CHAPTER 2—Descriptive Statistics: Tabular and Graphical Methods

2.1   Constructing either a frequency or a relative frequency distribution helps identify and quantify patterns in how often various categories occur.

L02-01

2.2   Relative frequency of any category is calculated by counting the number of occurrences of the category divided by the total number of observations. Percent frequency is calculated by multiplying relative frequency by 100.
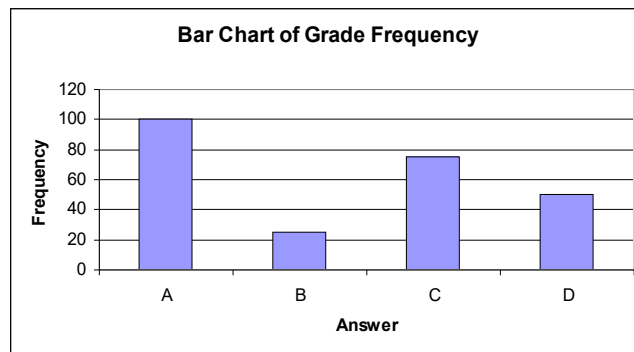
L02-01
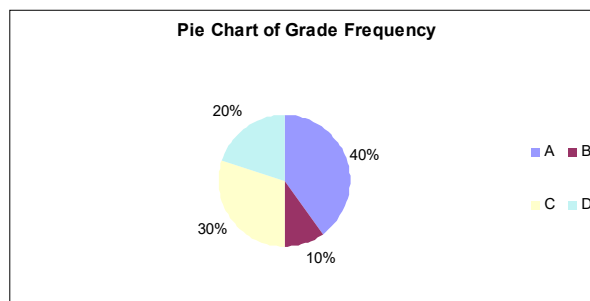
2.3   Answers and examples will vary.

L02-01

2.4   **a.**

Category / Class	Frequency	Relative Frequency	Percent Frequency
A	100	0.40	40%
B	25	0.10	10%
C	75	0.30	30%
D	50	0.20	20%

Bar Chart of Grade Frequency

   b.

L02-01

2.5   **a.**   (100 / 250) * 360 degrees = 144 degrees

   b.   (25 / 250) * 360 degrees = 36 degrees

Pie Chart of Grade Frequency

   c.
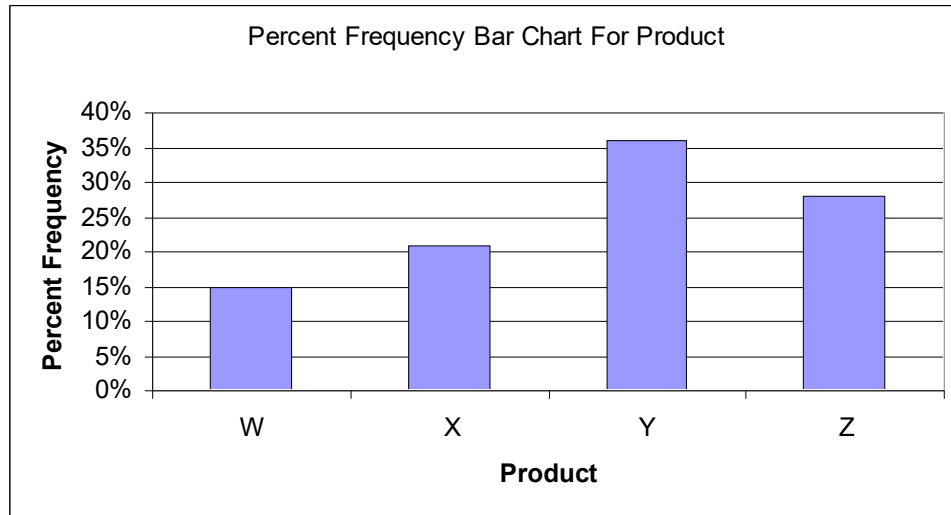
L02-01

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

      b. Product: W   X   Y   Z
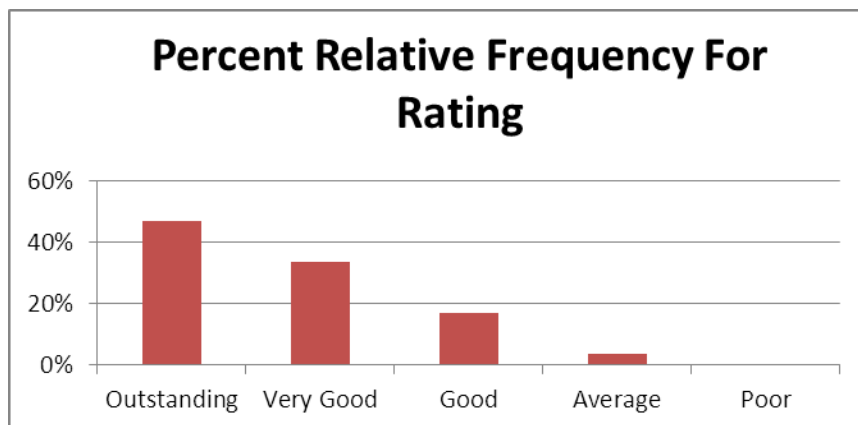   75   105   180   140

      c.

      d. Degrees for W would be 54, for X degrees would be 75.6, and for Y 129.6, and for Z 100.8.
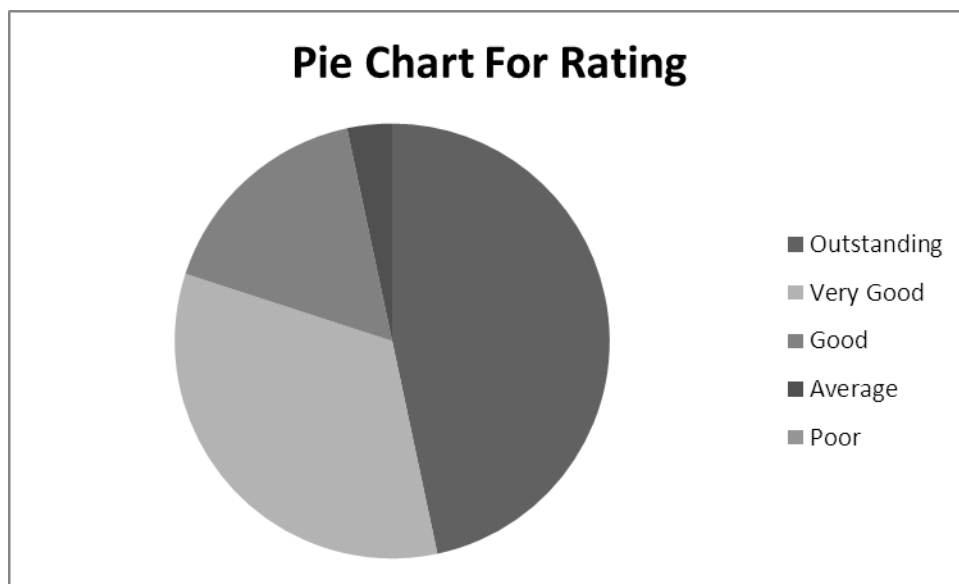
      L02-01

