50	3	0.20
51	5	0.33
52	3	0.20
53	2	0.13

d. $m = 51$ $m_{\bar{X}} = 51$

e. Tending toward normal



f. Sample means. Somewhat normal



(LO3)

39. a. The distribution will be normal.

b. $s_{\bar{x}} = \frac{5.5}{\sqrt{25}} = 1.1$

c. $z = \frac{36 - 35}{5.5/\sqrt{25}} = 0.91$ So probability is 0.1814, found by 0.5000 - 0.3186.

d. $z = \frac{34.5 - 35}{5.5/\sqrt{25}} = -0.45$ So probability is 0.6736, found by 0.5000 + 0.1736.

e. 0.4922, found by 0.3186 + 0.1736 **(LO5)**

40. a. The distribution will be normal.

b. $s_{\bar{x}} = \frac{8}{\sqrt{16}} = 2$

c. $z = \frac{140 - 135}{8/\sqrt{16}} = 2.5$ So probability is 0.0062, found by 0.5000 - 0.4938.

d. $z = \frac{128 - 135}{8/\sqrt{16}} = -3.5$ So probability is very close to 1.0.

e. 0.9938, found by 0.5000 + 0.4938. **(LO5)**

41. $z = \frac{335 - 350}{45/\sqrt{40}} = -2.11$ So probability is 0.9826, found by 0.5000 + 0.4826. **(LO5)**

42. a. $s_{\bar{x}} = \frac{40,000}{\sqrt{50}} = 5656.85$

b. The distribution will be normal

c. $z = \frac{112,000 - 110,000}{40,000/\sqrt{50}} = 0.35$ So probability is 0.3632, found by 0.5000 - 0.1368.

d. $z = \frac{100,000 - 110,000}{40,000/\sqrt{50}} = -1.77$ So probability is 0.9616, found by 0.5000 + 0.4616.

e. 0.5984, found by 0.4616 + 0.1368. **(LO5)**

43. $z = \frac{31.5 - 30.6}{2.5/\sqrt{60}}$ = 2.79. So probability is 0.9974, found by 0.5000 + 0.4974. **(LO5)**

44. a. $z_1 = \frac{17 - 18}{3.5/\sqrt{15}} = -1.11$ and $z_2 = \frac{20 - 18}{3.5/\sqrt{15}} = 2.21$

So probability is 0.8529, found by 0.3665 + 0.4864.

b. Since the sample size is small, you assume the population is normally distributed. **(LO5)**

45. Between 2679 and 2721, found by 2700 ± 1.96(68/$\sqrt{40}$). **(LO5)**

46. a. $z = \frac{25 - 23.5}{5/\sqrt{50}} = 2.12$ So probability is 0.0170, found by 0.5000 - 0.4830.

b. $z = \frac{22.5 - 23.5}{5/\sqrt{50}} = -1.41$ So probability is 0.9037, found by 0.4207 + 0.4830.

c. Between 22.33 and 24.67, found by $23.50 \pm 1.65\left(\frac{5}{\sqrt{50}}\right)$. **(LO5)**

47. $z = \frac{900 - 947}{205/\sqrt{60}} = -1.78$ So probability is 0.0375, found by 0.5000 - 0.4625. **(LO5)**

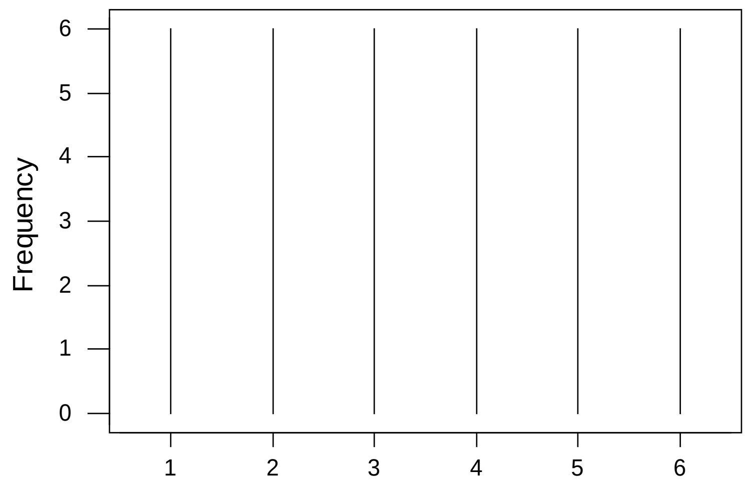
48. a. 36

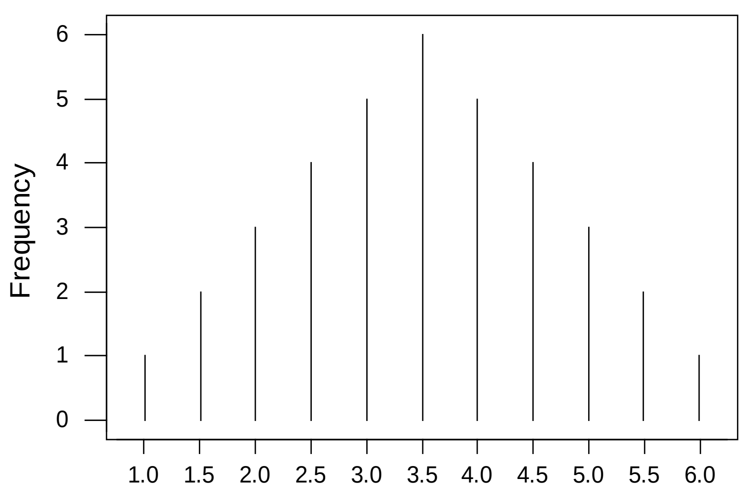
b.

Sample	1st roll	2nd roll	Mean
1	1	1	1
2	1	2	1.5
3	1	3	2
4	1	4	2.5
5	1	5	3
6	1	6	3.5
7	2	1	1.5
8	2	2	2
9	2	3	2.5
10	2	4	3
11	2	5	3.5
12	2	6	4
13	3	1	2
14	3	2	2.5
15	3	3	3
16	3	4	3.5
17	3	5	4
18	3	6	4.5
19	4	1	2.5
20	4	2	3
21	4	3	3.5
22	4	4	4
23	4	5	4.5

24	4	6	5
25	5	1	3
26	5	2	3.5
27	5	3	4
28	5	4	4.5
29	5	5	5
30	5	6	5.5
31	6	1	3.5
32	6	2	4
33	6	3	4.5
34	6	4	5
35	6	5	5.5
36	6	6	6

c.





d. Both means are 3.5. The standard deviation of individual rolls is 1.708, while the standard deviation of sample means is 1.208. **(LO3)**

49. $z = \frac{73 - 69}{12.5/\sqrt{50}}$ = 2.26

Area = 0.0119 found by 0.5 – 0.4881 **(LO5)**

50. z = (.5 - .35)/ $\sqrt{\frac{.35(1-.35)}{42}}$ = 2.04; then 0.5 - .4793 = 0.0207 **(LO6)**

51. z = (.3 - .25)/ $\sqrt{\frac{.25(1-.25)}{200}}$ = 1.63; then 0.5 + 0.4484 = 0.9484 **(LO6)**

52. a. z = (.10 - .05)/ $\sqrt{\frac{.05(1-.05)}{50}}$ = 1.62; then 0.5 - 0.4474 = 0.0526

b. z = (.01 - .05)/ $\sqrt{\frac{.05(1-.05)}{50}}$ = -1.30; then 0.5 - 0.4032 = 0.0968

c. .0526 + .0968 = .1494 **(LO6)**

53. a. z = (.85 - .90)/ $\sqrt{\frac{.90(1-.90)}{300}}$ = -2.892; then 0.5 + 0.4981 = 0.9981

b. z = (.92 - .90)/ $\sqrt{\frac{.90(1-.90)}{300}}$ = 1.15; then 0.5 - 0.3749 = 0.1251

c. 0.4981 + 0.3749 = 0.8730 **(LO6)**

54. a. $s_{\bar{x}} = \frac{2.1}{\sqrt{81}} = 0.2333$

b. $z_1 = \frac{6 - 6.5}{2.1/\sqrt{81}} = -2.14$ and $z_2 = \frac{7 - 6.5}{2.1/\sqrt{81}} = 2.14$

Probability is 0.9676, found by 0.4838 + 0.4838.

c. $z_1 = \frac{6.25 - 6.5}{2.1/\sqrt{81}} = -1.07$ and $z_2 = \frac{6.75 - 6.5}{2.1/\sqrt{81}} = 1.07$

Probability is 0.7154, found by 0.3577 + 0.3577.

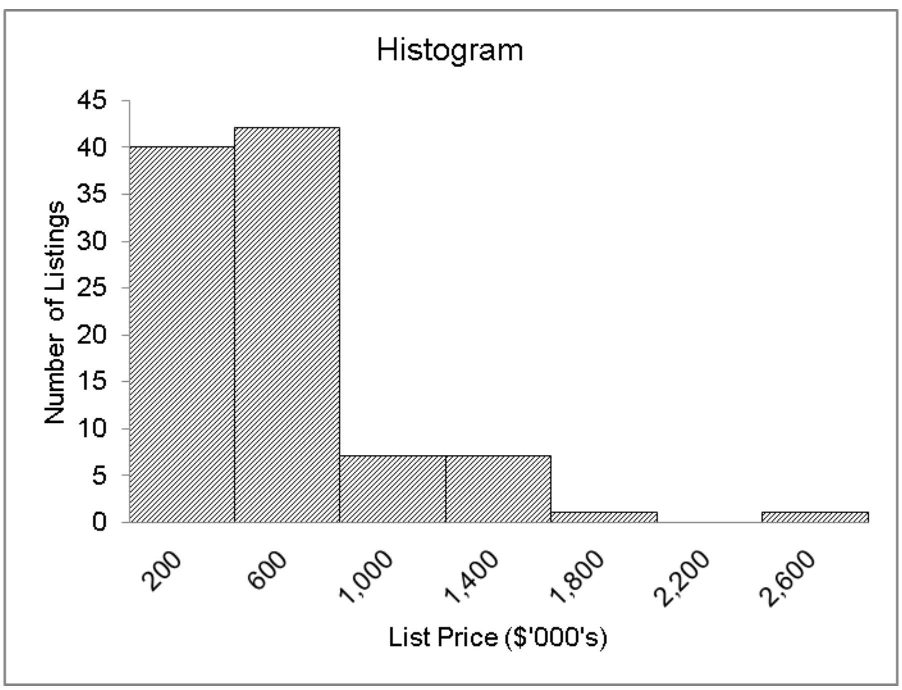
d. Probability is 0.0162, found by 0.5000 - 0.4838. **(LO5)**

55. a.

Part of this answer is from Chapter 2.

Descriptive statistics

	List Price
count	98
mean	567,496.76
population variance	171,676,803,157.51
population standard deviation	414,339.00



The distribution appears to be positively skewed.

b.

Sample list prices	
$	179,000
$	184,900
$	189,000
$	192,779
$	214,900
$	219,500
$	220,000
$	222,400
$	227,000
$	229,000
$	237,500
$	239,000
$	241,745
$	244,900
$	249,000
$	249,999
$	269,000
$	274,900
$	274,900
$	479,900
$	484,000
$	489,900
$	495,000
$	495,500
$	499,000
$	499,900
$	499,900
$	509,900
$	515,000
$	535,000
$	539,900
$	575,000
$	585,000
$	599,900
$	639,900

Descriptive statistics

	Sample list prices
count	35
mean	365,774.94

normal distribution

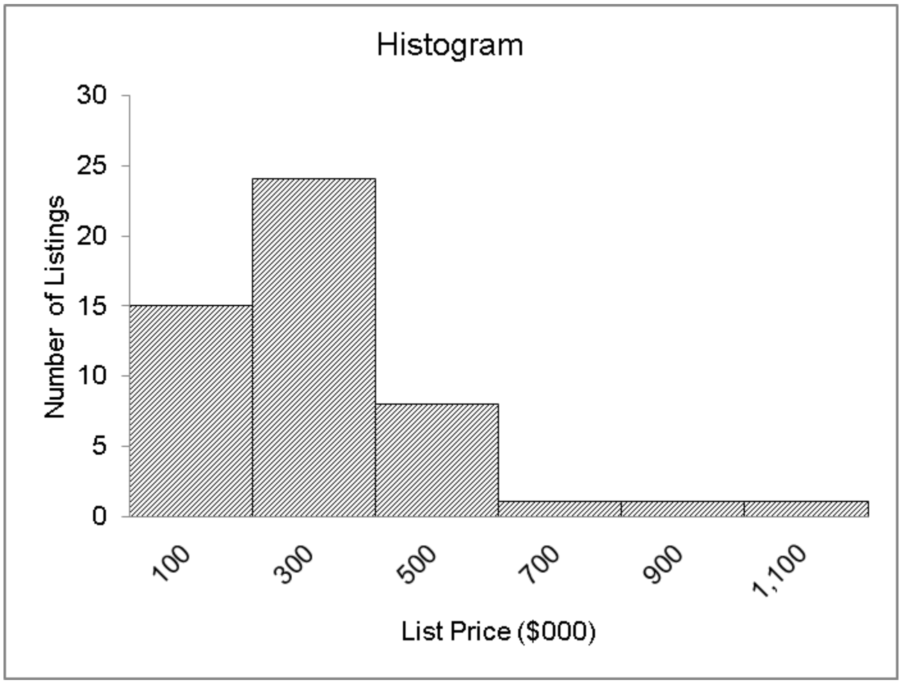
P(lower)	P(upper)	z	X	mean	std.dev
.0020	.9980	-2.88	365774.9	567496.8	70036.07

56. a.

Descriptive statistics

	List Price

count	50
mean	331,233.98
population variance	38,079,899,213.38
population standard deviation	195,140.72



The distribution appears to be positively skewed.

b.

Sample list prices	
$	154,900
$	154,900
$	154,900
$	159,900
$	162,900
$	164,900
$	165,000
$	169,900
$	189,900
$	199,900
$	219,000
$	219,000
$	219,900
$	221,500
$	222,500
$	319,900
$	324,900
$	329,900

Descriptive statistics

	Sample list prices
count	32
mean	294,359.34

$	329,900
$	332,900
$	339,900
$	346,999
$	349,000
$	364,900
$	364,900
$	377,900
$	399,900
$	419,900
$	469,900
$	509,900
$	509,900
$	549,900

normal distribution

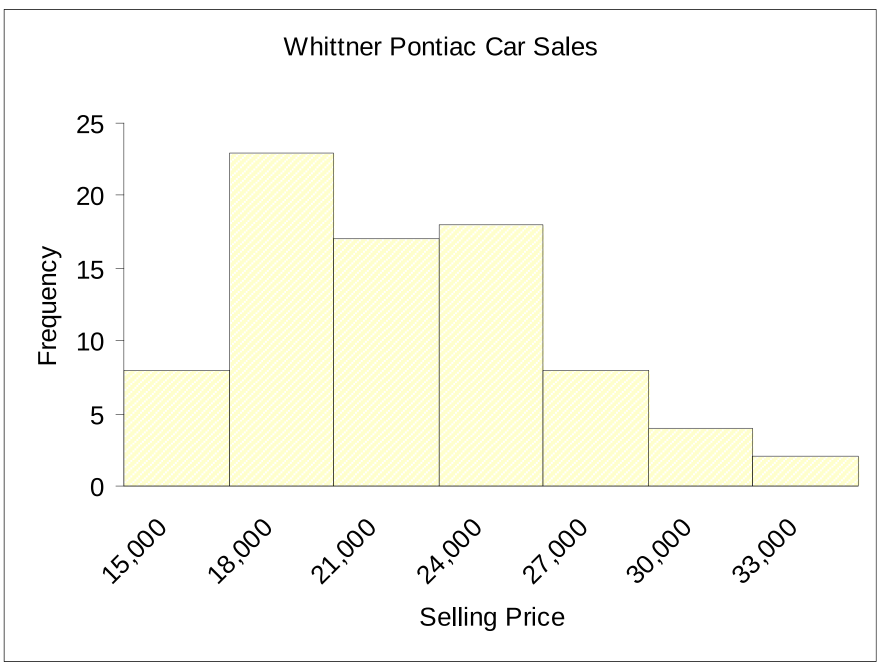
P(lower)	P(upper)	z	X	mean	std.dev
.1425	.8575	-1.07	294359.3	331234	34496.33

Case Study

Answers will vary. A suggested answer follows.

1.

mean	23,218.16
population standard deviation	4,327.14



The distribution follows an approximately normal distribution.

2.

17,891
20,155
20,445
19,331
19,688
20,642
20,895
23,657
20,203
21,639

mean	20,454.60
sample standard deviation	1,507.85

Output is from MegaStat.

normal distribution

p(lower)	p(upper)	z	x	mean	std.dev
.2615	.7385	-0.64	20454.6	23218.16	4327.14

CHAPTER 8

ESTIMATION AND CONFIDENCE INTERVALS

1. 51.314 and 58.686, found by $55 \pm 2.58\left(10/\sqrt{49}\right)$ **(LO3)**

2. 38.911 and 41.089, found by $40 \pm 1.96\left(5/\sqrt{81}\right)$ **(LO3)**

3. a. 1.581, found by $s_{\bar{X}} = 5/\sqrt{10}$
   b. The population is normally distributed and the population variance is known.
   c. 16.901 and 23.099, found by 20 ± 3.099 **(LO3)**

4. a. 3.953, found by $25/\sqrt{40}$
   b. 68.487 and 81.503, found by 75 ± 1.645(3.953)
   c. increase the confidence level **(LO3)**

5. a. 0.95, found by $4.75/\sqrt{25}$
   b. 14.64 to 19.064, found by 16.85 ± 2.33($4.75/\sqrt{25}$)
   c. decrease the confidence level **(LO3)**

6. 1.44 **(LO3)**

7. a. \$20. It is our best estimate of the population mean.
   b. \$18.60 and \$21.40, found by $\$20 \pm 1.96(\$5/\sqrt{49})$ About 95 percent of the intervals similarly constructed will include the population mean. **(LO3)**

8. a. \$18.775 and \$21.225, found by $\$20 \pm 1.96(\$5/\sqrt{64})$
   b. The confidence interval is based on the standard error computed by $s/\sqrt{n}$. As *n*, the sample size, increases (in this case from 49 to 64) the standard error decreases and the confidence interval becomes smaller. **(LO3)**

9. a. 40 litres
   b. 36.669 and 43.331, found by 40 ± 2.58(10 $\sqrt{60}$)
   c. If 100 such intervals were determined, the population mean would be included in about 99 intervals. **(LO3)**

10. 5.29 and 6.81 errors found by $6.05 \pm 1.96(2.44/\sqrt{40})$ **(LO3)**

11. a. 2.201
    b. 1.729
    c. 3.499 **(LO4)**

12. a. 2.145
b. 2.500
c. 1.796 **(LO4)**

13. a. 5.8697, found by 26.25/ $\sqrt{20}$
b. 64.85 to 85.15, found by 75 ± 1.729(5.8697)
c. increase **(LO4)**

14. a. 0.95, found by 4.75/$\sqrt{25}$
b. 14.483 to 19.217, found by 16.85 ± 2.492(0.95)
c. increase **(LO4)**

15. a. The population mean is unknown, but the best estimate is 20, the sample mean.
b. Use the *t* distribution as the population standard deviation is unknown. However, we must assume that the population is normally distributed.
c. 2.093
d. Between 19.06 and 20.94, found by $20 \pm 2.093\left(\frac{2}{\sqrt{20}}\right)$.
e. Neither value is reasonable because they are not inside the interval. **(LO4)**

16. a. The population mean is unknown, but the best estimate is 27, the sample mean.
b. Use the *t* distribution as the standard deviation is unknown and the population is normally distributed.
c. 1.753
d. Between 23.056 and 30.944, found by 27 ± 1.753(9 $\sqrt{16}$).
e. That value is reasonable because it is inside the interval. **(LO4)**

17. Between 95.39 and 101.81, found by $98.6 \pm 1.833\left(\frac{5.54}{\sqrt{10}}\right)$. **(LO4)**

18. Between 30.99 and 39.15, found by $35.07 \pm 2.624\left(\frac{6.02}{\sqrt{15}}\right)$. **(LO4)**

19. a. 0.375, found by 75/200
b. 0.0342, found by $\sqrt{\frac{(.375)(1-.375)}{200}}$ = 0.0342
c. 0.319 to 0.431, found by found by 0.375 ± 1.645(0.3423)
d. If 200 such intervals were determined, the population proportion would be included in about 180 intervals. **(LO5)**

20. a. 73% = 0.73
b. 0.01985, found by $\sqrt{\frac{(.73)(1-.73)}{500}}$ = 0.01985
c. 0.691 to 0.769, found by found by 0.73 ± 1.96(0.01985)

d. If 500 such intervals were determined, the population proportion would be included in about 475 intervals. **(LO5)**

21. a. 0.8, found by 80/100. **(LO5)**

b. 0.04, found by $\sqrt{\frac{0.8(1-0.8)}{100}}$.

c. Between 0.72 and 0.88, found by $0.8 \pm 1.96\left(\sqrt{\frac{0.8(1-0.8)}{100}}\right)$.

d. We are reasonably sure the population proportion is between 72 and 88 percent.

22. a. 0.75, found by 300/400. **(LO5)**

b. 0.0217, found by $\sqrt{\frac{0.75(1-0.75)}{400}}$.

c. Between 0.694 and 0.806, found by $0.75 \pm 2.58\left(\sqrt{\frac{0.75(1-0.75)}{400}}\right)$.

d. We are reasonably sure the population proportion is between 69 and 81 percent.

23. a. 0.625, found by 250/400. **(LO5)**

b. 0.0242, found by $\sqrt{\frac{0.625(1-0.625)}{400}}$.

c. Between 0.563 and 0.687, found by $0.625 \pm 2.58\left(\sqrt{\frac{0.625(1-0.625)}{400}}\right)$.

d. We are reasonably sure the population proportion is between 56 and 69 percent.

24. a. 0.05, found by 15/300.

b. Between 0.025 and 0.075, found by $0.05 \pm 1.96\left(\sqrt{\frac{0.05(1-0.05)}{300}}\right)$.

c. No, because 0.10 is not in the interval. **(LO5)**

25. 33.465 and 36.535, found by $35 \pm 1.96\left(\frac{5}{\sqrt{36}}\right)\sqrt{\frac{300-36}{300-1}}$ **(LO4)**

26. 36.846 and 43.154, found by $40 \pm 2.58\left(\frac{9}{\sqrt{49}}\right)\sqrt{\frac{500-49}{500-1}}$ **(LO4)**

27. 1.689 up to 2.031, found by $1.86 \pm 2.58\left(\frac{0.50}{\sqrt{50}}\right)\sqrt{\frac{400-50}{400-1}}$ **(LO4)**

28. 0.102 to 0.2978, found by $0.2 \pm 1.96\sqrt{\frac{.2(1-.2)}{60}}\sqrt{\frac{900-60}{900-1}}$ **(LO5)**

29. 0.0229 to 0.1104, found by $.066667 \pm 1.645\sqrt{\frac{.0667(1-.0667)}{75}}\sqrt{\frac{500-75}{500-1}}$

10/75 = 0.13; therefore, it would not be reasonable as 0.13 is not in the CI. **(LO5)**

30. 0.43 and 0.77, found by $0.60 \pm 1.96\sqrt{\frac{(.60)(.40)}{30}}\sqrt{\frac{300-30}{300-1}}$ **(LO5)**

31. 97, found by $\left(\frac{1.96 \times 10}{2}\right)^2 = 96.04$ **(LO6)**

32. 60, found by $\left(\frac{2.58 \times 15}{5}\right)^2 = 59.91$ **(LO6)**

33. 196, found by $n = 0.15(0.85)\left(\frac{1.96}{0.05}\right)^2 = 195.9216$ **(LO6)**

34. 165, found by $n = 0.45(0.55)\left(\frac{2.58}{0.10}\right)^2 = 164.75$ **(LO6)**

35. 554, found by $\left(\frac{1.96 \times 3}{0.25}\right)^2 = 553.19$ **(LO6)**

36. 26, found by $\left(\frac{1.96 \times 0.23}{0.09}\right)^2 = 25.09$ **(LO6)**

37. a. 577, found by $0.60(0.40)\left(\frac{1.96}{0.04}\right)^2 = 576.24$

b. 601, found by $0.50(0.50)\left(\frac{1.96}{0.04}\right)^2 = 600.25$ **(LO6)**

38. a. 5683, found by $0.30(0.70)\left(\frac{1.645}{0.01}\right)^2 = 5682.65$ **(LO6)**

b. Increase the allowable error from 0.01 to 0.05 (and/or reduce the confidence level).

Thus the sample size would be reduced to 228, found by 0.30(0.70) $\left(\frac{1.645}{0.05}\right)^2$ = 227.3

39. 6.13 years to 6.87 years, found by $6.5 \pm 1.989(1.7/\sqrt{85})$ **(LO4)**

40. a. 1.401 kg **(LO4)**

b. 1.398 to 1.404 kg, found by $1.401 \pm 2.030(.01/\sqrt{36})$

About 95 percent of similarly constructed intervals would include the population mean.

41. a. Between $1.27 to $1.33, found by 1.30 ±2.680(.07/$\sqrt{50}$)

b. $1.40 is not reasonable because it is outside of the confidence interval. **(LO4)**

42. Between 24.24 and 27.76, found by $26 \pm 2.01\left(\frac{6.2}{\sqrt{50}}\right)$. 28 is not reasonable because it is outside of the confidence interval. **(LO4)**

43. a. Between 7.2 and 8.8, found by $8 \pm 1.685\left(\frac{3}{\sqrt{40}}\right)$.

b. 9 is not reasonable because it is outside of the confidence interval. **(LO4)**

44. 2.54 up to 2.98 meals, found by $2.76 \pm 2.224(0.75/\sqrt{60})$ **(LO4)**

45. a. 65.5 up to 71.7 hours, found by 68.6 ± 2.680(8.2/$\sqrt{50}$)

b. The value suggested by the manager is included in the confidence interval. Therefore it is reasonable.

c. Changing the confidence level to 95% would reduce the width of the interval. The value of 2.608 would change to 2.010. **(LO4)**

46. a. From $1622 to $2018, found by $1820 ± 2.015(660/$\sqrt{45}$) **(LO4)**

b. The population mean could be $1700 because it is in the interval constructed above.

47. 62, found by n = $\left(\frac{1.96 \times 16}{4}\right)^2$.= 61.5 **(LO6)**

48. Between 231.50 and 348.50, found by $290 \pm 2.093\left(\frac{125}{\sqrt{20}}\right)$. **(LO4)**

49. Between $13,734 up to $15,028, found by $14{,}381 \pm 1.711\left(\frac{1892}{\sqrt{25}}\right)$. 15,000 is reasonable because it is inside of the confidence interval. **(LO4)**

50. Between 20.61 and 27.39, found by $24 \pm 2.624\left(\frac{5}{\sqrt{15}}\right)$. **(LO4)**

51. a. \$62.583, found by \$751/12.

b. Between \$60.54 and 64.63, found by $62.583 \pm 1.796\left(\frac{3.94}{\sqrt{12}}\right)$.

c. \$60 is not reasonable because it is outside of the confidence interval. **(LO4)**

52. a. 2408.8, found by 24088/10.

b. Between 2191.06 and 2626.54, found by $2408.8 \pm 2.262\left(\frac{304.4}{\sqrt{10}}\right)$. **(LO4)**

53. a. 89.4667, found by 1342/15.

b. Between 84.99 and 93.94, found by $89.4667 \pm 2.145\left(\frac{8.08}{\sqrt{15}}\right)$.

c. Yes, because even the lower limit of the confidence interval is above 80. **(LO4)**

54. Between 0.641 and 0.739, found by $0.69 \pm 2.58\left(\sqrt{\frac{0.69(1-0.69)}{600}}\right)$. **(LO5)**

55. Do not use PCF as n < 5%N. Between 0.647 and 0.753, found by $.7 \pm 2.58\left(\sqrt{\frac{0.7(1-0.7)}{500}}\right)$.

Yes, because even the lower limit of the confidence interval is above 0.500. **(LO5)**

56. a. p = 560/1000 = 0.560, from 0.53 up to 0.59, found by $0.56 \pm 1.96\sqrt{\frac{0.56(0.44)}{1000}}$

b. The lower point of the interval is greater than 0.50. So we can conclude the majority feel the President is doing a good job. **(LO5)**

57. \$52.56 and \$55.44, found by $\$54.00 \pm 1.96\frac{\$4.50}{\sqrt{35}}\sqrt{\frac{(500-35)}{500-1}}$ **(LO4)**

58. 0.345 and 0.695, found by $0.52 \pm 2.58\sqrt{\frac{(0.52)(0.48)}{50}}\sqrt{\frac{(650-50)}{650-1}}$ **(LO5)**

59. 369, found by $n = 0.60(1-0.60)[1.96/0.05]^2$ **(LO6)**

60. 133, found by $\{(1.645 \times 14)/2\}^2 = 132.595$ **(LO6)**

61. 97, found by $\left(\frac{1.96 \times 500}{100}\right)^2$ **(LO6)**

62. 865, found by $0.10(0.90)\left(\frac{1.96}{0.02}\right)^2$ **(LO6)**

63. a. 708.13, rounded up to 709, found by: $0.21(1-0.21)[1.96/0.03]^2$

b. 1068, found by $0.50(0.50)[1.96/0.03]^2$ **(LO6)**

64. a. 25%, found by 25/100

b. 0.172 to 0.328, found by $0.25 \pm 1.96\sqrt{\frac{(0.25)(0.75)}{100}}\sqrt{\frac{605-100}{604}}$ The interval is for those who do not use. Since 0.40 is not in the interval, then at least 0.40 will use the ATM.

c. 1.65, found by 165/100

d. 1.387 to 1.913, found by $1.65 \pm 1.96\frac{1.4659}{\sqrt{100}}\sqrt{\frac{505}{604}}$ Note: s = 1.4659

e. No, because 0 is not in the interval between 1.387 and 1.91. **(LO4)**

65. Between 0.573 and 0.653, found by $.613 \pm 2.58\left(\sqrt{\frac{0.613(1-0.613)}{1000}}\right)$. Yes, because even the lower limit of the confidence interval is above 0.500. **(LO5)**

66. Between \$29,690 and \$34,310, found by 32000 ± 1.690(8200/$\sqrt{36}$) **(LO4)**

67. Between 12.69 and 14.11, found by $13.4 \pm 1.96\left(\frac{6.8}{\sqrt{352}}\right)$. **(LO4)**

68. 2185, found by $.35(1-.35)\left(\frac{1.96}{.02}\right)^2$. **(LO5)**

69. a. 30, found by $180/\sqrt{36}$

b. Between \$355.10 and 476.90 mpg, found by $416 \pm 2.030\left(\frac{180}{\sqrt{36}}\right)$

c. 1245, found by $\left(\frac{1.96*180}{10}\right)^2$ **(LO6)**

70. a. Between 7849 and 8151, found by $8000 \pm 2.756\left(\frac{300}{\sqrt{30}}\right)$.

b. 554, found by $\{(1.96 \times 300)/25\}^2 = 553.2$ **(LO6)**

71. a. Answers will vary. **(LO4&5)**

Descriptive statistics

	List Price
Mean	764,432.08
sample standard deviation	588,551.67

Confidence interval – mean

95% confidence level

764432.075	Mean
588551.6667	std. dev.
40	N
2.023	t (df = 39)
188227.953	half-width
952,660.03	upper confidence limit
576,204.12	lower confidence limit

true mean	567,496.76

The true mean is not in the 95% confidence interval.

b. Answers will vary.

	Total Square Feet
Mean	2,134.53
sample standard deviation	1,534.76

Confidence interval – mean

95%	confidence level
2134.525	mean
1534.760651	std. dev.
40	n
2.023	t (df = 39)
490.840	half-width
2,625.37	upper confidence limit
1,643.68	lower confidence limit

true mean	1,710.01

The true mean is in the 95% confidence interval.

c. Answers will vary

	3 + Bedrooms
sample proportion	0.450

Confidence interval – proportion

95%	confidence level
0.45	proportion
40	n
1.960	z
0.154	half-width
0.60	upper confidence limit
0.30	lower confidence limit

true proportion	0.184

The true proportion is not in the 95% confidence interval.

72. This answer was obtained using MegaStat. Your answer may be slightly different due to rounding. **(LO4)**

a.

Descriptive statistics

	10-Aug
Mean	318,811.44
sample standard deviation	127,004.01

Confidence interval – mean

97%	confidence level
318811.4375	Mean
127004.0116	std. dev.
16	N
2.397	t (df = 15)
76,107.3138	half-width
394,918.75	upper confidence limit
242,704.12	lower confidence limit

b. It is unlikely that $400 000 is the population mean as the value is not in the confidence interval.

73. This answer was obtained using MegaStat. Your answer may be slightly different due to rounding. **(LO4)**

a.

Descriptive statistics

	11-Jan
Count	16
Mean	328,706.38
sample variance	21,081,255,802.65
sample standard deviation	145,193.86
confidence interval 80.% lower	280,044.45
confidence interval 80.% upper	377,368.30
half-width	48,661.92

b. It is possible that $300 000 is the population mean as the value is in the confidence interval.

74. This answer was obtained using MegaStat. Your answer may be slightly different due to rounding. **(LO4&5)**

a. Answers will vary

	List Price
Mean	241,830.00
sample standard deviation	72,719.67

Confidence interval – mean

90%	confidence level
241830	mean
72719.67156	std. dev.
20	n
1.729	t (df = 19)
28116.759	half-width
269946.76	upper confidence limit
213713.24	lower confidence limit

true mean	331,233.98

The true mean does not fall in the confidence interval.

b.

	Total Square Feet
Mean	1,013.80
sample standard deviation	178.61

Confidence interval – mean

90%	confidence level
1013.8	mean
178.614728	std. dev.
20	n
1.729	t (df = 19)
69.061	half-width
1082.86	upper confidence limit
944.74	lower confidence limit

true mean	1,250.78

The true mean does not fall in the confidence interval.

c.

	< 2 bedrooms
sample proportion	0

Confidence interval - proportion

90%	confidence level
0	proportion
20	n
1.645	z
0.000	half-width
0.00	upper confidence limit
0.00	lower confidence limit

true proportion	0.20

There is no confidence interval. The sample that we took did not contain any listings with less than 2 bedrooms.

CHAPTER 9

ONE-SAMPLE TESTS OF A HYPOTHESIS

1. a. $H_0 = 2$
   $H_1 \neq 2$
   b. $H_0 \geq 40$
   $H_1 < 40$
   c. $H_0 \leq \$1750$
   $H_1 > \$1750$
   d. $H_0 = 3.25$
   $H_1 \neq 3.25$ **(LO3)**

2. a. $H_0 \geq 1.5$
   $H_1 < 1.5$
   b. $H_0 \geq 1500$
   $H_1 < 1500$
   c. $H_0 = 2000$
   $H_1 \neq 2000$
   d. $H_0 = 500$
   $H_1 \neq 500$ **(LO3)**

3. a. Two-tailed
   b Reject H_o and accept H_1 when z does not fall in the region from – 1.96 and 1.96
   c. - 1.2, found by $z = \frac{49 - 50}{(5/\sqrt{36})}$
   d. Not enough evidence to reject H_o and conclude the mean is not different from 50.
   e. p = 0.2302, found by 2(0.5000 – 0.3849). A 23.02% chance of finding a z value this large when H_o is true. **(LO3,4&7)**

4. a. One-tail **(LO3,4&7)**
   b. Reject H_o when $z > 2.05$
   c. 4, found by $z = \frac{12 - 10}{(3/\sqrt{36})}$
   d. Reject H_o and conclude that the mean is greater than 10
   e. p-value is very close to 0, given a z value of 4.00. Very little chance H_o is true.

