CHAPTER 2—Descriptive Statistics: Tabular and Graphical Methods

§2.1 CONCEPTS

2.1 Constructing either a frequency or a relative frequency distribution helps identify and quantify patterns that are not apparent in the raw data.

LO02-01

2.2 Relative frequency of any category is calculated by dividing its frequency by the total number of observations. Percent frequency is calculated by multiplying relative frequency by 100.

LO02-01

2.3 Answers and examples will vary.

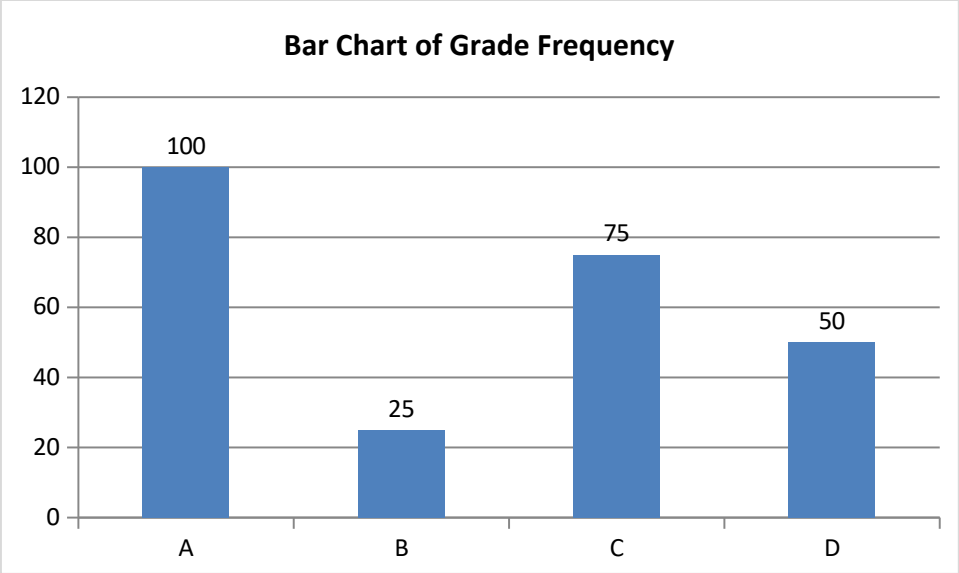
LO02-01

§2.1 METHODS AND APPLICATIONS

2.4 **a.**

Test Response	Frequency	Relative Frequency	Percent Frequency
A	100	0.4	40%
B	25	0.1	10%
C	75	0.3	30%
D	50	0.2	20%

b.

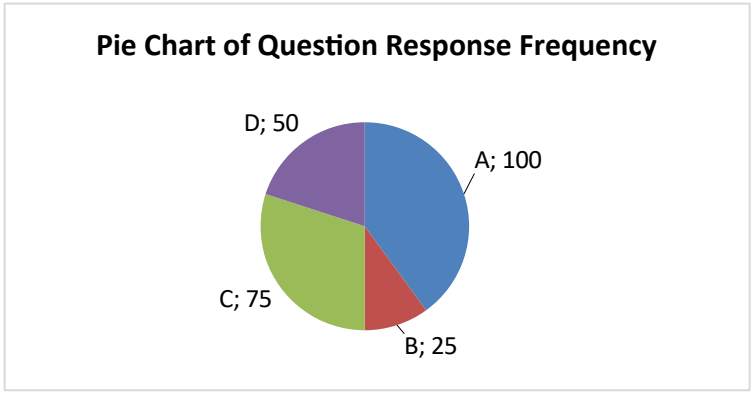


LO02-01

2.5 **a.** $(^{100}/_{250})$ • 360 degrees = 144 degrees for response (a)

b. $(^{25}/_{250})$ • 360 degrees = 36 degrees for response (b)

c.

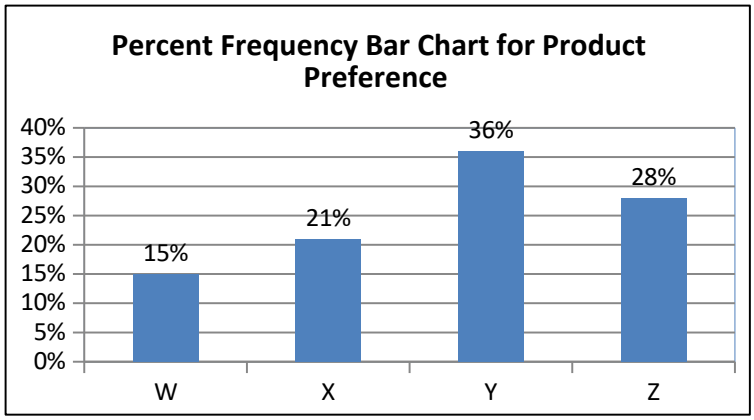


LO02-01

2.6 **a.** Relative frequency for product *x* is 1 – (0.15 + 0.36 + 0.28) = 0.21

b. Product: W X Y Z
frequency = relative frequency • N = 0.15 • 500 = 75 105 180 140

c.



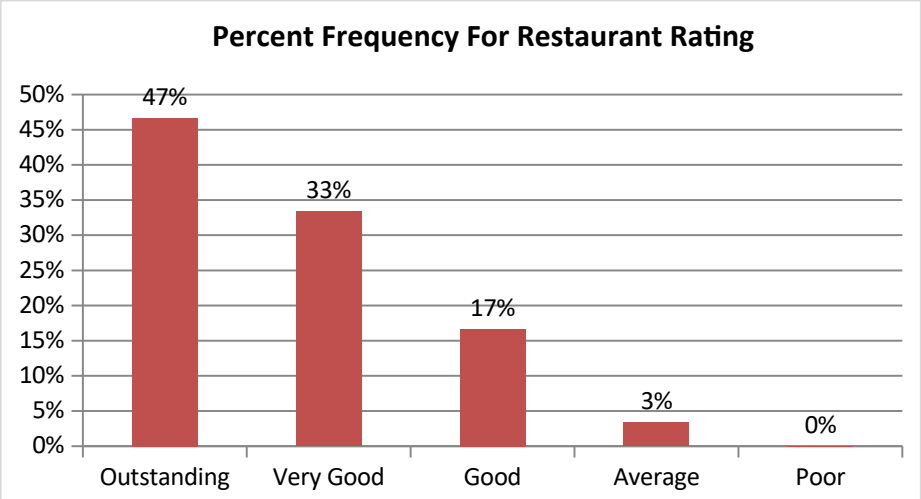
d. Degrees for W would be 0.15 • 360 = 54
for X 75.6
for Y 129.6
for Z 100.8.

LO02-01

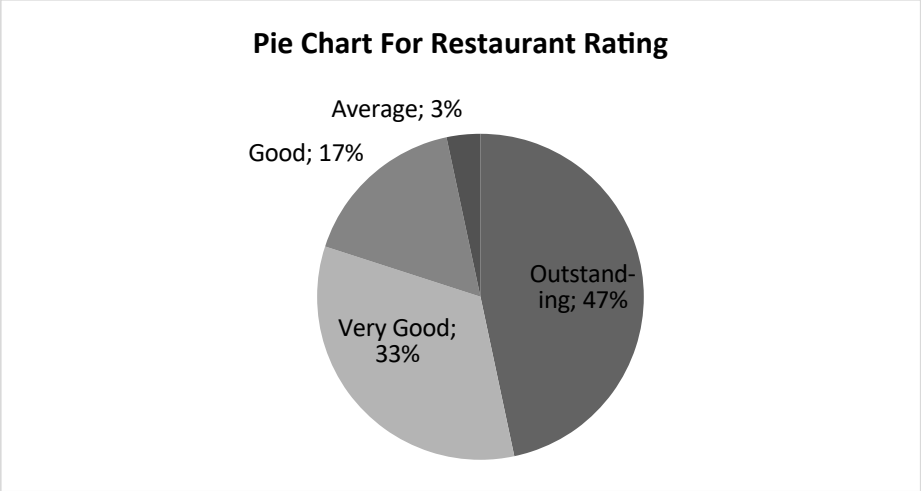
2.7 **a.**

Rating	Frequency	Relative Frequency
Outstanding	14	$^{14}/_{30} = 0.467$
Very Good	10	$^{10}/_{30} = 0.333$
Good	5	$^{5}/_{30} = 0.167$
Average	1	$^{1}/_{30} = 0.033$
Poor	0	$^{0}/_{30} = 0.000$
	$\sum = 30$	

b.



c.

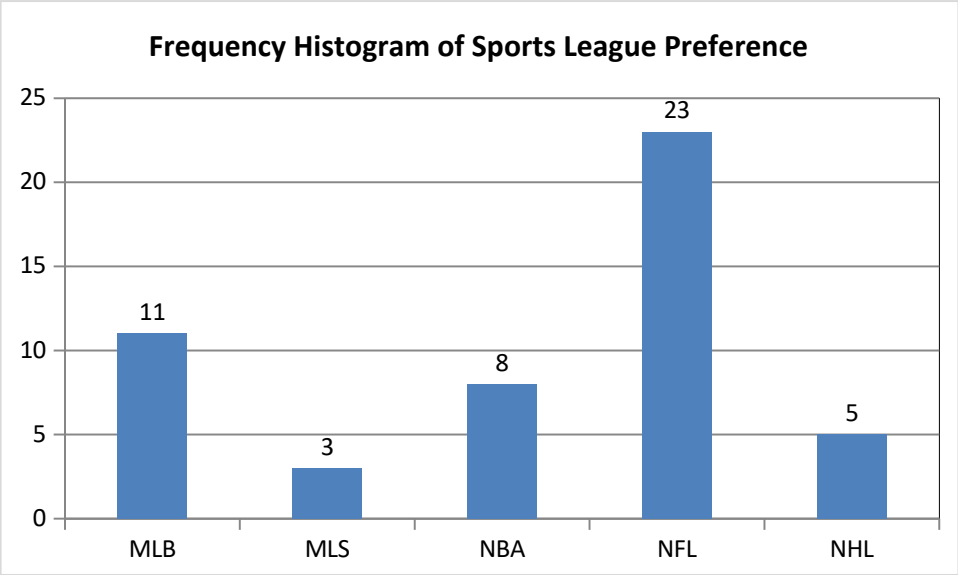


LO02-01

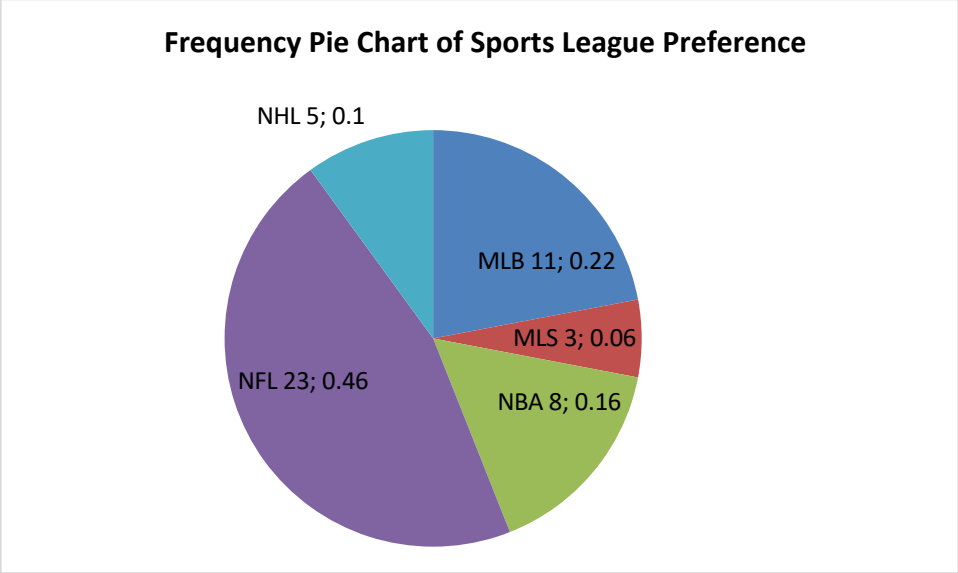
2.8 **a.** Frequency Distribution for ***Sports League Preference***

Sports League	*Frequency*	*Percent Frequency*	*Percent*
MLB	11	0.22	22%
MLS	3	0.06	6%
NBA	8	0.16	16%
NFL	23	0.46	46%
NHL	5	0.10	10%
	50		

b.



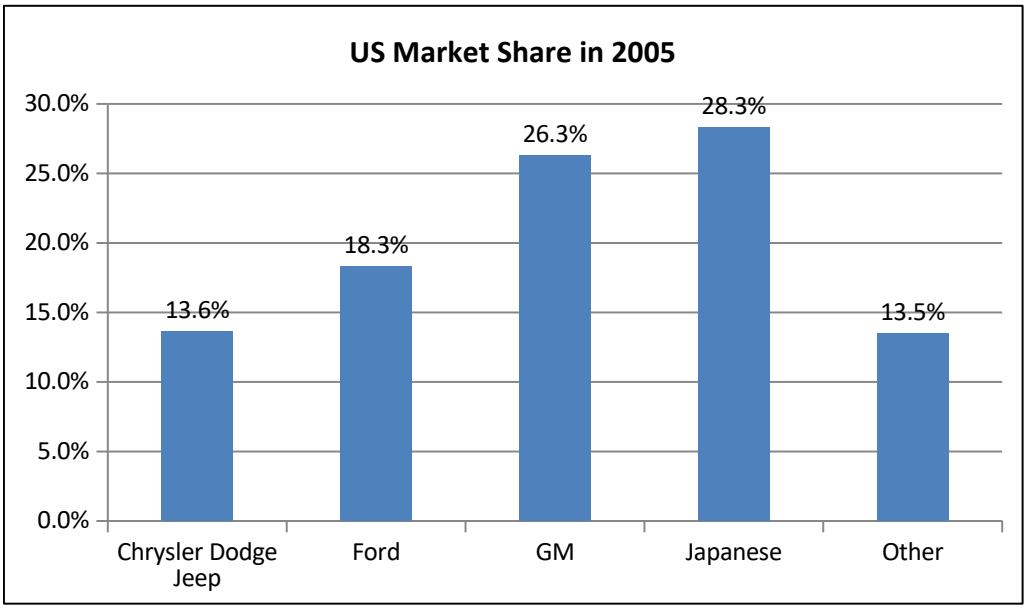
c.