2.7 **a.**

Rating	Frequency	Relative Frequency
Outstanding	14	0.467
Very Good	10	0.333
Good	5	0.167
Average	1	0.033
Poor	0	0.000

      b.

c.

Pie Chart For Rating

- Outstanding
- Very Good
- Good
- Average
- Poor

L02-01

2.8 a. Tally for Discrete Variables: Sports League

Sports League	Count	Rel. Freq.	Percent
MLB	11	0.22	22.00
MLS	3	0.06	6.00
NBA	8	0.16	16.00
NFL	23	0.46	46.00
NHL	5	0.10	10.00
N=	50		

b.

Chart of Sports League

c.

Pie Chart of Sports League
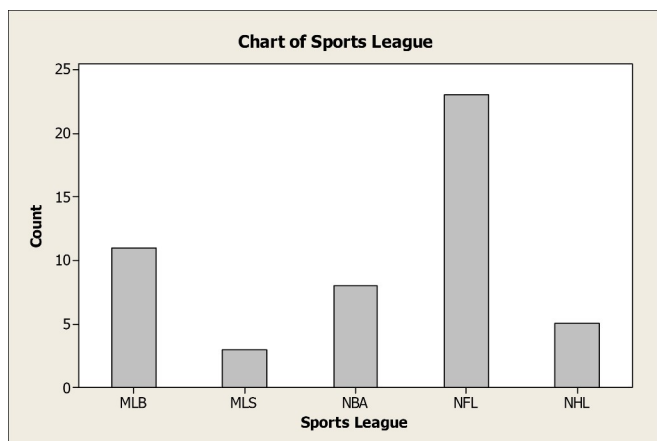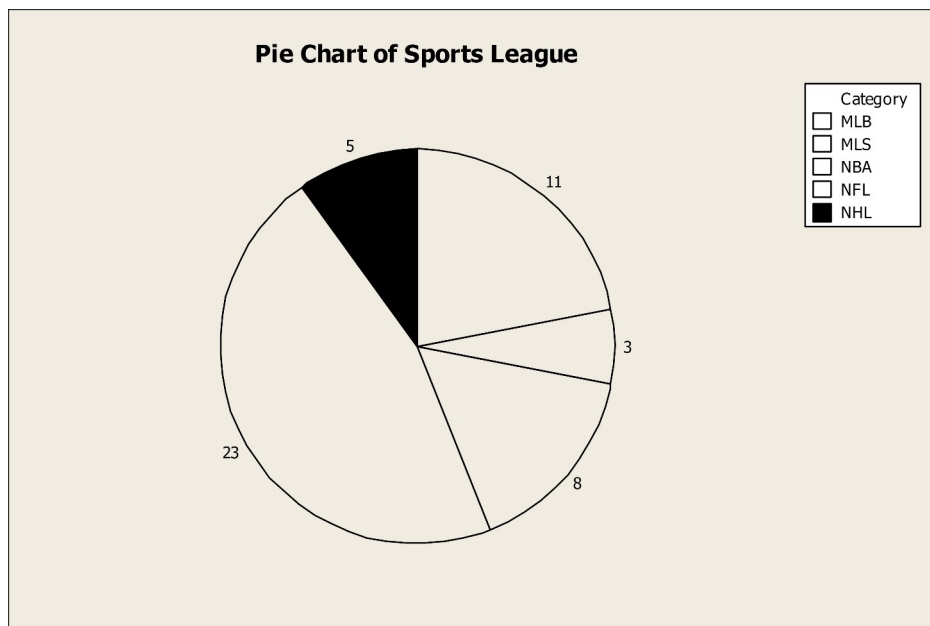
d. Most popular league is NFL and least popular is MLS.

