

Chapter 2 - Frequency Distributions and Graphs

EXERCISE SET 2-2

1. Frequency distributions are used to organize data in a meaningful way, to facilitate computational procedures for statistics, to make it easier to draw charts and graphs, and to make comparisons among different sets of data.

2. Categorical distributions are used with nominal or ordinal data, ungrouped distributions are used with data having a small range, and grouped distributions are used when the range of the data is large.

3.

a. $11.5 - 18.5, \frac{12+18}{2} = \frac{30}{2} = 15,$

$18.5 - 11.5 = 7$

b. $55.5 - 74.5, \frac{56+74}{2} = \frac{130}{2} = 65,$

$74.5 - 55.5 = 19$

c. $694.5 - 705.5, \frac{695+705}{2} = \frac{1400}{2} = 700,$

$705.5 - 694.5 = 11$

d. $13.55 - 14.75, \frac{13.6+14.7}{2} = \frac{28.3}{2} = 14.15,$

$14.75 - 13.55 = 1.2$

e. $2.145 - 3.935, \frac{2.15+3.93}{2} = \frac{6.08}{2} = 3.04,$

$3.935 - 2.145 = 1.79$

4. Five to twenty classes. Width should be an odd number so that the midpoint will have the same place value as the data.

5.

a. Class width is not uniform.

b. Class limits overlap, and class width is not uniform.

c. A class has been omitted.

d. Class width is not uniform.

6. An open-ended frequency distribution has either a first class with no lower limit or a last class with no upper limit. They are necessary to accommodate all the data.

7.	Class	Tally	f	Percent
	A		4	10%
	M		28	70%
	H		6	15%
	S		2	5%
			40	100%

8. $H = 36 \quad L = 7$

Range = $36 - 7 = 29$

Width = $29 \div 6 = 4.83$ or 5

Limits	Boundaries	f	cf
7 - 11	6.5 - 11.5	2	2
12 - 16	11.5 - 16.5	5	7
17 - 21	16.5 - 21.5	9	16
22 - 26	21.5 - 26.5	2	18
27 - 31	26.5 - 31.5	0	18
32 - 36	31.5 - 36.5	<u>1</u>	19

9.	Limits	Boundaries	f	cf
	19 - 21	18.5 - 21.5	2	2
	22 - 24	21.5 - 24.5	13	15
	25 - 27	24.5 - 27.5	11	26
	28 - 30	27.5 - 30.5	3	29
	31 - 33	30.5 - 33.5	<u>1</u>	30

The average speed is about 24.5 mph.

10. $H = 11,588 \quad L = 164$

Range = $11,588 - 164 = 11,424$

Width = $11,424 \div 10 = 1142.4$

Width = 1143

Limits	Boundaries	f	cf
164 - 1306	163.5 - 1306.5	25	25
1307 - 2449	1306.5 - 2449.5	12	37
2450 - 3592	2449.5 - 3592.5	5	42
3593 - 4735	3592.5 - 4735.5	5	47
4736 - 5878	4735.5 - 5878.5	1	48
5879 - 7021	5878.5 - 7021.5	0	48
7022 - 8164	7021.5 - 8164.5	0	48
8165 - 9307	8164.5 - 9307.5	1	49
9308 - 10450	9307.5 - 10450.5	0	49
10451 - 11593	10450.5 - 11593.5	<u>1</u>	50

The majority of the data is clustered in the first two classes.

11. $H = 780 \quad L = 746$

Range = $780 - 746 = 34$

Width = $34 \div 6 = 5.\bar{6}$ or 6; round up to 7

Chapter 2 - Frequency Distributions and Graphs

11. continued

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
745 - 751	744.5 - 751.5	4	4
752 - 758	751.5 - 758.5	5	9
759 - 765	758.5 - 765.5	7	16
766 - 772	765.5 - 772.5	11	27
773 - 779	772.5 - 779.5	2	29
780 - 786	779.5 - 786.5	<u>1</u>	30
		30	

12. H = 93 L = 0

Range = $93 - 0 = 93$

Width = $93 \div 7 \approx 13.29$ or 14

Use w = 15 for odd number.

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
0 - 14	-0.5 - 14.5	14	14
15 - 29	14.5 - 29.5	10	24
30 - 44	29.5 - 44.5	4	28
45 - 59	44.5 - 59.5	1	29
60 - 74	59.5 - 74.5	1	30
75 - 89	74.5 - 89.5	2	32
90 - 104	89.5 - 104.5	<u>1</u>	33
		33	

13. H = 70 L = 27

Range = $70 - 27 = 43$

Width = $43 \div 7 = 6.1$ or 7

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
27 - 33	26.5 - 33.5	7	7
34 - 40	33.5 - 40.5	14	21
41 - 47	40.5 - 47.5	15	36
48 - 54	47.5 - 54.5	11	47
55 - 61	54.5 - 61.5	3	50
62 - 68	61.5 - 68.5	3	53
69 - 75	68.5 - 75.5	<u>2</u>	55
		55	

14. H = 4040 L = 70

Range = $4040 - 70 = 3970$

Width = $3970 \div 8 = 496.25$ or 497

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
70 - 566	69.5 - 566.5	14	14
567 - 1063	566.5 - 1063.5	5	19
1064 - 1560	1063.5 - 1560.5	5	24
1561 - 2057	1560.5 - 2057.5	0	24
2058 - 2554	2057.5 - 2554.5	0	24
2555 - 3051	2554.5 - 3051.5	1	25
3052 - 3548	3051.5 - 3548.5	0	25
3549 - 4045	3548.5 - 4045.5	<u>2</u>	27
		27	

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
31 - 39	30.5 - 39.5	4	4
40 - 48	39.5 - 48.5	5	9
49 - 57	48.5 - 57.5	5	14
58 - 66	57.5 - 66.5	12	26
67 - 75	66.5 - 75.5	13	39
76 - 84	75.5 - 84.5	5	44
85 - 93	84.5 - 93.5	<u>3</u>	47
		47	

16. H = 775 L = 5

Width = $775 - 5 = 770$

Range = $770 \div 8 = 96.25$ or 97

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
5 - 101	4.5 - 101.5	17	17
102 - 198	101.5 - 198.5	6	23
199 - 295	198.5 - 295.5	6	29
296 - 392	295.5 - 392.5	2	31
393 - 489	392.5 - 489.5	2	33
490 - 586	489.5 - 586.5	3	36
587 - 683	586.5 - 683.5	1	37
684 - 780	683.5 - 780.5	<u>2</u>	39
		39	

17. H = 11,413 L = 150

Range = $11,413 - 150 = 11,263$

Width = $11,263 \div 10 = 1126.3$ or 1127

<i>Limits</i>	<i>Boundaries</i>	<i>f</i>	<i>cf</i>
150 - 1276	149.5 - 1276.5	2	2
1277 - 2403	1276.5 - 2403.5	2	4
2404 - 3530	2403.5 - 3530.5	5	9
3531 - 4657	3530.5 - 4657.5	8	17
4658 - 5784	4657.5 - 5784.5	7	24
5785 - 6911	5784.5 - 6911.5	3	27
6912 - 8038	6911.5 - 8038.5	7	34
8039 - 9165	8038.5 - 9165.5	3	37
9166 - 10,292	9165.5 - 10,292.5	3	40
10,293 - 11,419	10,292.5 - 11,419.5	<u>2</u>	42
		42	

18.