5. a. One-tailed
   b. Reject H_o and accept H_1 when $z > 1.645$
   c. 1.2, found by $z = \frac{21 - 20}{(5/\sqrt{36})}$
   d. Not enough evidence to reject H_o at the 0.05 significance level.
   e. p = 0.1151, found by 0.5000 – 0.3849. An 11.51% chance of finding a z-value this large or larger. **(LO3,4&7)**

6. a. One-tailed
   b. Reject H_o and accept H_1 where $z < -1.88$
   c. -2.67, found by $z = \dfrac{215 - 220}{(15/\sqrt{64})}$
   d. Reject H_o and conclude that the population mean is less than 220 at the 0.03 significance level.
   e. $p = 0.0038$, found by 0.5000 – 0.4962. Less than 0.5% chance H_o is true. **(LO3,4&7)**

7. a. H_o: $\mu = 96\,600$
   H_1: $\mu \neq 96\,600$
   b. Reject H_o if $z < -1.96$ or $z > 1.96$
   c. -0.69, found by $z = \dfrac{95795 - 96600}{8050/\sqrt{48}}$
   d. There is not enough evidence to reject H_o
   e. $p = 0.4902$, found by 2(0.5000 – 0.2549). Crosset's experience is not different from that claimed by the manufacturer. If the H_o is true, the probability of finding a value more extreme than this is 0.4902. **(LO3,4&7)**

8. a. H_o: $\mu \geq 3$ H_1: $\mu < 3$; choose $\alpha = 0.05$
   b. Reject H_o if $z < -1.645$
   c. -1.77, found by $z = \dfrac{2.75 - 3.0}{(1/\sqrt{50})}$
   d. Reject H_o
   e. $p = 0.0384$, found by (0.5000 – 0.4616). We conclude that the mean waiting time is less than three minutes. When H_o is true, the probability of obtaining a value smaller than -1.77 is 0.0384. **(LO3,4&7)**

9. a. H_o: $\mu \geq 6.8$ H_1: $\mu < 6.8$
   b. Reject H_o if $z < -2.33$
   c. -7.2, found by $z = \dfrac{6.2 - 6.8}{(0.5/\sqrt{36})}$
   d. Reject H_o at a significance level of .01.
   e. *p*-value is almost zero; the mean number of DVD's watched is less than 6.8 per month. If H_o is true, there is virtually no chance of getting a statistic this small. **(LO3,4&7)**

10. a. H_o: $\mu \leq 80$ H_1: $\mu > 80$
    b. Reject H_o if $z > 2.33$
    c. 8.49, found by $z = \dfrac{\$84.65 - 80.00}{\$3.24/\sqrt{35}}$
    d. Reject H_o at a significance level of .01.
    e. *p*-value is almost zero; the mean amount of tips per day is larger than \$80.00. If H_o is true, the probability of obtaining a sample mean this far above 80 is virtually zero. **(LO3,4&7)**

11. a. Reject H_o when $t > 1.833$

b. $t = \dfrac{12 - 10}{(3/\sqrt{10})} = 2.108$

c. Reject H_o; the mean is greater than 10. **(LO3&4)**

12. a. Reject H_o if $t < -3.106$ or $t > 3.106$

b. $t = \dfrac{407 - 400}{(6/\sqrt{12})} = 4.042$

c. Reject H_o, the mean does not equal 400. **(LO3&4)**

13. Choose α = .05

H_o: $\mu \leq 40$ H_1: $\mu > 40$; Reject H_o if $t > 1.703$

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

Reject H_o and conclude that the mean number of calls is greater than 40 per week. **(LO3&4)**

14. H_o: $\mu \geq 42.3$ H_1: $\mu < 42.3$ Reject H_o if $t < -1.319$

$t = \dfrac{40.6 - 42.3}{(2.7/\sqrt{24})} = -3.084$

Reject H_o. The mean assembly time is less than 42.3 minutes. **(LO3&4)**

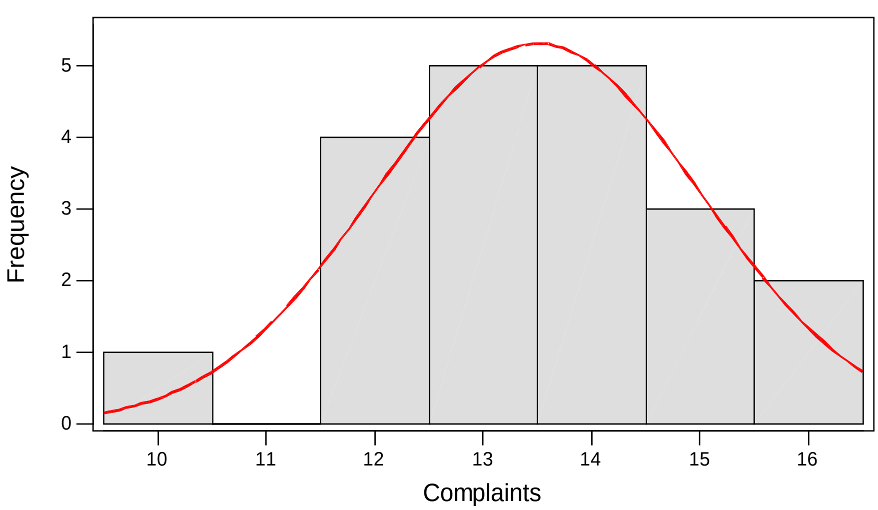
15. H_o: $\mu \leq 35\,600$ H_1: $\mu > 35\,600$ Reject H_o if $t > 1.740$

$t = \dfrac{37675 - 35600}{2415/\sqrt{18}} = 3.645$

Reject H_o and conclude that as the mean life of the spark plugs is greater than 35 600 km, the claim is true. **(LO3&4)**

16. a. The population of complaints follows a normal probability distribution.

Histogram of Complaints, with Normal Curve



b. The assumption of normality appears reasonable.

c. $H_0: \mu \geq 15$ $H_1: \mu < 15$ Reject H_0 if $t < -1.729$

$$t = \frac{(13.5 - 15)}{(1.504/\sqrt{20})} = -4.46$$

Reject the null hypothesis. The mean number of complaints is less than 15.

17. a. Reject H_o if $t < -3.747$

b. $\overline{X} = 17$ and $s = \sqrt{\dfrac{1495 - \dfrac{(85)^2}{5}}{5-1}} = 3.536$ $t = \dfrac{17-20}{3.536/\sqrt{5}} = -1.90$

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

d. Between 0.05 and 0.10, about 0.065; by computer, the *p*-value = .0639. **(LO3,4&7)**

18. a. Reject H_o if $t < -2.571$ or $t > 2.571$

b. $t = \dfrac{111.667 - 100}{6.055/\sqrt{6}} = 4.72$

c. Reject H_o. The population mean is not equal to 100

d. less than 0.01 (between 0.001 and 0.01); using a computer, the *p*-value = .0052

e. 105.31 to 118.02; the hypothesized mean does not fall in the confidence interval, and so H_o is rejected. **(LO3,4&7)**

19. H_o: $\mu \leq 4.35$ H_1: $\mu > 4.35$ Reject H_o if $t > 2.821$

$$t = \frac{4.368 - 4.35}{(0.0339/\sqrt{10})} = 1.68$$

Do not reject H_o. The additive did not increase the mean weight of the chickens. The *p*-value is between 0.10 and 0.05. Using a computer, the *p*-value = .0639 **(LO3,4&7)**

20. H_o: $\mu \leq 2160$ H_1: $\mu > 2160$ Reject H_o if $t > 2.306$

$$t = \frac{2172.44 - 2160}{(9.3823/\sqrt{9})} = 3.98$$

Reject H_o. The mean chlorine shelf life has increased. The *p*-value is less than 0.005. Using a computer, the *p*-value = .0020. **(LO3,4&7)**

21. H_o: $\mu \leq 4.0$ H_1: $\mu > 4.0$ Reject H_o if $t > 1.796$

$$t = \frac{4.50 - 4.0}{(2.68/\sqrt{12})} = 0.65$$

Do not reject H_o. The mean number of kms travelled has not been shown to be greater than 4.0. The *p*-value is greater than 0.10. Using a computer, the *p*-value = .2657. **(LO3,4&7)**

22. H_o: $\mu \leq 53$ H_1: $\mu > 53$ Reject H_o if $t > 1.761$

$$t = \frac{56.4 - 53.0}{(3.7378/\sqrt{15})} = 3.52$$

Reject H_o. The mean number of surveys conducted is greater than 53. The *p*-value is less than 0.005. Using a computer, the *p*-value = .0017. **(LO3,4&7)**

23. a. H_o is rejected if $z > 1.645$

b. 1.09, found by $z = \dfrac{0.75 - 0.70}{\sqrt{\dfrac{0.70(0.30)}{100}}}$

c. H_o is not rejected **(LO5)**

24. a. H_o is rejected if $z < -1.96$ or $z > 1.96$

b. -2.24, found by $z = \dfrac{0.30 - 0.40}{\sqrt{\dfrac{0.40(0.60)}{120}}}$

c. H_o is rejected **(LO5)**

25. a. H_o: $p \leq 0.52$ H_1: $p > 0.52$

b. H_o is rejected if $z > 2.33$

c. 1.62, found by $z = \dfrac{0.5667 - 0.52}{\sqrt{\dfrac{0.52(0.48)}{300}}}$

d. H_o is not rejected. We cannot conclude that the proportion of men driving on Hwy 400 is larger than 0.52. **(LO5)**

26. H_o: $p \leq 0.33$ H_1: $p > 0.33$ H_o is rejected if $z > 2.05$

2.00, found by $z = \dfrac{0.40 - 0.3333}{\sqrt{\dfrac{0.3333(0.6667)}{200}}}$

H_o is not rejected. The proportion of students with jobs is not larger. **(LO5)**

27. a. H_o: $p \geq 0.90$ H_1: $p < 0.90$

b. H_o is rejected if $z < -2.33$

c. -2.67, found by $z = \dfrac{0.82 - 0.90}{\sqrt{\dfrac{0.90(0.10)}{100}}}$

d. H_o is rejected. Less than 90% of the customers receive their orders in less than 10 minutes. **(LO5)**

28. Choose $\alpha = .05$

H_o: $p \geq 0.50$ H_1: $p < 0.50$

H_o is rejected if $z < -1.645$

-0.40, found by $z = \dfrac{0.48 - 0.50}{\sqrt{\dfrac{0.50(0.50)}{100}}}$

H_o is not rejected. The proportion of student changing their major has not changed. **(LO5)**

29. H_o: $\mu = \$60{,}000$ H_1: $\mu \neq \$60{,}000$

Reject H_o if $z < 1.645$ or $z > 1.645$

$$z = \frac{62\,500 - 60\,000}{6\,000/\sqrt{120}} = 4.56$$

Reject H_o. We can conclude that the mean salary is not \$60,000. The *p*-value is very close to zero. The confidence interval is from \$61 599 to \$63 401. As \$60 000 falls outside of the confidence interval, the hypothesized mean is rejected. **(LO3,4&7)**

30. a. H_o: $\mu \geq 50$ H_1: $\mu < 50$ Reject H_o if $z \leq -2.33$ $\overline{X} = 48.18$

$$z = \frac{48.18 - 50}{3/\sqrt{10}} = -1.92$$

H_o is not rejected. The mean weight is not less than 50 pounds.

b. Mr. Rutter can use the *z* distribution as the test statistic because the population follows a normal distribution and the population standard deviation ($\sigma = 3$) is known.

c. *p*-value = 0.5000 – 0.4726 = 0.0274 (by computer *p*-value = .0275) **(LO3,4&7)**

31. H_o: $\mu \geq 10$ H_1: $\mu < 10$

Reject H_o if $z < -1.645$

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

Reject H_o. The mean weight loss is less than 10 pounds. The *p*-value = 0.5000 – 0.4943 = 0.0057. **(LO3,4&7)**

32. H_o: $\mu \le 450$ H_1: $\mu > 450$

Reject H_o if $z > 1.645$ $$z = \frac{451.4 - 450}{.85/\sqrt{50}} = 11.65$$

Reject H_o. The cans are being overfilled. The p-value is very close to 0. **(LO3,4&7)**

33. H_o: $\mu \ge 7$ H_1: $\mu < 7$

$$t = \frac{(6.8 - 7)}{0.9/\sqrt{50}} = -1.57$$

The p-value is 0.0613, so at a significance level of 5%, do not reject the null hypothesis. University students sleep no less than the typical adult male. **(LO3,4&7)**

34. H_o: $\mu \le 90$ H_1: $\mu > 90$ Reject H_o if $t > 1.323$

$$t = \frac{94 - 90}{22 - \sqrt{100}} = 1.82$$

Reject H_o. At the 0.10 level we can conclude that the vacation selling time has increased. The p-value is .0360. **(LO3,4&7)**

35. H_o: $\mu \le 1.25$ H_1: $\mu > 1.25$ Reject H_o if $t > 1.691$

$$z = \frac{1.27 - 1.25}{0.05/\sqrt{35}} = 2.37$$

Reject H_o. The mean price of gasoline is greater than \$1.25. The p-value = .0119 (by computer) **(LO3,4&7)**

36. H_o: $\mu = 40$ H_1: $\mu \ne 40$ Reject H_o if $t < -1.671$

$$t = \frac{37.8 - 40}{12.2/\sqrt{60}} = -1.40$$

H_o is not rejected. The p-value = .1677 (by computer). We cannot conclude that the mean leisure time is untrue. **(LO3,4&7)**

37. H_o: $\mu \le 14$ H_1: $\mu > 14$ Reject H_o if $t > 2.821$

$$t = \frac{15.64 - 14}{1.561/\sqrt{10}} = 3.32$$

Reject H_o. The mean rate charged is greater than 14 percent. **(LO3&4)**

38. $H_0: \mu \ge 6$ $H_1: \mu < 6$ Reject H_0 if $t < -2.998$ [Assuming the population is normally distributed]

$$t = \frac{5.6375 - 6}{0.6346/\sqrt{8}} = -1.62$$

Do not reject H_0. The mean rate could be 6.0 percent. The p-value is 0.075. **(LO3,4&7)**

39. H_o: $\mu = 1.92$ H_1: $\mu \neq 1.92$ Reject H_o if $t < -2.201$ or $t > 2.201$

$\bar{X} = 2.08667$ $s_d = 0.40484$

$$t = \frac{2.08667 - 1.92}{.40484/\sqrt{12}} = 1.43$$

Do not reject H_o. There is not a difference in the mean amount of water consumed at the college surveyed and the national average.
The confidence interval is from 1.82944 to 2.34389. The hypothesized mean falls in the confidence interval, and so H_o is not rejected. **(LO3&4)**

40. H_o: $\mu \leq 25$ H_1: $\mu > 25$ Reject H_o if $t > 2.624$

$\bar{X} = 26.07$ $s = 1.5337$

$$t = \frac{26.07 - 25.00}{1.5337/\sqrt{15}} = 2.702$$

Reject H_o. The mean number of patients per day is more than 25. The *p*-value is less than 0.01. Using a computer, the *p*-value is .0087. This is lower than the significance level, and so, the null hypothesis is rejected. **(LO3,4&7)**

41. H_o: $\mu \geq 6.5$ H_1: $\mu < 6.5$ Reject H_o if $t < -2.718$ **(LO3&4)**

$\bar{X} = 5.1667$ $s = 3.1575$

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

Do not reject H_o. The mean is not less than 6.5.

42. H_o: $\mu \geq 3.5$ H_1: $\mu < 3.5$ Reject H_o if $t < -1.746$

$$t = \frac{2.9553 - 3.5}{0.5596/\sqrt{17}} = -4.013$$

Reject H_o. The mean time to complete a game is less than 3.5 hours. The *p*-value (by computer) = .0005. **(LO3,4&7)**

43. H_o: $\mu = 0$ H_1: $\mu \neq 0$ Reject H_o if $t < -2.110$ or $t > 2.110$

$\bar{X} = -0.2322$ $s = 0.3120$

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

Reject H_o. The mean gain or loss does not equal 0. The *p*-value is less than 0.01, but greater than 0.001, so the probability of no time gain or loss is very small. Using a computer, the *p*-value = .0057, and as this is less than the significance level, the null hypothesis is rejected. **(LO3,4&7)**

44. H_o: $\mu \leq 4.5\%$ H_1: $\mu > 4.5\%$ Reject H_o if $t > 1.796$

$\bar{X} = 4.5717$ $s = 0.2405$

$$t = \frac{4.5717 - 4.50}{0.2405/\sqrt{12}} = 1.033$$

Do not reject H_o. The mean number rate of return is not more than 4.5%. **(LO3&4)**

45. $H_0: \mu \leq 100$ $H_1: \mu > 100$ Reject H_0 if $t > 1.761$.

[Assuming the population is normally distributed]

$$t = \frac{109.4 - 100}{9.96/\sqrt{15}} = 3.655$$

Reject H_0. The mean is greater than 100. **(LO3&4)**

46. $H_o: \mu \le 267$ $H_1: \mu > 267$ Reject H_o if $t > 2.681$

$$t = \frac{288.31 - 267}{22.46/\sqrt{13}} = 3.421$$

Reject the null. The fare is higher. p-value < 0.005; using a computer, the p-value = .0025, and as this is less than the significance level, the null hypothesis is rejected. **(LO3&4)**

47. H_o: $p \le 0.60$ H_1: $p > 0.60$ H_o is rejected if $z > 2.33$

$$z = \frac{0.70 - 0.60}{\sqrt{\frac{0.60(0.40)}{200}}} = 2.89$$

H_o is rejected. Ms. Dennis is correct. More than 60% of the accounts are more than 3 months old. **(LO5)**

48. H_o: $p \le 0.55$ H_1: $p > 0.55$ H_o is rejected if $z > 1.645$

$$z = \frac{0.60 - 0.55}{\sqrt{\frac{0.55(0.45)}{70}}} = \frac{0.05}{0.0595} = 0.841$$

H_o is not rejected. We cannot conclude that more than 55% of the commuters would use the route. **(LO5)**

49. H_o: $p \le 0.44$ H_1: $p > 0.44$ H_o is rejected if $z > 1.645$

$$z = \frac{0.480 - 0.44}{\sqrt{\frac{0.44(0.56)}{1000}}} = 2.55$$

H_o is rejected. We conclude that there has been an increase in the proportion of people wanting to go to Europe. **(LO5)**

50. H_o: $p \le 0.10$ H_1: $p > 0.10$ H_o is rejected if $z > 1.645$

$$z = \frac{\frac{9}{50} - 0.10}{\sqrt{\frac{0.10(0.90)}{50}}} = 1.89$$

(LO5)

H_o is rejected. More than 10 percent of the sets need repair. p-value = 0.5000 – 0.4706 = 0.0294

51. H_o: $p \le 0.20$ H_1: $p > 0.20$ H_o is rejected if $z > 2.33$

$$z = \frac{\frac{56}{200} - 0.20}{\sqrt{\frac{0.20(0.80)}{200}}} = 2.83$$

H_o is rejected. More than 20 percent of the owners move during a particular year. The p-value = 0.5000 – 0.4977 = 0.0023 **(LO5)**

52. H_o: $\mu \ge 10$ H_1: $\mu < 10$ Reject H_o if $t < -1.895$

$$\bar{X} = \frac{78.3}{8} = 9.7875 \qquad s = \sqrt{\frac{5.889}{8-1}} = 0.9172$$

$$t = \frac{9.7875 - 10}{0.9172/\sqrt{8}} = -0.655$$

Do not reject H_o at a significance level of .05. We cannot conclude that the cost is not less than $10,000. **(LO3&4)**

53. H_o: $\mu \le 42$ H_1: $\mu > 42$ Reject H_o if $t > 1.796$

$$t = \frac{51 - 42}{8/\sqrt{12}} = 3.90$$

Reject H_o. The mean time for delivery is more than 42 days. The *p*-value is less than 0.005; using a computer, the *p*-value = .0012. **(LO3,4&7)**

54. H_o: $\mu \le 9$ H_1: $\mu > 9$ Reject H_o if $t > 2.998$

$\bar{X} = 9.488$ $s = 0.467$

$$t = \frac{9.488 - 9.00}{0.467/\sqrt{8}} = 2.95$$

Do not reject H_o. The mean waiting time does not exceed 9 minutes. The *p*-value is less than 0.025; using a computer, the *p*-value = .0107. **(LO3,4&7)**

55. H_o: $\mu \ge 8$ H_1: $\mu < 8$ Reject H_o if $t < -1.714$

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

Do not reject the null. The time is not less. **(LO3&4)**

56. H_o: $p = 0.50$ H_1: $p \ne 0.50$ Reject H_o if z is not between –1.96 and 1.96

$$z = \frac{0.482 - 0.500}{\sqrt{(0.5)(0.5)/1002}} = -1.14$$

Do not reject the null. There is not enough evidence to indicate that the results have changed. **(LO5)**

57. **(LO3,4&7)**

a)

251,338.03	confidence interval 95.% lower
406,074.72	confidence interval 95.% upper

b)

H_0: $\mu = 450\ 000$

H_1: $\mu \ne 450\ 000$

Hypothesis Test: Mean vs. Hypothesized Value

450,000.000	hypothesized value
328,706.375	mean Jan-11
145,193.856	std. dev.
36,298.464	std. error
16	n
15	df
-3.34	t

.0045	p-value (two-tailed)

Since the *p*-value < .02, the null hypothesis is rejected,
so we conclude that the mean could not be $450 000.

c)

H_0: μ = 275 000

H_1: μ ≠ 275 000

Since the value of $275 000 falls within the limits of the 95% CI from part a), there is not enough evidence to reject H_0 so we conclude that the mean could be $275 000.

d)

Hypothesis Test: Mean vs. Hypothesized Value

275,000.000	hypothesized value
328,706.375	mean Jan-11
145,193.856	std. dev.
36,298.464	std. error
16	N
15	Df
1.48	T
.1597	p-value (two-tailed)

Since the *p*-value > .05, there is not enough evidence to reject the null hypothesis,
so we conclude that the mean could be $275 000.

58. **(LO3,4, 5&7)**

a)

H_0: μ ≤ 600 000

H_1: μ > 600 000

Hypothesis Test: Mean vs. Hypothesized Value

600,000.000	hypothesized value
567,496.755	mean List Price
416,469.287	std. dev.
42,069.751	std. error
98	N
97	Df
-0.77	T
.7792	p-value (one-tailed, upper)

The *p*-value > .01, so there is not enough evidence to reject the null hypothesis.
We conclude that the list price is not greater than $600 000.

b)

H_0: μ ≤ 2100
H_1: μ > 2100

Hypothesis Test: Mean vs. Hypothesized Value

2,100.000	hypothesized value
1,710.010	mean Total Square Feet
1,103.162	std. dev.
111.436	std. error
98	N
97	Df
-3.50	T
.9996	p-value (one-tailed, upper)

Note: the *p*-value is very large, so the possibility of the null hypothesis being true is a certainty.
The *p*-value > .01, so there is not enough evidence to reject the null hypothesis.
We conclude that the mean number of square feet is less than or equal to 2100.

c)

# of apartments = 58 proportion = 58/98 =.5918

H_0: p ≤ .60
H_1: p > .60

Observed	*Hypothesized*	
0.5898	0.6	p (as decimal)
58/98	59/98	p (as fraction)
57.8	58.8	X
98	98	n
	0.0495	std. error
	-0.16	z
	.5816	p-value (one-tailed, upper)

The *p*-value > .05, so there is not enough evidence to reject the null hypothesis.
We conclude that the proportion ≤ .60.

Observed	*Hypothesized*	
0.5918	0.6	p (as decimal)
58/98	59/98	p (as fraction)
58.	58.8	X
98	98	n
	0.0495	std. error
	-0.16	z

.5655	p-value (one-tailed, upper)

d)

$H_{0:}$ p ≤ .60

$H_{1:}$ p > .60

with more than 2 bedrooms = 37 proportion = .378

Observed	Hypothesized	
0.378	0.6	p (as decimal)
37/98	59/98	p (as fraction)
37.044	58.8	X
98	98	n
	0.0495	std. error
	-4.49	z
	1.0000	p-value (one-tailed, upper)

Note: the *p*-value is very large, so the possibility of the null hypothesis being true is a certainty.
There is not enough evidence to reject the the null hypothesis.
We conclude that the proportion ≤ .60.

59. **(LO3,4, 5&7)**

a)

H_0: μ = 350 000

H_1: μ ≠ 350 000

Hypothesis Test: Mean vs. Hypothesized Value

350,000.000	hypothesized value
331,233.980	mean List Price
197,121.891	std. dev.
27,877.245	std. error
50	n
49	df
-0.67	t
.5040	p-value (two-tailed)
256,524.302	confidence interval 99.% lower
405,943.658	confidence interval 99.% upper
74,709.678	margin of error

The *p*-value > .01, so there is not enough evidence to reject the null hypothesis.
We conclude that the list price is $350 000.
The mean of $350 000 falls in the 99% CI, which confirms that we do not have enough evidence to reject the null hypothesis.

b)

H_0: μ = 1500

H_1: μ ≠ 1500

Hypothesis Test: Mean vs. Hypothesized Value

1,500.000	hypothesized value
1,250.780	mean Total Square Feet
730.875	std. dev.
103.361	std. error
50	n
49	df
-2.41	t
.0197	p-value (two-tailed)
1,043.068	confidence interval 95.% lower
1,458.492	confidence interval 95.% upper
207.712	margin of error

The *p*-value < .05, so we reject the null hypothesis.
We conclude that the list price is not 1500 square feet.
The mean of 1500 does not fall in the 95% CI, which confirms that we should reject the null hypothesis.

c)

\# of townhouses = 2 proportion = 0.04

H_0: p ≤ .10

H_1: p > .10

Observed	*Hypothesized*	
0.04	0.1	p (as decimal)
2/50	5/50	p (as fraction)
2.	5.	X
50	50	n
	0.0424	std. error
	-1.41	z
	.9214	p-value (one-tailed, upper)

Note: the *p*-value is very large, so the possibility of the null hypothesis being true is a certainty.
There is not enough evidence to reject the the null hypothesis.
We conclude that the proportion ≤ .10.

d)

$H_{0:}$ p ≤ .50

$H_{1:}$ p > .50

with more than 2 bedrooms = 23 proportion = .46

Observed	*Hypothesized*	
0.46	0.5	p (as decimal)
23/50	25/50	p (as fraction)
23.	25.	X
50	50	n
	0.0707	std. error
	-0.57	z
	.7142	p-value (one-tailed, upper)

Note: the *p*-value is large, so the possibility of the null hypothesis being true is good.
There is not enough evidence to reject the the null hypothesis.
We conclude that the proportion ≤ .50.

60. **(LO3,4, &7)**

a)

234,240.34	confidence interval 98.% lower
423,172.41	confidence interval 98.% upper

b)

$H_{0:}$ μ = 300 000

$H_{1:}$ μ ≠ 300 000

Hypothesis Test: Mean vs. Hypothesized Value

300,000.00000	hypothesized value
328,706.37500	mean Jan-11
145,193.85594	std. dev.
36,298.46398	std. error
16	n
15	df
0.79	t
.4414	p-value (two-tailed)

the *p*-value > .05, so we do not have enough evidence to reject the null hypothesis and conclude that the mean list price is = $300 000.

c)

$H_{0:}$ μ = 400 000
$H_{1:}$ μ ≠ 400 000

234,240.34	confidence interval 98.% lower
423,172.41	confidence interval 98.% upper

Since the hypothesized population mean of $400 000 does fall in the 98% CI, there is not enough evidence to reject the null hypothesis and conclude that the population mean is likely $400 000.

CHAPTER 10

TWO-SAMPLE TESTS OF HYPOTHESIS

1. a. Two-tailed test
   b. Reject H_o if $z < -2.05$ or $z > 2.05$
   c. 2.59, found by $z = \dfrac{102 - 99}{\sqrt{\dfrac{5^2}{40} + \dfrac{6^2}{50}}}$
   d. Reject H_o
   e. $p = 0.0096$, found by 2(0.5000 – 0.4952) **(LO1)**

2. a. One-tailed test
   b. Reject H_o if $z > 1.41$
   c. 0.607 found by $z = \dfrac{2.67 - 2.59}{\sqrt{\dfrac{(0.75)^2}{65} + \dfrac{(0.66)^2}{50}}}$
   d. Not enough evidence to reject H_o
   e. $p = 0.2709$, found by 0.5000 – 0.2291 **(LO1)**
      $p = 0.2719$ by computer

3. Step 1: H_o: $\mu_1 - \mu_2 \geq 0$ H_1: $\mu_1 - \mu_2 < 0$ **(LO1)**
   Step 2: The 0.05 significance level was chosen
   Step 3: Reject H_o if $z < -1.645$
   Step 4: -0.83, found by $z = \dfrac{3.5 - 3.7}{\sqrt{\dfrac{(1.05)^2}{40} + \dfrac{(1.3)^2}{55}}}$
   Step 5: There is not enough evidence to reject H_o. Babies using the Gibbs brand did not gain less weight. $p = 0.2033$ found by 0.5000 – 0.2967, which is > than the significance level, and so, there is not enough evidence to reject H_o. Using a computer, the p-value = .2037.

4. Step 1: H_o: $\mu_c - \mu_p = 0$ H_1: $\mu_c - \mu_p \neq 0$
   Step 2: The 0.05 significance level was chosen
   Step 3: Reject H_o if z less than -1.96 or greater than 1.96
   Step 4: -1.43, found by $z = \dfrac{595 - 610}{\sqrt{\dfrac{(48)^2}{35} + \dfrac{(42)^2}{40}}}$
   Step 5: not enough evidence to reject H_o. There is no difference in the mean number of kms traveled per month. The p-value is 0.126 found by 2(0.5000 – 0.4370), which is > than the significance level, and so, there is not enough evidence to reject H_o. Using a computer, the p-value = .1525. **(LO1)**

5. Two-tailed test. Because we are trying to show a difference exists between two means.
   H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$
   Reject H_o if $z < -2.58$ or $z > 2.58$

-2.66, found by $z = \dfrac{31.4 - 34.9}{\sqrt{\dfrac{(5.1)^2}{32} + \dfrac{(6.7)^2}{49}}}$ Reject H_o at the 0.01 level. There is a difference in the mean turnover rate. The *p*-value = 2(.5-.4961) = .0078. Since this is < than the significance level, we reject H_o **(LO1)**

6. H_o: $\mu_1 - \mu_2 \le 0$ H_1: $\mu_1 - \mu_2 > 0$ If $z > 2.05$, reject H_o.

$z = \dfrac{20.75 - 19.80}{\sqrt{\dfrac{(2.25)^2}{40} + \dfrac{(1.90)^2}{45}}} = 2.09$ Reject H_o. It is reasonable to conclude union nurses earn more.

The *p*-value is 0.0183, which is < than the significance level, and therefore, H_o is rejected. **(LO1)**

7. a. H_o is rejected if $z > 1.645$

b. 0.64, found by $p_c = \dfrac{70 + 90}{100 + 150}$

c. 1.61, found by $z = \dfrac{0.70 - 0.60}{\sqrt{\dfrac{(0.64)(0.36)}{100} + \dfrac{(0.64)(0.36)}{150}}}$

d. There is not enough evidence to reject $H_{o.}$

e. *p*-value = 0.5 – 0.4463 = 0.0537. The *p*-value is 0.0537, which is > than the significance level, and therefore, there is not enough evidence to reject $H_{o.}$ **(LO2)**

8. a. H_o is rejected if $z < -1.96$ or $z > 1.96$

b. 0.80, found by $p_c = \dfrac{170 + 110}{200 + 150}$

c. 2.70, found by $z = \dfrac{0.85 - 0.7333}{\sqrt{\dfrac{(0.80)(0.20)}{200} + \dfrac{(0.80)(0.20)}{150}}}$

d. H_o is rejected

e. *p*-value = 2(0.5 – 0.4965) = 0.0070. The *p*-value is < than the significance level, and therefore, H_o is rejected. **(LO2)**

9. a. H_o: $p_1 - p_2 = 0$ H_1: $p_1 - p_2 \ne 0$

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

c. $\bar{p}_c = \dfrac{24 + 40}{400 + 400} = 0.08$

d. -2.09, found by $z = \dfrac{0.06 - 0.10}{\sqrt{\dfrac{(0.08)(0.92)}{400} + \dfrac{(0.08)(0.92)}{400}}}$

e. H_o is rejected. The proportion infested is not the same in the two fields. The *p*-value = 2(0.5 – 0.4817) = 0.0366. The *p*-value is < than the significance level, and therefore, H_o is rejected. (computer *p*-value = .0371) **(LO2)**

10. a. H_o: $p_1 - p_2 \geq 0$ H_1: $p_1 - p_2 < 0$

b. 0.05 is chosen as the significance level. H_o is rejected if $z < -1.645$

$$\bar{p}_c = \frac{1530 + 2010}{3000 + 3000} = 0.59$$

c. -12.60, found by $z = \dfrac{0.51 - 0.67}{\sqrt{\dfrac{(0.59)(0.41)}{3000} + \dfrac{(0.59)(0.41)}{3000}}}$

d. H_o is rejected. The proportion of women who think men are thoughtful has declined.

e. the *p*-value is very close to 0, which is much smaller than the significance level, and therefore, H_o is rejected. **(LO2)**

11. H_o: $p_d - p_r \leq 0$ H_1: $p_d - p_r > 0$

H_o is rejected if $z > 2.05$ $\bar{p}_c = \dfrac{168 + 200}{800 + 1000} = 0.2044$

$$z = \frac{0.21 - 0.20}{\sqrt{\dfrac{(0.2044)(0.7956)}{800} + \dfrac{(0.2044)(0.7956)}{1000}}} = 0.52$$

There is not enough evidence to reject H_o. There is no difference in the proportion of Conservatives and Liberals who favor lowering the standards. The *p*-value = 0.5 – 0.1985 = 0.3015. The *p*-value is > than the significance level, and therefore, there is not enough evidence to reject H_o. **(LO2)**

12. H_o: $p_s - p_m = 0$ H_1: $p_s - p_m \neq 0$

H_o is rejected if $z < -1.96$ or $z > 1.96$ $\bar{p}_c = \dfrac{120 + 150}{400 + 600} = 0.27$

$$z = \frac{0.30 - 0.25}{\sqrt{\dfrac{(0.27)(0.73)}{400} + \dfrac{(0.27)(0.73)}{600}}} = 1.74$$

There is not enough evidence to reject H_o. There is no difference in the proportion of married and single drivers who have accidents. The *p*-value = 2(0.5 – 0.4591) = 0.0818. The *p*-value is > than the significance level, and therefore, there is not enough evidence to reject H_o. **(LO2)**

13. a. Reject H_o if $t > 2.120$ or $t < -2.120$ $df = 10 + 8 - 2 = 16$ **(LO1)**

b. $s_p^2 = \dfrac{(10-1)(4)^2 + (8-1)(5)^2}{10 + 8 - 2} = 19.9375$

c. $t = \dfrac{23 - 26}{\sqrt{19.9375\left(\dfrac{1}{10} + \dfrac{1}{8}\right)}} = -1.416$

d. There is not enough evidence to reject H_o.

e. *p*-value is greater than 0.10 and less than 0.20. The actual *p*-value = 0.1758.

14. a. Reject H_o if $t > 1.697$ or $t < -1.697$ $df = 17 + 15 - 2 = 30$ **(LO1)**

b. $$s_p^2 = \frac{(15-1)(12)^2 + (17-1)(15)^2}{15+17-2} = 187.20$$

c. $$t = \frac{350-342}{\sqrt{187.2\left(\frac{1}{15}+\frac{1}{17}\right)}} = 1.651$$

d. Do not reject H_o.

e. p-value is greater than 0.10 and less than 0.20, and so is > than the significance level; therefore, there is not enough evidence to reject H_o. The actual p-value = 0.1093.

15. H_o: $\mu_f - \mu_m \le 0$ H_1: $\mu_f - \mu_m > 0$ $df = 9 + 7 - 2 = 14$ Reject H_o if $t > 2.624$

$$s_p^2 = \frac{(7-1)(6.88)^2 + (9-1)(9.49)^2}{9+7-2} = 71.749$$

$$t = \frac{79-78}{\sqrt{71.749\left(\frac{1}{7}+\frac{1}{9}\right)}} = 0.234$$

There is not enough evidence to reject H_o. The mean grade of women is not higher than that of men. The p-value = .4090, which is > than the significance level and supports the decision. The actual p-value = 0.4091. **(LO1)**

16. H_o: $\mu_s - \mu_d \le 0$ H_1: $\mu_s - \mu_d > 0$ $df = 15 + 12 - 2 = 25$ Reject H_o if $t > 2.485$

$$s_p^2 = \frac{(15-1)(15.5)^2 + (12-1)(18.1)^2}{15+12-2} = 278.69$$

$$t = \frac{61-48.4}{\sqrt{278.69\left(\frac{1}{15}+\frac{1}{12}\right)}} = 1.949$$

There is not enough evidence to reject H_o. There is no difference in the mean amount of time spent watching television. The p-value is between 0.025 and 0.050, which confirms that there is not enough evidence to reject H_o. The actual p-value = 0.0313. **(LO1)**

17. H_o: $\mu_s - \mu_a \le 0$ H_1: $\mu_s - \mu_a > 0$ $df = 6 + 7 - 2 = 11$ Reject H_o if $t > 1.363$

$$s_p^2 = \frac{(6-1)(12.2)^2 + (7-1)(15.8)^2}{6+7-2} = 203.82$$

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

Reject H_o. The mean daily expenses are greater for the sales staff. The p-value is between 0.05 and 0.10, which is < than the significance level, and so H_o is rejected. The actual p-value = 0.0763. **(LO1)**

18. H_o: $\mu_t - \mu_n \le 0$ H_1: $\mu_t - \mu_n > 0$ $df = 12 + 8 - 2 = 18$ Reject H_o if $t > 2.552$

$$s_p^2 = \frac{(12-1)(22.8)^2 + (8-1)(34.4)^2}{12+8-2} = 777.88$$

$$t = \frac{526.8-535.8}{\sqrt{777.88\left(\frac{1}{8}+\frac{1}{12}\right)}} = -0.707$$

There is not enough evidence to reject H_o. The mean salary of nurses is higher. The p-value is greater than 0.10, which is > than the significance level, and so there is not enough evidence to reject H_o. The actual p-value = 0.2445. **(LO1)**

19. a. Reject H_o if $t > 2.353$

b. $\bar{d} = \frac{4}{4} = 1$ $s_d = \sqrt{\frac{38 - 4^2/4}{3}} = 3.367$

c. $t = \frac{1}{3.367/\sqrt{4}} = .59$

d. There is not enough evidence to reject the H_o. There is no difference in the defective parts produced on the day or afternoon shift. (The actual p-value = .2971) **(LO3)**

20. a. Reject H_o if $t < -2.776$ or $t > 2.776$

b. $\bar{d} = \frac{23}{5} = 4.6$ $s_d = \sqrt{\frac{115 - \frac{(23)^2}{5}}{4}} = 1.52$

c. $t = \frac{4.6}{1.52/\sqrt{5}} = 6.767$

d. Reject the H_o. There is a difference in the mean number of citations given by the two officers.

e. p-value is less than 0.01, but greater than 0.001. (The actual p-value = .0025) **(LO3)**

21. H_o: $\mu_d \le 0$ H_1: $\mu_d > 0$ Reject H_o if $t > 1.796$ **(LO3)**

$\bar{d} = 25.917$ $s_d = 40.791$

$t = \frac{25.917}{40.791/\sqrt{12}} = 2.20$ Reject H_o. The incentive plan resulted in an increase in daily income. The p-value is 0.250, which is < than the significance level, and so reject H_o.