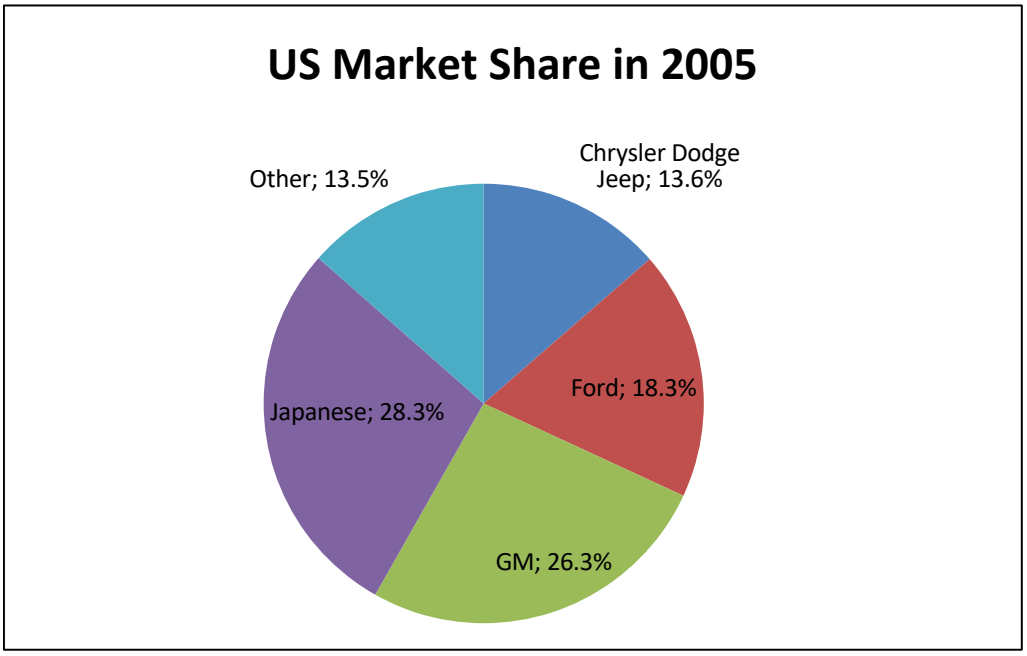


d. The most popular league is NFL and the least popular is MLS.

LO02-011

2.9



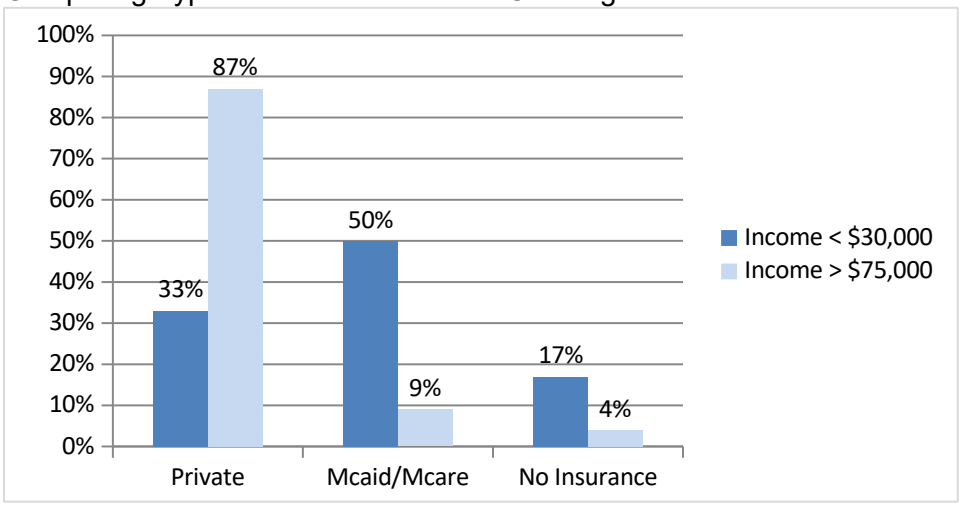


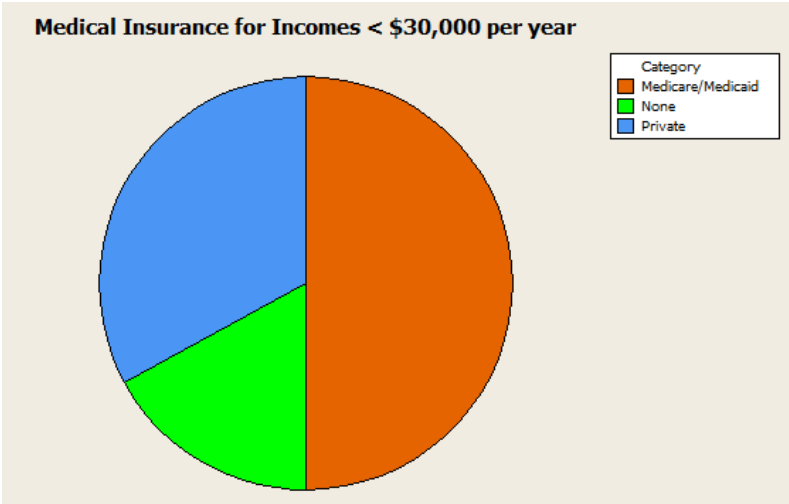
LO02-01

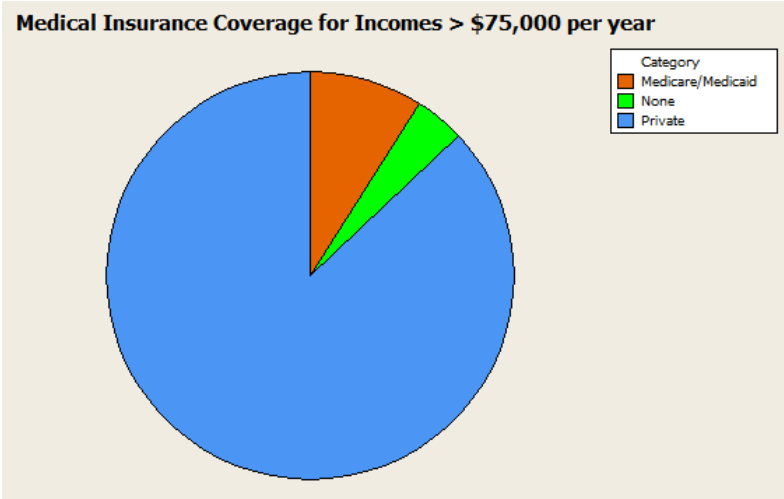
2.10 Comparing the pie chart above and the chart for 2010 in the text book shows that between 2005 and 2010, the three U.S. manufacturers, Chrysler, Ford and GM have all lost market share, while Japanese and other imported models have increased market share.

LO02-01

2.11 Comparing Types of Health Insurance Coverage Based on Income Level







LO02-01

2.12 **a.** Percent of calls that are require investigation or help = 28.12% + 4.17% = 32.29%

b. Percent of calls that represent a new problem = 4.17%

c. Only 4% of the calls represent a new problem to all of technical support, but one-third of the problems require the technician to determine which of several previously known problems this is and which solutions to apply. It appears that increasing training or improving the documentation of known problems and solutions will help.

LO02-02

§2.2 CONCEPTS

2.13 **a.** We construct a frequency distribution and a histogram for a data set so we can gain some insight into the shape, center, and spread of the data along with whether or not outliers exist.

b. A frequency histogram represents the frequencies for the classes using bars while in a frequency polygon the frequencies are represented by plotted points connected by line segments.

c. A frequency ogive represents a cumulative distribution while the frequency polygon does not represent a cumulative distribution. Also, in a frequency ogive, the points are plotted at the upper class boundaries; in a frequency polygon, the points are plotted at the class midpoints.

LO02-03

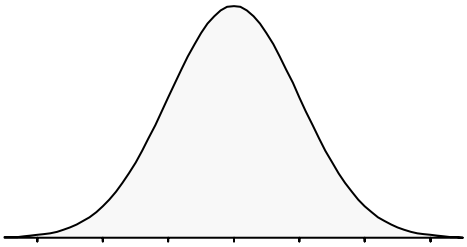
2.14 **a.** To find the frequency for a class, you simply count how many of the observations have values that are greater than or equal to the lower boundary and less than the upper boundary.

b. Once you determine the frequency for a class, the relative frequency is obtained by dividing the class frequency by the total number of observations (data points).

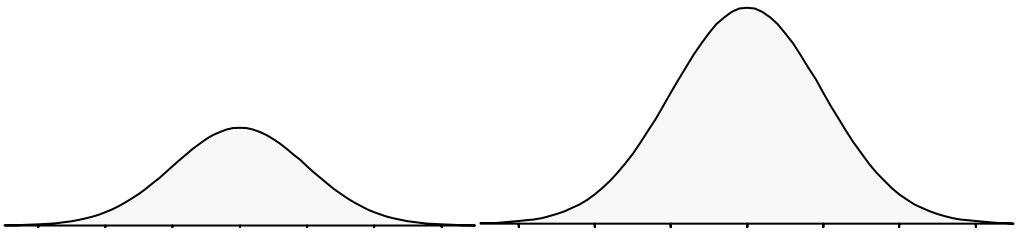
c. The percent frequency for a class is calculated by multiplying the relative frequency by 100.

LO02-03

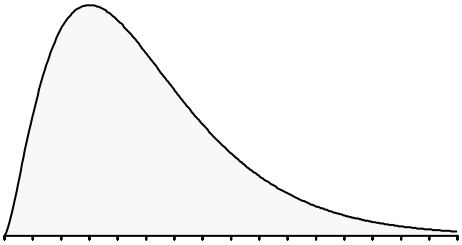
2.15 **a.** Symmetrical and mound shaped:
One hump in the middle; left side is a mirror image of the right side.



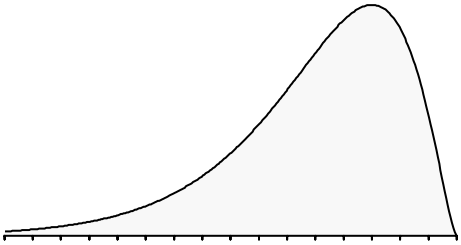
b. Double peaked:
Two humps, the left of which may or may not look like the right one, nor is each hump required to be symmetrical



c. Skewed to the Right:
Long tail to the right



d. Skewed to the left:
Long tail to the left



LO02-03

§2.2 METHODS AND APPLICATIONS

2.16 **a.** Since there are 28 points we use 5 classes (from Table 2.5).

b. Class Length (CL) = (largest measurement – smallest measurement) / #classes
= (46 – 17) / 5 = 6
(If necessary, round up to the same level of precision as the data itself.)

c. The first class's lower boundary is the smallest measurement, 17.

The first class's upper boundary is the lower boundary plus the Class Length, 17 + 3 = 23

The second class's lower boundary is the first class's upper boundary, 23

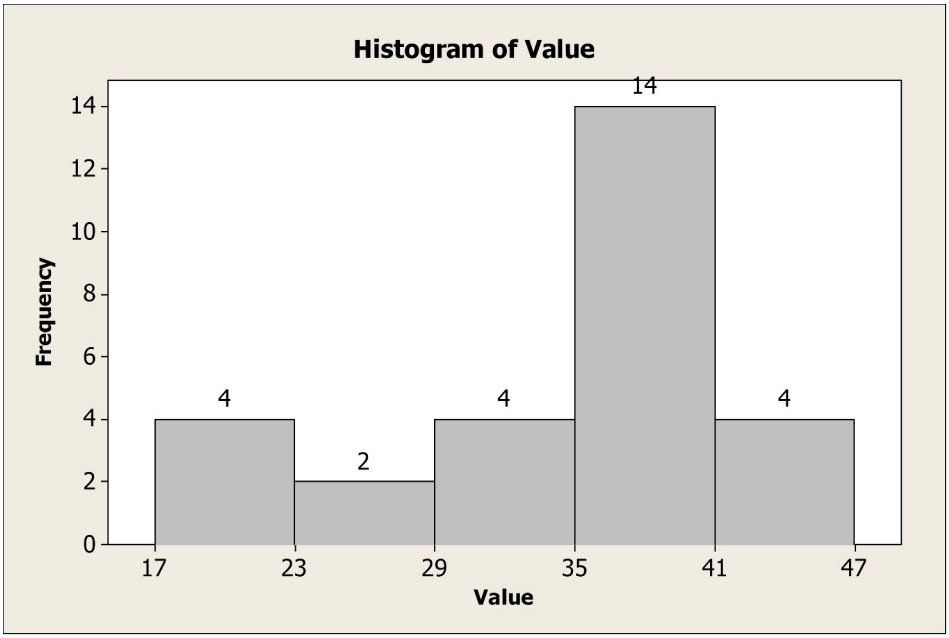
Continue adding the Class Length (width) to lower boundaries to obtain the 5 classes:

$17 \le x < 23$ | $23 \le x < 29$ | $29 \le x < 35$ | $35 \le x < 41$ | $41 \le x \le 47$

d. Frequency Distribution for ***Values***

lower		upper	midpoint	width	frequency	percent	cumulative frequency	cumulative percent
17	<	23	20	6	4	14.3	4	14.3
23	<	29	26	6	2	7.1	6	21.4
29	<	35	32	6	4	14.3	10	35.7
35	<	41	38	6	14	50.0	24	85.7
41	<	47	44	6	4	14.3	28	100.0
					28	100.0		

e.