H = 550 L = 306

Range = $550 - 306 = 244$

Width = $244 \div 8 = 30.5$ or 31

Chapter 2 - Frequency Distributions and Graphs

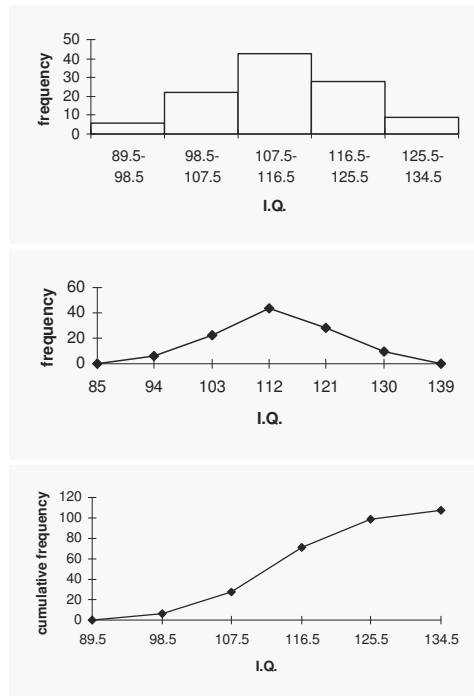
18. continued

$f_M, cf_M = \text{McGwire}$		$f_S, cf_S = \text{Sosa}$			
<i>Limits</i>	<i>Boundaries</i>	f_M	cf_M	f_S	cf_S
306 - 336	305.5 - 336.5	1	1	0	0
337 - 367	336.5 - 367.5	6	7	10	10
368 - 398	367.5 - 398.5	19	26	16	26
399 - 429	398.5 - 429.5	15	41	21	47
430 - 460	429.5 - 460.5	18	59	15	62
461 - 491	460.5 - 491.5	6	65	3	65
492 - 522	491.5 - 522.5	3	68	1	66
523 - 553	522.5 - 553.5	<u>2</u>	70	<u>0</u>	66
		70		66	

19. The percents add up to 101%. They should total 100% unless rounding was used.

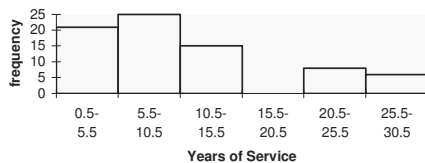
EXERCISE SET 2-3

1.



Eighty applicants do not need to enroll in the summer programs.

2.

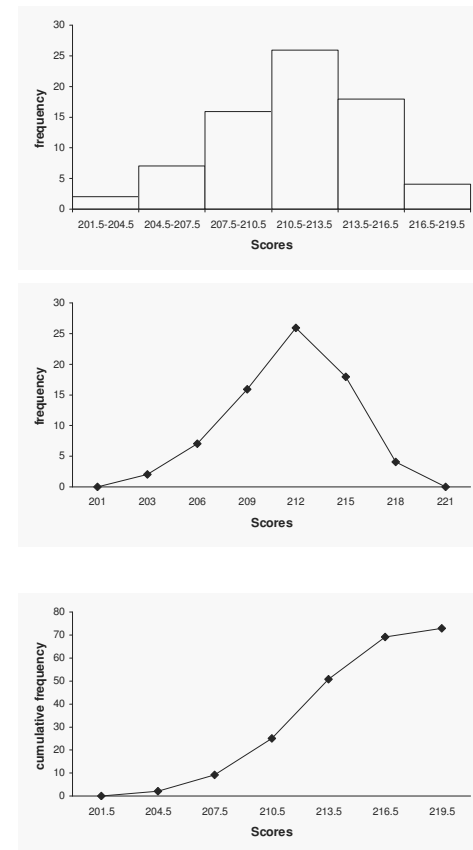


2. continued



The majority of employees have worked for less than 11 years.

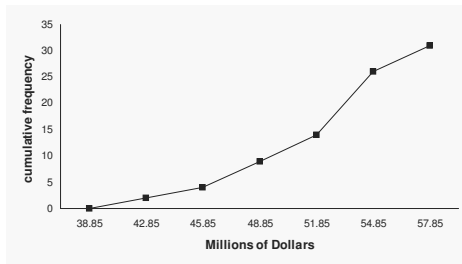
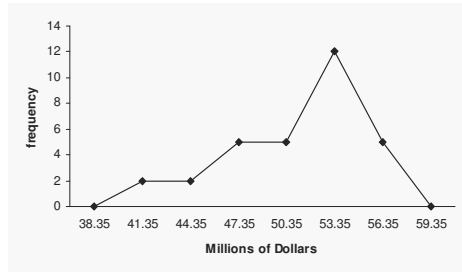
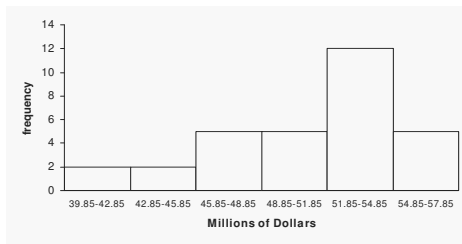
3.



The distribution appears to be slightly left skewed.

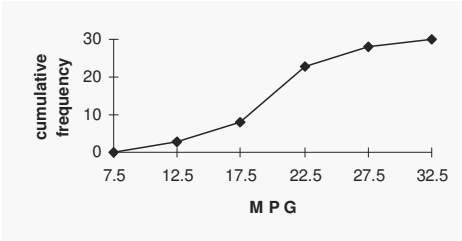
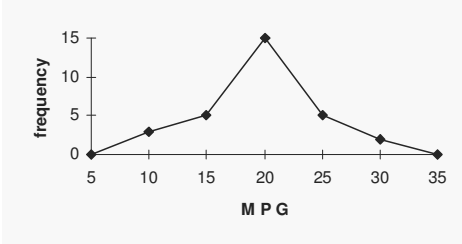
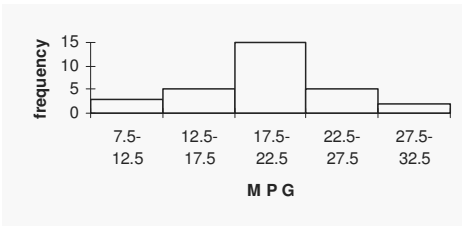
Chapter 2 - Frequency Distributions and Graphs

4.

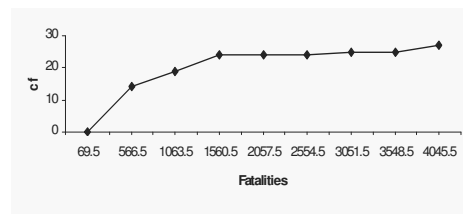
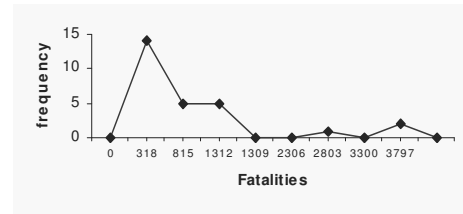


The distribution is left skewed or negatively skewed.

5.

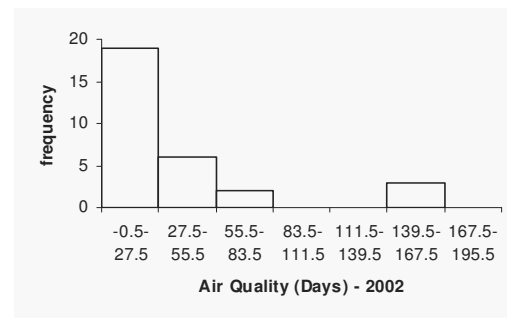
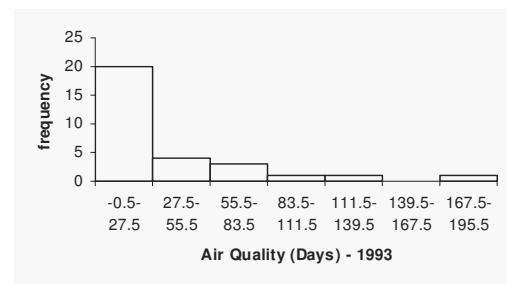


6.



