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. $.20(200) = 40$

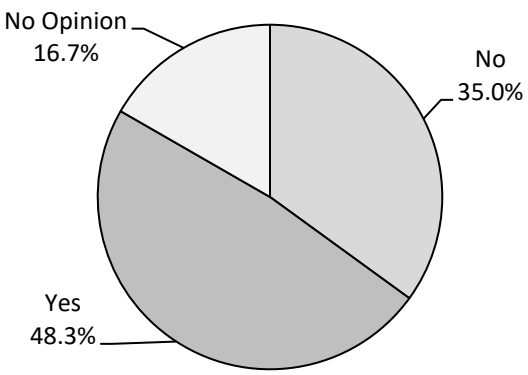
   c/d.

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

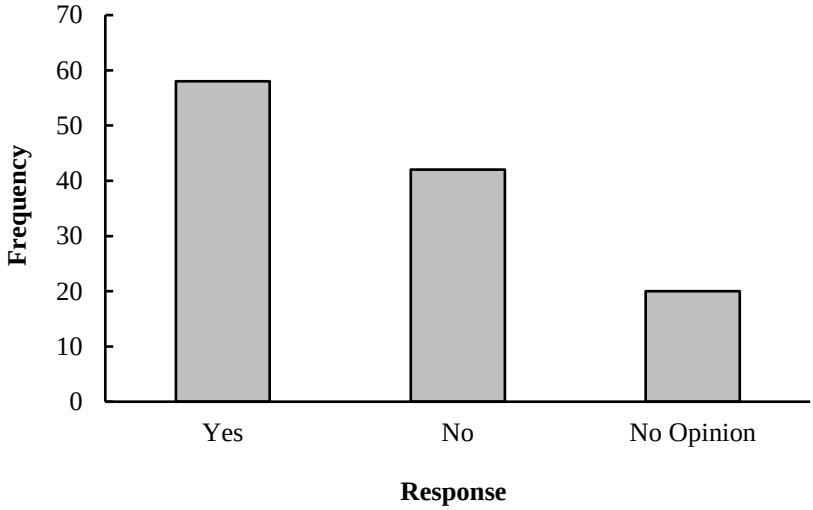
3. a. $360° \times 58/120 = 174°$

   b. $360° \times 42/120 = 126°$

c.



d.



4. a. These data are categorical.

b.

Website	Frequency	% Frequency
FB	8	16
GOOG	14	28

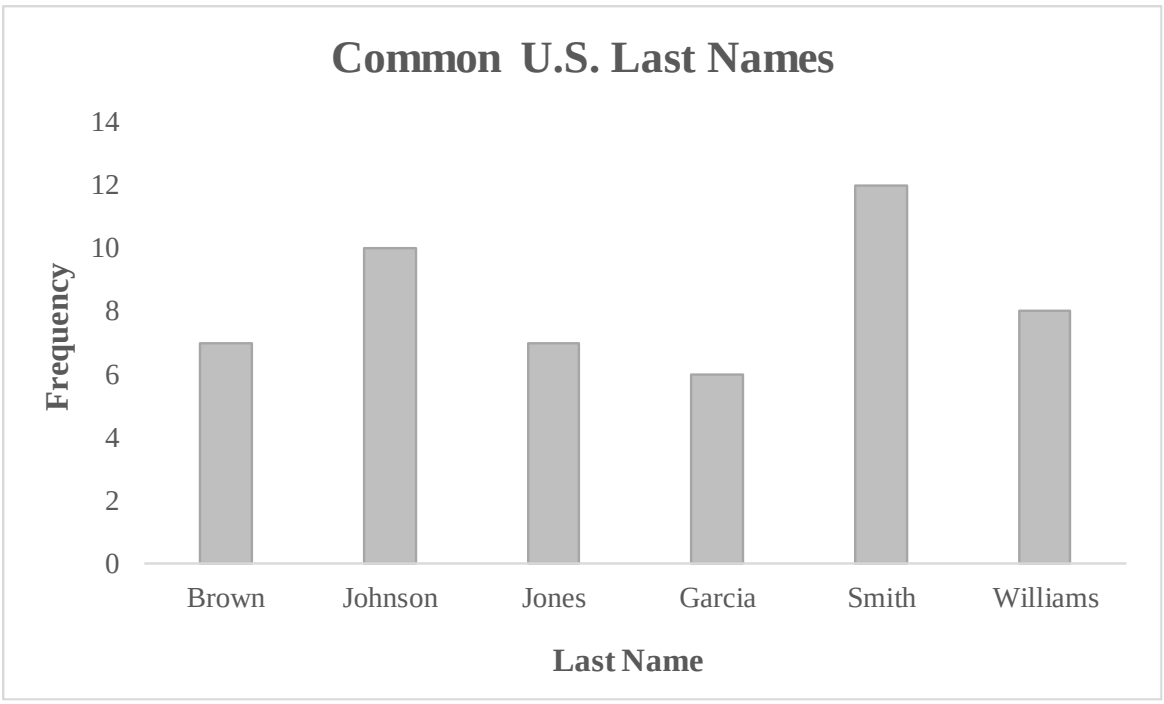
WIKI	9	18
YAH	13	26
YT	6	12
Total	50	100

c. The most frequently visited website is google.com (GOOG); the second is yahoo.com (YAH).

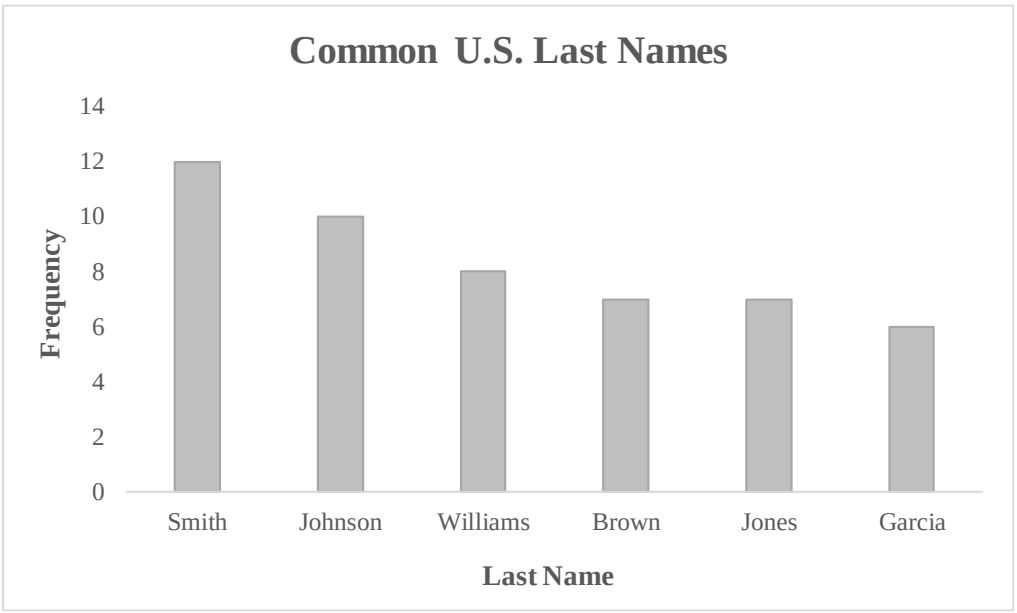
5. a.

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

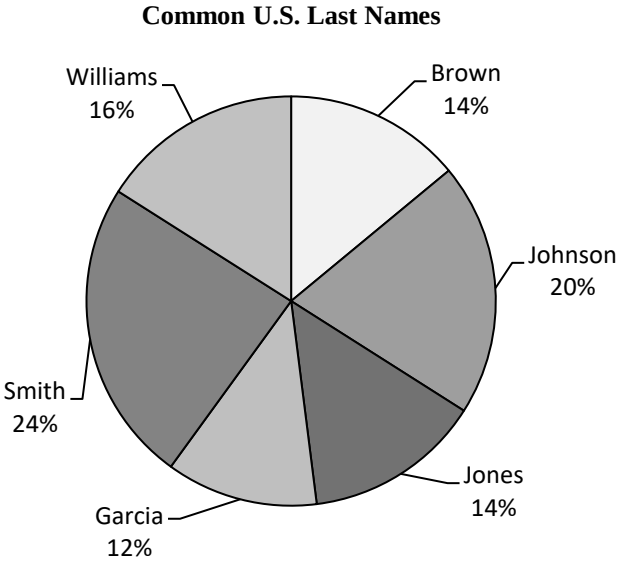
b.



c.



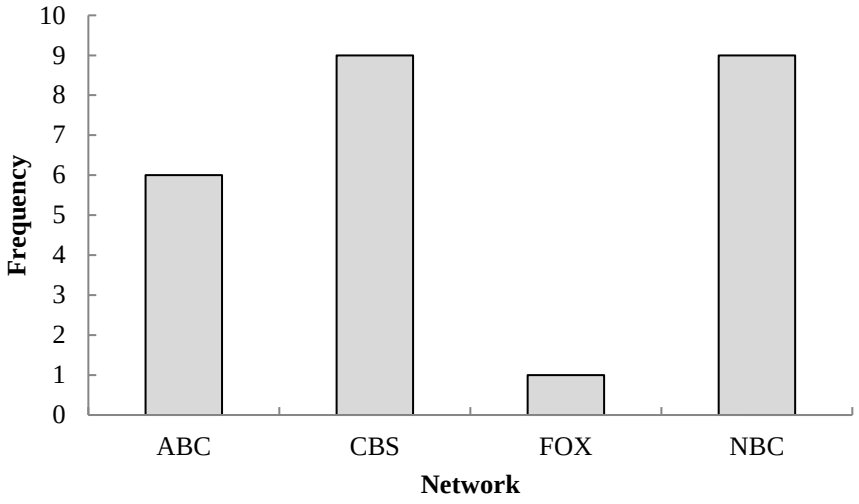
d.

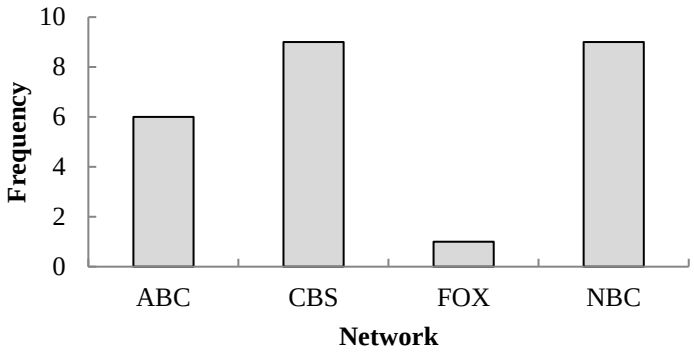


e. The three most common last names are Smith (24%), Johnson (20%), Williams (16%5). This is easily apparent from the sorted bar chart in c. Without the labeling of percentages, it is difficult to determine the most common names from the pie chart.

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 nine (36%) of the top 25. ABC is third with six (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, and 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. Although 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%)

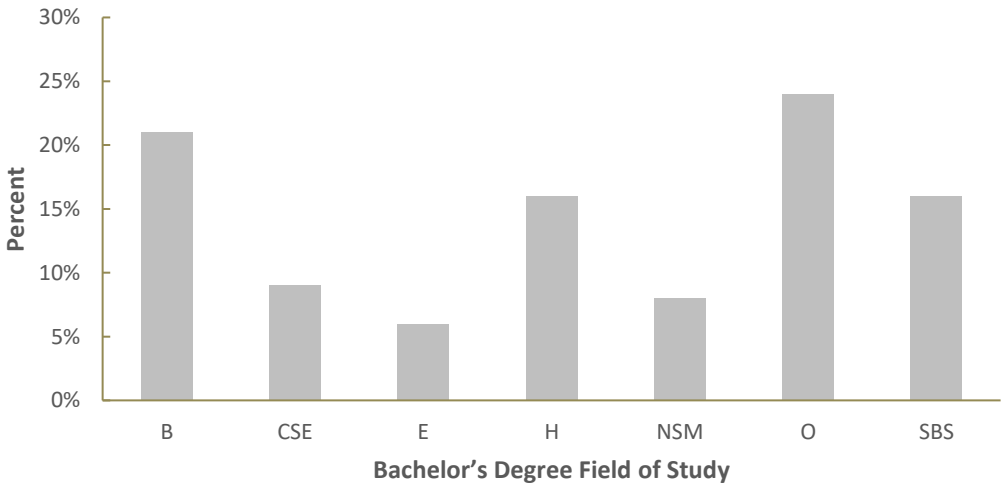
d. Right field (almost 13%)

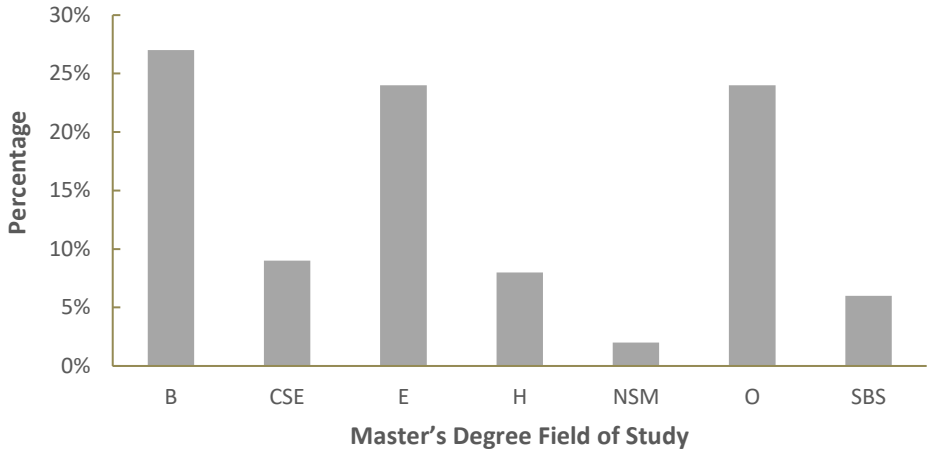
e. Infielders (16 or 29.1%) to outfielders (18 or 32.7%)

9. a.

	Bachelor's (%)	Master's (%)
B	21	27
CSE	9	9
E	6	24
H	16	8
NSM	8	2
SBS	16	6
O	24	24
Total	100	100

b.





c. The lowest percentage for a bachelor's is education (6%) and for master's in natural sciences and mathematics (2%).

d. The highest percentage for a bachelor's is other (24%) and for a master's in business (27%).

e.