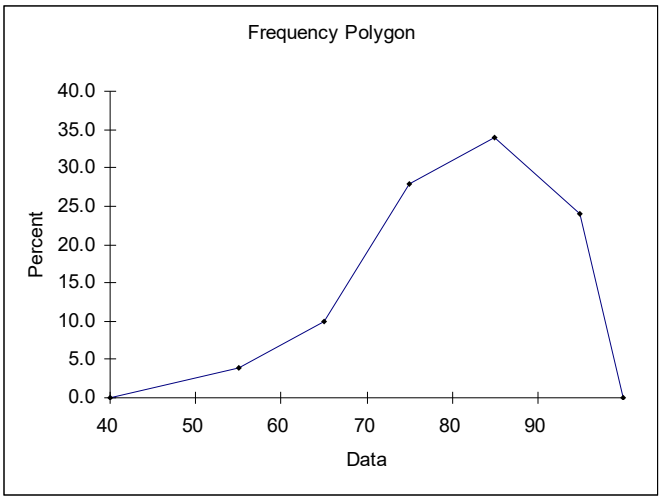
f. See output in answer to **d**.

LO02-03

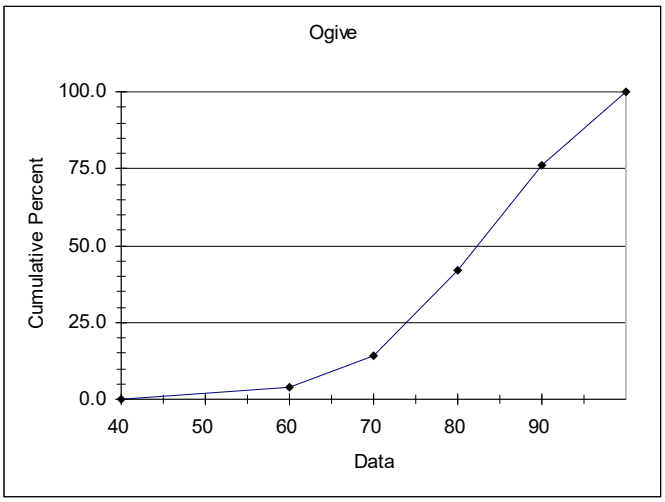
2.17 **a.** and **b.** *Frequency Distribution for* ***Exam Scores***

lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*relative frequency*	*cumulative frequency*	*cumulative percent*
50	<	60	55	10	2	4.0	0.04	2	4.0
60	<	70	65	10	5	10.0	0.10	7	14.0
70	<	80	75	10	14	28.0	0.28	21	42.0
80	<	90	85	10	17	34.0	0.34	38	76.0
90	<	100	95	10	12	24.0	0.24	50	100.0
					50	100.0			

c.



d.



LO02-03

2.18 **a.** Because there are 60 data points of design ratings, we use six classes (from Table 2.5).

b. Class Length (CL) = (Max – Min)/#Classes = (35 – 20) / 6 = 2.5 and we round up to 3, the level of precision of the data.

c. The first class's lower boundary is the smallest measurement, 20.

The first class's upper boundary is the lower boundary plus the Class Length, 20 + 3 = 23

The second class's lower boundary is the first class's upper boundary, 23

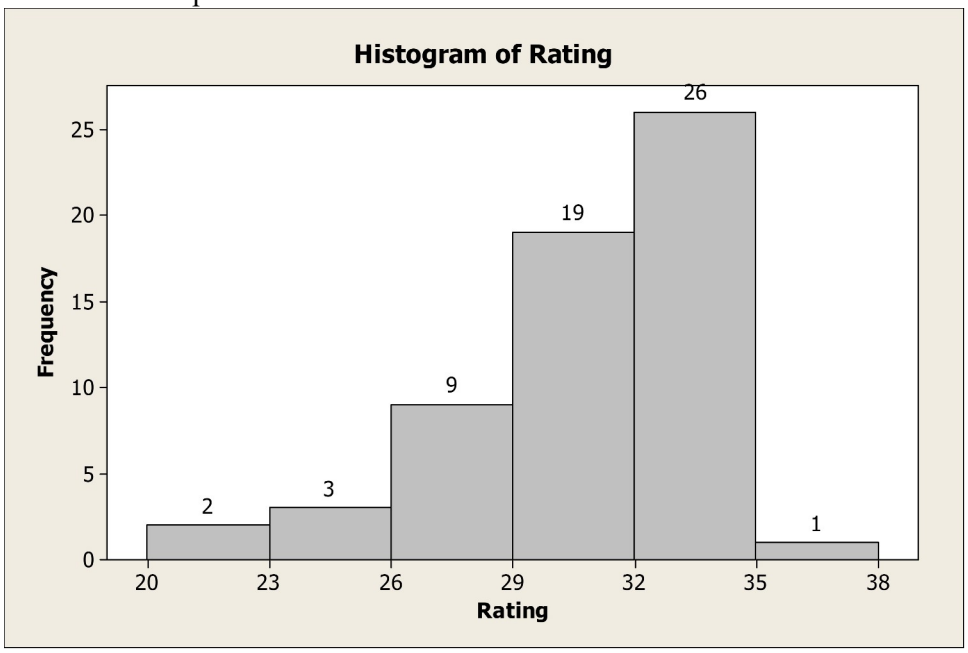
Continue adding the Class Length (width) to lower boundaries to obtain the 6 classes:

| 20 < 23 | 23 < 26 | 26 < 29 | 29 < 32 | 32 < 35 | 35 < 38 |

d. Frequency Distribution for ***Bottle Design Ratings***

lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
20	<	23	21.5	3	2	3.3	2	3.3
23	<	26	24.5	3	3	5	5	8.3
26	<	29	27.5	3	9	15	14	23.3
29	<	32	30.5	3	19	31.7	33	55
32	<	35	33.5	3	26	43.3	59	98.3
35	<	38	36.5	3	1	1.7	60	100
					60	100		

e. Distribution shape is skewed left.

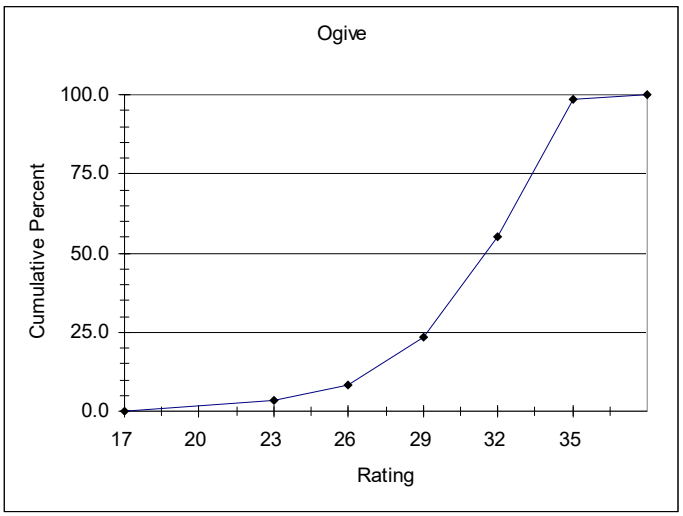


LO02-03

2.19 a & b. Frequency Distribution for ***Ratings***

lower		*upper*	*midpoint*	*width*	*relative frequency*	*percent*	*cumulative relative frequency*	*cumulative percent*
20	<	23	21.5	3	0.033	3.3	0.033	3.3
23	<	26	24.5	3	0.050	5.0	0.083	8.3
26	<	29	27.5	3	0.150	15.0	0.233	23.3
29	<	32	30.5	3	0.317	31.7	0.550	55.0
32	<	35	33.5	3	0.433	43.3	0.983	98.3
35	<	38	36.5	3	0.017	1.7	1.000	100.0
					1.000	100		

c.



LO02-03

2.20 a. Because we have the annual pay of 25 celebrities, we use five classes (from Table 2.5).

Class Length (CL) = (290 – 28) / 5 = 52.4 and we round up to 53 since the data are in whole numbers.

The first class's lower boundary is the smallest measurement, 28.

The first class's upper boundary is the lower boundary plus the Class Length, 28 + 53 = 81

The second class's lower boundary is the first class's upper boundary, 81

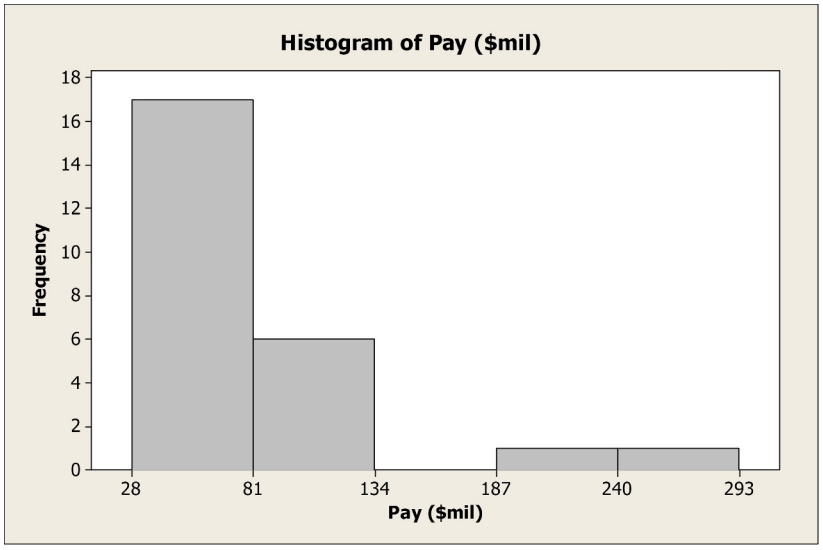
Continue adding the Class Length (width) to lower boundaries to obtain the 5 classes:

| 28 < 81 | 81 < 134 | 134 < 187 | 187 < 240 | 240 < 293 |

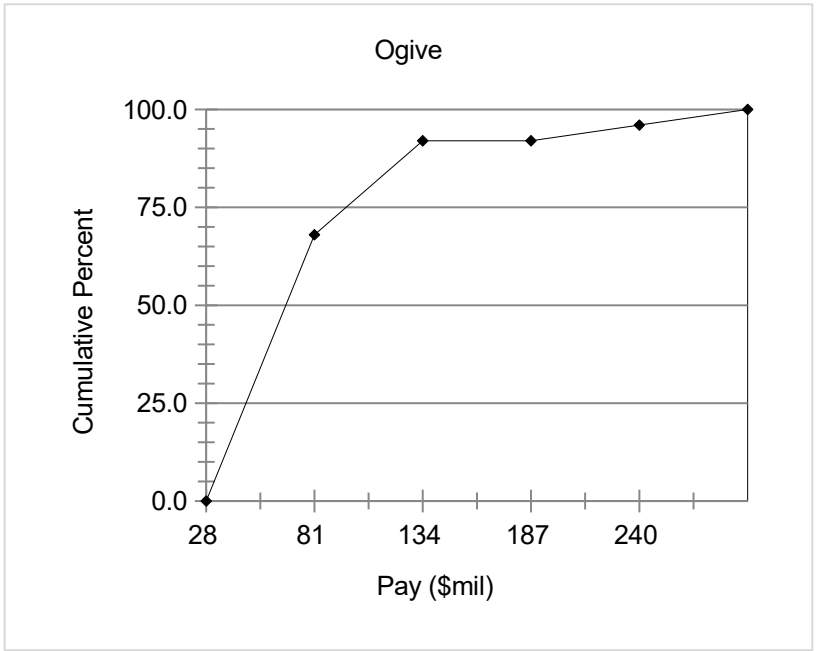
2.20 a. (*cont.*)

Frequency Distribution for ***Celebrity Annual Pay($mil)***

lower	<	*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
28	<	81	54.5	53	17	34.0	17	34.0
81	<	134	107.5	53	6	12.0	23	46.0
134	<	187	160.5	53	0	0.0	23	46.0
187	<	240	213.5	53	1	2.0	24	48.0
240	<	293	266.5	53	1	2.0	25	50.0
					25	50.0		



c.



LO02-03

2.21 **a.** The video game satisfaction ratings are concentrated between 40 and 46.

b. Shape of distribution is slightly skewed left. Recall that these ratings have a minimum value of 7 and a maximum value of 49. This shows that the responses from this survey are reaching near to the upper limit but significantly diminishing on the low side.

c.

Class:	1	2	3	4	5	6	7
Ratings:	34<x≤36	36<x≤38	38<x≤40	40<x≤42	42<x≤44	44<x≤46	46<x≤48

d.

