Chapter 2
Descriptive Statistics: Tabular and Graphical Displays

Learning Objectives

1. Learn how to construct and interpret summarization procedures for qualitative data such as: frequency and relative frequency distributions, bar graphs and pie charts.

2. Learn how to construct and interpret tabular summarization procedures for quantitative data such as: frequency and relative frequency distributions, cumulative frequency and cumulative relative frequency distributions.

3. Learn how to construct a dot plot and a histogram as graphical summaries of quantitative data.

4. Learn how the shape of a data distribution is revealed by a histogram. Learn how to recognize when a data distribution is negatively skewed, symmetric, and positively skewed.

5. Be able to use and interpret the exploratory data analysis technique of a stem-and-leaf display.

6. Learn how to construct and interpret cross tabulations, scatter diagrams, side-by-side and stacked bar charts.

7. Learn best practices for creating effective graphical displays and for choosing the appropriate type of display.

Solutions:

1.

Class	Frequency	Relative Frequency
A	60	60/120 = 0.50
B	24	24/120 = 0.20
C	36	36/120 = 0.30
	120	1.00

2. a. $1 - (.22 + .18 + .40) = .20$

   b. $.20(200) = 40$

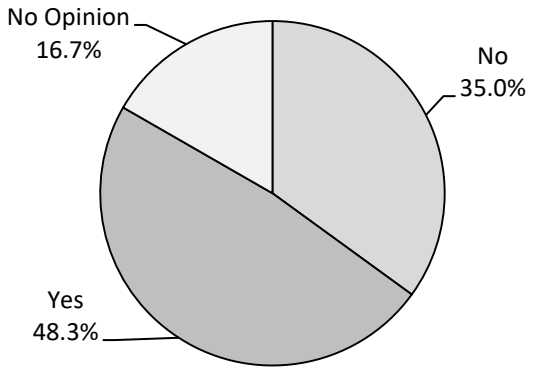
   c/d.

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

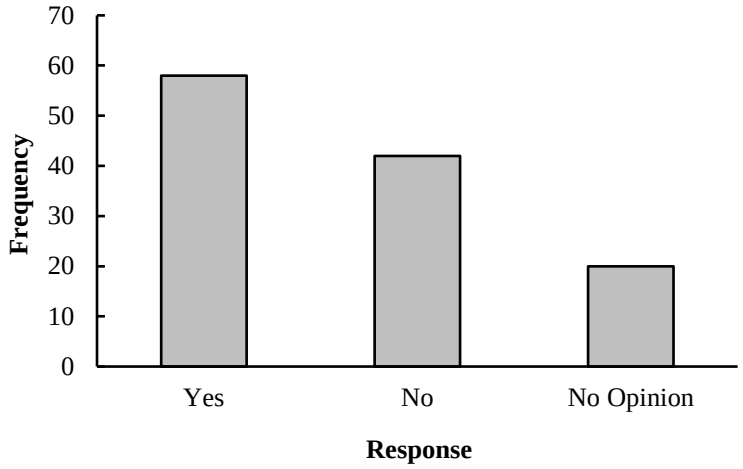
3. a. $360° \times 58/120 = 174°$

   b. $360° \times 42/120 = 126°$

   c.



d.

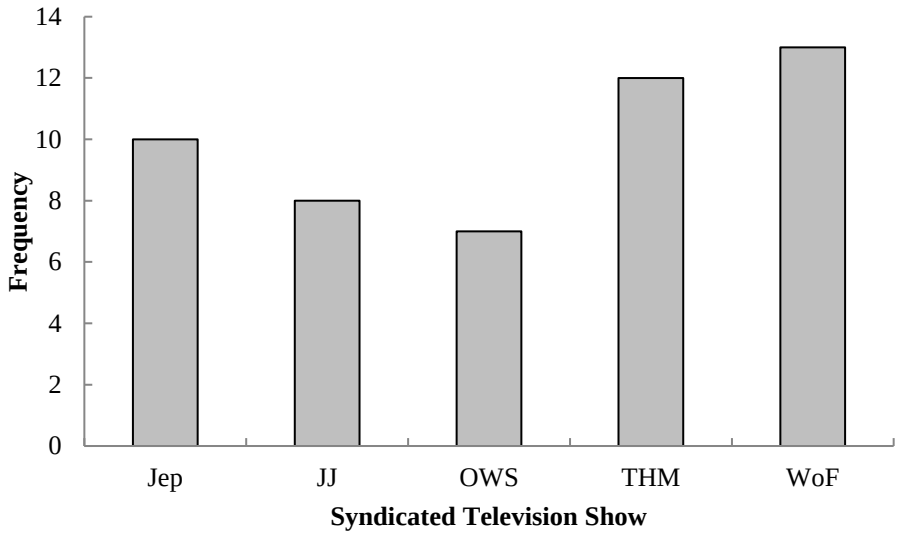


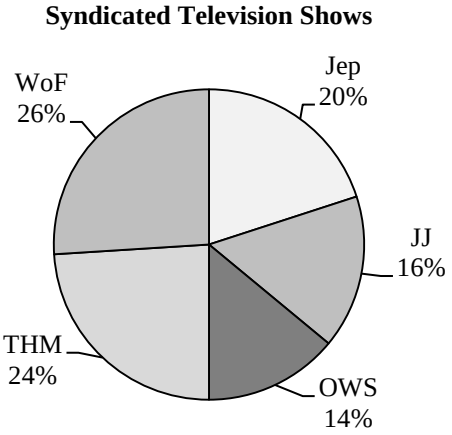
4. a. These data are categorical.

b.

Show	Relative Frequency	% Frequency
Jep	10	20
JJ	8	16
OWS	7	14
THM	12	24
WoF	13	26
Total	50	100

c.



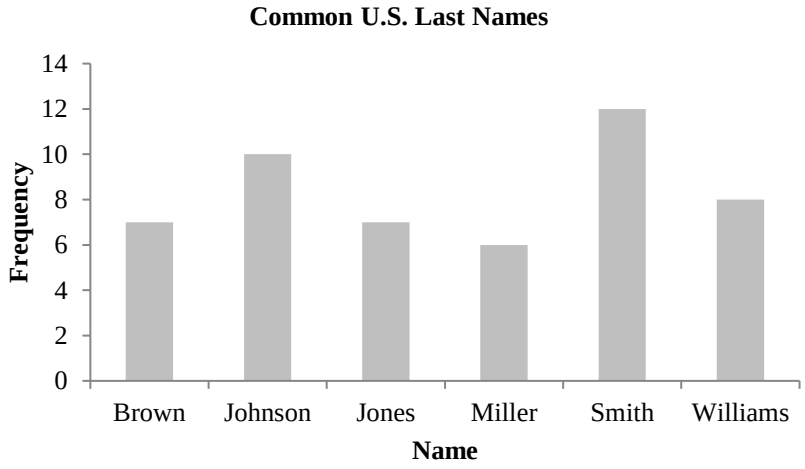


d. The largest viewing audience is for *Wheel of Fortune* and the second largest is for *Two and a Half Men*.

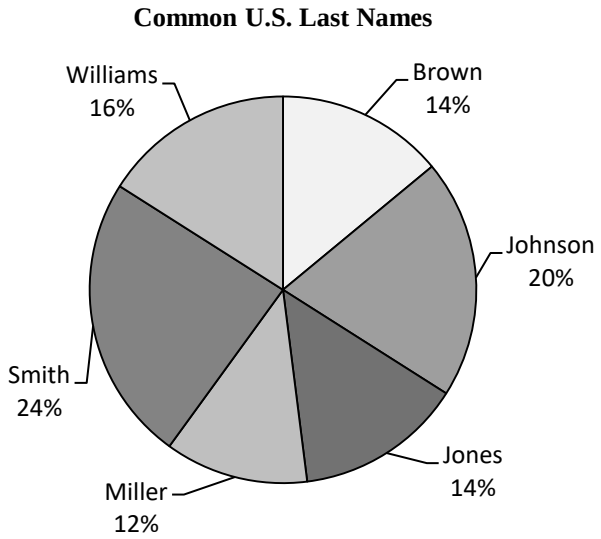
5. a.

Name	Frequency	Relative Frequency	Percent Frequency
Brown	7	0.14	14%
Johnson	10	0.20	20%
Jones	7	0.14	14%
Miller	6	0.12	12%
Smith	12	0.24	24%
Williams	8	0.16	16%
Total:	50	1	100%

b.



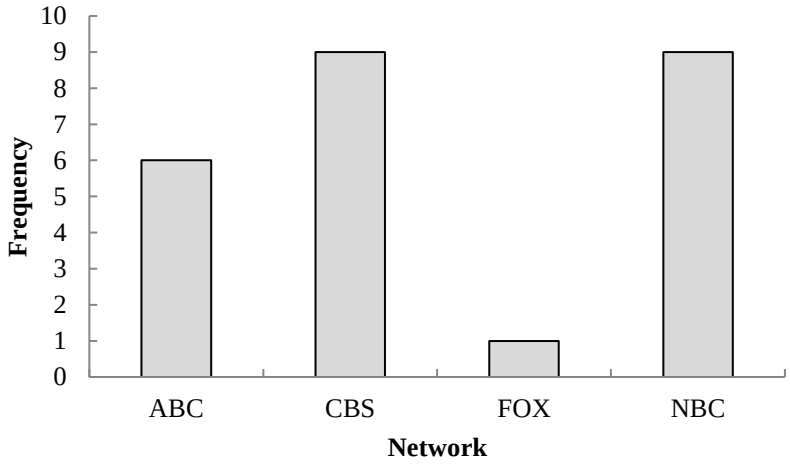
c.



d. The three most common last names are Smith (24%), Johnson (20%), Williams (16%5)

6. a.

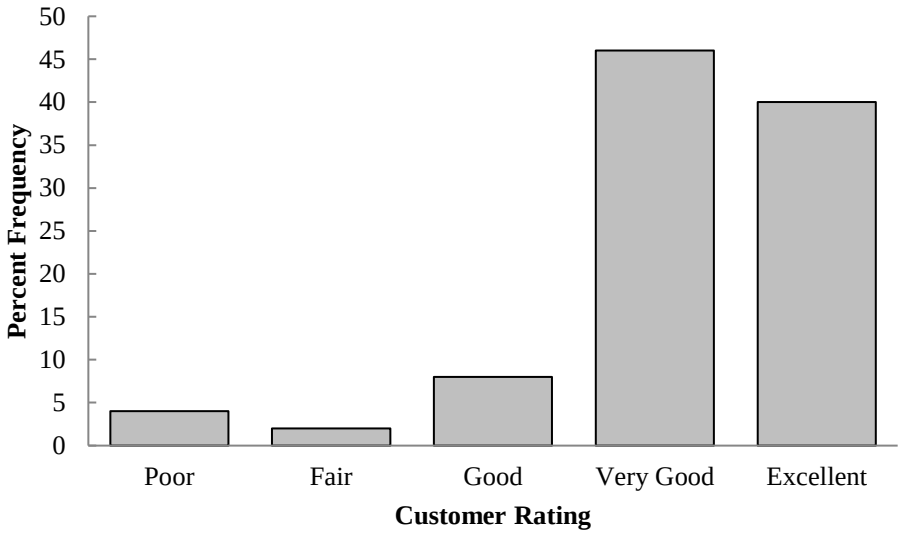
Network	Relative Frequency	% Frequency
ABC	6	24
CBS	9	36
FOX	1	4
NBC	9	36
Total:	25	100



b. For these data, NBC and CBS tie for the number of top-rated shows. Each has 9 (36%) of the top 25. ABC is third with 6 (24%) and the much younger FOX network has 1(4%).

7. a.

Rating	Frequency	Percent Frequency
Excellent	20	40
Very Good	23	46
Good	4	8
Fair	1	2
Poor	2	4
	50	100



Management should be very pleased with the survey results. 40% + 46% = 86% of the ratings are very good to excellent. 94% of the ratings are good or better. This does not look to be a Delta flight where significant changes are needed to improve the overall customer satisfaction ratings.

b. While the overall ratings look fine, note that one customer (2%) rated the overall experience with the flight as Fair and two customers (4%) rated the overall experience with the flight as Poor. It might be insightful for the manager to review explanations from these customers as to how the flight failed to meet expectations. Perhaps, it was an experience with other passengers that Delta could do little to correct or perhaps it was an isolated incident that Delta could take steps to correct in the future.

8. a.