	Bachelor's (%)	Master's (%)	Difference (%)
B	21	27	6
CSE	9	9	0
E	6	24	18
H	16	8	8
NSM	8	2	6
SBS	16	6	10
O	24	24	0

Education has the largest increase in percent: 18%.

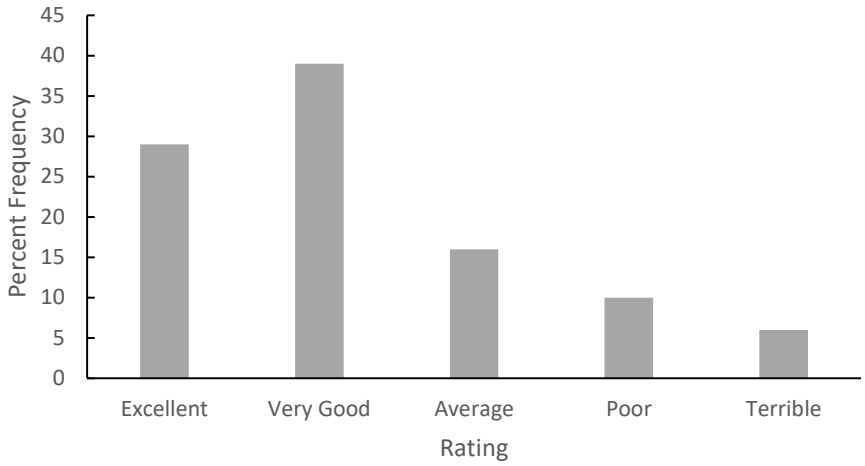
10. a.

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. At the Lakeview Lodge, 29% + 39% = 68% of the guests 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 the Timber Hotel follows:

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

At the Lakeview Lodge, 48% + 31% = 79% of the guests 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 the Timber Hotel provided somewhat better ratings than guests at the Lakeview Lodge.

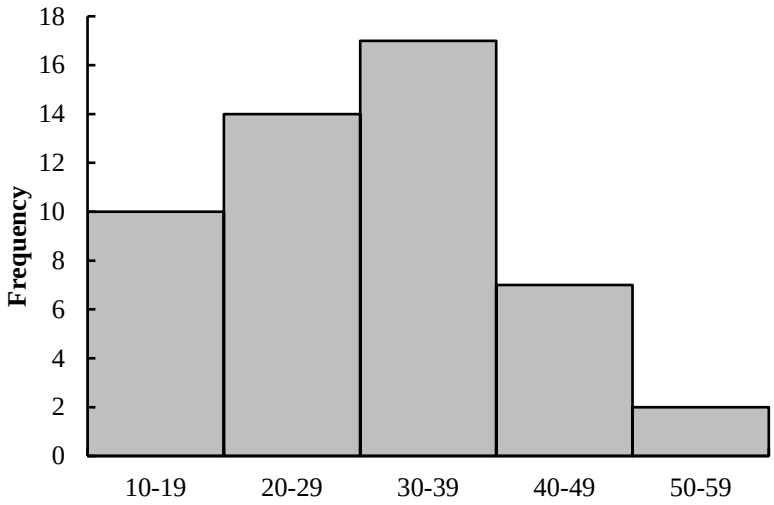
11.

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

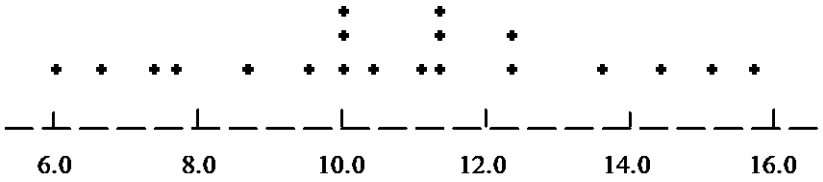
12.

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.

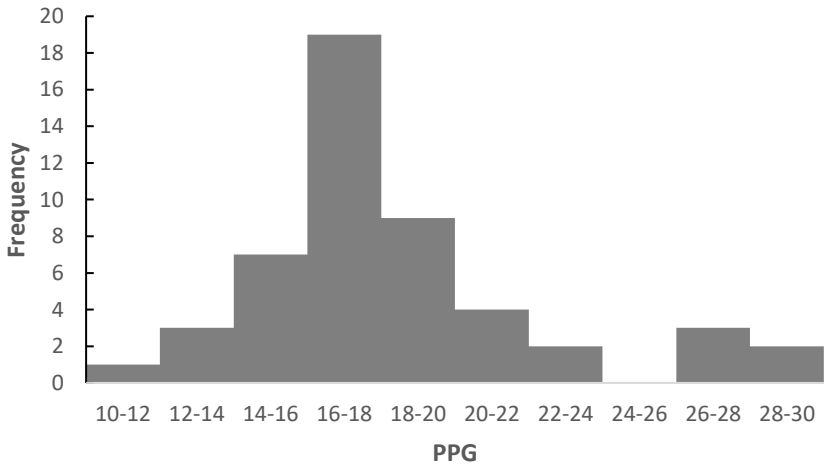
PPG	Relative Frequency
10–12	0.02
12–14	0.06
14–16	0.14
16–18	0.38
18–20	0.18
20–22	0.08
22–24	0.04
24–26	0.00
26–28	0.06
28–30	0.04
Total	1.00

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%

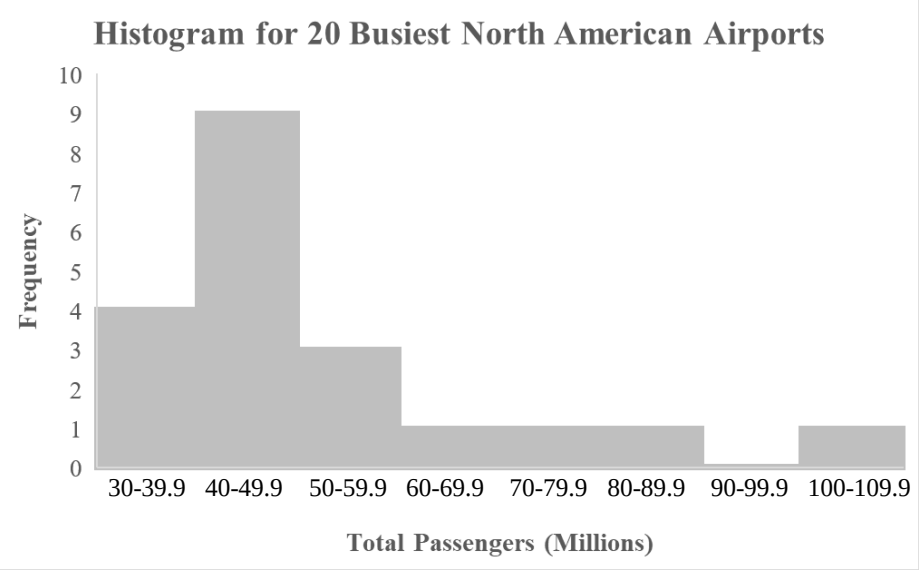
19. a. The busiest airport is Hartsfield-Jackson Atlanta (ATL) with 104.2 million total passengers.

The least busy airport is Detroit Metropolitan (DTW) with 34.4 million total passengers.

b.

Total Passengers (Millions)	Frequency
30–39.9	4
40–49.9	9
50–59.9	3
60–69.9	1
70–79.9	1
80–89.9	1
90–99.9	0
100–109.9	1

c.



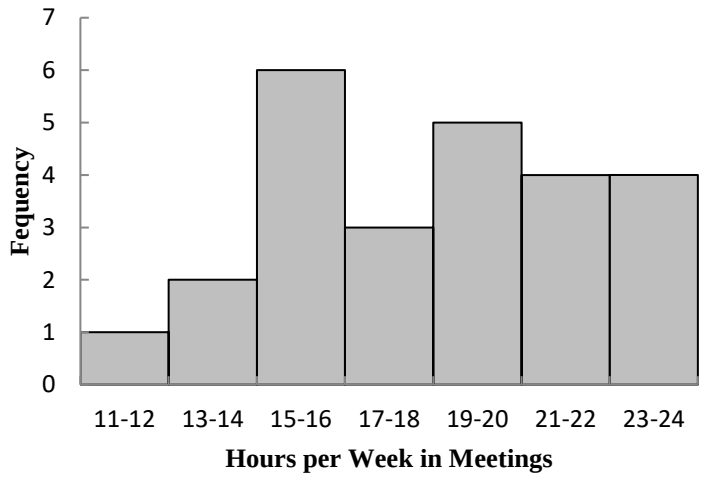
Most of the top 20 busiest North American airports service fewer than 60 million passengers. Only four of the 20 airports have more than 60 million passengers.

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.

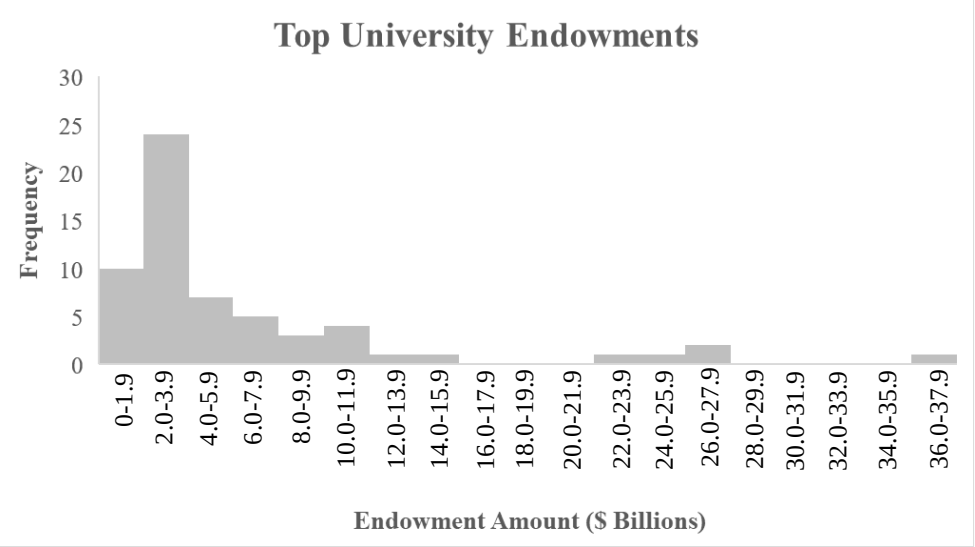
21. a/b/c/d.

Endowment Amount ($ Billions)	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0–1.9	10	0.17	10	0.17
2.0–3.9	24	0.40	34	0.57
4.0–5.9	7	0.12	41	0.68
6.0–7.9	5	0.08	46	0.77
8.0–9.9	3	0.05	49	0.82
10.0–11.9	4	0.07	53	0.88
12.0–13.9	1	0.02	54	0.90
14.0–15.9	1	0.02	55	0.92
16.0–17.9	0	0.00	55	0.92
18.0–19.9	0	0.00	55	0.92
20.0–21.9	0	0.00	55	0.92

22.0–23.9	1	0.02	56	0.93
24.0–25.9	1	0.02	57	0.95
26.0–27.9	2	0.03	59	0.98
28.0–29.9	0	0.00	59	0.98
30.0–31.9	0	0.00	59	0.98
32.0–33.9	0	0.00	59	0.98
34.0–35.9	0	0.00	59	0.98
36.0–37.9	1	0.02	60	1.00
Total	60	1.00		

e. Most universities (55) have endowments of less than $16 billion. Only five have endowments larger than $16 billion. We see that .92, or 92%, of the universities have endowments of less than $16 billion, and only .08, or 8%, of the universities have endowments larger than $16 billion.

f.



The histogram shows the distribution is skewed to the right with five university endowments in the $22 billion to $38 billion range.

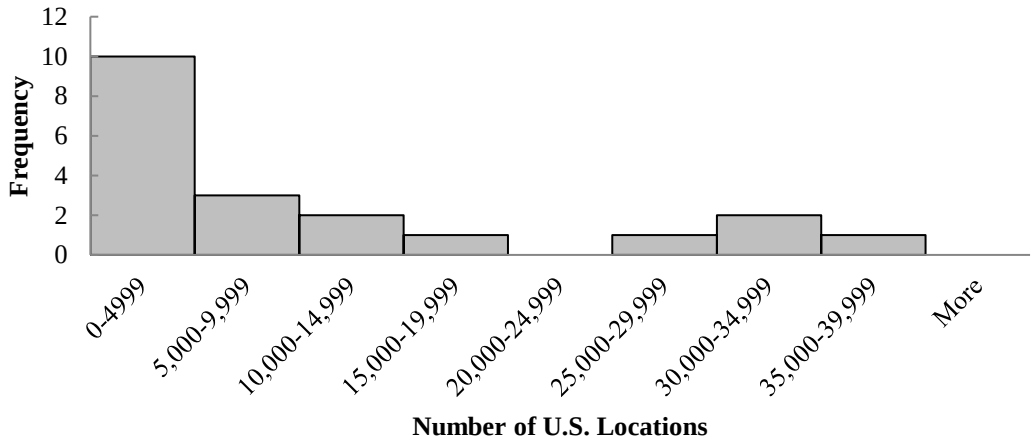
g. Harvard University has the largest endowment at $~~16~~ 36 billion. All other universities have endowments less than $28 billion. Most (92%) have endowments less than $16 billion.

22. a.

No. U.S. Locations	Frequency	Percent Frequency
0–4,999	10	50
5,000–9,999	3	15
10,000–14,999	2	10
15,000–19,999	1	5
20,000–24,999	0	0
25,000–29,999	1	5

30,000–34,999	2	10
35,000–39,999	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.

