Chapter 2
Descriptive Statistics: Tabular and Graphical Presentations

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, a histogram, and an ogive 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 and scatter diagrams of bivariate data.

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

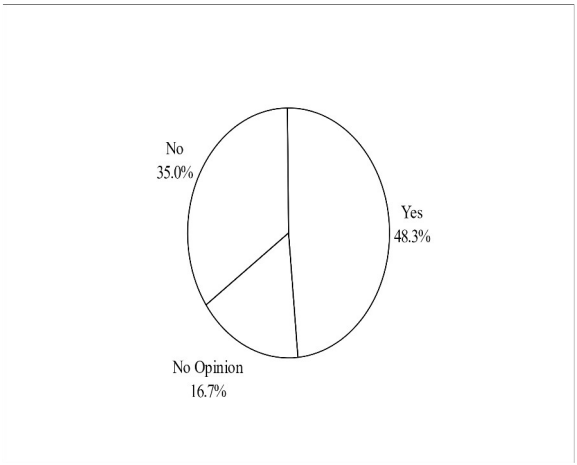
   c/d.

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

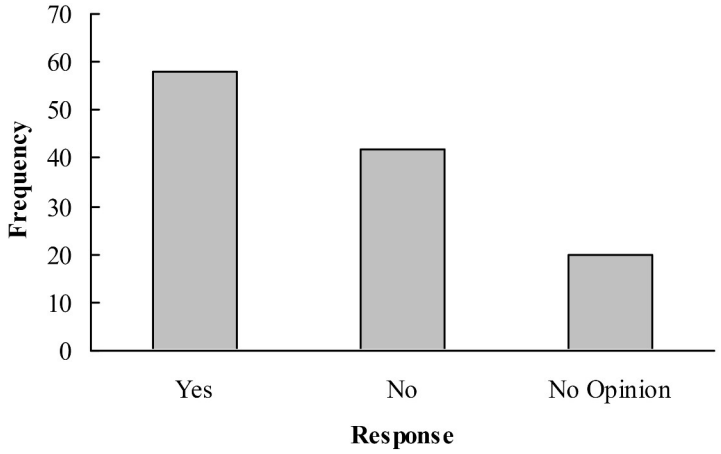
3. a. 360° x 58/120 = 174°

   b. 360° x 42/120 = 126°

   c.



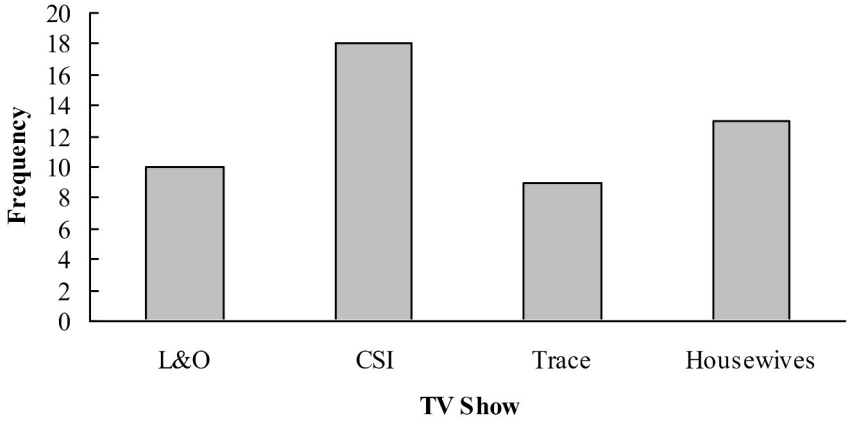
   d.

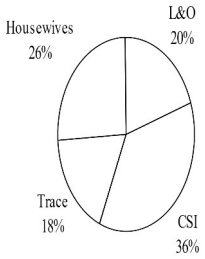


4. a. Categorical

b.

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



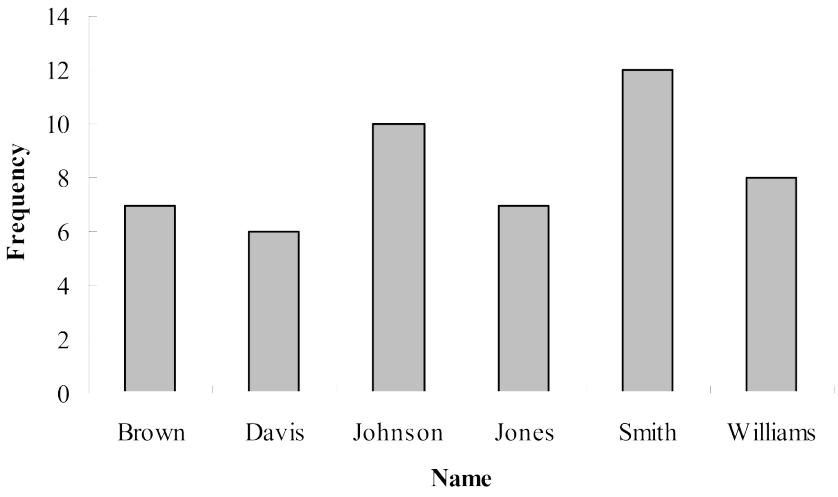


d. CSI had the largest viewing audience. Desperate Housewives was in second place.

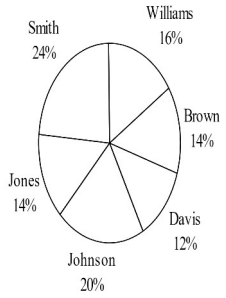
5. a.

Name	Frequency	Relative Frequency	Percent Frequency
Brown	7	.14	14%
Davis	6	.12	12%
Johnson	10	.20	20%
Jones	7	.14	14%
Smith	12	.24	24%
Williams	8	.16	16%
	50	1.00	

b.



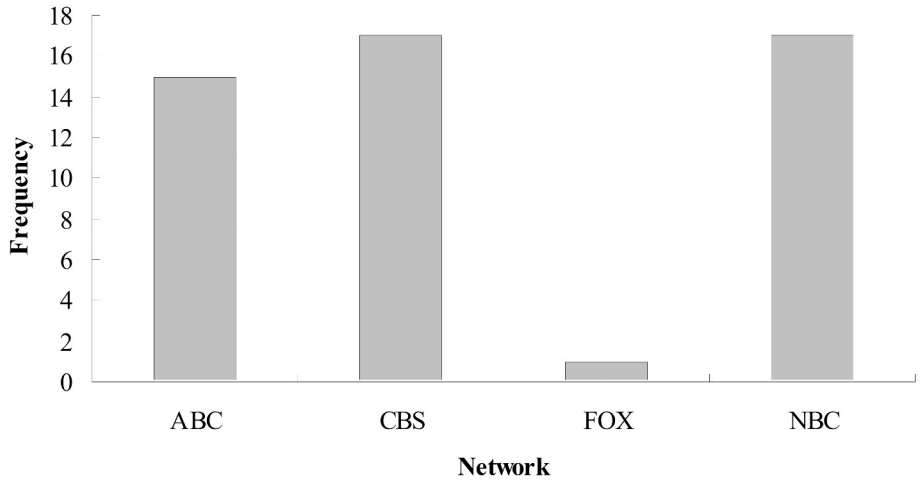
c. Brown .14 x 360 = 50.4°
Davis .12 x 360 = 43.2°
Johnson .20 x 360 = 72.0°
Jones .14 x 360 = 50.4°
Smith .24 x 360 = 86.4°
Williams .16 x 360 = 57.6°



d. Most common: Smith, Johnson and Williams

6. a.

Network	Frequency	Percent Frequency
ABC	15	30%
CBS	17	34%
FOX	1	2%
NBC	17	34%
Total	50	100%



b. CBS and NBC are tied, each with 17 of the top rated television shows. ABC is a close third with 15. The fact that the three networks are so close is surprising. FOX, the newest television network, does not have the history to compete with the other three networks in term of the top rated shows in television history.