Position	Frequency	Relative Frequency
Pitcher	17	0.309
Catcher	4	0.073
1st Base	5	0.091
2nd Base	4	0.073
3rd Base	2	0.036
Shortstop	5	0.091
Left Field	6	0.109
Center Field	5	0.091
Right Field	7	0.127
	55	1.000

b. Pitchers (Almost 31%)

c. 3rd Base (3 – 4%)

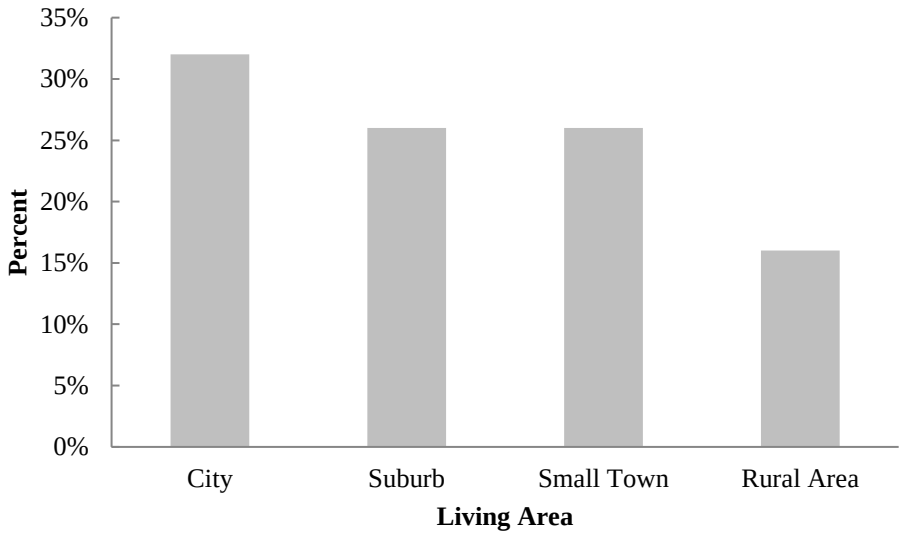
d. Right Field (Almost 13%)

e. Infielders (16 or 29.1%) to Outfielders (18 or 32.7%)

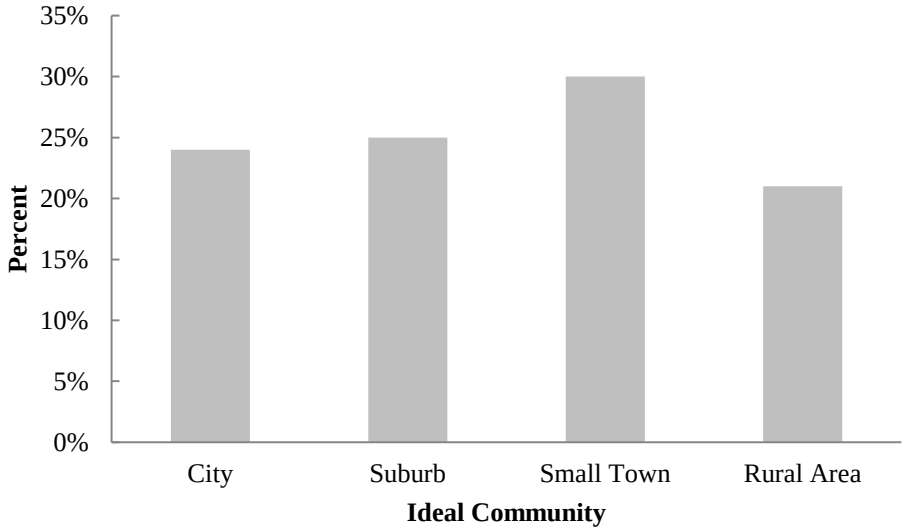
9. a.

Living Area	Live Now	Ideal Community
City	32%	24%
Suburb	26%	25%
Small Town	26%	30%
Rural Area	16%	21%
Total	100%	100%

b. Where do you live now?



What do you consider the ideal community?



c. Most adults are now living in a city (32%).

d. Most adults consider the ideal community a small town (30%).

e. Percent changes by living area: City –8%, Suburb –1%, Small Town +4%, and Rural Area +5%. Suburb living is steady, but the trend would be that living in the city would decline while living in small towns and rural areas would increase.

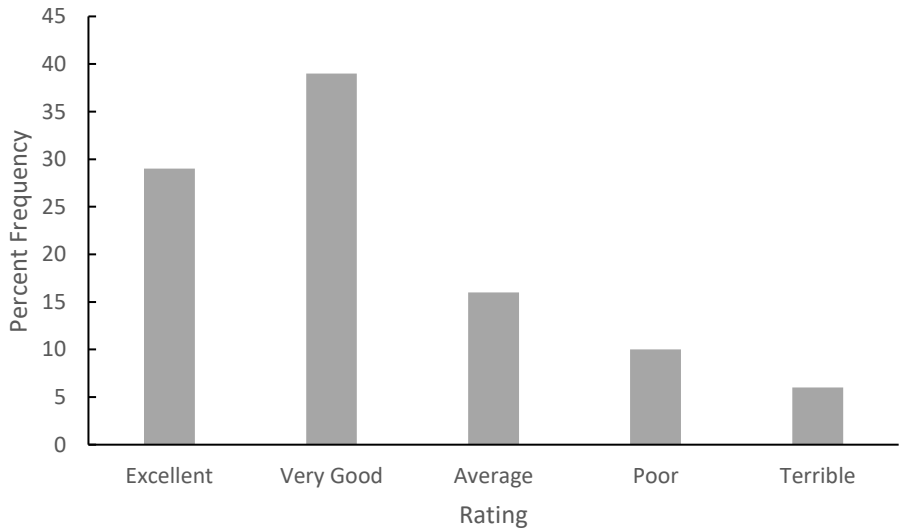
10. a.

Rating	Frequency
Excellent	187
Very Good	252
Average	107
Poor	62
Terrible	41
Total	649

b.

Rating	Percent Frequency
Excellent	29
Very Good	39
Average	16
Poor	10
Terrible	6
Total	100

c.



d. 29% + 39% = 68% of the guests at the Sheraton Anaheim Hotel rated the hotel as Excellent or Very Good. But, 10% + 6% = 16% of the guests rated the hotel as poor or terrible.

e. The percent frequency distribution for Disney's Grand Californian follows:

Rating	Percent Frequency
Excellent	48
Very Good	31
Average	12
Poor	6
Terrible	3
Total	100

48% + 31% = 79% of the guests at the Sheraton Anaheim Hotel rated the hotel as Excellent or Very Good. And, 6% + 3% = 9% of the guests rated the hotel as poor or terrible.

Compared to ratings of other hotels in the same region, both of these hotels received very favorable ratings. But, in comparing the two hotels, guests at Disney's Grand Californian provided somewhat better ratings than guests at the Sheraton Anaheim Hotel.

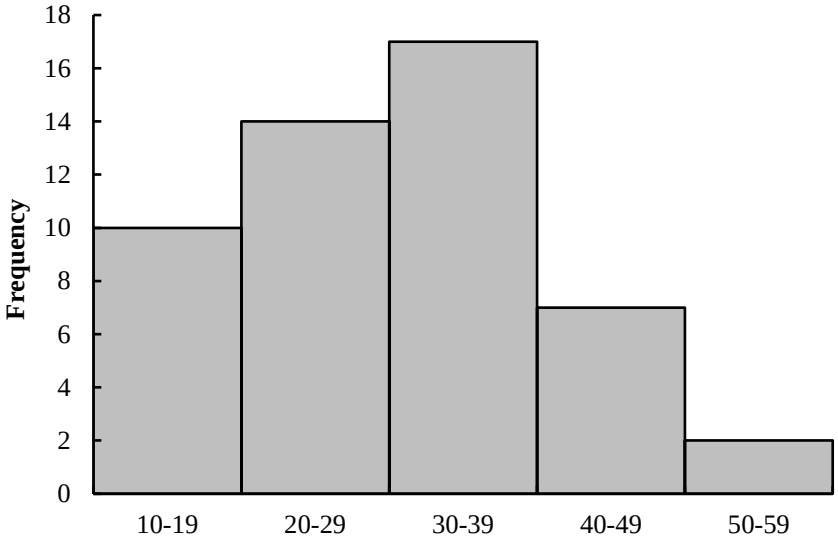
11.

Class	Frequency	Relative Frequency	Percent Frequency
12–14	2	0.050	5.0
15–17	8	0.200	20.0
18–20	11	0.275	27.5
21–23	10	0.250	25.0
24–26	9	0.225	22.5
Total	40	1.000	100.0

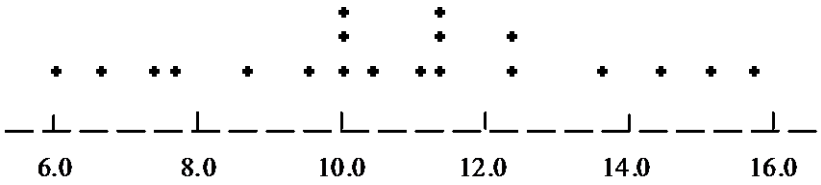
12.

Class	Cumulative Frequency	Cumulative Relative Frequency
less than or equal to 19	10	.20
less than or equal to 29	24	.48
less than or equal to 39	41	.82
less than or equal to 49	48	.96
less than or equal to 59	50	1.00

13.



14. a.



b/c.

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

15. Leaf Unit = .1

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

16. Leaf Unit = 10

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

17. a/b.

Waiting Time	Frequency	Relative Frequency
0 – 4	4	0.20
5 – 9	8	0.40
10 – 14	5	0.25
15 – 19	2	0.10
20 – 24	1	0.05
Totals	20	1.00

c/d.

Waiting Time	Cumulative Frequency	Cumulative Relative Frequency
Less than or equal to 4	4	0.20
Less than or equal to 9	12	0.60
Less than or equal to 14	17	0.85
Less than or equal to 19	19	0.95
Less than or equal to 24	20	1.00

e. 12/20 = 0.60

18. a.

PPG	**Frequency**
10-12	1
12-14	3
14-16	7
16-18	19
18-20	9
20-22	4
22-24	2
24-26	0
26-28	3
28-30	2
Total	50

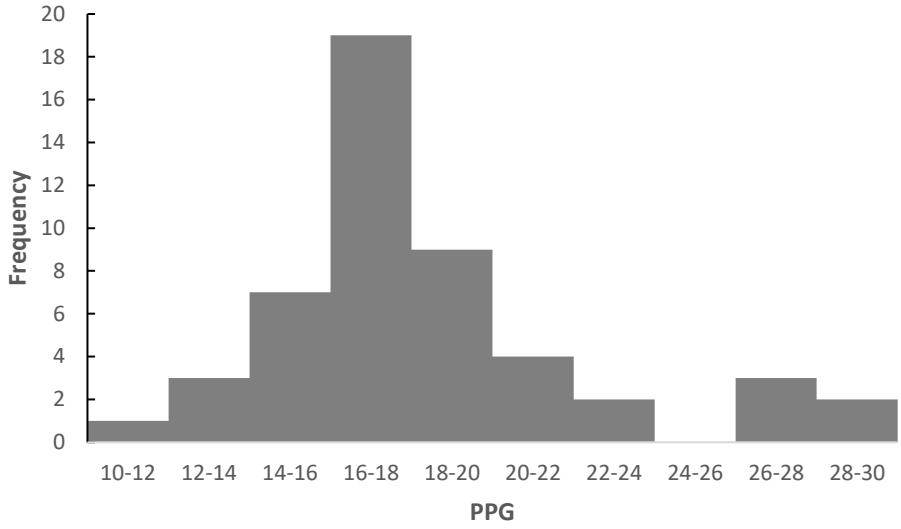
b.

PPG	Relative Frequency
10-12	0.02
12-14	0.06
14-16	0.14
16-18	0.38
18-20	0.18
20-22	0.08
22-24	0.04
24-26	0.00
26-28	0.06
28-30	0.04
Total	1.00

c.

PPG	Cumulative Percent Frequency
less than 12	2
less than 14	8
less than 16	22
less than 18	60
less than 20	78
less than 22	86
less than 24	90
less than 26	90
less than 28	96
less than 30	100

d.



e. There is skewness to the right.