Cum Freq:	1	4	13	25	45	61	65

LO02-03

2.22 **a.** The bank wait times are concentrated between 4 and 7 minutes.

b. The shape of distribution is slightly skewed right. Waiting time has a lower limit of 0 and stretches out to the high side where there are a few people who have to wait longer.

c. The class length is 1 minute.

d. Frequency Distribution for ***Bank Wait Times***

lower	<	*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
-0.5	<	0.5	0	1	1	1%	1	1%
0.5	<	1.5	1	1	4	4%	5	5%
1.5	<	2.5	2	1	7	7%	12	12%
2.5	<	3.5	3	1	8	8%	20	20%
3.5	<	4.5	4	1	17	17%	37	37%
4.5	<	5.5	5	1	16	16%	53	53%
5.5	<	6.5	6	1	14	14%	67	67%
6.5	<	7.5	7	1	12	12%	79	79%
7.5	<	8.5	8	1	8	8%	87	87%
8.5	<	9.5	9	1	6	6%	93	93%
9.5	<	10.5	10	1	4	4%	97	97%
10.5	<	11.5	11	1	2	2%	99	99%
11.5	<	12.5	12	1	1	1%	100	100%
					100			

LO02-03

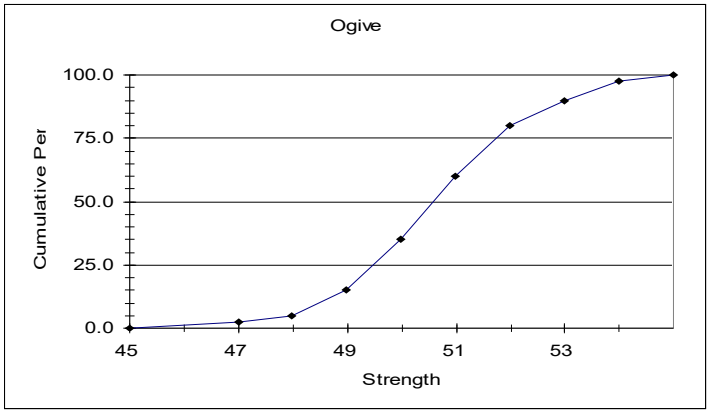
2.23 **a.** The trash bag breaking strengths are concentrated between 48 and 53 pounds.

b. The shape of distribution is symmetric and bell shaped.

c. The class length is 1 pound.

d.

Class:	46<47	47<48	48<49	49<50	50<51	51<52	52<53	53<54	54<55
Cum Freq.	2.5%	5.0%	15.0%	35.0%	60.0%	80.0%	90.0%	97.5%	100.0%

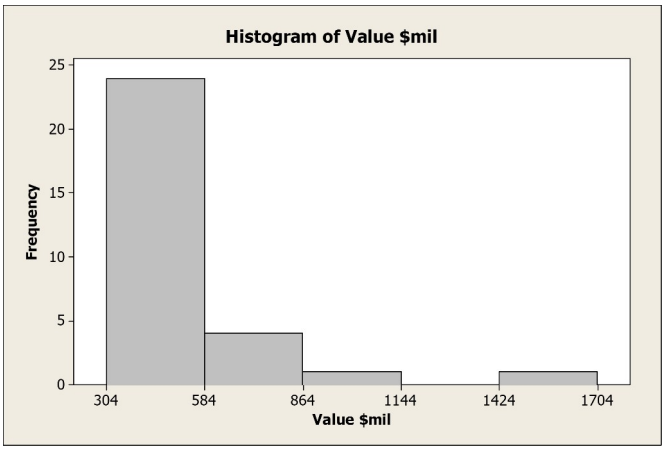


LO02-03

2.24 **a.** Because there are 30 data points, we will use 5 classes (Table 2.5). The class length will be (1700-304)/5= 279.2, rounded to the same level of precision as the data, 280.

Frequency Distribution for ***MLB Team Value ($mil)***

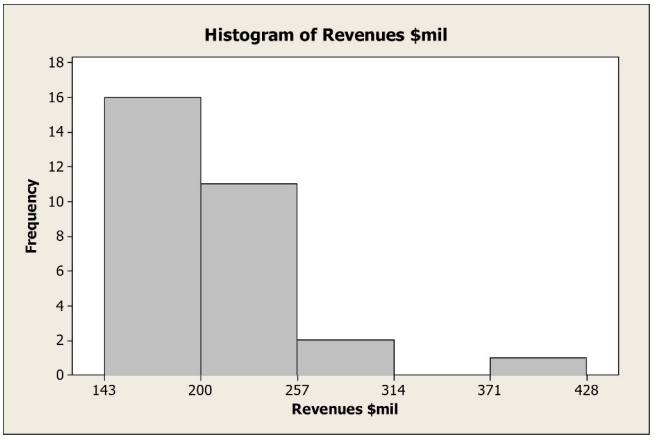
lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
304	<	584	444	280	24	80.0	24	80.0
584	<	864	724	280	4	13.3	28	93.3
864	<	1144	1004	280	1	3.3	29	96.7
1144	<	1424	1284	280	0	0.0	29	96.7
1424	<	1704	1564	280	1	3.3	30	100.0
					30	100.0		



Distribution is skewed right and has a distinct outlier, the NY Yankees.

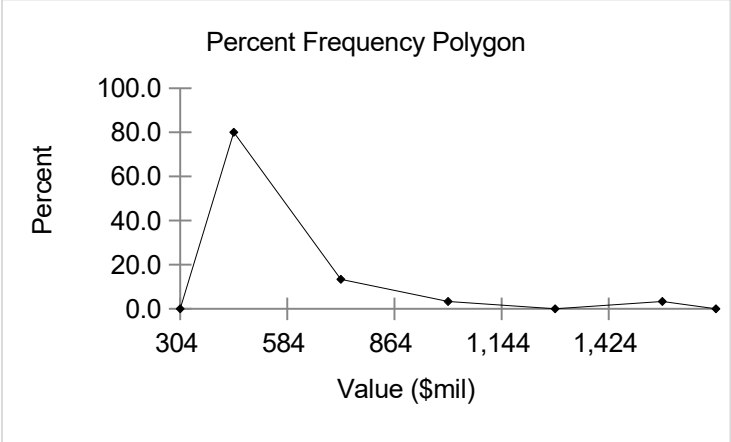
2.24 b. Frequency Distribution for ***MLB Team Revenue***

lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
143	<	200	171.5	57	16	53.3	16	53.3
200	<	257	228.5	57	11	36.7	27	90.0
257	<	314	285.5	57	2	6.7	29	96.7
314	<	371	342.5	57	0	0.0	29	96.7
371	<	428	399.5	57	1	3.3	30	100.0
					30	100.0		



The distribution is skewed right.

c.

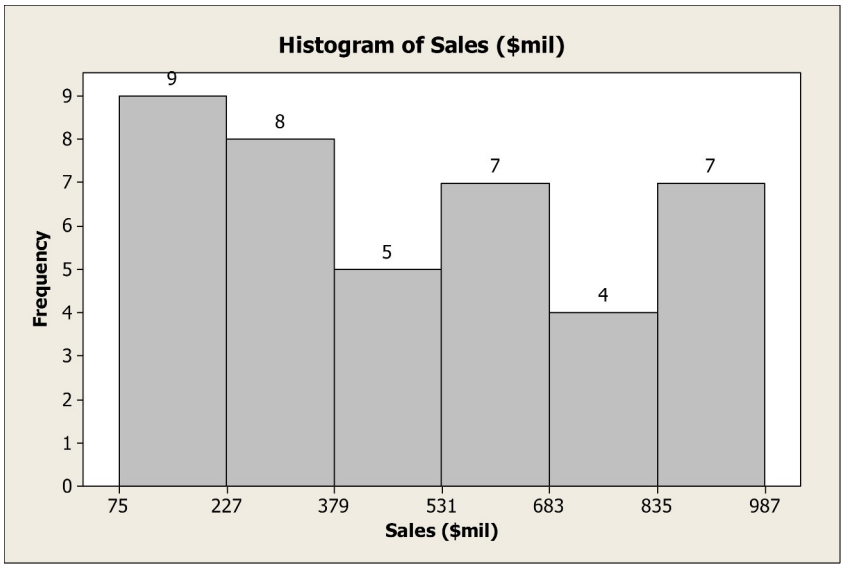


LO02-03

2.25 a. Because there are 40 data points, we will use 6 classes (Table 2.5). The class length will be (986-75)/6= 151.83. Rounding up to the same level of precision as the data gives a width of 152. Beginning with the minimum value for the first lower boundary, 75, add the width, 152, to obtain successive boundaries.

Frequency Distribution for ***Sales ($mil)***

lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
75	<	227	151	152	9	22.5	9	22.5
227	<	379	303	152	8	20.0	17	42.5
379	<	531	455	152	5	12.5	22	35.0
531	<	683	607	152	7	17.5	29	60.0
683	<	835	759	152	4	10.0	33	70.0
835	<	987	911	152	7	17.5	40	87.5
					40	100.0		

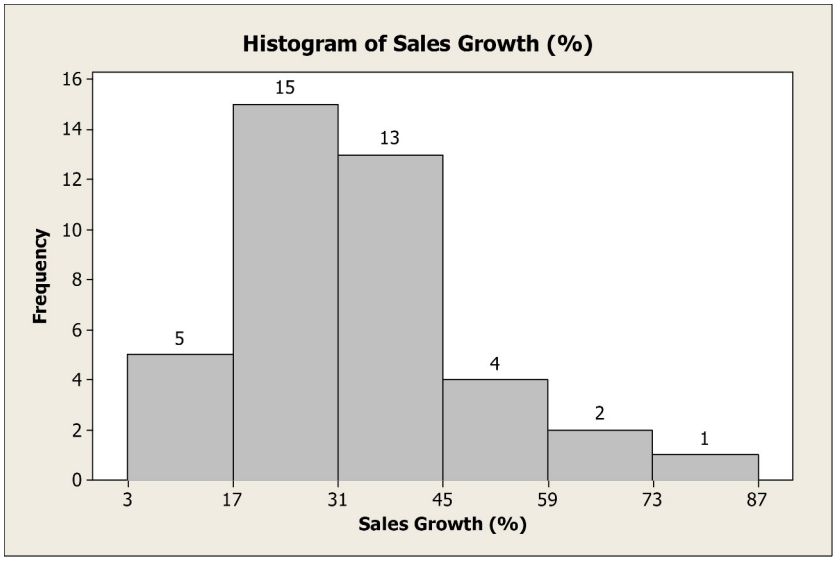


The distribution is relatively flat, perhaps mounded.

2.25 b. Again, we will use 6 classes for 40 data points. The class length will be (86-3)/6= 13.83. Rounding up to the same level of precision gives a width of 14. Beginning with the minimum value for the first lower boundary, 3, add the width, 14, to obtain successive boundaries.

Frequency Distribution for ***Sales Growth (%)***

lower		*upper*	*midpoint*	*width*	*frequency*	*percent*	*cumulative frequency*	*cumulative percent*
3	<	17	10	14	5	12.5	5	12.5
17	<	31	24	14	15	37.5	20	50.0
31	<	45	38	14	13	32.5	33	82.5
45	<	59	52	14	4	10.0	37	92.5
59	<	73	66	14	2	5.0	39	97.5
73	<	87	80	14	1	2.5	40	100.0
					40	100.0		



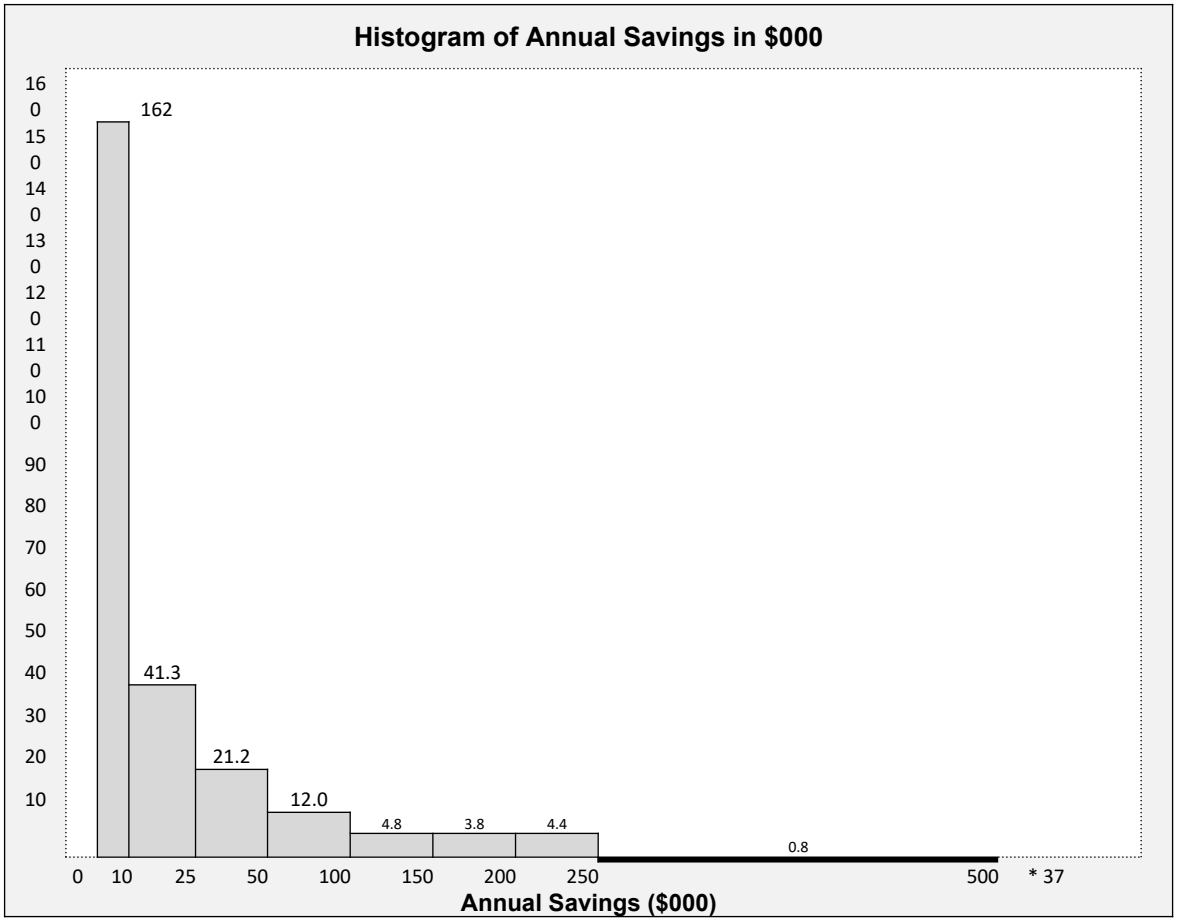
The distribution is skewed right.