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%)

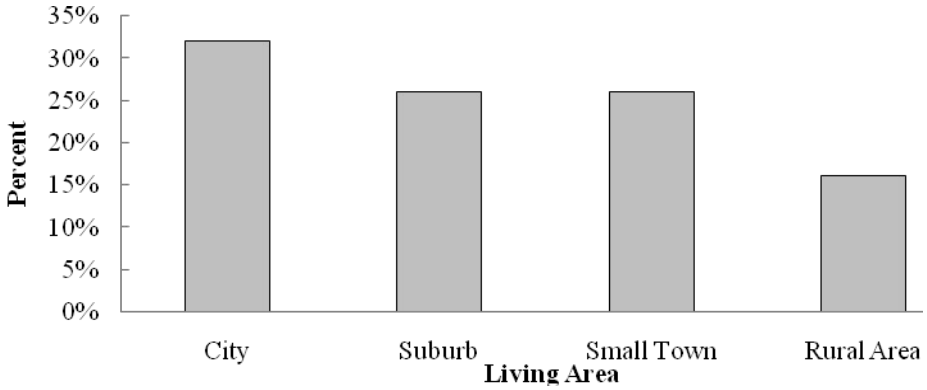
d. Right Field (Almost 13%)

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%	23%
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	20
Good	101
Fair	528
Bad	244
Terrible	122
Total	1015

b.

Rating	Percent Frequency
Excellent	2
Good	10
Fair	52
Bad	24
Terrible	12
Total	100

c.

d. 24% + 12% = 36% of adults in the United Sates think the Federal Bank is doing a bad or a terrible job in handling the credit problems. Only 10% + 2% = 12% think the Federal Bank is doing a good or excellent job.

e. 40% + 10% = 50% of adults in Spain think the European Central Bank is doing a bad or terrible job in handling the credit problems. Only 4% of adults in Spain think the European Central Bank is doing a good or excellent job.

Both countries show pessimism and relatively low confidence in how the banks are handling the credit problems in the financial markets. But in comparing the two countries, adults in Spain show more concern and more pessimism about the bank's ability compared to adults in the United States.

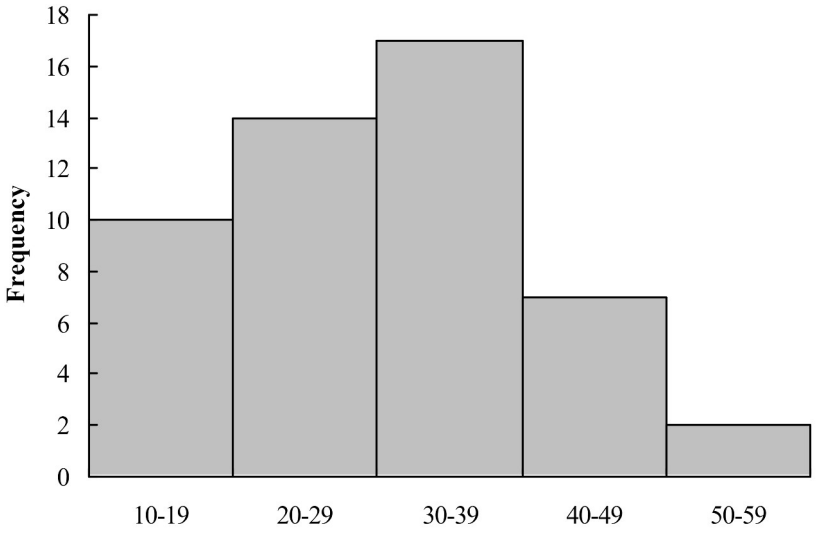
11.

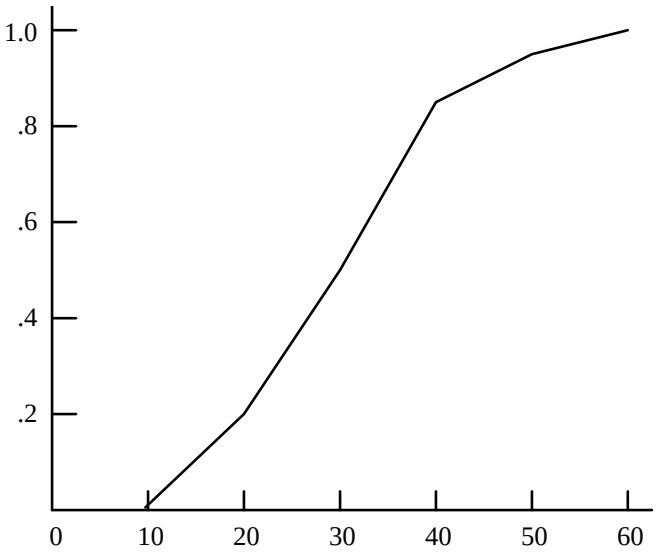
Class	Frequency	Relative Frequency	Percent Frequency
12-14	2	0.050	5.0
15-17	8	0.200	20.0
18-20	11	0.275	27.5
21-23	10	0.250	25.0
24-26	9	0.225	22.5
Total	40	1.000	100.0

12.

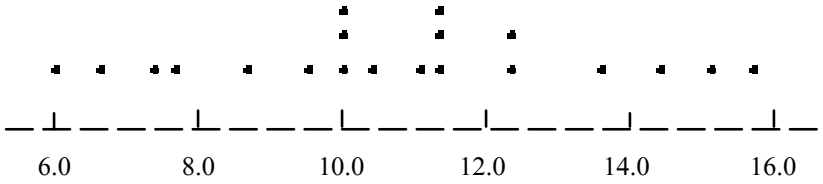
Class	Cumulative Frequency	Cumulative Relative Frequency
less than or equal to 19	10	.20
less than or equal to 29	24	.48
less than or equal to 39	41	.82
less than or equal to 49	48	.96
less than or equal to 59	50	1.00

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. 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

16. a.

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

b.

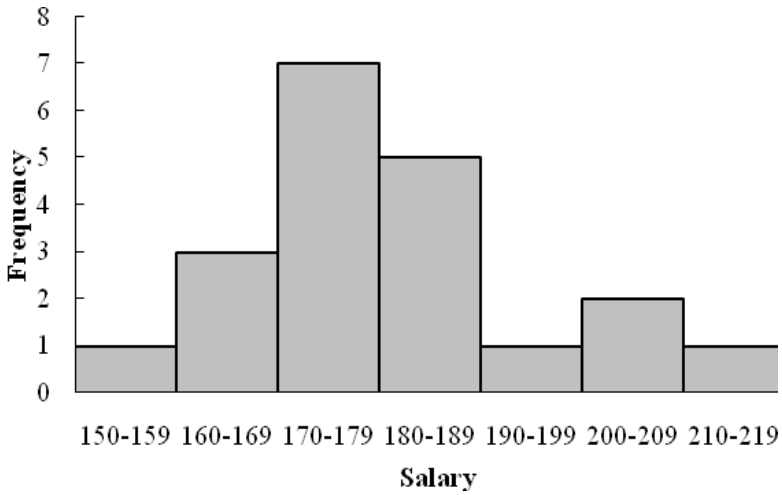
Salary	**Percent Frequency**
150-159	5
160-169	15
170-179	35
180-189	25
190-199	5

200-209	10
210-219	5
Total	100

c.

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

d.



e. There is skewness to the right.

f. (3/20)(100) = 15%

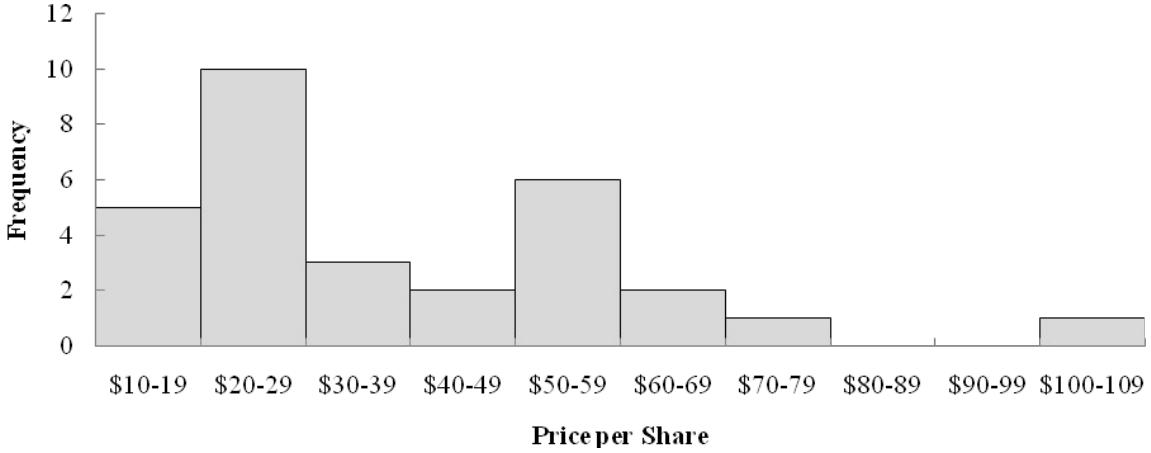
17. a. The highest price stock is for IBM with a price of $109 per share.
The lowest price stock is for Alcoa with a price of $11 per share.