L02-01

2.9

US Market Share In 2005

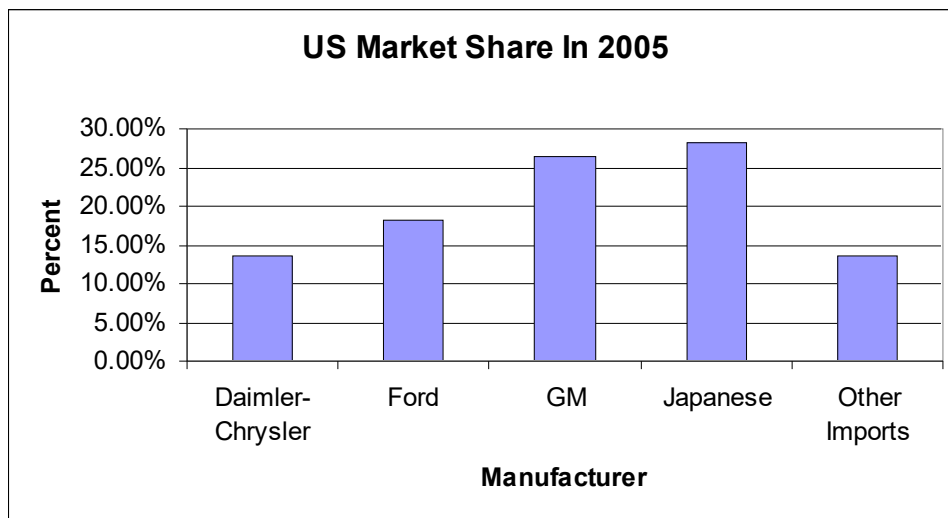**US Market Shares In 2005**

14%   14%

Daimler-Chrysler Jeep

18%

Ford

GM

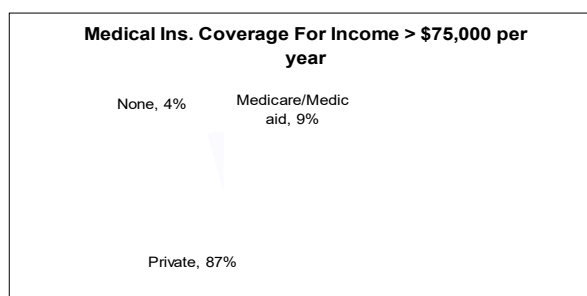28%

Japanese

26%

Other Imports
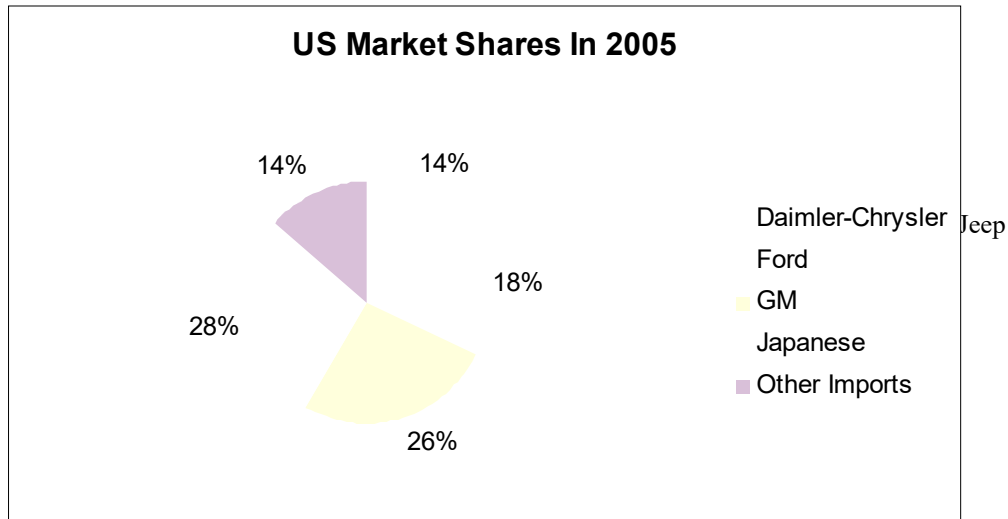
L02-01

2.10   Comparing the two pie charts they show that since 2005 Ford & GM, have lost market share, while Chrysler and Japanese models have increased market share.

L02-01

2.11

Comparing Types of Health Insurance Based on Income Level

100%
80%   87%
60%
40%   33%   50%
20%   17%
0%   4%
       Priv. Ins.   Medi /   No Ins.
                     Medi

◼ Inc. < $30,000
◼ Inc. > $75,000

9%

Medical Ins. Coverage For Income < $30,000 per year

None, 17%   Medicare/Medic aid, 33%

Private, 50%

Medical Ins. Coverage For Income > $75,000 per year

None, 4%   Medicare/Medic aid, 9%

Private, 87%
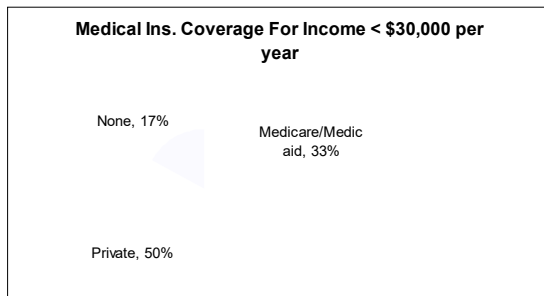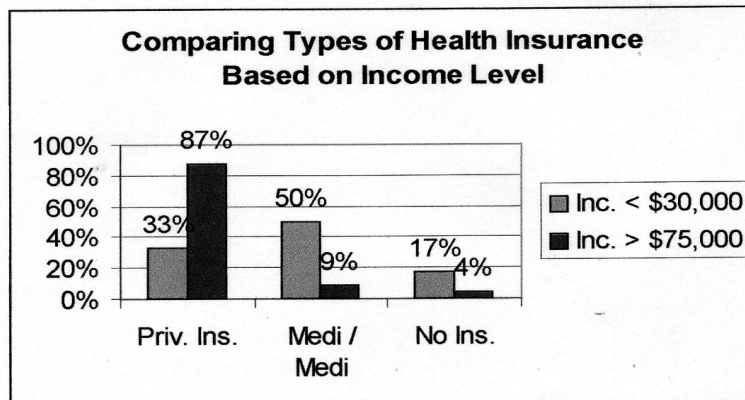
L02-01

2.12   **a.**   32.29%

         b.   4.17%

         c.   Explanations will vary

L02-02

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 frequency in a class using bars while in a frequency polygon the frequencies in consecutive classes are connected by a line.

         c.   A frequency ogive represents a cumulative distribution while the frequency polygon is not a cumulative distribution. Also, in a frequency polygon the lines connect the class midpoints while in a frequency ogive the lines connect the upper boundaries of the classes.
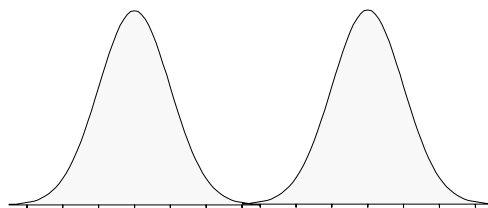
L02-03

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

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

         c.   Percent frequency for a class is calculated by multiplying the relative frequency by 100.
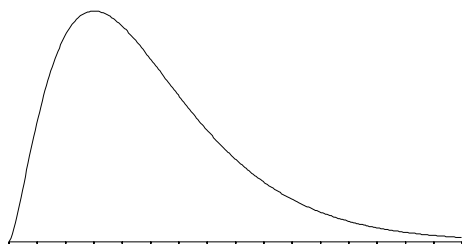
L02-03