22. H_o: $\mu_d \ge 0$ H_1: $\mu_d < 0$ Reject H_o if $t < -2.998$

$\bar{d} = 3.625$ $s_d = 4.8385$

$t = \frac{3.625}{4.8385/\sqrt{8}} = 2.12$ Do not reject H_o. The p-value is about 0.035 **(LO3)**

23. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \ne 0$ Reject H_0 if $t < -2.645$ or $t > 2.645$

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

$t = \frac{24.51 - 22.69}{\sqrt{17.31\left(\frac{1}{35} + \frac{1}{40}\right)}} = 1.87$

There is not enough evidence to reject the null hypothesis. There is no difference in the means. The p-value is 0.0627, which is > than the significance level, and confirms that there is not enough evidence to reject H_o. **(LO1)**

24. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \ne 0$ Reject H_0 if $t < -2.011$ or $t > 2.011$

$s_p^2 = \frac{(25-1)(5)^2 + (25-1)(3)^2}{25+25-2} = 17.0$

$$t = \frac{30 - 28}{\sqrt{17.0\left(\frac{1}{25} + \frac{1}{25}\right)}} = 1.71$$

There is not enough evidence to reject the null hypothesis. The mean of the 2 programs is not different. The p-value (.0928) is > than the significance level (< 0.05) which supports the decision. **(LO1)**

25. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Reject H_o if $z < -2.58$ or $z > 2.58$

$$z = \frac{36.2 - 37.0}{\sqrt{\frac{(1.14)^2}{35} + \frac{(1.30)^2}{40}}} = -2.84$$

Reject H_o. There is a difference in the useful life of the two brands of paint. The p-value is 0.0046, found by 2(0.5000 – 0.4977). Since the p-value is < than the significance level, H_o is rejected. **(LO1)**

26. H_o: $\mu_1 - \mu_2 \geq 0$ H_1: $\mu_1 - \mu_2 < 0$ Reject H_o if $z < -1.645$

$$z = \frac{345 - 351}{\sqrt{\frac{(21)^2}{54} + \frac{(28)^2}{60}}} = -1.30$$

There is not enough evidence to reject H_o. There is not enough evidence to conclude that more units are produced on the afternoon shift. (The p-value is .0964) **(LO1)**

27. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Reject H_o if $z < -1.96$ or $z > 1.96$

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

There is not enough evidence to reject H_o. There is no difference in the mean number of calls. The p-value = 2(0.5000 – 0.3508) = 0.2984, which is > than the significance level, and so, there is not enough evidence to reject H_o. **(LO1)**

28. H_o: $\mu_1 - \mu_2 \geq 0$ H_1: $\mu_1 - \mu_2 < 0$ Reject H_o if $z < -2.33$

$$z = \frac{4.35 - 5.84}{\sqrt{\frac{(1.20)^2}{50} + \frac{(1.36)^2}{40}}} = -5.44$$

H_o is rejected. The mean number of cups for regular coffee drinkers is less. The p-value is almost zero. **(LO1)**

29. H_0: $\mu_A - \mu_B \geq 0$ H_1: $\mu_A - \mu_B < 0$ Reject H_0 if $t < -1.668$

$$s_p^2 = \frac{(40-1)(9{,}200)^2 + (30-1)(7{,}100)^2}{40 + 30 - 2} = 70{,}041{,}912 \qquad t = \frac{57000 - 61000}{\sqrt{70041912\left(\frac{1}{40} + \frac{1}{30}\right)}} = -1.98$$

Reject H_0. The mean income of those selecting Plan B is larger. The p-value is 0.0259, which is smaller than the significance level, and confirms that the null hypothesis should be rejected. The coefficients of skewness do not affect the results. The sample sizes are both at least 30. So the Central Limit Theorem ensures the sample means follow a normal distribution. (Computer *p*-value = .0200) **(LO1)**

30. H_0: $\mu_1 - \mu_2 \le 0$ H_1: $\mu_1 - \mu_2 > 0$ Reject H_0 if t > 1.665

$$s_p^2 = \frac{(35-1)(4.2)^2 + (45-1)(3.9)^2}{35+45-2} = 16.27$$

$$t = \frac{18-15.5}{\sqrt{16.27\left(\frac{1}{35}+\frac{1}{45}\right)}} = 2.75$$

Reject H_0. Software issues take longer on average. The *p*-value is 0.0037, which is smaller than the significance level, and confirms that the null hypothesis should be rejected. (Computer *p*-value = .0032) **(LO1)**

31. H_o: $p_1 - p_2 \le 0$ H_1: $p_1 - p_2 > 0$ Reject H_o if $z > 1.645$

$$p_c = \frac{180+261}{200+300} = 0.882 \qquad z = \frac{0.90 - 0.87}{\sqrt{\frac{0.882(0.118)}{200} + \frac{0.882(0.118)}{300}}} = 1.019$$

There is not enough evidence to reject H_o. There is no difference in the proportions that found relief in the new and the old drugs. The p-value is 0.5 – 0.3461 = 0.1539, which is > than the significance level, and confirms that there is not enough evidence to reject H_o. (Computer p-value = .1542) **(LO2)**

32. H_o: $p_1 - p_2 \ge 0$ H_1: $p_1 - p_2 < 0$ If $z < -1.645$, reject H_o.

$$p_c = \frac{160+170}{300+290} = 0.56 \qquad z = \frac{0.5333 - 0.5862}{\sqrt{\frac{(0.56)(0.44)}{300} + \frac{(0.56)(0.44)}{290}}} = -1.29$$

There is not enough evidence to reject the null. We cannot conclude an increased proportion believe the economy is expanding. The p-value is 0.5 – 0.4015 = 0.0985, which is > than the significance level, and confirms that there is not enough evidence to reject H_o. (Computer p-value = .0980) **(LO2)**

33. H_o: $p_1 - p_2 \le 0$ H_1: $p_1 - p_2 > 0$ If $z > 2.33$, reject H_o.

$$p_c = \frac{990+970}{1500+1600} = 0.63 \qquad z = \frac{0.6600 - 0.60625}{\sqrt{\frac{(0.63)(0.37)}{1500} + \frac{(0.63)(0.37)}{1600}}} = 3.10$$

Reject the null. We can conclude the proportion of men who believe the division is fair is greater. The p-value is virtually zero. (Computer p-value = .0010) **(LO2)**

34. $H_o : p_1 - p_2 = 0$ $H_1 : p_1 - p_2 \neq 0$ If z is not between -2.58 and 2.58, reject H_o.

$$p_c = \frac{450+352}{500+400} = 0.891 \qquad z = \frac{0.90 - 0.88}{\sqrt{\frac{(0.891)(0.109)}{500} + \frac{(0.891)(0.109)}{400}}} = 0.96$$

There is not enough evidence to reject the null. There is no difference in the percentages. The p-value = 2(0.5 – 0.3315) = 0.3370, and confirms that there is not enough evidence to reject H_o. (Computer p-value = .3385) **(LO2)**

35. H_o: $p_1 - p_2 = 0$ $H_1 : p_1 - p_2 \neq 0$

H_o is rejected if $z < -1.96$ or $z > 1.96$

$$p_c = \frac{68+45}{98+85} = 0.617$$

$$z = \frac{0.6939 - 0.5294}{\sqrt{\frac{(0.617)(0.383)}{98} + \frac{(0.617)(0.383)}{85}}} = 2.28$$

H_o is rejected. The proportions are not the same. **(LO2)**

36. H_o: $p_1 - p_2 = 0$ $H_1 : p_1 - p_2 \neq 0$

H_o is rejected if $z < -1.96$ or $z > 1.96$

$$\bar{p}_c = \frac{56+52}{270+203} = 0.22833$$

$$z = \frac{0.2074 - 0.2562}{\sqrt{\frac{(0.228)(0.772)}{270} + \frac{(0.228)(0.772)}{203}}} = -1.25$$

H_o is not rejected. The proportions are the same. **(LO2)**

37. H_o: $p_1 - p_2 \geq 0$ H_1: $p_1 - p_2 < 0$ If $z < -1.645$, reject H_o. **(LO2)**

$$p_c = \frac{13+7}{180+67} = 0.081 \qquad z = \frac{0.07222 - 0.10448}{\sqrt{\frac{(0.081)(0.919)}{180} + \frac{(0.081)(0.919)}{67}}} = -0.826$$

Do not reject the null. We cannot determine the 32-bit machines are inferior to the 8-bit ones.

38. H_o: $p_1 - p_2 \leq 0$ H_1: $p_1 - p_2 > 0$ If $z > 1.645$, reject H_o.

$$p_c = \frac{39+11}{388+307} = 0.07194 \qquad z = \frac{0.10052 - 0.03583}{\sqrt{\frac{(0.07194)(0.92806)}{388} + \frac{(0.07194)(0.92806)}{307}}} = 3.277$$

Reject the null.
The Banking Services division has significantly more female top executives. **(LO2)**

39. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Reject H_o if $t < -2.086$ or $t > 2.086$

$$s_p^2 = \frac{(10-1)(10.5)^2 + (12-1)(14.25)^2}{10+12-2} = 161.2969 \qquad t = \frac{83.55 - 78.8}{\sqrt{161.2969\left(\frac{1}{10} + \frac{1}{12}\right)}} = 0.874$$

There is not enough evidence to reject H_o. There is no difference in the mean number of hamburgers sold at the two locations. The p-value is > than 0.20, which is > than the significance level, and so the decision is supported. (Computer p-value = .3928) **(LO1)**

40. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ If t is between -2.086 and 2.086, reject H_o.

$$s_p^2 = \frac{(10-1)(4.44)^2 + (12-1)(2.68)^2}{10+12-2} = 12.82 \qquad t = \frac{27.46 - 25.69}{\sqrt{12.82\left(\frac{1}{10} + \frac{1}{12}\right)}} = 1.15$$

There is not enough evidence to reject H_o. The mean waiting times are not different.
P-value = .2617 **(LO1)**

41. H_o: $\mu_1 - \mu_2 \leq 0$ H_1: $\mu_1 - \mu_2 > 0$ Reject H_o if $t > 2.567$

$$s_p^2 = \frac{(8-1)(2.2638)^2 + (11-1)(2.4606)^2}{8+11-2} = 5.672 \qquad t = \frac{10.375 - 5.636}{\sqrt{5.672\left(\frac{1}{8} + \frac{1}{11}\right)}} = 4.28$$

Reject H_o. The mean number of transaction by the young adults is more than for the senior citizens. P-value = .0003 **(LO1)**

42. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Reject H_o if $t > 2.086$ or $t < -2.086$

$\bar{X}_1 = 12.17$ $s_1 = 1.0563$ $\bar{X}_2 = 14.875$ $s_2 = 2.2079$

$$s_p^2 = \frac{(10-1)(1.0563)^2 + (12-1)(2.2079)^2}{10+12-2} = 3.1832 \qquad t = \frac{12.17 - 14.875}{\sqrt{3.1832\left(\frac{1}{10}+\frac{1}{12}\right)}} = -3.541$$

Reject H_o. There is a difference in the mean race times. p-value = .0021 **(LO1)**

43. H_o: $\mu_1 - \mu_2 \le 0$ H_1: $\mu_1 - \mu_2 > 0$ Reject H_o if $t > 2.650$

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

$$s_p^2 = \frac{(8-1)(15.094)^2 + (7-1)(19.914)^2}{8+7-2} = 305.708 \qquad t = \frac{125.125 - 117.714}{\sqrt{305.708\left(\frac{1}{8}+\frac{1}{7}\right)}} = 0.819$$

There is not enough evidence to reject $H_{o.}$ There is no increase in the mean number sold at the regular price and the mean number sold at reduced price. P-value = .2138 **(LO1)**

44. H_o: $\mu_d \geq 0$ H_1: $\mu_d < 0$ Reject H_o if $t < -2.998$

$\bar{d} = -2.5$ $s_d = 2.928$ $t = \dfrac{2.5}{2.928\sqrt{8}} = -2.415$

There is not enough evidence to reject H_o. The mean number of accidents has not been reduced.
P-value = .0232 **(LO3)**

45. H_o: $\mu_d \leq 0$ H_1: $\mu_d > 0$ Reject H_o if $t > 1.895$

$\bar{d} = 1.75$ $s_d = 2.9155$ $t = \dfrac{1.75}{2.9155/\sqrt{8}} = 1.698$

There is not enough evidence to reject H_o. There is no difference in the mean number of absences. The *p*-value is greater than 0.05. (actual *p*-value = .0667) **(LO3)**

46. H_o: $\mu_d = 0$ H_1: $\mu_d \neq 0$ Reject H_o if z < - 1.761 or *z* > 1.761

$\bar{d} = -246$ $s_d = 547$ $t = \dfrac{-247.67}{548.04/\sqrt{15}} = -1.742$

There is not enough evidence to reject H_o. There is no difference in the mean insurance price.
P-value = .1030 **(LO3)**

47. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ If *t* is not between - 2.024 and 2.024, reject H_o.

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

$$t = \frac{150000 - 180000}{\sqrt{1{,}157{,}894{,}737\left(\frac{1}{15} + \frac{1}{25}\right)}} = -2.70$$ **(LO1)**

Reject the null hypothesis. The population means are different. The *p*-value is almost zero.

48. H_o: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ If *t* is not between - 2.228 and 2.228, reject H_o.

$$s_p^2 = \frac{(5-1)(1.73)^2 + (7-1)(0.9)^2}{5+7-2} = 1.683 \qquad t = \frac{3.000 - 1.857}{\sqrt{1.683\left(\frac{1}{5} + \frac{1}{7}\right)}} = 1.50$$

There is not enough evidence to reject H_o. We conclude that there is not a difference in the taste ratings. P-value = .1637 **(LO1)**

49. H_o: $\mu_d \le 0$ H_1: $\mu_d > 0$ Reject H_o if $t > 1.895$

$\bar{d} = 3.11$ $s_d = 2.91$ $t = \dfrac{3.11}{2.91/\sqrt{8}} = 3.02$

Reject H_o. The mean contamination rate is lower. P-value = .0096 **(LO3)**

50. H_o: $\mu_1 - \mu_2 \le 0$ H_1: $\mu_1 - \mu_2 > 0$ Reject H_o if $t > 1.397$

$$s_p^2 = \frac{(5-1)(5.95)^2 + (5-1)(4.02)^2}{5+5-2} = 25.78 \qquad t = \frac{16.67 - 14.81}{\sqrt{25.78\left(\frac{1}{5}+\frac{1}{5}\right)}} = 0.58$$

There is not enough evidence to reject H_o. P-value = .2892 **(LO1)**

51. **(LO1)**

a.

H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$

List Price	Group 2	
416,154.63	986,598.04	Mean
235,580.14	516,451.58	std. dev.
72	26	N

96	Df
-570,443.413	difference (List Price - Group 2)
110,504,312,251.332	pooled variance
332,421.889	pooled std. dev.
76,058.845	standard error of difference
0	hypothesized difference
-7.50	T
3.20E-11	p-value (two-tailed)

Reject the null hypothesis as the p-value is very close to zero. There is a difference in the list price of homes with more than 2000 square feet.

b.

H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$

List Price	Group 2	
402,400.54	839,682.41	mean
231,549.45	505,624.19	std. dev.
61	37	N

96	Df

-437,281.864	difference (List Price - Group 2)
129,380,399,062.459	pooled variance
359,694.869	pooled std. dev.
74,951.702	standard error of difference
0	hypothesized difference
-5.83	T
7.29E-08	p-value (two-tailed)

Reject the null hypothesis as the *p*-value is very close to zero. There is a difference in the list price of homes with more than 2 bedrooms.

52. **(LO1)**

a.

H_0: $\mu_d \leq 0$ H_1: $\mu_d > 0$

Hypothesis Test: Paired Observations

0.000	hypothesized value
328,706.375	mean Jan-11
314,036.938	mean Jan-10
14,669.438	mean difference (Jan-11 - Jan-10)
34,540.440	std. dev.
8,635.110	std. error
16	N
15	df
1.70	T
.0550	p-value (one-tailed, upper)

There is not enough evidence to reject the null hypothesis. We cannot conclude that the mean list price is more.

b.

.0550 p-value (one-tailed, lower)

The *p*-value > the significance level and therefore supports the decision not to reject the null hypothesis.

53. **(LO1)**

a.

H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$

Hypothesis Test: Independent Groups (t-test, pooled variance)

Winnipeg	Calgary	
92.860	87.180	mean
9.628	10.853	std. dev.

5	5	n

8	Df
5.6800	difference (Winnipeg - Calgary)
105.2425	pooled variance
10.2588	pooled std. dev.
6.4882	standard error of difference
0	hypothesized difference

0.88	t
.4069	p-value (two-tailed)

There is not enough evidence to reject the null hypothesis.
There is not a difference in the mean gas prices.
Note: the *p*-value > alpha (.05).

b.

$H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 \neq 0$

Halifax	Saint John	
100.240	95.040	mean
14.395	12.959	std. dev.
5	5	n

8	df
5.2000	difference (Halifax - Saint John)
187.5730	pooled variance
13.6957	pooled std. dev.
8.6619	standard error of difference
0	hypothesized difference

0.60	T
.5649	p-value (two-tailed)

There is not enough evidence to reject the null hypothesis.
There is not a difference in the mean gas prices.
Note: the *p*-value > alpha (.01).

c.

$H_0: \mu_1 - \mu_2 \geq 0$ $H_1: \mu_1 - \mu_2 < 0$

Toronto	Montreal	
93.360	100.280	Mean
13.557	14.721	std. dev.
5	5	N

8	df
-6.9200	difference (Toronto - Montreal)
200.2525	pooled variance
14.1511	pooled std. dev.
8.9499	standard error of difference
0	hypothesized difference

-0.77	t
.2308	p-value (one-tailed, lower)

There is not enough evidence to reject the null hypothesis.

There is not a difference in the mean gas prices.
Note: the *p*-value > alpha (.05).

d.

$H_o: \mu_1 - \mu_2 \geq 0$ $H_1: \mu_1 - \mu_2 < 0$

Edmonton	St John's	
85.660	105.860	Mean
10.295	12.554	std. dev.
5	5	N

8	df
-20.2000	difference (Edmonton - St John's)
131.7955	pooled variance
11.4802	pooled std. dev.
7.2607	standard error of difference
0	hypothesized difference

-2.78	t
.0119	p-value (one-tailed, lower)

We reject the null hypothesis as
the *p*-value (.0119) < alpha (.02).
There is a difference in the mean gas prices
between Edmonton and St John's.

54. **(LO1)**
a.

$H_0: \mu_1 - \mu_2 = 0$ $H_0: \mu_1 - \mu_2 \neq 0$

List Price	Group 2	
213,650.00	397,374.97	mean
81,059.39	212,772.66	std. dev.
18	32	n

48	df
-183,724.969	difference (List Price - Group 2)
31,565,394,286.479	pooled variance
177,666.526	pooled std. dev.
52,345.502	standard error of difference
0	hypothesized difference
-3.51	t
.0010	p-value (two-tailed)

Reject the null hypothesis as the *p*-value is very close to zero. There is a difference in the list price of homes with more than 1000 square feet.

b.

$H_0: \mu_1 - \mu_2 = 0$ $H_0: \mu_1 - \mu_2 \neq 0$

List Price	Group 2	
278,252.61	499,008.33	Mean
130,650.03	274,355.49	std. dev.
38	12	N

48	Df
-220,755.728	difference (List Price - Group 2)
30,407,274,300.203	pooled variance
174,376.817	pooled std. dev.
57,741.934	standard error of difference
0	hypothesized difference
-3.82	T
.0004	p-value (two-tailed)

Reject the null hypothesis as the *p*-value is very close to zero. There is a difference in the list price of homes with a half bath.

CHAPTER 11

ANALYSIS OF VARIANCE

1. 9.01 from Appendix A.7 **(LO1)**

2. 9.78 **(LO1)**

3. Reject H_o if $F > 10.5$, where *df* numerator are 7 and 5 in the denominator.

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

   There is not enough evidence to reject H_o. There is no difference in the variations of the two populations. **(LO2)**

4. Reject H_o if $F > 9.15$, where *df* in numerator are 4 and 6 in the denominator.

   $F = 2.94$, found by: $F = \frac{s_1^2}{s_2^2} = \frac{(12)^2}{(7)^2} = 2.94$

   There is not enough evidence to reject H_o. There is no difference in the variations of the two populations. **(LO2)**

5. H_o: $\sigma_1^2 - \sigma_2^2 = 0$ H_1: $\sigma_1^2 - \sigma_2^2 \neq 0$

   Reject H_o when $F > 3.10$ (3.10 is about halfway between 3.14 and 3.07)

   $F = 1.44$, found by $F = \frac{(12)^2}{(10)^2} = 1.44$

   There is not enough evidence to reject H_o. There is no difference in the variations of the two populations. **(LO2)**

6. H_o: $\sigma_1^2 - \sigma_2^2 \leq 0$ H_1: $\sigma_1^2 - \sigma_2^2 > 0$

   Reject H_o when $F > 3.68$ $F = 1.24$, found by $F = \frac{(3.9)^2}{(3.5)^2} = 1.24$

   There is not enough evidence to reject H_o. There is less variation in the stocks. **(LO2)**

7. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same.

   b. Reject H_o if $F > 4.26$

   c. 62.17, 12.75, 74.92

   d.

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

   e. Reject H_o. The treatment means are not all the same. **(LO4&5)**

8. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same.

   b. Reject H_o if $F > 3.89$

   c. 70.4, 82.53, 152.93

   d.

Source	*SS*	*df*	*MS*	*F*
Treatment	70.40	2	35.20	5.12

Error	82.53	12	6.88
Total	152.93	14	

e. Reject H_o. The treatment means are not all the same. **(LO4&5)**

9. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same.
Reject H_o if $F > 4.26$

Source	*SS*	*df*	*MS*	*F*
Treatment	276.50	2	138.25	14.18
Error	87.75	9	9.75	
Total	364.25	11		

Reject H_o. The treatment means are not all the same. **(LO4)**

10. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same. **(LO4)**
Reject H_o if $F > 3.89$

Source	*SS*	*df*	*MS*	*F*
Treatment	22.93	2	11.4560	5.7325
Error	24.00	12	2.0	
Total	46.93	14		

Reject H_o 5.7325 > 3.89. The mean number of hours spent on a terminal are not equal.

11. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same. **(LO4&6)**
b. Reject H_o if $F > 4.26$
c. SST = 107.20 SSE = 9.47 SS total = 116.67
d.

Source	*SS*	*df*	*MS*	*F*
Treatment	107.20	2	53.600	50.96
Error	9.47	9	1.052	
Total	116.67	11		

e. Since 50.96 > 4.26, H_o is rejected. At least one of the means differs.
f. $(\bar{X}_1 - \bar{X}_2) \pm t\sqrt{MSE(1/n_1 + 1/n_2)}$

$(9.667 - 2.20) \pm 2.262\sqrt{1.052(1/3 + 1/5)}$

7.467 ± 1.69

[5.777, 9.157] Yes, we can conclude that the treatments 1 and 2 have different means.

12. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all the same. **(LO4&6)**
b. Reject H_o if $F > 3.47$
c. SST = 46.96 SSE = 53.00 SS total = 99.96
d.

Source	*SS*	*df*	*MS*	*F*
Treatment	46.96	2	23.48	9.30
Error	53.00	21	2.52	
Total	99.96	23		

e. Since 9.30 > 3.47, H_o is rejected. At least one of the means differ.
f. $(\bar{X}_1 - \bar{X}_2) \pm t\sqrt{MSE(1/n_1 + 1/n_2)}$

$(6.0 - 4.25) \pm 2.080\sqrt{2.52(1/10 + 1/8)}$

1.75 ± 1.57

[0.18, 3.32] Yes, we can conclude that the treatments 2 and 3 are different.

13 H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: Treatment means are not all equal,
Reject H_o if $F > 3.71$

Source	*SS*	*df*	*MS*	*F*
Treatment	32.33	3	10.77	2.36
Error	45.67	10	4.567	
Total	78.00	13		

Since 2.36 is less than 3.71, there is not enough evidence to reject H_o. There is no difference in the mean number of months. **(LO4)**

14. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: At least one mean differs Reject H_o if $F > 3.81$

Source	*SS*	*df*	*MS*	*F*
Treatment	86.49	2	43.245	13.09
Error	42.95	13	3.3038	
Total	129.44	15		

Since 13.09 > 3.81, H_o is rejected. At least one mean rate of return differs.

b. No, you would need more information, such as confidence intervals to make that decision. **(LO4)**

15. H_o: $\sigma_1^2 \le \sigma_2^2$ H_1: $\sigma_1^2 > \sigma_2^2$ **(LO1)**

$df_1 = 21 - 1 = 20$ $df_2 = 18 - 1 = 17$ H_o is rejected if $F > 3.16$

$F = \frac{(45,600)^2}{(21,330)^2} = 4.57$ Reject H_o. There is more variation in selling price of waterfront homes.

16. H_o: $\sigma_n^2 \le \sigma_o^2$ H_1: $\sigma_n^2 > \sigma_o^2$

$df_1 = 16 - 1 = 15$ $df_2 = 16 - 1 = 15$ H_o is rejected if $F > 2.40$

$F = \frac{(22)^2}{(12)^2} = 3.36$ Reject H_o. There is more variation in processing time of the new machine. **(LO1)**

17. Sharkey: $n = 7$ s = 14.79

White: $n = 8$ s = 22.95

H_o: $\sigma_w^2 \le \sigma_s^2$ H_1: $\sigma_w^2 > \sigma_s^2$

$df_s = 7 - 1 = 6$ $df_w = 8 - 1 = 7$ H_o is rejected if $F > 8.26$

$F = \frac{(22.95)^2}{(14.79)^2} = 2.41$ There is not enough evidence to reject H_o. There is no difference in the variation of the weekly sales. **(LO1)**

18. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all equal.

b. $\alpha = 0.05$ Reject if $F > 3.89$

c.

Source	*SS*	*df*	*MS*	*F*
Treatment	40	2	20	4
Error	60	12	5	
Total	100	14		

d. Reject H_o. The treatment means are not all equal. **(LO4)**

19. a. H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: Treatment means are not all equal.

b. $\alpha = 0.05$ Reject if $F > 3.10$

c.

Source	*SS*	*df*	*MS*	*F*
Treatment	50	3	50/3	1.67
Error	200	20	10	

Total	250	23

d. Do not reject H_o. There is not a difference in the treatment means. **(LO4)**

20.

Source	*SS*	*df*	*MS*	*F*
Treatment	320	2	160	8.00
Error	180	9	20	
Total	500	11		

a. 3
b. 12
c. 4.26
d. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Not all means are equal.
e. H_o is rejected. The treatment means differ. **(LO4)**

21. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Not all means are equal.
H_o is rejected if $F > 3.89$

Source	*SS*	*df*	*MS*	*F*
Treatment	63.33	2	31.667	13.38
Error	28.40	12	2.367	
Total	91.73	14		

H_o is rejected. There is a difference in the treatment means. **(LO4)**

22. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Not all means are equal.
H_o is rejected if $F > 3.89$

Source	*SS*	*df*	*MS*	*F*
Treatment	26.13	2	13.067	13.52
Error	11.60	12	0.967	
Total	37.73	14		

H_o is rejected since 13.52 > 3.89. There is a difference in the mean weight loss among the three diets. **(LO4)**

23. H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: Not all means are equal. **(LO4)**
H_o is rejected if $F > 3.10$

Source	*SS*	*df*	*MS*	*F*
Factor	87.79	3	29.26	9.12
Error	64.17	20	3.21	
Total	151.96	23		

Since computed F of 9.12 > 3.10, the null hypothesis of no difference is rejected at the 0.05 level.

24. One-way ANOVA: Clothes, Food, Toys
H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Not all means are equal.

```
Source  DF      SS      MS      F      P
Factor   2  3182.0  1591.0  35.56  0.000
Error   29  1297.5    44.7
Total   31  4479.5
S = 6.689   R-Sq = 71.04%   R-Sq(adj) = 69.04%
                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level     N    Mean  StDev  -------+---------+---------+---------+--
Clothes   9  28.000  8.139  (----*----)
Food     12  46.417  6.186                    (---*---)
Toys     11  52.636  5.887                           (---*---)
                            -------+---------+---------+---------+--
```

```
                                        30        40        50        60

Pooled StDev = 6.689
```

The hypothesis of identical means can definitely be rejected since the p-value is virtually zero. As seen from the non-overlapping confidence intervals, clothes have a mean attention span at least ten minutes below the other groups. **(LO4,5&6)**

25. a. H_o: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$ Critical value of F = 4.75

Source	*SS*	*df*	*MS*	*F*
Treatment	219.43	1	219.43	23.10
Error	114.00	12	9.5	
Total	333.43	13		

b. $$t = \frac{19 - 27}{\sqrt{9.5\left(\frac{1}{6} + \frac{1}{8}\right)}} = -4.81$$

Since $t^2 = F$. That is $(-4.81)^2 \approx 23.10$ (actually 23.14, difference due to rounding).

c. H_o is rejected. There is a difference in the mean scores. **(LO4)**

26. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: The means are not all equal. Critical value of F = 3.44

Mean	*n*	*Std. Dev*	
49.0	7	7.81	High-sch
74.7	9	15.31	Under
78.3	9	15.58	Master
68.8	25	18.32	Total

ANOVA table

Source	*SS*	*Df*	*MS*	*F*	*p-value*
Treatment	3,872.00	2	1,936.000	10.18	.0007
Error	4,182.00	22	190.091		
Total	8,054.00	24			

Since the computed value of 10.18 exceeds the critical value of 3.44, the null hypothesis is rejected. The mean salary for those with high school or less is $49 000, it is $74 670 for those with an undergraduate degree, and $78 300 for those with a Master's degree or more. The salary for those with only high school differs from both the other groups. The salaries for those with college work do no differ. The confidence interval for the difference between high school and undergraduate is computed as follows:

$$(49.00 - 74.67) \pm 2.074\sqrt{190\left(\frac{1}{7} + \frac{1}{9}\right)} = -25.67 \pm 14.41$$

for $\bar{X}_1$ and $\bar{X}_3$: $(\bar{X}_1 - \bar{X}_3) \pm 2.074\sqrt{190\left(\frac{1}{7} + \frac{1}{9}\right)} = -29.333 \pm 14.41$

Note that zero is not in the interval.
Confidence intervals for 2 & 3 follow. Note that zero is in the interval. **(LO4,5&6)**

upper	9.813125
lower	-17.1465

27. H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_1: Treatment means are not all equal.
Reject H_o if F > 3.49. The computed value of F is 9.61.
The null hypothesis of equal means is rejected because the F statistic (9.61) is greater than the critical value (3.49). The p- value (0.0016) is also less than the significance level (0.05). The mean waiting times are different. **(LO4)**

28. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all equal.
Reject H_o if F > 5.61. The computed value of F is 8.26.
The null hypothesis of equal means is rejected because the F statistic (8.26) is greater than the critical value (5.61) using the 0.01 significance level. The p- value (0.0019) is also less than the significance level (0.01). The mean gas mileages are different. **(LO4)**

29. a. H_o: $\mu_1 = \mu_2 = \mu_3$ H_1: Treatment means are not all equal.
Reject H_o if F > 6.36. The computed value of F is 11.33.
The null hypothesis of equal means is rejected because the F statistic (11.33) is greater than the critical value (6.36). The p- value (0.0010) is also less than the significance level (0.01). The mean production rates are different.

b. $$(43.33 - 41.5) \pm 2.947\sqrt{0.5444\left(\frac{1}{6} + \frac{1}{6}\right)}$$

This reduces to 1.83 ± 1.26. So the difference is between 0.57 and 3.09. **(LO4&6)**

30. a. There is not enough evidence to reject the null hypothesis of equal means because the F statistic (3.41) is less than the critical value (5.49). The p-value (0.0478) is also greater than the significance level (0.01). The mean amounts of money withdrawn are not different.

b. $$(82.5 - 38.2) \pm 1.703\sqrt{1477.633\left(\frac{1}{10} + \frac{1}{10}\right)}$$

This reduces to 44.3 ± 29.3. So the difference is between 15.0 and 73.6. **(LO4&6)**

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

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Treatment	0.03478	5	0.006956	3.86
Error	0.10439	58	0.001800	
Total	0.13917	63		

H_o is rejected. There is a difference in the mean weight of the colors. **(LO4&5)**

32. a. Recall that $\bar{X} = \sum X / n$ so $\bar{X}(n) = \sum X$ For the first treatment
$\bar{X}(n) = 51.43(10)$, so $\sum X = 514.3$ SST = 306.934, found by

$$SST = \frac{(514.3)^2}{10} + \frac{(446.4)^2}{10} + \frac{(472.0)^2}{10} + \frac{(508.5)^2}{10} - \frac{(1941.2)^2}{40}$$

b. 650.75 – 306.934 = 343.816

c.

Source	*SS*	*df*	*MS*	*F*
Treat	306.934	3	100.215	10.304
Error	343.816	36	9.725	
Total	650.750			

d. 10.304 > 2.89, so reject H_0. There is a difference in the treatment means.

e. The 95% CI: (51.43 – 50.85) ± 2.03 $\sqrt{9.725(1/10+1/10)}$ = 0.58 ± 2.831

[- 2.251, 3.411] Zero is in the interval, and so, we cannot conclude that the number of minutes of music differ between $\bar{X}_1$ and $\bar{X}_4$ **(LO5&6)**

33. **(LO4)**

a.

$H_0: \mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

$\alpha = .05$

One factor ANOVA

	Mean	*n*	*Std. Dev*	
	459,074.7	58	292,294.38	Apartment
	904,248.0	26	552,948.91	House
	391,278.5	14	137,470.32	Townhouse
	567,496.8	98	416,469.29	Total

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Treatment	4,064,986,860,294.07	2	2,032,493,430,147.030	15.13	1.97E-06
Error	12,759,339,849,142.10	95	134,308,840,517.285		
Total	16,824,326,709,436.10	97			

The *p*-value is very close to zero and less than a significance level of .01, and so, we reject the null hypothesis and conclude that the variances of list prices of the home styles are different.

b.

$H_0: \mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal α = .05

One factor ANOVA

Mean	n	Std. Dev	
402,400.5	61	231,549.45	1 and 2 bedrooms
586,994.7	19	308,194.90	3 bedrooms
1,106,408.3	18	542,334.67	> bedrooms
567,496.8	98	416,469.29	Total

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Treatment	6,897,546,936,345.69	2	3,448,773,468,172.850	33.01	1.31E-11
Error	9,926,779,773,090.43	95	104,492,418,664.110		
Total	16,824,326,709,436.10	97			

The *p*-value is very close to zero and less than a significance level of .01, and so, we reject the null hypothesis and conclude that the variances of list prices of homes with 1 and 2 bedrooms, 3 bedrooms, and more than 3 bedrooms are different.

c.

$$H_0 : s_1^2 - s_2^2 = 0$$

$$H_1 : s_1^2 - s_2^2 ,^1\ 0$$

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
less than 1500 sq ft	59	22358533	378958.2	3.77E+10
more than 1500 sq ft	39	33256149	852721.8	2.46E+11

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5.27E+12	1	5.27E+12	43.78668	2.09E-09	3.940163
Within Groups	1.16E+13	96	1.2E+11			

Total	1.68E+13	97

The p-value of the F-test is very close to zero, so we reject the null hypothesis, and conclude that there is a difference in the variability of the list prices of the homes.

34. **(LO1)**

$$H_0: \sigma_1^2 - \sigma_2^2 = 0$$

$$H_1: \sigma_1^2 - \sigma_2^2 \neq 0$$

F-Test Two-Sample for Variances

	40544	*40179*
Mean	328706.375	314036.9375
Variance	2.1081E+10	15621985400
Observations	16	16
Df	15	15
F	1.34946073	
P(F<=f) one-tail	0.28442497	
F Critical one-tail	2.40344707	

The *p*-value is greater than the significance level, so there is not enough evidence to reject the null hypothesis. We conclude that the variances of list prices of the homes listed January 2011 and January 2010 are not different.

35. **(LO4&5)**

a.

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1: not all population means are equal

One factor ANOVA

Mean	*n*	*Std. Dev*	
92.86	5	9.628	Winnipeg
87.18	5	10.853	Calgary
94.86	5	9.808	Saskatoon
91.63	15	9.948	Total

ANOVA table

Source	SS	df	MS	F	p-value
Treatment	158.741	2	79.3707	0.78	.4819
Error	1,226.752	12	102.2293		
Total	1,385.493	14			

The *p*-value is > the significance level of .05,
so, there is not enough evidence to reject the
null hypothesis and we conclude that the variances of gas prices of the three cities are not different.

b.