LO02-03

2.26 a. Frequency Distribution for ***Annual Savings in $000***

					width =*factor*	*frequency* =*height*
lower	*upper*	*midpoint*	*width*	*frequency*	*base*	*factor*
0	< 10	5.0	10	162	10 / 10 =1.0	162 / 1.0 =162.0
10	< 25	17.5	15	62	15 / 10 =1.5	62 / 1.5 =41.3
25	< 50	37.5	25	53	25 / 10 =2.5	53 / 2.5 =21.2
50	< 100	75.0	50	60	50 / 10 =5.0	60 / 5.0 =12
100	< 150	125.0	50	24	50 / 10 =5.0	24 / 5.0 =4.8
150	< 200	175.0	50	19	50 / 10 =5.0	19 / 5.0 =3.8
200	< 250	225.0	50	22	50 / 10 =5.0	22 / 5.0 =4.4
250	< 500	375.0	250	21	250 / 10 =25.0	21 / 25.0 =0.8
500				37		
				460		

2.26 b. and **2.27**



LO02-03

§2.3 CONCEPTS

2.28 The horizontal axis spans the range of measurements, and the dots represent the measurements.

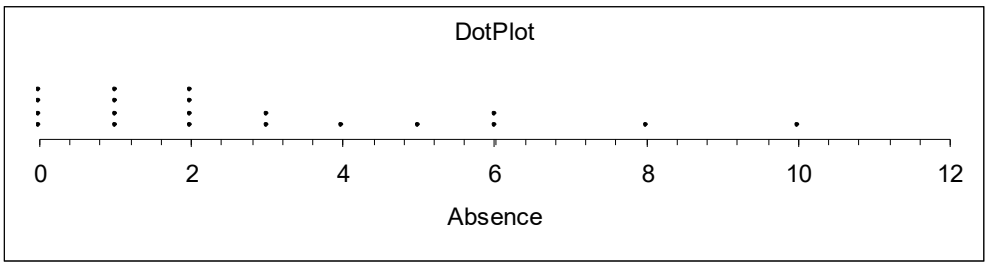
LO02-04

2.29 A dot plot with 1,000 points is not practical. Group the data and use a histogram.

LO02-03, LO02-04

§2.3 METHODS AND APPLICATIONS

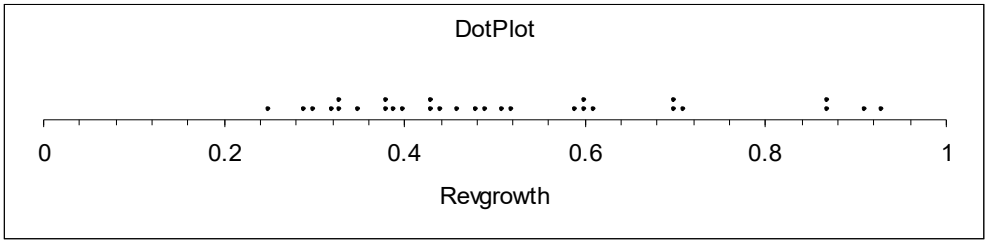
2.30



The distribution is concentrated between 0 and 2 and is skewed to the right. Eight and ten are probably high outliers.

LO02-04

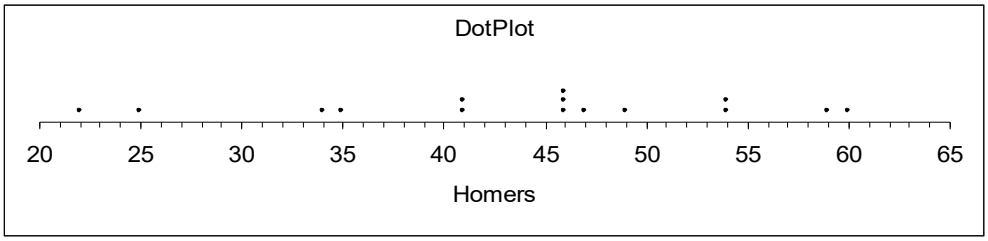
2.31



Most growth rates are no more than 71%, but 4 companies had growth rates of 87% or more.

LO02-04

2.32



Without the two low values (they might be outliers), the distribution is reasonably symmetric.

LO02-04

§2.4 CONCEPTS

2.33 Both the histogram and the stem-and-leaf show the shape of the distribution, but only the stem-and-leaf shows the values of the individual measurements.

LO02-03, LO02-05

2.34 Several advantages of the stem-and-leaf display include that it:
-Displays all the individual measurements.
-Puts data in numerical order
-Is simple to construct

LO02-05

2.35 With a large data set (e.g., 1,000 measurements) it does not make sense to do a stem-and-leaf because it is impractical to write out 1,000 data points. Group the data and use a histogram..

LO02-03, LO02-05

§2.4 METHODS AND APPLICATIONS

2.36 Stem Unit = 10, Leaf Unit = 1 ***Revenue Growth in Percent***

Frequency	*Stem*	*Leaf*
1	2	8
4	3	0 2 3 6
5	4	2 2 3 4 9
5	5	1 3 5 6 9
2	6	3 5
1	7	0
1	8	3
1	9	1
20		

LO02-05

2.37 Stem Unit = 1, Leaf Unit =.1 ***Profit Margins (%)***

Frequency	*Stem*	*Leaf*
2	10	4 4
0	11	
1	12	6
3	13	2 8 9
4	14	0 1 4 9
4	15	2 2 8 9
4	16	1 1 4 8
0	17	
0	18	
0	19	
0	20	
0	21	
1	22	2
0	23	
0	24	
1	25	2
20		

LO02-05

2.38 Stem Unit = 1000, Leaf Unit = 100 ***Sales ($mil)***

Frequency	*Stem*	*Leaf*
5	1	2 4 4 5 7
5	2	0 4 7 7 8
4	3	3 3 5 7
2	4	2 6
1	5	4
2	6	0 8
1	7	9

LO02-05

2.39 **a.** The Payment Times distribution is skewed to the right.

b. The Bottle Design Ratings distribution is skewed to the left.

LO02-05

2.40 **a.** The distribution is symmetric and centered near 50.7 pounds.

b. 46.8, 47.5, 48.2, 48.3, 48.5, 48.8, 49.0, 49.2, 49.3, 49.4

LO02-05

2.41 Stem unit = 10, Leaf Unit = 1 ***Home Runs***

Leaf	*Stem*	*Leaf*
Roger Maris		Babe Ruth
8	0	
6 4 3	1	
8 6 3	2	2 5
9 3	3	4 5
	4	1 1 6 6 6 7 9
	5	4 4 9
1	6	0

The 61 home runs hit by Maris would be considered an outlier for him, although an exceptional individual achievement.

LO02-05

2.42 **a.** Stem unit = 1, Leaf Unit = 0.1 ***Bank Customer Wait Time***

Frequency	*Stem*	*Leaf*
2	0	4 8
6	1	1 3 4 6 8 8
9	2	0 2 3 4 5 7 8 9 9
11	3	1 2 4 5 6 7 7 8 8 9 9
17	4	0 0 1 2 3 3 3 4 4 5 5 5 6 7 7 8 9
15	5	0 1 1 2 2 3 4 4 5 6 6 7 8 8 8
13	6	1 1 2 3 3 3 4 5 5 6 7 7 8
10	7	0 2 2 3 4 4 5 7 8 9
7	8	0 1 3 4 6 6 7
6	9	1 2 3 5 8 9
3	10	2 7 9
1	11	6
100		

b. The distribution of wait times is fairly symmetrical, may be slightly skewed to the right.

LO02-05

2.43 **a.** Stem unit = 1, Leaf Unit = 0.1 ***Video Game Satisfaction Ratings***

Frequency	*Stem*	*Leaf*
1	36	0
0	37	
3	38	0 0 0
4	39	0 0 0 0
5	40	0 0 0 0 0
6	41	0 0 0 0 0 0
6	42	0 0 0 0 0 0
8	43	0 0 0 0 0 0 0 0
12	44	0 0 0 0 0 0 0 0 0 0 0 0
9	45	0 0 0 0 0 0 0 0 0
7	46	0 0 0 0 0 0 0
3	47	0 0 0
1	48	0
65		

b. The video game satisfaction ratings distribution is slightly skewed to the left.

c. Since 19 of the 65 ratings (29%) are below 42 indicating very satisfied, it would *not* be accurate to say that almost all purchasers are very satisfied.

LO02-05

§2.5 CONCEPTS

2.44 Contingency tables are used to study the association between two variables.

LO02-06

2.45 We fill each cell of the contingency table by counting the number of observations that have both of the specific values of the categorical variables associated with that cell.

LO02-06

2.46 A row percentage is calculated by dividing the cell frequency by the total frequency for that particular row and by expressing the resulting fraction as a percentage.

A column percentage is calculated by dividing the cell frequency by the total frequency for that particular column and by expressing the resulting fraction as a percentage.

Row percentages show the distribution of the column categorical variable for a given value of the row categorical variable.

Column percentages show the distribution of the row categorical variable for a given value of the column categorical variable.

LO02-06

§2.5 METHODS AND APPLICATIONS

2.47 Cross tabulation of Brand Preference vs. Purchase History

Brand Preference		Purchased? No	Yes	*Total*
Koka	**Observed**	**14**	**2**	***16***
	% of row	87.5%	12.5%	100%
	% of column	66.7%	10.5%	40%
	% of total	35.0%	5.0%	40%
Rola	**Observed**	**7**	**17**	***24***
	% of row	29.2%	70.8%	100%
	% of column	33.3%	89.5%	60%
	% of total	17.5%	42.5%	60%
Total	***Observed***	***21***	***19***	***40***
	% of row	52.5%	47.5%	100%
	% of column	100.0%	100.0%	100%
	% of total	52.5%	47.5%	100%

a. 17 shoppers who preferred Rola-Cola had purchased it before.

b. 14 shoppers who preferred Koka-Cola had not purchased it before.

c. If you have purchased Rola previously you are more likely to prefer Rola.
If you have not purchased Rola previously you are more likely to prefer Koka.

LO02-06

2.48 Cross tabulation of Brand Preference vs. Sweetness Preference

Brand Preference		Sweetness Preference Very Sweet	Sweet	Not So Sweet	*Total*
Koka	**Observed**	**6**	**4**	**6**	***16***
	% of row	37.5%	25.0%	37.5%	100%
	% of column	42.9%	30.8%	46.2%	40%
	% of total	15.0%	10.0%	15.0%	40%
Rola	**Observed**	**8**	**9**	**7**	***24***
	% of row	33.3%	37.5%	29.2%	100%
	% of column	57.1%	69.2%	53.8%	60%
	% of total	20.0%	22.5%	17.5%	60%
Total	***Observed***	***14***	***13***	***13***	***40***
	% of row	35.0%	32.5%	32.5%	100%
	% of column	100.0%	100.0%	100.0%	100%
	% of total	35.0%	32.5%	32.5%	100%

a. 8 + 9 = 17 shoppers who preferred Rola-Cola also preferred their drinks Sweet or Very Sweet.

b. 6 shoppers who preferred Koka-Cola also preferred their drinks not so sweet.

c. Rola drinkers may prefer slightly sweeter drinks than Koka drinkers.

LO02-06

2.49 Cross tabulation of Brand Preference vs. Number of 12-Packs Consumed Monthly

Brand Preference		Consumption 0 to 5	6 to 10	>10	Total
Koka	Observed	12	3	1	*16*
	% of row	75.0%	18.8%	6.3%	100%
	% of column	60.0%	17.6%	33.3%	40%
	% of total	30.0%	7.5%	2.5%	40%
Rola	Observed	8	14	2	*24*
	% of row	33.3%	58.3%	8.3%	100%
	% of column	40.0%	82.4%	66.7%	60%
	% of total	20.0%	35.0%	5.0%	60%
Total	*Observed*	*20*	*17*	*3*	*40*
	% of row	50.0%	42.5%	7.5%	100%
	% of column	100.0%	100.0%	100.0%	100%
	% of total	50.0%	42.5%	7.5%	100%

a. 8 + 14 = 22 shoppers who preferred Rola-Cola purchase 10 or fewer 12-packs.

b. 3 + 1 = 4 shoppers who preferred Koka-Cola purchase 6 or more 12-packs.

c. People who drink more cola seem more likely to prefer Rola.

LO02-06

2.50 **a.** 16%, 56%

b.