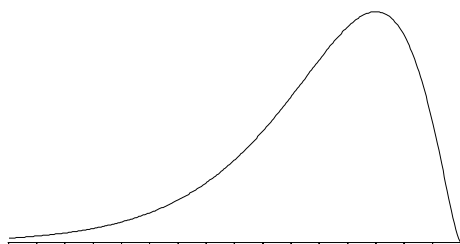**2.15**   **a.**   One hump in the middle; left side looks like right side.
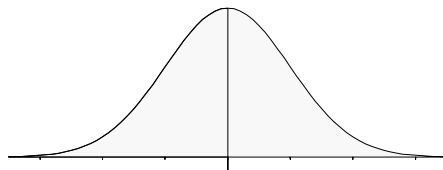
         b.   Two humps, left side may or may not look like right side.

c. Long tail to the right

d. Long tail to the left

L02-03

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

b. Class Length (CL) = (46 − 17) / 5 = 6

c. $17 \le x < 23$, $23 \le x < 29$, $29 \le x < 35$, $35 \le x < 41$, $41 \le x < 47$

d.

Frequency Distribution -
Quantitative

| | *Data* | | | | *frequency* | | cumulative | |
lower		*upper*	*midpoint*	*width*		*percent*	*frequency*	*percent*
12	<	18	15	6	1	3.6	1	3.6
18	<	24	21	6	3	10.7	4	14.3
24	<	30	27	6	2	7.1	6	21.4
30	<	36	33	6	5	17.9	11	39.3
36	<	42	39	6	14	50.0	25	89.3
42	<	48	45	6	3	10.7	28	100.0
					28	100.0		

e.
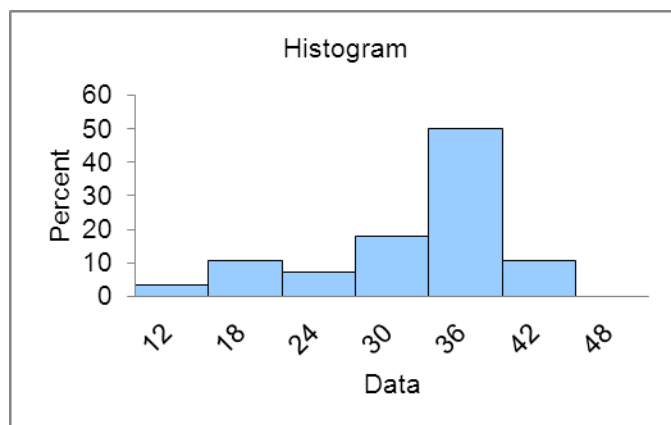
f. See output in answer to d.

L02-03

2.17a & b.

Class	Frequency	Cum Frequency	Percent Frequency	Cum % Frequency
50 < 60	2	2	4%	4%
60 < 70	5	7	10%	14%
70 < 80	14	21	28%	42%
80 < 90	17	38	34%	76%
90 < 100	12	50	24%	100%
Total	50	50	100%	

c.

Frequency Polygon

d.

Ogive
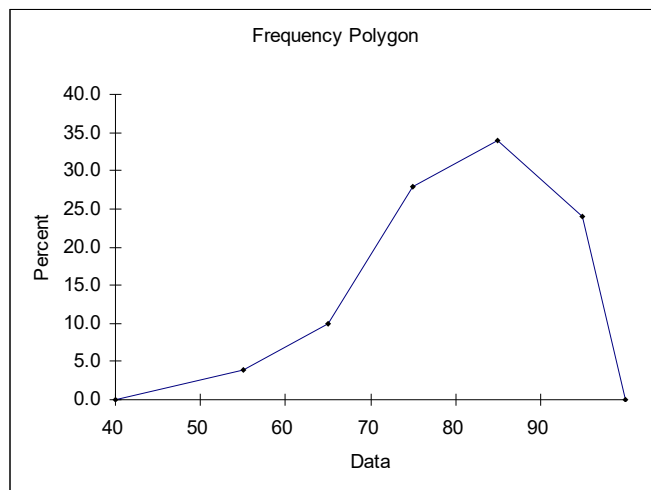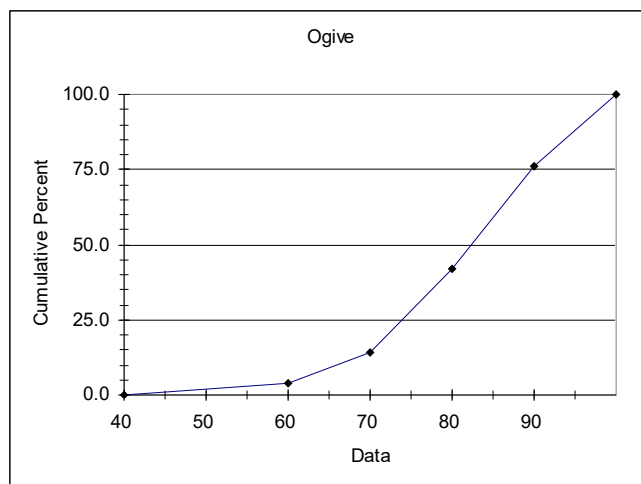
L02-03

2.18 a.   6 classes because there are 60 data points (from Table 2.5).

b.   Class Length (CL) = (35 − 20) / 6 = 2.5 and we round up to 3.

c. $20 \leq x < 23, 23 \leq x < 26, 26 \leq x < 29, 29 \leq x < 32, 32 \leq x < 35, 35 \leq x < 38$

d.

Rating			*midpoint*	*width*	*frequency*	*Percent*	cumulative	
lower		*upper*					*frequency*	*percent*
20	<	23	21.5	3	2	3.3	2	3.3

23	<	26	24.5	3	3	5.0	5	8.3
26	<	29	27.5	3	9	15.0	14	23.3
29	<	32	30.5	3	19	31.7	33	55.0
32	<	35	33.5	3	26	43.3	59	98.3
35	<	38	36.5	3	1	1.7	60	100.0
					60	100.0		

e.

Distribution shape is skewed left.

L02-03

2.19a & b.

Rating			midpoint	width	frequency	Percent	cumulative frequency	percent
lower		upper						
20	<	23	21.5	3	2	3.3	2	3.3
23	<	26	24.5	3	3	5.0	5	8.3
26	<	29	27.5	3	9	15.0	14	23.3
29	<	32	30.5	3	19	31.7	33	55.0
32	<	35	33.5	3	26	43.3	59	98.3
35	<	38	36.5	3	1	1.7	60	100.0
					60	100.0		

c.