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

Price per Share	Frequency
$10-19	5
$20-29	10
$30-39	3
$40-49	2
$50-59	6
$60-69	2
$70-79	1
$80-89	0
$90-99	0

$100-109	1

c.



The general shape of the distribution is skewed to the right. Half of the companies (15) have a price per share less than $30. A mid-priced stock appears to be in the $30 to $49 range, while the most frequently priced stock is in the $20 to $29 range.

Five stocks are less than $20 per share (Alcoa, Bank of America, General Electric, Intel and Pfizer). Four stocks are $60 or more per share (3M, Chevron, ExxonMobil and IBM).

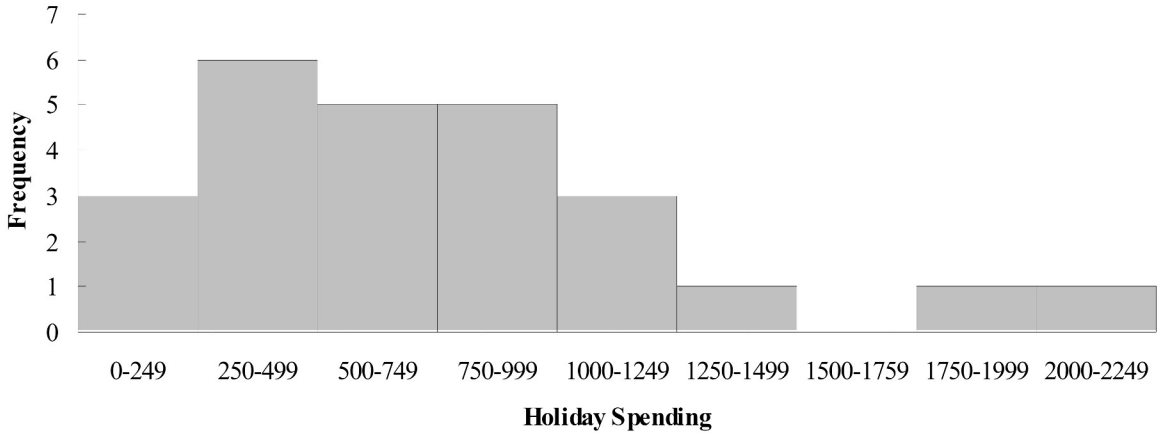
d. A variety of comparisons are possible depending upon when the study is done.

18. a. The lowest holiday spending is $180; the highest $2050.

b.

Spending	Frequency	Percent
0-249	3	12
250-499	6	24
500-749	5	20
750-999	5	20
1000-1249	3	12
1250-1499	1	4
1500-1759	0	0
1750-1999	1	4
2000-2249	1	4
Total	25	100

c. The distribution shows a positive skewness.



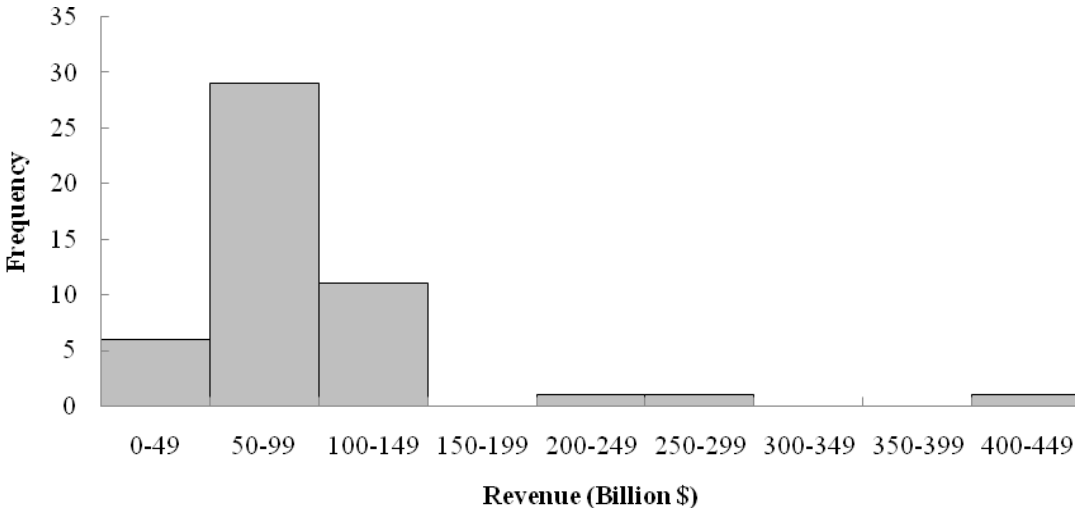
d. The holiday spending ranges from $0 to less than $2250. The majority of the spending is between $250 and $1000 with 16 of the 25 customers, 64%, in this range. The middle or average spending is around $750 per customer. The distribution has a positive skewness with two consumers above $1750. One consumer is above $2000.

19. 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.



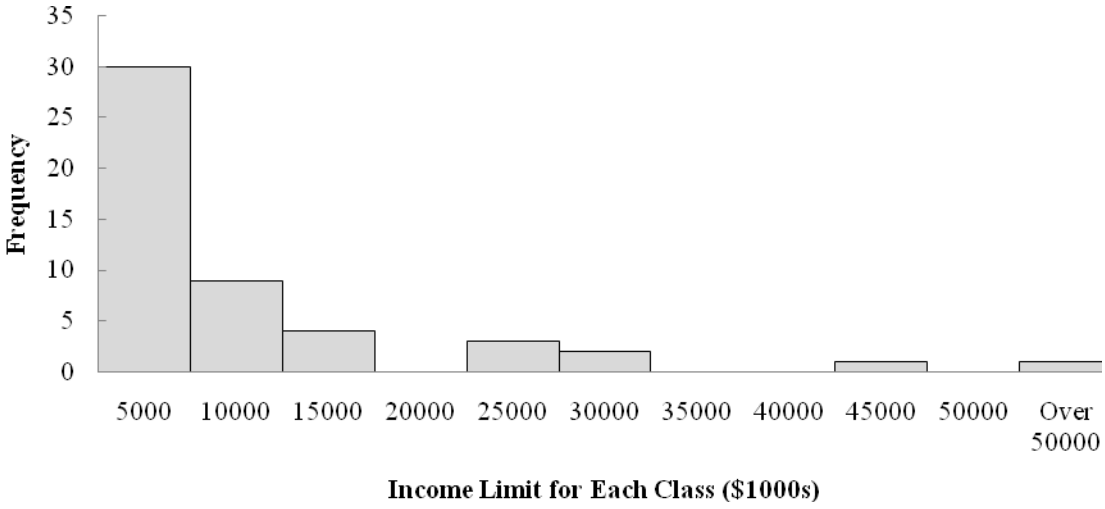
The histogram shows the distribution is skewed to the right with four corporations in the $200 to $449 billion range.

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.

20. a.

Off-Course Income ($1000s)	Frequency	Percent Frequency
0-4,999	30	60
5,000-9,999	9	18
10,000-14,999	4	8
15,000-19,999	0	0
20,000-24,999	3	6
25,000-29,999	2	4
30,000-34,999	0	0
35000-39,999	0	0
40,000-44,999	1	2
45,000-49,999	0	0
Over 50,000	1	2
Total	50	100

b. Histogram of Off-Course Income



Note: The first class is labeled 5000 and provides the golfers who had an off-course income in the range 0 to 4999 or *less than* 5000. These were the golfers with less than $5 million in off-course income.