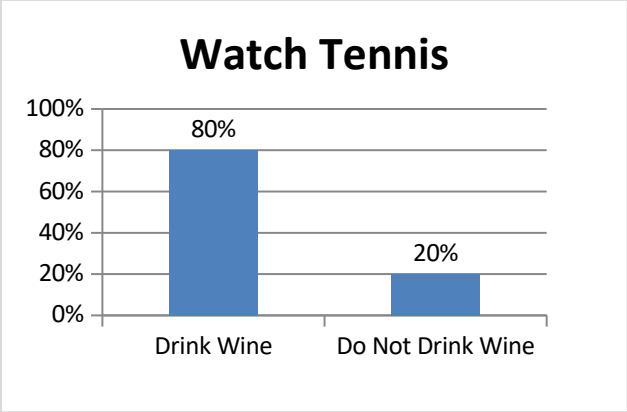
Row Percentage Table	Watch Tennis	Do Not Watch Tennis	Total
Drink Wine	40%	60%	100%
Do Not Drink Wine	6.7%	93.3%	100%

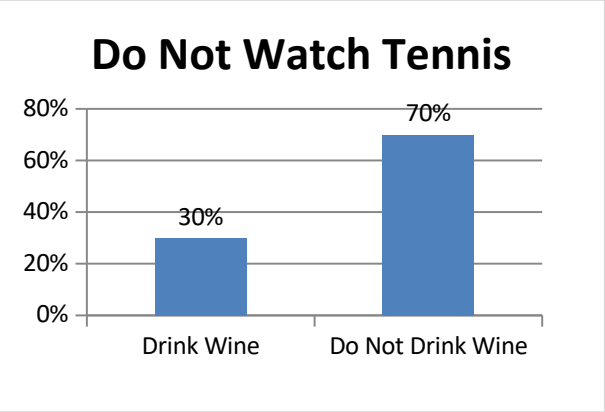
c.

Column Percentage Table	Watch Tennis	Do Not Watch Tennis
Drink Wine	80%	30%
Do Not Drink Wine	20%	70%
Total	100%	100%

d. People who watch tennis are more likely to drink wine than those who do not watch tennis..

e.





LO02-01, LO02-06

2.51 **a.**

	TV Violence		
TV Quality	Increased	Not Increased	Total
Worse	362	92	454
Not Worse	359	187	546
Total	721	279	1000

b. Row percentages

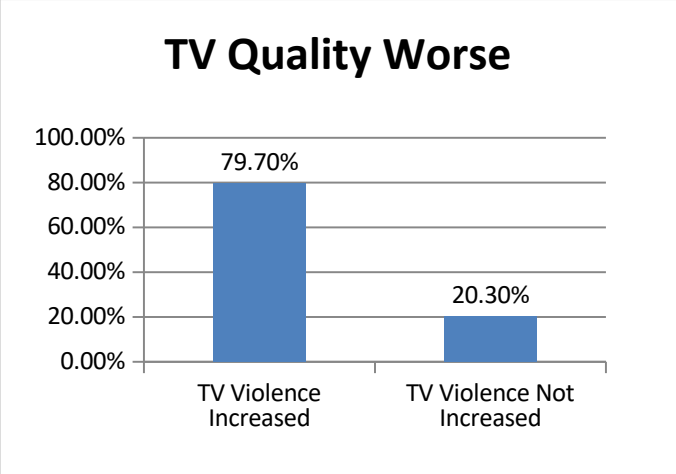
	TV Violence		
TV Quality	Increased	Not Increased	Total
Worse	79.7%	20.3%	100%
Not Worse	65.8%	34.2%	100%

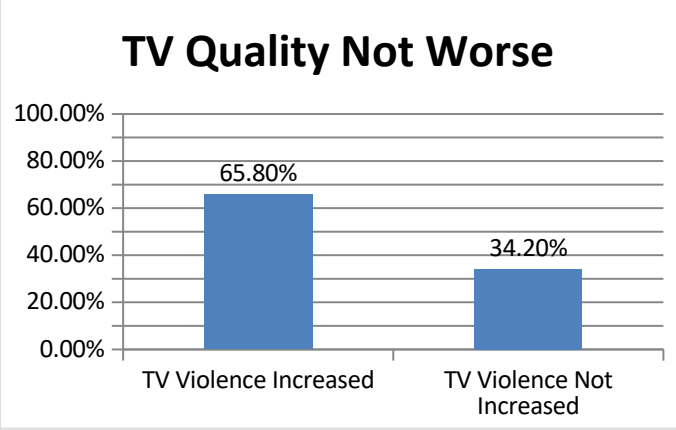
c. Column percentages

	TV Violence	
TV Quality	Increased	Not Increased
Worse	50.2%	33.0%
Not Worse	49.8%	67.0%
Total	100.0%	100.0%

d. Those people who think TV violence has increased are more likely to think TV quality has gotten worse.

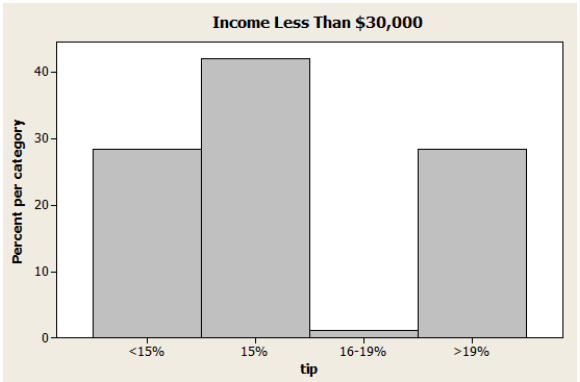
e.

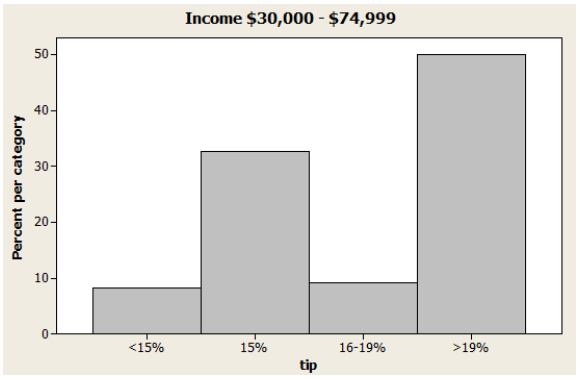


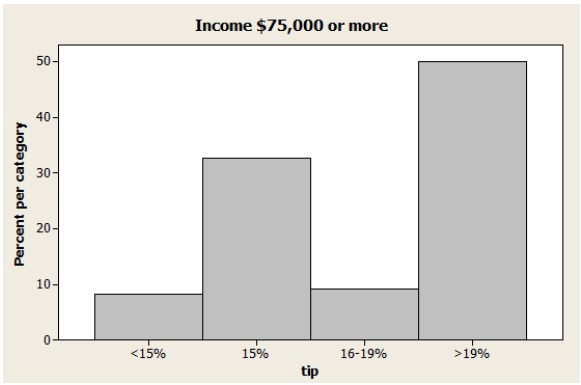


LO02-01, LO02-06

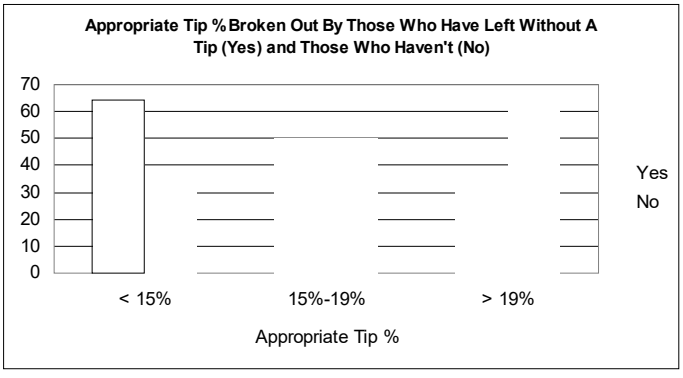
2.52 a. As income rises the percent of people seeing larger tips as appropriate also rises.







b. People who have left at least once without leaving a tip are more likely to think a smaller tip is appropriate.



LO02-01, LO02-06

§2.6 CONCEPTS

2.53 A scatterplot is used to look at the relationship between two quantitative variables.

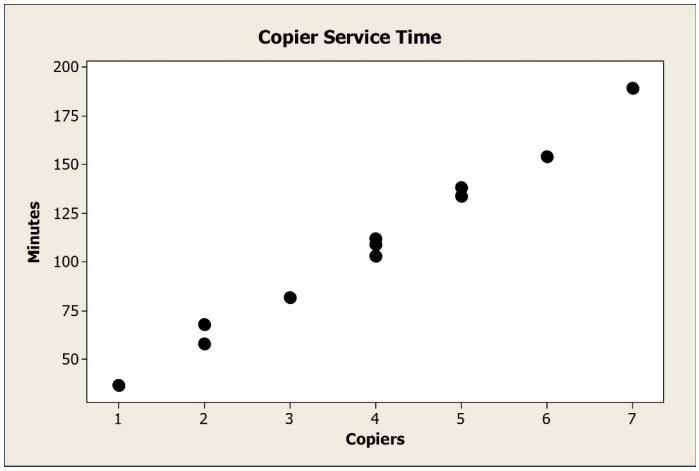
LO02-07

2.54 On a scatter plot, each value of y is plotted against its corresponding value of x.
On a times series plot, each individual process measurement is plotted against its corresponding time of occurrence.

LO02-07

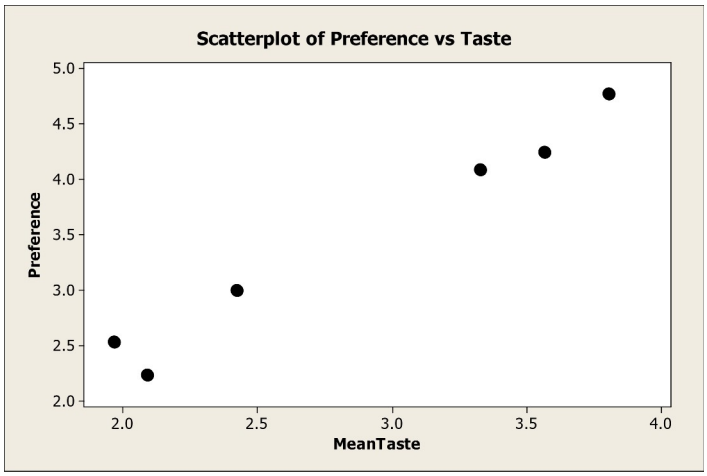
§2.6 METHODS AND APPLICATIONS

2.55 As the number of copiers increases, so does the service time.

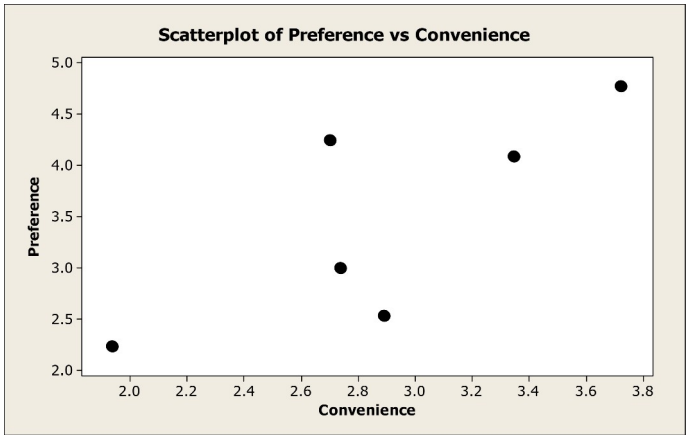


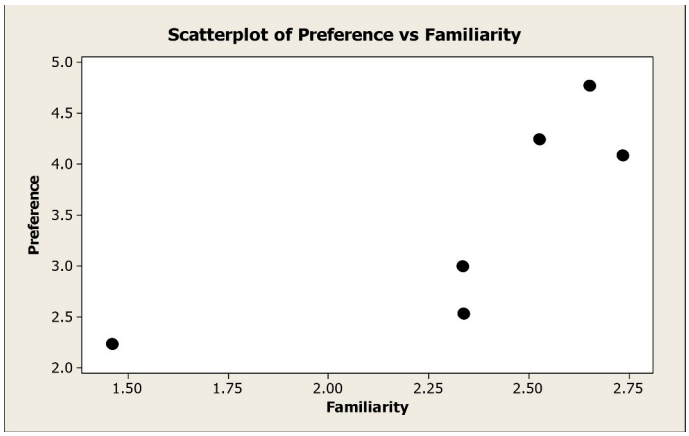
LO02-07

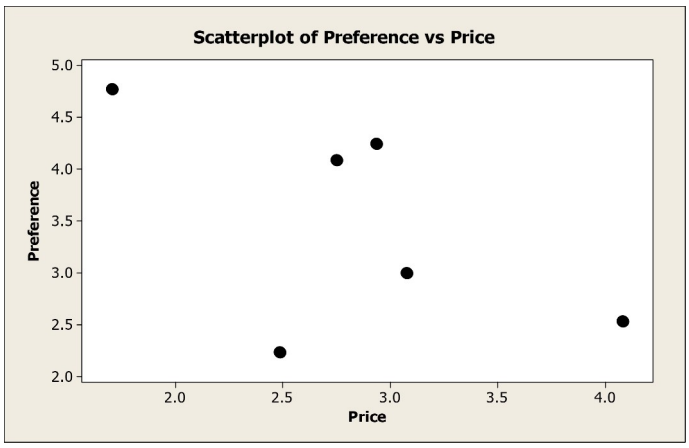
2.56 The scatterplot shows that the average rating for taste is related to the average rating for preference in a positive linear fashion. This relationship is fairly strong.



2.56 (*cont*.) The scatterplots below show that average convenience, familiarity, and price are all approximately linearly related to average preference in a positive, positive, and negative fashion (respectively). These relationships are not as strong as the one between taste and preference.







LO02-07

2.57 Cable rates decreased in the early 1990's in an attempt to compete with the newly emerging satellite business. As the satellite business was increasing its rates from 1995 to 2005, cable was able to do the same.

LO02-07

§2.7 CONCEPTS

2.58 When the vertical axis does not start at zero, the bars themselves will not be as tall as if the bars had started at zero. Hence, the relative differences in the heights of the bars will be more pronounced.

LO02-08

2.59 Examples and reports will vary.

LO02-08

§2.7 METHODS AND APPLICATIONS

2.60 The administration's compressed plot indicates a steep increase of nurses' salaries over the four years, while the union organizer's stretched plot shows a more gradual increase of the same salaries over the same time period.