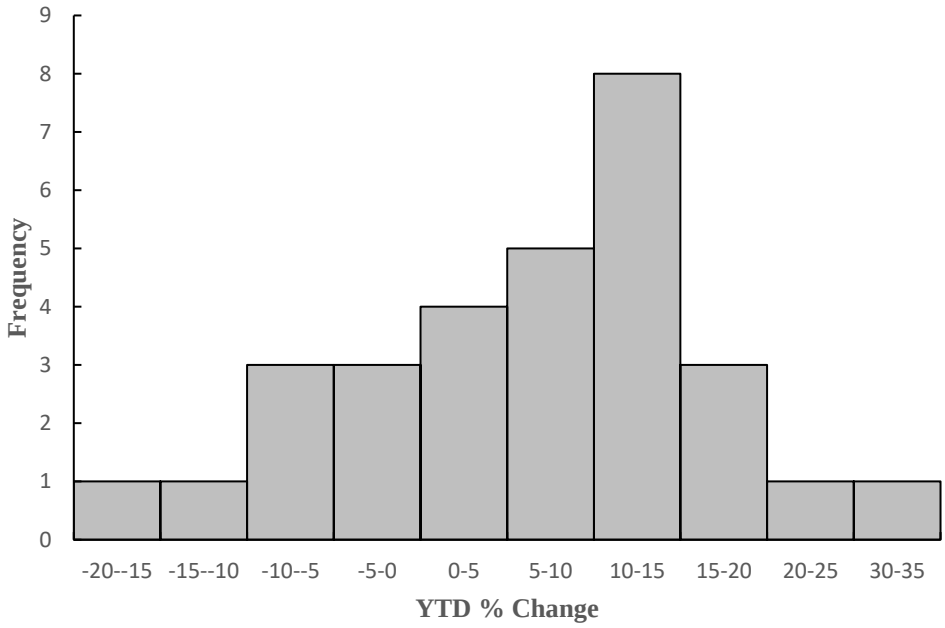
23. a. The highest positive YTD % change for Japan's Nikkei Index with a YTD % change of 31.4%.

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 left. 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.

Starting Median Salary

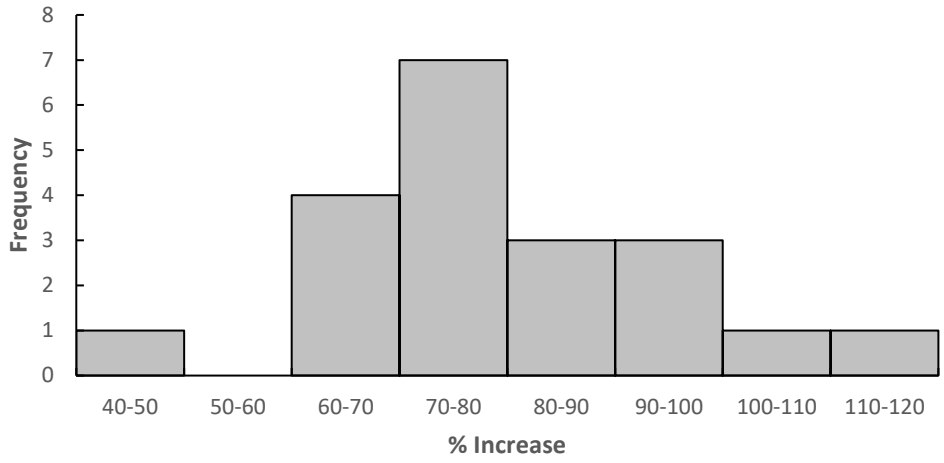
4	6 8
5	1 2 3 3 5 6 8 8
6	0 1 1 1 2 2
7	1 2 5

Mid-Career Median Salary

8	0 0 4
9	3 3 5 6 7
10	5 6 6
11	0 1 4 4 4
12	2 3 6

There is a wider spread in the mid-career median salaries than in the starting median salaries. Also, as expected, the mid-career median salaries are higher that the starting median salaries. The mid-career median salaries were mostly in the $93,000 to $114,000 range while the starting median salaries were mostly in the $51,000 to $62,000 range.

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 it 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 the result of new regulations that take into account driver safety, fan safety, the environmental impact, and fuel consumption during races.

31. a. The cross-tabulation of condition of the greens by gender follows.

	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 who say the greens are too fast: 40/100 = 40%. Of male golfers, 35/100 = 35% say 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. When the cross-tabulations are aggregated, however, 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 follow.

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 Higher	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.

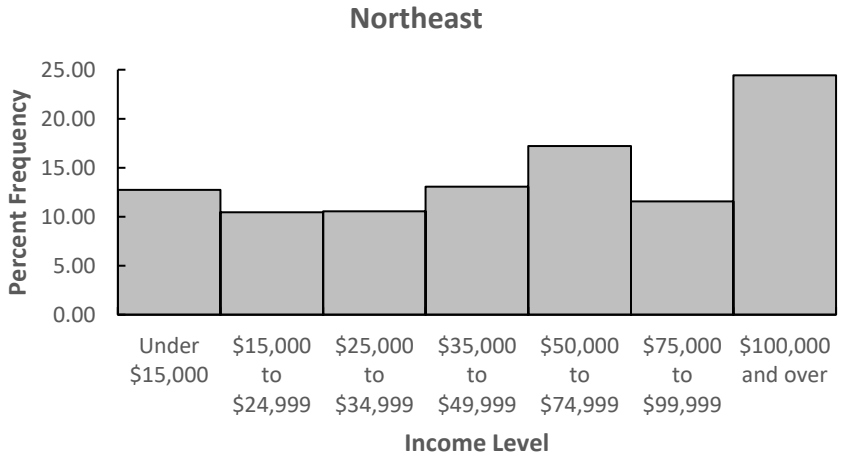
For example, the percent frequency distribution of the West region is as follows:

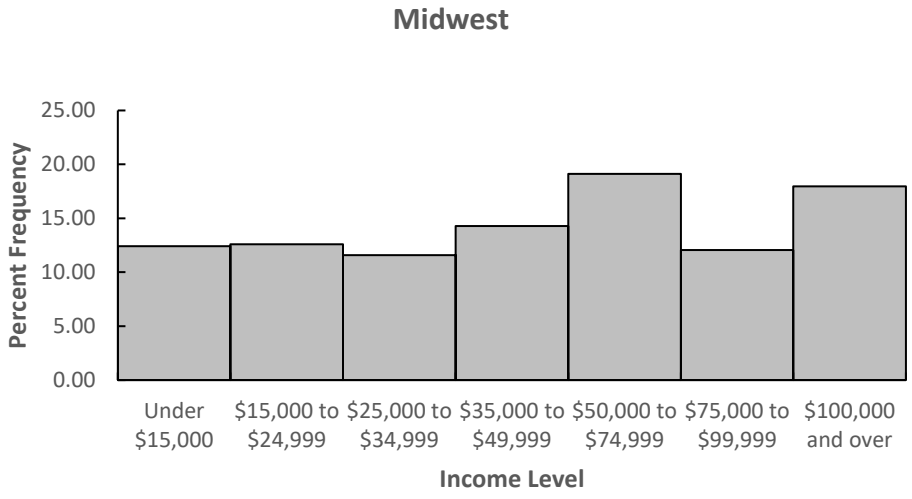
Income Level	Percent Frequency
Under $15,000	11.84
$15,000 to $24,999	10.73
$25,000 to $34,999	10.15
$35,000 to $49,999	13.65
$50,000 to $74,999	18.44
$75,000 to $99,999	11.77
$100,000 and over	23.43
Total	100.00

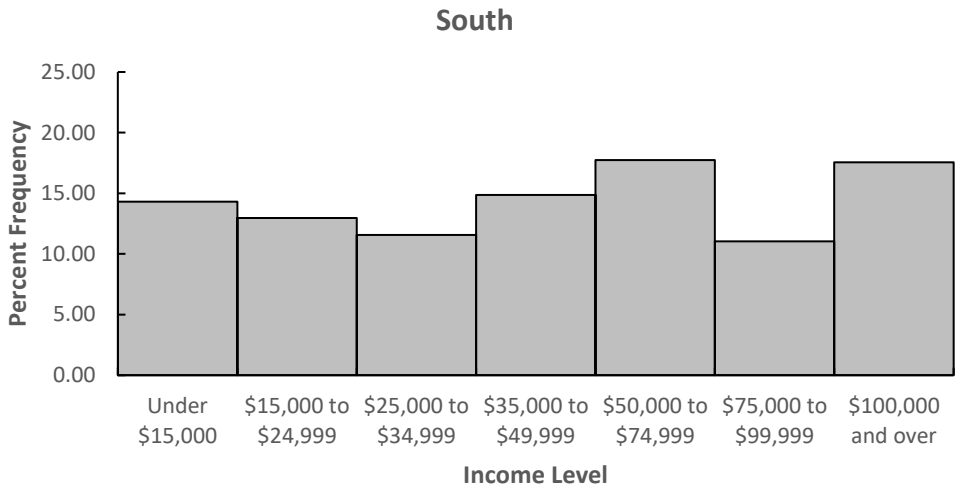
b. West: 18.44 + 11.77 + 23.43 = 53.64%

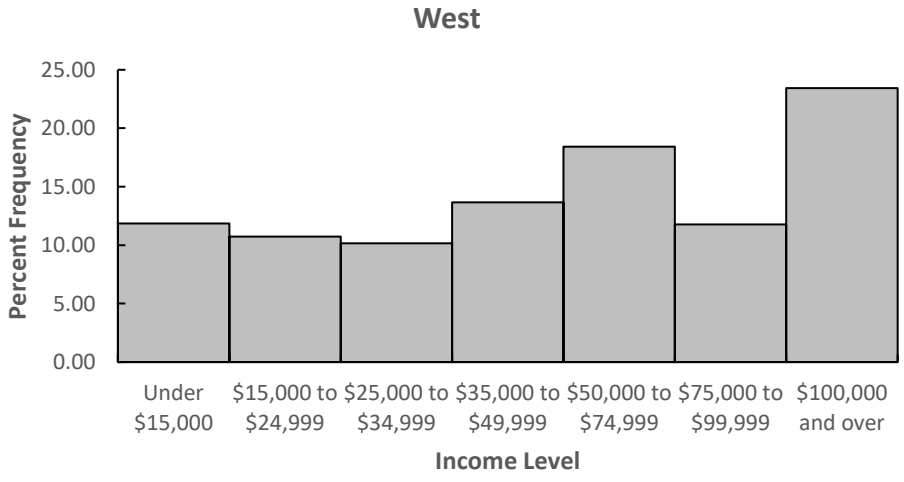
South: 17.73 + 11.04 + 17.57 = 46.34%

c.









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

d. Column percentages follow.

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 Higher
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 and 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 and 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 cross-tabulation shows that four 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 and 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 ($ Billion)	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.

	One-Year Value Change (%)						
Industry	–60–41	-40–21	–20–1	0–19	20–39	40–60	Total
Automotive and 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.

One-Year 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 one-year value change (%). The technology brands had the greatest variability.

35. a.

Hwy MPG

Size	20–24	25–29	30–34	35–39	40–44	Total
Compact	13	25	49	29	6	122
Large	10	31	19	11	1	72
Midsize	15	35	61	29	7	147
Total	~~91~~38	~~129~~91	~~69~~129	~~14~~69	~~91~~14	341

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	Total
A	3	43	57	5		108
F		8	48	82	16	154
R	10	33	32	4		79
Total	13	84	137	91	16	341

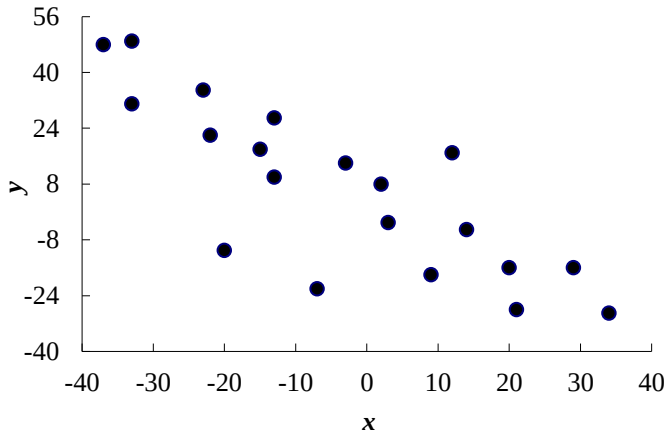
d. Higher fuel efficiencies are associated with front-wheel-drive cars.

e.

	City MPG					
Fuel Type	10–14	15–19	20–24	25–29	30–34	Total
P	13	58	94	16	1	182
R		26	43	75	15	159
Total	13	84	137	91	16	341

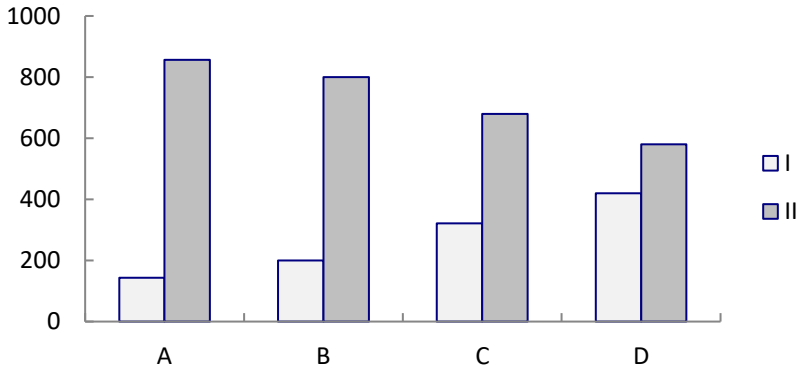
f. Higher fuel efficiencies are associated with cars that use regular gas.

36. a.



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

37. a.

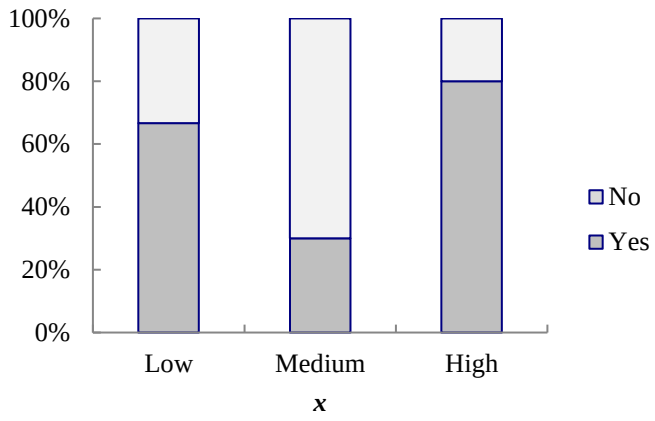


b. As X goes from A to D the frequency for I increases and the frequency of II decreases.

		y		
		Yes	No	
	Low	66.667	33.333	100
x	Medium	30.000	70.000	100

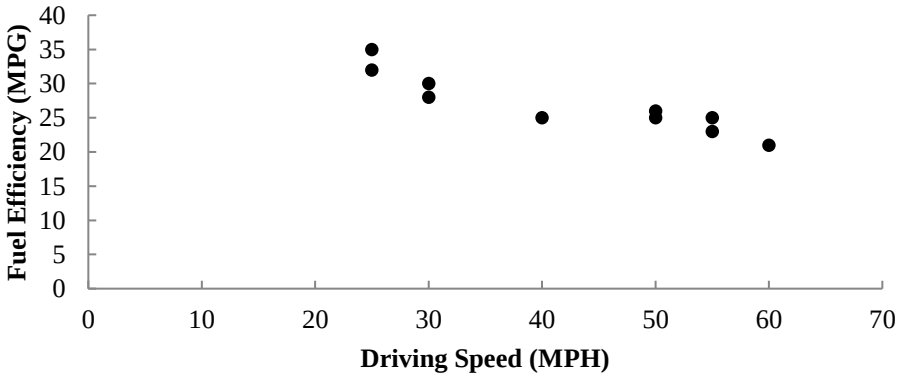
High	80.000	20.000	100

38. a.