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1: not all population means are equal

One factor ANOVA

Mean	n	Std. Dev	
100.24	5	14.395	Halifax
87.18	5	10.853	Calgary
95.04	5	12.959	Saint John
94.15	15	13.104	Total

ANOVA table

Source	SS	df	MS	F	p-value
Treatment	432.305	2	216.1527	1.32	.3044
Error	1,971.712	12	164.3093		
Total	2,404.017	14			

The *p*-value is > the significance level of .01,
so, there is not enough evidence to reject the
null hypothesis and we conclude that the variances of gas prices of the three cities are not different.

c.

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1: not all population means are equal

One factor ANOVA

Mean	n	Std. Dev	
93.36	5	13.557	Toronto
103.24	5	12.508	Vancouver
100.28	5	14.721	Montreal
98.96	15	13.323	Total

ANOVA table

Source	SS	df	MS	F	p-value
Treatment	257.104	2	128.5520	0.69	.5193
Error	2,227.772	12	185.6477		
Total	2,484.876	14			

The *p*-value is > the significance level of .05, so, there is not enough evidence to reject the null hypothesis and we conclude that the variances of gas prices of the three cities are not different.

36. **(LO4&5)**

a.

$H_0: \mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

$\alpha = .05$

One factor ANOVA

Mean	n	Std. Dev	
223,391.7	24	105,650.77	Apartment
446,620.8	24	211,770.89	House
240,700.0	2	27,152.90	Townhouse
331,234.0	50	197,121.89	Total

ANOVA table

Source	SS	df	MS	F	p-value
Treatment	615,050,743,509.69	2	307,525,371,754.844	11.21	.0001
Error	1,288,944,217,159.29	47	27,424,345,045.942		
Total	1,903,994,960,668.98	49			

The *p*-value is close to zero and less than the significance level of .05, and so, we reject the null hypothesis and conclude that the variances of list prices of the home styles are different.

b.

$H_0: \mu_1 = \mu_2 = \mu_3$
H_1:not all population means are equal

$\alpha = .05$
One factor ANOVA

	Mean	n	Std. Dev	
	225,400.0	27	99,680.26	1 and 2 bedrooms
	405,312.5	8	149,231.97	3 bedrooms
	482,226.6	15	239,263.55	> 3 bedrooms
	331,234.0	50	197,121.89	Total

ANOVA table

Source	SS	df	MS	F	p-value
Treatment	688,305,032,131.38	2	344,152,516,065.690	13.31	2.64E-05
Error	1,215,689,928,537.60	47	25,865,743,160.375		
Total	1,903,994,960,668.98	49			

The *p*-value is very close to zero and less than the significance level of .05, and so, we reject the null hypothesis and conclude that the variances of list prices of homes with 1 and 2 bedrooms,
3 bedrooms, and more than 3 bedrooms are different.

c.

$H_0 : s_1^2 - s_2^2 = 0$

$H_1 : s_1^2 - s_2^2 ,^1\ 0$

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
less than 1500 sq ft	41	10982299	267861	1.34E+10

1500 sq ft +	9	5579400	619933.3	5.66E+10

ANOVA

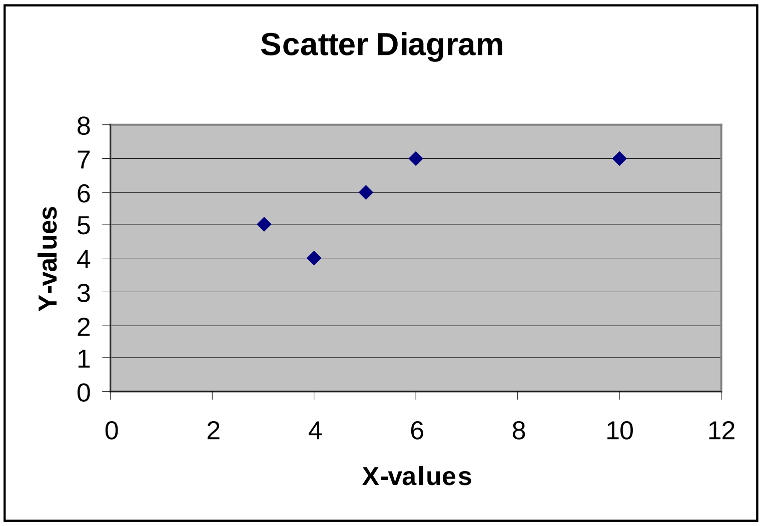
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	9.15E+11	1	9.15E+11	44.38888	2.42E-08	4.042652
Within Groups	9.89E+11	48	2.06E+10			
Total	1.9E+12	49				

The p-value of the F-test is very close to zero, so we reject the null hypothesis, and conclude that there is a difference in the variability of the list prices of the homes.

CHAPTER 12

LINEAR REGRESSION AND CORRELATION

1. a.

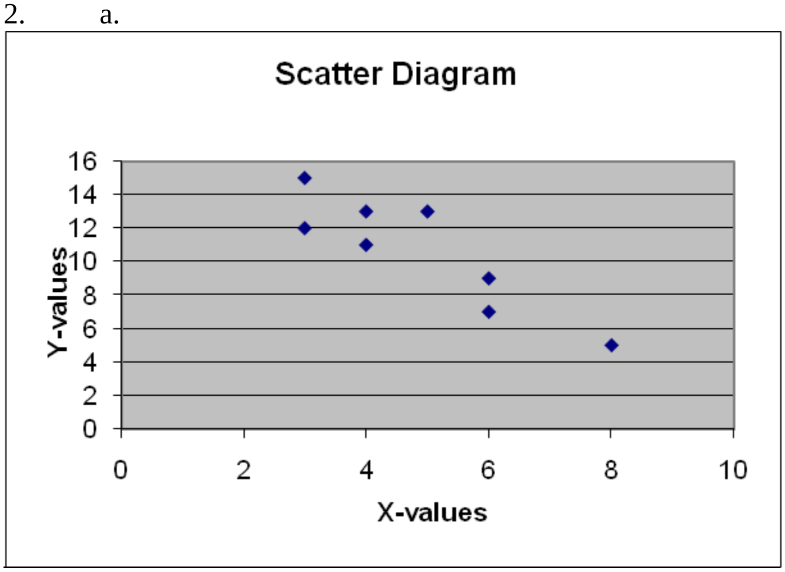


b. $Y' = 3.7671 + 0.3630X$

$$b = \frac{5(173) - (28)(29)}{5(186) - (28)^2} = 0.3630 \qquad a = \frac{29}{5} - (0.363)\frac{28}{5} = 3.7671$$

c. 6.3081, found by $Y' = 3.7671 + 0.3630(7)$ **(LO1&3)**

2. a.

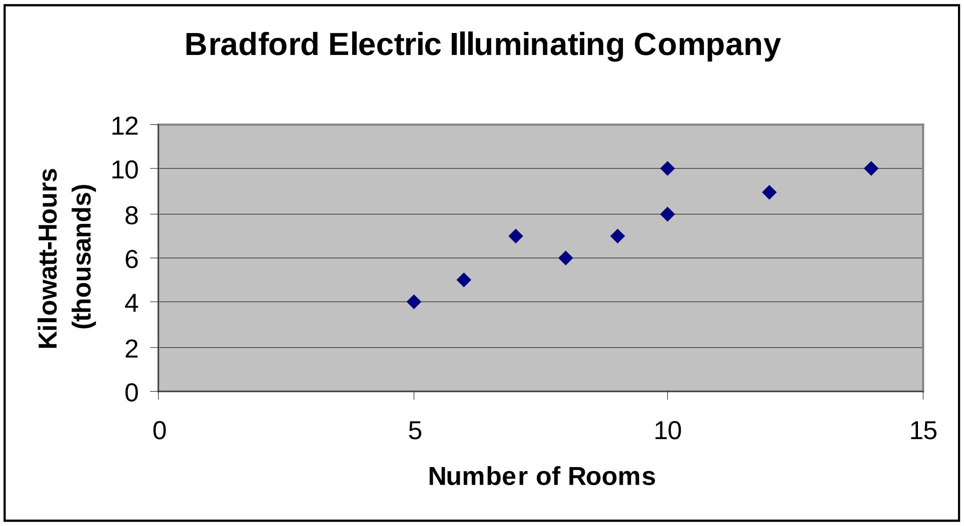


b. $Y' = 19.1198 – 1.7425X$

$$b = \frac{8(378) - (39)(85)}{8(211) - (39)^2} = -1.7425 \qquad a = \frac{85}{8} - (-1.74)\frac{39}{8} = 19.1198$$

c. 6.9223, found by 19.1198 – 1.7425(7) **(LO1&3)**

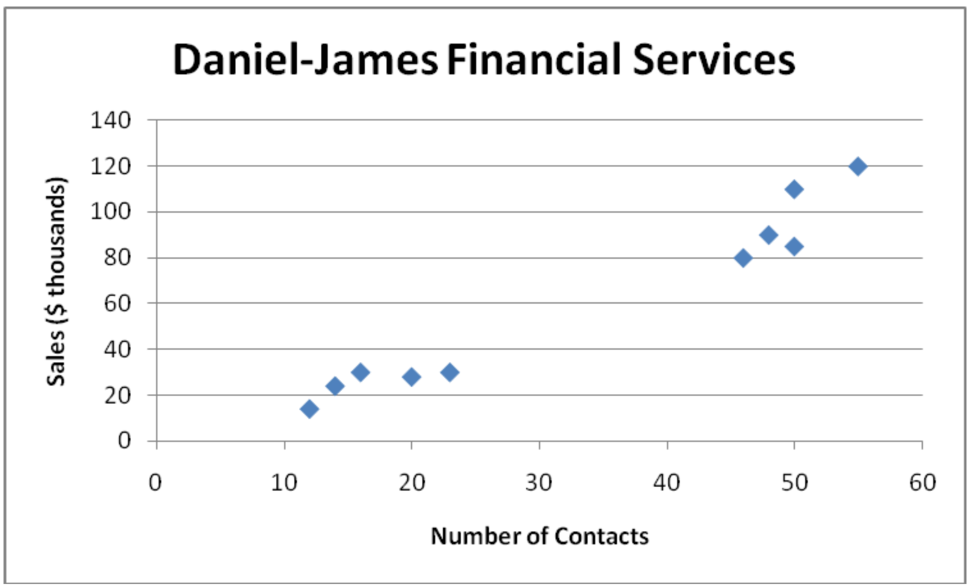
3. a.



b. $b = \dfrac{10(718) - (91)(74)}{10(895) - (91)^2} = \dfrac{446}{669} = 0.667$ $\qquad a = \dfrac{74}{10} - (0.667)\dfrac{91}{10} = 1.333$

c. Y¢= 1.333 + 0.667(6) = 5.335 **(LO1&3)**

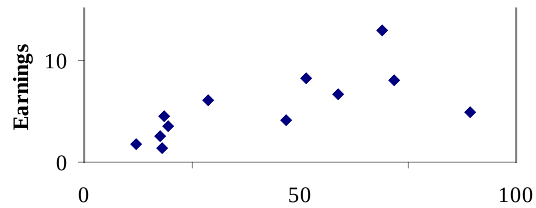
4. a.



b. $b = \dfrac{10(26{,}584) - (334)(611)}{10(13{,}970) - (334)^2} = 2.1946$ $\qquad a = \dfrac{611}{10} - (2.1946)\dfrac{334}{10} = -12.201$

c. Y¢= - 12.201 + 2.1946(40) = 75.583 **(LO1&3)**

5. a.



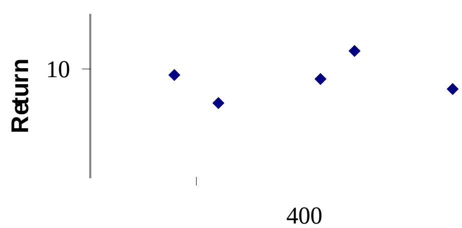
b. $$b = \frac{12(3306.35) - (501.1)(64.1)}{5(28,459) - (501.1)^2} = 0.0836$$

$$a = \frac{64.1}{12} - (0.0836)\frac{501.10}{12} = 1.8507$$

c. $Y' =$ 1.8507 + 0.0836(50.0) = 6.0307 ($ million)

Note: calculator or computer values may be slightly different due to rounding. **(LO1&3)**

6. a.



b. $$b = \frac{9(34,111) - (3504.5)(87.9)}{9(1,659,866) - (3504.5)^2} = -0.000393$$

$$a = \frac{87.9}{9} - (-0.000393)\frac{3504.5}{9} = 9.9197$$

c. $Y' = 9.9197 - 0.000393(400.0) = 9.7625$. **(LO1&3)**

7. a. Police is the independent variable and crime is the dependent variable

b.

c. $$b = \frac{8(1502) - (146)(95)}{8(2906) - (146)^2} = -0.95963 \quad a = \frac{95}{8} - (-0.95963)\left(\frac{146}{8}\right) = 29.3882$$

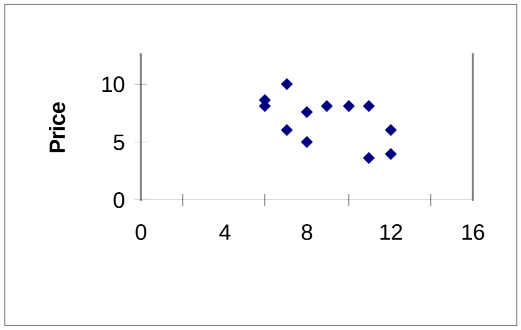
$Y' = 29.3882 - 0.95963\,Police$

d. $Y' = 29.3882 - 0.95963(20) = 10.196$

e. Inverse relationship. As the number of police increase, crime decreases. **(LO1&3)**

8. a. Age is the independent variable and price is the dependent variable.

b.



c. $n = 12 \quad \Sigma X = 107 \quad \Sigma Y = 82.9 \quad \Sigma X^2 = 1009 \quad \Sigma XY = 712.9 \quad \Sigma Y^2 = 615.29$

$$b = \frac{12(712.9) - 107(82.9)}{12(1009) - (107)^2} = -.4788$$

$$a = \frac{82.9}{12} - (-.4788)(\frac{107}{12}) = 11.177$$

d. $y' = 11.177 - .4788(10) = 6.39$

e. Inverse relationship. As the age of the car increases, the selling price decreases. **(LO1&3)**

9. a. 0.993, found by $\sqrt{\frac{175 - 3.767(29) - 0.363(173)}{5 - 2}}$

b. $Y' \pm 0.993$ **(LO5)**

10 a. 1.6578, found by $s_{y \cdot x} = \sqrt{\frac{983 - 19.1197(85) - (-1.7425)(378)}{8 - 2}}$

b. $Y' \pm 3.3156$ **(LO5)**

11. a. 0.898, found by $\sqrt{\frac{584 - 1.333(74) - 0.667(718)}{10 - 2}}$

b. $Y' \pm 1.796$ **(LO5)**

12. a. 9.3178, found by $\sqrt{\frac{51,581 - (-12.201)(611) - 2.1946(26,584)}{10 - 2}}$

b. $Y' \pm 2(9.3178)$ **(LO5)**

13. 3.379, found by $\sqrt{\frac{1419 - 29.3877(95) - (-0.9596)(1502)}{8 - 2}}$ **(LO5)**

14. 1.733, found by $\sqrt{\frac{615.29 - 11.179(82.9) - (-0.479)(712.9)}{12 - 2}}$ **(LO5)**

15. a. $6.308 \pm 3.182(0.993)\sqrt{\frac{1}{5} + \frac{(7 - 5.6)^2}{29.2}} = 6.308 \pm 1.633$ **(LO4)**

$= [4.675, 7.941]$

b. $6.308 \pm (3.182)(0.933)\sqrt{1 + 1/5 + 0.0671} = [2.751, 9.865]$ **(LO4)**

16. a. $6.9222 \pm 2.447(1.6578)\sqrt{\frac{1}{8} + \frac{(7 - 4.875)^2}{20.875}} = 6.9222 \pm 2.37$ **(LO4)**

$= [4.5522, 9.2922]$

b. $6.9222 \pm 2.447(1.6578)\sqrt{1 + \frac{1}{8} + \frac{(7 - 4.875)^2}{20.875}} = 6.9222 \pm 4.6982$

$= [2.2238, 11.6208]$

17. a. [4.495, 6.171]

b. [3.440, 7.226] **(LO4)**

18. a. [66.42, 84.75]

b. [47.10, 104.07] **(LO4)**

19. $\sum X = 28$ $\sum Y = 29$ $\sum X^2 = 186$ $\sum XY = 173$ $\sum Y^2 = 175$

$$r = \frac{5(173) - (28)(29)}{\sqrt{[5(186) - (28)^2][5(175) - (29)^2]}} = 0.75$$

The 0.75 coefficient indicates a rather strong positive correlation between *X* and *Y*. The coefficient of determination is 0.5625, found by $(0.75)^2$. *X* accounts for more than 56 percent of the variation in *Y*. **(LO5)**

20. $\sum X = 39$ $\sum Y = 85$ $\sum X^2 = 211$ $\sum XY = 378$ $\sum Y^2 = 983$

$$r = \frac{8(378) - (39)(85)}{\sqrt{[8(211) - (39)^2][8(983) - (85)^2]}} = -0.89$$

The −0.89 indicates a very strong negative relationship between *X* and *Y*. The coefficient of determination is 0.7921, found by $(-0.89)^2$. *X* accounts for close to 80 percent of the variation in *Y*. **(LO5)**

21. a. $n = 5$ $\sum X = 20$ $\sum Y = 85$ $\sum X^2 = 90$ $\sum XY = 376$ $\sum Y^2 = 1595$

$$r = \frac{5(376) - (20)(85)}{\sqrt{[5(90) - (20)^2][5(1595) - (85)^2]}} = 0.9295$$

b. $r^2 = (.9295)^2 = 0.864$

c. The 0.9295 indicates a very strong positive relationship between *X* and *Y*. The coefficient of determination is 0.864. *X* accounts for about 86.4 percent of the variation in *Y*. **(LO5)**

22. a. $n = 5$ $\sum X = 15$ $\sum Y = 120$ $\sum X^2 = 55$ $\sum XY = 430$ $\sum Y^2 = 3450$

$$r = \frac{5(430) - (15)(120)}{\sqrt{[5(55) - (15)^2][5(3450) - (120)^2]}} = 0.927$$

b. The r^2 is 0.8593, so about 86 percent of the variation in production is explained by the variation in the number of assemblers. **(LO5)**

23. a. $n = 8$ $\sum X = 146$ $\sum Y = 95$ $\sum X^2 = 2906$ $\sum XY = 1502$ $\sum Y^2 = 1419$

$$r = \frac{8(1502) - (146)(95)}{\sqrt{[8(2906) - (146)^2][8(1419) - (95)^2]}} = -0.874$$

b. 0.76, found by $(-0.874)^2$

c. -0.874 indicates a strong inverse relationship. As the number of police increase, the crime decreases. The coefficient of determination is 0.76. *X* accounts for about 76 percent of the variation in *Y*. **(LO5)**

24. a. $n = 12 \quad \sum X = 107 \quad \sum Y = 82.9 \quad \sum X^2 = 1009 \quad \sum XY = 712.9 \quad \sum Y^2 = 615.29$

$$r = \frac{12(712.9) - (107)(82.9)}{\sqrt{[12(1009) - (107)^2][12(615.29) - (82.9)^2]}} = -0.544$$

b. 0.296, found by $(-0.544)^2$

c. Moderate negative (inverse) correlation between age of car and selling price. So, 29.6 percent of the variation in the selling price is explained by the variation in the age of the car. **(LO5)**

25. Reject H_o if $t > 1.812$

$$t = \frac{0.32\sqrt{12-2}}{\sqrt{1-(0.32)^2}} = 1.07$$ Do not reject H_o. **(LO6)**

26. Reject H_o if $t < -1.771$

$$t = \frac{-0.46\sqrt{15-2}}{\sqrt{1-(-0.46)^2}} = -1.868$$ Reject H_o. **(LO6)**

27. H_o: $\rho \geq 0$ H_1: $\rho < 0$ Reject H_o if $t < -2.552$ $df = 18$

$$t = \frac{-0.78\sqrt{20-2}}{\sqrt{1-(-0.78)^2}} = -5.288$$

Reject H_o. There is a negative correlation between litres sold and the pump price. **(LO6)**

28. H_o: $\rho \leq 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.734$ $df = 18$

$$t = \frac{0.86\sqrt{20-2}}{\sqrt{1-(0.86)^2}} = 7.150$$

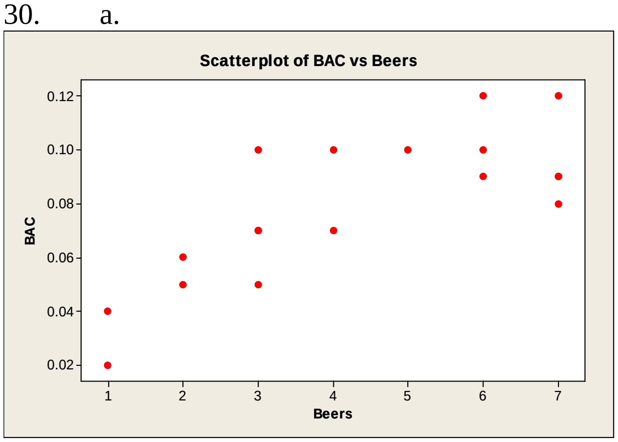
Reject H_o. There is a positive correlation between assets and pretax profit. **(LO6)**

29. H_o: $\rho \leq 0$ H_1: $\rho > 0$ Reject H_o if $t > 2.650$ $df = 13$

$$t = \frac{0.667\sqrt{15-2}}{\sqrt{1-(0.667)^2}} = 3.23$$

Reject H_o. There is a positive correlation between passengers and cost. **(LO6)**

30. a.



b.

Beers	BAC	$(X-\overline{X})$	$(Y-\overline{Y})$	$(X-\overline{X})^2$	$(Y-\overline{Y})^2$	$(X-\overline{X})(Y-\overline{Y})$
6	0.1	1.722	0.0211	2.966	0.00045	0.0364
7	0.09	2.722	0.0111	7.411	0.00012	0.0302
7	0.09	2.722	0.0111	7.411	0.00012	0.0302
4	0.1	-0.278	0.0211	0.077	0.00045	-0.0059
5	0.1	0.722	0.0211	0.522	0.00045	0.0152
3	0.07	-1.278	-0.0089	1.633	0.0000	0.0114
3	0.1	-1.278	0.0211	1.633	0.00045	-0.0270
6	0.12	1.722	0.0411	2.966	0.00169	0.0708
6	0.09	1.722	0.0111	2.966	0.00012	0.0191
3	0.07	-1.278	-0.0089	1.633	0.0001	0.0114
3	0.05	-1.278	-0.0289	1.633	0.00083	0.0369
7	0.08	2.722	0.0011	7.411	0.00000	0.0030
1	0.04	-3.278	-0.0389	10.744	0.00151	0.1275
4	0.07	-0.278	-0.0089	0.077	0.0001	0.0025
2	0.06	-2.278	-0.0189	5.188	0.00036	0.0430
7	0.12	2.722	0.0411	7.411	0.00169	0.1119
2	0.05	-2.278	-0.0289	5.188	0.00083	0.0658
1	0.02	-3.278	-0.0589	10.744	0.00347	0.1930
77	1.42			77.611	0.01278	0.7756

$$\overline{X} = \frac{77}{18} = 4.278 \qquad \overline{Y} = \frac{1.42}{18} = 0.0789$$

$$s_x = \sqrt{\frac{77.61}{17}} = 2.1367 \qquad s_y = \sqrt{\frac{0.01278}{17}} = 0.0274$$

$$r = \frac{0.7756}{(18-1)(2.1367)(0.0274)} = 0.779$$

c. 0.607, found by $(0.779)^2$

d. H_0: $\rho \le 0$ H_1: $\rho > 0$ Reject H_0 if $t > 2.583$ $df = 16$

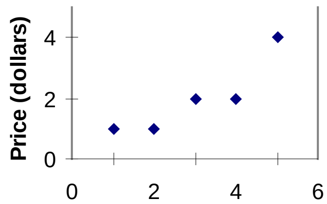
$$t = \frac{0.779\sqrt{18-2}}{\sqrt{1-(0.779)^2}} = 4.97$$

Reject H_0. There is a positive correlation between beers consumed and BAC. p-value = .0001, which is very close to zero, and confirms that H_0 should be rejected. **(LO5&6)**

31. Coefficient of correlation $r = 0.8944$, found by $\frac{(5)(340)-(50)(30)}{\sqrt{[(5)(600)-(50)^2][(5)(200)-(30)^2]}}$

Then $(0.8944)^2 = 0.80$, the coefficient of determination. **(LO5)**

32. a.



b. 0.8167, found by (6 – 1.1)/6 $\sum (Y - Y')^2 = 1.1$ and $\sum (Y - \bar{Y})^2 = 6$

also, $r = \dfrac{(5)(37) - (15)(10)}{\sqrt{[(5)(55) - (15)^2][(5)(26) - (10)^2]}} = 0.9037$ and $r^2 = 0.8167$

c. Turnover accounts for 81.67 percent of the variation in price. **(LO1&5)**

33. a. $r^2 = 1000/1500 = 0.667$

b. 0.82, found by $\sqrt{0.667}$

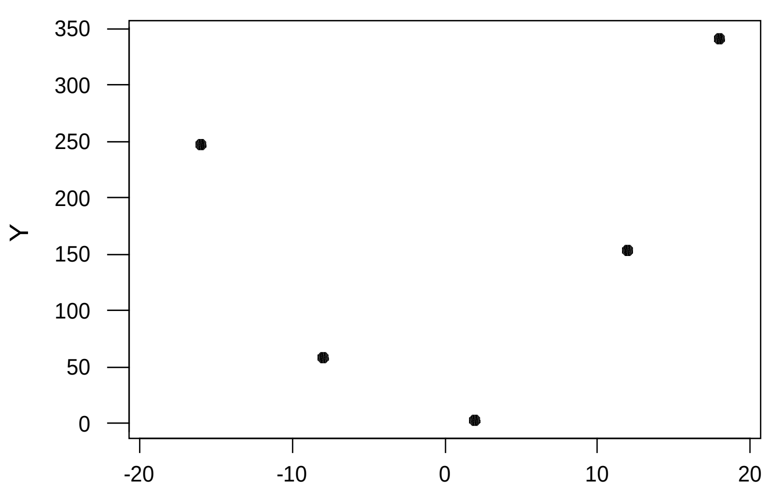
c. 6.20, found by $s_e = \sqrt{\dfrac{500}{15 - 2}}$ **(LO5)**

34.

Source	*SS*	*df*	*MS*
Regression	7,200	1	7,200
Residual	1,800	18	100
Total	9,000	19	

(LO5)

35.

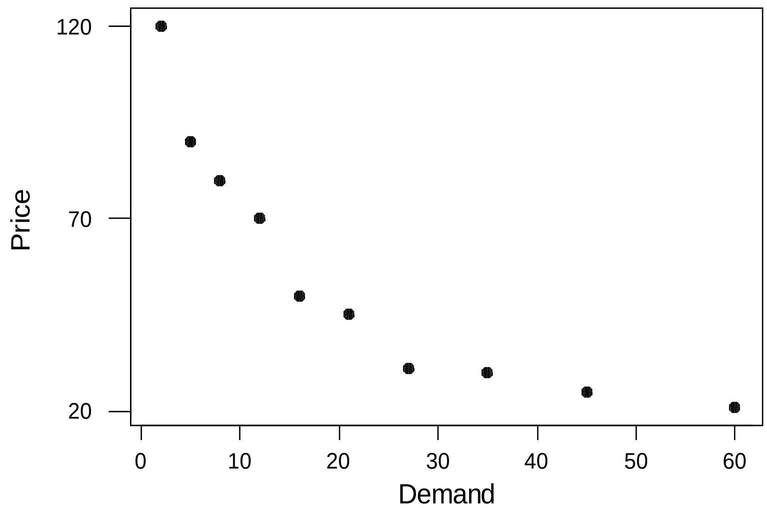


a. The relationship does not appear to be linear.

b. The correlation of *X* and *Y* is 0.752. There is a positive high correlation between the variables.

c. 151.388 **(LO1&5)**

36. a. The correlation of Demand and Price is -0.865 which indicates a fairly strong inverse linear relationship.



b. **17.428**

(LO1&5)

37. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.714$

$$t = \frac{0.94\sqrt{25-2}}{\sqrt{1-(0.94)^2}} = 13.213$$

Reject H_o. There is a positive correlation between passengers and weight of luggage. **(LO6)**

38. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 2.552$

$$t = \frac{0.40\sqrt{20-2}}{\sqrt{1-(0.40)^2}} = 1.852$$

Do not reject H_o. We cannot conclude that there is a positive correlation between GPA and family income. **(LO6)**

39. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 2.764$

$$t = \frac{0.47\sqrt{12-2}}{\sqrt{1-(0.47)^2}} = 1.684$$

Do not reject H_o. There is not a positive correlation between engine size and performance. The p-value is greater than 0.05, but less than 0.10. **(LO6)**

40. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.734$

$$t = \frac{0.21\sqrt{20-2}}{\sqrt{1-(0.21)^2}} = 0.911$$

Do not reject H_o. There is not a positive correlation between shots attempted and shots scored. The p-value is greater than 0.10. **(LO6)**

41. H_o: $\rho \ge 0$ H_1: $\rho < 0$ Reject H_o if $t < -1.701$ $df = 28$

$$t = \frac{-0.45\sqrt{30-2}}{\sqrt{1-0.2025}} = -2.67$$

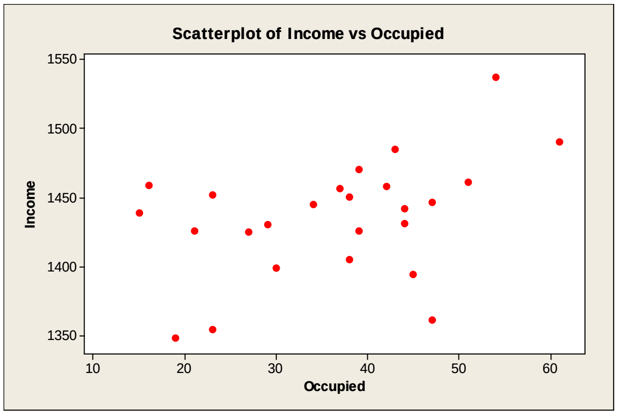
Reject H_o. There is a negative correlation between the selling price and the number of miles driven. **(LO6)**

42. a. Yes, because the correlation coefficient is positive.
    b. 0.119 or 11.9%, found by $(0.345)^2$.
    c. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.771$

$$t = \frac{0.345\sqrt{15-2}}{\sqrt{1-(0.345)^2}} = 1.325$$

Do not reject H_o. There is not a positive correlation between fat grams consumed and cholesterol level. The p-value is greater than 0.10. **(LO6)**

43. a.



Revenue increases slightly as the number of occupied rooms increases.

b. Pearson correlation of Income and Occupied = 0.423

c. H_o: *r* £ 0 H_1: *r* > 0 Reject H_o if $t > 1.319$ $df = 23$

$$t = \frac{0.423\sqrt{25-2}}{\sqrt{1-(0.423)^2}} = 2.24$$

Reject H_o.
There is a positive correlation between revenue and occupied rooms.

d. 17.9%, found by $(0.423)^2$, of the variation in revenue is explained by variation in occupied rooms. **(LO1,5&6)**

44. a. Decrease, because the correlation coefficient is negative.

b. 0.287 or 28.7%, found by $(–0.536)^2$

c. H_o: *r* ³ 0 H_1: *r* < 0 Reject H_o if $t < -1.714$

$$t = \frac{-0.536\sqrt{25-2}}{\sqrt{1-(-0.536)^2}} = -3.045$$

Reject H_o. There is a negative correlation between job satisfaction and stress. **(LO5&6)**

45 a. No, the coefficient is -0.5170 which indicates a negative relationship between the variables.

b. SSR = 3119.4256/3960 = 78.77%

c. $r = \sqrt{0.77877} = -0.8824$; strong, negative

d. $Y' = 45.85$; 46 units; yes, this is reasonable. **(LO5)**

46. a. Yes, the coefficient of Hours of Study is positive.

b. SSR = 1569.0299/2733 = 57.41%

c. $r = \sqrt{0.5741} = 0.7577$; strong, positive relationship

d. Y¢$= 5(4.3284) + 60.1418 = 81.78$; yes, it is reasonable. **(LO5)**

47. a. $r = 0.589$

b. $r^2 = (0.589)^2 = 0.3469$

c. H_o: *r* £ 0 H_1: *r* > 0 Reject H_o if $t > 1.860$

$$t = \frac{0.589\sqrt{10-2}}{\sqrt{1-(0.589)^2}} = 2.062$$

H_o is rejected. There is a positive association between family size and the amount spent on food. **(LO5&6)**

48. a. $r = 0.333$ **(LO5&6)**

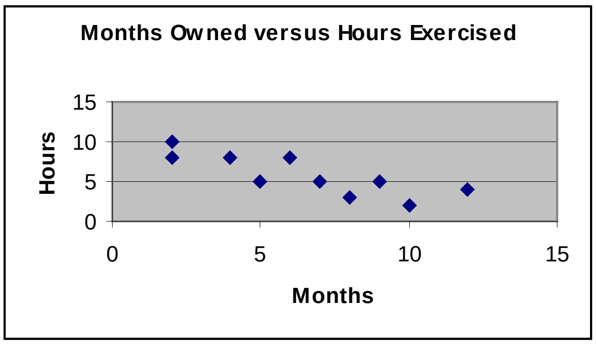
b. $r^2 = (0.333)^2 = 0.111$

c. H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.812$

$$t = \frac{0.333\sqrt{12-2}}{\sqrt{1-(.333)^2}} = 1.12$$

H_o is not rejected. We have not shown this to be a relationship between the variables.

49. a. It looks to be an inverse relationship between the variables. As the months owned increases the number of hours exercised decreases.



b. $$r = \frac{10(313)-(65)(58)}{\sqrt{[10(523)-(65)^2][10(396)-(58)^2]}} = -0.827$$

c. H_o: $\rho \ge 0$ H_1: $\rho < 0$ Reject H_o if $t < -2.896$

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

Reject $H_{o.}$ We can conclude that there is a negative association between months owned and hours exercised. **(LO1,5&6)**

50. **(LO5)**

Source	*SS*	*df*	*MS*	*F*
Regression	300	1	300	54.0
Error	100	18	5.556	
Total	400	19		

a. $$s_e = \sqrt{\frac{100}{18}} = 2.3570$$

b. $r^2 = 300/400 = 0.75$

c. $r = \sqrt{0.75} = -0.866$ The sign of r is negative because the sign of b is negative.

51. a.

Source	*SS*	*df*	*MS*	*F*
Regression	50	1	50	2.5556
Error	450	23	19.5652	
Total	500	24		

b. $n = 25$

c. $s_e = \sqrt{19.5652} = 4.4233$

d. $r^2 = 50/500 = 0.10$ **(LO5)**

52. a. $Y' = 17.08 + 0.16(50) = 25.08$

b. $25.08 \pm 3.182(4.05)\sqrt{1 + \frac{1}{5} + \frac{(50 - 42)^2}{1030}} = 25.08 \pm 14.48$

$= [10.60, 39.56]$ **(LO3&4)**

53. a. $n = 15$ $\sum X = 107$ $\sum X^2 = 837$ $\sum Y = 118.6$

$\sum Y^2 = 969.92$ $\sum XY = 811.60$ $s_{y \cdot x} = 1.114$

$$b = \frac{15(811.60) - (107)(118.6)}{15(837.0) - (107)^2} = \frac{-516.2}{1106.0} = -0.4667$$

$$a = \frac{118.6}{15} - (-0.4667)\left(\frac{107}{15}\right) = 11.2358$$

More bidders decrease winning bid.

b. $Y' = 11.2358 - 0.4667(7.0) = 7.9689$

c. $$7.9689 \pm (2.160)(1.114)\sqrt{1 + \frac{1}{15} + \frac{(7 - 7.1333)^2}{837 - \frac{(107)^2}{15}}} = 7.9689 \pm 2.4854$$

[5.4835,10.4543]

d. $r^2 = 0.499$. The number of bidders explains nearly 50 percent of the variation in the amount of the bid. **(LO3,4&5)**

54. a. $$b = \frac{15(13{,}114.64) - (1193.8)(163.60)}{15(126{,}252.04) - (1193)^2} = 0.0030$$

$$a = \frac{163.6}{15} - (0.0030)\left(\frac{1193.8}{15}\right) = 10.6678$$

b. $r^2 = (0.466)^2 = 0.2172$; no it is not a strong predictor **(LO3&4)**

55. a. $$b = \frac{30(18{,}924) - (320.33)(1575.6)}{30(4292.5) - (320.33)^2} = 2.41 \qquad a = \frac{1575.6}{30} - 2.41\left(\frac{320.33}{30}\right) = 26.8$$

The regression equation is: Price = 26.8 + 2.41 dividend. Foe each additional dollar of dividend, the price increases by $2.41.

b. $r^2 = \frac{5057.6}{7682.7} = 0.658$ 65.8% of the variation in price is explained by the dividend.

c. $r = \sqrt{.658} = 0.811$ $H_0\text{: } \rho \le 0$ $H_1\text{: } \rho > 0$

At the 5% level, reject H_o when $t > 1.701$. $t = \frac{0.811\sqrt{30-2}}{\sqrt{1-(0.811)^2}} = 7.34$

Thus H_o is rejected. The population correlation is positive. **(LO3,5&6)**

56. a. 20, found by one more than the total degrees of freedom (19 + 1

b. 25.88, found by the square root of mean square error $\left(\sqrt{670}\right)$.

c. 0.46, found by SSR/SStotal $\frac{10,354}{22,408}$.

d. 0.68, found by $\sqrt{0.46}$

e. Yes, because the *t*- value (3.93) is greater than the critical value (1.734) and the *p*- value (0.001) is less than the significance level (0.05). **(LO5&6)**

57. a. 35, found by one more than the total degrees of freedom (34 + 1)

b. 5457, found by the square root of MSE (mean square error) $\left(\sqrt{29,778,406}\right)$.