The data are clustered to the left. There are two gaps in the histogram, one between 1560.5 and 2554.5 and the other between 3051.5 and 3548.5. The ogive shows a sharp increase in values at the beginning, followed by a leveling off of values.

7.

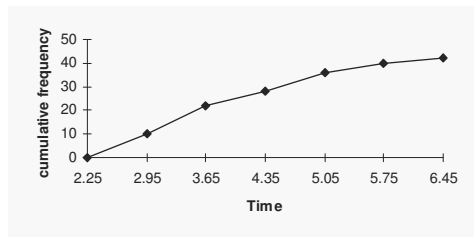
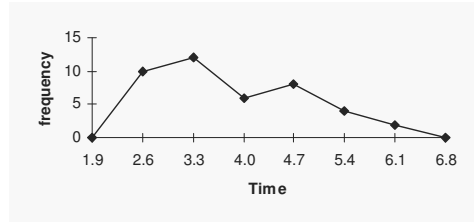
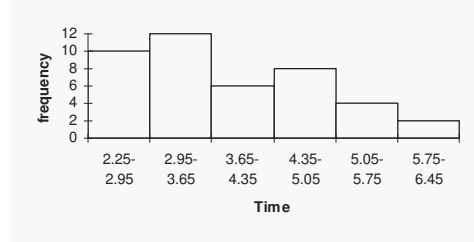


Chapter 2 - Frequency Distributions and Graphs

7. continued

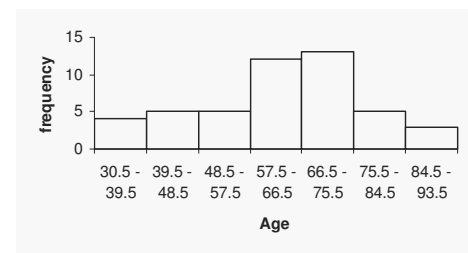
Both graphs are similar in that they are positively skewed. Also, it looks as if the air quality has improved somewhat in that there are slightly more smaller values in 2002, which means fewer days with unacceptable levels of pollution.

8.

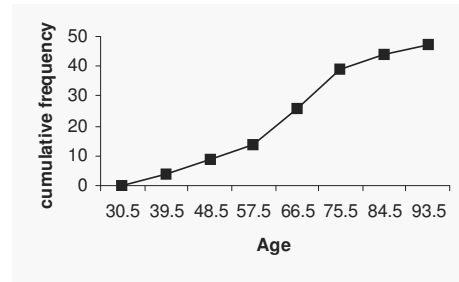
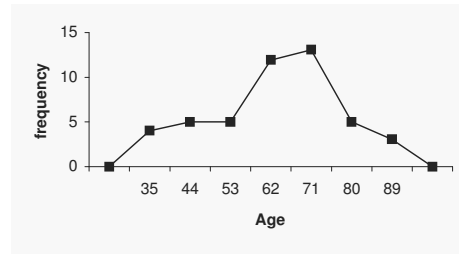


The data values fall somewhat on the left side of the distribution. The histogram is right skewed. There are no gaps in the histogram.

9.

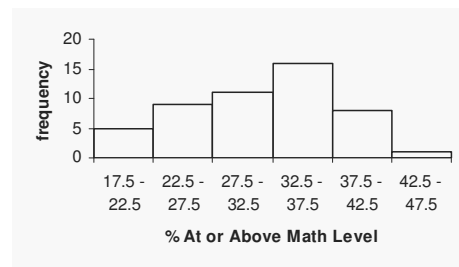
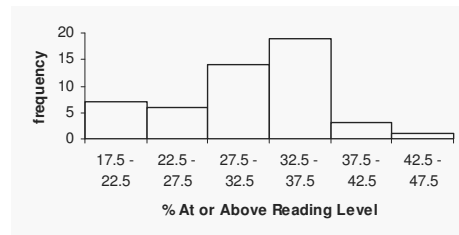


9. continued



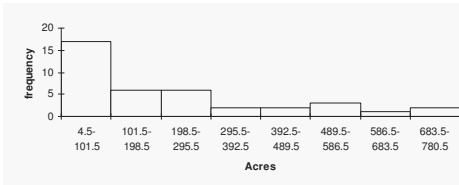
The histogram has a peak at the class of 66.5 – 75.5 and is somewhat negatively skewed.

10.



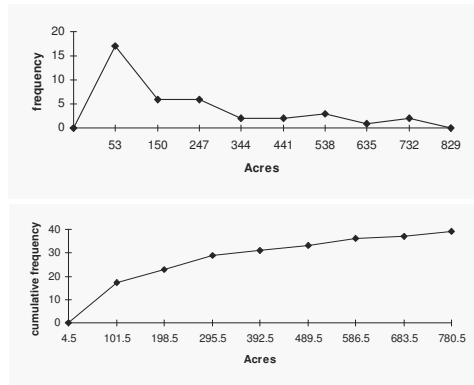
The distribution of math percentages is more bell-shaped than the distribution of reading percentages, and its peak in the class of 32.5 – 37.5 is not as high as the peak of the reading percentages.

11.



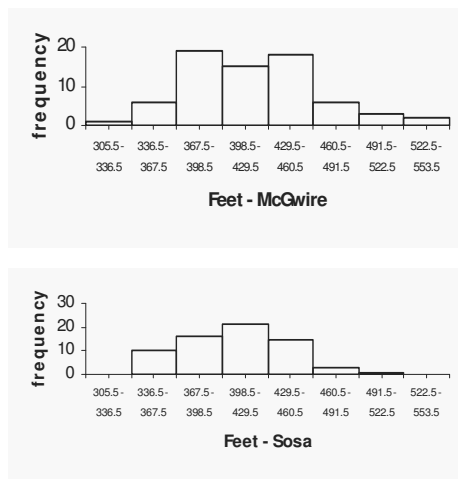
Chapter 2 - Frequency Distributions and Graphs

11. continued



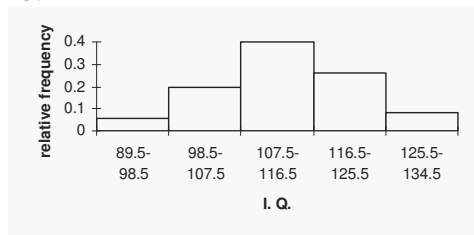
The peak is in the first class, and then the histogram is rather uniform after the first class. Most of the parks have less than 101.5 thousand acres as compared with any other class of values.

12.

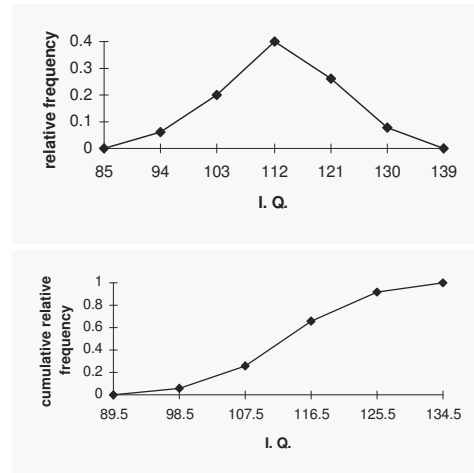


The histograms show that the distances of McGwire's homeruns are more variable (spread out) than Sosa's homerun distances.