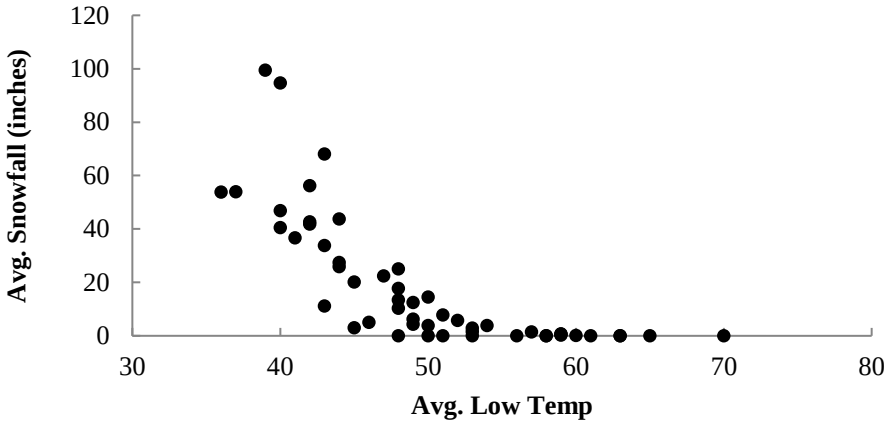
b.

39. a.



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

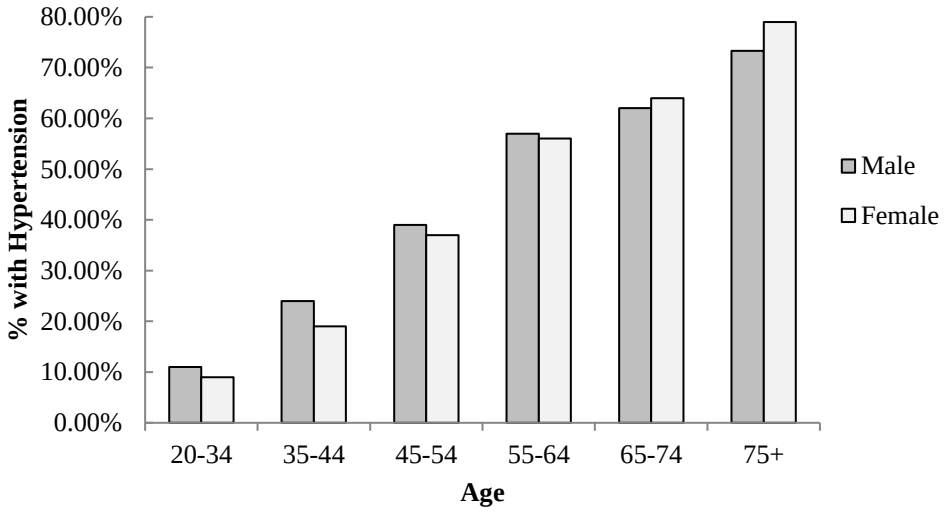
40. a.



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, New York, and Rochester, New York. Both are located near large lakes in the state.

41. a.

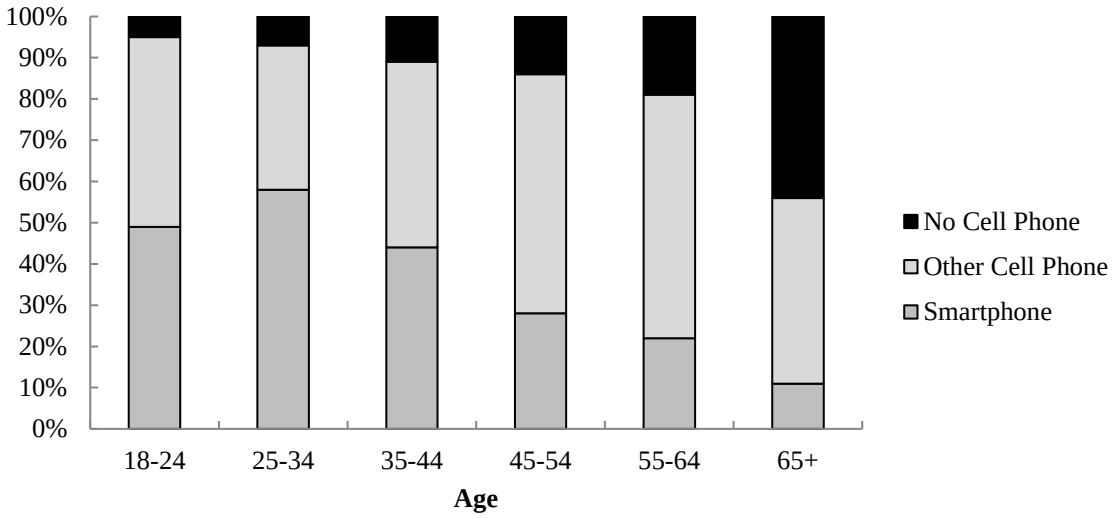


b. The percentage of people with hypertension increases with age.

c. For ages before 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 for

males.

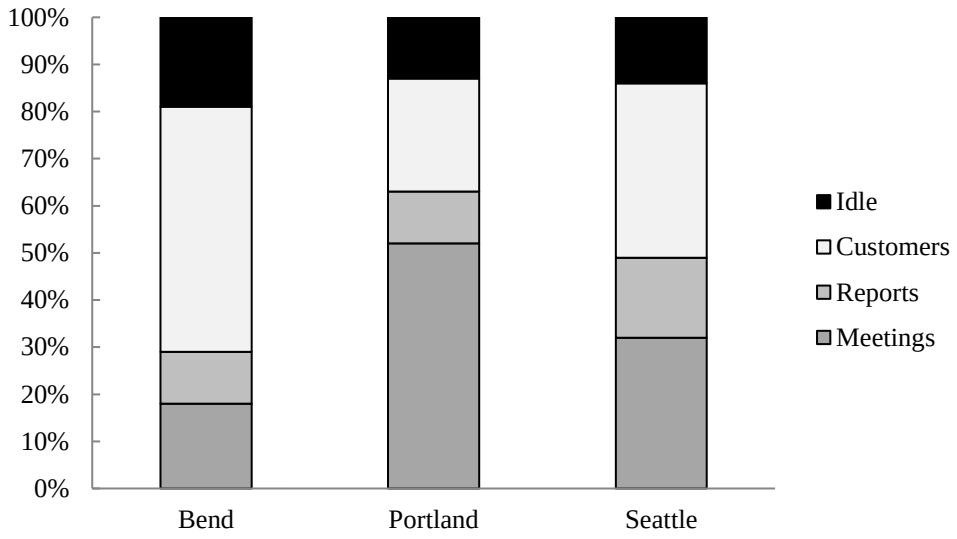
42. a.



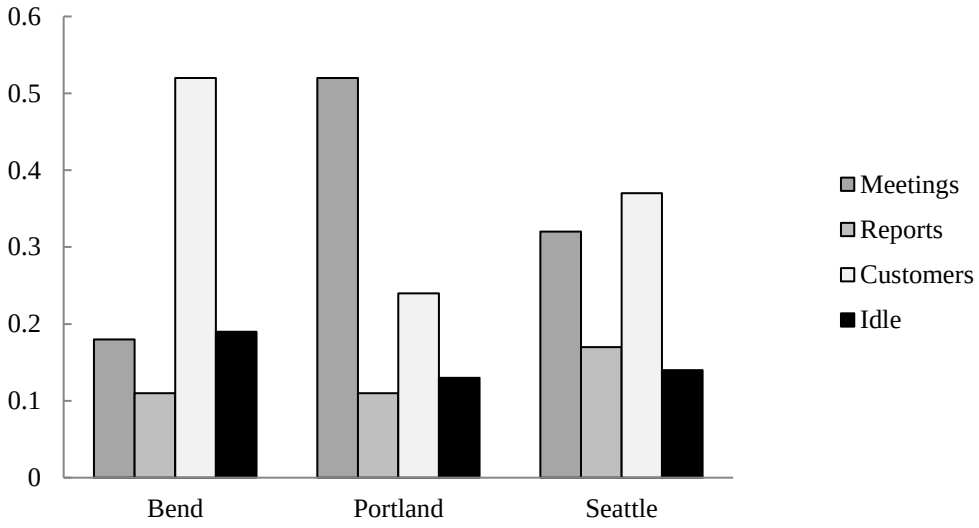
b. After increasing in ages 25–34, smartphone ownership decreases with increasing age. The percentage of people with no cell phone increases with age. There is less variation across age groups in the percentage who own other cell phones.

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 necessity than luxury.

43. a.



b.

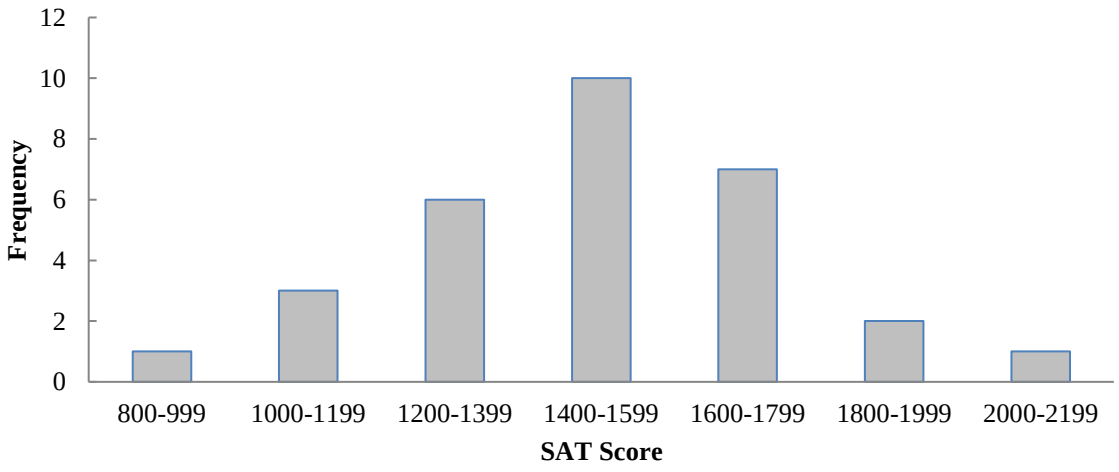


c. The stacked bar chart seems simpler than the side-by-side bar chart and more easily conveys the differences in store managers' use of time.

44. a.

Class	Frequency
800–999	1

1000–1199	3
1200–1399	6
1400–1599	10
1600–1799	7
1800–1999	2
2000–2199	2
Total	30



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

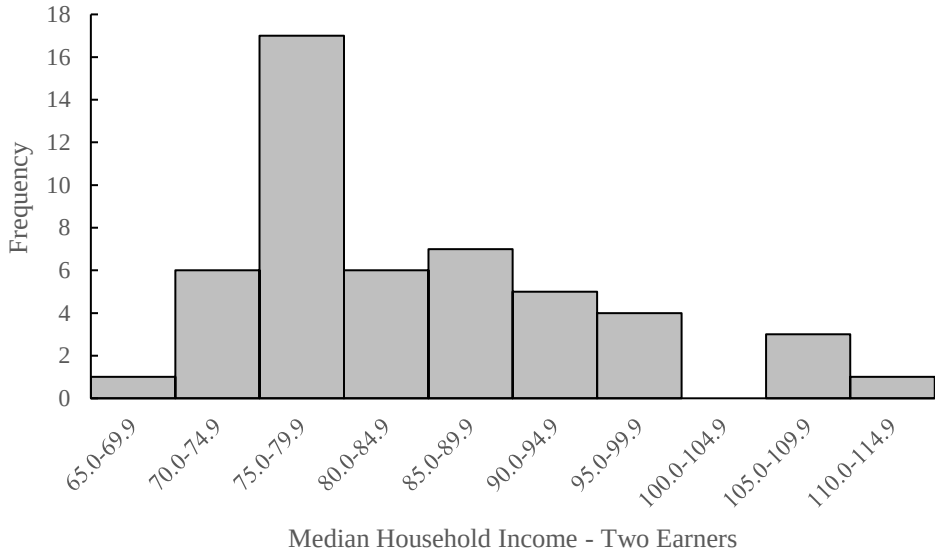
c. Ten of 30, or 33%, of the scores are between 1400 and 1599. The average SAT score looks to be slightly more than 1500. Scores below 800 or above 2200 are unusual.

45. a.

Median Household Income	Frequency	Percent Frequency
65.0–69.9	1	2

70.0–74.9	6	12
75.0–79.9	17	34
80.0–84.9	6	12
85.0–89.9	7	14
90.0–94.9	5	10
95.0–99.9	4	8
100.0–104.9	0	0
105.0-109.9	3	6
110.0-114.9	1	2
	50	100%

b.



c. The distribution is skewed to the right. There is a gap in the \$100.0–\$104.9 range.

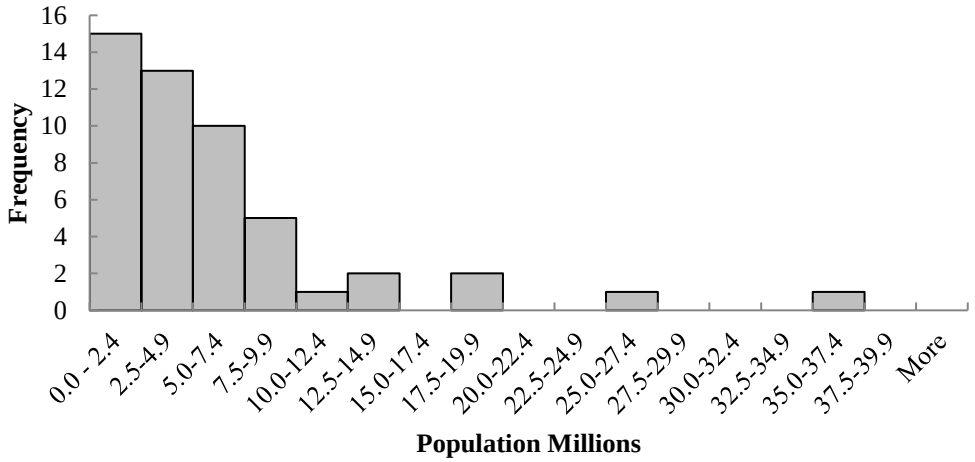
The most frequent range for the median household income is \$75.0–\$79.9 thousand.

d. New Jersey \$110.7 thousand

e. Idaho \$67.1 thousand

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. Fifteen states (30%) have a population less than 2.5 million. More than half of the states have populations of less than 5 million (28 states, or 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 (600.000).