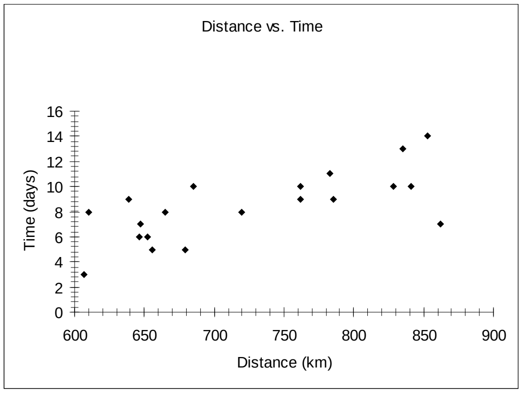
c. 0.93, found by SSR/SStotal $\frac{13,548,662,082}{14,531,349,474}$.

d. 0.97, found by $\sqrt{0.93}$

e. H_o: $\rho \leq 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.692$

Yes, because the *t*- value (21.33) is greater than the critical value (1.692) and the *p*- value (0.000) is less than the significance level (0.05). Reject the null hypothesis. There is a positive association between market value and size of the home. **(LO5&6)**

58. a. There is a lot of scatter in the graph but there appears to be a positive relationship between the two variables. As the distance increases, so does the shipping time.



b. $r = \frac{12(\Sigma 5548) - (\Sigma 28)(\Sigma 2356)}{\sqrt{\left[12(74) - (\Sigma 28)^2\right]\left[12(465226) - (\Sigma 2356)^2\right]}} = .333$

H_o: $\rho \le 0$ H_1: $\rho > 0$ Reject H_o if $t > 1.734$

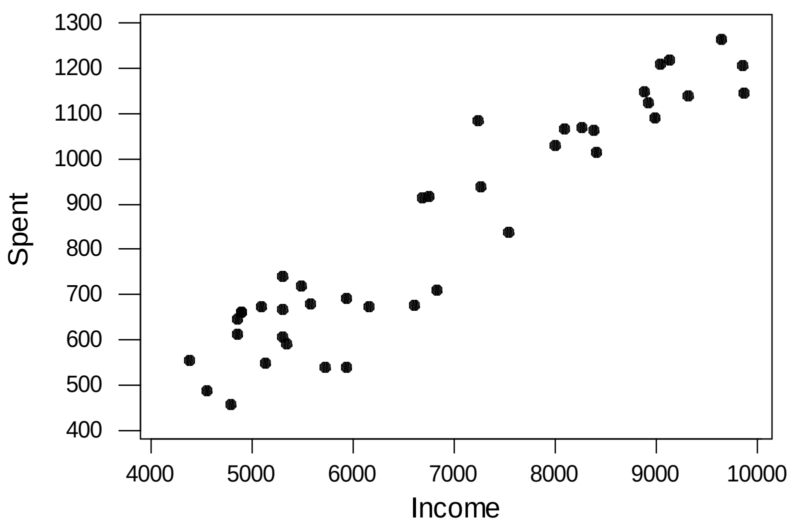
$$t = \frac{0.692\sqrt{20-2}}{\sqrt{1-(0.692)^2}} = 4.067 \quad t = \frac{0.333\sqrt{12-2}}{\sqrt{1-(.333)^2}} = 1.117$$

Do not reject H_o. There is not positive association between shipping distance and shipping time.

c. $r^2 = 0.110$; only 11% of the variation in shipping time is explained by shipping distance.

d. $S_e = 15.390$ **(LO1,5&6)**

59. a.



b. $$b = \frac{40(245,795,835) - (273,387)(33,625)}{40(1,987,875,615) - (273,387)^2} = 0.13388$$

$$a = \frac{33,625}{40} - 0.13388\left(\frac{273,387}{40}\right) = -74.4$$

The regression equation is Spent = - 74.4 + 0.134 Income. For each additional dollar of income, 13.4 cents more is spent on groceries.

c. $$r = \frac{40(245,795,835) - (273,387)(33,625)}{\sqrt{[40(1,987,875,615) - (273,387)^2][40(30,662,885) - (33,625)^2]}} = 0.945$$

H_o: $\rho \le 0$ H_1: $\rho > 0$ At the 5% level, reject H_o when $t > 1.686$.

$$t = \frac{0.945\sqrt{40-2}}{\sqrt{1-(0.945)^2}} = 17.8$$

Thus H_o is rejected. The population correlation is positive.

d. We know is that there is a strong positive association between income and groceries; however, other factors such as location and growth in the area need to be considered. **(LO1,3,5&6)**

60. **(LO3,4,5&6)**

a.

Regression Analysis

r^2	0.662	n	98
r	0.813	k	1
Std. Error	243550.879	Dep. Var.	**List Price**

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Regression	11,129,891,751,449.2000	1	11,129,891,751,449.2000	187.63
Residual	5,694,434,957,986.9500	96	59,317,030,812.3640	
Total	16,824,326,709,436.1000	97		

Regression output

variables	*coefficients*	*std. error*	*t (df=96)*	*p-value*
Intercept	42,424.7503			
Total Square Feet	307.0578	22.4163	13.698	2.59E-24

Predicted values for: List Price

		95% Confidence Interval		*95% Prediction Interval*
Total Square Feet	*Predicted*	*lower*	*upper*	*lower*
2,200	717,951.953	664,470.752	771,433.155	231,558.100

Y' = 42 424.75 + 307.06 Square Feet
Predicted value = $717,952
95% CI: $664 471 to $771 433
95% PI: $231 558 to $1 204 346

b.

Regression Analysis

r^2	0.409	n	98
r	0.640	k	1
Std. Error	321821.210	Dep. Var.	**List Price**

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Regression	6,881,713,174,946.5900	1	6,881,713,174,946.5900	66.45
Residual	9,942,613,534,489.5400	96	103,568,890,984.2660	
Total	16,824,326,709,436.1000	97		

Regression output

variables	*coefficients*	*std. error*	*t (df=96)*	*p-value*
Intercept	55,108.7106			
Number of Bedrooms	205,795.1982	25,246.5254	8.151	1.36E-12

Predicted values for: List Price

		95% Confidence Interval		95% Prediction Interval
Number of Bedrooms	*Predicted*	*lower*	*upper*	*lower*
3	672,494.305	603,083.901	741,904.709	29,924.412

Y' = 55 108.71 + 205 795.20 Number of Bedrooms
Predicted value = $672 494
95% CI: $603 084 to $741 905
95% PI: $29 924 to $1 315 064

c.

Total Square Feet *p-value* 2.59E-24
r 0.813
The size of the house and the list price are positively correlated.
The p-value is very close to zero, and therefore, the variable size is significant.

Number of Bedrooms *p-value* 1.36E-12
r 0.640
The number of bedrooms and the list price are positively correlated.
The p-value is very close to zero, and therefore, the number of bedrooms is significant.

61. **(LO3,4,5&6)**
a.
Regression Analysis

r^2	0.722	n	50
r	0.850	k	1
Std. Error	105034.932	Dep. Var.	**List Price**

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Regression	1,374,442,787,920.0700	1	1,374,442,787,920.0700	124.58
Residual	529,552,172,748.9060	48	11,032,336,932.2689	
Total	1,903,994,960,668.9800	49		

Regression output

variables	*coefficients*	*std. error*	*t (df=48)*	*p-value*
Intercept	44,616.1734			
Total Square Feet	229.1513	20.5302	11.162	6.14E-15

Predicted values for: List Price

	95% Confidence Interval	95% Prediction Interval

Total Square Feet	*Predicted*	*lower*	*upper*	*lower*
2,000	502,918.683	459,924.894	545,912.472	287,399.856

Y' = 44 616.17 + 229.15 Square Feet
Predicted value = $502 919
95% CI: $459 925 to $545 912
95% PI: $287 400 to $718 438

b.

Regression Analysis

r^2	0.292	n	50
r	0.541	k	1
Std. Error	167552.835	Dep. Var.	**List Price**

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Regression	556,445,242,159.6050	1	556,445,242,159.6050	19.82
Residual	1,347,549,718,509.3700	48	28,073,952,468.9453	
Total	1,903,994,960,668.9800	49		

Regression output

Variables	*coefficients*	*std. error*	*t (df=48)*	*p-value*
Intercept	106,724.2293			
Number of Bedrooms	85,690.7445	19,247.5067	4.452	.0001

Predicted values for: List Price

		95% Confidence Interval		95% Prediction Interval
Number of Bedrooms	*Predicted*	*lower*	*upper*	*lower*
4	449,487.207	377,918.906	521,055.509	105,081.572

Y' = 106 724.23 + 85 690.74 Number of Bedrooms
Predicted value = $449 487
95% CI: $377 919 to $521 056
95% PI: $105 082 to $793 893

c.

Total Square Feet *p-value* 6.14E-15
r 0.850

The size of the house and the list price are positively correlated.
The p-value is very close to zero, and therefore, the variable size is significant.

Number of Bedrooms	*p-value*	.0001
	r	0.541

The number of bedrooms and the list price are positively correlated.
The p-value is .0001, and therefore, the variable number of bedrooms is significant.

62. **(LO5&6)**

Correlation Matrix

	World Box Office ($ million)	*Adjusted Budget ($ million)*
World Box Office ($ million)	1.000	
Adjusted Budget ($ million)	.027	1.000

$H_0: \rho \leq 0$ $H_1: \rho > 0$

At the 5% level, reject H_o if $t > 1.677$

$$t = \frac{0.027\sqrt{50 - 2}}{\sqrt{1 - (0.027)^2}} = 0.187$$

There is not enough evidence to reject Ho.
The population correlation is not necessarily negative.
Big budget movies do not always lead to large box office returns.

CHAPTER 13

MULTIPLE REGRESSION AND CORRELATION ANALYSIS

1. a. Multiple regression equation
   b. the Y-intercept
   c. \$374 748 found by $Y' = 64\,100 + 0.394(796\,000) + 9.6(6940) - 11\,600(6.0)$ **(LO1)**

2. a. Multiple regression equation
   b. One dependent, four independent
   c. A regression coefficient
   d. 0.002
   e. 105.014, found by $Y' = 11.6 + 0.4(6) + 0.286(280) + 0.112(97) + 0.002(35)$ **(LO1)**

3. a. 497.736, found by $Y' = 16.24 + 0.017(18) + 0.0028(26{,}500) + 42(3) + 0.0012(156{,}000) + 0.19(141) + 26.8(2.5)$
   b. Two more social activities. Income added only 28 to the index; social activities added 53.6. **(LO1)**

4. a. 21.72 cubic feet
   b. reduction of .582 cubic feet
   c. It is logical, the more insulation the less heat needed, the higher the temperature the less heat needed **(LO1)**

5. a. 65
   b. 2
   c. 1
   d. $s_{Y.12} = \sqrt{\dfrac{SSE}{n-(k+1)}} = \sqrt{\dfrac{583.693}{65-(2+1)}} = \sqrt{9.414} = 3.068$

      95% of the residuals will be between ±6.136, found by 2(3.068)
   e. $R^2 = \dfrac{SSR}{SStotal} = \dfrac{77.907}{661.6} = .118$

      The independent variables explain 11.8% of the variation. **(LO2&3)**

6. a. 52
   b. 5
   c. 1
   d. $s_{Y.12345} = \sqrt{\dfrac{SSE}{n-(k+1)}} = \sqrt{\dfrac{2647.38}{52-(5+1)}} = \sqrt{57.55} = 7.59$

      95% of the residuals will be between ±15.17, found by 2(7.59)
   e. $R^2 = \dfrac{SSR}{SStotal} = \dfrac{3710}{6357.38} = .584$

      The independent variables explain 58.4% of the variation. **(LO2&3)**

7. a. $Y' = 20 - X_1 + 12X_2 - 15X_3$
   b. $Y' = 20 - (4) + 12(6) - 15(8) = -32$

c. n = 22; 3 independent variables

d.

Source	*SS*	*DF*	*MS*	*F*
Regression	7,500	3	2500	18
Error	2,500	18	138.89	
Total	10,000	21		

e. H_o: $b_1 = b_2 = b_3 = 0$ H_1: Not all b's are 0 Reject H_o if $F > 3.16$
F = 18.0 Reject H_o. Not all net regression coefficients equal zero.

f.

For X_1	for X_2	for X_3
H_o: $b_1 = 0$	H_o: $b_2 = 0$	H_o: $b_3 = 0$
H_1: b_1 ¹ 0	H_1: b_2 ¹ 0	H_1: b_3 ¹ 0
$t = -4.00$	$t = 1.50$	$t = -3.00$

Reject H_o if $t > 2.101$ or $t < -2.101$
Delete variable 2, keep 1 and 3 **(LO3,4&5)**

8. a. $Y' = -150 + 2000\,X_1 - 25\,X_2 + 5\,X_3 - 300\,X_4 + 0.60\,X_5$

b. $Y' = -150 + 2000(4) – 25(6) + 5(8) - 300\,(6) + 0.60(8) = 5944.8$

c. n = 21; 5 independent variables

d.

Source	*SS*	*DF*	*MS*	*F*
Regression	1500	5	300	9.00
Error	500	15	33.33	
Total	2000	20		

e. H_o: $b_1 = b_2 = b_3 = b_4 = b_5 = 0$ H_1: Not all b's are 0
Reject H_o if $F > 2.90$
Reject H_o. Not all of the regression coefficients are zero.

f.

For X_1	for X_2	for X_3	for X_4	for X_5
H_o: $b_1 = 0$	H_o: $b_2 = 0$	H_o: $b_3 = 0$	H_o: $b_4 = 0$	H_o: $b_5 = 0$
H_1: b_1 ¹ 0	H_1: b_2 ¹ 0	H_1: b_3 ¹ 0	H_1: b_4 ¹ 0	H_1: b_5 ¹ 0
$t = 4.00$	$t = -0.833$	$t = 1.00$	$t = -3.00$	$t = 4.00$

Reject H_o if $t > 2.131$ or $t < -2.131$
Delete variable 2 then rerun, perhaps delete variable 3 also. **(LO3,4&5)**

9. a. X_4 at -.819 had the strongest correlation with the dependent variable

b. X_2, X_3 *and* X_4 have the strongest correlation with the dependent variable

c. yes, between X_3 and X_4. **(LO7)**

10. a. Income at .806 is the most highly correlated with Sales.

b. Income (.806) and Population (.510)

c. No; there is no evidence of multicollinearity. **(LO7)**

11. a. Horsepower is the most highly correlated with speed at 0.83.

b. It is reasonable as the weight of a car would probably slow it down.

c. No, none of the independent variables is highly correlated with each other. **(LO7)**

12. a. Location is the most highly correlated with the selling price.

b. Yes, the number of bedrooms and the number of bathrooms are highly correlated. If both are left in the equation, incorrect conclusions about statistical significance may be reached, so one variable should be dropped. **(LO7)**

13. a. n = 40

b. 4

c. $R^2 = 750/1250 = 0.60$

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

e. H_o: $b_1 = b_2 = b_3 = b_4 = 0$ H_1: Not all b's equal 0

H_o is rejected if $F > 2.65$

$$F = \frac{750/4}{500/35} = 13.125$$

H_o is rejected. At least one b_I does not equal zero **(LO2,3&4)**

14. H_o: $b_1 = 0$ H_o: $b_2 = 0$

H_1: b_1 ¹ 0 H_1: b_2 ¹ 0

H_o is rejected if $t < -2.074$ or $t > 2.074$

$$t = \frac{2.676}{0.56} = 4.779 \qquad t = \frac{-0.880}{0.710} = -1.239$$

The second variable can be deleted. **(LO4&5)**

15. a. $n = 26$

b. $R^2 = 100/140 = 0.7143$

c. 1.4142, found by $\sqrt{2}$

d. H_o: $b_1 = b_2 = b_3 = b_4 = b_5 = 0$ H_1: Not all b's are 0 Reject H_o if $F > 2.71$

Computed $F = 10.0$. Reject H_o. At least one regression coefficient is not zero.

e. H_o is rejected in each case if $t < -2.086$ or $t > 2.086$. X_1 and X_5 should be dropped.

(x1=1.33 do not reject; x2=15 reject; x3=4 reject; x4=-2.5 reject; x5=.75 do not reject)

(LO2,3,4&5)

16. H_o: $b_1 = b_2 = b_3 = b_4 = b_5 = 0$ H_1: Not all b's equal zero.
$df_1 = 5$ $df_2 = 20 - (5 + 1) = 14$, so H_o is rejected if $F > 2.96$

Source	*SS*	*df*	*MSE*	*F*
Regression	448.28	5	89.656	17.58
Error	71.40	14	5.10	
Total	519.68	19		

So H_o is rejected. Not all the regression coefficients equal zero. **(LO2&4)**

17. a. $28,000

b. 0.5809 found by $R^2 = \frac{SSR}{SStotal} = \frac{3050}{5250}$

c. 9.199 found by $\sqrt{84.62}$

d. H_o is rejected if $F > 2.97$ (approximately)   Computed $F = 1016.67/84.62 = 12.01$
H_o is rejected. At least one regression coefficient is not zero.

e. If computed t is to the left of - 2.056 or to the right of 2.056, the null hypothesis in each of these cases is rejected. Computed t for X_2 and X_3 exceed the critical value. Thus, "population" and "advertising expenses" should be retained and "number of competitors," X_1 dropped. **(LO2,3,4&5)**

18. a. The strongest correlation is between GPA and HS Marks. There is no problem with multicollinearity.

b. $R^2 = \frac{3.8967}{5.0631} = .7696$

c. H_o is rejected if $F > 5.41$
$F = 1.2989/0.2333 = 5.57$. At least one coefficient is not zero.

d. Any H_o is rejected if $t < -2.571$ or $t > 2.571$. It appears that only HS Marks is significant. Verbal and Math could be eliminated. Also, Verbal and Math could be eliminated as their p-values are > .05.

e. $R^2 = 3.7159/5.0631 = 0.7339$, which not too different from the model in b where Math and Verbal were included.

f. The normality assumption seems reasonable since the graph is fairly linear.

g. There appears to be no violation of homoscedasticity. **(LO3,4,5&6)**

19. a. The strongest relationship is between sales and income (0.964). A problem could occur if both "outlets" and "income" (.825) and "cars" and "outlets" (.775) are part of the final solution (.775). This is called multicollinearity.

b. Y′ = -19.6715-.0006(outlets)+1.7399(cars)+.4099(income)+2.0357(age) -.0344(supervisor); $R^2 = \frac{1593.81}{1602.89} = 0.9943$

c. H_o is rejected. At least one regression coefficient is not zero. The computed value of F is 140.36. Critical value = 6.26; therefore, reject the null hypothesis (Note that the p-value = .0001, which is < than the significance level, and supports the decision to reject the null.

d. Delete "outlets" and "supervisors". Critical values are - 2.776 and 2.776. Note that "age" is also insignificant.

e. $R^2 = \frac{1593.66}{1602.89} = 0.9942$ There was little change in the coefficient of determination. Note that age is now a significant variable.

f. The normality assumption seems reasonable since the graph is fairly linear.

g. There appears to be no violation of homoscedasticity. **(LO3,4,5&6)**

20. a. The correlation matrix is:

	Salary	Years	Rating
Years	0.868		
Rating	0.547	0.187	
Master	0.311	0.208	0.458

Years has the strongest correlation with salary. There does not appear to be a problem with multicollinearity.

b.

Regression output

variables	coefficients	std. error	t (df=16)	p-value
Intercept	43.9152	1.9163	5.174	.0001
Years	0.8994	0.0877	10.258	1.93E-08
Rating	0.1539	0.0314	4.895	.0002
Masters	-0.6673	1.2139	-0.550	.5901

The regression equation is: Y¢= 43.9152 + 0.8994X_1 + 0.1539X_2 - 0.6673 X_3

Years	Rating	Masters	Predicted
5	60	0	57.6470488

Y¢= 57.647 or $57 647

c. H_o is rejected if $F > 3.24$; Computed $F = 301.06/5.71 = 52.72$

H_o is rejected. At least one regression coefficient is not zero.

d. A regression coefficient is dropped if computed t is to the left of - 2.120 or right of 2.120. Keep "years" and "rating"; drop "masters." (See the regression output above).

e. Dropping "masters", we have:

Salary = 44.1157 + 0.8926(years) + 0.1464(rating)

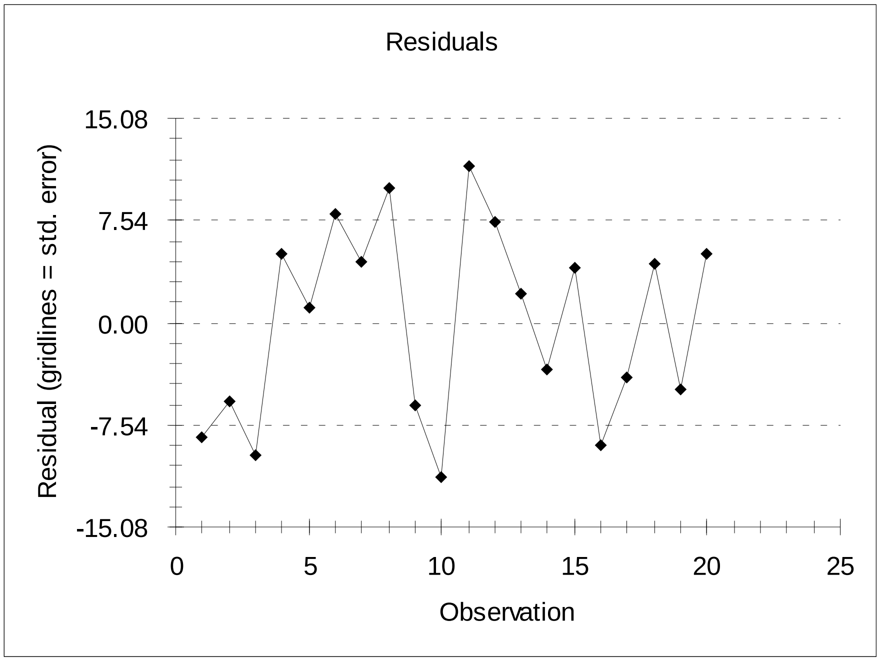
Regression output

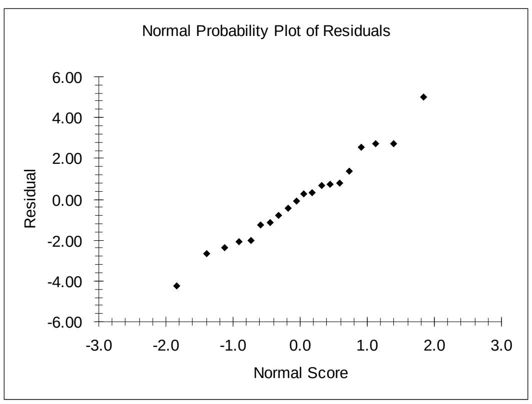
variables	coefficients	std. error	t (df=17)	p-value
Intercept	44.1157	1.8422	5.491	3.98E-05
Years	0.8926	0.0850	10.499	7.54E-09
Rating	0.1464	0.0277	5.283	.0001

f.

Observation	Y	Predicted	Residual
1	55.10	56.38	-1.28
2	57.60	54.87	2.73
3	53.30	53.37	-0.07
4	67.00	66.29	0.71
5	62.60	64.62	-2.02
6	69.00	68.32	0.68
7	66.00	63.27	2.73
8	60.80	58.27	2.53
9	72.60	71.80	0.80
10	55.70	59.97	-4.27
11	49.70	49.40	0.30
12	54.60	55.02	-0.42
13	75.80	76.94	-1.14
14	70.70	70.42	0.28
15	62.40	64.78	-2.38
16	57.60	60.27	-2.67

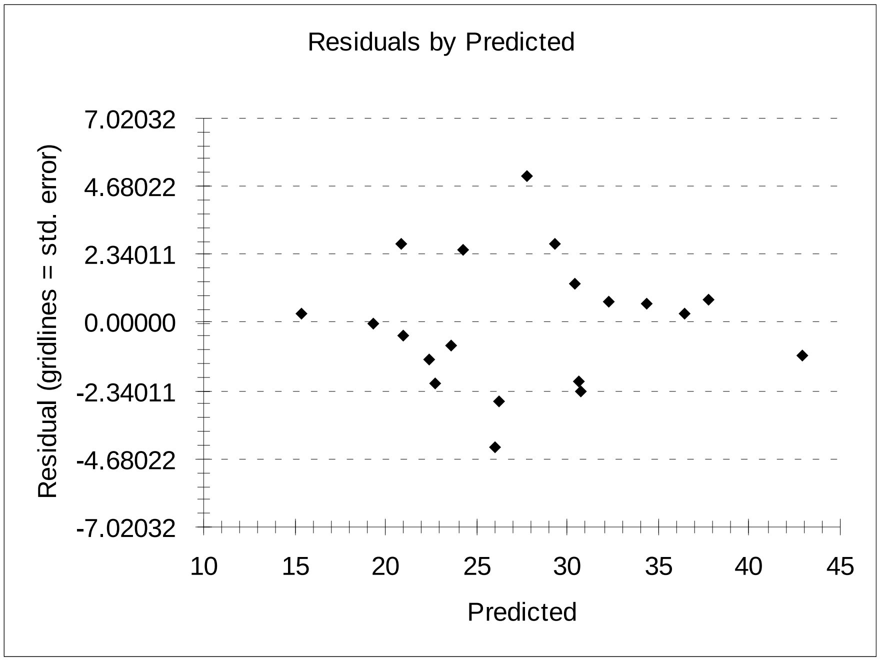
17	65.80	64.43	1.37
18	54.70	56.76	-2.06
19	56.80	57.61	-0.81
20	66.80	61.80	5.00





The distribution of the residuals is approximately normal.

g.



The plot does not appear to violate any assumptions of regression. **(LO3,4,5,6&7)**

21. a. The correlation matrix is:

	Cars	Adv	Sales
Adv	0.808		
Sales	0.872	0.537	
City	0.639	0.713	0.389

Size of sales force (0.872) has strongest correlation with cars sold. Fairly strong relationship between location of dealership and advertising (0.713). Could be a problem.

b. The regression equation is: Y' = 31.1328 + 2.1516adv + 5.0140sales + 5.6651city

Y' = 31.1328 + 2.1516(15) + 5.0140(20) + 5.6651(1) = 169.352

c. H_o: $\beta_1 = \beta_2 = \beta_3 = 0$ H_1: Not all β's are 0

Reject H_o if computed $F > 4.07$

Analysis of Variance

Source	*SS*	*DF*	*MS*
Regression	5504.4	3	1834.8
Error	420.2	8	52.5
Total	5924.7	11	

F = 1834.8/52.5 = 34.95. Reject H_o. At least one regression coefficient is not 0.

d. H_o is rejected in all cases if $t < -2.306$ or if $t > 2.306$. Advertising and sales force should be retained, city dropped. (Note that dropping city removes the problem with multicollinearity.)

Predictor	*Coef*	*Stdev*	*t-ratio*	*P*
Constant	31.13	13.40	2.32	0.049
Adv	2.1516	0.8049	2.67	0.028
Sales	5.0140	0.9105	5.51	0.000
City	5.665	6.332	0.89	0.397

e. The new output is Y' = 25.30 + 2.6187adv + 5.0233sales

Predictor	*Coef*	*Stdev*	*t-ratio*

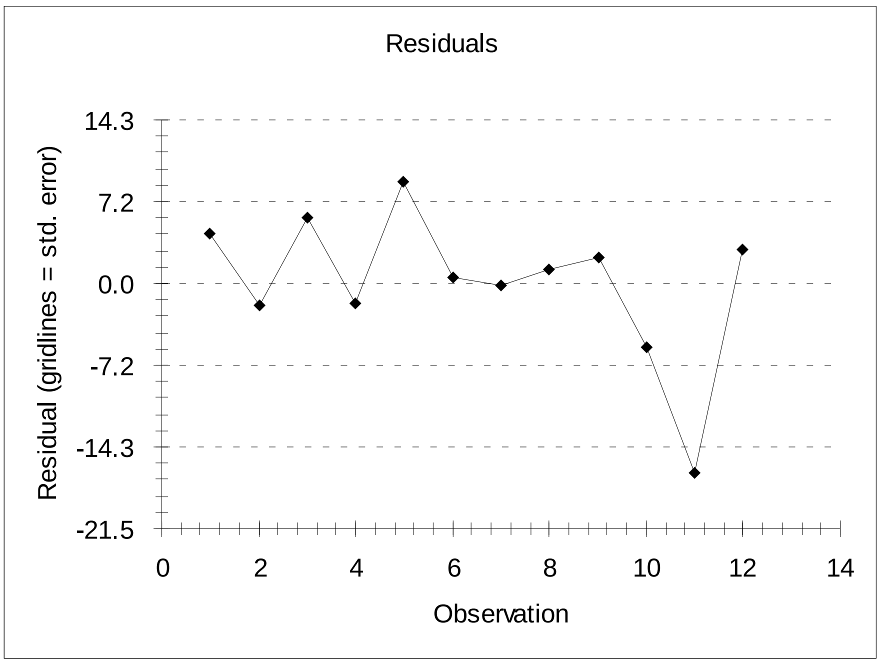
Constant	25.30	11.57	2.19
Adv	2.6187	0.6057	4.32
Sales	5.0233	0.9003	5.58

Analysis of Variance

Source	*SS*	*DF*	*MS*
Regression	5462.4	2	2731.2
Error	462.3	9	51.4
Total	5924.7	11	

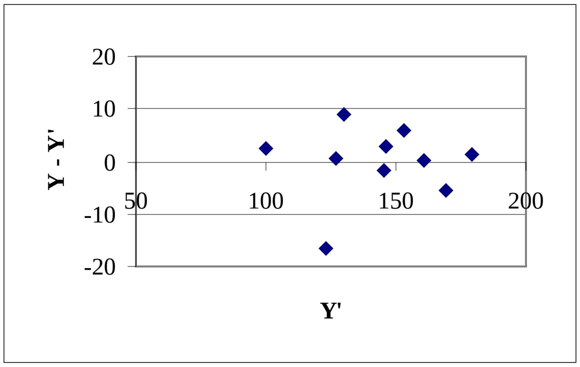
f.

Observation	*Cars*	*Predicted*	*Residual*
1	127.0	122.7	4.3
2	138.0	139.9	-1.9
3	159.0	153.2	5.8
4	144.0	145.8	-1.8
5	139.0	130.1	8.9
6	128.0	127.5	0.5
7	161.0	161.1	-0.1
8	180.0	178.8	1.2
9	102.0	99.7	2.3
10	163.0	168.5	-5.5
11	106.0	122.7	-16.7
12	149.0	146.0	3.0



The normality assumption is reasonable.

g. The critical value could be a problem. However, with a small sample the residual plot is acceptable.



(LO3,4,5,6&7)

22. a. The regression equation is: $Y' = 1480.7 + 0.7315X_1 + 9.991X_2 - 2.308X_3$

b. R^2 = 83.5% or 0.835

c. H_o: $b_1 = b_2 = b_3 = 0$ H_1: Not all b_i's = 0 Reject H_o if $F > 3.59$

$$F = \frac{10{,}057.7/3}{1982.3/11} = 18.60$$

H_o is rejected. Some of the net regression coefficients do not equal zero.

d. H_o: $b_1 = 0$ H_o: $b_2 = 0$ H_o: $b_3 = 0$

H_1: $b_1 \neq 0$ H_1: $b_2 \neq 0$ H_1: $b_3 \neq 0$

Reject H_o if $t < -2.201$ or $t > 2.201$

Reject H_o for area and spaces, do not reject for income. Delete income.

e. $R^2 = 0.804$, $Y' = 1342.49 + 0.7727X_1 + 11.634X_2$ **(LO1,3,4,&5)**

23. a. The regression equation is: $Y' = 965.3 + 2.865X_1 + 6.75X_2 + 0.2873X_3$ = \$2,458,780.

b. Analysis of Variance

Source	*SS*	*DF*	*MS*
Regression	45510101	3	15170034
Error	12215892	12	1017991
Total	57725994	15	

$F = 15170032/1017991 = 14.902$

H_o is rejected because computed F of 14.9 is greater than the critical value of 3.49. At least one of the regression coefficients is not zero.

c. H_o: $b_1 = 0$ H_o: $b_2 = 0$ H_o: $b_3 = 0$

H_1: $b_1 \neq 0$ H_1: $b_2 \neq 0$ H_1: $b_3 \neq 0$

Reject H_o if $t < -2.179$ or $t > 2.179$

variables	*t (df=12)*	*p-value*
Intercept		
X1	1.810	.0954
X2	0.657	.5236
X3	2.586	.0238

X_1=1.810 do not reject; x_2=.657 do not reject; x_3=2.586 reject. Both workers and dividends are not significant variables. Inventory is significant. Delete dividends and rerun the regression equation.

d. The regression equation (if we used X_1 and X_3) is: $Y¢= 1134.8 + 3.258X_1 + 0.3099X_3$

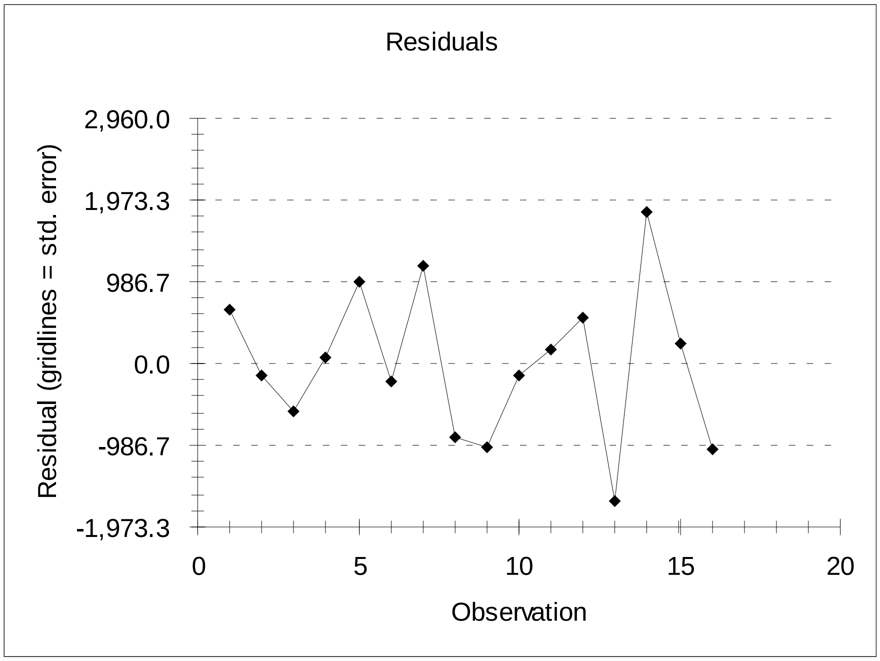
Predictor	*Coef*	*Stdev*	*t-ratio*
Constant	1134.8	418.6	2.71
Workers	3.258	1.434	2.27
Inv	0.3099	0.1033	3.00

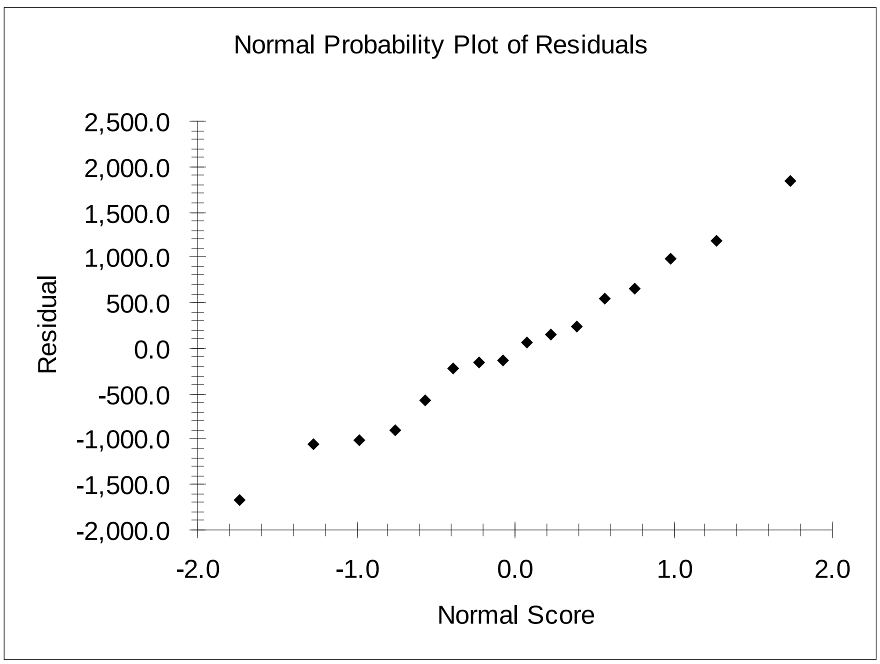
Analysis of Variance

Source	*SS*	*DF*	*MS*	*F*
Regression	45070638	2	22535319	23.15
Error	12655356	13	973489	
Total	57725994	15		

e.