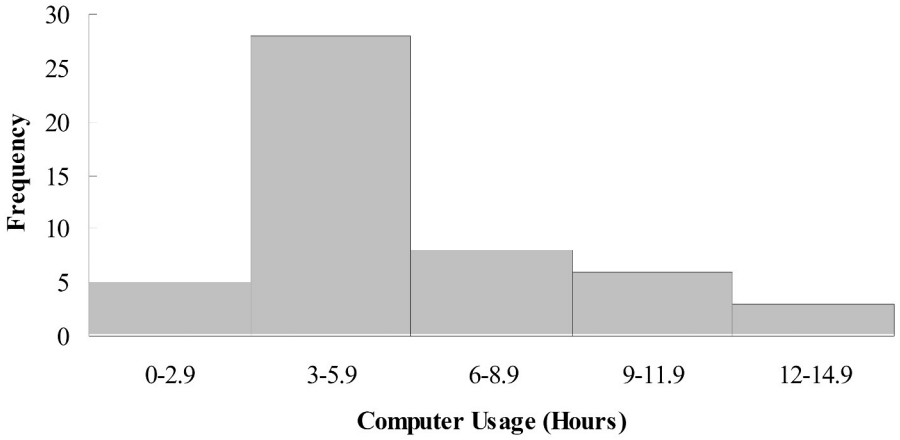
c. Off-course income is skewed to the right. Only Tiger Woods earns over $50 million.

d. Considering the top 50 golfers, the majority (60%) earn less than $5 million in off-course income per year. 60% + 18% = 78% earn less than $10 million. Five golfers (10%) earn between $20 million and $30 million. Tiger Woods with $99.8 million and Phil Mickelson with $40.2 million in off-course income are clearly the leaders in this income category.

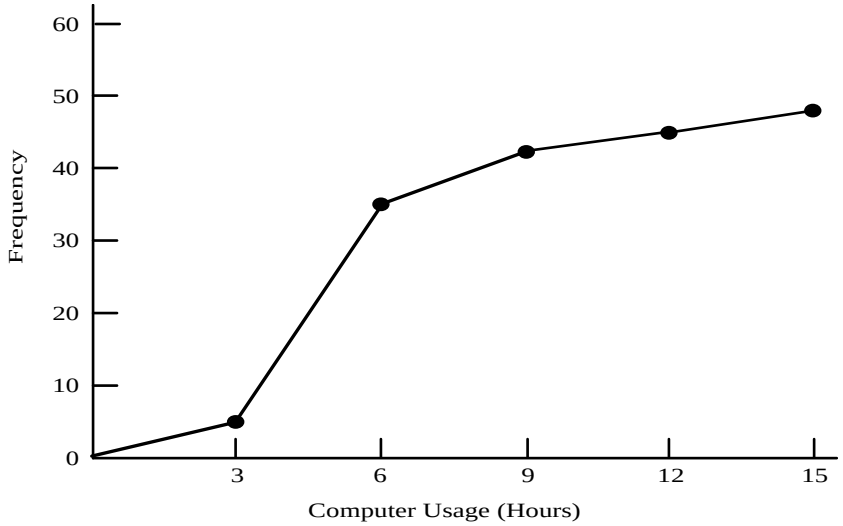
21. a/b.

Computer Usage (Hours)	Frequency	Relative Frequency
0.0 - 2.9	5	0.10
3.0 - 5.9	28	0.56
6.0 - 8.9	8	0.16
9.0 - 11.9	6	0.12
12.0 - 14.9	3	0.06
Total	50	1.00

c.



d.



e. The majority of the computer users are in the 3 to 6 hour range. Usage is somewhat skewed toward the right with 3 users in the 12 to 14.9 hour range.

22.

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

23. Leaf Unit = .1

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

24. Leaf Unit = 10

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

25.

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

26. 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.

27. a.

7	5 9
8	3 6
9	5 6 8
10	0 4 4
11	1 5
12	
13	7
14	5 5

b. Observations such as the following can be made using the stem-and-leaf display.

- The daily rate varies from $75 to $145
- Typical mid-priced daily rates are $95 to $115 with the average daily rate around $100.
- A daily rate in excess of $115 should be considered relatively high. High daily rates of $137 and $145 were found at three ski resorts.

28. 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

d. 4/40 = 10% of the runners were "20-something." With only 10% of the registrants "20-something," the article pointed out that surprisingly few registrants were in this age group. One suggested reason was that "20-somethings" don't have the time to train for a 13.1 mile race. For "20-somethings," college, starting careers, and starting families may take priority over training for long distance races.

29. a.

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

b.

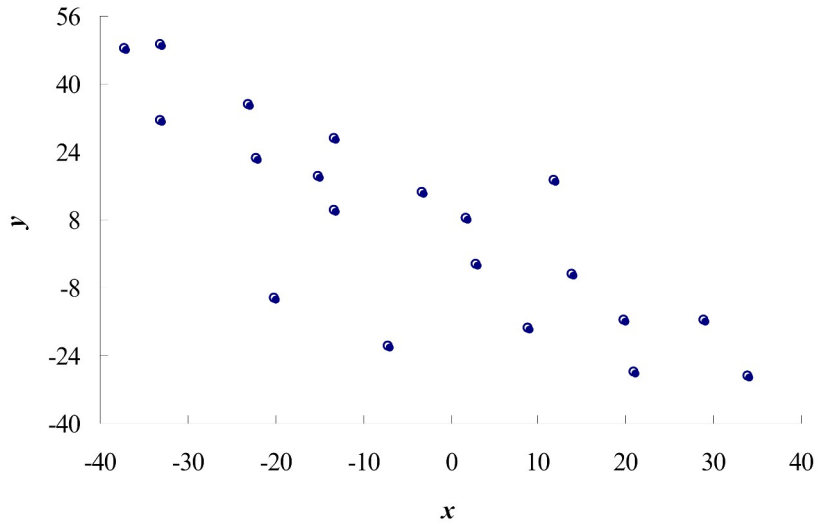
		y		
		1	2	Total
x	A	100.0	0.0	100.0
	B	84.6	15.4	100.0
	C	16.7	83.3	100.0

c.

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

d. 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*.

30. a.



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

31. a. Row Percentages:

Education Level	Household Income ($1000s)					Total
	Under 25	25.0-49.9	50.0-74.9	75.0-99.9	100 or More	
Not H.S. Graduate	42.23	34.73	13.94	5.41	3.68	100.00
H.S. Graduate	22.25	31.00	22.75	11.93	12.07	100.00
Some College	13.99	26.20	23.31	16.20	20.30	100.00
Bachelor's Degree	6.42	15.19	20.66	18.72	39.02	100.00
Beyond Bach. Deg.	3.71	10.60	16.29	15.87	53.54	100.00
Total	17.77	25.08	20.64	13.90	22.62	100.00

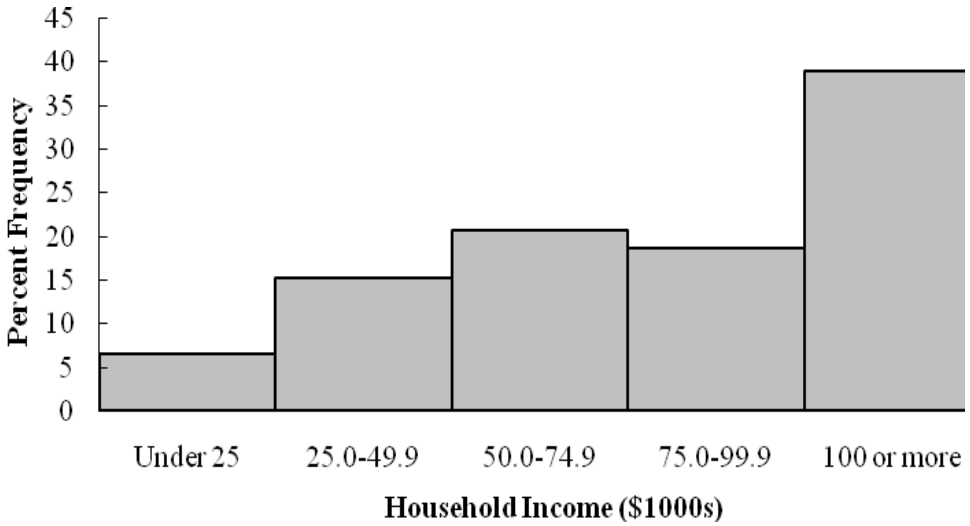
There are six percent frequency distributions in this table with row percentages. The first five give the percent frequency distribution of income for each educational level. The total row provides an overall percent frequency distribution for household income.