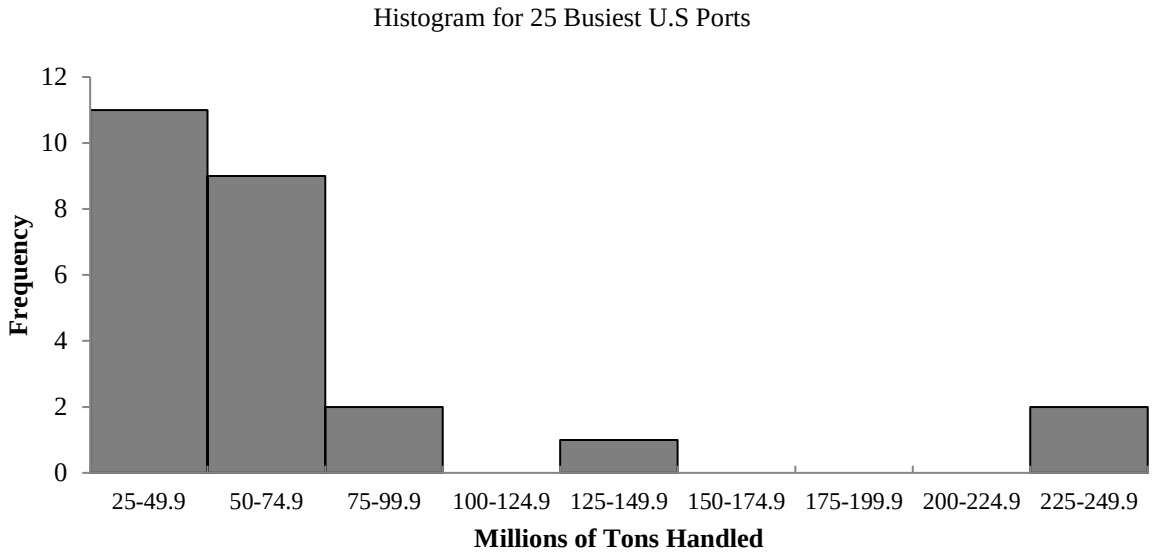
f. (11/50)(100) = 22%

19. a. The largest number of tons is 236.3 million (South Louisiana). The smallest number of tons is 30.2 million (Port Arthur).

b.

Millions Of Tons	Frequency
25-50	11
50-75	9
75-100	2
100-125	0
125-150	1
150-175	0
175-200	0
200-225	0
225-250	2

c.



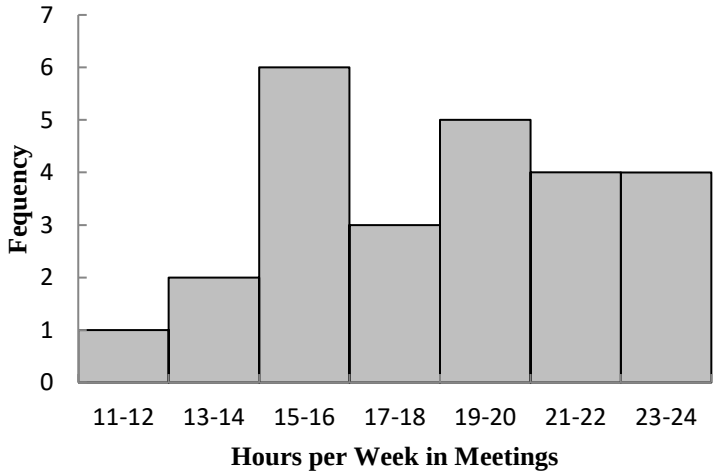
Most of the top 25 ports handle less than 75 million tons. Only five of the 25 ports handle above 75 million tons.

20. a. Lowest = 12, Highest = 23

b.

Hours in Meetings per Week	Frequency	Percent Frequency
11-12	1	4%
13-14	2	8%
15-16	6	24%
17-18	3	12%
19-20	5	20%
21-22	4	16%
23-24	4	16%
	25	100%

c.



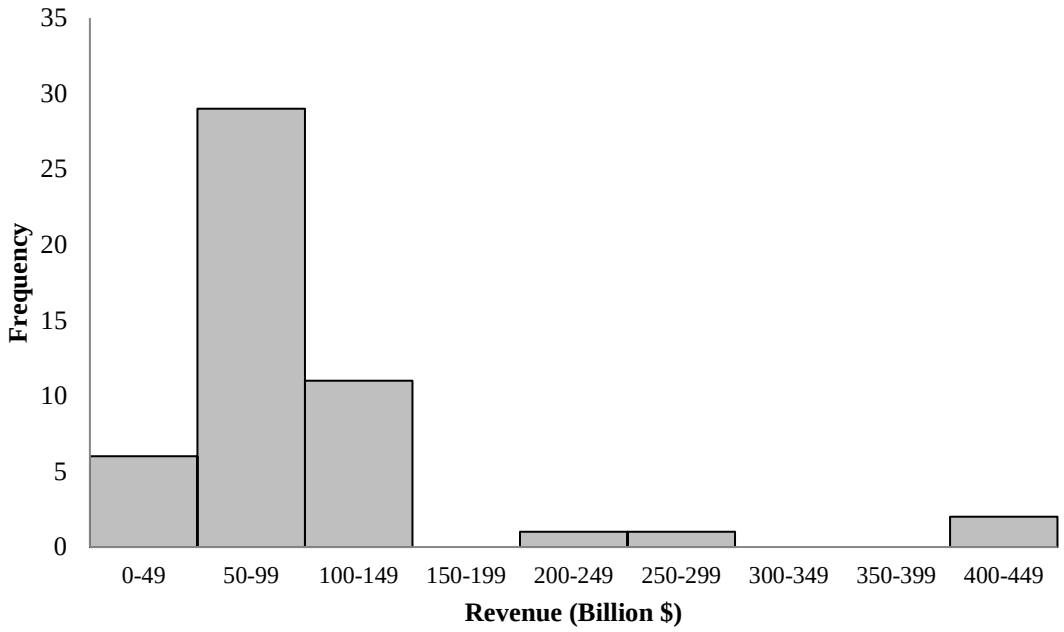
The distribution is slightly skewed to the left.

21. a/b/c/d.

Revenue	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0-49	6	.12	6	.12
50-99	29	.58	35	.70
100-149	11	.22	46	.92
150-199	0	.00	46	.92
200-249	1	.02	47	.94
250-299	1	.02	48	.96
300-349	0	.00	48	.96
350-399	0	.00	48	.96
400-449	2	.04	50	1.00
Total	50	1.00		

e. The majority of the large corporations (40) have revenues in the $50 billion to $149 billion range. Only 4 corporations have revenues of over $200 billion and only 2 corporations have revenues over $400 billion. .70, or 70%, of the corporations have revenues under $100 billion. .30, or 30%, of the corporations have revenues of $100 billion or more.

f.



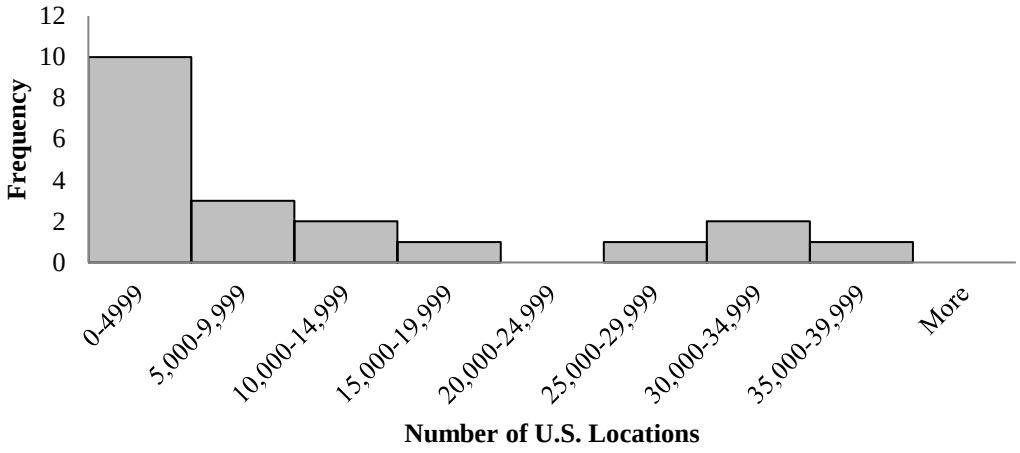
The histogram shows the distribution is skewed to the right with four corporations in the $200 to $449 billion range.

g. Exxon-Mobil is America's largest corporation with an annual revenue of $443 billion. Walmart is the second largest corporation with an annual revenue of $406 billion. All other corporations have annual revenues less than $300 billion. Most (92%) have annual revenues less than $150 billion.

22. a.

# U.S. Locations	Frequency	Percent Frequency
0-4999	10	50
5000-9999	3	15
10000-14999	2	10
15000-19999	1	5
20000-24999	0	0
25000-29999	1	5
30000-34999	2	10
35000-39999	1	5
Total:	20	100

b.



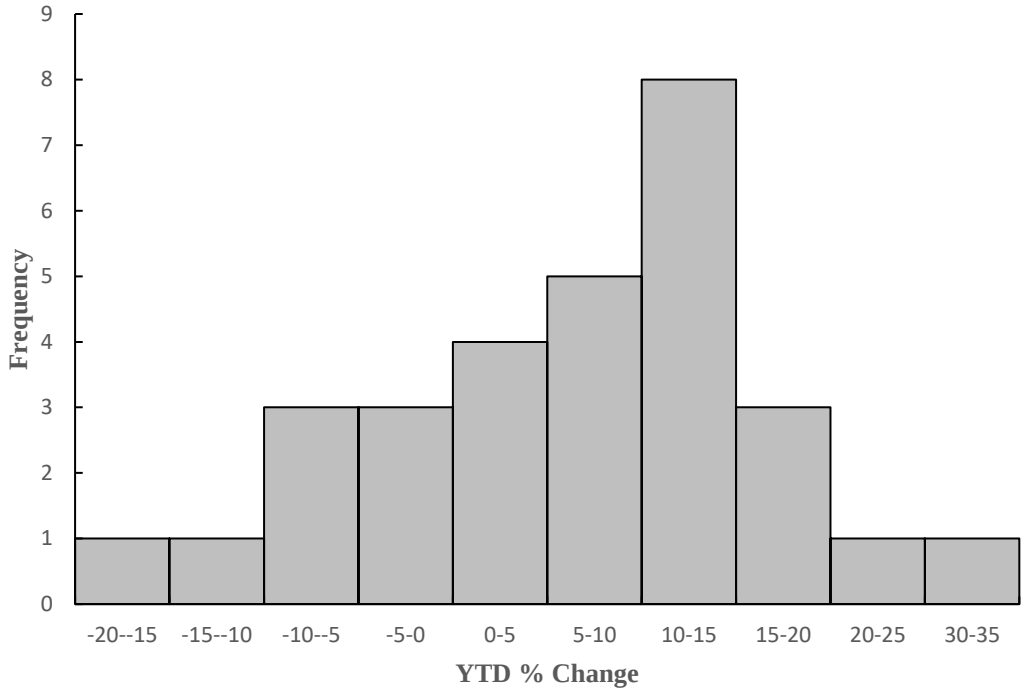
c. The distribution is skewed to the right. The majority of the franchises in this list have fewer than 20,000 locations (50% + 15% + 15% = 80%). McDonald's, Subway and 7-Eleven have the highest number of locations.

23. a. The highest positive YTD % Change for Japan's Nikkei index with a YTD % Change of 31.4%.

b. A class size of 10 results in 10 classes.

YTD % Change	Frequency
-20--15	1
-15--10	1
-10--5	3
-5-0	3
0-5	4
5-10	5
10-15	8
15-20	3
20-25	1
30-35	1

c.



The general shape of the distribution is skewed to the right. Twenty two of the 30 indexes have a positive YTD % Change and 13 have a YTD % Change of 10% or more. Eight of the indexes had a negative YTD % Change.

d. A variety of comparisons are possible depending upon when the study is done.

24. Median Pay

6	6 7 7
7	2 4 6 7 7 8 9
8	0 0 1 3 7
9	9
10	0 6
11	0
12	1

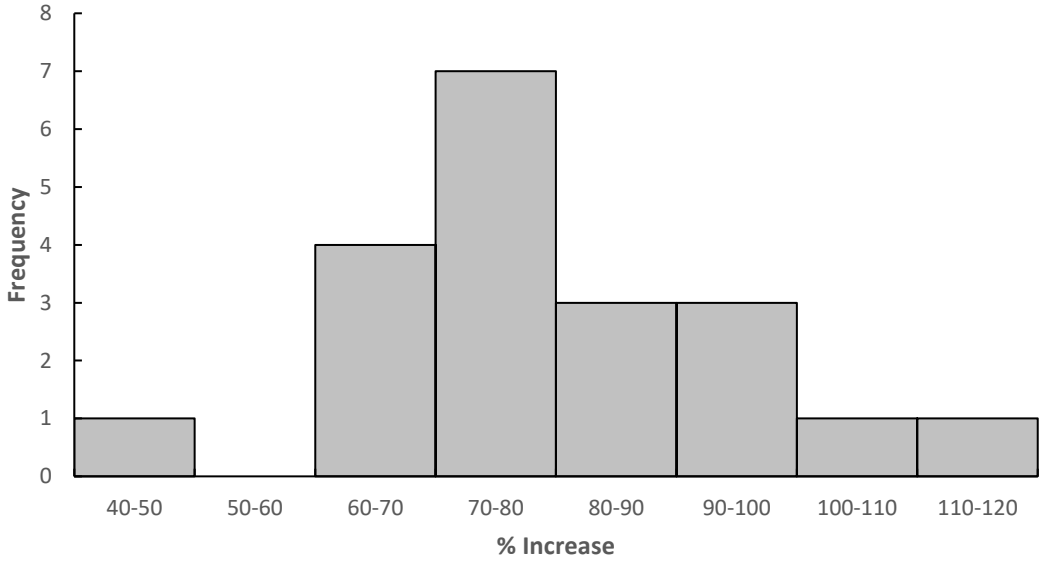
The median pay for these careers is generally in the $70 and $80 thousands. Only four careers have a median pay of $100 thousand or more. The highest median pay is $121 thousand for a finance director.