47. a.

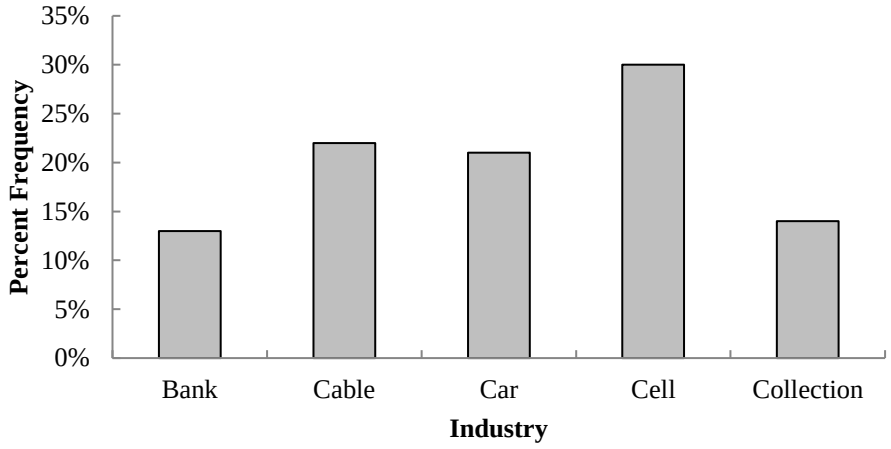
```
 1    8
 2    014
 3    18
 4    007899
 5    012444578
 6    00139
 7    237888
 8    011
 9    1
10    3
11    0289
12    9
13    01
14
15    46
16    68
17
18
19    2
20
21
22
23
24
25
26
27    2
```

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.

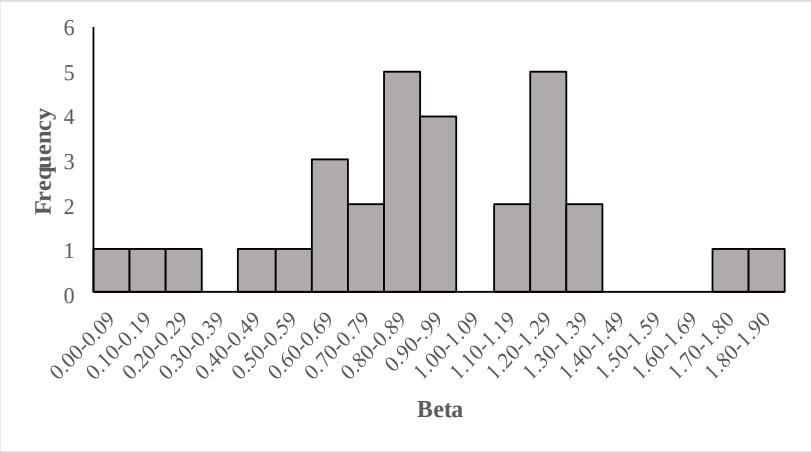
d. The percentage frequency distribution shows that the two financial industries (banks and collection agencies) had about the same number of complaints. Also, new car dealers and cable and satellite television companies also had about the same number of complaints.

49. a.

Beta	Frequency	Percent Frequency
0.00–0.09	1	3.3
0.10–0.19	1	3.3
0.20–0.29	1	3.3
0.30–0.39	0	0.0
0.40–0.49	1	3.3
0.50–0.59	1	3.3
0.60–0.69	3	10.0
0.70–0.79	2	6.7
0.80–0.89	5	16.7
0.90–.99	4	13.3
1.00–1.09	0	0.0
1.10–1.19	2	6.7
1.20–1.29	5	16.7
1.30–1.39	2	6.7
1.40–1.49	0	0.0
1.50–1.59	0	0.0
1.60–1.69	0	0.0

1.70–1.80	1	3.3
1.80–1.90	1	3.3
Total	30	100.0%

b.



c. The distribution is somewhat skewed to the left.

d. The stock with the highest beta is JP Morgan Chase & Company with a beta of 1.84. The stock with the lowest beta is Verizon Communications, Inc., with a beta of .04.

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
Less than $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 higher	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 Higher
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 less than $25,000 and households with an income of $100,000 or more. For

instance, 75.26% of households with an income of less than $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, she should receive the scholarship offer.

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

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

Based on this cross-tabulation, 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, she 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 on separate or unaggregated cross-tabulations, the conclusion can be reversed when the cross-tabulations are grouped or aggregated. When Simpson's paradox is present, the decision maker will have to decide whether the unaggregated or aggregated form of the cross-tabulation is more helpful in identifying the desired conclusion. *Note:* The authors prefer the recommendation to offer the scholarship to Emily Janson because it is based on 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 unaggregated 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. Cross-tabulation 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. Cross-tabulation 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. Twelve companies had 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 even though 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.

	Tuition and Fees ($)								
Year Founded	1–5,000	10,001–15,000	15,001–20,000	20,001–25,000	25,001–30,000	30,001–35,000	35,001–40,000	40,001–45,000	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 and Fees ($)								
Year Founded	1–5,000	10,001–15,000	15,001–20,000	20,001–25,000	25,001–30,000	30,001–35,000	35,001–40,000	40,001–45,000	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.

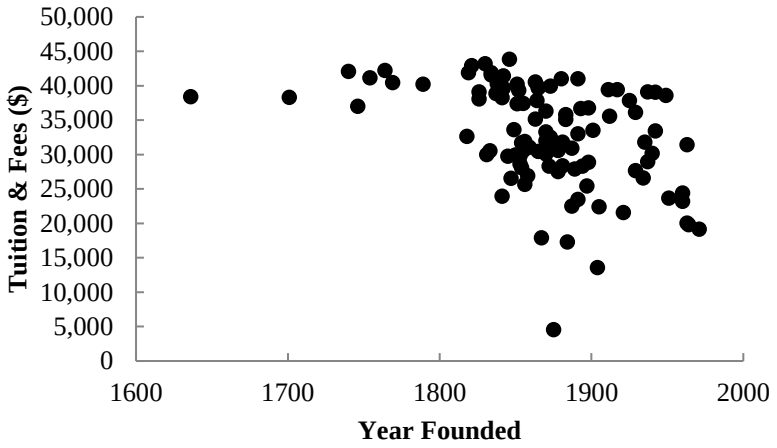
	Percent Graduating													
Year Founded	35–40	40–45	45–50	50–55	55–60	60–65	65–70	70–75	75–80	80–85	85–90	90–95	95–100	Grand Total
1600–1649													1	1
1700–1749													3	3
1750–1799												1	3	4
1800–1849						1	2	4	2	3	4	3	2	21
1850–1899			1	2	4	3	11	5	9	6	3	4	1	49
1900–1949	1	1	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

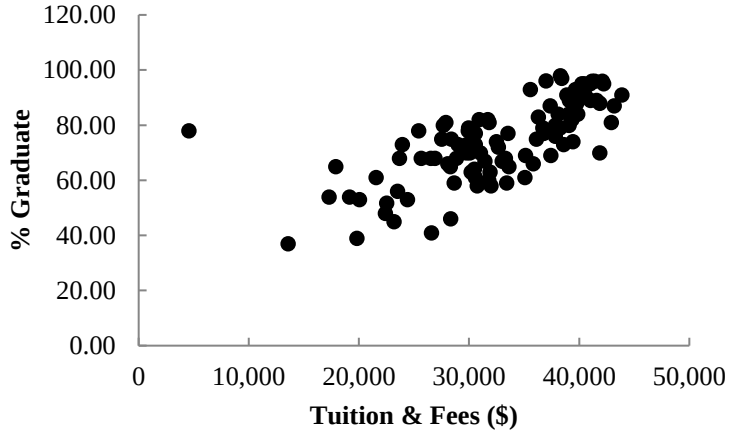
c. Older colleges and universities tend to have higher graduation rates.

55. a.



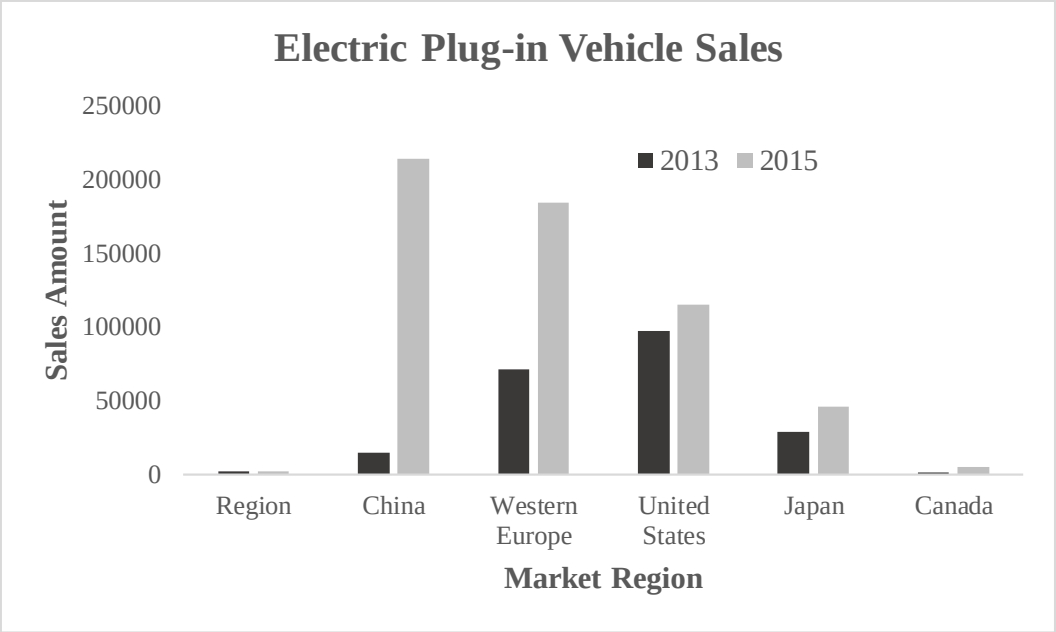
b. Older colleges and universities tend to be more expensive.

56. a.



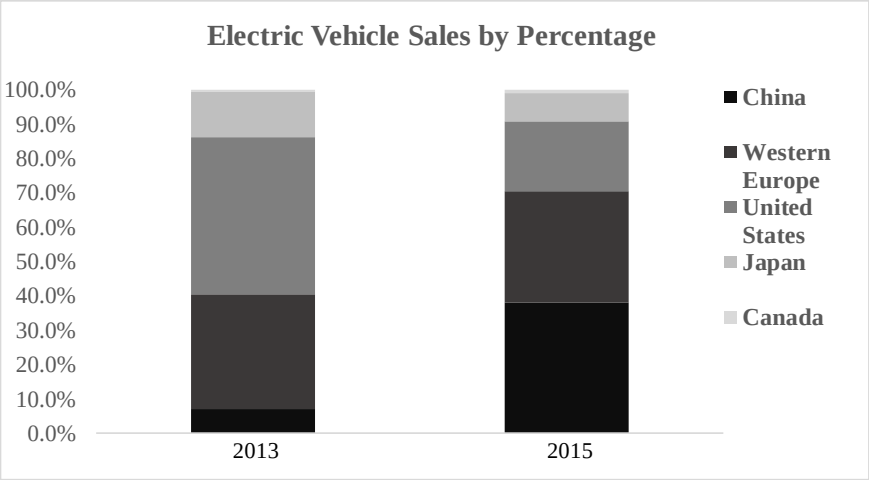
b. There appears to be a strong positive relationship between Tuition and Fees and Percent Graduating.

57. a.



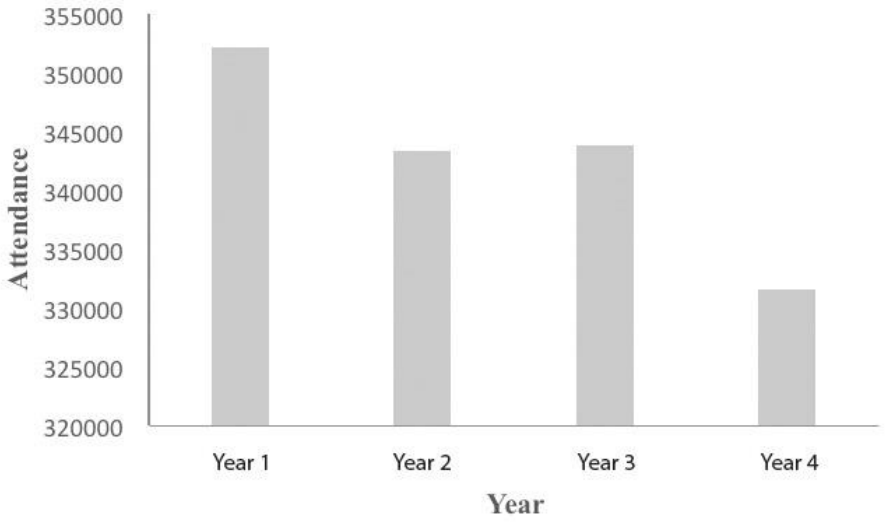
b.