L02-03

2.20a & b & c.

Frequency Distribution

| Pay ($mil) | | | | | frequency | | cumulative frequency | |
low er		upper	midpoint	width	y	percent	y	percent
20	<	40	30	20	2	8.0	2	8.0
40	<	60	50	20	7	28.0	9	36.0
60	<	80	70	20	7	28.0	16	64.0
80	<	100	90	20	3	12.0	19	76.0
100	<	120	110	20	2	8.0	21	84.0
120	<	140	130	20	2	8.0	23	92.0
140	<	160	150	20	0	0.0	23	92.0
160	<	180	170	20	0	0.0	23	92.0
180	<	200	190	20	1	4.0	24	96.0
200	<	220	210	20	0	0.0	24	96.0
220	<	240	230	20	0	0.0	24	96.0
240	<	260	250	20	0	0.0	24	96.0
260	<	280	270	20	0	0.0	24	96.0
280	<	300	290	20	1	4.0	25	100.0
					25	100.0		
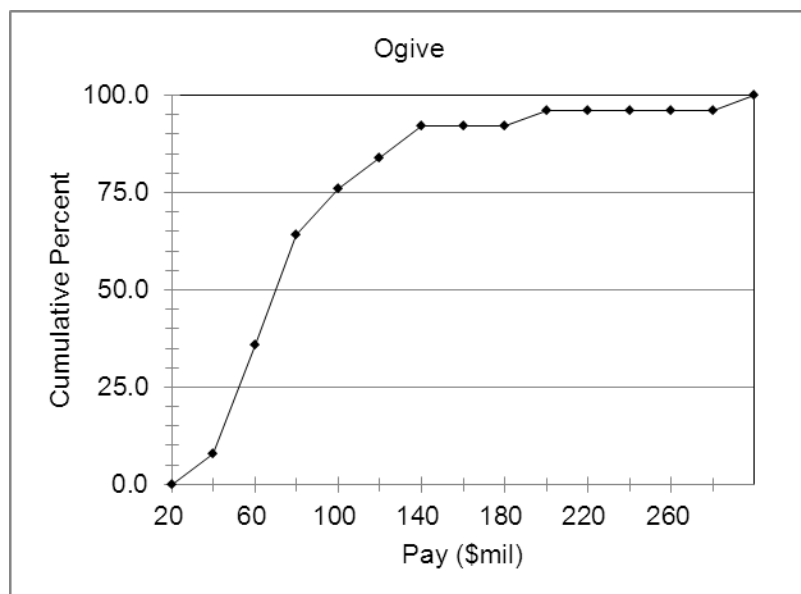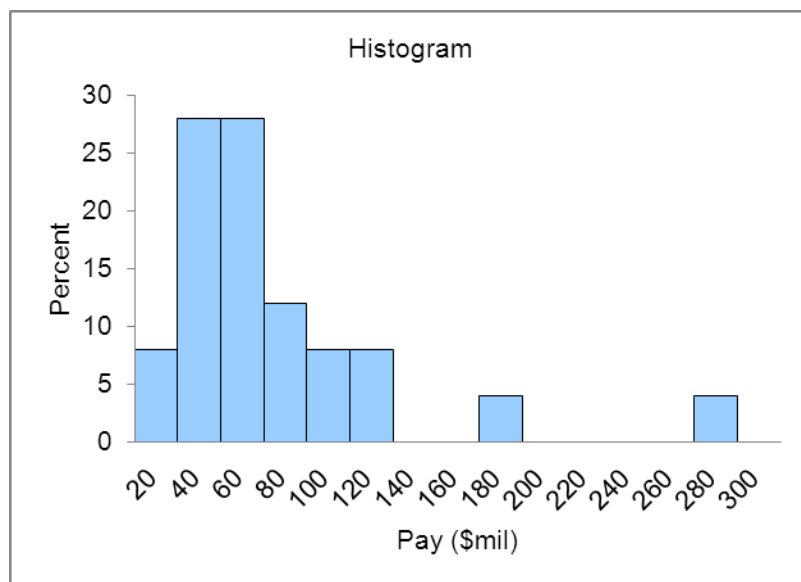
L02-03

2.21 a. Concentrated between 42 and 46.

b. Shape of distribution is slightly skewed left. Ratings have an upper limit but stretch out to the low side.

c.

Class	1	2	3	4	5	6	7	8
	$34 < x \le 36$,	$36 < x \le 38$,	$38 < x \le 40$,	$40 < x \le 42$,	$42 < x \le 44$,	$44 < x \le 46$,	$46 < x \le 48$,	more

d.

Class	1	2	3	4	5	6	7	8
Cum Freq	1	4	13	25	45	61	65	65

L02-03

2.22 **a.** Concentrated between 3.5 and 5.5.

b. 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.

d.
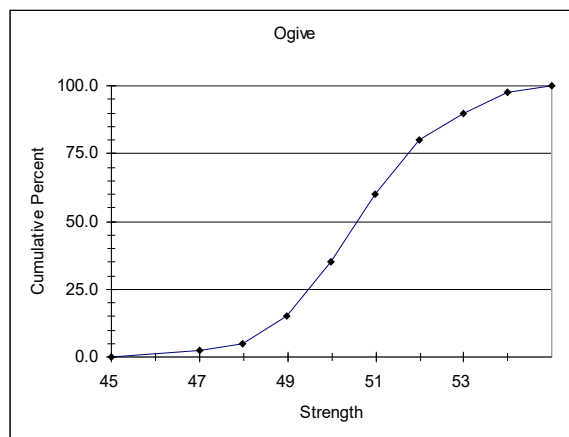| Class	Cum Frequency
-0.5< 0 .5	1
0.5< 1.5	5
1.5< 2.5	12
2.5< 3.5	20
3.5< 4.5	37
4.5< 5.5	53
5.5< 6.5	67
6.5< 7.5	79
7.5< 8.5	87
8.5< 9.5	93
9.5<10.5	97
10.5<11.5	99
11.5<12.5	100

L02-03

2.23 **a.** Concentrated between 48 and 53.

b. Shape of distribution is symmetric and bell shaped.

c. Class length is 1.