The second row, labeled H.S. Graduate, is the percent frequency distribution for households headed by high school graduates. The fourth row, labeled Bachelor's Degree, is the percent frequency distribution for households headed by bachelor's degree recipients.

b. The percentage of households headed by high school graduates earning $75,000 or more is 11.93% + 12.07 = 24.00%. The percent of households headed by bachelor's degree recipients earning $75,000 or more is 18.72% + 39.02% = 57.74%.

c. The percent frequency histogram for high school graduates.

The percent frequency distribution for college graduates with a bachelor’s degree.



The histograms show that households headed by a college graduate with a bachelor's degree earn more than households headed by a high school graduate. Yes, there is a positive relationship between education level and income.

32. a. Column Percentages:

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

There are six percent frequency distributions in this table of column percentages. The first five columns give the percent frequency distributions for each income level. The percent frequency distribution in the "Total" column gives the overall percent frequency distributions for educational level. From that percent frequency distribution we see that 13.51% of the heads of households did not graduate from high school.

b. The column percentages show that 25.11% of households earning $100,000 or more were headed by persons having schooling beyond a bachelor's degree. The row percentages show that 53.54% of the households headed by persons with schooling beyond a bachelor's degree earned $100,000 or more. These percentages are different because they came from different percent frequency distributions and provide different kinds of information.

c. Compare the "under 25" percent frequency distributions to the "Total" percent frequency distributions. We see that for this low income level the percentage with lower levels of education is higher than for the overall population and the percentage with higher levels of education is lower than for the overall population.

Compare the "100 or more" percent frequency distribution to "Total" percent frequency distribution. We see that for this high income level the percentage with lower levels of education is lower than for

the overall population and the percentage with higher levels of education is higher than for the overall population.

From the comparisons it is clear that there is a positive relationship between household incomes and the education level of the head of the household.

33. 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.

34. a.

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

b.

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

c.

Fund Type	Frequency
DE	27

FI	10
IE	8
Total	45

d. The right margin shows the frequency distribution for the fund type variable and the bottom margin shows the frequency distribution for the 5 year average return variable.

e. Higher returns are associated with International Equity funds and lower returns are associated with Fixed Income funds.

35. a.

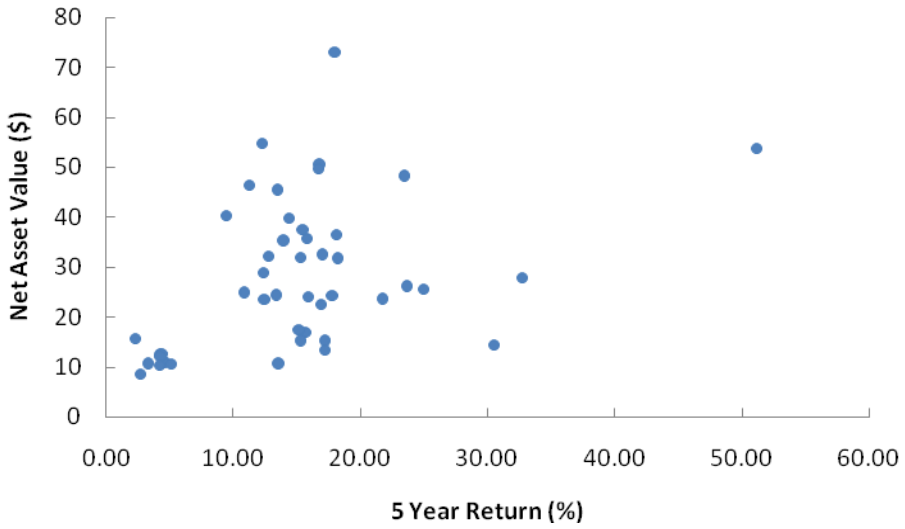
	Expense Ratio (%)						
Fund Type	0-0.24	0.25-0.49	0.50-0.74	0.75-0.99	1.00-1.24	1.25-1.49	Total
DE	1	1	3	5	10	7	27
FI	2	4	3	0	0	1	10
IE	0	0	1	2	4	1	8
Total	3	5	7	7	14	9	45

b.

Expense Ratio (%)	Frequency	Percent
0-0.24	3	6.7
0.25-0.49	5	11.1
0.50-0.74	7	15.6
0.75-0.99	7	15.6
1.00-1.24	14	31.0
1.25-1.49	9	20.0
Total	45	100

c. Higher expense ratios are associated with Domestic Equity funds and lower expense ratios are associated with Fixed Income fund

36. a. The scatter diagram is shown below:



b. There is some indication that higher 5-year returns are associated with higher net asset values.

37. a.

	Highway MPG					
Size	15-19	20-24	25-29	30-34	35-39	Total
Compact	26	76	9	0	0	111
Midsize	0	0	85	46	4	135
Large	0	0	65	0	0	65
Total	26	76	159	46	4	311

b. Higher fuel efficiencies are associated with midsize cars. In fact, for these data compact cars had the lowest fuel efficiencies.

c.

	City MPG						
Drive	5-9	10-14	15-19	20-24	25-29	30-35	Total
4	0	10	51	8	0	0	69
F	0	2	80	74	9	2	167
R	1	23	50	1	0	0	75
Total	1	35	181	83	9	2	311

d. Higher fuel efficiencies are associated with front wheel drive cars. Rear wheel drive cars had somewhat lower fuel efficiencies than four wheel drive cars.

e.

	City MPG	

Fuel Type	5-9	10-14	15-19	20-24	25-29	30-35	Total
P	1	33	105	18	0	0	157
R	0	2	76	65	9	2	154
Total	1	35	181	83	9	2	311

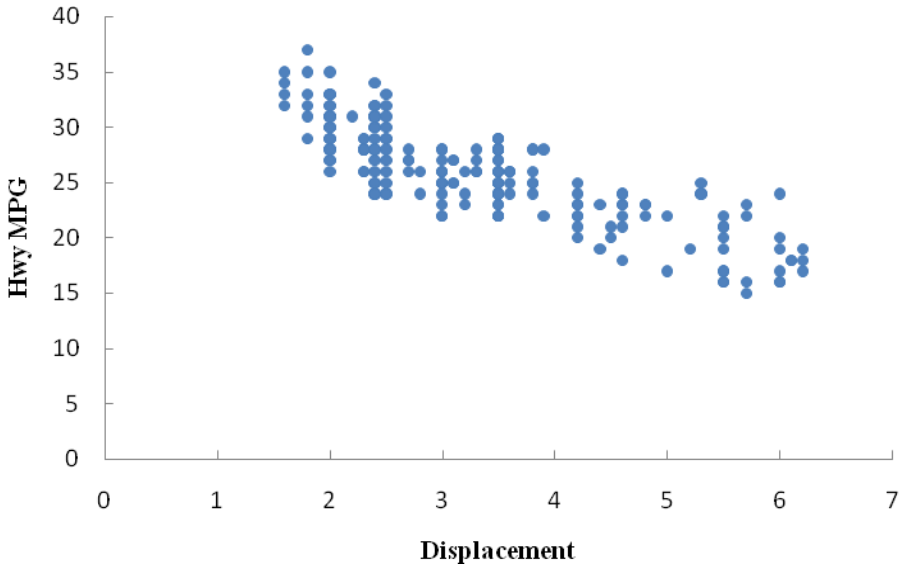
f. Higher fuel efficiencies are associated with cars that use regular fuel.

38. a.

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

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

c. The scatter diagram is shown below:



d. The scatter diagram shows that lower fuel efficiencies are associated with larger displacement engines.