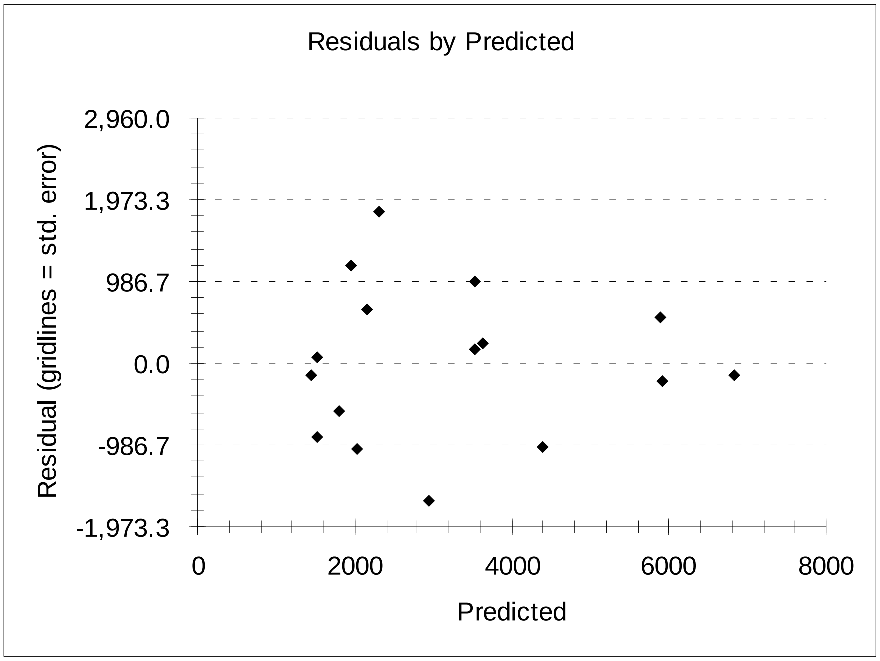
Observation	*Profit*	*Predicted*	*Residual*
1	2,800.0	2,148.6	651.4
2	1,300.0	1,445.7	-145.7
3	1,230.0	1,812.3	-582.3
4	1,600.0	1,532.9	67.1
5	4,500.0	3,520.8	979.2
6	5,700.0	5,920.3	-220.3
7	3,150.0	1,960.0	1,190.0
8	640.0	1,534.7	-894.7
9	3,400.0	4,402.7	-1,002.7
10	6,700.0	6,833.4	-133.4
11	3,700.0	3,539.9	160.1
12	6,440.0	5,885.9	554.1
13	1,280.0	2,939.9	-1,659.9
14	4,160.0	2,322.7	1,837.3
15	3,870.0	3,624.0	246.0
16	980.0	2,026.2	-1,046.2





The normality assumption is reasonable.

f.



(LO3,4,5,&6)

24. a. $Y' = 28.2 + 0.0287X_1 + 0.650X_2 - 0.049X_3 - 0.0004X_4 + 0.723X_5$

b. $R^2 = 0.750$

c & d.

Predictor	*Coef*	*Stdev*	*t-ratio*	*P*
Constant	28.242	2.986	9.46	0.000
Value	0.028669	0.004970	5.77	0.000
Years	0.6497	0.2412	2.69	0.014
Age	- 0.04895	0.3126	- 1.57	0.134
Mortgage	- 0.000405	0.001269	- 0.32	0.753
Gender	0.7227	0.2491	2.90	0.009

$s = 0.5911$ R-sq = 75.0% R-sq(adj) = 68.4%

Analysis of Variance

Source	*DF*	*SS*	*MS*	*F*	*p*
Regression	5	19.8914	3.9783	11.39	0.000
Error	19	6.6390	0.3494		
Total	24	26.5304			

Consider dropping the variables age and mortgage.

e. Drop mortgage first.

R²	0.748		
Adjusted R²	0.698	n	25
R	0.865	k	4
Std. Error	0.578	Dep. Var.	**Income**

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Regression	19.8559	4	4.9640	14.87	8.62E-06
Residual	6.6745	20	0.3337		

Total	26.5304	24

Regression output					confidence interval	
variables	*coefficients*	*std. error*	*t (df=20)*	*p-value*	*95% lower*	*95% upper*
Intercept	28.0638	2.8661	9.791	4.50E-09	22.0851	34.0424
Value	0.0281	0.0046	6.173	4.96E-06	0.0186	0.0376
Years	0.6587	0.2342	2.813	.0107	0.1702	1.1471
Gender	0.7385	0.2386	3.096	.0057	0.2409	1.2362
Age	-0.0490	0.0305	-1.602	.1247	-0.1127	0.0148

The new regression equation is: *Y*¢= 28.0638 + 0.02815Value + 0.6587Years + 0.7358Gender - .0490Age. The *R*-square value is 0.748. Age is still an insignificant variable.

Drop Age.

R^2	0.717		
Adjusted R^2	0.661	n	25
R	0.847	k	4
Std. Error	0.612	Dep. Var.	**Income**

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Regression	19.0344	4	4.7586	12.70	2.65E-05
Residual	7.4960	20	0.3748		
Total	26.5304	24			

Regression output					confidence interval	
variables	*coefficients*	*std. error*	*t (df=20)*	*p-value*	*95% lower*	*95% upper*
Intercept	29.9896	2.8683	10.456	1.49E-09	24.0064	35.9728
Value	0.0259	0.0048	5.385	2.86E-05	0.0159	0.0359
Years	0.3974	0.1860	2.137	.0452	0.0094	0.7853
Gender	0.6920	0.2572	2.690	.0141	0.1554	1.2285
Mortgage	-0.00040490	0.0013	-0.308	.7612	-0.00314625	0.00233646

The new regression equation is: *Y*¢= 29.9896 + 0.0259Value + 0.3974Years + 0.6920Gender + .0004Mortgage. The *R*-square value is 0.717. Mortgage is still an insignificant variable.

Drop age and mortgage.

R^2	0.716		
Adjusted R^2	0.676	n	25
R	0.846	k	3
Std. Error	0.599	Dep. Var.	**Income**

ANOVA

table

Source	SS	df	MS	F	p-value
Regression	18.9989	3	6.3330	17.66	5.90E-06
Residual	7.5316	21	0.3586		
Total	26.5304	24			

Regression output

Variables	coefficients	std. error	t (df=21)	p-value	confidence interval 95% lower	95% upper
Intercept	29.8109	2.7479	10.849	4.57E-10	24.0964	35.5254
Value	0.0253	0.0044	5.803	9.25E-06	0.0163	0.0344
Years	0.4063	0.1797	2.262	.0344	0.0327	0.7800
Gender	0.7079	0.2465	2.872	.0091	0.1952	1.2205

The new regression equation is: $Y¢$= 29.8109 + 0.0253Value + 0.4063Years + 0.7079Gender. The *R*-square value is 0.716. All variables in the equation are significant. **(LO2,3,4&5)**

25. a. The correlation of Income and Median Age = 0.721. So there does appear to be a linear relationship between the two.

b. Income is the "dependent" variable.

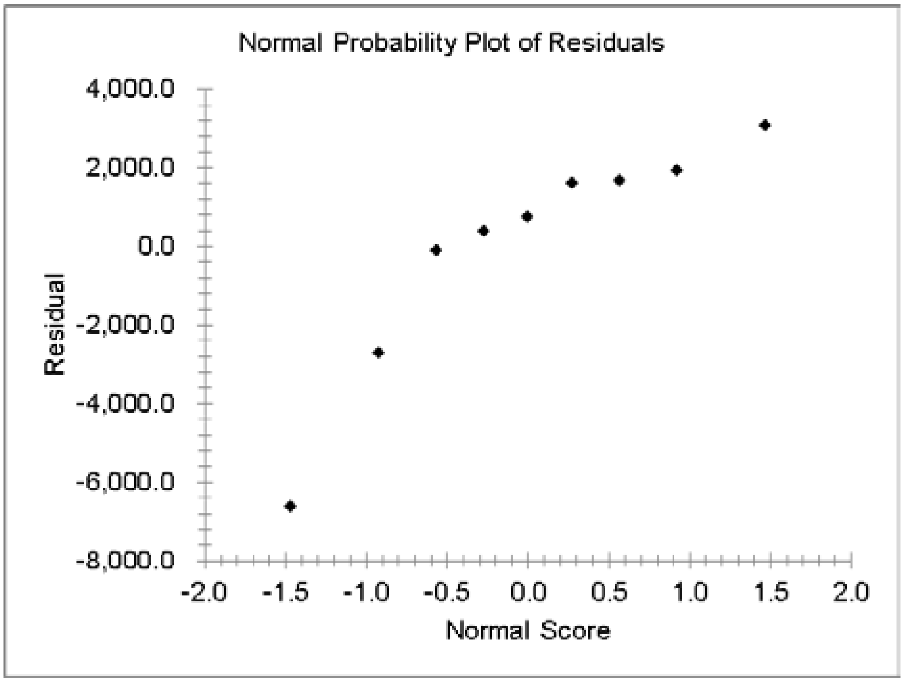
c. The regression equation is Income = 22805 + 362 Median Age. For each year increase in age, the income increases $362 on average.

d. Using an indicator variable for Population > 400 000, the regression equation is Income = 24865 + 251 Median Age + 6888 Population.
That changes the estimated effect of an additional year of age to $251 and the effect of the population > 400 000 as adding $6888 to income.

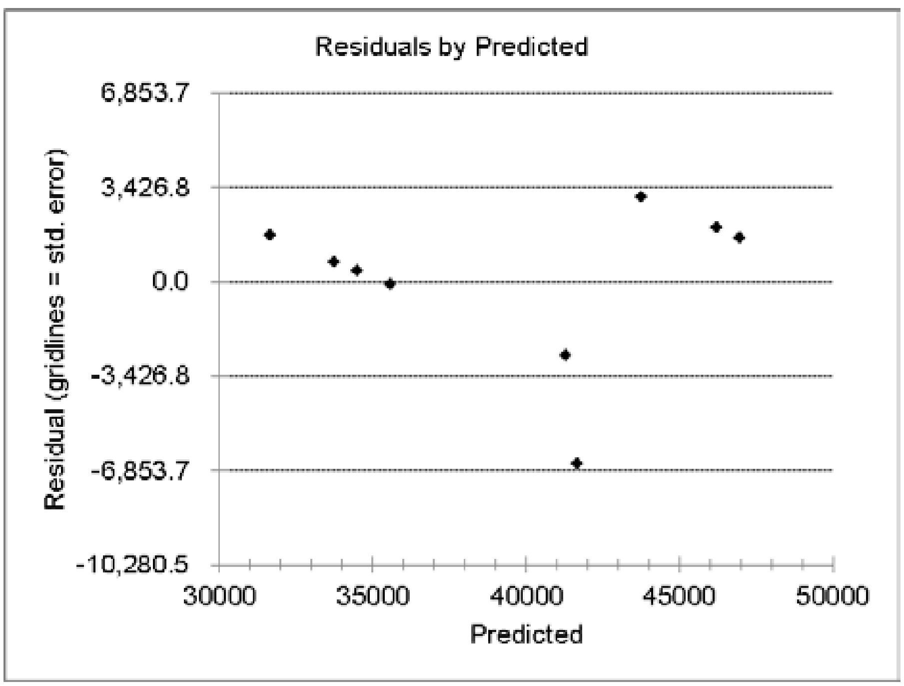
e. Notice the p-values are not more than 5%. This indicates that the independent variables are significant influences on income.

```
Predictor    Coef  SE Coef     T      P
Constant    24865     4552  5.46  0.002
med age     250.5    102.4  2.45  0.050
pop          6888     2501  2.75  0.033
```

f. A histogram of the residuals is close to a normal distribution.



g The variation is about the same across the different fitted values.



(LO1,4,5,6&7)

26. a. The correlation matrix is:

	Food	*Income*
Income	0.156	
Size	0.876	- 0.098

Income and size are not related. There is no multicollinearity.

b. The regression equation is: Food = 2.84 + 0.00613 Income + 0.425 Size. Another dollar of income leads to an increased food bill of 0.6 cents. Another member of the family adds $425 to the food bill.

c. $R^2 = \frac{18.41}{22.29} = 0.826$

To conduct the global test: H_o: = $b_1 = b_2 = 0$ versus H_1: Not all b_i's = 0

At the 0.05 significance level, H_o is rejected if $F > 3.44$.

Source	*DF*	*SS*	*MS*	*F*	*p*
Regression	2	18.4095	9.2048	52.14	0.000
Error	22	3.8840	0.1765		
Total	24	22.2935			

The computed value of F is 52.14, so H_o is rejected. Some of the regression coefficients and R^2 are not zero.

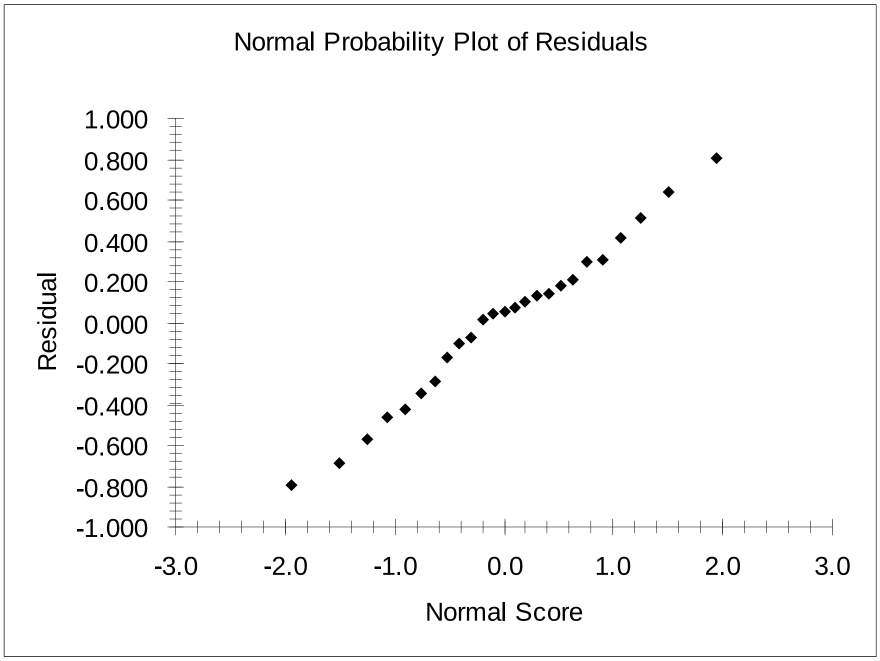
d. Since the *p*-values are less than 0.05, there is no need to delete variables.

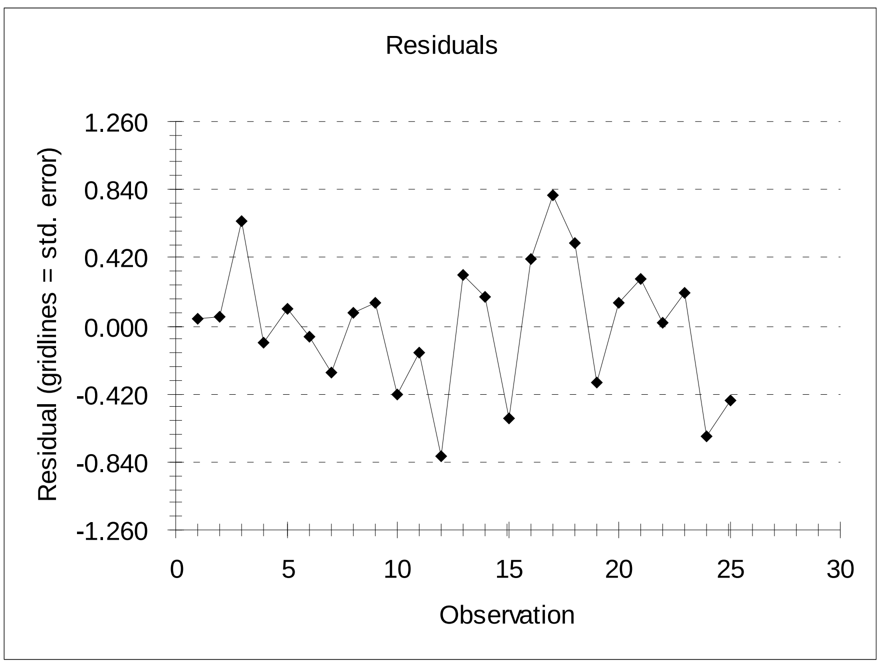
Predictor	*Coef*	*SE Coef*	*T*	*P*
Constant	2.8436	0.2618	10.86	0.000
Income	0.006129	0.002250	2.72	0.012
Size	0.42476	0.04222	10.06	0.000

e. The residuals appear normally distributed.

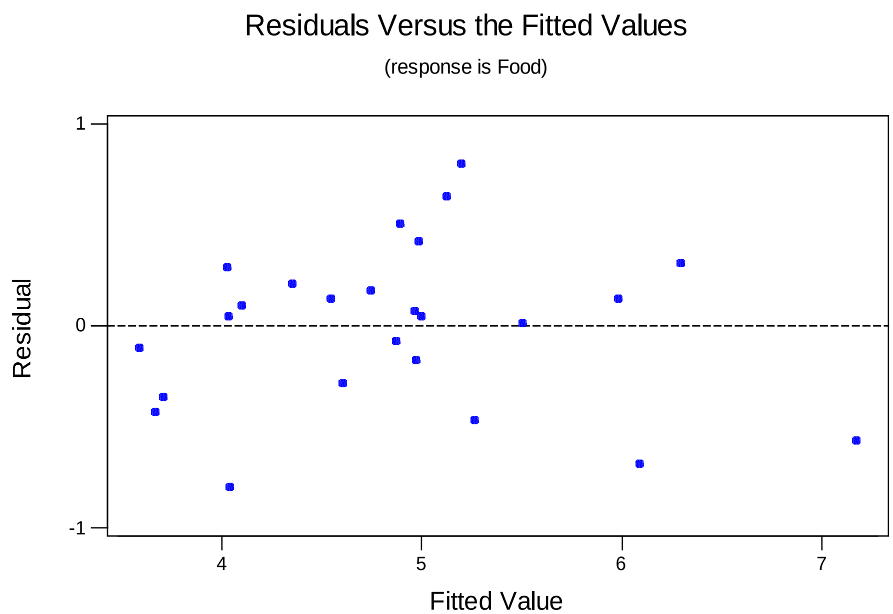
Observation	*Food*	*Predicted*	*Residual*
1	5.040	4.996	0.044
2	4.080	4.030	0.050
3	5.760	5.120	0.640
4	3.480	3.587	-0.107
5	4.200	4.096	0.104
6	4.800	4.871	-0.071
7	4.320	4.606	-0.286
8	5.040	4.963	0.077
9	6.120	5.982	0.138
10	3.240	3.666	-0.426
11	4.800	4.966	-0.166
12	3.240	4.040	-0.800
13	6.600	6.292	0.308
14	4.920	4.743	0.177
15	6.600	7.169	-0.569
16	5.400	4.984	0.416
17	6.000	5.194	0.806
18	5.400	4.891	0.509
19	3.360	3.709	-0.349
20	4.680	4.544	0.136
21	4.320	4.026	0.294
22	5.520	5.505	0.015
23	4.560	4.352	0.208

24	5.400	6.085	-0.685
25	4.800	5.264	-0.464





f. The variance is the same as we move from small values to large. So there is no homoscedasticity problem.



(LO1,3,4,5,6&7)

27. a.

	Salary	*GPA*	*Business*
Salary	1.000		
GPA	.902	1.000	
Business	.911	.851	1.000

Yes; multicollinearity occurs between GPA and Business (.851)

b. Y′ = 23.4474 + 2.7748GPA + 1.3071Business

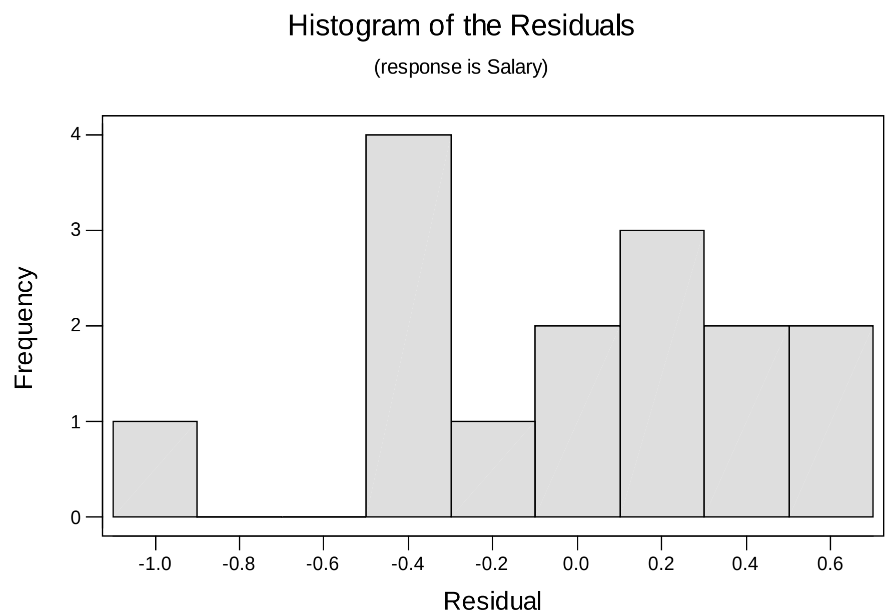
As GPA increases by one point salary increases by $2775. The average business school graduate makes $1307 more than a corresponding non-business graduate. Estimated salary is $33,079; found by $23,447 + 2775(3.00) + 1307(1).

c. $R^2 = \frac{21.182}{23.857} = 0.888$ so, 88.8% of the variation in the *y*-variable is explained or accounted for

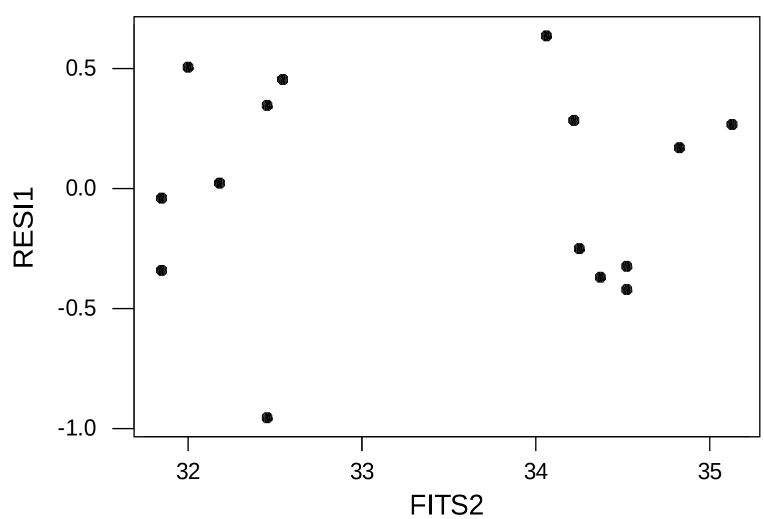
d. Since the *p*-values are less than 0.05, there is no need to delete variables.

Predictor	*Coef*	*SE Coef*	*T*	*P*
Constant	23.447	3.490	6.72	0.000
GPA	2.775	1.107	2.51	0.028
Business	1.3071	0.4660	2.80	0.016

e. The residuals appear normally distributed.



f. The variance is the same as we move from small values to large. So there is no homoscedasticity problem.



(LO1,3,4,5,6&7)

28. a. The regression equation is P/E = 29.9 – 5.32 EPS + 1.45 Yield.

b. Here is part of the software output:

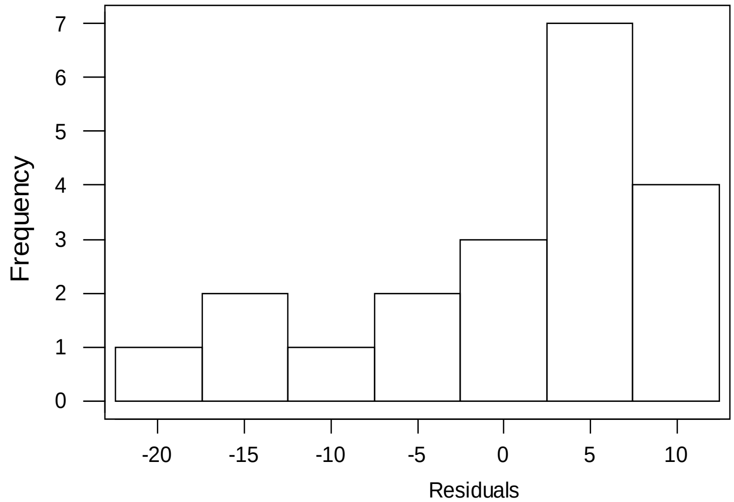
Predictor	*Coef*	*SE Coef*	*T*	*P*
Constant	29.913	5.767	5.19	0.000
EPS	–5.324	1.634	–3.26	0.005
Yield	1.449	1.798	0.81	0.431

Thus EPS has a significant relationship with P/E while Yield does not.

c. As EPS increase by one, P/E decreases by 5.324 and when Yield increases by one, P/E increases by 1.449.

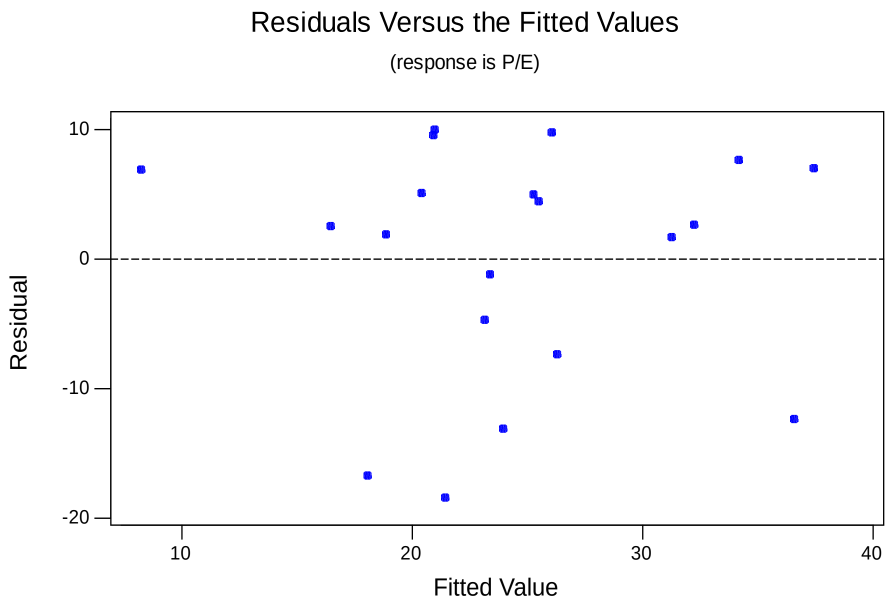
d. The second stock has a residual of -18.43, signifying that it's fitted P/E is around 21.

e.



The distribution of residuals may show a negative skewness.

f.



There does not appear to be any problem with homoscedasticity.

g. The correlation matrix is

	P/E	*EPS*
EPS	–0.602	
Yield	0.054	0.162

The correlation between the independent variables is shown in the correlation matrix. There is no multicollinearity.

(LO1,3,6&7)

29. a. The regression equation is

Sales = 1.02 + 0.0829 Infomercials

Predictor	*Coef*	*SE Coef*	*T*	*P*
Constant	1.0188	0.3105	3.28	0.006

Infomercials	0.08291	0.01680	4.94	0.000	

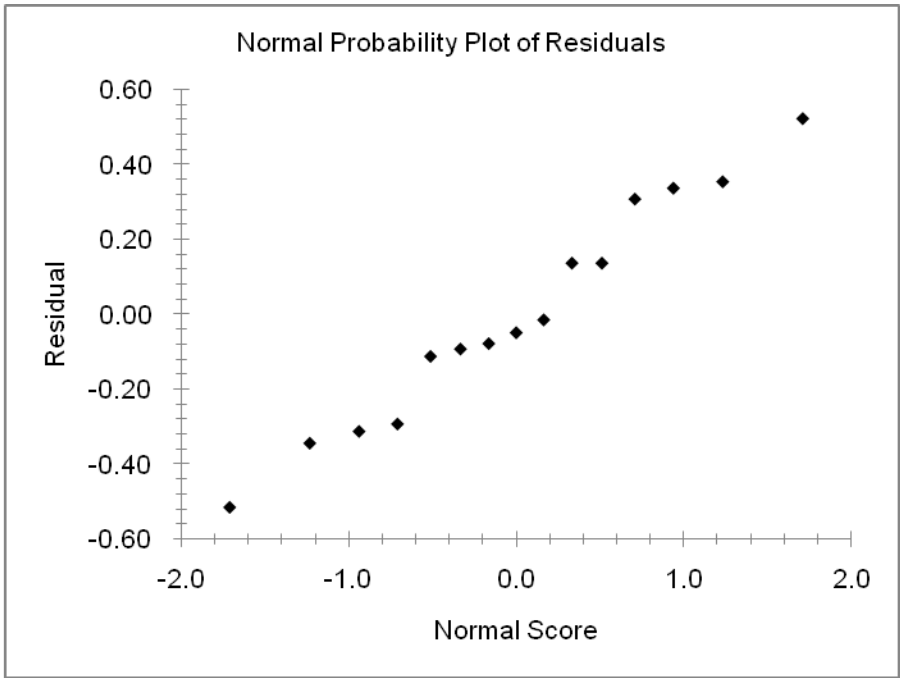
S = 0.308675 *R*-sq = 65.2% *R*-sq(adj) = 62.5%

Analysis of Variance

Source	*DF*	*SS*	*MS*	*F*	*P*
Regression	1	2.3214	2.3214	24.36	0.000
Residual Error	13	1.2386	0.0953		
Total	14	3.5600			

b. The global test on F demonstrates there is a substantial connection between sales and the number of commercials.

c.



d. The residuals appear to be fairly normally distributed.

(LO1,3,5,6&7)

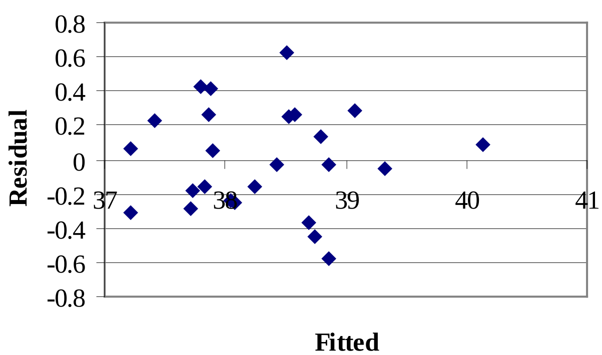
30. a. $\hat{Y} = -5.7328 + 0.00754X_1 + 0.0509X_2 + 1.0974X_3$

b. H_o: $b_1 = b_2 = b_3 = 0$ H_i: Not all b_i's = 0 Reject H_o if $F > 3.07$

$$F = \frac{11.3437/3}{2.2446/21} = 35.38$$

c. All coefficients are significant. Do not delete any.

d. The residuals appear to be random. No problem.



e. The histogram appears to be normal. No problem.

Histogram of Residual $N = 25$

Midpoint	*Count*	
- 0.6	1	*
- 0.4	3	***
- 0.2	6	******
- 0.0	6	******
0.2	6	******
0.4	2	**
0.6	1	*

(LO1,4,5&6)

31. **(LO1,3,4,5,6&7)**

a.

Regression output

variables	*coefficients*	*std. error*	*t (df=93)*	*p-value*	*confidence interval 95% lower*	*95% upper*
Intercept	-39,826.2	63,708.2	-0.625	.5334	-166,338.0	86,685.7
Number of Bedrooms	-22,566.6	34,591.9	-0.652	.5158	-91,259.1	46,126.0
Full Baths	97,362.9	50,802.9	1.916	.0584	-3,521.6	198,247.5
Total Square Feet	334.5	47.9	6.987	4.15E-10	239.4	429.6
Type	-196,886.9	66,993.3	-2.939	.0042	-329,922.3	-63,851.4

List price = -39 826.2 -22 566.6Number of Bedrooms + 97 362.9Full Baths + 334.5Total Square Feet

The number of Full Baths and the Total Square Feet have positive coefficients. The others are negative. We would expect Type to be negative, but not the other variables.

b.

R^2 0.730

73.0% of the variation in the y-variable is explained ot accounted for.

c.

Correlation Matrix

	List Price	*Number of Bedrooms*	*Full Baths*	*Total Square Feet*	*Type*
List Price	1.000				
Number of Bedrooms	.640	1.000			
Full Baths	.736	.646	1.000		

Total Square Feet	.813	.846	.755	1.000	
Type	.315	.652	.264	.619	1.000

Strong correlation: Full baths (.736) and Total Square Feet (.813).

Multicollinearity: exists between Bedrooms and Total Square Feet (.846), and Full baths and Total Square Feet (.755).

d.

H_o: $b_1 = b_2 = b_3 = b_4 = 0$

H_i: Not all b's = 0

p-value
1.24E-25

The *p*-value < .05, so reject the null hypothesis.

e.

Regresson output

variables	*coeficients*	*std. error*	*t (df=93)*	*p-value*
Intercept	-39,826.2	63,708.191	-0.625	.5334
Number of Bedrooms	-22,566.6	34,591.857	-0.652	.5158
Full Baths	97,362.9	50,802.921	1.916	.0584
Total Square Feet	334.5	47.871	6.987	4.15E-10
Type	-196,886.9	66,993.299	-2.939	.0042

H_o: $b_1 = 0$

H_i: $b_1 \neq 0$

The Number of bedrooms is not significant to the model, and should be removed.

f.

Regression output

variables	*coefficients*	*std. error*	*t (df=94)*	*p-value*
Intercept	-53,924.1	59,747.5682	-0.903	.3691
Full Baths	92,038.6	49,989.6131	1.841	.0688
Total Square Feet	318.9	41.3781	7.708	1.30E-11
Type	-211,781.7	62,789.7878	-3.373	.0011

Full Baths should also be removed from the equation

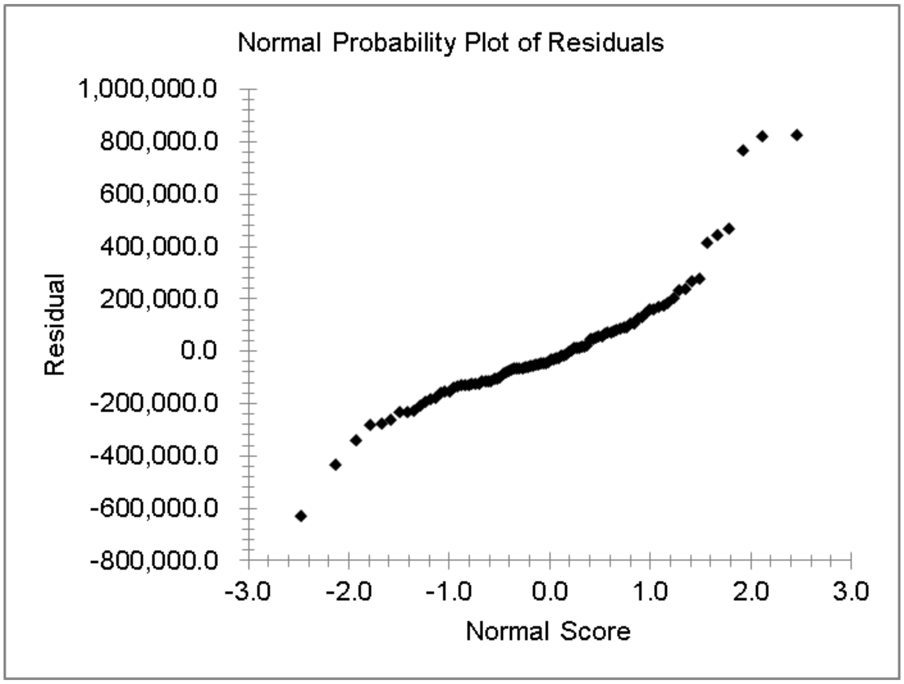
Regression output

variables	coefficients	std. error	t (df=95)	p-value
Intercept	25,442.336	41,888.2426	0.607	.5450
Total Square Feet	378.473	26.1398	14.479	8.98E-26
Type	-257,589.540	58,370.8391	-4.413	2.69E-05

All variables are now significant

g.

The plot appears to be normal.



32. **(LO1,3,4,5,6&7)**

a.

Regression output

variables	coefficients	std. error	t (df=45)	p-value
Intercept	7,736.7	36,730.4542	0.211	.8341
Number of Bedrooms	-16,853.4	19,218.1872	-0.877	.3852
Full Baths	95,210.9	30,620.4307	3.109	.0032
Total Square Feet	136.2	30.5300	4.460	.0001
Type	93,835.8	44,198.9152	2.123	.0393

Y' = 7736.7 -16 853.4Number of Bedrooms + 95 210.9Full Baths + 136.2Square feet + 93 835.8Type

Number of bedrooms is negative, and the other variables are positive.

b.

R^2 0.802

80.2% of the variation in the y-variable is explained ot accounted for.

c.

Correlation Matrix

	List Price	*Number of Bedrooms*	*Full Baths*	*Total Square Feet*	*Type*
List Price	1.000				
Number of Bedrooms	.541	1.000			
Full Baths	.824	.553	1.000		
Total Square Feet	.850	.494	.807	1.000	
Type	.531	.809	.449	.420	1.000

Strong correlation: Full Baths (.824) and Total Square Feet (.850).