Region	2013	2015
China	7.0%	37.9%
Western Europe	33.4%	32.6%
United States	45.6%	20.4%
Japan	13.5%	8.2%
Canada	0.4%	0.9%
Total:	100.0%	100.0%



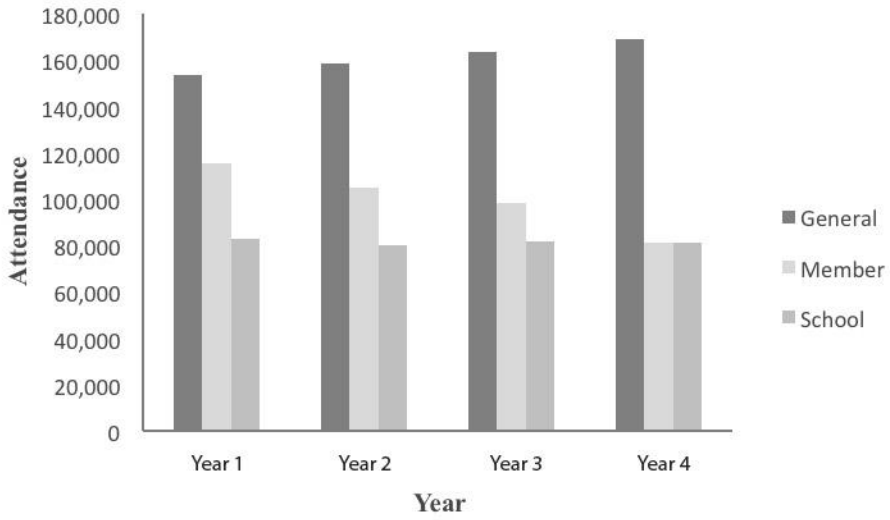
c. The graph in~~s~~ part a is more insightful because is shows the change in vehicle sales over time for each market region.

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.

Case Solutions

Chapter 2

Descriptive Statistics: Tabular and Graphical Presentations

Case Problem 1 Pelican Stores

1. There were 70 promotional customers and 30 regular customers. Because there are 100 observations in the sample, the frequency and percent frequency distribution are the same. Percent frequency distributions for many of the variables are given.

No. of Items	Percent Frequency
1	29
2	27
3	10
4	10
5	9
6	7
7 or more	8
Total:	100

Net Sales	Percent Frequency
0.00–24.99	9
25.00–49.99	30
50.00–74.99	25
75.00–99.99	10
100.00–124.99	12
125.00–149.99	4
150.00–174.99	3
175.00–199.99	3
200 or more	4
Total:	100

Method of Payment	Percent Frequency
American Express	2
Discover	4
MasterCard	14
Proprietary Card	70
Visa	10
Total:	100

Gender	Percent Frequency
Female	93
Male	7
Total:	100

Martial Status	Percent Frequency
Married	84
Single	16
Total:	100

Age	Percent Frequency
20–29	10
30–39	30

40–49	33
50–59	16
60–69	7
70–79	4
Total:	100

These percent frequency distributions provide a profile of Pelican's customers. Many observations are possible, including:

- A large majority of the customers use National Clothing's proprietary credit card.

- More than half of the customers purchase one or two items, but a few make numerous purchases.

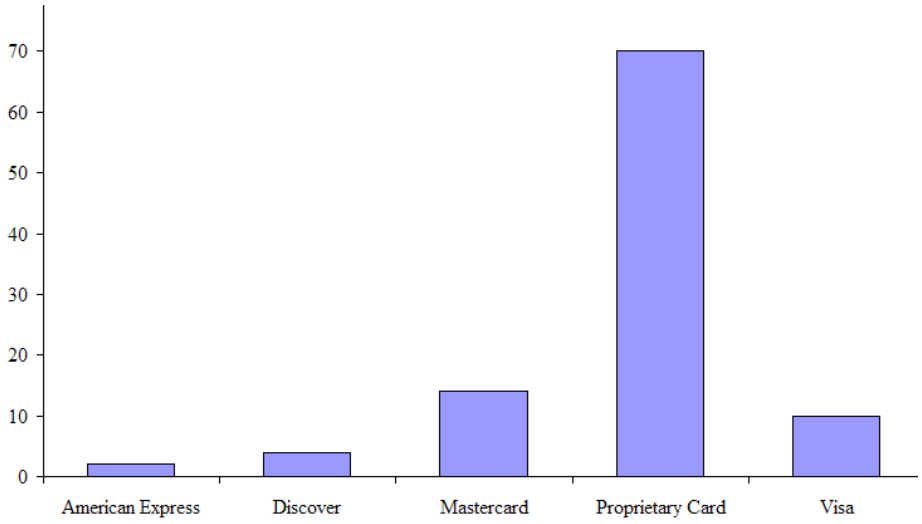
- The percent frequency distribution of net sales shows that 61% of the customers spent $50 or more.

- Customers are distributed across all adult age groups.

- The overwhelming majority of customers are female.

- Most of the customers are married.

2.

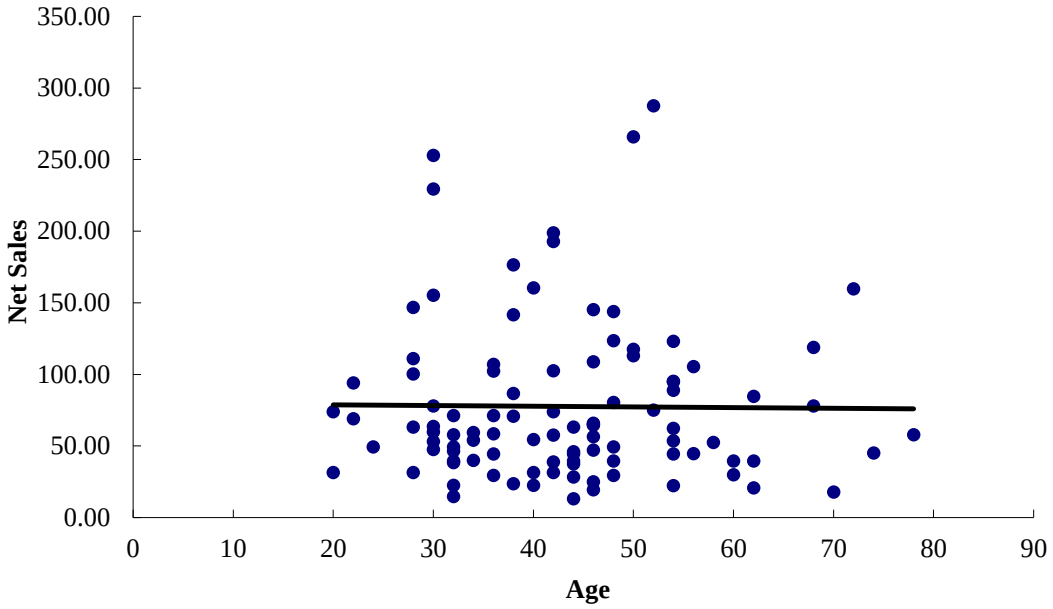


3. A crosstabulation of type of customer versus net sales is shown.

	Net Sales									
Customer	0–25	25–50	50–75	75–100	100–125	125–175	175–200	200–225	225–250	250-
Promotional	7	17	17	8	9	3	2	3	1	2
Regular	2	13	8	2	3	1	1			
Total	9	30	25	10	12	4	3	3	1	2

From the crosstabulation it appears that net sales are larger for promotional customers.

4. A scatter diagram of Net Sales versus Age is shown as follows. A trend line has been fitted to the data. From this, it appears that there is no relationship between net sales and age.



Age is not a factor in determining net sales.

Case Problem 2 Movie Theater Releases

This case provides the student with the opportunity to use tabular and graphical presentations to analyze data from the movie industry. Developing and interpreting frequency distributions, percent frequency distributions and scatter diagrams are emphasized. The interpretations and insights can be quite varied. We illustrate some below.

Frequency Distribution and Percent Frequency Distribution

The choice of the classes for frequency distributions or percent frequency distributions can be expected to vary. The frequency distributions we developed are as follows:

Opening Gross Sales (Millions)	Frequency (or Percentage)
$0–9.99	14
10–19.99	34
20–29.99	22
30–39.99	10
40–49.99	5
50–59.99	3
60–69.99	1
70–79.99	2
80–89.99	1
90–99.99	0
100–109.99	2

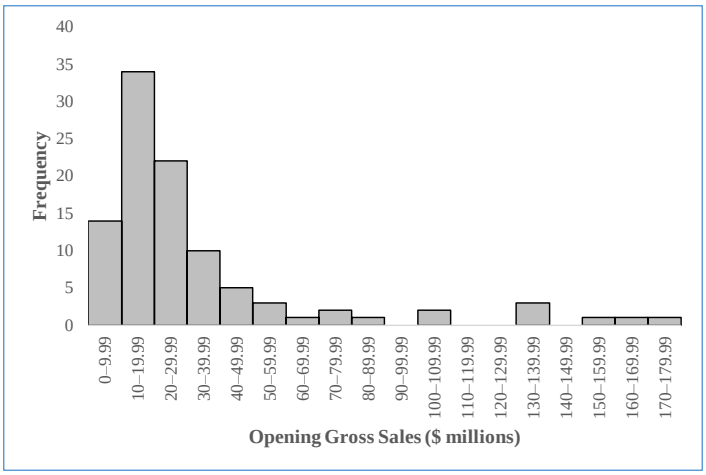
110–119.99	0
120–129.99	0
130–139.99	3
140–149.99	0
150–159.99	1
160–169.99	1
170–179.99	1

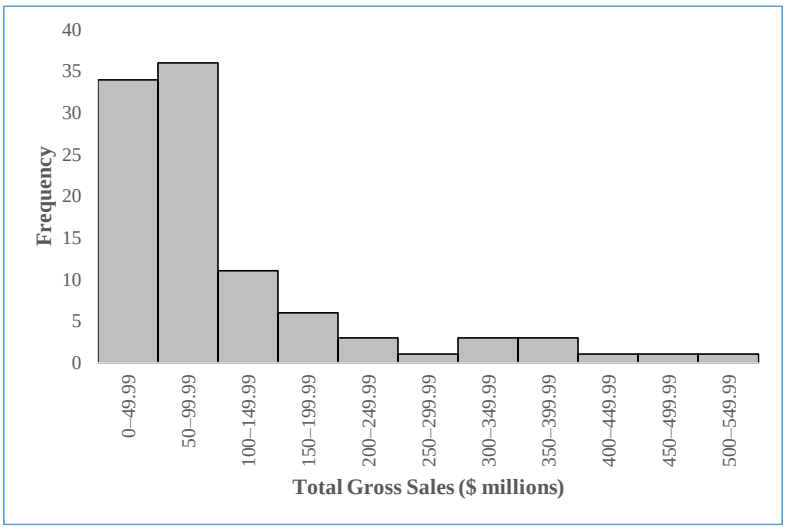
Total Gross Sales (Millions)	Frequency (or Percentage)
$0–49.99	34
50–99.99	36
100–149.99	11
150–199.99	6
200–249.99	3
250–299.99	1
300–349.99	3
350–399.99	3
400–449.99	1
450–499.99	1
500–549.99	1
Total	100

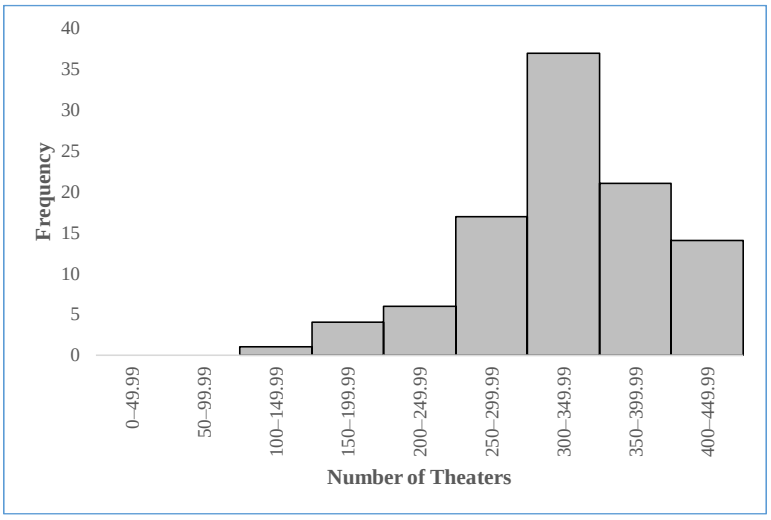
Number of Theaters	Frequency (or Percentage)
0–499	0
500–999	0
1,000–1,499	1
1,500–1,999	4
2,000–2,499	6
2,500–2,999	17
3,000–3,499	37
3,500–3,999	21
4,000–4,499	14
100	

Number of Weeks in Release	Frequency (or Percentage)
0–4	0
5–9	15
10–14	43
15–19	23
20–24	14
25–29	4
30–34	0
35–39	0
40–44	0
45–49	1
100	

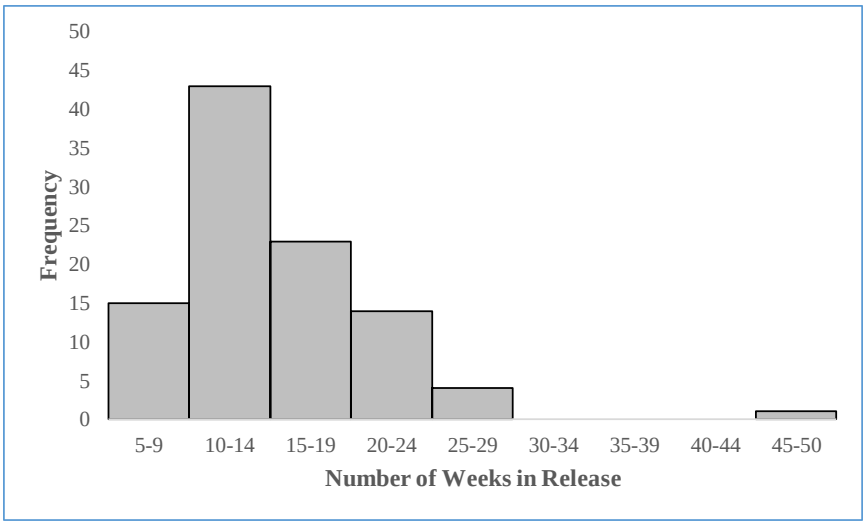
Histograms







The following histograms are based on the frequency distributions shown above.



Interpretation

Opening Weekend Gross Sales The distribution is skewed to the right. Numerous movies have somewhat low opening weekend gross sales, while a relatively few (8%) have an opening weekend gross sales of $100 million or more. Only 3% had opening weekend gross sales of $150 million or more. Eighty percent of the movies had opening weekend gross sales less than $40 million, and 92% of the movies had opening weekend gross sales less than $100 million.

Total Gross Sales This distribution is also skewed to the right. Again, the majority of the movies have relatively low total gross sales with 70% of movies having gross sales less than $100 million and 91% less than $300 million. Highly successful blockbuster movies are rare. Total gross sales of more than $400 million occurred only 3% of the time, and gross sales of more than $500 million occurred only 1% of the time. Unless there is something unusually attractive about the movie, a total gross sales less than $100 million appears typical.