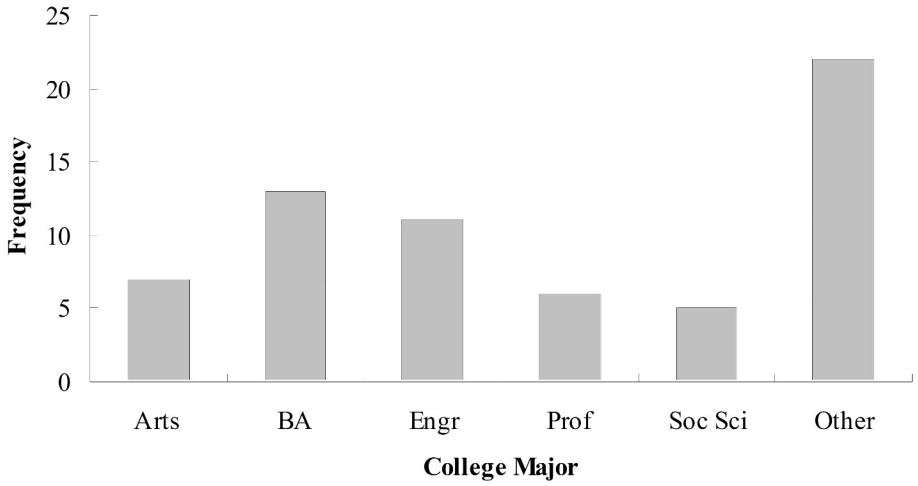
e. It is easier to see the relationship between the two variables using the scatter diagram.

39. a.

Major	Frequency	Percent Frequency
Arts/Humanities	7	10.9

Business Administration	13	20.3
Engineering	11	17.2
Professional	6	9.4
Social Science	5	7.8
Other	22	34.4
Total	64	100.0

b.



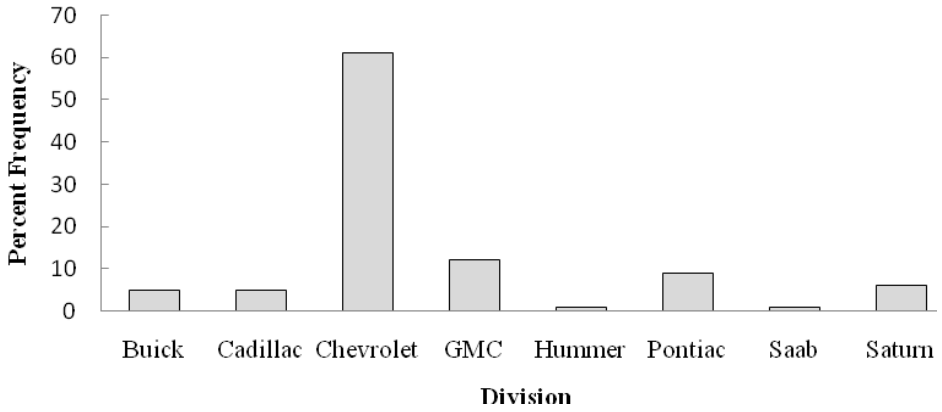
c. 34.4% select another major. So 100% - 34.4% = 65.6% select one of the five most popular majors.

d. Business Administration is the most popular major selected by incoming freshmen, 20.3%

40. a. Frequency distribution and percent frequency distribution of sales by division.

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

b.



c. Chevrolet is General Motors leading division with 61% of the vehicles sold. This is considered General Motors most important division.

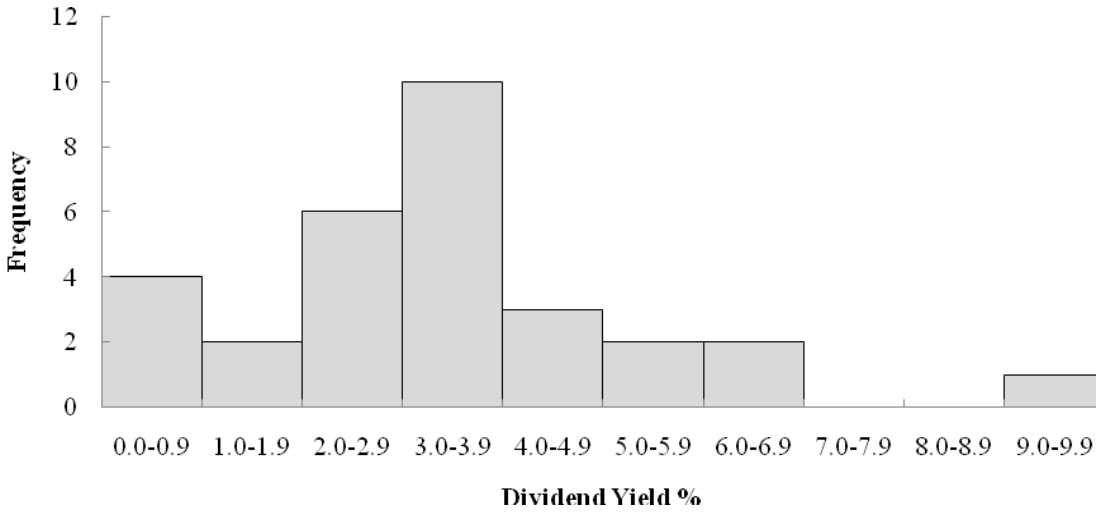
d. Based on the percentages shown, the Hummer division at 1% and Saab division at 1% would be good candidates for General Motors to consider discontinuing. Chevrolet at 61% and GMC at 12% account for 73% of the total vehicles sold. General Motors would be almost certain to maintain these two divisions.

Pontiac remains a solid contributor with 9% of vehicles sold. At the time it was doubtful than General Motors would be able to maintain all three of the other divisions. Some elimination or merging of divisions was anticipated for Saturn 6%, Buick 5%, and Cadillac 5%.

41. 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.



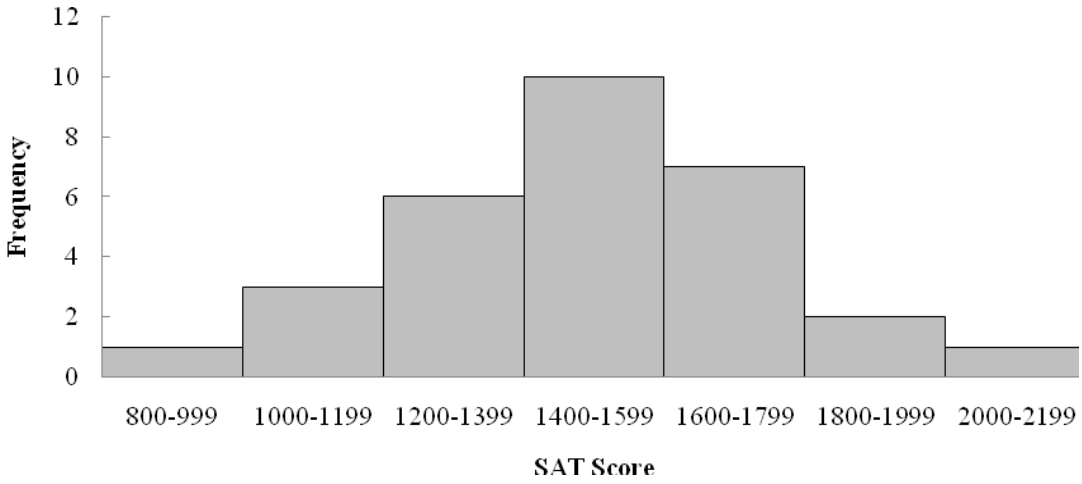
c. The distribution is skewed to the right.

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.

42. 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

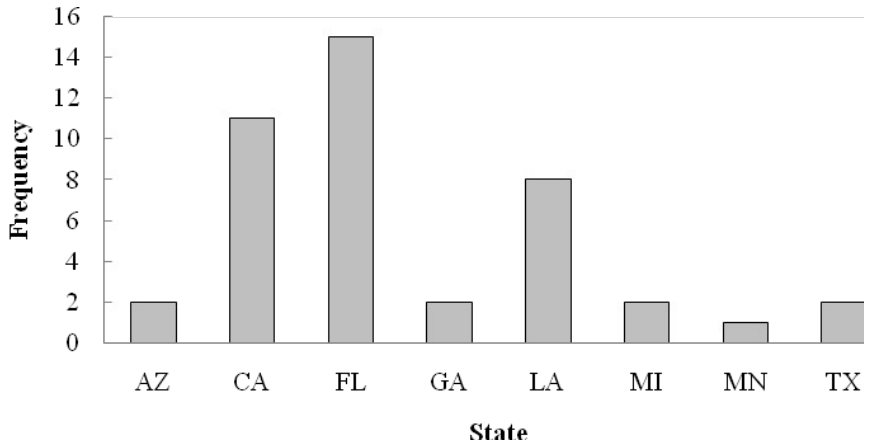


b. The distribution if nearly symmetrical. It could be approximated by a bell-shaped curve.