Top Pay

10	0 6 9
11	1 6 9
12	2 5 6
13	0 5 8 8
14	0 6
15	2 5 7
16	
17	
18	
19	
20	
21	4
22	1

The most frequent top pay is in the $130 thousand range. However, the top pay is rather evenly distributed between $100 and $160 thousand. Two unusually high top pay values occur at $214 thousand for a finance director and $221 thousand for an investment banker. Also, note that the top pay has more variability than the median pay.

25. a.



b. The histogram is skewed to the right.

c.

4	3
5	
6	1 3 7 9
7	1 3 4 5 7 7 9
8	2 4 7
9	0 3 6
10	0
11	3

d. Rotating the stem-and-leaf display counterclockwise onto its side provides a picture of the data that is similar to the histogram in shown in part (a). Although the stem-and-leaf display may appear to offer the same information as a histogram, it has two primary advantages: the stem-and-leaf display is easier to construct by hand; and the stem-and-leaf display provides more information than the histogram because the stem-and-leaf shows the actual data.

26. a.

2	1 4
2	6 7
3	0 1 1 1 2 3
3	5 6 7 7
4	0 0 3 3 3 3 3 4 4
4	6 6 7 9
5	0 0 0 2 2
5	5 6 7 9
6	1 4
6	6
7	2

b. Most frequent age group: 40-44 with 9 runners

c. 43 was the most frequent age with 5 runners

27. a.

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

b.

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

c.

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

d. Category A values for *x* are always associated with category 1 values for *y*. Category B values for *x* are usually associated with category 1 values for y. Category C values for *x* are usually associated with category 2 values for *y*.

28. a.

		y				
		20-39	40-59	60-79	80-100	Grand Total
	10-29			1	4	5
x	30-49	2		4		6
	50-69	1	3	1		5
	70-90	4				4
	Grand Total	7	3	6	4	20

b.

		y				
		20-39	40-59	60-79	80-100	Grand Total
	10-29			20.0	80.0	100
x	30-49	33.3		66.7		100
	50-69	20.0	60.0	20.0		100
	70-90	100.0				100

c.

		y			
		20-39	40-59	60-79	80-100
	10-29	0.0	0.0	16.7	100.0
x	30-49	28.6	0.0	66.7	0.0
	50-69	14.3	100.0	16.7	0.0
	70-90	57.1	0.0	0.0	0.0
	Grand Total	100	100	100	100

d. Higher values of x are associated with lower values of y and vice versa

29. a.

	Average Miles per Hour					
Make	130-139.9	140-149.9	150-159.9	160-169.9	170-179.9	Total
Buick	100.00	0.00	0.00	0.00	0.00	100.00
Chevrolet	18.75	31.25	25.00	18.75	6.25	100.00
Dodge	0.00	100.00	0.00	0.00	0.00	100.00
Ford	33.33	16.67	33.33	16.67	0.00	100.00

b. 25.00 + 18.75 + 6.25 = 50 percent

c.

	Average Miles per Hour				
Make	130-139.9	140-149.9	150-159.9	160-169.9	170-179.9
Buick	16.67	0.00	0.00	0.00	0.00
Chevrolet	50.00	62.50	66.67	75.00	100.00
Dodge	0.00	25.00	0.00	0.00	0.00
Ford	33.33	12.50	33.33	25.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

d. 75%

30. a.

	Year					
Average Speed	1988-1992	1993-1997	1998-2002	2003-2007	2008-2012	Total
130-139.9	16.7	0.0	0.0	33.3	50.0	100
140-149.9	25.0	25.0	12.5	25.0	12.5	100
150-159.9	0.0	50.0	16.7	16.7	16.7	100
160-169.9	50.0	0.0	50.0	0.0	0.0	100
170-179.9	0.0	0.0	100.0	0.0	0.0	100

b. It appears that most of the faster average winning times occur before 2003. This could be due to new regulations that take into account driver safety, fan safety, the environmental impact, and fuel consumption during races.

31. a. The crosstabulation of condition of the greens by gender is below.

	Green Condition		
Gender	Too Fast	Fine	Total
Male	35	65	100
Female	40	60	100
Total	75	125	200

The female golfers have the highest percentage saying the greens are too fast: 40/100 = 40%. Male golfers have 35/100 = 35% saying the greens are too fast.

b. Among low handicap golfers, 1/10 = 10% of the women think the greens are too fast and 10/50 = 20% of the men think the greens are too fast. So, for the low handicappers, the men show a higher percentage who think the greens are too fast.

c. Among the higher handicap golfers, 39/51 = 43% of the woman think the greens are too fast and 25/50 = 50% of the men think the greens are too fast. So, for the higher handicap golfers, the men show a higher percentage who think the greens are too fast.

d. This is an example of Simpson's Paradox. At each handicap level a smaller percentage of the women think the greens are too fast. But, when the crosstabulations are aggregated, the result is reversed and we find a higher percentage of women who think the greens are too fast.

The hidden variable explaining the reversal is handicap level. Fewer people with low handicaps think the greens are too fast, and there are more men with low handicaps than women.

32. a. Row percentages are shown below.

Region	**Under $15,000**	**$15,000 to $24,999**	**$25,000 to $34,999**	**$35,000 to $49,999**	**$50,000 to $74,999**	**$75,000 to $99,999**	**$100,000 and over**	**Total**
Northeast	12.72	10.45	10.54	13.07	17.22	11.57	24.42	100.00
Midwest	12.40	12.60	11.58	14.27	19.11	12.06	17.97	100.00
South	14.30	12.97	11.55	14.85	17.73	11.04	17.57	100.00
West	11.84	10.73	10.15	13.65	18.44	11.77	23.43	100.00

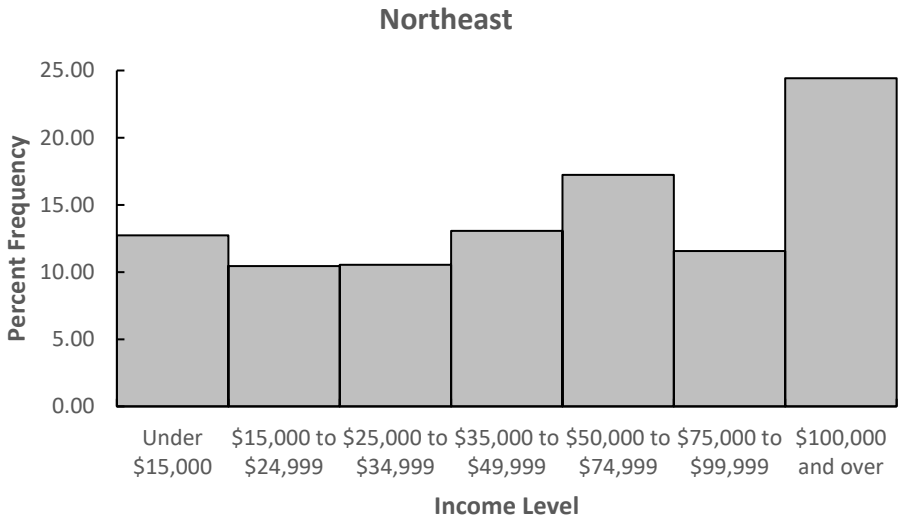
The percent frequency distributions for each region now appear in each row of the table. For example, the percent frequency distribution of the West region is as follows:

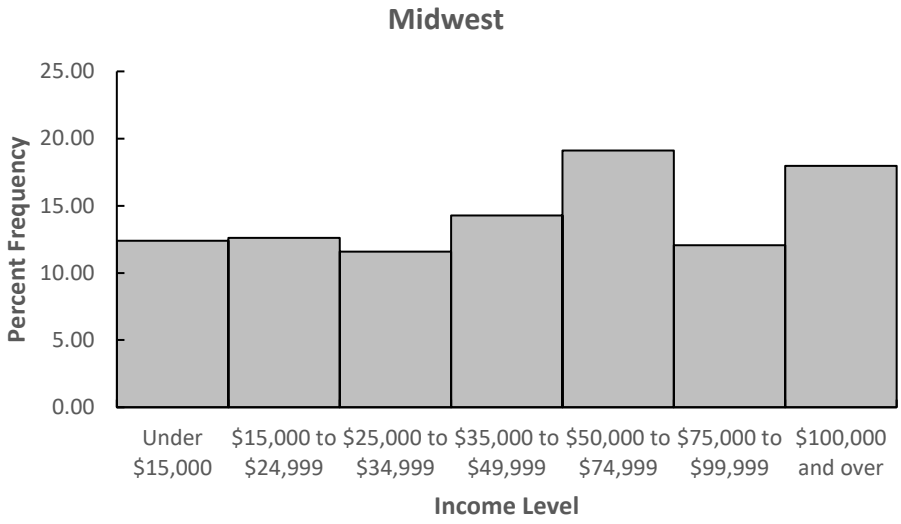
Income Level	Percent Frequency
Under $15,000	11.84
$15,000 to $24,999	10.73
$25,000 to $34,999	10.15
$35,000 to $49,999	13.65
$50,000 to $74,999	18.44
$75,000 to $99,999	11.77
$100,000 and over	23.43
Total	100.00

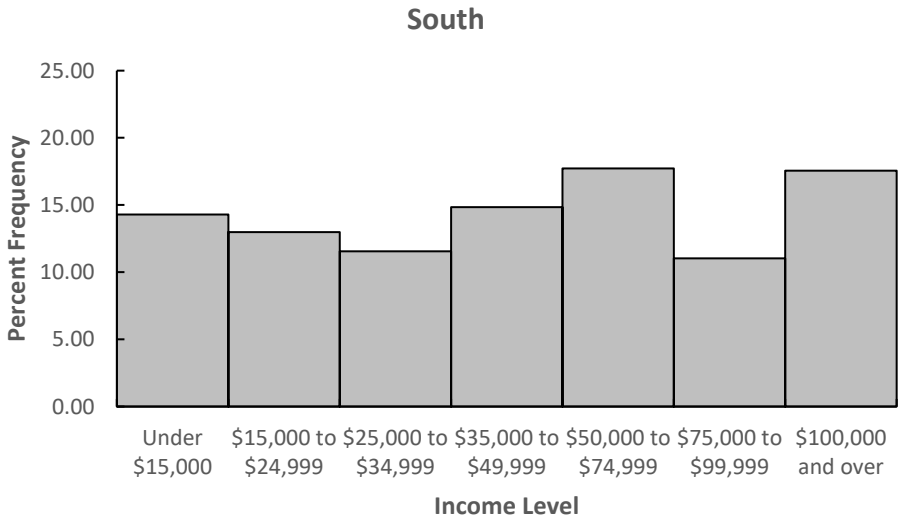
b. West: 18.44 + 11.77 + 23.43 = 53.64%

South: 17.73 + 11.04 + 17.57 = 46.34%

c.







The largest difference appears to be a higher percentage of household incomes of $100,000 and over for the Northeast and West regions.

d. Column percentages are shown below.

Region	Under $15,000	$15,000 to $24,999	$25,000 to $34,999	$35,000 to $49,999	$50,000 to $74,999	$75,000 to $99,999	$100,000 and over
Northeast	17.83	16.00	17.41	16.90	17.38	18.35	22.09
Midwest	21.35	23.72	23.50	22.68	23.71	23.49	19.96
South	40.68	40.34	38.75	39.00	36.33	35.53	32.25
West	20.13	19.94	20.34	21.42	22.58	22.63	25.70
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Each column is a percent frequency distribution of the region variable for one of the household income categories. For example, for an income level of $35,000 to $49,999 the percent frequency distribution for the region variable is as follows:

Region	Percent Frequency
Northeast	16.90
Midwest	22.68
South	39.00
West	21.42
Total	100.00

33. a.

	Brand Value ($ billions)						
Industry	0-10	10-20	20-30	30-40	40-50	50-60	Total
Automotive & Luxury	10	4	1				15
Consumer Packaged Goods	7	5					12
Financial Services	11	3					14
Other	14	10		2			26
Technology	7	4		1	1	2	15
Total	49	26	1	3	1	2	82

b.