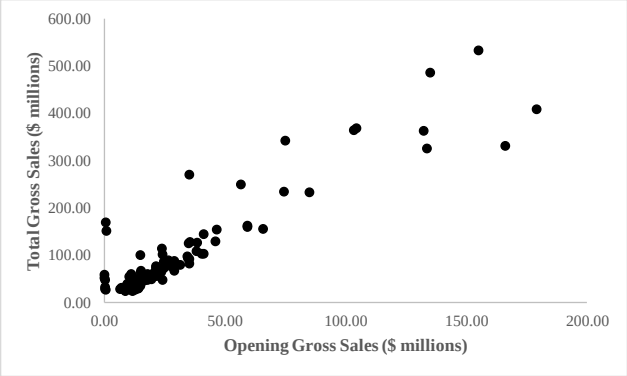
Number of Theaters This distribution is skewed to the left. The number of theaters range from slightly more than 1,000 to almost 4,500. Eighty-nine percent of the movies had large market exposure, playing in 2,500 or more theaters. No movies were in fewer than 1,000 theaters, and only 11% were in fewer than 2,500 theaters. Most top movies in 2016 appeared to receive large market exposure in 2,500 or more theaters.

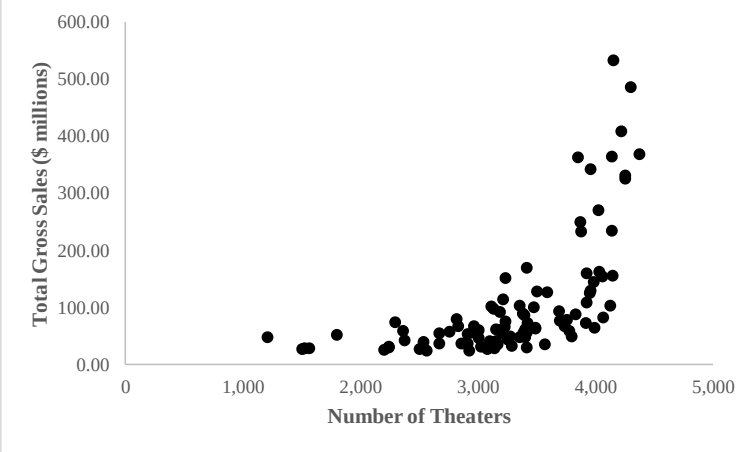
Number of Weeks in Release This distribution is skewed to the right, but not as much as the distributions on sales. Almost all movies in 2016 spent at least 10 weeks in release. Only 15% of movies in 2016 spent fewer than 10 weeks in release. One movie (*Hidden Figures*) spent much longer in release than any other movie at 46 weeks.

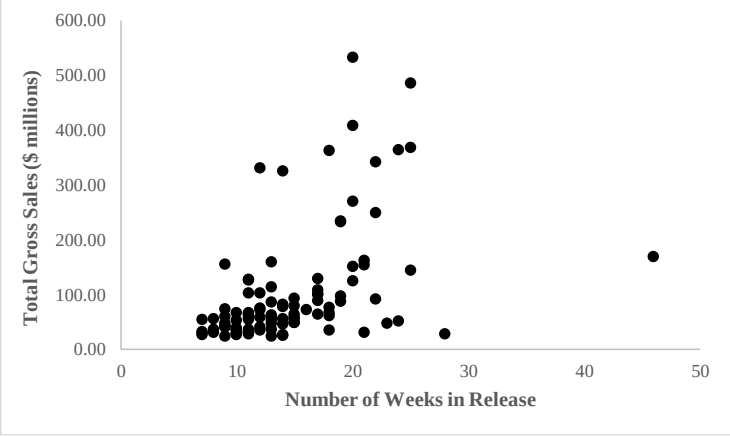
General Observations The data show there are relatively few high-end, highly successful movies. The financial rewards are there for the pictures that make the blockbuster level. But the majority of movies will have relatively low opening weekend gross sales and low total gross sales. Movies being shown in more than 2500 theaters and movies that spend at least 10 weeks in release are common.

Scatter Diagrams

Three scatter diagrams are suggested to show how Total Gross Sales is related to each of the other three variables.







Interpretation

Opening Weekend Gross Sales The scatter plot of total gross sales and opening weekend gross sales shows a strong positive relationship. Movies with the highest total gross sales were

those with the highest opening gross sales. How a movie does during its opening weekend should be a strong predictor of how the movie will do in terms of total gross sales. Note in the scatter diagram that the majority of the movies show a low opening weekend gross sales and a low total gross sales.

Number of Theaters The scatter plot of the total gross sales and number of theaters also shows a positive relationship. For movies playing in fewer than 3,500 theaters, the total gross sales were significantly less than those movies playing in more than 3,500 theaters. If the movie is shown in more theaters, higher total gross sales are anticipated. For movies playing in more than 3,500 theaters, the positive relationship is especially strong. This scatter chart also appears to show a nonlinear relationship because movies playing in the most theaters increase in total gross sales rapidly compared to those playing in fewer theaters.

Number of Weeks in Release The scatter plot of the total gross sales and number of weeks in release shows a positive relationship, but this relationship appears to be the weakest of the three relationships studied. Generally, the more successful movies with higher gross sales are in release for more weeks. However, this is not always the case. The longest released movie (*Hidden Figures*) had less in total gross sales than many movies that had shorter release times. And many movies that were in release for more than 20 weeks had less total gross revenue than those with fewer than 20 weeks in release. This suggests that in some cases blockbuster movies with high gross sales may run their course quickly and not have an excessively long run in release. At the same time, perhaps quality movies with a limited audience may not generate the high total gross sales but may still show a run of 20 or more weeks. The number of weeks in release does not appear to the best predictor of total gross sales.

Case Problem 3 Queen City

This case provides the student with the opportunity to use basic tabular and graphical presentations to describe data from the annual expenditures for the city of Cincinnati, Ohio. The data set is large relative to others in the text. It contains 5,427 records of expenditures. As such, one point of this case is to expose students to a larger data set and help them understand that the pivot tables and charts can be used on a larger data set. In some cases, the student will have to copy, paste, and aggregate data to create the desired tables and charts. Style of presentation may vary by student (for example, vertical versus horizontal bar charts may be used). We illustrate with results and comments below.

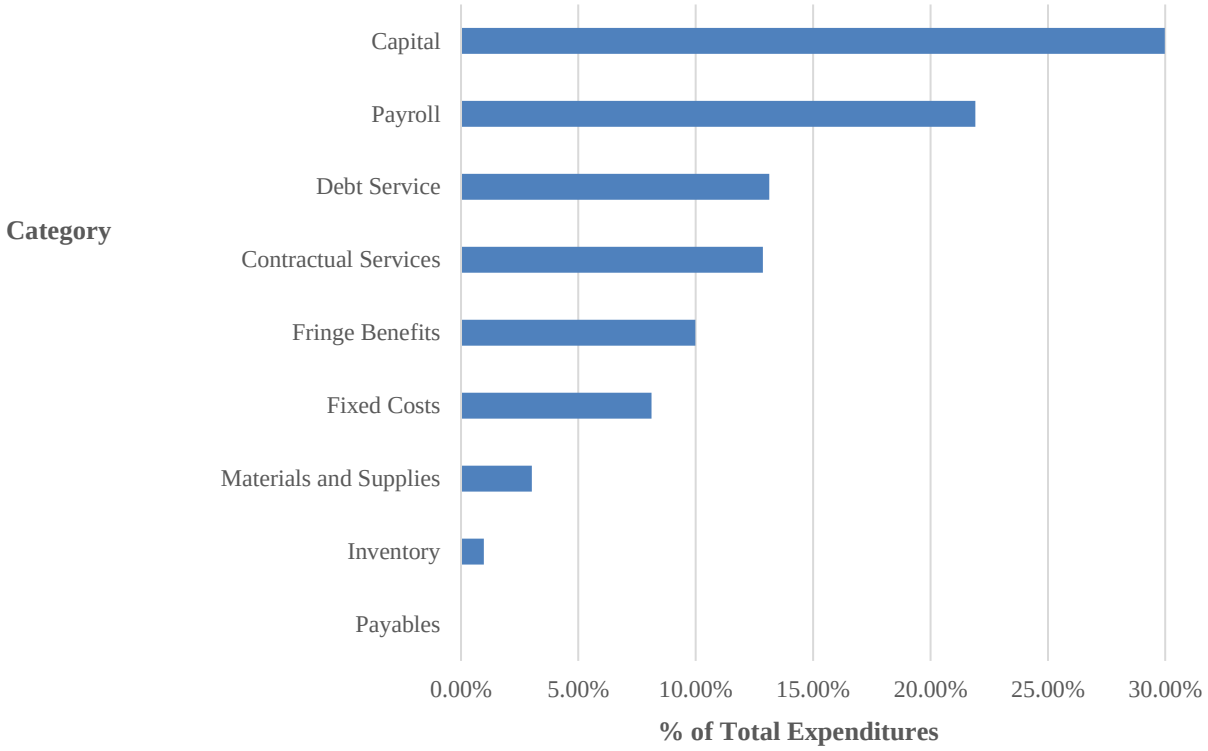
Expenditures by Category

The pivot table shows expenditures and percentage of total expenditures by category. The bar chart shows percentage of total expenditures by category (both the table and the bar chart are sorted in descending order). Capital expenditures and payroll account for more than 50% of all expenditures. Total expenditures are more than $660 million. Debt Service seems somewhat high with more than 10% of total expenditures.

Category	Total Expenditures ($)	% of Total Expenditures
Capital	198,365,854	29.98

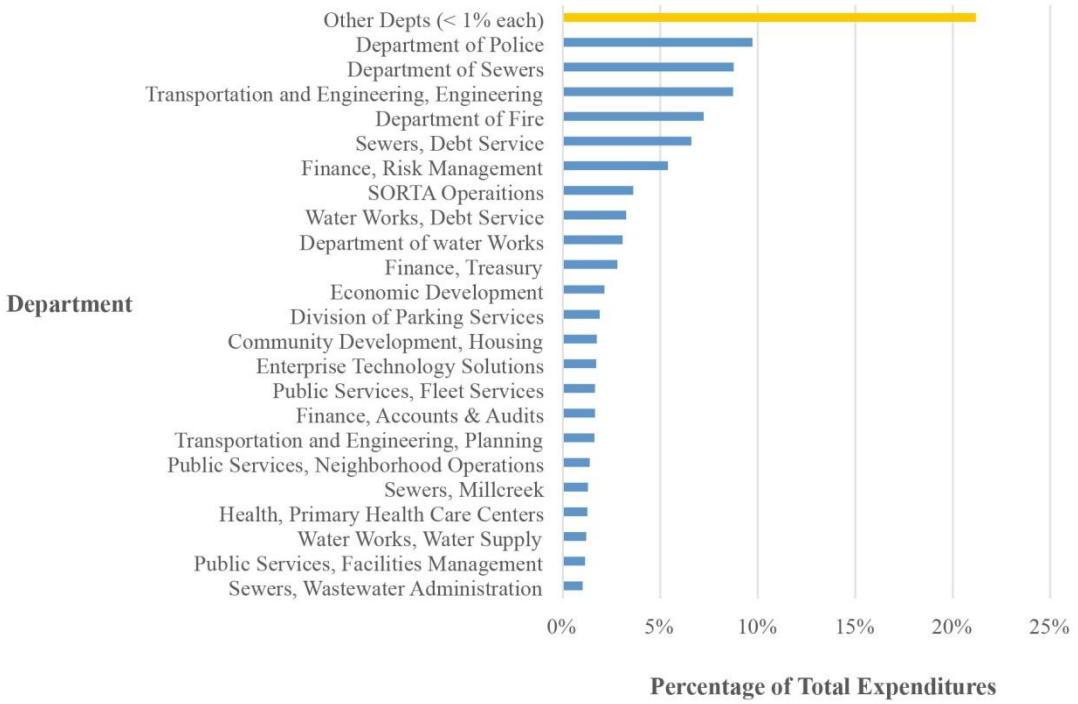
Payroll	145,017,555	21.92
Debt Service	86,913,978	13.14
Contractual Services	85,043,249	12.85
Fringe Benefits	66,053,340	9.98
Fixed Costs	53,732,177	8.12
Materials and Supplies	19,934,710	3.01
Inventory	6,393,394	0.97
Payables	180,435	0.03
Grand Total	661,634,693	100.0

Expenditures by Department



The following table and bar chart show the percentages of total expenditures incurred by department. Note that we have combined all departments that individually incurred less than 1% of the total expenditures. Of all 119 departments, 96 each account for less than 1% of the total expenditures. As shown as follows, only six individual departments incur 5% or more of the total expenditures. These include Police, Sewers, Transportation Engineering (Engineering). Fire, Sewer Debt Service, and Finance and Risk Management. Debt service on sewers as a percentage of total expenditures appears to be especially high.