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.

43. a.

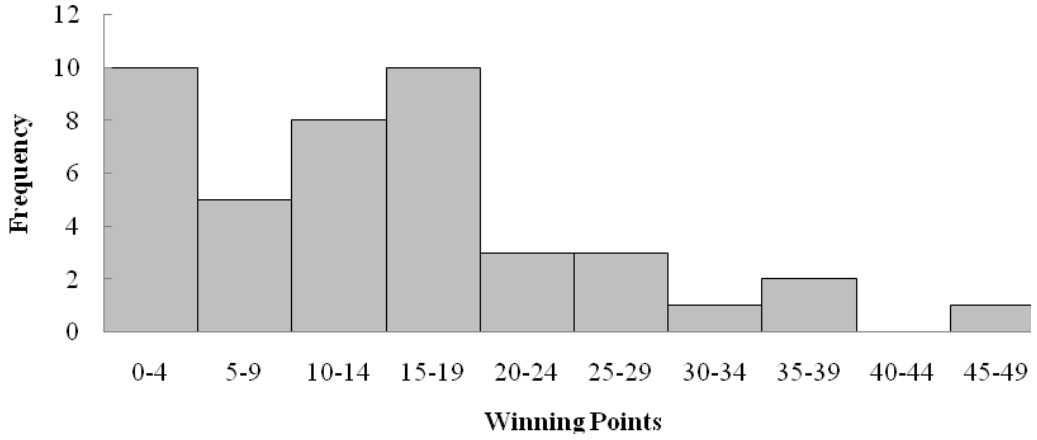
State	Frequency
Arizona	2
California	11
Florida	15
Georgia	2
Louisiana	8
Michigan	2
Minnesota	1
Texas	2
Total	43



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.

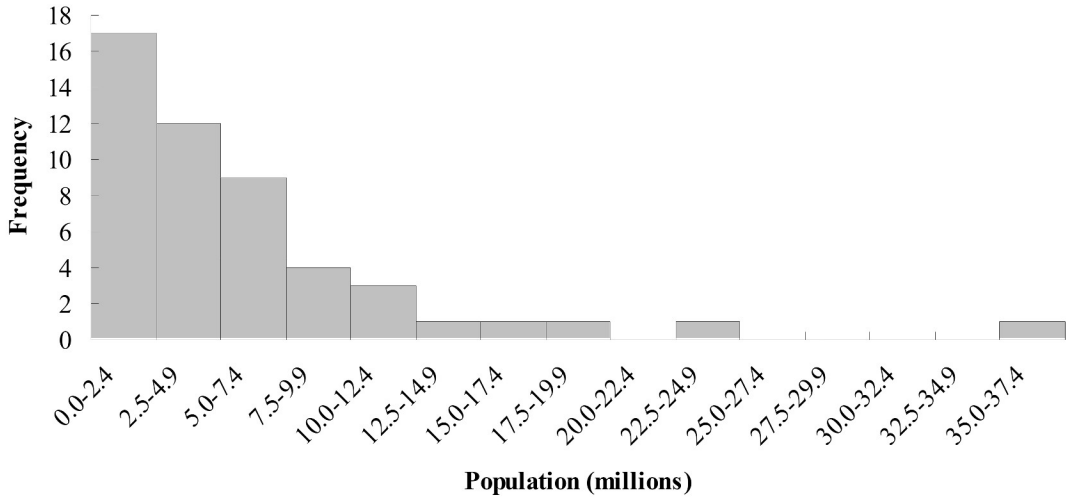
25th Super Bowl: 1991 New York Giants 20 Buffalo Bills 19, Tampa Stadium, Tampa, FL
24th Super Bowl: 1990 San Francisco 49ers 55 Denver Broncos 10, Superdome, New Orleans, LA

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.

44. a.

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

b.



c. High positive skewness.

d. 17 states (34%) have a population less than 2.5 million. Over half of the states have population less than 5 million (29 states – 58%). Only eight states have a population greater than 10 million (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania and Texas). The largest state is California (35.9 million) and the smallest state is Wyoming (500 thousand).

45. a.

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

b. Smallest roughly $3 billion or less; medium $7-$8 billion; largest $11-$12 billion.

c. CVS ($12,700) and Walgreens ($11,660)

46. a& b.

High Temperature		Low Temperature	
1		1	1
2		2	1 2 6 7 9
3	0	3	1 5 6 8 9
4	1 2 2 5	4	0 3 3 6 7
5	2 4 5	5	0 0 4
6	0 0 0 1 2 2 5 6 8	6	5
7	0 7	7	
8	4	8	

c. The most frequent range for temperature was in the 60s (9 of 20). Only one low temperature was above 54. High temperatures were mostly 41 to 68, while low temperatures were mostly 21 to 47.

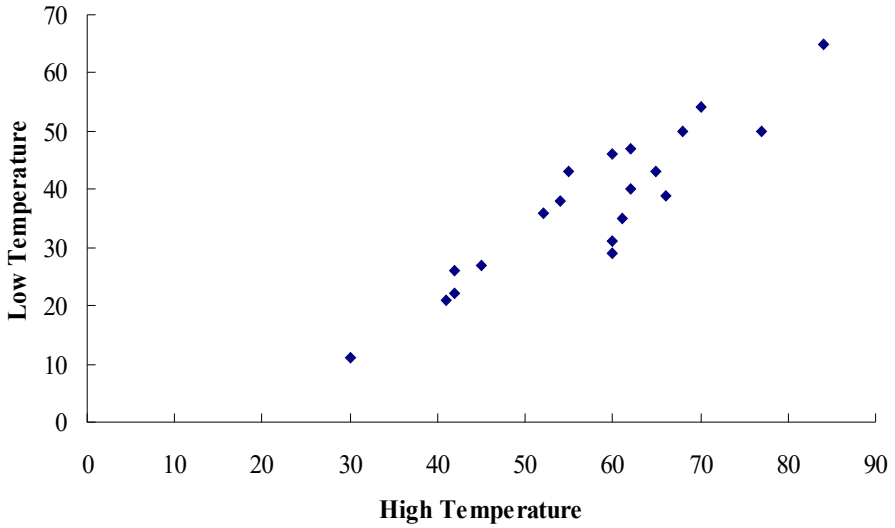
Low was 11; High was 84.

d.

High Temp	Frequency	Low Temp	Frequency

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

47. a.



b. There is a positive relationship between high temperature and low temperature for these cities. As one goes up so does the other.

48. a.

Level of Support	Percent Frequency
Strongly favor	1617/5372 = 30.10
Favor more than oppose	1871/5372 = 34.83
Oppose more than favor	1135/5372 = 21.13
Strongly oppose	749/5372 = 13.94
Total	100.00

The results show support for a higher tax. Note that 30.10% + 34.83% = 64.93% of the respondents said they strongly favor or favor more than oppose a higher tax on higher carbon emission cars.

b.

Country	Percent Frequency
Great Britain	1087/5372 = 20.2
Italy	1045/5372 = 19.5
Spain	1109/5372 = 20.6
Germany	1111/5372 = 20.7
United States	1020/5372 = 19.0
Total	100.0

The poll had an approximately equal representation of the five countries with roughly 20% of the poll respondents coming from each country.

c. Converting the entries in the crosstabulation into column percentages provides the following results:

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

Considering the percentage of respondents who favor the higher tax by either saying "strongly favor" or "favor more than oppose", we have the following favorable support for the higher tax in each country.

Great Britain	31.00 + 34.04 = 65.04%
Italy	31.96 + 39.04 = 71.00%
Spain	45.99 + 32.01 = 78.00%
Germany	19.98 + 36.99 = 56.97%
United States	20.98 + 32.06 = 53.04%

More that 50% of the respondents favor the higher tax for the higher carbon emission cars in all five countries. But the support for the higher tax is greater in the European countries. Spain and Italy have the greatest support for the higher tax with 78% and 71% respectively. Germany is close in views to the United States with 56.97% expressing favor for the higher tax. United States shows the lowest level of support for the higher tax with 53.04%. Note that United States ranks first in terms of the response "strongly oppose" the higher tax with 20% of the respondents providing this opinion.