13.

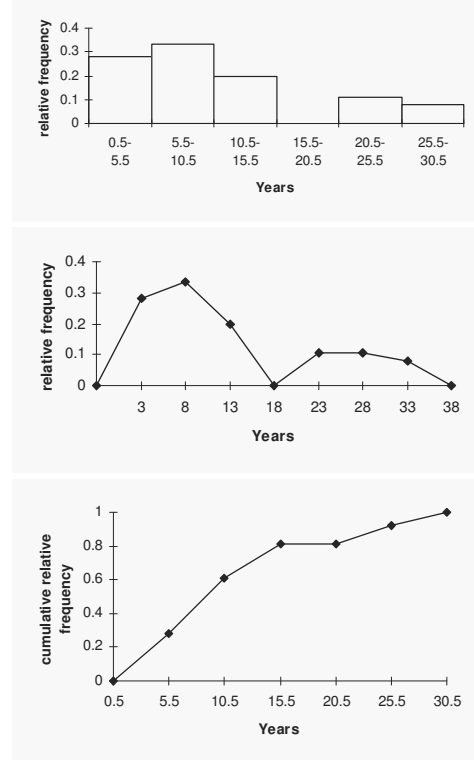


13. continued



The proportion of applicants who need to enroll in a summer program is 0.26 or 26%.

14.



The proportion of employees who have been with the company longer than 20 years is 0.187 or 18.7%.

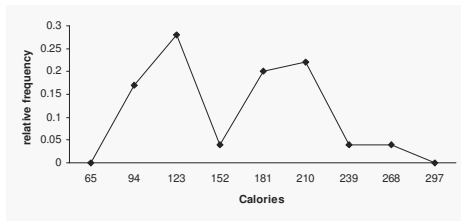
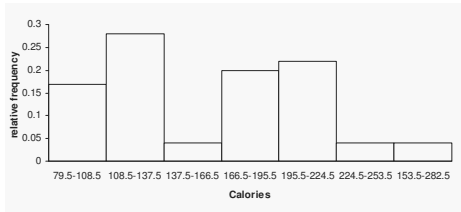
15. $H = 270$ $L = 80$
 $\text{Range} = 270 - 80 = 190$
 $\text{Width} = 190 \div 7 = 27.1$ or 28
 Use width = 29 (rule 2)

Chapter 2 - Frequency Distributions and Graphs

15. continued

Limits	Boundaries	f	rf	crf
80 - 108	79.5 - 108.5	8	0.17	0.17
109 - 137	108.5 - 137.5	13	0.28	0.45
138 - 166	137.5 - 166.5	2	0.04	0.49
167 - 195	166.5 - 195.5	9	0.20	0.69
196 - 224	195.5 - 224.5	10	0.22	0.91
225 - 253	224.5 - 253.5	2	0.04	0.95
254 - 282	253.5 - 282.5	2	<u>0.04</u>	0.99*
		40	0.99	

*due to rounding



The histogram has two peaks.

16.

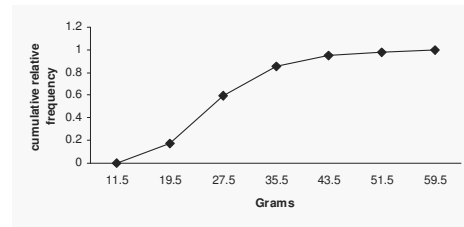
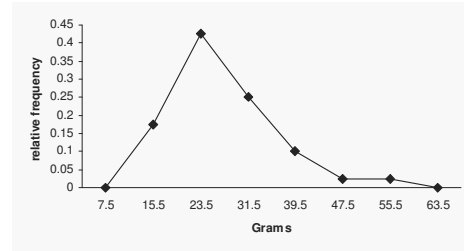
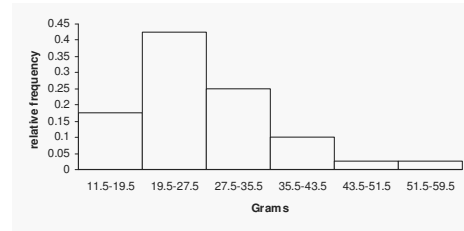
$$H = 57 \quad L = 12$$

$$\text{Range} = 57 - 12 = 45$$

$$\text{Width} = 45 \div 6 = 7.5 \text{ or } 8$$

Limits	Boundaries	f	rf	crf
12 - 19	11.5 - 19.5	7	0.175	0.175
20 - 27	19.5 - 27.5	17	0.425	0.600
28 - 35	27.5 - 35.5	10	0.25	0.850
36 - 43	35.5 - 43.5	4	0.10	0.950
44 - 51	43.5 - 51.5	1	0.025	0.975
52 - 59	51.5 - 59.5	<u>1</u>	<u>0.025</u>	1.000
		40	1.000	

16. continued



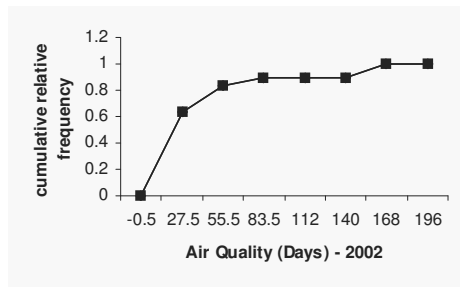
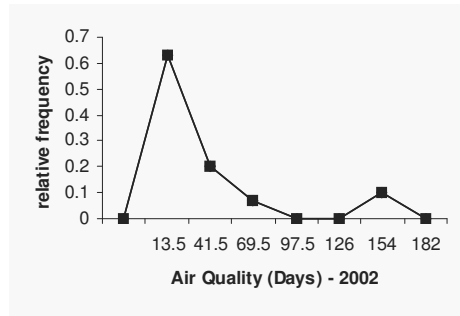
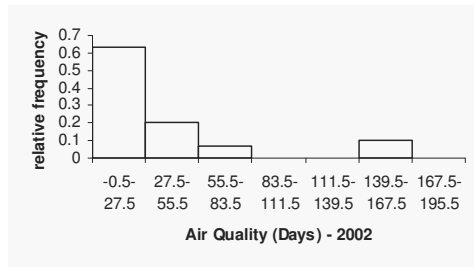
The histogram is positively skewed.

17.

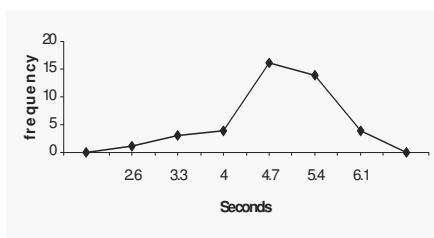
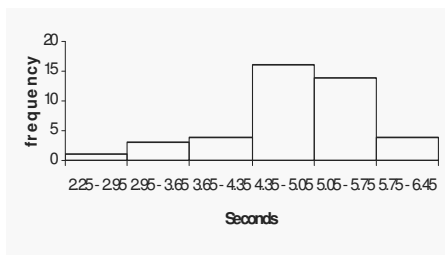
Boundaries	rf	crf
-0.5 - 27.5	0.63	0.63
27.5 - 55.5	0.20	0.83
55.5 - 83.5	0.07	0.90
83.5 - 111.5	0.00	0.90
111.5 - 139.5	0.00	0.90
139.5 - 167.5	0.10	1.00
167.5 - 195.5	<u>0.00</u>	1.00
	100.0	

Chapter 2 - Frequency Distributions and Graphs

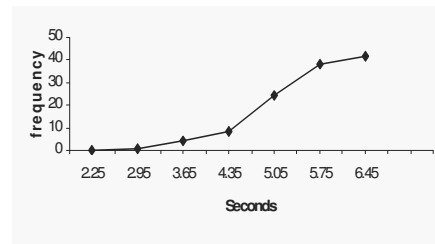
17. continued



18.