Industry	Total
Automotive & Luxury	15
Consumer Packaged Goods	12
Financial Services	14
Other	26
Technology	15
Total	82

c.

Brand Value ($ billions)	**Frequency**
0-10	49
10-20	26
20-30	1
30-40	3
40-50	1
50-60	2
Total	82

d. The right margin shows the frequency distribution for the fund type variable and the bottom margin shows the frequency distribution for the brand value.

e. Higher brand values are associated with the technology brands. For instance, the crosstabulation shows that 4 of the 15 technology brands (approximately 27%) had a brand value of $30 billion or higher.

34. a.

	Brand Revenue ($ billions)						
Industry	0-25	25-50	50-75	75-100	100-125	125-150	Total
Automotive & Luxury	10	1	1		1	2	15
Consumer Packaged Goods	12						12
Financial Services	2	4	2	2	2	2	14
Other	13	5	3	2	2	1	26
Technology	4	4	4	1	2		15
Total	41	14	10	5	7	5	82

b.

Brand Revenue ($ billions)	**Frequency**
0-25	41
25-50	14
50-75	10
75-100	5
100-125	7
125-150	5
Total	82

c. Consumer packaged goods have the lowest brand revenues; each of the 12 consumer packaged goods brands in the sample data had a brand revenue of less than $25 billion. Approximately 57% of the financial services brands (8 out of 14) had a brand revenue of $50 billion or greater, and 47% of the technology brands (7 out of 15) had a brand revenue of at least $50 billion.

d.

	1-Yr Value Change (%)						
Industry	-60--41	-40--21	-20--1	0-19	20-39	40-60	Total
Automotive & Luxury				11	4		15
Consumer Packaged Goods			2	10			12
Financial Services		1	6	7			14
Other			2	20	4		26
Technology	1	3	4	4	2	1	15
Total	1	4	14	52	10	1	82

e.

1-Yr Value Change (%)	Frequency
-60--41	1
-40--21	4
-20--1	14
0-19	52
20-39	10
40-60	1
Total	82

f. The automotive & luxury brands all had a positive 1-year value change (%). The technology brands had the greatest variability.

35. a.

	Hwy MPG						
Size	15-19	20-24	25-29	30-34	35-39	40-44	Total
Compact	3	4	17	22	5	5	56
Large	2	10	7	3	2		24
Midsize	3	4	30	20	9	3	69
Total	8	18	54	45	16	8	149

b. Midsize and Compact seem to be more fuel efficient than Large.

c.

	City MPG						
Drive	10-14	15-19	20-24	25-29	30-34	40-44	Total
A	7	18	3				28
F		17	49	19	2	3	90
R	10	20		1			31
Total	17	55	52	20	2	3	149

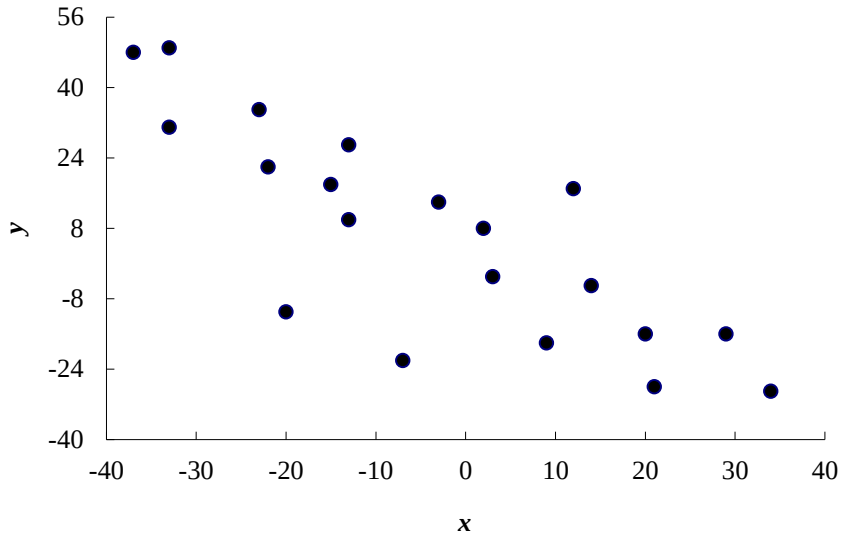
d. Higher fuel efficiencies are associated with front wheel drive cars.

e.

	City MPG						
Fuel Type	15-19	20-24	25-29	30-34	35-39	40-44	Total
P	8	16	20	12			56
R		2	34	33	16	8	93
Total	8	18	54	45	16	8	149

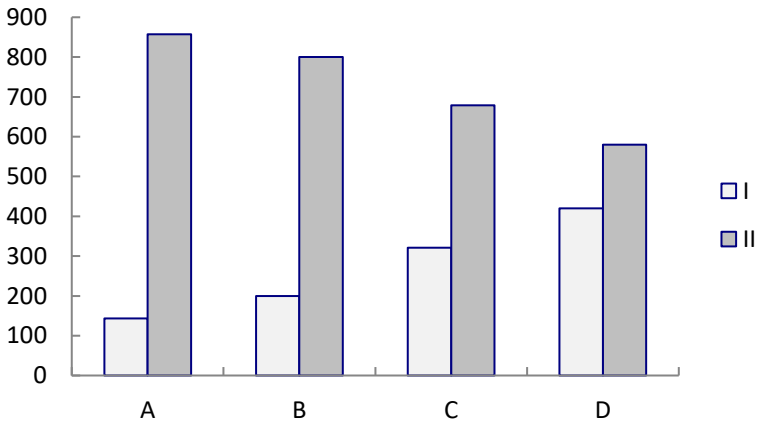
f. Higher fuel efficiencies are associated with cars that use regular gas.

36. a.



b. There is a negative relationship between *x* and *y; y* decreases as *x* increases.

37. a.

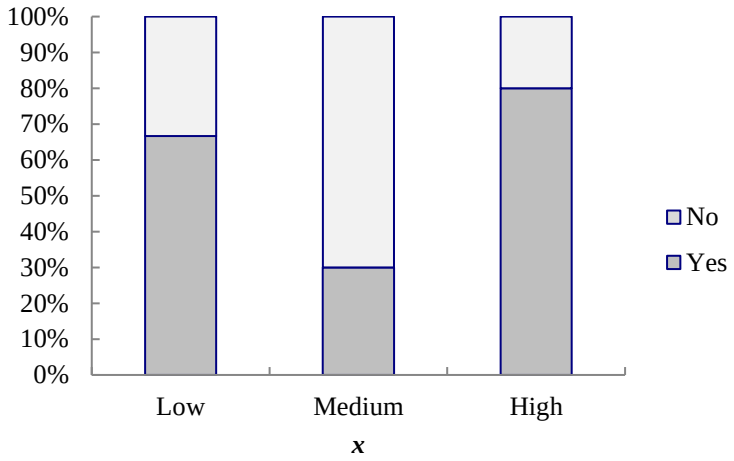


b. As X goes from A to D the frequency for I increases and the frequency of II decreases.

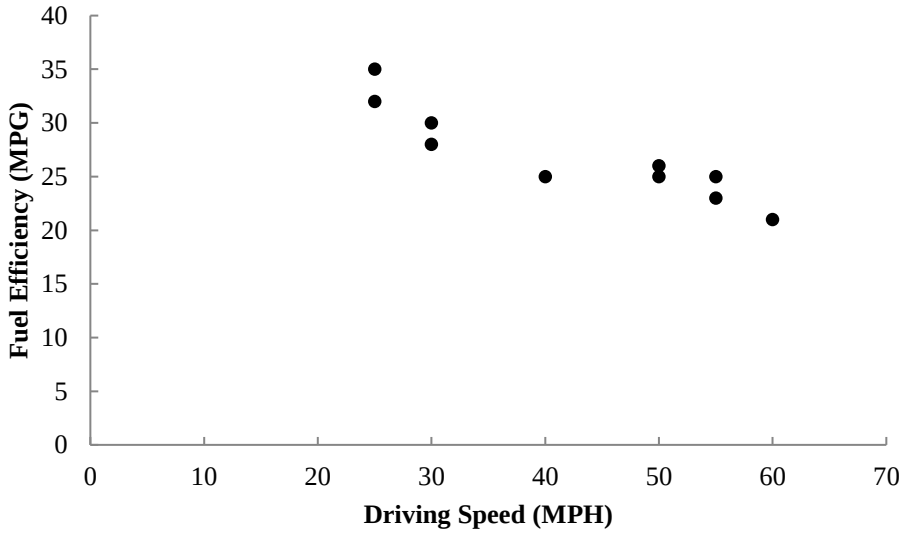
38. a.

		y		
		Yes	No	
x	Low	66.667	33.333	100
	Medium	30.000	70.000	100
	High	80.000	20.000	100

b.

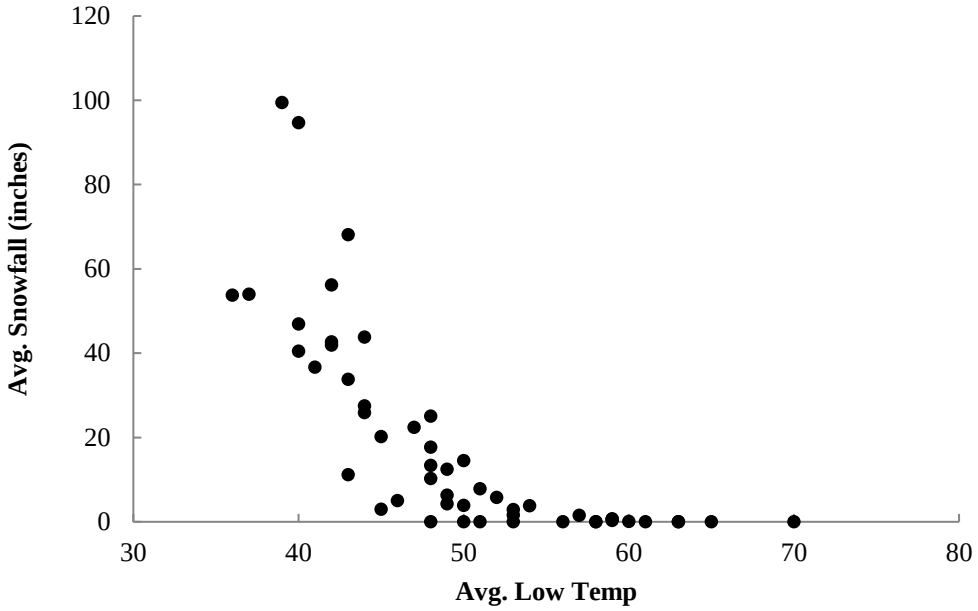


39. a.



b. For midsized cars, lower driving speeds seem to yield higher miles per gallon.

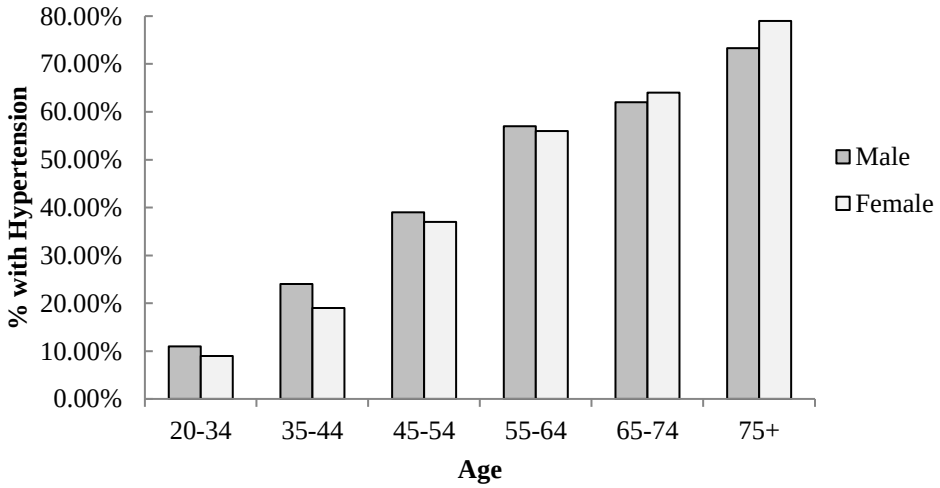
40. a.



b. Colder average low temperature seems to lead to higher amounts of snowfall.

c. Two cities have an average snowfall of nearly 100 inches of snowfall: Buffalo, N.Y and Rochester, NY. Both are located near large lakes in New York.