Multicollinearity: exists between Bedrooms and Type (.809) and Full Baths and Total Square Feet (.807).

d.

H_o: $b_1 = b_2 = b_3 = b_4 = 0$

H_i: Not all b's = 0

p-value
3.00E-15

The *p*-value < .05, so reject the null hypothesis.
The model is significant.

e.

Regression output

variables	*p-value*
Intercept	.8341
Number of Bedrooms	.3852

H_o: $b_1 = 0$

H_i: $b_1 \neq 0$

Full Baths	.0032
Total Square Feet	.0001
Type	.0393

The Number of bedrooms is not significant to the model, and should be

H_o: $b_1 = 0$

H_i: $b_1 \neq 0$

The Number of bedrooms is not significant to the model, and should be removed.

f.

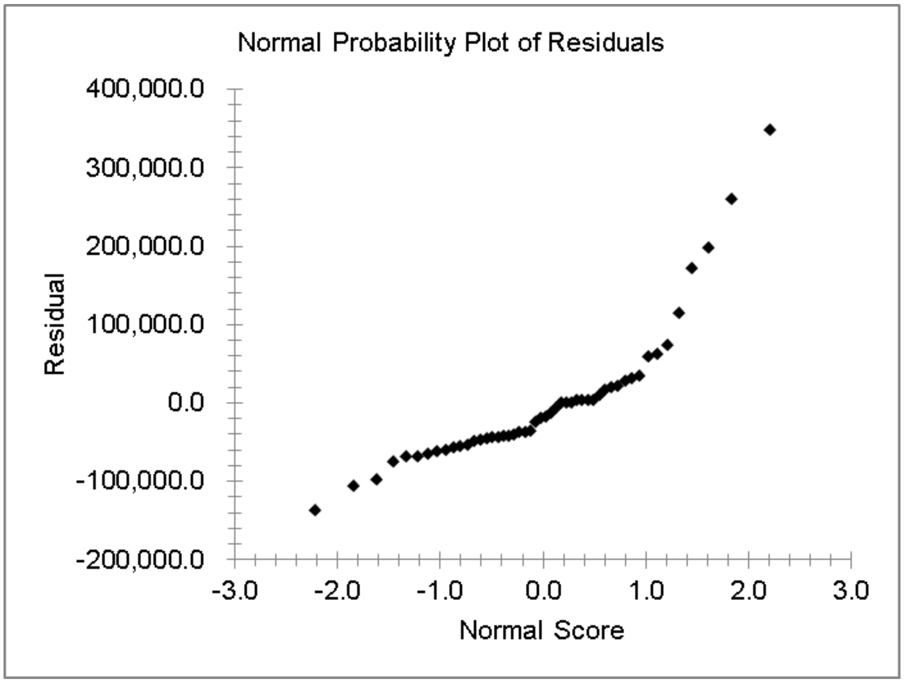
Regression output

variables	*coefficients*	*std. error*	*t (df=46)*	*p-value*
Intercept	-10,990.3053	29,809.7827	-0.369	.7141
Full Baths	89,066.2216	29,733.0534	2.996	.0044
Total Square Feet	135.6088	30.4463	4.454	.0001
Type	64,738.9538	29,126.9911	2.223	.0312

All independent variables are now significant.

g.

The plot appears to be normal.



CHAPTER 14

CHI-SQUARE APPLICATIONS FOR NOMINAL DATA

1. a. 3
   b. 7.815 **(LO1)**

2. a. 5
   b. 15.086 **(LO1)**

3. a. Reject H_o if $\chi^2 > 5.991$
   b. $\chi^2 = \frac{(10-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(30-20)^2}{20} = 10.0$
   c. Reject H_o. The proportions are not equal. **(LO1)**

4. a. Reject H_o if $\chi^2 > 7.815$
   b. $\chi^2 = \frac{(10-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(30-20)^2}{20} + \frac{(20-20)^2}{20} = 10.0$
   c. Reject H_o. The categories are not equal. **(LO1)**

5. H_o: The outcomes are the same H_1: The outcomes are not the same
   Reject H_o if $\chi^2 > 9.236$
   $\chi^2 = \frac{(3-5)^2}{5} + \ldots + \frac{(7-5)^2}{5} = 7.60$
   Do not reject H_o. Cannot reject H_o that outcomes are the same. **(LO1)**

6. H_o: Rounds played is the same for each day H_1: Rounds played is not the same
   Reject H_o if $\chi^2 > 9.488$
   $\chi^2 = \frac{(124-104)^2}{104} + \ldots + \frac{(120-104)^2}{104} = 15.308$
   Reject H_o. The number of rounds played is not the same for each day. **(LO1)**

7. H_o: There is no difference in the proportions H_1: There is a difference in the proportions
   Reject H_o if $\chi^2 > 15.086$
   $\chi^2 = \frac{(47-40)^2}{40} + \ldots + \frac{(34-40)^2}{40} = 3.400$ **(LO1)**
   Do not reject H_o. There is no difference in the proportions.

8. H_o: The proportions are the same H_1: The proportions are not the same
   Reject H_o if $\chi^2 > 18.475$
   $\chi^2 = \frac{(6-10)^2}{10} + \ldots + \frac{(19-10)^2}{10} + \frac{(6-10)^2}{10} = 24.60$
   Reject H_o. The accidents are not evenly distributed throughout the day. **(LO1)**

9. a. Reject H_o if $\chi^2 > 9.210$

b. $$\chi^2 = \frac{(30-24)^2}{24} + \frac{(20-24)^2}{24} + \frac{(10-12)^2}{12} = 2.50$$

c. Do not reject H_o **(LO1)**

10. H_o: The proportions are as stated H_1: The proportions are not as stated

Reject H_o if $\chi^2 > 7.824$

$$\chi^2 = \frac{(60-50)^2}{50} + \frac{(30-25)^2}{25} + \frac{(10-25)^2}{25} = 12.00$$

Reject H_o. Proportions are not as stated. **(LO1)**

11. H_o: The proportions are as stated H_1: The proportions are not as stated

Reject H_o if $\chi^2 > 11.345$

$$\chi^2 = \frac{(50-25)^2}{25} + \frac{(100-75)^2}{75} + \frac{(190-125)^2}{125} + \frac{(160-275)^2}{275} = 115.22$$

Reject H_o. Proportions are not as stated. **(LO1)**

12. H_o: The proportions are as stated H_1: The proportions are not as stated

Reject H_o if $\chi^2 > 7.815$

$$\chi^2 = \frac{(165-150)^2}{150} + \frac{(140-150)^2}{150} + \frac{(125-150)^2}{150} + \frac{(70-50)^2}{50} = 14.333$$

Reject H_o. Proportion of viewers is not as stated **(LO1)**

13. H_o: There is no relationship between size and section read H_1: There is a relationship

Reject H_o if $\chi^2 > 9.488$

$$\chi^2 = \frac{(170-157.50)^2}{157.50} + \ldots + \frac{(88-83.62)^2}{83.62} = 7.340$$

Do not reject H_o. There is no relationship between size and section read. **(LO3)**

14. H_o: There is no relationship between quality and manufacturer

H_1: There is a relationship

Reject H_o if $\chi^2 > 7.815$

$$\chi^2 = \frac{(12-9)^2}{9} + \frac{(8-9)^2}{9} + \ldots + \frac{(89-91)^2}{91} = 3.663$$

Do not reject H_o. There is no relationship between quality and manufacturer. **(LO3)**

15. H_o: No relationship between error rates and item type

H_1: There is a relationship between error rates and item type

Reject H_o if $\chi^2 > 9.21$

$$\chi^2 = \frac{(20-14.1)^2}{14.1} + \frac{(10-15.9)^2}{15.9} + \ldots + \frac{(200-199.75)^2}{199.75} + \frac{(225-225.25)^2}{225.25} = 8.033$$

Do not reject H_o. There is no relationship between error rates and item type.

(Computer value = 8.036) **(LO3)**

16. H_o: No relationship between phone use and accidents
H_1: There is a relationship between phone use and accidents
Reject H_o if $\chi^2 > 3.841$

$$\chi^2 = \frac{(25 - 31.45)^2}{31.45} + \frac{(300 - 293.55)^2}{293.55} + \frac{(50 - 43.55)^2}{43.55} + \frac{(400 - 406.45)^2}{406.45} = 2.523$$

Do not reject H_o. There is no relationship between phone use and accidents. **(LO3)**

17. H_o: $p_s = 0.50$, $p_r = p_e = 0.25$ H_1: Distribution is not as given above.
$df = 2$ Reject H_o if $\chi^2 > 4.605$.

Turn	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
Straight	112	100	12	1.44
Right	48	50	-2	0.08
Left	40	50	-10	2.00
Total	200	200		3.52

H_o is not rejected. The proportions are as given in the null hypothesis. **(LO1)**

18. H_o: $p_s = p_c = p_e$ H_1: The proportions are not equal.
$df = 2$ Reject H_o if $\chi^2 > 5.991$.

Gift	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
Sweatshirt	183	166.67	16.33	1.6000
Coffee Cup	175	166.67	8.33	0.4163
Earrings	142	166.67	-24.67	3.6516
Total	500	500.00	0	5.6679

H_o is not rejected. There is not a preference for the gifts. **(LO1)**

19. H_o: There is no preference with respect to TV stations.
H_1: There is a preference with respect to TV stations.
$df = 3 - 1 = 2$ H_o is rejected if χ^2 is greater than 5.991

TV Station	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
CTV	53	50	3	9	0.18
Global	64	50	14	196	3.92
Citytv	33	50	-17	289	5.78
Total	150	150			9.88

H_o is rejected. There is a preference for TV stations. **(LO3)**

20. H_o: $p_1 = p_2 = p_3 = p_4$ H_1: The proportions are not equal.
df = 3 Reject H_o if $\chi^2 > 11.345$.

Entrance	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
Main	140	100	40	16.00
Broad	120	100	20	4.00
Cherry	90	100	-10	1.00
Walnut	50	100	-50	25.00
Total	400	400	0	46.00

H_o is rejected. The entrances are not equally likely. **(LO1)**

21. H_o: $p_{ne} = 0.21$, $p_m = 0.24$, $p_s = 0.35$ and $p_w = 0.20$
H_1: The distribution is not as given.
Reject H_o if $\chi^2 > 11.345$.

Area	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2/f_e$
New York State	68	84	-16	3.0476
Midwest	104	96	8	0.6667
South	155	140	15	1.6071
West	73	80	-7	0.6125
Total	400	400	0	5.9339

There is not enough evidence to reject H_o. The geographical distribution of her club members has not changed. **(LO3)**

22. H_o: Distribution with Poisson with $\mu = 2$. H_1: Distribution is not Poisson with $\mu = 2$
Decision rule: If $\chi^2 > 9.488$ reject H_o.

Applications	f_o	f_e	$(f_o - f_e)^2/f_e$
0	50	40.59	2.1815
1	77	81.21	0.2183
2	81	81.21	0.0005
3	48	54.12	0.6921
4	31	27.06	0.5737
5 or more	13	15.78	0.4898
Total	300		4.1558

Do not reject H_o. It is reasonable to conclude the distribution is Poisson with $\mu=2$. **(LO1)**

23. H_o: $p_0 = 0.4, p_1 = 0.3, p_2 = 0.2, p_3 = 0.1$
H_1: The proportions are not as given
Reject H_o if $\chi^2 > 7.815$

Applications	f_o	f_e	$(f_o - f_e)^2/f_e$
0	46	48	0.083
1	40	36	0.444
2	22	24	0.167
3	12	12	0.000
Total	120		0.694

Do not reject H_o. Evidence does not show a change in the accident distribution. **(LO1)**

24. H_o: Store size and advertising amount are not related
H_1: Store size and advertising amount are related
Reject H_o if $\chi^2 > 5.991$

$$\chi^2 = \frac{(40 - 56.97)^2}{56.97} + \ldots + \frac{(32 - 37.70)^2}{37.70} = 18.177$$

Reject H_o. Store size and advertising amount are related. **(LO3)**

25. H_o: Level of management and concern regarding the environment are not related
H_1: Level of management and concern regarding the environment are related
Reject H_o if $\chi^2 > 16.812$

$$\chi^2 = \frac{(15 - 14)^2}{14} + \ldots + \frac{(31 - 28)^2}{28} = 1.550$$

Do not reject H_o. Levels of management and environmental concerns are not related. **(LO3)**

26. H_o: Age and pressure are not related H_1: Age and pressure are related

Reject H_o if $\chi^2 > 16.812$

$$\chi^2 = \frac{(20-19.44)^2}{19.44} + \ldots + \frac{(43-40.32)^2}{40.32} = 2.191$$

Do not reject H_o. Age and pressure are not related. **(LO3)**

27. H_o: Whether a claim is filed and age is not related

H_1: Whether a claim is filed and age is related

Reject H_o if $\chi^2 > 7.815$

$$\chi^2 = \frac{(170-203.33)^2}{203.33} + \frac{(74-40.67)^2}{40.67} + \ldots + \frac{(24-35.67)^2}{35.67} = 53.639$$

Reject H_o. Age is related to whether a claim is filed. **(LO3)**

28. H_o: Pension plan preference and job class are not related

H_1: Pension plan preference and job class are related

Reject H_o if $\chi^2 > 13.277$

$$\chi^2 = \frac{(10-17.33)^2}{17.33} + \ldots + \frac{(22-33.94)^2}{33.94} = 84.04$$

Reject H_o. There is a relationship between pension plan preference and job class. **(LO3)**

29. H_o: $p_0 = 0.55, p_1 = 0.28 p_2 = 0.17$

H_1: The proportions are not as given

Reject H_o if $\chi^2 > 5.991$

Applications	f_o	f_e	$(f_o - f_e)^2/f_e$
0	220	247.5	3.056
1	158	126	8.127
2	72	76.5	0.265
Total	450		11.448

Reject H_o. Young adults differ from the general population. **(LO1)**

30. H_o: The distributions are the same for human resource and technical personnel.

H_1: The distributions are different.

$df = 3$ Reject H_o if $\chi^2 > 11.345$

$$\chi^2 = \frac{(35-59.1111)^2}{59.1111} + \ldots + \frac{(16-25.5556)^2}{25.5556} = 35.944$$

Reject H_o. The distributions are different. **(LO3)**

31. H_o: The proportions are the same H_1: The proportions are not the same

Reject H_o if $\chi^2 > 16.919$

f_o	f_e	$f_o - f_e$	$(f_o\text{-}f_e)$^2	$(f_o\text{-}f_e)^2/f_e$
44	28	16	256	9.143
32	28	4	16	0.571
23	28	-5	25	0.893
27	28	-1	1	0.036
23	28	-5	25	0.893

24	28	-4	16	0.571
31	28	3	9	0.321
27	28	-1	1	0.036
28	28	0	0	0.000
21	28	-7	49	1.750
				14.214

Do not reject H_o. The digits are evenly distributed. **(LO1)**

32. **(LO3)**

a.

Ho: There is no relationship between type and list price

H_1: There is a relationship between type and list price

List Price (thousands$)

Type	**less than 300**	**300 to under 500**	**500 to under 700**	**more than 700**	**Total**
Apartment	19	4	1	0	24
Townhouse	2	0	0	0	2
House	4	12	6	2	24
					50

Chi-square Contingency Table Test for Independence

	less than 300	300 to under 500	500 to under 700	more than 700	Total
Apartment	**19**	**4**	**1**	**0**	24
Townhouse	**2**	**0**	**0**	**0**	2
House	**4**	**12**	**6**	**2**	24
Total	25	16	7	2	50

21.43	chi-square
6	df
.0015	p-value

Reject H_o. There is an association between the variables type and list price.

b.

$H_{o:}$ The number of bedrooms and list price are not related.

H_1: The number of bedrooms and list price are related.

Type	**less than 300**	**300 to under 500**	**500 to under 700**	**700 +**	**Total**
1-3 Bedrooms	24	8	3	0	35
4+ bedrooms	1	8	4	2	15
					50

Chi-square Contingency Table Test for Independence

	less than 300	300 to under 500	500 to under 700	700 +	Total
1-3 Bedrooms	**24**	**8**	**3**	**0**	35
4+ bedrooms	**1**	**8**	**4**	**2**	15
Total	25	16	7	2	50

18.22	chi-square
3	df
.0004	p-value

Reject H_o. There is an association between the number of bedrooms and list price.

33. **(LO3)**

a.

Ho: There is no relationship between type and list price

H_1: There is a relationship between type and list price

List Price (thousands$)

Type	**less than 300**	**300 to under 500**	**500 to under 700**	**more than 700**	**Total**
Apartment/THouse	21	4	1	0	26
House	4	12	6	2	24
	25	16	7	2	50

Chi-square Contingency Table Test for Independence

	less than 300	300 to under 500	500 to under 700	more than 700	Total
Apartment/THouse	**21**	**4**	**1**	**0**	26
House	**4**	**12**	**6**	**2**	24
Total	25	16	7	2	50

21.09	chi-square
3	df
.0001	p-value

Reject H_o. There is an association between the type and list price.

b.

$H_{o:}$ The number of bedrooms and list price are not related.

H_1: The number of bedrooms and list price are related.

List Price (thousands$)

Type	**less than 400**	**400 to under 600**	**600 to under 800**	**800 +**	**Total**
1-3 Bedrooms	39	29	8	4	80
4+ bedrooms	1	2	3	12	18
					98

	less than 400	400 to under 600	600 to under 800	more than 800	Total
1-3 Bedrooms	**39**	**29**	**8**	**4**	80
4+ bedrooms	**1**	**2**	**3**	**12**	18
Total	40	31	11	16	98

44.46	chi-square
3	df
1.21E-09	p-value

Reject H_o. There is an association between the number of bedrooms and list price.

34. **(LO3)**

a.

H_0: Type and price are not related.

H_1: Type and price are related.

Type		**$25 000 > $30 000**	**$30 000 > $35 000**	**over $35 000**	**Total**
0	16	8	4	2	30
1	40	10	0	0	50
Total	56	18	4	2	80

12.28	chi-square
3	df
.0065	p-value

Reject H_o. Conclude that code and price are related.

b.

H_0: Age and price are not related.

H_1: Age and price are related.

Age	Under $25 000	$25 000 > $35 000	over $35 000	Total
20 to under 30	6	0	0	6
30 to under 40	19	5	0	24
40 to under 50	22	11	0	33
over 50	9	6	2	17
Total	56	22	2	80

Chi-square Contingency Table Test for Independence

	Under $25 000	$25 000 > $35 000	over $35 000	Total
20 to under 30	**6**	**0**	**0**	**6**
30 to under 40	**19**	**5**	**0**	**24**
40 to under 50	**22**	**11**	**0**	**33**
over 50	**9**	**6**	**2**	**17**
Total	56	22	2	80

12.05	chi-square
9	df
.2104	p-value

Do not reject H_o. Conclude that age and price are not related.

These are not the results we would expect. As you can see above, there are several cells with frequencies less than 5. It would make sense to combine some of the categories and do the test again.

CHAPTER 15

INDEX NUMBERS

1.

average 2007-2009: $166,424

		Index
Jan-11	$221,933	133.4
Jan-10	$206,454	124.1
Jan-09	$183,873	110.5
Jan-08	$169,668	101.9
Jan-07	$145,731	87.6
Jan-06	$140,748	84.6
Jan-05	$119,728	71.9

2. a.

Year	All benefits	Index	Found by
2004	**296.87**	99.0	296.87/300.01*100
2005	**300.01**	100.0	300.01/300.01*100
2006	**308.73**	102.9	308.73/300.01*100
2007	**317.65**	105.9	317.65/300.01*100
2008	**330.88**	110.3	330.88/300.01*100
2009	**337.56**	112.5	337.56/300.01*100
2010	**341.22**	113.7	341.22/300.01*100

b.

average 2005-2007 = 308.80

Year	All benefits	Index	Found by
2004	**296.87**	96.1	296.87/308.80*100
2005	**300.01**	97.2	300.01/308.80*100
2006	**308.73**	100.0	308.73/308.80*100
2007	**317.65**	102.9	317.65/308.80*100
2008	**330.88**	107.2	330.88/308.80*100
2009	**337.56**	109.3	337.56/308.80*100
2010	**341.22**	110.5	341.22/308.80*100

3. a.

	List Price	Index	Found by
Jan-11	$343,675	150.0	343 675/229 184*100
Jan-10	$328,728	143.4	328 728/229 184*100
Jan-09	$274,711	119.9	274 711/229 184*100
Jan-08	$309,448	135.0	309 448/229 184*100
Jan-07	$282,420	123.2	282 420/229 184*100
Jan-06	$258,274	112.7	258 274/229 184*100
Jan-05	$229,184	100.0	229 184/229 184*100

Hose prices have increased by 50% over the time period.

b.

Average = $288,860

	List Price	Index
Jan-11	$343,675	119.0
Jan-10	$328,728	113.8
Jan-09	$274,711	95.1
Jan-08	$309,448	107.1
Jan-07	$282,420	97.8
Jan-06	$258,274	89.4
Jan-05	$229,184	79.3

4. 275.9 found by (5.49/1.99)(100) There was a 175.9 percent increase.

5. a. $P_t = \frac{2.69}{2.49}(100) = 108.03$ $P_s = \frac{3.59}{3.29}(100) = 109.12$

$P_c = \frac{2.79}{1.79}(100) = 155.9$ $P_a = \frac{3.79}{2.29}(100) = 165.5$

b. $P = \frac{12.86}{9.86}(100) = 130.4$

c. $P = \frac{(2.69)(6) + (3.59)(4) + (2.79)(2) + (3.79)(3)}{(2.49)(6) + (3.29)(4) + (1.79)(2) + (2.29)(3)}(100) = 123.1$

d. $P = \frac{(2.69)(6) + (3.59)(5) + (2.79)(3) + (3.79)(4)}{(2.49)(6) + (3.29)(5) + (1.79)(3) + (2.29)(4)}(100) = 125.5$

e. $I = \sqrt{(123.1)(125.5)} = 124.3$

6. a. $P_b = \frac{0.49}{0.23}(100) = 213.0$ $P_G = \frac{0.27}{0.29}(100) = 93.10$ $P_A = \frac{0.35}{0.35}(100) = 100$

$P_b = \frac{1.99}{1.02}(100) = 195.1$ $P_b = \frac{2.99}{0.89}(100) = 336.0$

b. $P = \frac{6.09}{2.78}(100) = 219.1$

c. $P = \frac{(.49)(100) + (.27)(50) + (.35)(85) + (1.99)(8) + (2.99)(6)}{(.23)(100) + (.29)(50) + (.35)(85) + (1.02)(8) + (.89)(6)}(100) = 156.2$

d. $P = \frac{(.49)(120) + (.27)(55) + (.35)(85) + (1.99)(10) + (2.99)(8)}{(.23)(120) + (.29)(55) + (.35)(85) + (1.02)(10) + (.89)(8)}(100) = 162.5$

e. $I = \sqrt{(156.2)(162.5)} = 159.3$

7. a. $P_W = \frac{0.10}{0.07}(100) = 142.9$ $P_b = \frac{0.10}{0.04}(100) = 250.0$

$P_b = \frac{0.18}{0.15}(100) = 120.0$ $P_H = \frac{0.10}{0.08}(100) = 125.0$

b. $P = \frac{0.48}{0.34}(100) = 141.2$

c. $P = \frac{(.10)(17000) + (.10)(125000)(.18)(40000) + (.10)(62000)}{(.07)(17000) + (.04)(125000) + (.15)(40000) + (.08)(62000)}(100)$ = 160.9

d. $P = \frac{(.10)(20000) + (.10)(130000)(.18)(42000) + (.10)(65000)}{(.07)(20000) + (.04)(130000) + (.15)(42000) + (.08)(65000)}(100)$ = 160.6

e. $I = \sqrt{(160.9)(160.6)}$ = 160.7

8. a. $P_P \frac{1.10}{0.90}(100) = 122.2$ $P_P = \frac{0.80}{0.65}(100) = 123.1$

$P_E = \frac{0.55}{0.45}(100) = 122.2$ $P_P = \frac{1.09}{0.89}(100) = 122.5$

$P_{PP} = \frac{4.99}{5.99}(100) = 83.3$ $P_{PC} = \frac{19.99}{15.99}(100) = 125.0$

b. $P = \frac{28.52}{24.87}(100) = 114.7$

c.

$$P = \frac{(1.10)(50) + (.80)(50) + (.55)(250) + (1.09)(500) + (4.99)(300) + (19.99)(150)}{(.90)(50) + (.65)(50) + (.45)(250) + (.89)(500) + (5.99)(300) + (15.99)(150)}(100)$$

= 109.2

d.

$$P = \frac{(1.10)(55) + (.80)(60) + (.55)(275) + (1.09)(750) + (4.99)(450) + (19.99)(200)}{(.90)(55) + (.65)(60) + (.45)(275) + (.89)(750) + (5.99)(450) + (15.99)(200)}(100)$$

= 108.1

e. $I = \sqrt{(109.2)(108.1)}$ = 108.6

9. $V = \frac{1.87(214) + 2.05(489) + 1.48(203) + 3.29(106)}{1.52(200) + 2.10(565) + 1.48(291) + 3.05(87)}(100) = 93.8$

10. $V = \frac{28.80(4259) + 3.08(62,949) + 0.48(22,370)}{23.60(1760) + 2.96(86,450) + 0.40(9460)}(100) = 108.7$

11. The increase in the CPI is 16.5%, so 600(1.165) = $699

12. % increase(boots) = (150 – 125)/125(100) = 20%. Therefore, the boots increased more than the CPI.

13. *X* = ($47 500)/1.165 = $40 772.53 Salary increased $40 772.53 - $39 000 = $1 772.53.

14. Plumbers = $\frac{159.4}{133.8}(100) = 119.1$ Electrician = $\frac{158.7}{126.0}(100) = 126.0$

Plumbers wages have increased 19.1 percent, where as electrician has increased 26.0 percent.

15.

Year	*Mercury*	*Mercury*	*Industry Index*
1995	$26 650	100.0	100.0
2000	$31 972	120.0	122.5
2005	$36 382	136.5	136.9
2008	$37 269	139.8	144.9
2011	$39 500	148.2	146.0

The Mercury plant workers received increases slightly less than the industry average with the exception of 2011.

16.

Year	*Wages*	*Index (1995 = 100.0)*	*Wage(1990 $)*
2006	$175 000	148.3	118 004
2007	175 000	140.6	124 467
2008	150 000	120.9	124 069
2009	120 000	110.2	108 893
2010	120 000	105.3	113 960
2011	130 000	105.0	123 810

Sam's real wage in 1995 dollars went up and down as did the industry, but appears to be rebounding now.

17.

Region	**Jan-11**	**Index**
National Average	**$343,675**	100.0
Vancouver	$762,562	221.9
Calgary	$394,655	114.8
Edmonton	$315,483	91.8
Toronto	$427,159	124.3
Ottawa	$329,640	95.9
Saint John	$171,788	50.0
Halifax	$252,141	73.4

The value of home listings is highest in Vancouver, at 121.9 percent higher than the national average. Saint John is the lowest at 50.0 percent less than the national average.

18.

Jan-07	**Index**
$282,420	100.0
$530,695	187.9
$375,746	133.0
$303,820	107.6
$353,724	125.2
$260,898	92.4
$143,971	51.0
$197,246	69.8

The value of home listings is highest in Vancouver, at 87.9 percent higher than the national average. Saint John is the lowest at 49.0 percent less than the national average.

19.

Jan-04	Index
$212,757	100.0
$354,135	166.5
$219,774	103.3
$170,960	80.4
$310,190	145.8
$229,313	107.8
$123,552	58.1
$169,320	79.6

The value of home listings is highest in Vancouver, at 66.5 percent higher than the national average. Saint John is the lowest at 41.9 percent less than the national average.

20.

	Jan-11	Index
Average	**$490,900**	100.0
Vancouver	$762,562	155.3
Calgary	$394,655	80.4
Edmonton	$315,483	64.3
Toronto	$427,159	87.0
Ottawa	$329,640	67.2
Saint John	$171,788	35.0
Halifax	$252,141	51.4

The value of home listings is highest in Vancouver, and is the only city which is higher than the average of Vancouver, Calgary and Edmonton.

21.

	Jan-11	Index
Average	**$378,400**	100.0
Vancouver	$762,562	201.5
Calgary	$394,655	104.3
Edmonton	$315,483	83.4
Toronto	$427,159	112.9
Ottawa	$329,640	87.1
Saint John	$171,788	45.4
Halifax	$252,141	66.6

The value of home listings is highest in Vancouver at 101.5 percent higher than the average of Toronto and Ottawa.

22.

	Jan-11	Index
Average	**$211,965**	100.0
Vancouver	$762,562	359.8
Calgary	$394,655	186.2
Edmonton	$315,483	148.8
Toronto	$427,159	201.5
Ottawa	$329,640	155.5

Saint John	$171,788	81.0
Halifax	$252,141	119.0

The value of home listings is highest in Vancouver at 259.8 percent higher than the average of Saint John and Halifax.

23.

Date	Closing Price	Index
Mar-11	$55.65	165.6
Mar-10	75.25	224.0
Mar-09	54.49	162.2
Mar-08	115.49	343.7
Mar-07	157.50	468.8
Mar-06	98.89	294.3
Mar-05	92.71	275.9
Mar-04	122.4	364.3
Mar-03	19.08	56.8
Mar-02	44.34	132.0
Mar-01	33.60	100.0

The index for Mar 07 is 468.8 which is 368.8 percent higher than the base price.
Note that the stock split on Aug 15, 2007 was 3:1.
The index for Mar 04 is 364.3 which is 264.3 percent higher than the base price.
Note that the stock split on May 25, 2004 was 2:1.

24.

Date	Adjusted Closing Price	Index
Mar-11	$55.65	993.8
Mar-10	75.25	1343.8
Mar-09	54.49	973.0
Mar-08	115.49	2062.3
Mar-07	52.50	937.5
Mar-06	32.96	588.6
Mar-05	30.90	551.8
Mar-04	20.40	364.3
Mar-03	3.18	56.8
Mar-02	7.39	132.0
Mar-01	5.60	100.0

The index for Mar 08 is the highest and Mar 03, the lowest.

25.

Date	Closing Price	Index
Mar-11	$55.65	53.2
Mar-10	75.25	71.9
Mar-09	54.49	52.1
Mar-08	115.49	110.3
Mar-07	157.50	150.5

Mar-06	98.89	94.5
Mar-05	92.71	88.6
Mar-04	122.4	116.9
Mar-03	19.08	18.2
Mar-02	44.34	42.4
Mar-01	33.60	32.1

average = 104.7

The index for Mar 07 is 150.5 which is 50.5 percent higher than the base period.
Note that the stock split on Aug 15, 2007 was 3:1.

26.

Date	Adjusted Closing Price	Index
Mar-11	$55.65	254.6
Mar-10	75.25	344.2
Mar-09	54.49	249.3
Mar-08	115.49	528.3
Mar-07	52.50	240.2
Mar-06	32.96	150.8
Mar-05	30.90	141.4
Mar-04	20.40	93.3
Mar-03	3.18	14.5
Mar-02	7.39	33.8
Mar-01	5.60	25.6

average = 21.86

The index for Mar 10 is 344.2 which is 244.2 percent higher than the base period.

27. $P_M = \frac{2.39}{.81}(100) = 295.1$ $P_S = \frac{1.49}{.84}(100) = 177.4$

$P_M = \frac{3.79}{1.44}(100) = 263.2$ $P_P = \frac{3.99}{2.91}(100) = 137.1$

28. $P_P = \frac{11.66}{6.00}(100) = 194.3$

29. $P = \frac{(2.39)(18) + (1.49)(5) + (3.79)(70) + (3.99)(27)}{(.81)(18) + (.84)(5) + (1.44)(70) + (2.91)(27)}(100) = 213.7$

30. $P = \frac{(2.39)(27) + (1.49)(9) + (3.79)(65) + (3.99)(33)}{(.81)(27) + (.84)(9) + (1.44)(65) + (2.91)(33)}(100) = 208.1$

31. $I = \sqrt{(213.7)(208.1)} = 210.9$

32. $P = \frac{(2.39)(27) + (1.49)(9) + (3.79)(65) + (3.99)(33)}{(.81)(18) + (.84)() + (1.44)(70) + (2.91)(65)}(100) = 230.1$

33. $P_R = \frac{0.60}{0.50}(100) = 120$ $P_S = \frac{0.90}{1.20}(100) = 75.0$ $P_W = \frac{1.00}{0.85}(100) = 117.65$

34. $P = \frac{2.50}{2.55}(100) = 98.04$

35. $P = \frac{(.60)(320) + (.90)(110) + (1.00)(230)}{(.50)(320) + (1.20)(110) + (.85)(230)}(100) = 106.9$

36. $P = \frac{(.60)(340) + (.90)(130) + (1.00)(250)}{(.50)(340) + (1.20)(130) + (.85)(250)}(100) = 106.0$

37. $I = \sqrt{(106.9)(106.0)} = 106.5$

38. $P = \frac{(.60)(340) + (.90)(130) + (1.00)(250)}{(.50)(320) + (1.20)(110) + (.85)(230)}(100) = 117.1$

39. $P_C = \frac{.90}{.60}(100) = 150.0$ $P_C = \frac{.69}{.49}(100) = 140.8$

$P_P = \frac{2.99}{1.99}(100) = 150.3$ $P_P = \frac{1.29}{.89}(100) = 144.9$

40. $P_P = \frac{5.87}{3.97}(100)$ = 147.9

41. $P = \frac{(.90)(2000)+(.69)(200)+(2.99)(400)+(1.29)(100)}{(.60)(2000)+(.49)(200)+(1.99)(400)+(.89)(100)}(100)$ = 149.5

42. $P = \frac{(.90)(1500)+(.69)(200)+(2.99)(500)+(1.29)(200)}{(.60)(1500)+(.49)(200)+(1.99)(500)+(.89)(200)}(100)$ = 149.3

43. $I = \sqrt{(149.5)(149.3)}$ = 149.4

44. $P = \frac{(.90)(1500)+(.69)(200)+(2.99)(500)+(1.29)(200)}{(.60)(2000)+(.49)(200)+(1.99)(400)+(.89)(100)}(100)$ = 148.5

45. $P_{PC} = \frac{5.99}{4.99}(100)$ = 120.0 $P_{PL} = \frac{.99}{.89}(100)$ = 111.2

$P_{PP} = \frac{1.19}{.99}(100)$ = 120.2 $P_{PC} = \frac{1.79}{1.49}(100)$ = 120.1

46. $P = \frac{9.96}{8.36}(100)$ = 119.1

47. $P = \frac{(5.99)(400)+(.99)(1000)+(1.19)(850)+(1.79)(350)}{(4.99)(400)+(.89)(1000)+(.99)(850)+(1.49)(350)}(100)$ = 118.2

48. $P = \frac{(5.99)(500)+(.99)(1200)+(1.19)(1000)+(1.79)(350)}{(4.99)(500)+(.89)(1200)+(.99)(1000)+(1.49)(350)}(100)$ = 118.2

49. $I = \sqrt{(118.2)(118.2)}$ = 118.2

50. $P = \frac{(5.99)(500)+(.99)(1200)+(1.19)(1000)+(1.79)(350)}{(4.99)(400)+(.89)(1000)+(.99)(850)+(1.49)(350)}(100)$ = 141.2

51. $I = \frac{1971.0}{1159.0}(0.20) + \frac{91}{87}(0.10) + \frac{114.7}{110.6}(0.40) + \frac{1501}{1214}(0.30) = 1.2305*100 = 123.05$

The economy is up 23.05 percent from the base year to the current year.

52. a. 2002 = 100.0 (base period)

Year	*CPI*	*Photo Supplies*	*Index*	*Photo Services*	*Index*
2004	104.7	175		65	
2005	107.0	205	117.1	70	107.7
2006	109.1	300	171.4	72	110.8
2007	111.5	310	177.1	86	132.3
2008	114.1	315	180.0	92	141.5

2009	114.4	318	181.7	92.5	142.3
2010	116.5	320	182.9	93	143.1

b. The increase in the CPI is 11.3% from 2004 to 2010. The increase in the sales of photographic supplies is 82.9%, and the increase in photographic services is 43.1% for the same period.

53. February: $I = \frac{6.8}{8.0}(0.40) + \frac{23}{20}(0.35) + \frac{303}{300}(100) = .9950 * 100 = 99.5$

March: $I = \frac{6.4}{8.0}(0.40) + \frac{21}{20}(0.35) + \frac{297}{300}(100) = .9350 * 100 = 93.5$

54. a. $X = \frac{1}{1.115} = 0.89686$

b. $2511.21, found by (0.89686)($2800)

c. $X = \frac{1}{1.165} = 0.85837$

d. $2489.28, found by (0.85837)($2900)

55. a.

Year	S&P/TSX	Index	NASDAQ	Index
2001	7833.24	100.0	1950.4	100.0
2002	6772.66	86.5	1335.51	68.5
2003	8293.70	105.9	2003.37	102.7
2004	9246.65	118.0	2175.44	111.5
2005	11272.36	143.9	2205.32	113.1
2006	12908.39	164.8	2415.29	123.8
2007	13778.58	175.9	2652.28	136.0
2008	8987.70	114.7	1577.03	80.9
2009	11746.11	150.0	2269.15	116.3
2010	13443.22	171.6	2652.87	136.0

S&P/TSX increased by 71.6% over the time period.
NASDAQ increased by 36% over the time period.

b.