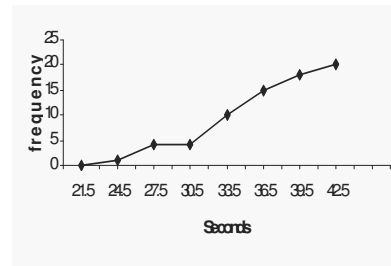
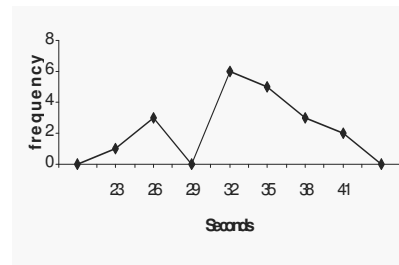
18. continued



Based on the histograms, the older dogs have longer reaction times. Also, the reaction times for older dogs is more variable.

19.

Limits	Boundaries	X_m	f	cf
22 - 24	21.5 - 24.5	23	1	1
25 - 27	24.5 - 27.5	26	3	4
28 - 30	27.5 - 30.5	29	0	4
31 - 33	30.5 - 33.5	32	6	10
34 - 36	33.5 - 36.5	35	5	15
37 - 39	36.5 - 39.5	38	3	18
40 - 42	39.5 - 42.5	41	<u>2</u>	<u>20</u>



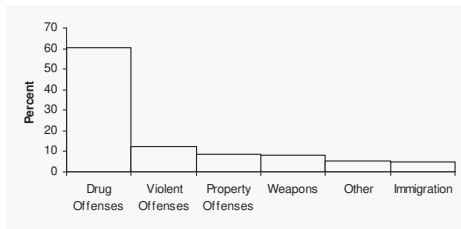
20.

- a. 0
- b. 14
- c. 10
- d. 16

Chapter 2 - Frequency Distributions and Graphs

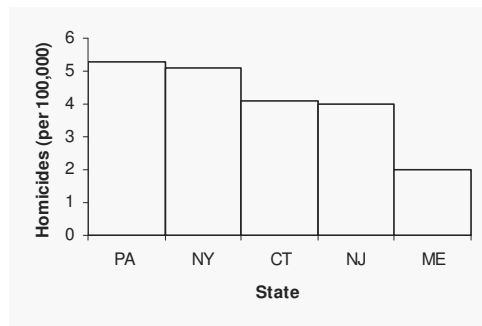
EXERCISE SET 2-4

1.

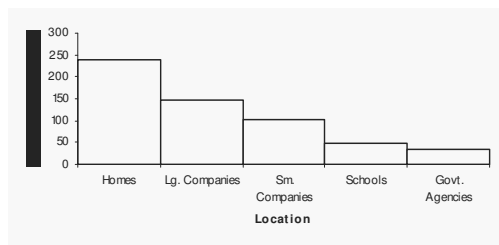


The majority of the money should be spent for drug rehabilitation.

2.

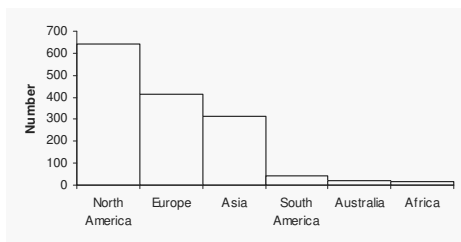


3.

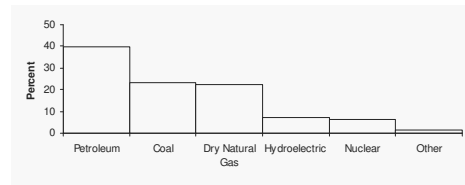


The best place to market products would be to residential users.

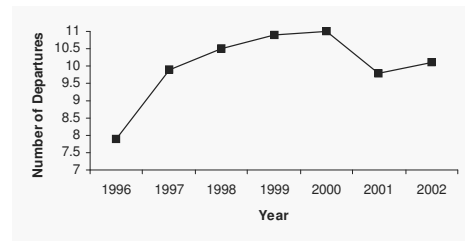
4.



5.

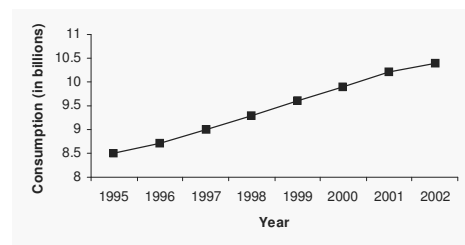


6.



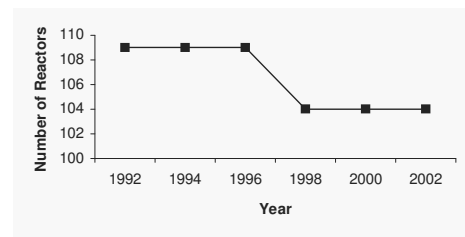
Departures increased until 2000, decreased in 2001, then increased in 2002.

7.

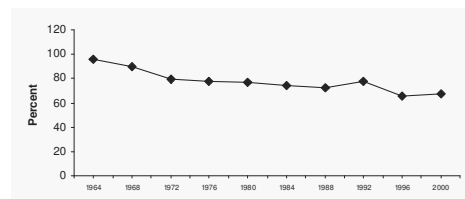


There is a steady increase in consumption of tobacco products.

8.



9.

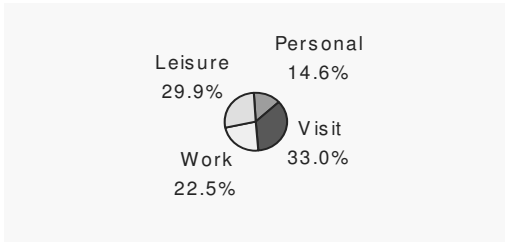


The graph shows a decline in the percentages of registered voters voting in presidential elections.

Chapter 2 - Frequency Distributions and Graphs

10.

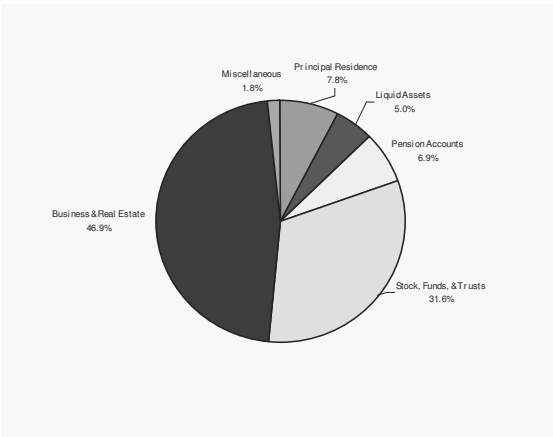
Personal Business	146	14.6%	52.56°
Visit friends or family	330	33.0%	118.8°
Work-related	225	22.5%	81.0°
Leisure	299	29.9%	107.64°
	1000	100%	360°



About $\frac{1}{3}$ of the travelers visit friends or relatives, with the fewest travelling for personal business.

11.