LO02-08

2.61 **a.** No. The graph of the number of private elementary schools is showing only a very slight (if any) increasing trend when scaled with public schools.

b. Yes. The graph of the number of private elementary schools is showing strong increasing trend, particularly after 1950.

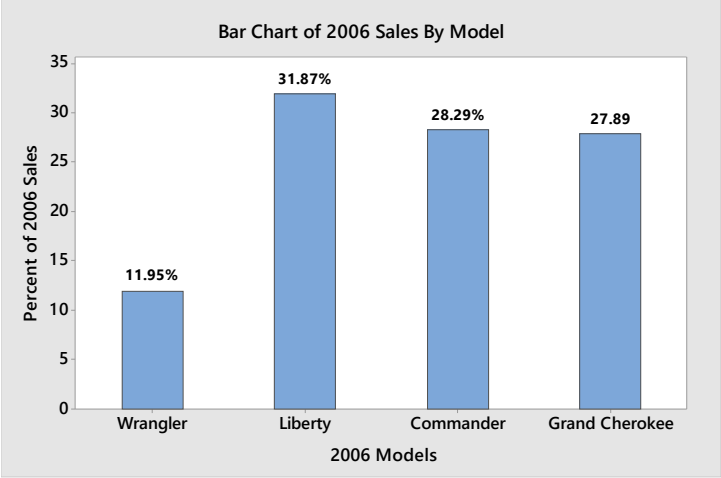
c. The line graph is more appropriate because it shows growth.

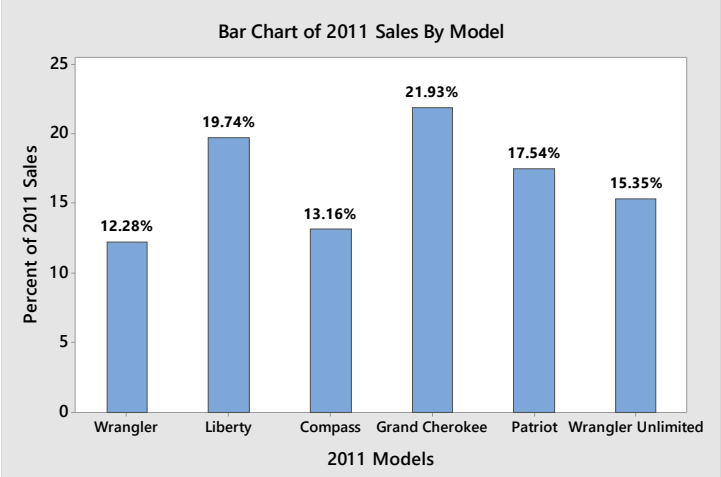
d. Neither graph gives an accurate understanding of the changes spanning a half century. Because of the very large difference in scale between private and public schools, a comparison of growth might be better described using percent increase.

LO02-08

SUPPLEMENTARY EXERCISES

2.62





Reports will vary but should mention that although Liberty sales declined, this is not surprising since Liberty was one of 4 models in 2006 but one of 6 in 2011. As the dealer's second most popular model in 2011, it is still an important part of his sales.

LO02-01

2.63 A large portion of manufacturers are rated 3 for Overall Mechanical Quality. No US cars received ratings above 3.

Overall Mechanical Quality	*frequency*
2	6
3	23
4	2
5	2
	33

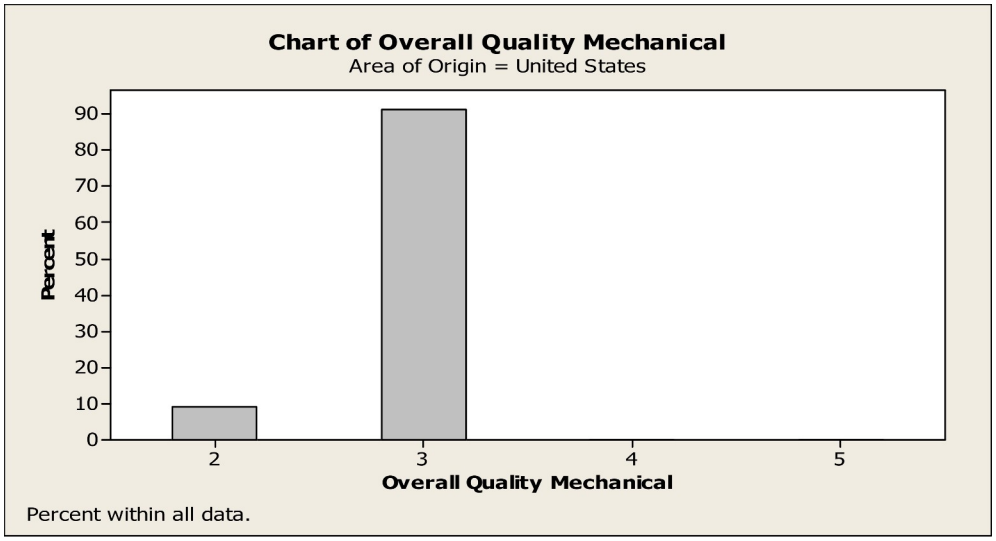
LO02-01

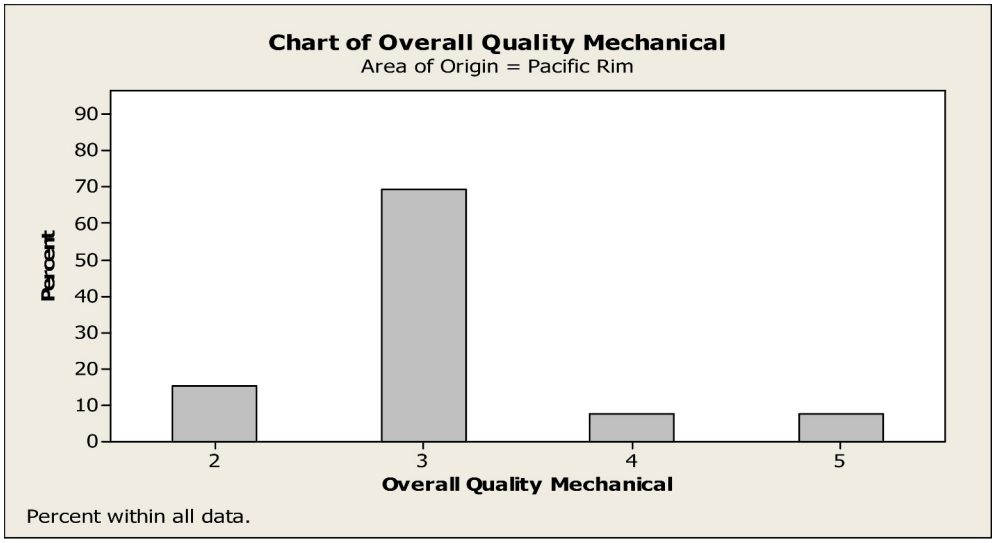
2.64 No Pacific Rim company received a 2 while US companies received 3 of the 4 ratings of 2 for overall design quality.

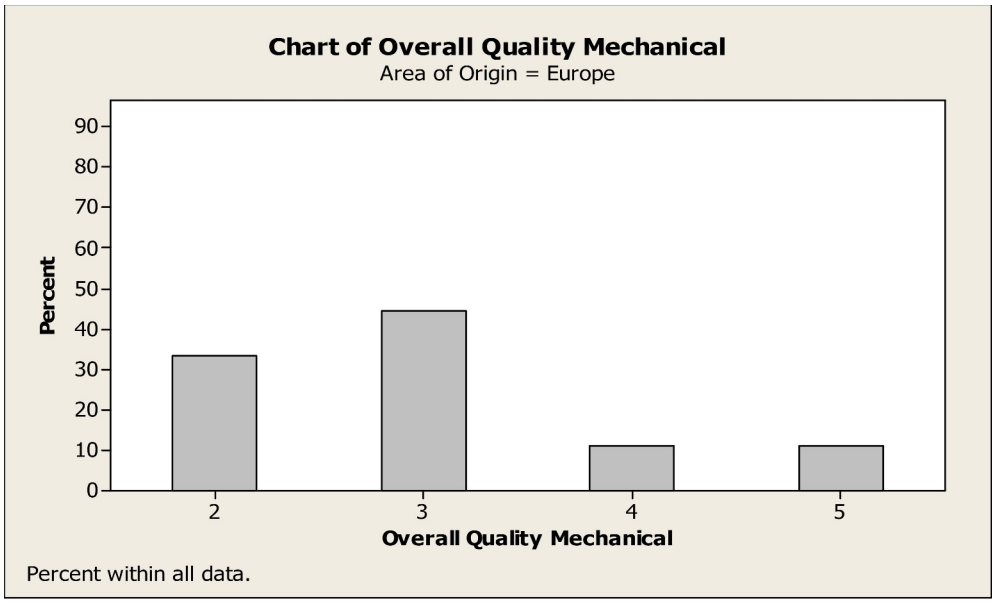
Overall Design Quality	*frequency*	*relative frequency*
2	4	0.12
3	22	0.67
4	6	0.18
5	1	0.03
	33	100.00

LO02-01

2.65 Average was the most frequent rating for all 3 regions. 10 of 11 US ratings were average; better than average ratings went only to Pacific Rim & European companies, but each region had more than 1 in the below average category.

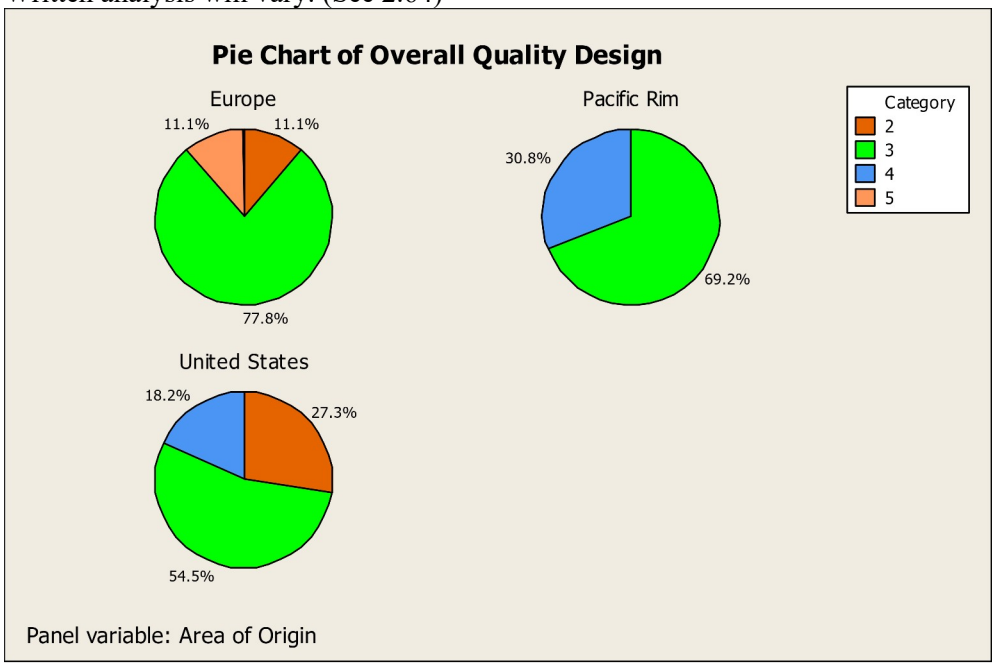






LO02-01

2.66 Written analysis will vary. (See 2.64)



LO02-01

2.67 & 2.68

Overall Quality Mechanical

Area of Origin	Among the Best	Better than Most	About Average	The Rest	Total
Europe	1	1	4	3	9
	11.11%	11.11%	44.44%	33.33%	100%
Pacific Rim	1	1	9	2	13
	7.69%	7.69%	69.23%	15.38%	100%
United States	0	0	10	1	11
	0%	0%	90.91%	9.09%	100%
Total	2	2	23	6	33
	6.06%	6.06%	69.70%	18.18%	100%

Only Europe and the Pacific Rim have above average ratings, but the US is the least likely to receive the lowest rating. LO02-06

2.68 Written reports will vary. See 2.65 for percentage bar charts. See 2.67 for row percentages.

LO02-06

2.69 & 2.70

Overall Quality Design

Area of Origin	2	3	4	5	Total
Europe	1	7	0	1	9
	11.11%	77.78%	0%	11.11%	100%
Pacific Rim	0	9	4	0	13
	0%	69.23%	30.77%	0%	100%
United States	3	6	2	0	11
	27.27%	54.55%	18.18%	0%	100%
Total	4	22	6	1	33
	12.12%	66.67%	18.18%	3.03%	100%

LO02-06

2.70 Written reports will vary. See 2.66 for pie charts. See 2.69 for row percentages.

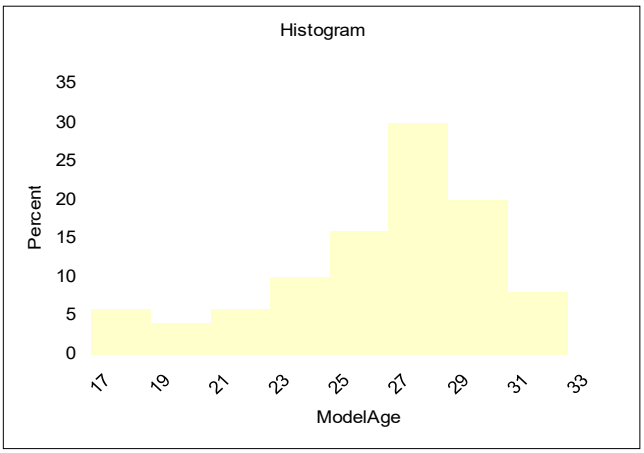
LO02-06

2.71 **a.** Frequency Distribution for ***Model Age***

Lower		*Upper*	*Midpoint*	*Width*	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
17	<	19	18	2	3	6	3	6
19	<	21	20	2	2	4	5	10
21	<	23	22	2	3	6	8	16
23	<	25	24	2	5	10	13	26
25	<	27	26	2	8	16	21	42
27	<	29	28	2	15	30	36	72
29	<	31	30	2	10	20	46	92
31	<	33	32	2	4	8	50	100
					50	100		

While the 2^K rule suggests using 6 classes, we are using 8 as suggested in the problem.