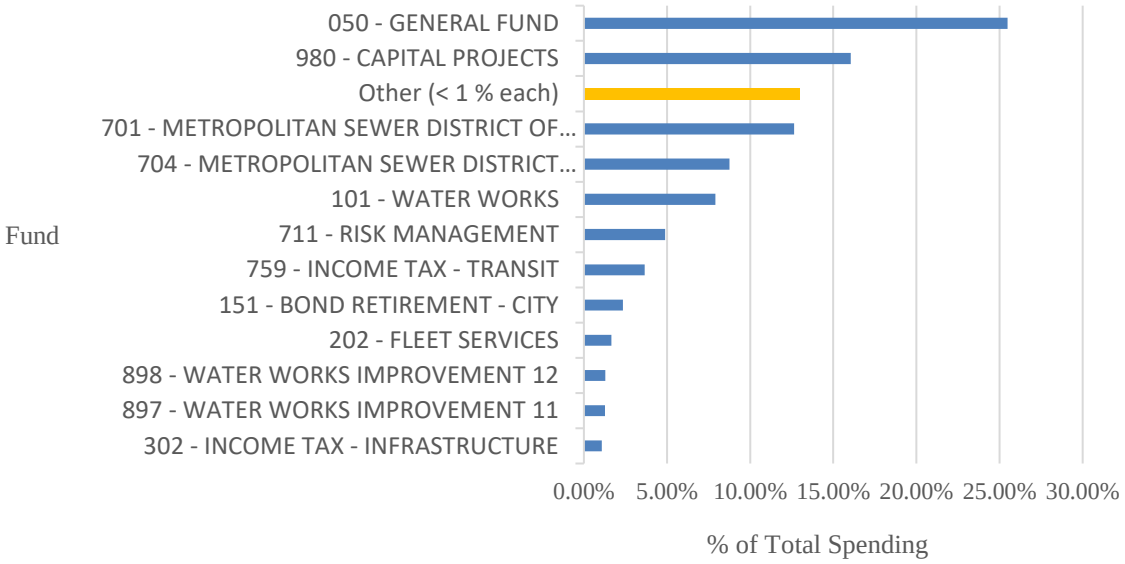
Department	% of Total Expenditures
Department of Police	9.7
Department of Sewers	8.8
Transportation and Engineering (Engineering)	8.7
Department of Fire	7.2
Sewer Debt Service	6.6
Finance, Risk Management	5.4
SORTA Operations	3.6
Water Works, Debt Service	3.2
Department of Water Works	3.1
Finance, Treasury	2.8
Economic Development	2.1
Division of Parking Services	1.9
Community Development, Housing	1.7
Enterprise Technology Solutions	1.7
Public Services, Fleet Services	1.7
Finance, Accounts and Audits	1.7
Transportation and Engineering, Planning	1.6
Public Services, Neighborhood Operations	1.4
Sewers, Millcreek	1.3
Health, Primary Health Care Centers	1.2
Water Works, Water Supply	1.2
Public Services, Facilities Management	1.1
Sewers, Wastewater Administration	1.0
Other Depts. (< 1% each)	21.2%
Total	100.0%



Expenditures by Fund

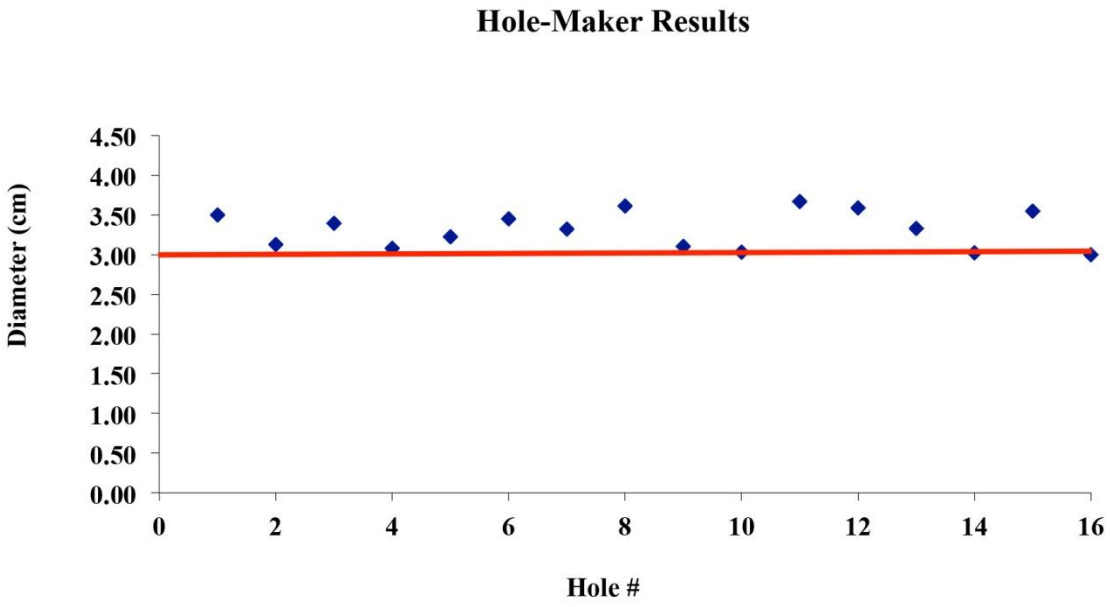
The following table and bar chart show the percentages of total expenditures charged by the fund used to pay. Note that we have combined those funds that each cover less than 1% of the total expenditures. Of 129 funds in the data base, 117 each account for less than 1% of total expenditures.

Fund	% of Total Expenditures Covered
050–GENERAL FUND	25.5
980–CAPITAL PROJECTS	16.0
701–METROPOLITAN SEWER DISTRICT OF GREATER CINCINNATI	12.7
704–METROPOLITAN SEWER DISTRICT CAPITAL IMPROVEMENTS	8.8
101–WATER WORKS	7.9
711–RISK MANAGEMENT	4.9
759–INCOME TAX–TRANSIT	3.7
151–BOND RETIREMENT–CITY	2.4
202–FLEET SERVICES	1.7
898–WATER WORKS IMPROVEMENT 12	1.3
897–WATER WORKS IMPROVEMENT 11	1.3
302–INCOME TAX–INFRASTRUCTURE	1.1
Other (< 1 % each).	12.9
Total	100.0%



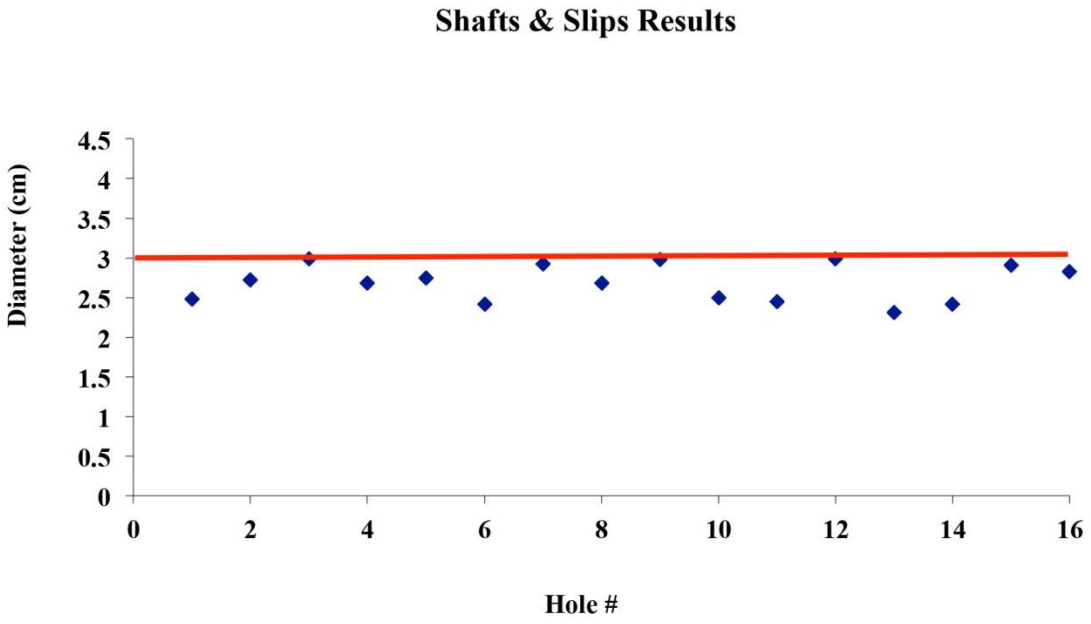
Other Points Of 5,427 records of expenditures in the database, 235 (4.3%) are negative.

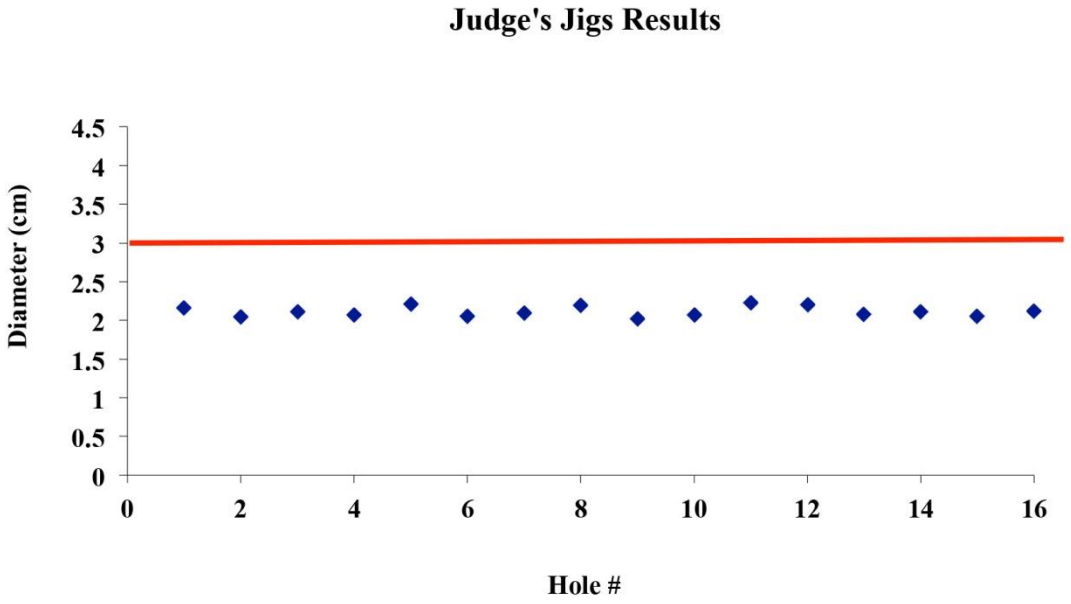
Case Problem 4 Cut-Rate Machining, Inc.



A scatter diagram of the results for Hole-Maker in the order the holes were drilled shows that this machine consistently overdrills and is moderately consistent.

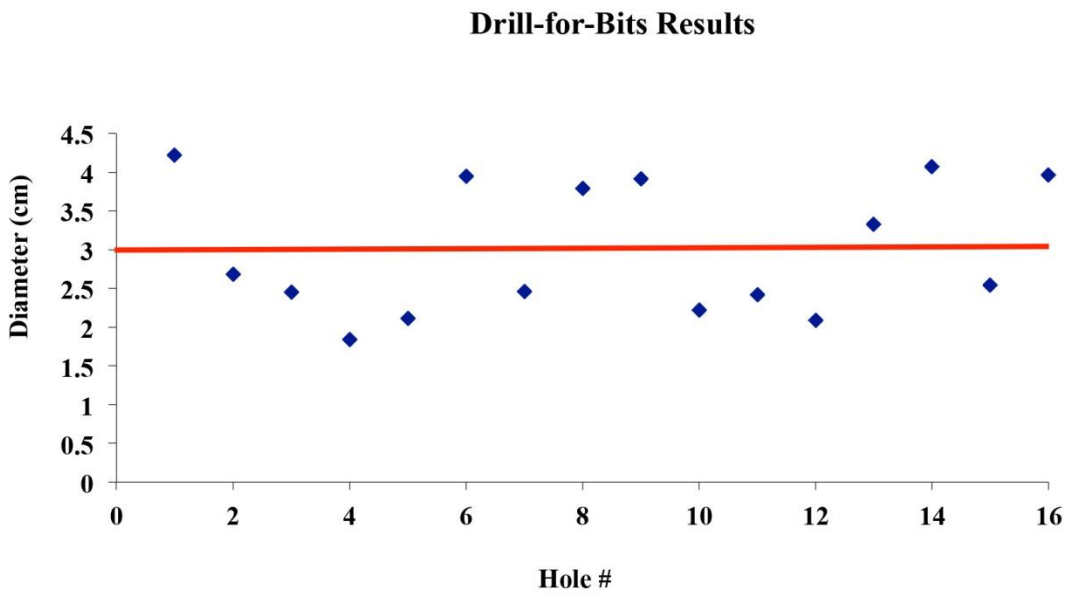
A scatter diagram of the results for Shafts & Slips in the order the holes were drilled shows that this machine consistently underdrills and is moderately consistent.





A scatter diagram of the results for Judge's Jigs in the order the holes were drilled shows that on average this machine this machine consistently underdrills and is extremely consistent.

A scatter diagram of the results for Drill-for-Bits in the order the holes were drilled shows that an average diameter of approximately 3 centimeters. However, this machine this machine is extremely inconsistent.



If we focus solely on the average performance of a drill, we would purchase Drill-for-Bits as the diameters of holes drilled by this vendor's drill appear to be centered at approximately 3 centimeters. However, the diameters of the holes drilled by Drill-for-Bits' machine are extremely inconsistent—several are more than ½ centimeter too wide and several are more than ½ centimeter to narrow.

The diameters of holes drilled by the machine provided by Hole-Maker are more consistent than those drilled by the machine provided by Drill-for-Bits, and this machine did not drill a single hole that is too narrow. If holes that are slightly too wide are acceptable, we should consider purchasing our drill from Hole-Maker.

The diameters of holes drilled by the machine provided by Shafts & Slips are similar in consistency to the holes by the machine provided by Hole-Maker, and this machine did not drill a single hole that is too wide. If holes that are slightly too small are acceptable, we should consider purchasing our drill from Shafts & Slips.

The diameters of holes drilled by the machine provided by Judge's Jigs are far more consistent than holes by the machine provided any of the other vendors, but these holes are far too narrow. We should determine if this drill can be recalibrated to that, then the mean size of holes drilled is approximately 3 centimeters. If this can be done, we should consider purchasing our drill from Judge's Jigs and recalibrating the drill; this would give us a machine that consistently drills holes of approximately 3 centimeters.

However, we should scrutinize the way these data were collected before we make a decision. We were told that Weideman started all four machines at 8 A.M. and let them warm up for two hours. We also see from the data that the drill provided by Hole-Maker was tested from 10 A.M. to noon, the drill provided by Shafts & Slips, Inc. was tested from noon to 2 P.M., the drill provided by Judge's Jigs was tested from 2 P.M. to 4 P.M., and the drill provided by Drill-for-Bits was

tested from 4 P.M. to 6 P.M. Were all drills allowed to keep running after the 8 A.M. to 10 A.M. warm-up period? Either way, this could bias the results.

We also see from the data that Ms. Ames ran the test drills from 10 A.M. to 4 P.M. when the drills provided by Hole-Maker, Shafts & Slips, and Judge's Jigs were tested. Mr. Silver ran the test drill from 4 P.M. to 6 P.M. when the drill provided by Drill-for- Bits was tested. If these two employees are not equally competent, then this could bias the results. Furthermore, did Ms. Ames become fatigued as the day progressed? Did she take a break for lunch or take a break at any other time?

We also note that we only tested one drill for each vendor. If the drill provided by a vendor is not representative of the drills that vendor produced, then this also could bias the results.

The data for this test should have been collected through an experimental study in which the four machine were all warmed up for the same amount of time and then left running as eight holes were drilled by each employee using the drill provided by each vendor in a random order. A design such as this would have eliminated the potential sources of bias we have identified and led to the collection of more reliable data, which would lead to a superior decision.