49. 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).

50. a.

		Fuel Type				
Year Constructed	Elec	Nat. Gas	Oil	Propane	Other	Total
1973 or before	40	183	12	5	7	247
1974-1979	24	26	2	2	0	54
1980-1986	37	38	1	0	6	82
1987-1991	48	70	2	0	1	121
Total	149	317	17	7	14	504

b.

Year Constructed	Frequency	Fuel Type	Frequency
1973 or before	247	Electricity	149
1974-1979	54	Nat. Gas	317
1980-1986	82	Oil	17
1987-1991	121	Propane	7
Total	504	Other	14
		Total	504

c. Crosstabulation of Column Percentages

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

d. Crosstabulation of row percentages.

	Fuel Type

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

e. Observations from the column percentages crosstabulation

For those buildings using electricity, the percentage has not changed greatly over the years. For the buildings using natural gas, the majority were constructed in 1973 or before; the second largest percentage was constructed in 1987-1991. Most of the buildings using oil were constructed in 1973 or before. All of the buildings using propane are older.

Observations from the row percentages crosstabulation

Most of the buildings in the CG&E service area use electricity or natural gas. In the period 1973 or before most used natural gas. From 1974-1986, it is fairly evenly divided between electricity and natural gas. Since 1987 almost all new buildings are using electricity or natural gas with natural gas being the clear leader.

51. a. Crosstabulation for stockholder's equity and profit.

	Profits ($000)						
Stockholders' Equity ($000)	0-200	200-400	400-600	600-800	800-1000	1000-1200	Total
0-1200	10	1				1	12
1200-2400	4	10			2		16
2400-3600	4	3	3	1	1	1	13
3600-4800					1	2	3
4800-6000		2	3	1			6
Total	18	16	6	2	4	4	50

b. Crosstabulation of Row Percentages.

	Profits ($000)						
Stockholders' Equity ($1000s)	0-200	200-400	400-600	600-800	800-1000	1000-1200	Total
0-1200	83.33	8.33	0.00	0.00	0.00	8.33	100
1200-2400	25.00	62.50	0.00	0.00	12.50	0.00	100
2400-3600	30.77	23.08	23.08	7.69	7.69	7.69	100
3600-4800		0.00	0.00	0.00	33.33	66.67	100
4800-6000	0.00	33.33	50.00	16.67	0.00	0.00	100

c. Stockholder's equity and profit seem to be related. As profit goes up, stockholder's equity goes up. The relationship, however, is not very strong.

52. a. Crosstabulation of market value and profit.

	Profit ($1000s)				
Market Value ($1000s)	0-300	300-600	600-900	900-1200	Total

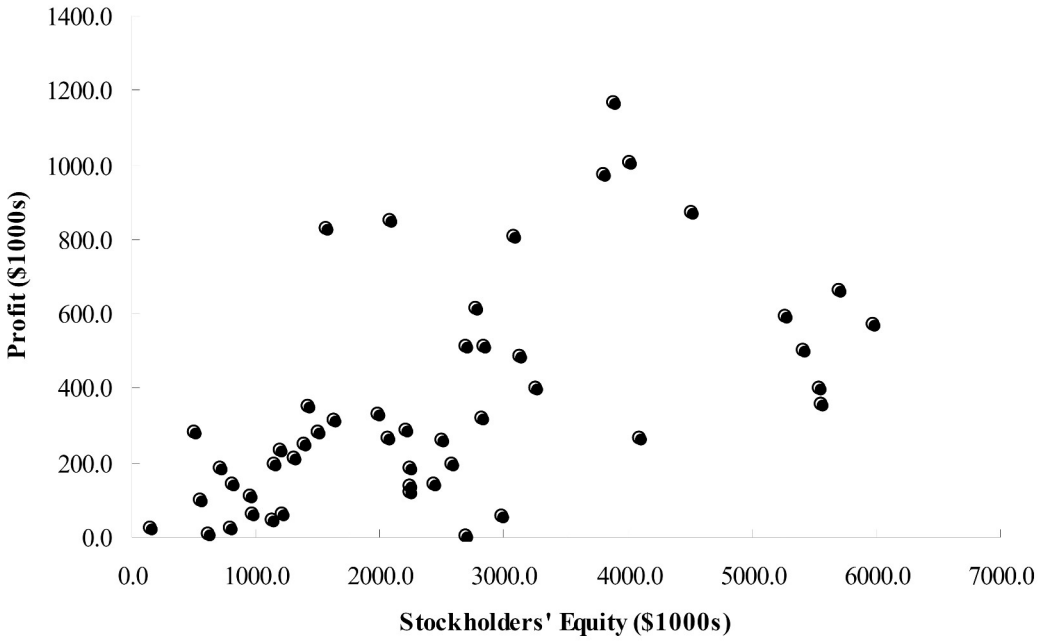
0-8000	23	4			27
8000-16000	4	4	2	2	12
16000-24000		2	1	1	4
24000-32000		1	2	1	4
32000-40000		2	1		3
Total	27	13	6	4	50

b. Crosstabulation of Row Percentages.

	Profit ($1000s)				
Market Value ($1000s)	0-300	300-600	600-900	900-1200	Total
0-8000	85.19	14.81	0.00	0.00	100
8000-16000	33.33	33.33	16.67	16.67	100
16000-24000	0.00	50.00	25.00	25.00	100
24000-32000	0.00	25.00	50.00	25.00	100
32000-40000	0.00	66.67	33.33	0.00	100

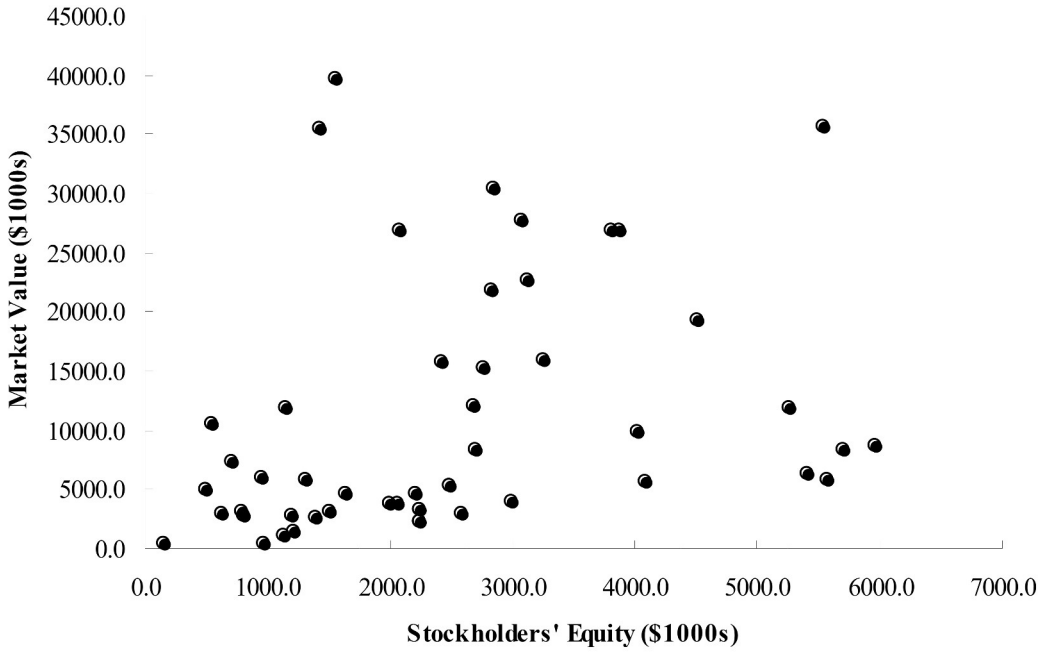
c. There appears to be a positive relationship between Profit and Market Value. As profit goes up, Market Value goes up.

53. a. Scatter diagram of Profit vs. Stockholders' Equity.



b. Profit and Stockholders' Equity appear to be positively related.

54. a. Scatter diagram of Market Value and Stockholders' Equity.



b. There is a positive relationship between Market Value and Stockholders' Equity.