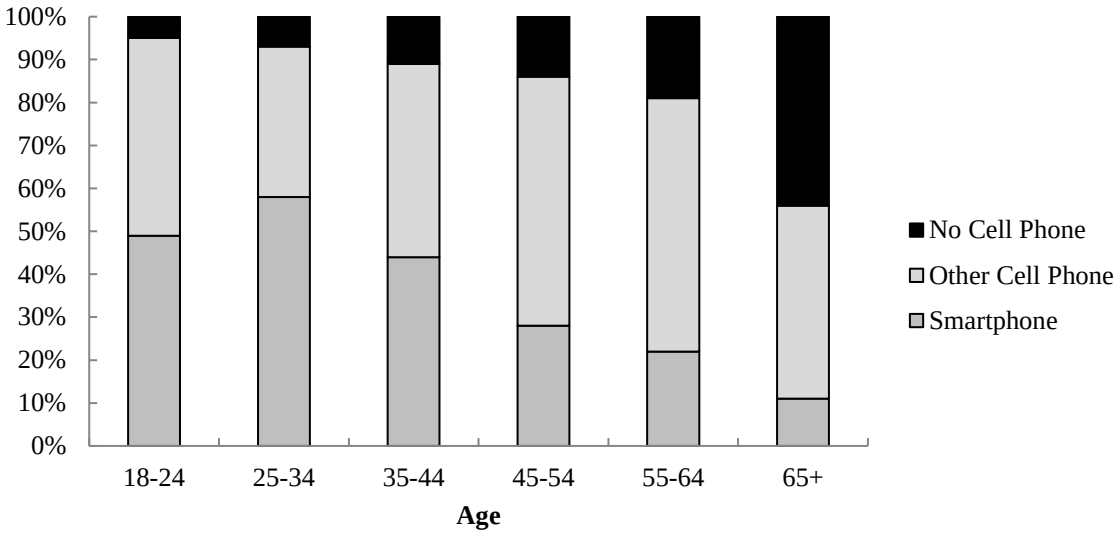
41. a.



b. The percentage of people with hypertension increases with age.

c. For ages earlier than 65, the percentage of males with hypertension is higher than that for females. After age 65, the percentage of females with hypertension is higher than that for males.

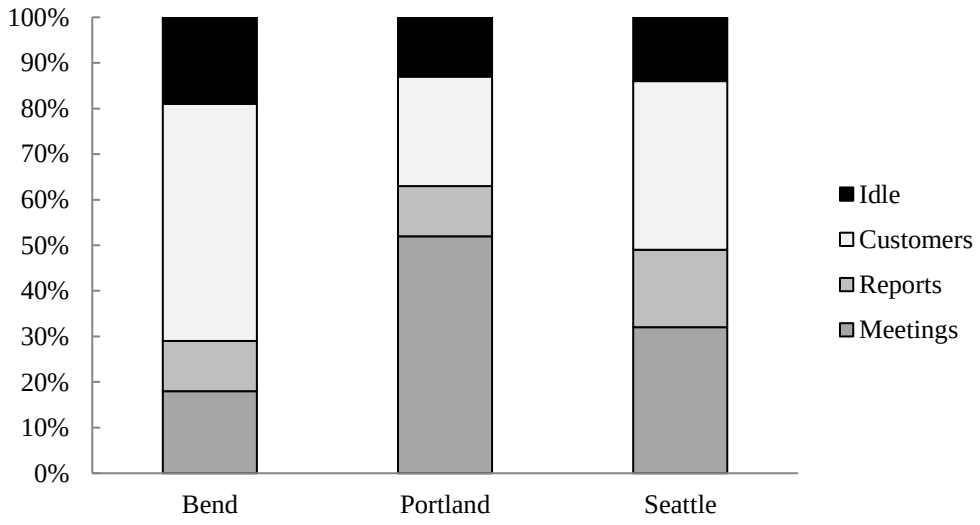
42. a.



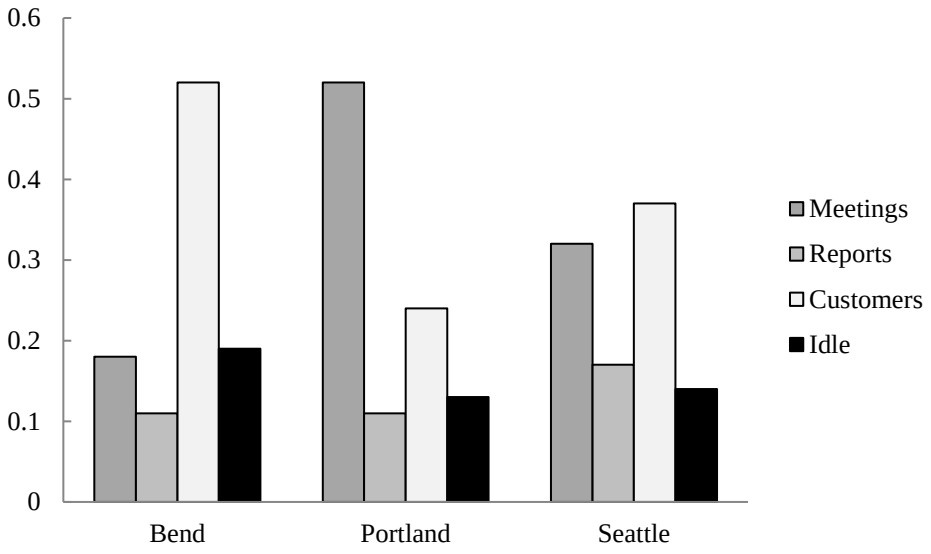
b. After an increase in age 25-34, smartphone ownership decreases as age increases. The percentage of people with no cell phone increases with age. There is less variation across age groups in the percentage who own other cell phones.

c. Unless a newer device replaces the smartphone, we would expect smartphone ownership would become less sensitive to age. This would be true because current users will become older and because the device will become to be seen more as a necessity than a luxury.

43. a.



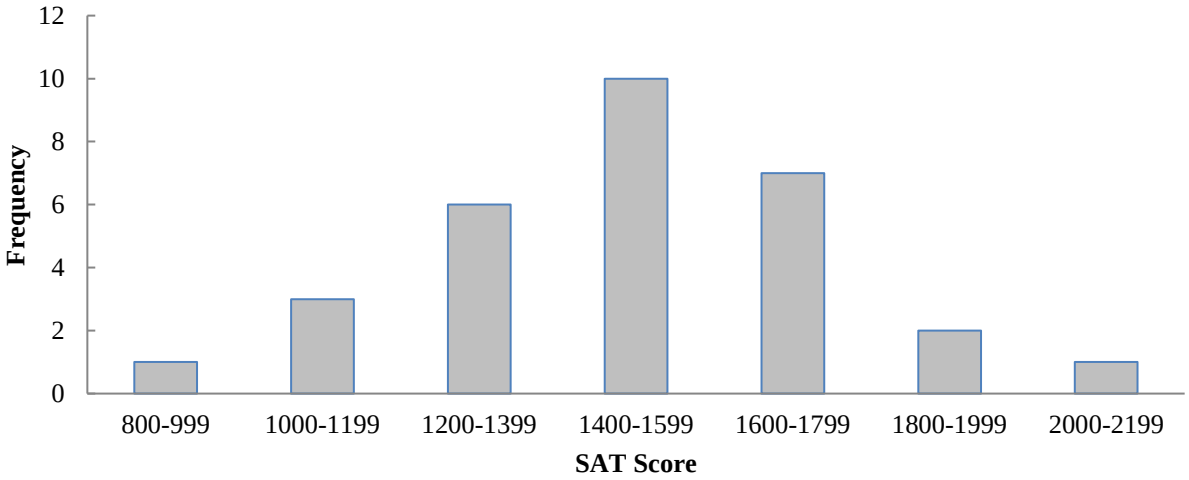
b.



c. The stacked bar chart seems simpler than the side-by-side bar chart and more easily conveys the differences in store managers' use of time.

44. a.

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

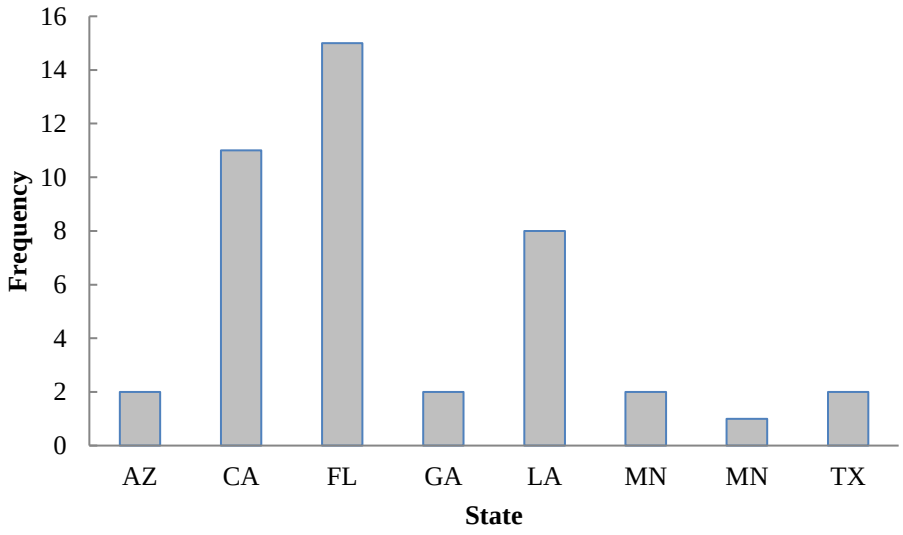


b. The distribution if nearly symmetrical. It could be approximated by a bell-shaped curve.

c. 10 of 30 or 33% of the scores are between 1400 and 1599. The average SAT score looks to be a little over 1500. Scores below 800 or above 2200 are unusual.

45. a.

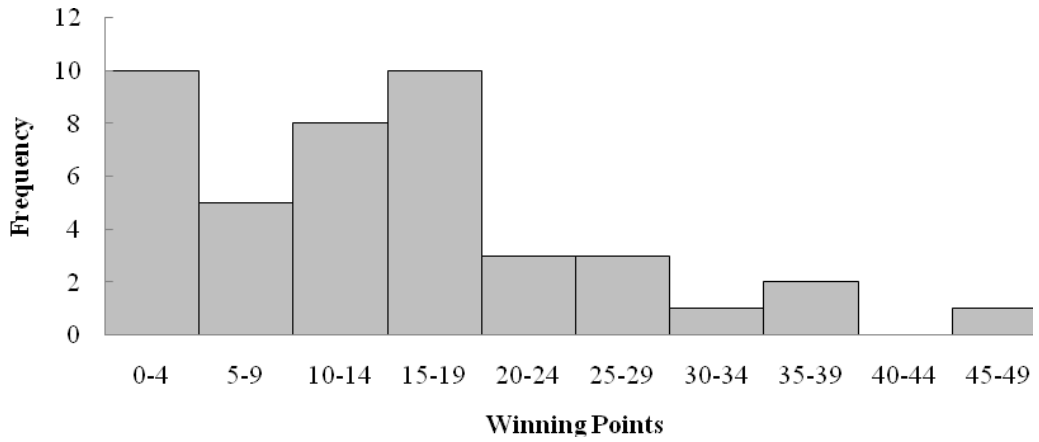
State	Frequency
Arizona	2
California	11
Florida	15
Georgia	2
Louisiana	8
Michigan	2
Minnesota	1
Texas	2
Total	43



b. Florida has had the most Super Bowl with 15, or 15/43(100) = 35%. Florida and California have been the states with the most Super Bowls. A total of 15 + 11 = 26, or 26/43(100) = 60%. Only 3 Super Bowls, or 3/43(100) = 7%, have been played in the cold weather states of Michigan and Minnesota.

c.

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



d. The most frequent winning points have been 0 to 4 points and 15 to 19 points. Both occurred in 10 Super Bowls. There were 10 close games with a margin of victory less than 5 points, 10/43(100) = 23% of the Super Bowls. There have also be 10 games, 23%, with a margin of victory more than 20 points.