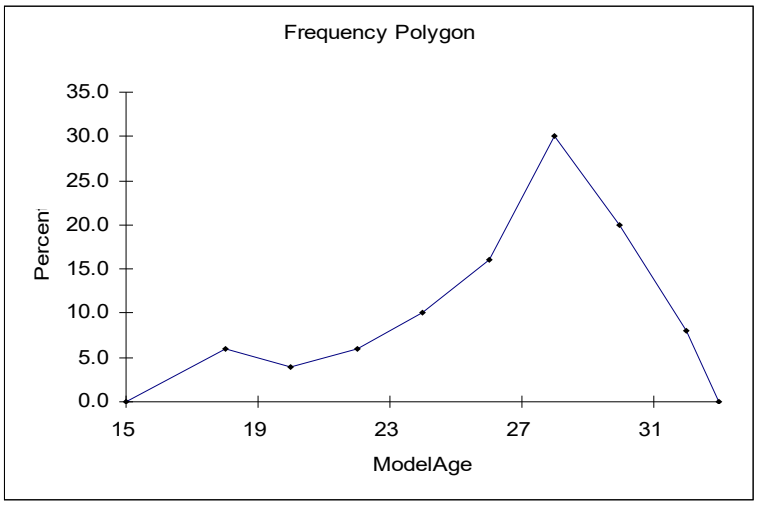
b.



c. This distribution is skewed to the left.

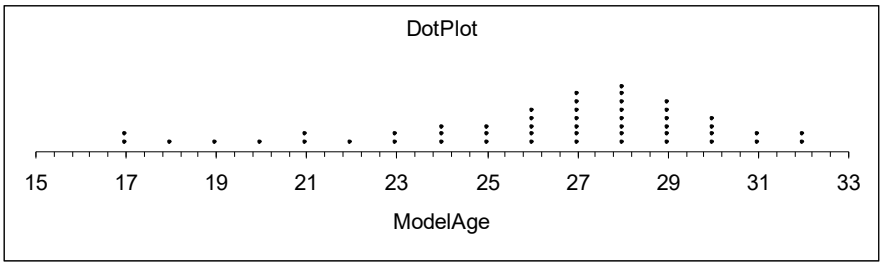
LO02-03

2.72



LO02-03

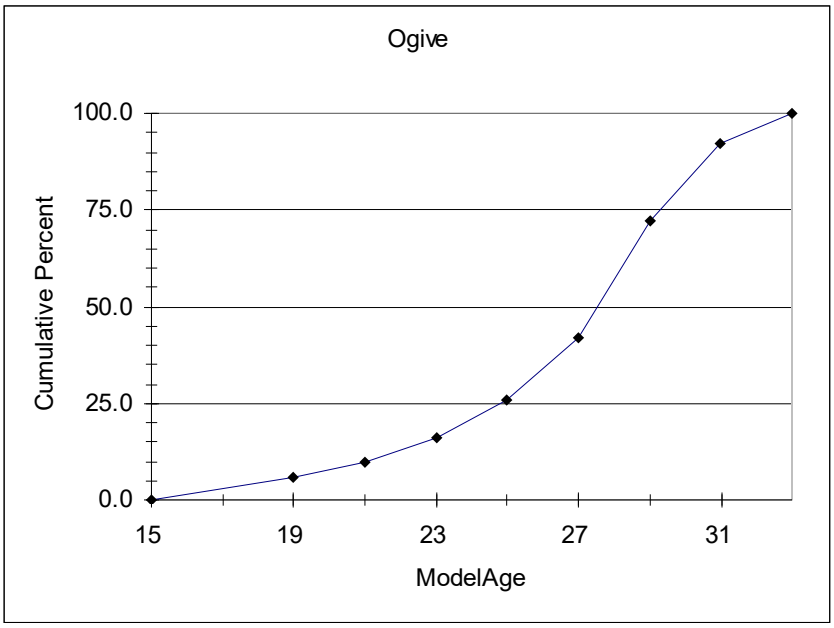
2.73 26% of the perceived ages are below 25. This is probably too high.



LO02-04

2.74 **a & b & c.** See table in 2.71

d.

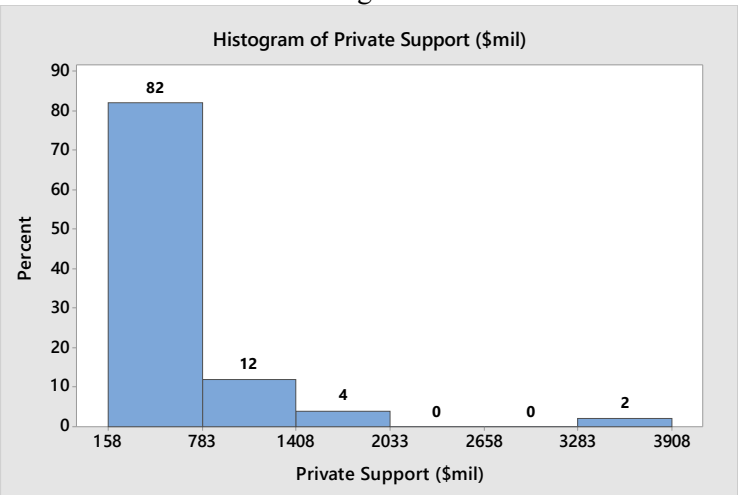


e. 36 out of 50 = 72%

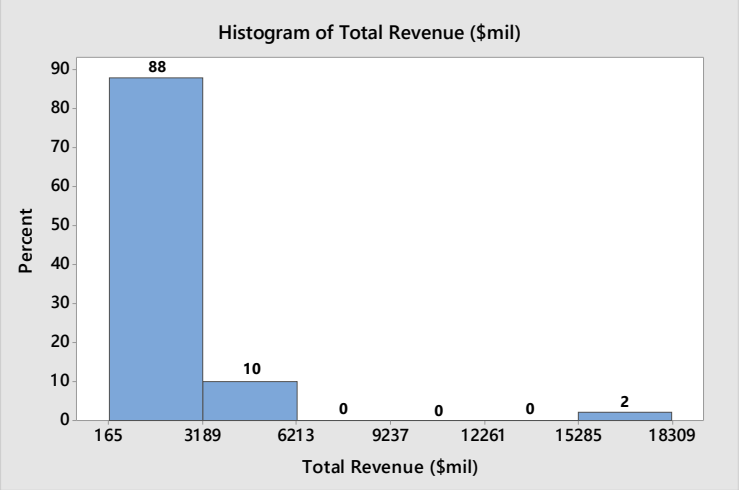
f. 8 out of 50 = 16%

LO02-03

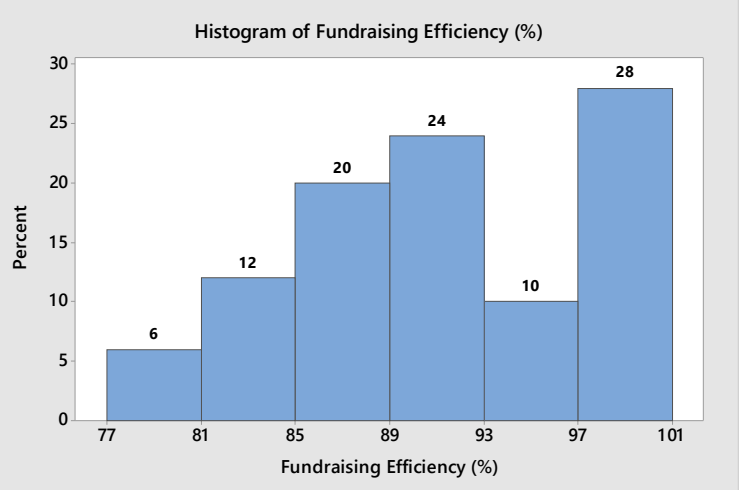
2.75 Distribution is skewed to the right



Distribution is skewed to the right



Distribution is skewed to the left



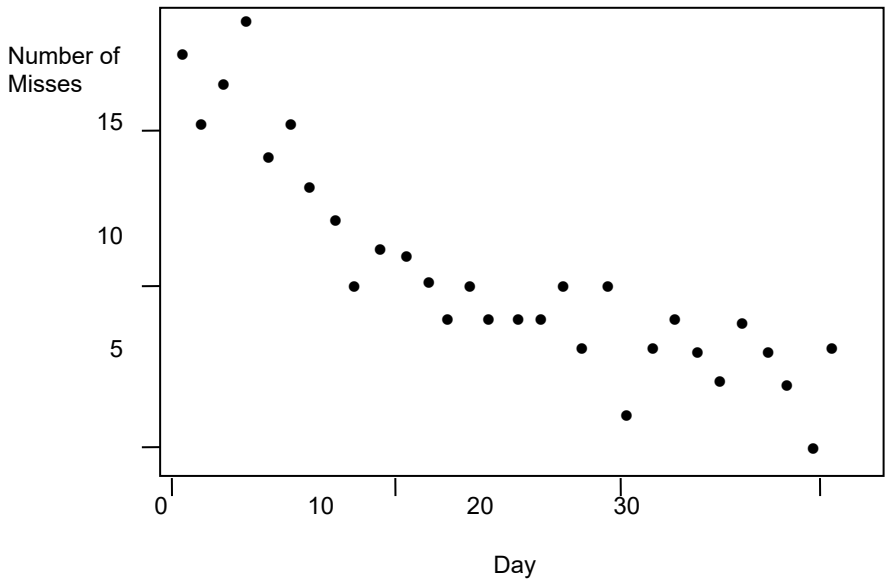
LO02-03

2.76 Distribution has one high outlier and with or without the outlier is skewed right.

LO02-04

2.77 Stem Unit = 1, Leaf Unit = 0.1 ***Shots Missed***.

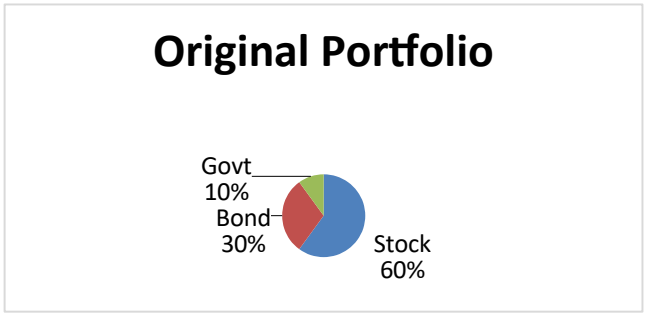
Frequency	Stem	Leaf
1	5	0
2	6	0
4	7	0 0
9	8	0 0 0 0 0
15	9	0 0 0 0 0 0
15	10	0 0 0 0 0
10	11	0 0
8	12	0
7	13	0
6	14	0
5	15	0 0
3	16	0
2	17	0
1	18	0
30		



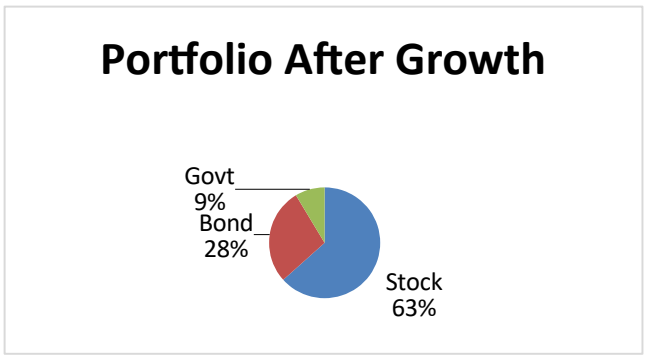
The time series plot shows that the player is improving over time. Therefore the stem-and-leaf display does not predict how well the player will shoot in the future.

LO02-05

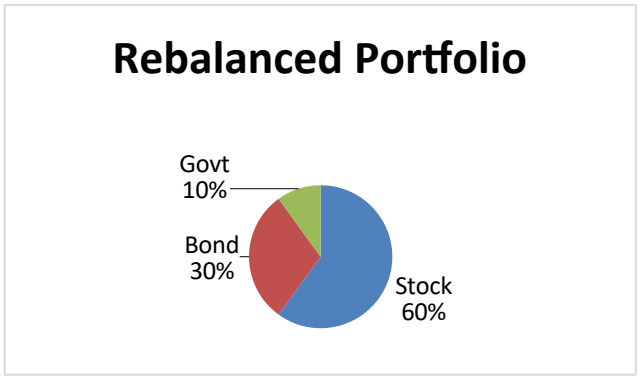
2.78 **a.** Stock funds: $60,000; bond funds: $30,000; govt. securities: $10,000



b. Stock funds: $78,000 (63.36%); Bond funds: $34,500 (28.03%); Govt. securities: $10,600 (8.61%)



c. Stock funds: $73,860; Bond funds: $36,930; Govt. securities: $12,310



LO02-01

2.79 The graph indicates that Chevy trucks far exceed Ford and Dodge in terms of resale value, but the y-axis scale is misleading.

LO02-08

INTERNET EXERCISES

2.80 Answers will vary depending on which poll(s) the student refers to.

LO02-01 – LO02-08