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%

L02-03

2.24   a.

	Value		midpoint	width	frequency
lower		upper			
300	<	400	350	100	11
400	<	500	450	100	8
500	<	600	550	100	5
600	<	700	650	100	1
700	<	800	750	100	2
800	<	900	850	100	1
900	<	1,000	950	100	1
1,000	<	1,100	1,050	100	0
1,100	<	1,200	1,150	100	0
1,200	<	1,300	1,250	100	0
1,300	<	1,400	1,350	100	0
1,400	<	1,500	1,450	100	0
1,500	<	1,600	1,550	100	0
1,600	<	1,700	1,650	100	0
1,700	<	1,800	1,750	100	1

30

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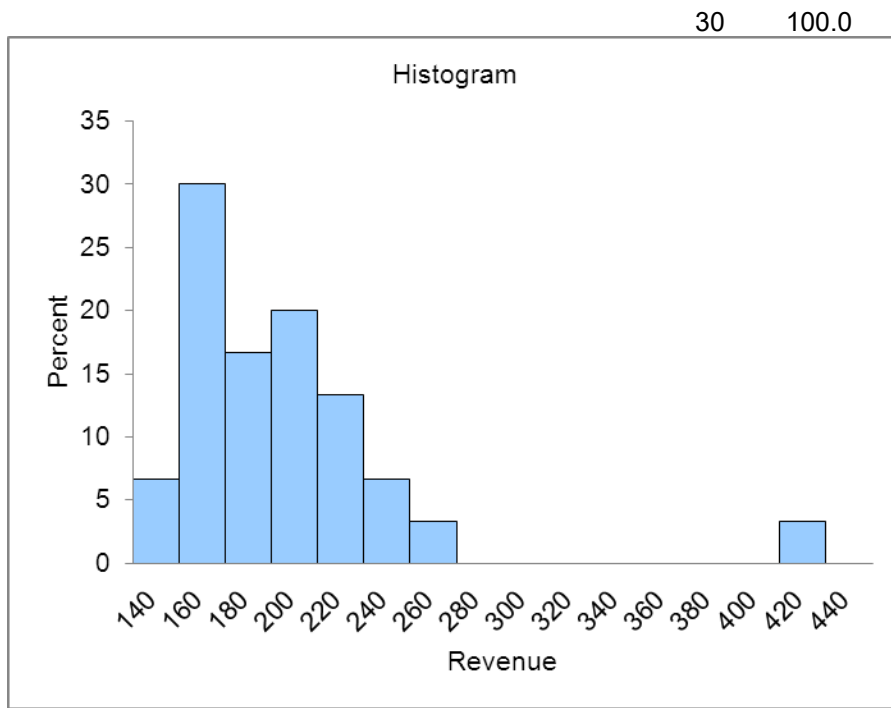
b.

Frequency Distribution - Quantitative

Revenue							cumulative	
low er		*uppe r*	*midpoint*	*width*	*frequenc y*	*percent*	*frequenc y*	*percent*
140	<	160	150	20	2	6.7	2	6.7
160	<	180	170	20	9	30.0	11	36.7
180	<	200	190	20	5	16.7	16	53.3
200	<	220	210	20	6	20.0	22	73.3
220	<	240	230	20	4	13.3	26	86.7
240	<	260	250	20	2	6.7	28	93.3
260	<	280	270	20	1	3.3	29	96.7
280	<	300	290	20	0	0.0	29	96.7
300	<	320	310	20	0	0.0	29	96.7
320	<	340	330	20	0	0.0	29	96.7
340	<	360	350	20	0	0.0	29	96.7
360	<	380	370	20	0	0.0	29	96.7
380	<	400	390	20	0	0.0	29	96.7
400	<	420	410	20	0	0.0	29	96.7
420	<	440	430	20	1	3.3	30	100.0
					30	100.0		

Distribution is skewed right.

c.

Frequency Polygon
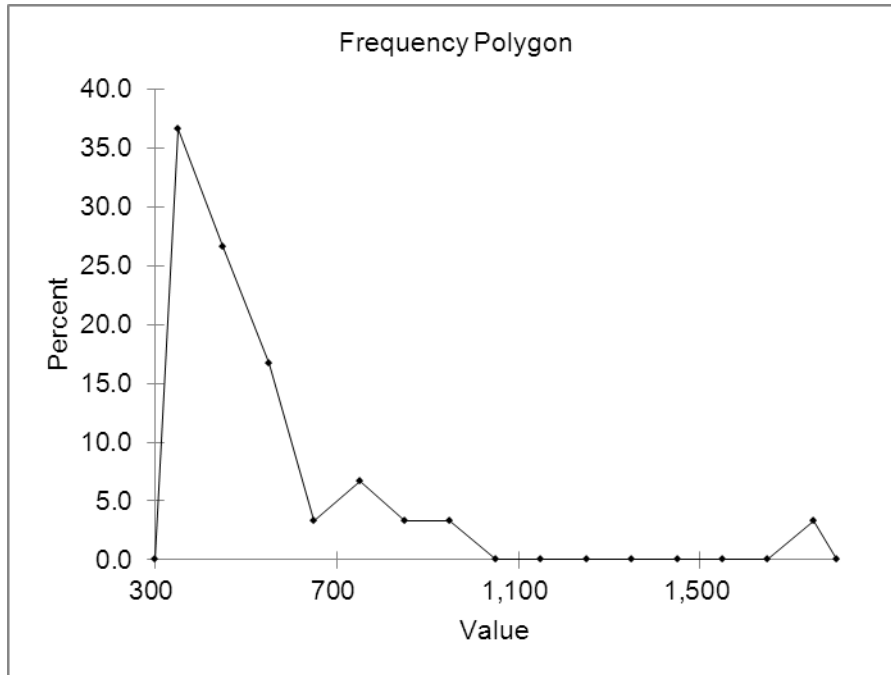
L02-03

2.25 a.