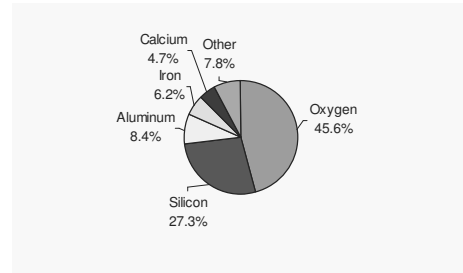
Principal Residence	7.8%	28.08°
Liquid Assets	5.0%	18.0°
Pension Accounts	6.9%	24.84°
Stocks, Funds, and Trusts	31.6%	113.76°
Business & Real Estate	46.9%	168.84°
Miscellaneous	<u>1.8%</u>	<u>6.48°</u>
	100.0%	360.00°



12.

Oxygen	45.6%	164.16°
Silicon	27.3%	98.28°
Aluminum	8.4%	30.24°
Iron	6.2%	22.32°
Calcium	4.7%	16.92°
Other	<u>7.8%</u>	<u>28.08°</u>
	100.0%	360.00°

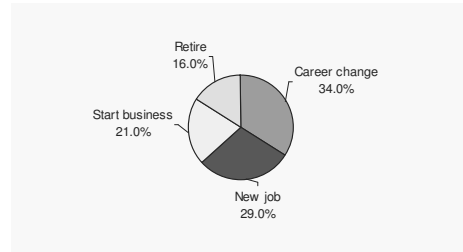
12. continued



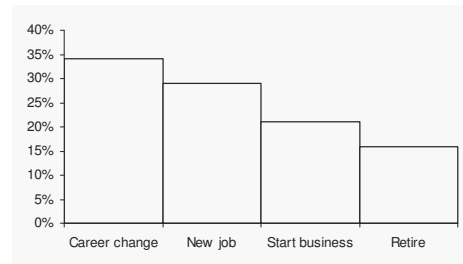
13.

Career change	34%	122.4°
New job	29%	104.4°
Start business	21%	75.6°
Retire	<u>16%</u>	<u>57.6°</u>
	100%	360.0°

Pie chart:



Pareto chart:



The pie graph better represents the data since we are looking at parts of a whole.

14.

- time series graph
- pie graph
- Pareto chart
- pie graph
- time series graph
- Pareto chart

Chapter 2 - Frequency Distributions and Graphs

15.

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

The distribution is somewhat symmetric and unimodal. The majority of the Presidents were in their 50's when inaugurated.

16.

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

The majority of automobile thefts occurred in the 50's and 80's. The data is grouped towards the higher end of the distribution.

17.

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

The distributions are similar but variety 2 seems to be more variable than variety 1.

18.

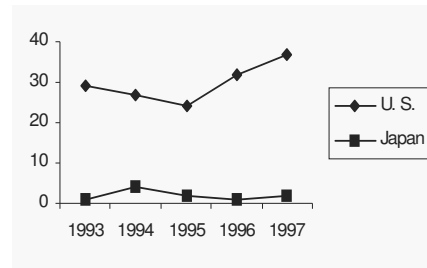
Females										Males									
										5	0	3							
										1	5	9							
										2	2								
										3	1	1							
										6	4	1	4	6	6				
										5	2	6	6	6	9				
										8	5	6	0	0	6	6			
										7	2	0	7	7					
										8	7	8							
8	7	6	6	0	0					9	6	8							

The distribution for unemployed males is more variable than the distribution for unemployed females. There are more unemployed females than males world-wide.

19.

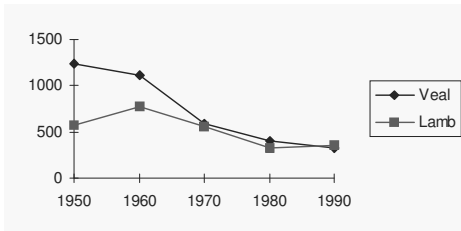
1	3	4	8	9															
2	5	8	9																
3	2	8																	
4	1																		

20.



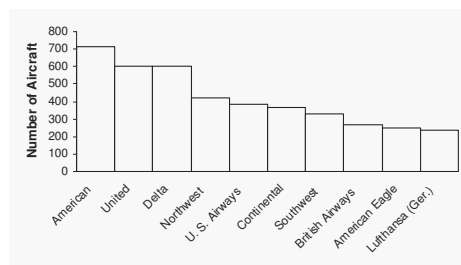
The United States has many more launches than Japan. The number of launches is relatively stable for Japan, while launches varied more for the U. S. The U. S. launches decreased slightly in 1995 and increased after that year.

21.



In 1950, veal production was considerably higher than lamb. By 1970, production was approximately the same for both.

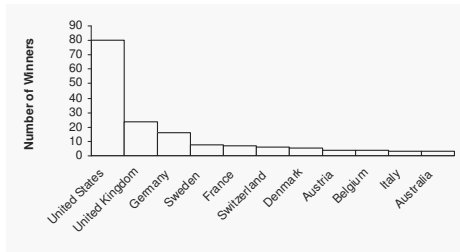
22.



A Pareto chart is most appropriate.

Chapter 2 - Frequency Distributions and Graphs

23.



24. The bottle for 2004 is much wider, giving a distorted view of the difference since only the heights of the bottles should be compared.

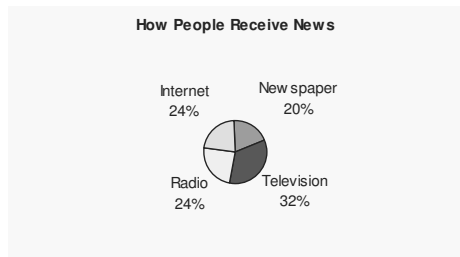
25. The values on the y axis start at 3.5. Also there are no data values shown for the years 2004 through 2011.

REVIEW EXERCISES - CHAPTER 2

1.

Class	f
Newspaper	10
Television	16
Radio	12
Internet	<u>12</u>
	50

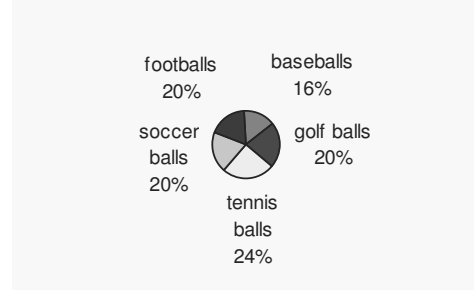
2.



3.

Class	f
baseball	4
golf ball	5
tennis ball	6
soccer ball	5
football	<u>5</u>
	25

4.

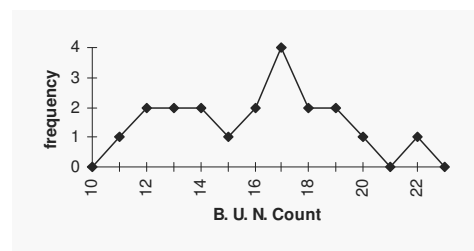
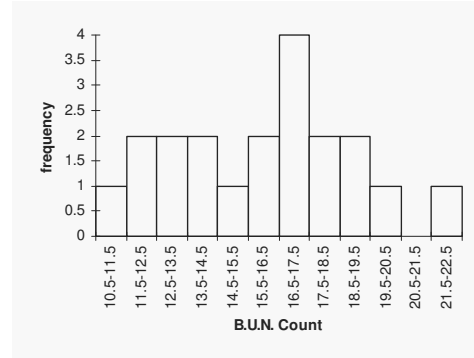


More tennis balls were sold than any other type of ball.

5.

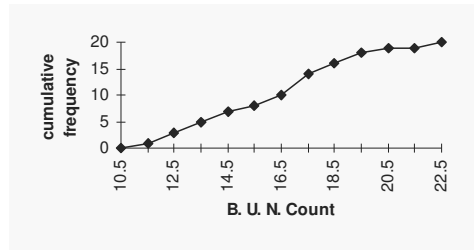
Class	f	cf
11	1	1
12	2	3
13	2	5
14	2	7
15	1	8
16	2	10
17	4	14
18	2	16
19	2	18
20	1	19
21	0	19
22	<u>1</u>	20
	20	

6.