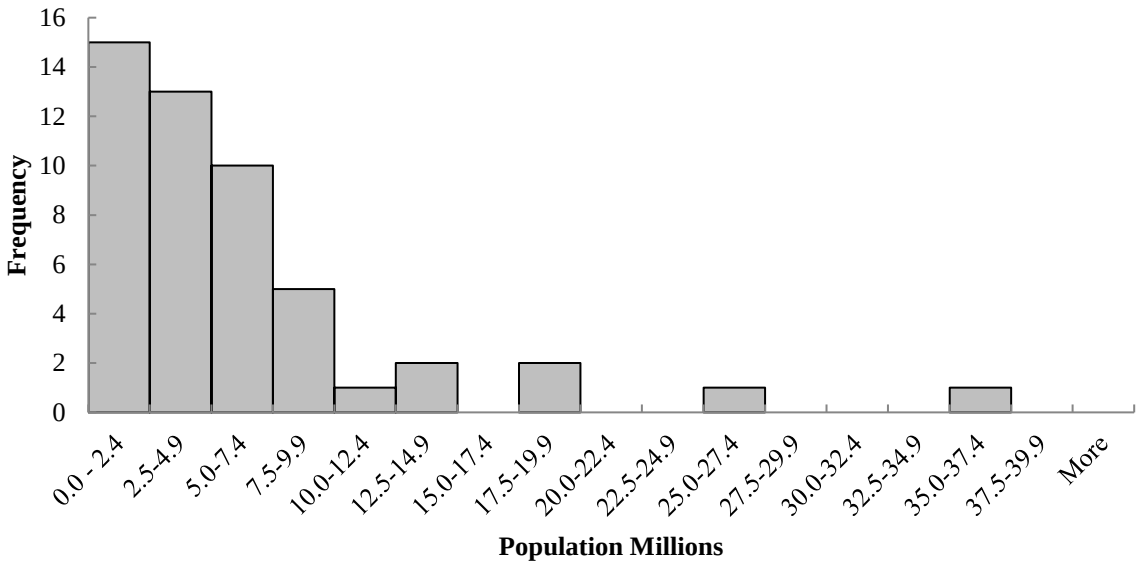
e. The closest games was the 25th Super Bowl with a 1 point margin. It was played in Florida. The largest margin of victory occurred one year earlier in the 24th Super Bowl. It had a 45 point margin and was played in Louisiana. More detailed information not available from the text information.

25th Super Bowl: 1991 New York Giants 20 Buffalo Bills 19, Tampa Stadium, Tampa, FL
24th Super Bowl: 1990 San Francisco 49ers 55 Denver Broncos 10, Superdome, New Orleans, LA

Note: The data set SuperBowl contains a list of the teams and the final scores of the 43 Super Bowls. This data set can be used in Chapter 2 and Chapter 3 to provide interesting data summaries about the points scored by the winning team and the points scored by the losing team in the Super Bowl. For example, using the median scores, the median Super Bowl score was 28 to 13.

46. a.

Population in Millions	Frequency	% Frequency
0.0 - 2.4	15	30.0%
2.5-4.9	13	26.0%
5.0-7.4	10	20.0%
7.5-9.9	5	10.0%
10.0-12.4	1	2.0%
12.5-14.9	2	4.0%
15.0-17.4	0	0.0%
17.5-19.9	2	4.0%
20.0-22.4	0	0.0%
22.5-24.9	0	0.0%
25.0-27.4	1	2.0%
27.5-29.9	0	0.0%
30.0-32.4	0	0.0%
32.5-34.9	0	0.0%
35.0-37.4	1	2.0%
37.5-39.9	0	0.0%
More	0	0.0%



b. The distribution is skewed to the right.

c. 15 states (30%) have a population less than 2.5 million. Over half of the states have population less than 5 million (28 states – 56%). Only seven states have a population greater than 10 million (California, Florida, Illinois, New York, Ohio, Pennsylvania and Texas). The largest state is California (37.3 million) and the smallest states are Vermont and Wyoming (600 thousand).

47. a.

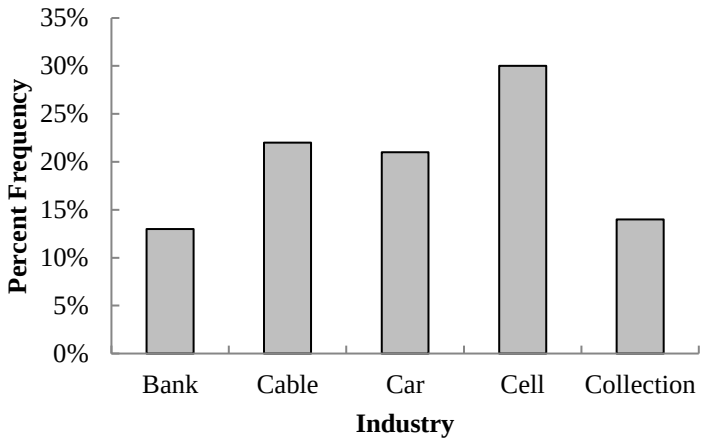
```
 1   8
 2   014
 3   18
 4   007899
 5   012444578
 6   00139
 7   237888
 8   011
 9   1
10   3
11   0289
12   9
13   01
14
15   46
16   68
17
18
19   2
20
21
22
23
24
25
26
27   2
```

b. The majority of the start-up companies in this set have less than $90 million in venture capital. Only 6 of the 50 (12%) have more than $150 million.

48. a.

Industry	**Frequency**	**% Frequency**
Bank	26	13%
Cable	44	22%
Car	42	21%
Cell	60	30%
Collection	28	14%
Total	200	100%

b.



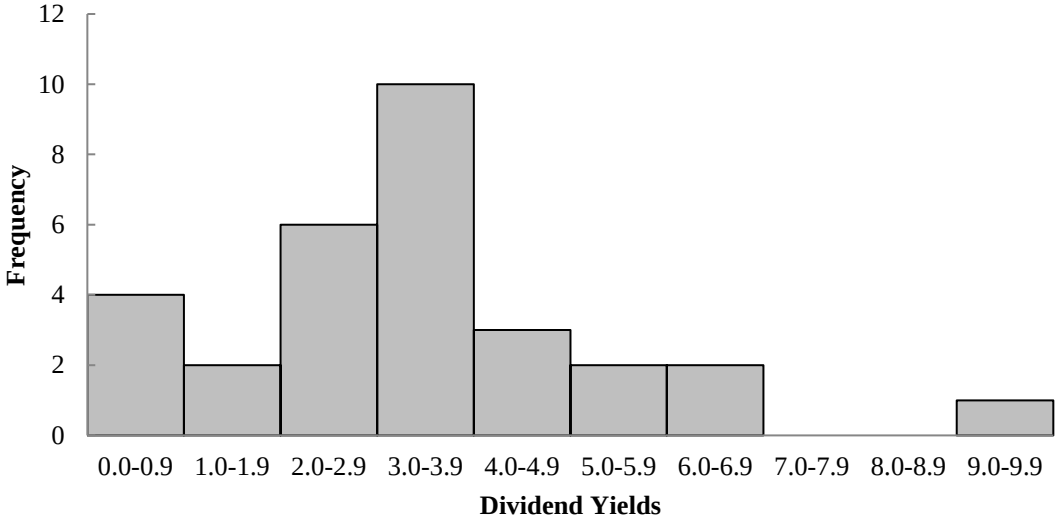
c. The cellular phone providers had the highest number of complaints.

d. The percentage frequency distribution shows that the two financial industries (banks and collection agencies) had about the same number of complaints. Also, new car dealers and cable and satellite television companies also had about the same number of complaints.

49. a.

Yield%	Frequency	Percent Frequency
0.0-0.9	4	13.3
1.0-1.9	2	6.7
2.0-2.9	6	20.0
3.0-3.9	10	33.3
4.0-4.9	3	10.0
5.0-5.9	2	6.7
6.0-6.9	2	6.7
7.0-7.9	0	0.0
8.0-8.9	0	0.0
9.0-9.9	1	3.3
Total	30	100.0

b.



c. The distribution is skewed to the right.

d. Dividend yield ranges from 0% to over 9%. The most frequent range is 3.0% to 3.9%. Average dividend yields looks to be between 3% and 4%. Over 50% of the companies (16) pay from 2.0 % to 3.9%. Five companies (AT&T, DuPont, General Electric, Merck, and Verizon) pay 5.0% or more. Four companies (Bank of America, Cisco Systems, Hewlett-Packard, and J.P. Morgan Chase) pay less than 1%.

e. General Electric had an unusually high dividend yield of 9.2%. 500 shares at $14 per share is an investment of 500($14) = $7,000. A 9.2% dividend yield provides .092(7,000) = $644 of dividend income per year.

50. a.

Level of Education	**Percent Frequency**
High School graduate	32,773/65,644(100) = 49.93
Bachelor's degree	22,131/65,644(100) = 33.71
Master's degree	9003/65,644(100) = 13.71
Doctoral degree	1737/65,644(100) = 2.65
Total	100.00

13.71 + 2.65 = 16.36% of heads of households have a master's or doctoral degree.

b.

Household Income	**Percent Frequency**
Under $25,000	13,128/65,644(100) = 20.00
$25,000 to $49,999	15,499/65,644(100) = 23.61
$50,000 to $99,999	20,548/65,644(100) = 31.30
$100,000 and over	16,469/65,644(100) = 25.09
Total	100.00

31.30 + 25.09 = 56.39% of households have an income of $50,000 or more.

c.

	Household Income			
Level of Education	Under $25,000	$25,000 to $49,999	$50,000 to $99,999	$100,000 and over
High School graduate	75.26	64.33	45.95	21.14
Bachelor's degree	18.92	26.87	37.31	47.46
Master's degree	5.22	7.77	14.69	24.86
Doctoral degree	0.60	1.03	2.05	6.53
Total	100.00	100.00	100.00	100.00

There is a large difference between the level of education for households with an income of under $25,000 and households with an income of $100,000 or more. For instance, 75.26% of households with an income of under $25,000 are households in which the head of the household is a high school graduate. But, only 21.14% of households with an income level of $100,000 or more are households in which the head of the household is a high school graduate. It is interesting to note, however, that 45.95% of households with an income of $50,000 to $99,999 are households in which the head of the household his a high school graduate.

51. a. The batting averages for the junior and senior years for each player are as follows:

Junior year:

Allison Fealey	15/40 = .375
Emily Janson	70/200 = .350

Senior year:

Allison Fealey	75/250 = .300
Emily Janson	35/120 = .292

Because Allison Fealey had the higher batting average in both her junior year and senior year, Allison Fealey should receive the scholarship offer.

b. The combined or aggregated two-year crosstabulation is as follows:

	Combined 2-Year Batting	
Outcome	**A. Fealey**	**E. Jansen**
Hit	90	105
No Hit	200	215
Total At Bats	290	320

Based on this crosstabulation, the batting average for each player is as follows:

Combined Junior/Senior Years

Allison Fealey	90/290 = .310
Emily Janson	105/320 = .328

Because Emily Janson has the higher batting average over the combined junior and senior years, Emily Janson should receive the scholarship offer.

c. The recommendations in parts (a) and (b) are not consistent. This is an example of Simpson's Paradox. It shows that in interpreting the results based upon separate or un-aggregated crosstabulations, the conclusion can be reversed when the crosstabulations are grouped or aggregated. When Simpson's Paradox is present, the decision maker will have to decide whether the un-aggregated or the aggregated form of the crosstabulation is the most helpful in identifying the desired conclusion. Note: The authors prefer the recommendation to offer the scholarship to Emily Janson because it is based upon the aggregated performance for both players over a larger number of at-bats. But this is a judgment or personal preference decision. Others may prefer the conclusion based on using the un-aggregated approach in part (a).

52 a.

	Size of Company			
Job Growth (%)	**Small**	**Midsized**	**Large**	**Total**
-10-0	4	6	2	12
0-10	18	13	29	60
10-20	7	2	4	13
20-30	3	3	2	8
30-40	0	3	1	4
60-70	0	1	0	1
Total	32	28	38	98

b. Frequency distribution for growth rate.

Job Growth (%)	**Total**
-10-0	12
0-10	60
10-20	13
20-30	8
30-40	4
60-70	1
Total	98

Frequency distribution for size of company.

Size	**Total**
Small	32
Medium	28
Large	38
Total	98

c. Crosstabulation showing column percentages.

	Size of Company		
Job Growth (%)	**Small**	**Midsized**	**Large**
-10-0	13	21	5
0-10	56	46	76
10-20	22	7	11
20-30	9	11	5
30-40	0	11	3
60-70	0	4	0
Total	100	100	100

d. Crosstabulation showing row percentages.

	Size of Company			
Job Growth (%)	**Small**	**Midsized**	**Large**	**Total**
-10-0	33	50	17	100
0-10	30	22	48	100
10-20	54	15	31	100
20-30	38	38	25	100
30-40	0	75	25	100
60-70	0	4	0	100

e. 12 companies had a negative job growth: 13% were small companies; 21% were midsized companies; and 5% were large companies. So, in terms of avoiding negative job growth, large companies were better off than small and midsized companies. But, although 95% of the large companies had a positive job growth, the growth rate was below 10% for 76% of these companies. In terms of better job growth rates, midsized companies performed better than either small or large companies. For instance, 26% of the midsized companies had a job growth of at least 20% as compared to 9% for small companies and 8% for large companies.

53. a.