Frequency Distribution - Quantitative

| Sales ($mil)			midpoint	width	frequency	percent	cumulative frequency	percent
lower		upper						
0	<	100	50	100	3	7.5	3	7.5
100	<	200	150	100	4	10.0	7	17.5
200	<	300	250	100	6	15.0	13	32.5
300	<	400	350	100	5	12.5	18	45.0
400	<	500	450	100	4	10.0	22	55.0
500	<	600	550	100	4	10.0	26	65.0
600	<	700	650	100	4	10.0	30	75.0
700	<	800	750	100	1	2.5	31	77.5
800	<	900	850	100	4	10.0	35	87.5
900	<	1,000	950	100	5	12.5	40	100.0
					40	100.0		

Histogram
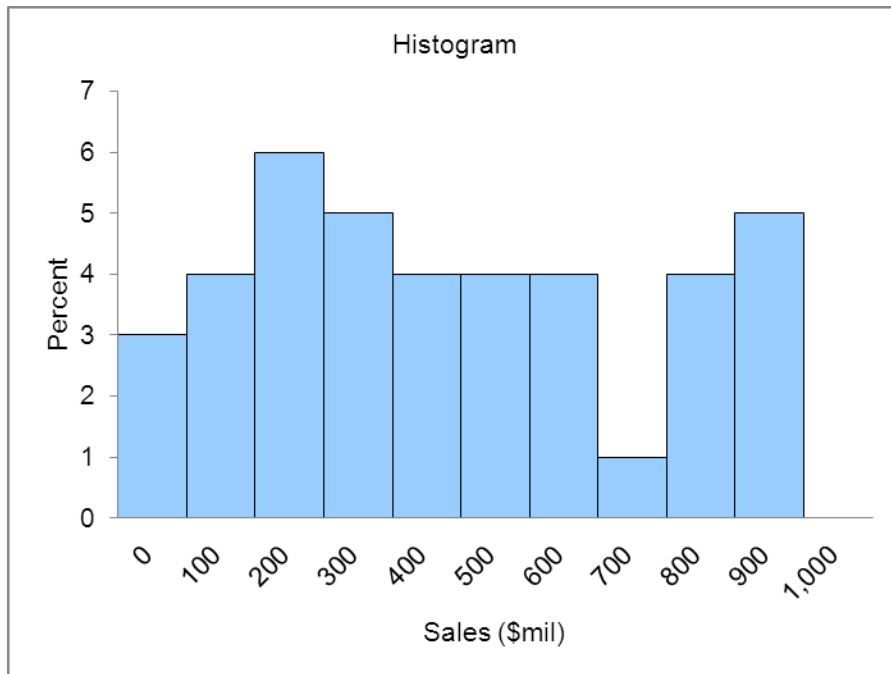
Distribution is relatively flat, perhaps two humped.

b.

Frequency Distribution - Quantitative

Sales Growth (%) lo wer		upp er	midpoi nt	widt h	frequen cy	percen t	cumulative frequen cy	perce nt
0	<	10	5	10	4	10.0	4	10.0
10	<	20	15	10	3	7.5	7	17.5
20	<	30	25	10	13	32.5	20	50.0
30	<	40	35	10	11	27.5	31	77.5
40	<	50	45	10	4	10.0	35	87.5
50	<	60	55	10	2	5.0	37	92.5
60	<	70	65	10	2	5.0	39	97.5
70	<	80	75	10	0	0.0	39	97.5
80	<	90	85	10	1	2.5	40	100.0
					40	100.0		
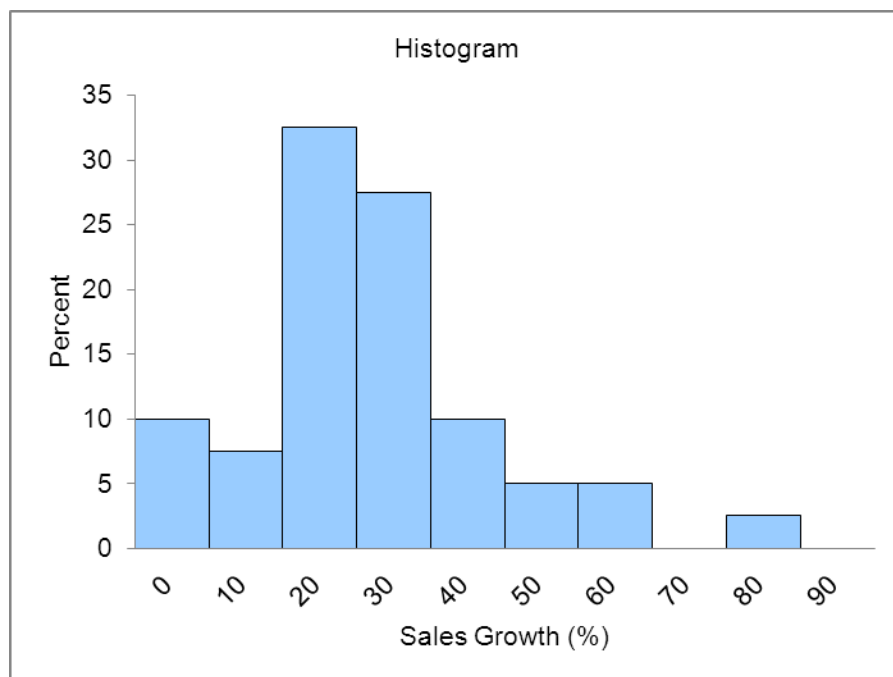
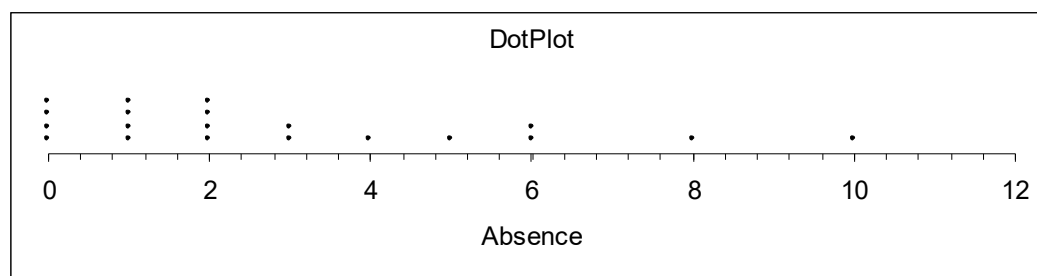Distribution is skewed right.