Chapter 2 - Frequency Distributions and Graphs

6. continued

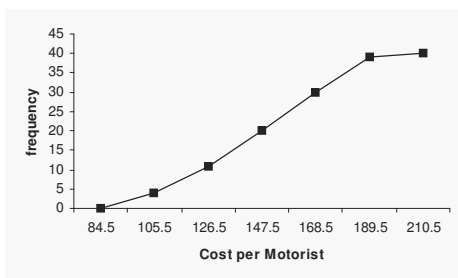
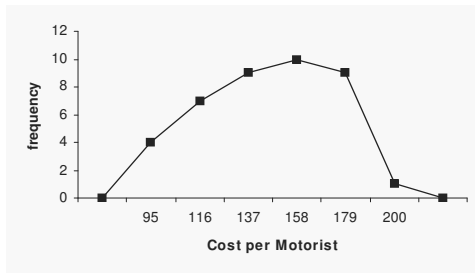
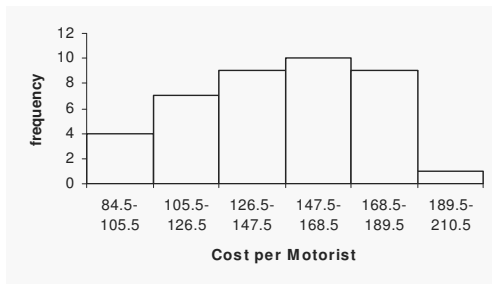


The distribution is somewhat uniform, with a slight peak in the 16.5 - 17.5 class. There is a gap in the 20.5 - 21.5 class.

7.

Limits	Boundaries	f	cf
85 - 105	84.5 - 105.5	4	4
106 - 126	105.5 - 126.5	7	11
127 - 147	126.5 - 147.5	9	20
148 - 168	147.5 - 168.5	10	30
169 - 189	168.5 - 189.5	9	39
190 - 210	189.5 - 210.5	1	40

8.



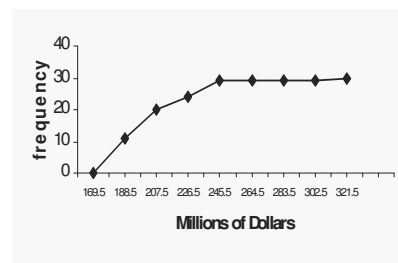
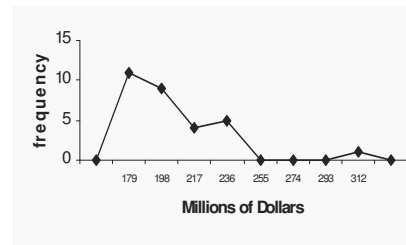
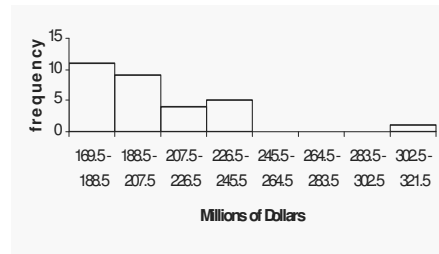
8. continued

The distribution is negatively skewed with most of the data in the two classes 147.5 - 168.5 and 168.5 - 189.5.

9.

Limits	Boundaries	f	cf
170 - 188	169.5 - 188.5	11	11
189 - 207	188.5 - 207.5	9	20
208 - 226	207.5 - 226.5	4	24
227 - 245	226.5 - 245.5	5	29
246 - 264	245.5 - 264.5	0	29
265 - 283	264.5 - 283.5	0	29
284 - 302	283.5 - 302.5	0	29
303 - 321	302.5 - 321.5	1	30
		30	

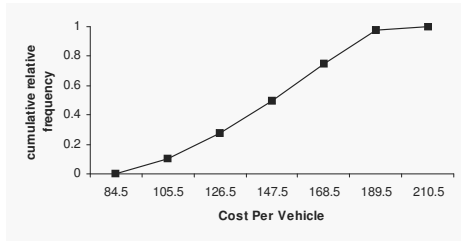
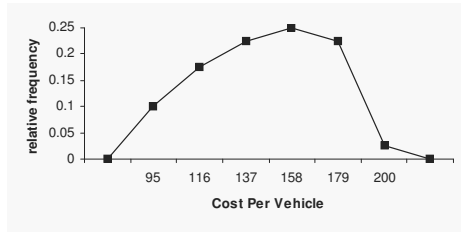
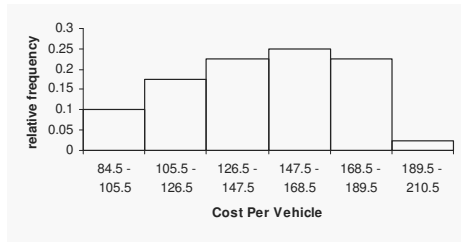
10.



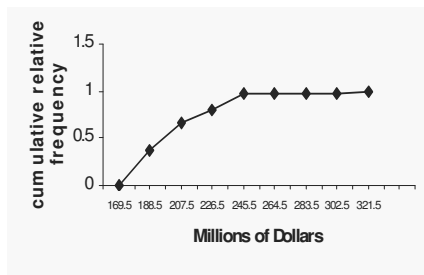
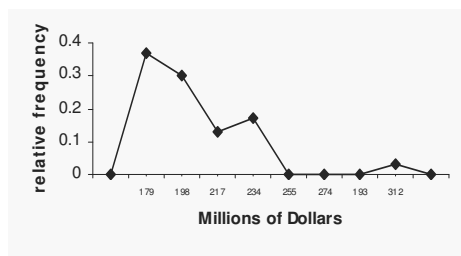
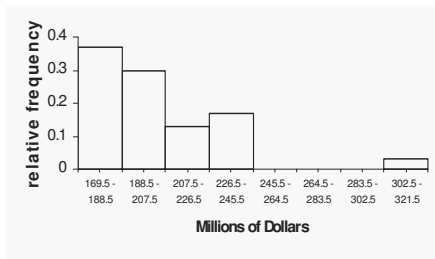
The typical value of the franchises is between \$169.5 - \$188.5 million. All but one of the franchises are valued between \$169.5 and \$245.5 million.

Chapter 2 - Frequency Distributions and Graphs

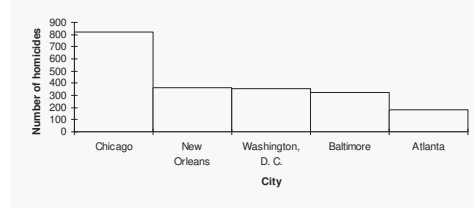
11.



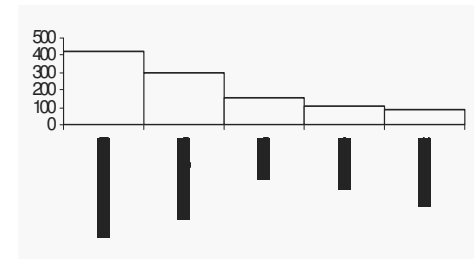
12.



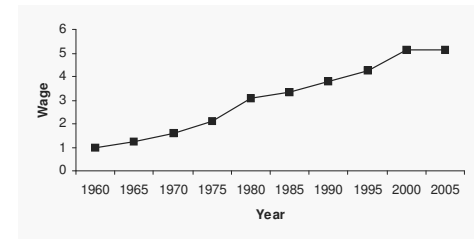
13.