	Tution & Fees ($)								
Year Founded	1-5000	10001-15000	15001-20000	20001-25000	25001-30000	30001-35000	35001-40000	40001-45000	Total
1600-1649							1		1
1700-1749							2	1	3
1750-1799								4	4
1800-1849				1	3	3	6	8	21
1850-1899	1		2	2	13	14	13	4	49
1900-1949		1		2	3	4	8		18
1950-2000			2	4		1			7
Total	1	1	4	9	19	22	30	17	103

b.

	Tuition & Fees ($)								
Year Founded	1-5000	10001-15000	15001-20000	20001-25000	25001-30000	30001-35000	35001-40000	40001-45000	Grand Total
1600-1649							100.00		100
1700-1749							66.67	33.33	100
1750-1799								100.00	100
1800-1849				4.76	14.29	14.29	28.57	38.10	100
1850-1899	2.04		4.08	4.08	26.53	28.57	26.53	8.16	100
1900-1949		5.56		11.11	16.67	22.22	44.44		100
1950-2000			28.57	57.14		14.29			100

c. Colleges in this sample founded before 1800 tend to be expensive in terms of tuition.

54. a.

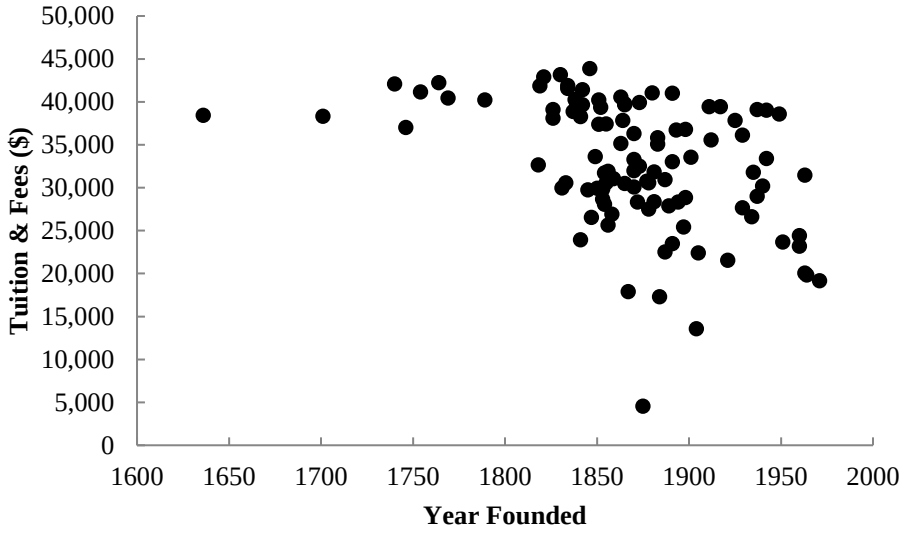
	% Graduate												
Year Founded	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90-95	95-100	Grand Total
1600-1649													1	1
1700-1749													3	3
1750-1799												1	3	4
1800-1849						1	2	4	2	3	4	3	2	21
1850-1899			1	2	4	3	11	5	9	6	3	4	1	49
1900-1949	1	1	1		1	3		3	2	4	1	1		18
1950-2000	1		1	3			2							7
Grand Total	2	1	3	5	5	7	15	12	13	13	8	9	10	103

b.

	% Graduate												
Year Founded	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80	80-85	85-90	90-95	95-100	Grand Total
1600-1649													100.00	100
1700-1749													100.00	100
1750-1799												25.00	75.00	100
1800-1849							4.76	9.52	19.05	9.52	14.29	19.05	14.29	9.52	100
1850-1899				2.04	4.08	8.16	6.12	22.45	10.20	18.37	12.24	6.12	8.16	2.04	100
1900-1949		5.56	5.56	5.56		5.56	16.67		16.67	11.11	22.22	5.56	5.56		100
1950-2000		14.29		14.29	42.86			28.57							100

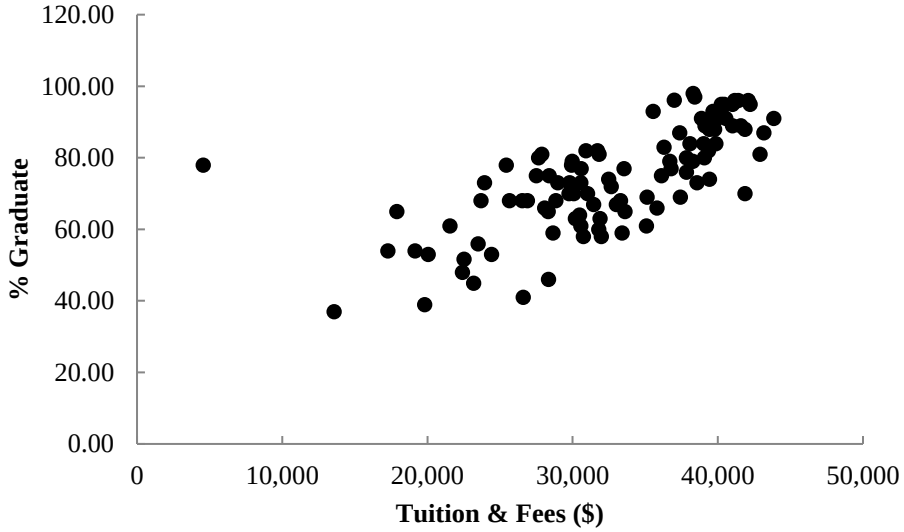
c. Older colleges and universities tend to have higher graduation rates.

55. a.



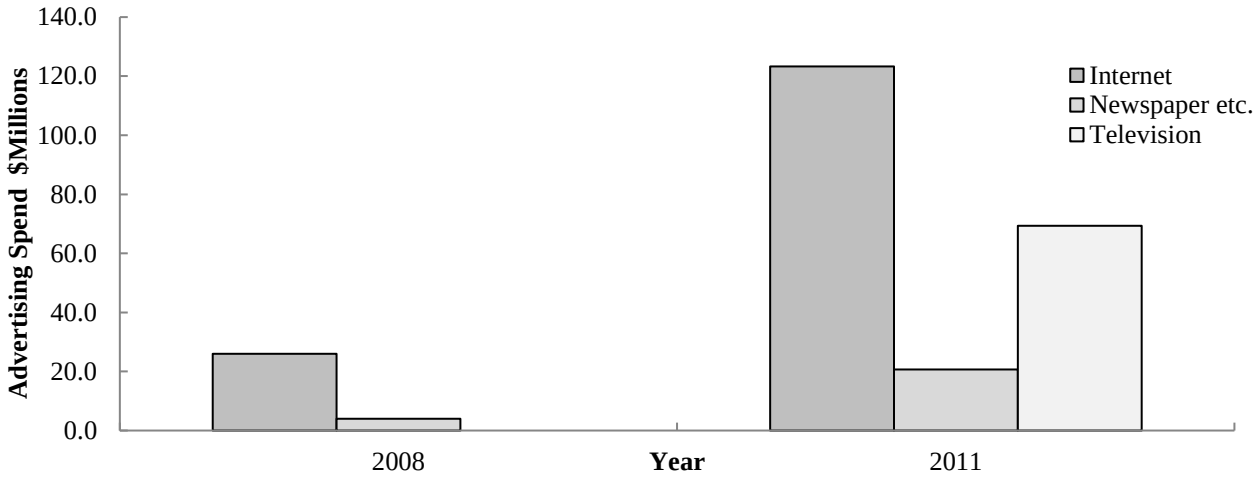
b. Older colleges and universities tend to be more expensive.

56. a.



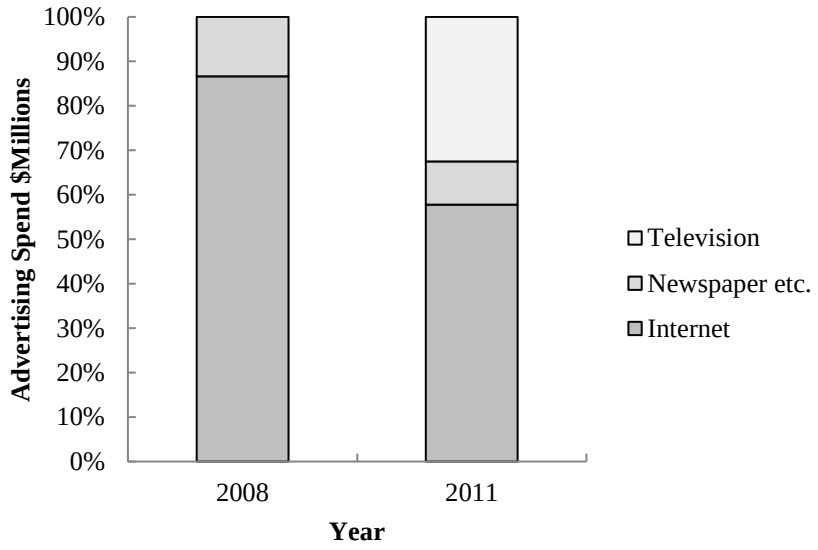
b. There appears to be a strong positive relationship between Tuition & Fees and % Graduation.

57. a.



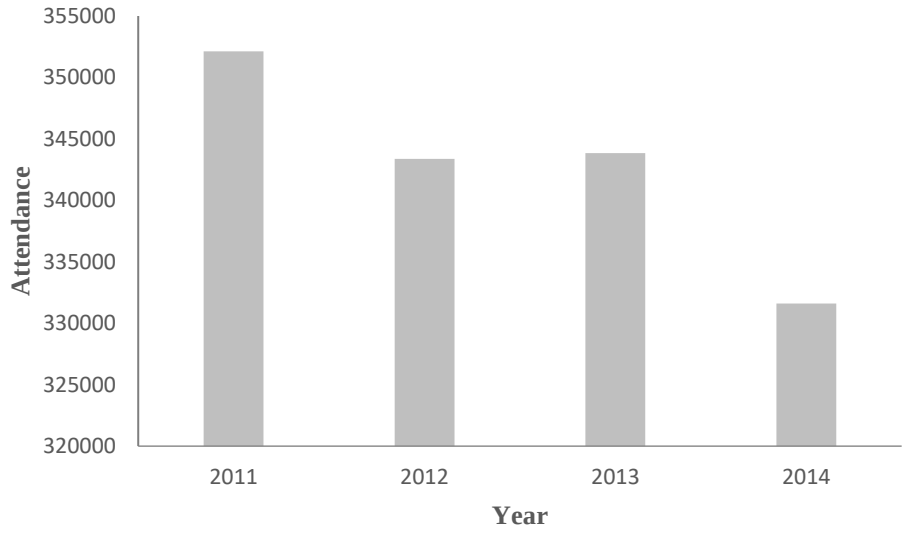
b.

	2008	2011
Internet	86.7%	57.8%
Newspaper etc.	13.3%	9.7%
Television	0.0%	32.5%
Total	100.0%	100.0%



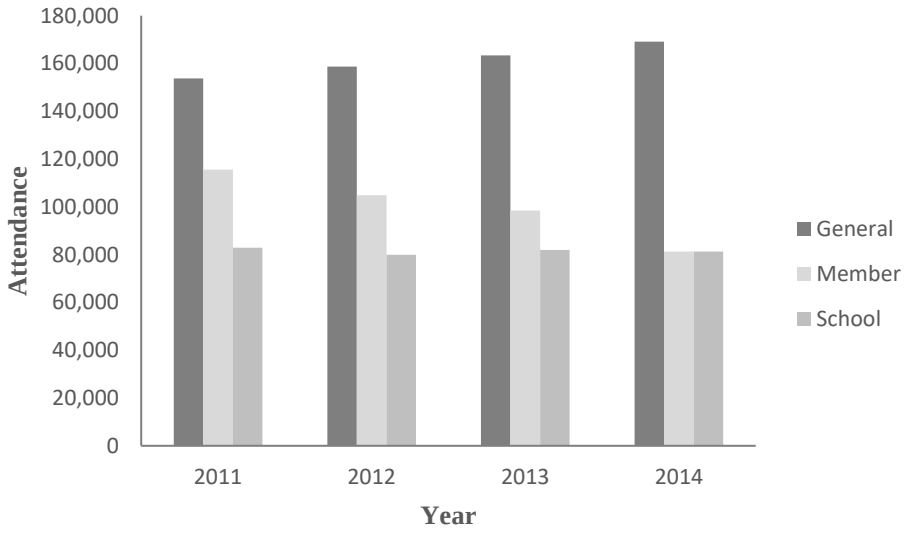
c. The graph is part a is more insightful because is shows the allocation of the budget across media, but also dramatic increase in the size of the budget.

58. a.



Zoo attendance appears to be dropping over time.

b.



c. General attendance is increasing, but not enough to offset the decrease in member attendance. School membership appears fairly stable.