Year	S&P/TSX	Index	S&P 500	Index
2001	7833.24	100.0	1148.08	100.0
2002	6772.66	86.5	879.82	76.6
2003	8293.70	105.9	1111.92	96.9
2004	9246.65	118.0	1211.92	105.6
2005	11272.36	143.9	1248.29	108.7

2006	12908.39	164.8	1418.3	123.5
2007	13778.58	175.9	1468.36	127.9
2008	8987.70	114.7	903.25	78.7
2009	11746.11	150.0	1115.1	97.1
2010	13443.22	171.6	1257.64	109.5

S&P/TSX increased by 71.6% over the time period.
S&P 500 increased by 9.5% over the time period.

c.

Year	**S&P/TSX Venture**	**Index**	**NASDAQ**	**Index**
2001	1036.59	100.0	1950.4	100.0
2002	1074.08	103.6	1335.51	68.5
2003	1751.28	168.9	2003.37	102.7
2004	1825.47	176.1	2175.44	111.5
2005	2236.55	215.8	2205.32	113.1
2006	2987.08	288.2	2415.29	123.8
2007	2839.66	273.9	2652.28	136.0
2008	797.02	76.9	1577.03	80.9
2009	1520.72	146.7	2269.15	116.3
2010	2287.85	220.7	2652.87	136.0

S&P/TSX Venture increased by 120.7% over the time period.
NASDAQ increased by 36% over the time period.

d.

Year	**S&P/TSX Venture**	**Index**	**S&P 500**	**Index**
2001	1036.59	100.0	1148.08	100.0
2002	1074.08	103.6	879.82	76.6
2003	1751.28	168.9	1111.92	96.9
2004	1825.47	176.1	1211.92	105.6
2005	2236.55	215.8	1248.29	108.7
2006	2987.08	288.2	1418.3	123.5
2007	2839.66	273.9	1468.36	127.9
2008	797.02	76.9	903.25	78.7
2009	1520.72	146.7	1115.1	97.1
2010	2287.85	220.7	1257.64	109.5

S&P/TSX Venture increased by 120.7% over the time period.
S&P 500 increased by 9.5% over the time period.

CHAPTER 16

TIME SERIES AND FORECASTING

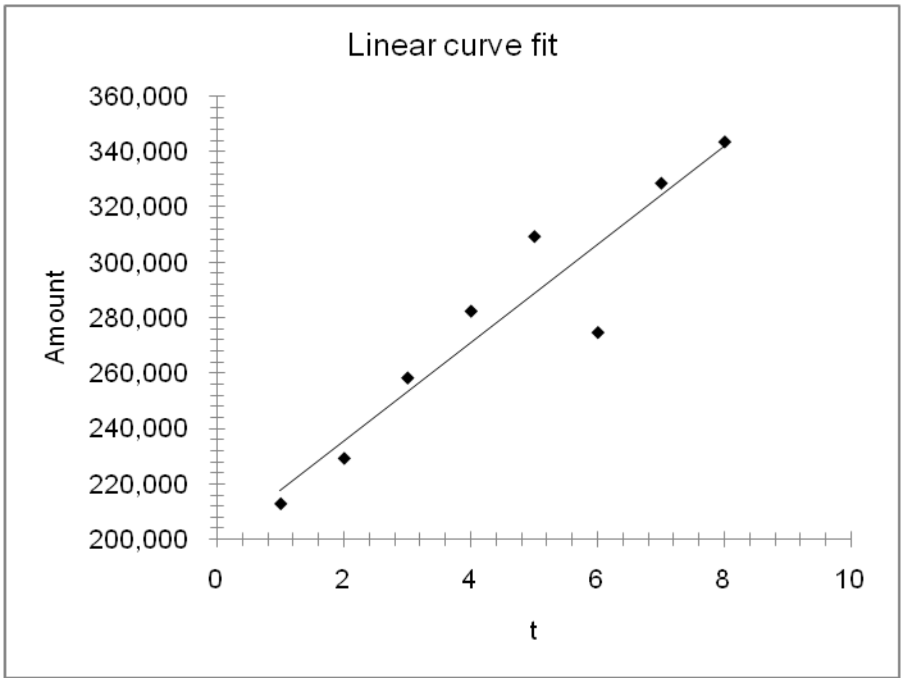
1. $b = \dfrac{2469 - (721)(15)/5}{55 - (15)^2/5} = \dfrac{306}{10} = 30.6$ $\quad a = \dfrac{72}{5} - 30.6\left(\dfrac{15}{5}\right) = 52.4$

for 2013, $t = 7$; $\quad Y' = 52.4 + 30.6t = 52.4 + 30.6(7) = 266.6$

2. $b = \dfrac{9(5024.6) - (90.1)(45)}{9(285) - (45)^2} = 2.0683 \quad a = \dfrac{980.1}{9} - 2.0683\left(\dfrac{45}{9}\right) = 98.5585$

$Y' = 98.5585 + 2.0683t$ $\qquad Y' = 98.5585 + 2.0683(12) = 123.3781$

3. a.



b. $b = \dfrac{8(10\,821\,629) - (2\,239\,197)(36)}{8(204) - (204)^2} = 17\,743.6890$

$a = \dfrac{2\,239\,197}{8} - (17\,743.6890)\left(\dfrac{36}{8}\right) = 200\,052.2143$

$Y' = 200\,052.2143 + 17\,743.690t$

c. Every year, expect the amount to increase by about $17 744.

d. 2015 ($t = 12$) = $412 979

e. 2009 ($t = 6$) = $306 515; the predicted value is an estimate while the given value is the actual value.

4. The regression equation is: $\hat{Y}$ = 4425 + 4947*t*.
For 2010, t = 18 and $\hat{Y}$ = 4425 + 4947(18) = 93 471
For 2011, t = 19 and $\hat{Y}$ = 4425 + 4947(19) = 98 418

5. Y' = 1.7143 + 0.75*t*; for t = 9: Y' = 1.7143+ 0.75(9) = 8.5

6. The regression equation is: $\hat{Y}$ = 15.77 + 1.73*t*

For 2012, $t = 9$ and $\hat{Y} = 15.77 + 1.73\,(8) = 29.61$

7. a. $b = \frac{5.274318 - (1.390087)(15)/5}{55 - (15)^2/5} = \frac{1.104057}{10} = 0.1104057$

$$a = \frac{1.390087}{5} - 0.1104057\left(\frac{15}{5}\right) = -0.0531997$$

b. 28.95%, found by 1.28945 – 1.0

c. $Y' = -0.0531997 + 0.1104057t$ for 2014, $t = 8$

$Y' = -0.0531997 + 0.1104057(8) = 0.8300459$ Antilog of 0.8300459 = 6.76

8. a. $Y' = 1.92333 + 0.0415302t$, where the first year is coded 1.

b. 10.03%, found by taking the antilog of 0.0415302

c. $Y' = 2.3801622$, the antilog is 239.97

9.

Quarter	*Average SI Component*	*Seasonal Index*
1	0.6859	0.6911
2	1.6557	1.6682
3	1.1616	1.1704
4	0.4732	0.4768

Note: Excel/MegaStat answers may be slightly different due to rounding.

10.

Quarter	*Average SI Component*	*Seasonal Index*
1	0.9122	0.9077
2	0.7647	0.7609
3	1.1318	1.1261
4	1.2159	1.2098

Note: Excel/MegaStat answers may be slightly different due to rounding.

11.

t	*estimated pairs (millions)*	*Seasonal index*	*Quarterly forecast (millions)*
21	40.05	110.0	44.055
22	41.80	120.0	50.160
23	43.55	80.0	34.840
24	45.30	90.0	40.770

12. Sales for each quarter are 500, found by 2000/4. The estimated sales for the second quarter are 725, found by 500(1.45).

13. $Y' = 5.1658 + 0.37805t$. The following are the sales estimates.

Estimate	*Index*	*Seasonally adjusted*
10.080	0.6911	6.966
10.458	1.6682	17.446
10.837	1.1704	12.684
11.215	0.4768	5.343

14. $Y' = 5.48 - 0.0112t$. The following are the quarterly estimates.

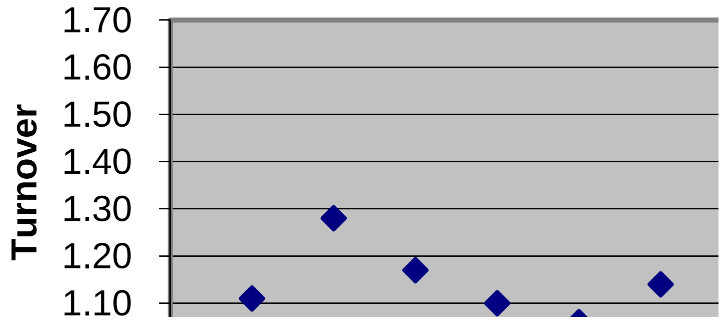
Fitted	*Index*	*Forecast*

5.293	0.9077	4.8045
5.282	0.7609	4.0191
5.270	1.1261	5.9345
5.259	1.2089	6.3575

15. a. *Y*¢= 18,000 - 400*t*, assuming the line starts at 18,000 in 1991 and goes down to 10,000 in 2011.
    b. 400
    c. 8000, found by 18,000 – 400(25)

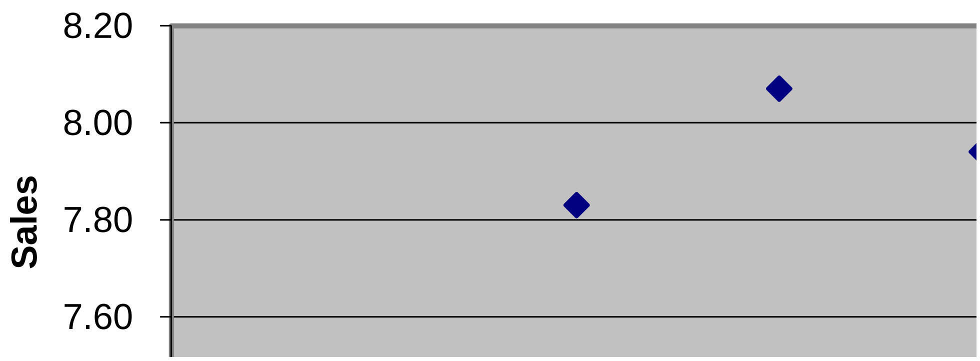
16. a. *Y*¢= 4000 + 933*t* assuming a straight line goes from 4000 to 18,000.
    b. $933

17. a.



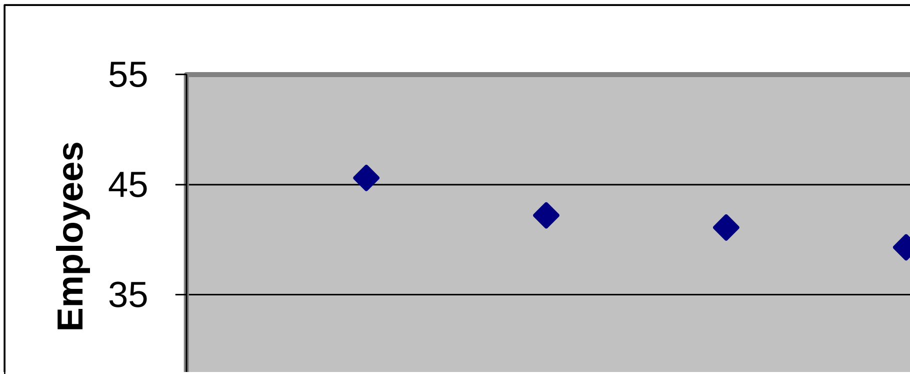
    b. *Y*¢= 1.00455 + 0.04409*t*, using *t* = 1 for 2001
    c. for 2004, *Y*¢= 1.18091, and for 2009 *Y*¢= 1.40136
    d. for 2016, *Y*¢= 1.70
    e. Each asset turned over 0.044 times

18. a.



    b. *Y*¢= 7.6340 + 0.05457*t*, using *t* = 1 for 2006
    c. For 2008, *Y*¢= 7.798 and for 2010, *Y*¢= 7.907
    d. *Y*¢= 8.125
    e. Sales increased 0.05457 billion dollars per year

19. a.



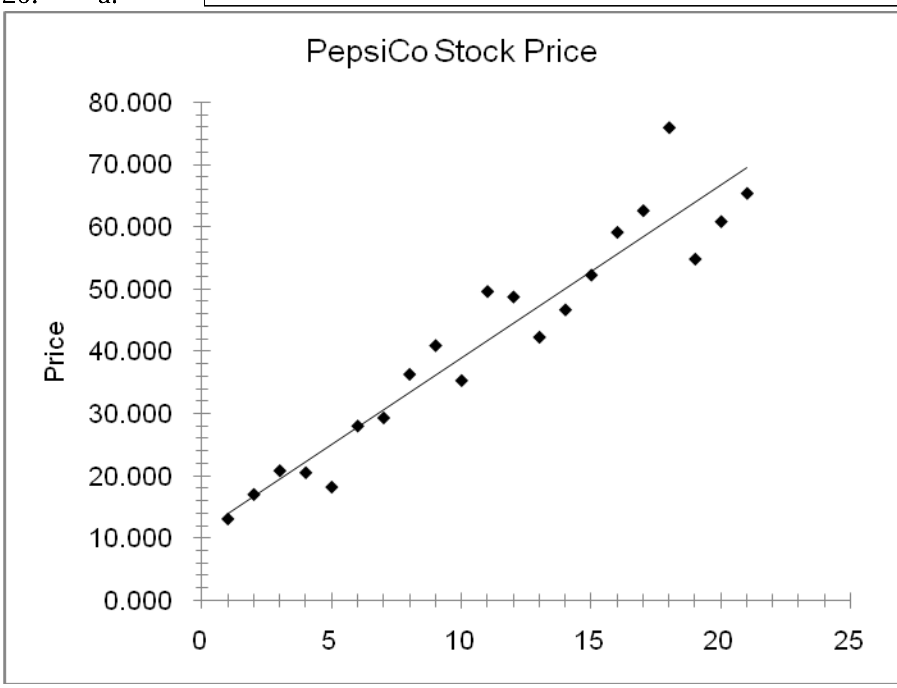
b. $Y' = 49.140 - 2.9829t$

c. for 200, $Y' = 40.1914$ and for 2010, $Y' = 34.2257$

d. for 2014 $Y' = 22.2943$

e. The number of employees decreases at a rate of 2982.9 per year.

20. a.



b. $Y' = 11.1720 + 2.7789t$

c. 11.1720 + 2.7789 (6) = 27.85; 11.1720 + 2.7789 (11) = 41.74

d. 11.1720 + 2.7789 (22) = 72.31. Yes because $R^2 = 0.908$

e. 2.779 per year

21. a. Log $Y' = 0.790231 + 0.113669t$

b. Log $Y' = 1.244907$, antilog is 17.575

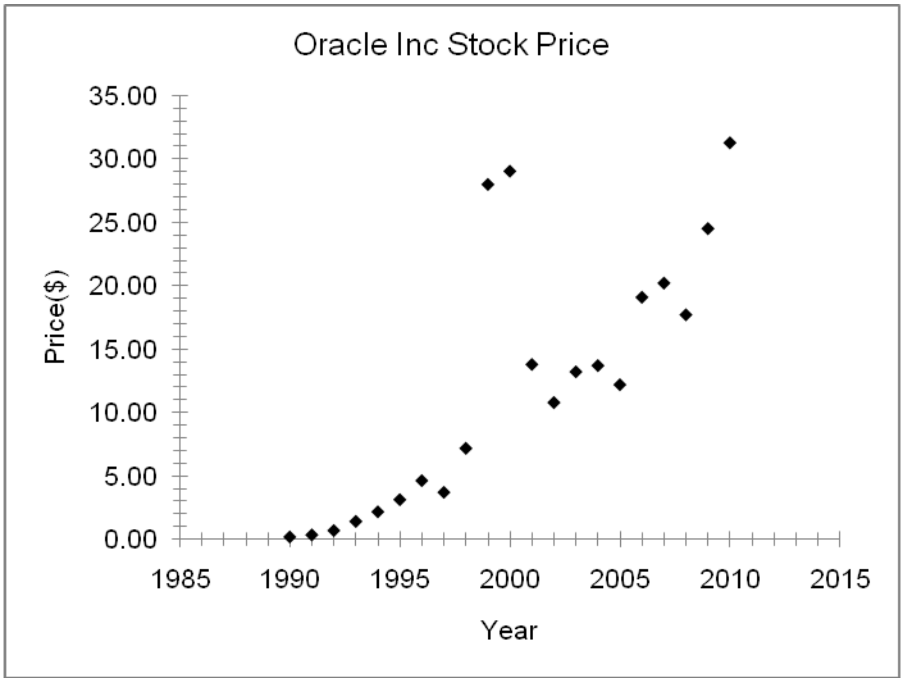
Log $Y' = 1.813252$, antilog is 65.05

c. 29.92, which is the antilog of 0.113669 minus 1

d. Log $Y' = 2.154258$, antilog is 142.65

22. a. Log Y¢= 1.92767 + 0.0272809*t*
b. Log Y¢= 2.3096026, antilog is 204.0
c. 6.48 percent, which is the antilog of 0.0272809 minus 1

23. a.



The following results are using MegaStat.

b. The equations are Y¢= –2.0533 + 1.3004*t* and/or log Y¢ = –0.1955 + 0.0910*t*.
The equation using the logarithm appears better because R^2 is larger.

c. log Y¢ = –0.1955 + 0.0910(4) = 0.16856; antilog = $1.47.
log Y¢ = –0.1955 + 0.0910 (9) = 0.62366; antilog = $4.20.

d. log Y¢ = –0.1955 + 0.0910 (24) = 1.98896, antilog is $97.49. The linear estimate is Y¢= –0.7528 + 1.3004 (24) = $30.46.

e. The annual rate of increase is 23.3 percent, found by the antilog of 0.0910 minus 1.

24. a. July 44.2, August 72.3, September 197.5

b.

Month	*Total*	*Mean*	*Seasonal*
Jan.	345.3	86.325	86.5
Feb.	424.3	106.075	106.3
March	697.8	174.450	174.8
April	483.9	120.975	121.2
May	239.2	59.800	59.9
June	190.3	47.575	47.7
July	180.6	45.150	45.2
August	295.6	73.900	74.0
Sept.	798.5	199.625	200.0
Oct.	351.9	87.975	88.1
Nov.	424.6	106.15	106.4
Dec.	358.6	89.65	89.8

1197.65

Correction = 1200/1197.65 = 1.001962

c. Sales for September and March are considerably above average and below average for May, June, and July.

25. a. July 87.5, August 92.9, September 99.3, October 109.1

b.

Month	*Total*	*Mean*	*Seasonal*
July	348.9	87.225	86.777
Aug.	368.1	92.025	91.552
Sept.	395.0	98.750	98.242
Oct.	420.4	105.100	104.560
Nov.	496.2	124.050	123.412
Dec.	572.3	143.075	142.340
Jan.	333.5	83.375	82.946
Feb.	297.5	74.375	73.993
March	347.3	86.825	86.379
April	481.3	120.325	119.707
May	396.2	99.050	98.541
June	368.1	92.025	91.552
		1206.200	

Correction = 1200/1206.2 = 0.99486

c. April, November, and December are periods of high sales, while February is low.

26. a.

Seasonal Index by Quarter

Quarter	*Average SI Component*	*Seasonal Index*
1	0.7577	0.7558
2	0.9949	0.9924
3	1.4095	1.4060
4	0.8526	0.8505

Note: Excel answers may be different due to rounding.

b. The third quarter is more than 40% above a typical quarter. The production activity is below average in the first and fourth quarters.

c. *Y*¢= 10.0989 + 0.14213*t*

d. The projections for 2012 are as follows:

Period	*Production*	*Index*	*Forecast*
21	13.084	0.7558	9.889
22	13.226	0.9924	13.125
23	13.368	1.4060	18.795
24	13.510	0.8505	11.490

27. a. *Seasonal Index by Quarter*

Quarter	*Average SI Component*	*Seasonal Index*
1	0.5014	0.5027
2	1.0909	1.0936
3	1.7709	1.7753
4	0.6354	0.6370

b. The production is the largest in the third quarter. It is 77.5% above the average quarter. The second quarter is also above average. The first and fourth quarters are well below average, with the first quarter at about 50% of a typical quarter.

28. a. *Seasonal Index by Quarter*

Quarter	*Average SI Component*	*Seasonal Index*
1	1.1909	1.1939
2	1.1215	1.1243
3	0.4350	0.4361
4	1.2516	1.2548

Note: Excel answers may be different due to rounding.

b. *Y*¢= 163.208 + 4.1253*t*

Period	*Sales*	*Index*	*Forecast*
29	282.8417	1.1939	337.6847
30	286.9670	1.1243	322.6370
31	291.0923	0.4361	126.9454
32	295.2176	1.2548	370.4390

29 a. *Seasonal Index by Quarter*

Quarter	*Average SI Component*	*Seasonal Index*
1	0.5549	0.5577
2	0.8254	0.8296
3	1.5102	1.5178
4	1.0973	1.1029

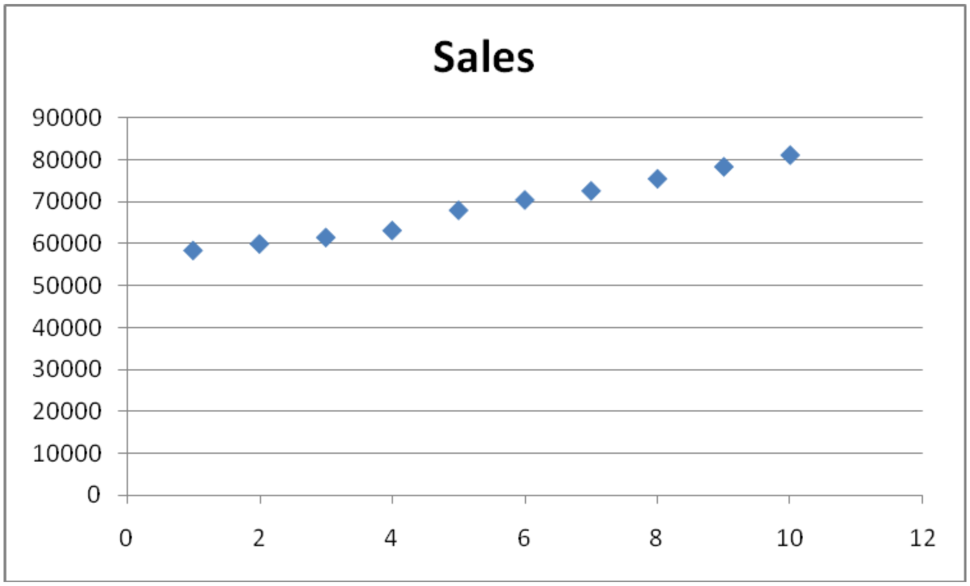
Note: Excel answers may be different due to rounding.

b. *Y*¢= 7.667 + 0.0023*t*

c.

Period	*Production*	*Index*	*Forecast*
21	7.7153	0.5577	4.3028
22	7.7176	0.8296	6.4025
23	7.7199	1.5178	11.7173
24	7.72221.1029		8.5168

30. a.



b. The regression equation is $Y' = 2624.576t + 54\,474.733$

c. $Y' = 2624.576(3) + 54\,474.733 = 62\,348.5$ and $Y' = 2624.576(8) + 54\,474.733 = 75\,471.3$

d. $Y' = 2624.576(13) + 54\,474.733 = 88\,594.2$. Since R^2 for the linear regression is nearly 1.0, the fit is excellent.

e. Sales are increasing by about $2625 per year.

31. *Seasonal Index by Quarter*

Quarter	*Average SI Component*	*Seasonal Index*
1	1.1962	1.2053
2	1.0135	1.0212
3	0.6253	0.6301
4	1.1371	1.1457

The regression equation is: *Y*¢= 43.611 + 7.21153*t*

Period	*Visitors*	*Index*	*Forecast*
29	252.86	1.2053	304.77
30	260.07	1.0212	265.58
31	267.29	0.6301	168.42
32	274.50	1.1457	314.50

In 2011 there were a total of 928 visitors. A ten percent increase in 2012 means there will be 1021 visitors. The quarterly estimates are 1021/4 = 255.25 visitors per quarter.

Period	*Visitors*	*Index*	*Forecast*
Winter	255.25	1.2053	307.65
Spring	255.25	1.0212	260.66
Summer	255.25	0.6301	160.83
Fall	255.25	1.1457	292.44

The regression approach is probably superior because the trend is considered.

32. a. *Seasonal Index by Quarter*

Quarter	*Average SI Component*	*Seasonal Index*
1	1.1900	1.1896
2	1.1044	1.1040
3	0.4449	0.4447
4	1.2621	1.2617

b. Fall quarter enrollment is the largest, about 26% above the average and the summer quarter is the lowest, about 44% of average enrollment.

c. $Y' = 1574 + 49.3t$

Period	*Visitors*	*Index*	*Forecast*
21	2609.3	1.1896	3104
22	2658.6	1.1040	2935
23	2707.9	0.4447	1204
24	2757.2	1.2617	3479

33. Purse regression equation is Purse = 134,740 + 57,651*t
Prize regression equation is Prize = 20,211 + 8648*t.

Notice both the slope and the intercept of the second equation are 15% of the corresponding part of the first equation. The prize is always 15% of the purse. The projected purse for 2011 is $1.52 million, found by 134,740 + 57,651(24) and the fitted prize is $227,755.

34.

a	linear	r^2	0.575		*t*	*Predicted*
	log	r^2	0.593	use log	12	15,315.371
b	linear	r^2	0.129		*t*	*Predicted*
	log	r^2	0.127	use linear	12	11,841.403
c	linear	r^2	0.107		*t*	*Predicted*
	log	r^2	0.089	use linear	12	2,363.888
d	linear	r^2	0.283		*t*	*Predicted*
	log	r^2	0.250	use linear	12	2,610.831
e	linear	r^2	0.059		*t*	*Predicted*
	log	r^2	0.055	use linear	12	1275.96

CHAPTER 17

AN INTRODUCTION TO DECISION MAKING

1. $EMV(A_1) = 0.30(\$50) + 0.50(\$70) + 0.20(\$100) = \$70$
   $EMV(A_2) = 0.30(\$90) + 0.50(\$40) + 0.20(\$80) = \$63$
   $EMV(A_3) = 0.30(\$70) + 0.50(\$60) + 0.20(\$90) = \$69$
   Decision: Choose alternative 1

2. Choose returnables because *EMV* (returnables) is higher
   EMV (returnables) = \$80(0.70) + \$40(0.30) = \$68 thousand
   EMV (nonreturnables) = \$25(0.70) + \$60(0.30) = \$35.5 thousand
   Use returnable bottles

3. *Opportunity loss*

	S_1	S_2	S_3
A_1	\$40	\$ 0	\$ 0
A_2	0	30	20
A_3	20	10	10

4. *Opportunity loss (\$000)*

Type of bottle	*Passed*	*Not passed*
Returnable	0	\$20
Nonreturnable	\$55	0

5. Answers in \$000
   EOL(A_1) = 0.30(\$40) + 0.50(\$0) + 0.20(\$0) = \$12
   EOL(A_2) = 0.30(\$0) + 0.50(\$30) + 0.20(\$20) = \$19
   EOL(A_3) = 0.30(\$20) + 0.50(\$10) + 0.20(\$10) = \$13

6. Answers in \$000
   EOL (Returnables) = \$0(0.70) + \$20(0.30) = \$6 thousand
   EOL (Nonreturnables) = \$55(0.70) + 0(0.30) = \$38.5 thousand

7. Expected value under conditions of certainty is \$82, found by 0.30(\$90) + 0.50(\$70) + 0.20(\$100) *EVPI* = \$82 - \$70 = \$12

8. Condition of Certainty = \$80(0.70) + \$60(0.30) = \$74 thousand
   Then *EVPI* = \$74 - \$68 = \$6 thousand

9. Yes, it changes the decision. Choose alternative 2 (answers in \$000)
   EMV(A_1) = 0.50(\$50) + 0.20(\$70) + 0.30(\$100) = \$69
   EMV(A_2) = 0.50(\$90) + 0.20(\$40) + 0.30(\$80) = \$77
   EMV(A_3) = 0.50(\$70) + 0.20(\$60) + 0.30(\$90) = \$74

10. Choose returnables. Does not alter decision.
    EMV (returnables) = \$80(0.30) + \$40(0.70) = \$52 thousand
    EMV (nonreturnables) = \$25(0.30) + \$60(0.70) = \$49.5 thousand

11. a. Answers in ($000)

EMV (neither) = 0.30($0) + 0.50($0) + 0.20($0) = $0
EMV (1) = 0.30($125) + 0.50($65) + 0.20($30) = $76
EMV (2) = 0.30($105) + 0.50($60) + 0.20($30) = $67.50
EMV (both) = 0.30($220) + 0.50($110) + 0.20($40) = $129

b. Choose both

c. *Opportunity loss*

	S_1	S_2	S_3
Neither	$220	$110	$40
1	95	45	10
2	115	50	10
Both	0	0	0

d. *EOL*(neither) = 0.30($220) + 0.50($110) + 0.20($40) = $129.00
EOL(1) = 0.30($95) + 0.50($45) + 0.20($10) = $53.00
EOL(2) = 0.30($115) + 0.50($50) + 0.20($10) = $61.50
EOL(both) = 0.30($0) + 0.50($0) + 0.20($0) = $0

e. *EVPI* = $0, found by $129 - $129
Certainty = 0.30($220) + 0.50($110) + 0.20($40) = $129

12.

	Weather	
Transportation	*Good*	*Bad*
Plane	$366.25	$372.50
Train	490	502.50
Car	370	445

EMV(plane) = $366.25(0.40) + 372.50(0.60) = $370.00
EMV(train) = $490(0.40) + $502.5(0.60) = $497.50
EMV(car) = $370(0.40) + $445(0.60) = $415.00
Choose the plane because $370 is the least cost.
Certainty = $366.25(0.40) + 372.50(0.60) = $370.00 *EPVI* = $370 - $370 = $0

13. The payoff table is as follows in $000

	Recession	*No Recession*
	S_1	S_2
Production	- $10.0	$15.0
Stock	- 5.0	12.0
CD	6.0	6.0

a. Purchase CD

b. Increase production

c. (Answers in $000)
EMV(Prod.) = 0.20(- 10) + 0.80(15.0) = 10.0
EMV(Stock) = 0.20(- 5) + 0.80(12.0) = 8.6
EMV(CD) = 0.20(6) + 0.80(6) = 6.0
Expand Production

d. *EVPI* = [0.20(6) + 0.80(15)] - [10.0] = 13.2 – 10.0 = 3.2

14. Payoff Table

a.

	S_1	S_2	S_3
No Inspection	$7.2	$14.4	$21.60
Inspection	10.0	10.0	10.00

b. *EMV*(No inspect) = 0.70(7.2) + 0.20(14.4) + 0.1(21.60) = $10.08

Inspect = $10.0

c. *EVPI* = [0.70(7.2) + 0.20(10) + 0.10(10)] – 10.0 = - 1.96

15. a.

	Event				
Act	*10*	*11*	*12*	*13*	*14*
10	$500	$500	$500	$500	$500
11	200	550	550	550	550
12	- 100	250	600	600	600
13	- 400	- 50	300	650	650
14	- 700	- 350	0	350	700

b.

Act	*Expected profit*
10	$500.00
11	504.50
12	421.50
13	233.50
14	- 31.50

Order 11 mobile homes because expected profit of $504.50 is the highest.

c.

	Opportunity Loss				
Supply	*10*	*11*	*12*	*13*	*14*
10	$0	$50	$100	$150	$200
11	300	0	50	100	150
12	600	300	0	50	100
13	900	600	300	0	50
14	1200	900	600	300	0

d.

			Act		
	10	11	12	13	14
Expect. Opp. Loss	$95.50	$91	$174	$362	$627

Decision: Order 11 homes because the opportunity loss of $91 is the smallest.

e. $91, found by $595.50 – 504.50 = $91.00 value of perfect information

16. a.

	Event			
Act	*7*	*8*	*9*	*10*
7	$35	$35	$35	$35
8	15	40	40	40
9	- 5	20	45	45
10	- 25	0	25	50

b. Expected profits are:

Demand	*Expected Payoff*
7	$35.00
8	37.50
9	33.75
10	18.75

The computation for 9 snowmobiles is:

Event	*Probability*	*Payoff*	*Expected Profit*
X	*P(X)*	*Y*	*P(X)Y*
7	0.10	$- 5	$- 0.50
8	0.25	20	5.00
9	0.45	45	20.25
10	0.20	45	9.00
	Total		33.75

c. Lease 8 snowmobiles because the expected profit of $37.50 is the highest.

d. Opportunity loss table is:

Act	*Opportunity loss*			
Number available	*7*	*8*	*9*	*10*
7	0	5	10	15
8	20	0	5	10
9	40	20	0	5
10	60	40	20	0

e. Computations for 8 snowmobiles:

Demand	*Probability*	*Opp.Loss*	*Expected Loss*
X	*P(X)*	*OL*	*P(X)OL*
7	0.10	$20	$2.00
8	0.25	0	0.00
9	0.45	5	2.25
10	0.20	10	2.00
Total			$6.25

The expected opportunity losses are:

	Number available			
	7	8	9	10
Expected Opp. Loss	$8.75	$6.25	$10.00	$25.00

f. Lease 8 snowmobiles

g. The expected value of perfect information is $6.25. Profit under certainty is $43.75, found by:

Event	*Profit*	*Total Profit*	*Probability*	*Expected Profit*
7	$5	$35	0.10	$3.50
8	5	40	0.25	10.00
9	5	45	0.45	20.25
10	5	50	0.20	10.00
Total				$43.75

Then, $43.75 - $37.50 = $6.25

h. All evidence indicates that leasing 8 snowmobiles would be the most profitable.

17. a.

	Event					
Act	*41*	*42*	*42*	*44*	*45*	*46*
41	$410	$410	$410	$410	$410	$410
42	405	420	420	420	420	420
43	400	415	430	430	430	430
44	395	410	425	440	440	440
45	390	405	420	435	450	450
46	385	400	415	430	445	460

b. Expected profits are:

Act	*Expected Payoff*
41	$410.00
42	419.10
43	426.70
44	432.20
45	431.70
46	427.45

c. Order 44 because $432.20 is the largest expected profit.

d. Expected opportunity loss:

41	42	43	44	45	46
$28.30	$19.20	$11.60	$6.10	$6.60	$10.85

e. Order 44 because the opportunity loss of $6.10 is the smallest. Yes, it agrees.

f. $6.10, found by $438.30 - $432.20 = $6.10 value of perfect information
The maximum we should pay for perfect information is $6.10.

18. a. The cost per car is $4000, found by $6000 - $2000. If Tim purchased 20 cars and he can rent 20 cars the payoff is $12,500. It is computed as follows:

(20 cars)(5 days)(50 weeks)($20 - $1.50) - (20 cars)($4000) = $92,500 - $80,000 = $12,500 The other payoffs are computed in a similar fashion.

Payoff (in$000)
States of Nature

		20	*21*	*22*	*23*	*EMV(A_1)*
A_1	20	12.5	12.500	12.500	12.500	12.5000
A_2	21	8.5	13.125	13.125	13.125	12.6625
A_3	22	4.5	9.125	13.750	13.750	11.9000
A_4	23	0.5	5.125	9.750	14.375	8.8250

$EMV(A_1) = 1(12.5) = 12.50$
$EMV(A_2) = 0.10(8.5)+0.90(13.125)=12.6625$
$EMV(A_3) = 0.10(0.45) + 0.20(9.125) + 0.70(13.75) = 11.90$
$EMV(A_4) = 0.10(0.5) + 0.20(5.125) + 0.50(9.75) + 0.20(14.375)) = 8.825$

b. $EVCP = 0.1(12.5) + 0.20(13.125) + 0.50(13.75) + 0.20(14.375) = 13.625$
$EVPI – 13.625 – 12.6625 = 0.9625$
$962.50

19. a.

	Event			
Option	*100*	*300*	*500*	*700*
1	$29.99	$39.99	$59.99	$79.99
2	34.99	34.99	44.99	64.99
3	59.99	59.99	59.99	59.99

b. Expected costs are:

Option	*Expected Cost*	
1	$52.49	found by .25(29.99) +.25(39.99) +.25(59.99) +.25(79.99)
2	44.99	found by .25(34.99) +.25(34.99) +.25(44.99) +.25(64.99)
3	59.99	found by .25(59.99) +.25(59.99) +.25(59.99) +.25(59.99)

Option 2 is best.

c. Option 1, because 29.99 is lower than 34.99 or 59.99.

d. Option 3, because 59.99 is lower than 79.99 or 64.99.

e.

	Event			
Option	*100*	*300*	*500*	*700*
1	$0	$5	$15	$20
2	5	0	0	5
3	30	25	15	0

f. Option 2, because 5 is lower than 20 or 30.

g. *EVPI* = 44.99 – [.25(29.99) +.25(34.99) +.25(44.99) +.25(59.99)]
= 44.99 – 42.49 = 2.50

20.

	Engine Condition	
Test	*Out of tune*	*Okay*
Yes	$80	$20
No	100	0
Probability	0.30	0.70

EMV(Test) = $0.30(80) + 0.70(20) = $38
EMV(No Test) = $0.30(100) + 0.70(0) = $30

It seems the test is not advisable because $30 is less than $38.