14.

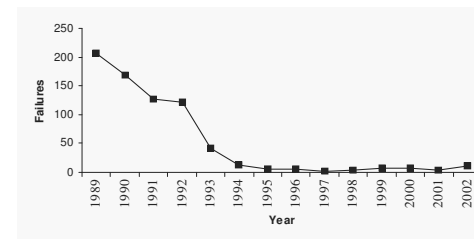


15.



The minimum wage has increased over the years with the largest increase occurring between 1975 and 1980.

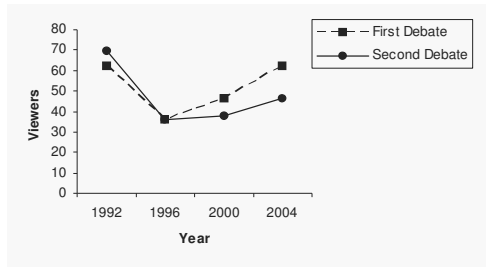
16.



Failures decreased to only one failure in 1997, increased slightly from 1998 to 1999, decreasing through 2001, then increasing in 2002.

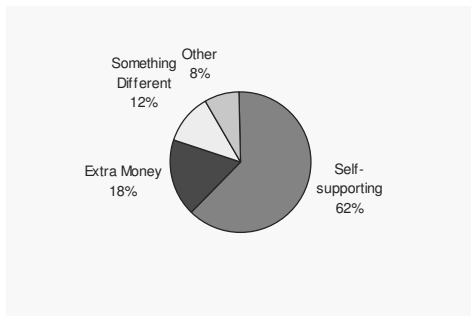
Chapter 2 - Frequency Distributions and Graphs

17.



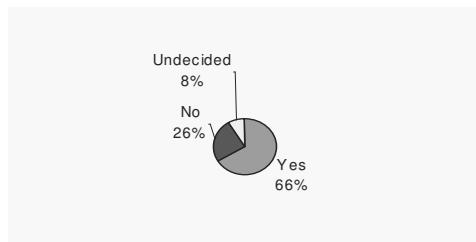
About the same number of people watched the first and second debates in 1992 and 1996. After that more people watched the first debate than watched the second debate.

18.



The majority of women worked to support themselves or their families.

19.



The majority of people surveyed would like to spend the rest of their careers with their present employer.

20.

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

21.

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

The peak of the distribution is in the range of 25 – 29.

22.

20	0	4	9
21	0	1	2 7 8 8
22	2	7	7 7 8
23	0	1	3 7 8
24	1	2	2 3 7
25	1	1	3 4 6
26	0		

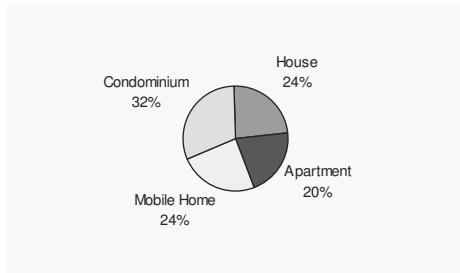
The distribution of aptitude scores is fairly uniform.

CHAPTER 2 QUIZ

- False
- False
- False
- True
- True
- False
- False
- c.
- c.
- b.
- b.
- Categorical, ungrouped, grouped
- 5, 20
- categorical
- time series
- stem and leaf plot
- vertical or y
- | Class | f | cf |
|-------|----|----|
| H | 6 | 6 |
| A | 5 | 11 |
| M | 6 | 17 |
| C | 8 | 25 |
| | 25 | |

Chapter 2 - Frequency Distributions and Graphs

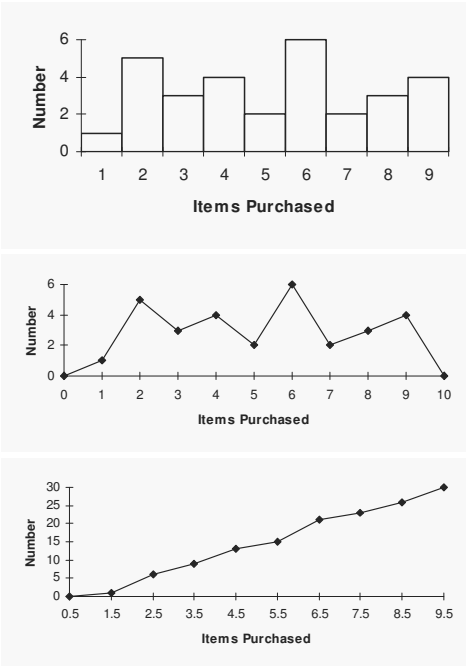
19.



20.

Class	f	cf
0.5 – 1.5	1	1
1.5 – 2.5	5	6
2.5 – 3.5	3	9
3.5 – 4.5	4	13
4.5 – 5.5	2	15
5.5 – 6.5	6	21
6.5 – 7.5	2	23
7.5 – 8.5	3	26
8.5 – 9.5	4	30
	30	

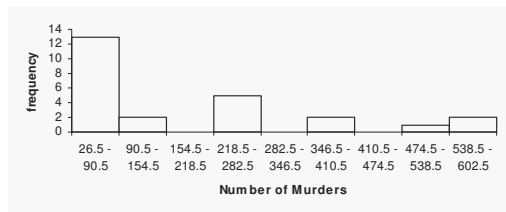
21.



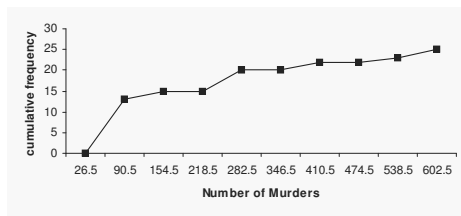
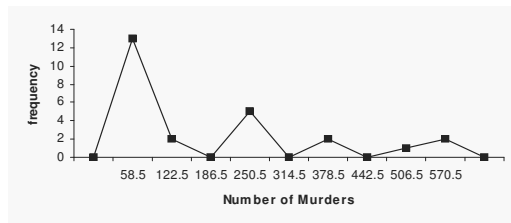
22.

Limits	Boundaries	f	cf
27 - 90	26.5 - 90.5	13	13
91 - 154	90.5 - 154.5	2	15
155 - 218	154.5 - 218.5	0	15
219 - 282	218.5 - 282.5	5	20
283 - 346	282.5 - 346.5	0	20
347 - 410	346.5 - 410.5	2	22
411 - 474	410.5 - 474.5	0	22
475 - 538	474.5 - 538.5	1	23
539 - 602	538.5 - 602.5	2	25
		25	

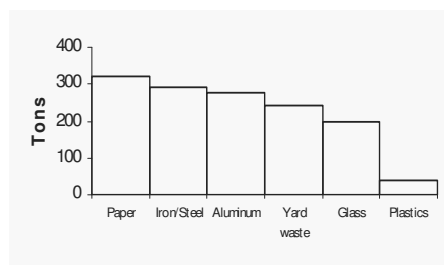
23.



The distribution is positively skewed with one more than half of the data values in the lowest class.

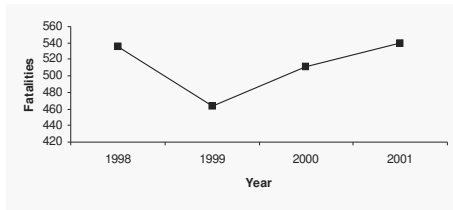


24.



Chapter 2 - Frequency Distributions and Graphs

25.



Fatalities decreased in 1999 and then increased the next two years.

26.

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