L02-03

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

L02-04

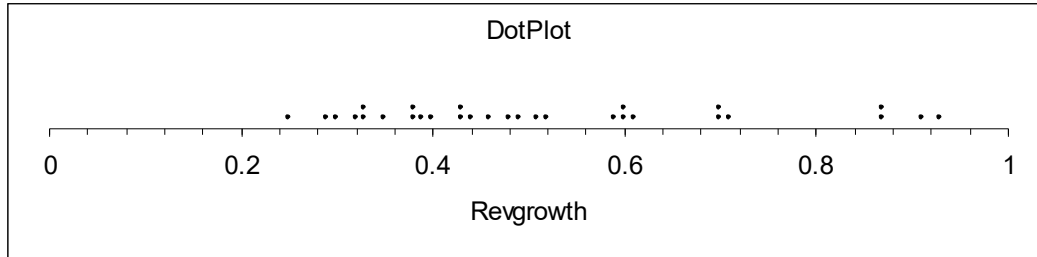**2.27**   A dot plot with a 1000 points is not practical. Use a histogram.

L02-03, L02-04

2.28

Distribution is concentrated between 0 and 2 and is skewed to the right. 10 and 8 are probably high outliers.

L02-04

2.29

DotPlot

[DotPlot of Revgrowth with axis labeled 0, 0.2, 0.4, 0.6, 0.8, 1]

Revgrowth

High outliers greater than 80%. Eliminating the high outliers the distribution is reasonably symmetric.

L02-04

2.30

DotPlot

[DotPlot of Homers with axis labeled 20, 25, 30, 35, 40, 45, 50, 55, 60, 65]

Homers
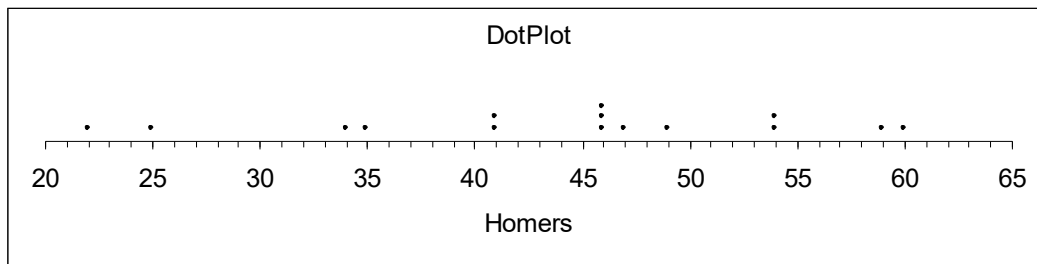
Low outliers 22 and 25. Without outliers distribution is reasonably symmetric.

L02-04

2.31   A stem & leaf enables one to see the shape of the distribution and still see all the measurements where in a histogram you cannot see the values of the individual measurements.

L02-03, L02-05

2.32   --Displays all the individual measurements.
--Puts data in numerical order
--Simple to construct

L02-05

2.33   With a large data set (eg 1000 measurements) it does not make sense to do a stem & leaf because it is impractical to write out 1000 leafs. Should use a histogram.

L02-03, L02-05

2.34

Stem Unit = 10, Leaf Unit = 1

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
		3
2	6	5
1	7	0
1	8	3
1	9	1
20		

L02-05

2.35

Stem Unit = 1, Leaf Unit = .1

Frequency	Stem	Leaf
		4
2	10	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		

L02-05

2.36   Rounding each measurement to the nearest hundred yields the following stem & leaf

Stem unit = 1000, Leaf Unit = 100

Frequency	Stem	Lea f
5	1	2 4 4 5 7
5	2	0 4 7 7 8
4	3	3 3 5 7
		2
2	4	6
1	5	4
		0
2	6	8
1	7	9
20		

L02-05

2.37 a.   Payment times distribution is skewed to the right.

b.   Bottle design ratings distribution is skewed to the left.

L02-05

2.38 a.   Distribution is symmetric

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

L02-05

2.39

```
Roger Maris │ 0 │ Babe Ruth
          8 │ 0 │
        4 3 │ 1 │
          6 │ 1 │
          3 │ 2 │ 2
        8 6 │ 2 │ 5
```

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

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

L02-05

2.40 a.

	WaitTime
count	100

Stem and Leaf plot
for   WaitTime
stem unit =   1
leaf unit =   0.1

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
<u>1</u>	11	6
100		

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

L02-05

2.41 a.

Stem and Leaf plot for   Ratings
stem unit =   1
leaf unit =   0.1

Descriptive statistics

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. Distribution is slightly skewed to the left.

c. Since 19 of the ratings are below 42 it would not be accurate to say that almost all purchasers are very satisfied.

L02-05

2.42 Cross tabulation tables are used to study association between categorical variables.

L02-06

2.43 Each cell is filled with the number of observations that have the specific values of the categorical variables associated with that cell.

L02-06

2.44 Row percentages are calculated by dividing the cell frequency by the total frequency for that particular row. Column percentages are calculated by dividing the cell frequency by the total frequency for that particular column. 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.
L02-06

2.45

Crosstabulation

			Purchased?		
			No	Yes	Total
	Koka	Observed	**14**	**2**	16
		% of row	87.5%	12.5%	100.0%
		% of column	66.7%	10.5%	40.0%
Preference		% of total	35.0%	5.0%	40.0%
	Rola	Observed	**7**	**17**	24
		% of row	29.2%	70.8%	100.0%
		% of column	33.3%	89.5%	60.0%

	% of total	17.5%	42.5%	60.0%
Total	Observed	21	19	40
	% of row	52.5%	47.5%	100.0%
	% of column	100.0%	100.0%	100.0%
	% of total	52.5%	47.5%	100.0%

a. 17 **b.** 14

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.

L02-06

2.46

Crosstabulation

			Preference			
			Very Sweet	Sweet	Not So Sweet	Total
Preference	Koka	Observed	**6**	**4**	**6**	16
		% of row	37.5%	25.0%	37.5%	100.0%
		% of column	42.9%	30.8%	46.2%	40.0%
		% of total	15.0%	10.0%	15.0%	40.0%
	Rola	Observed	**8**	**9**	**7**	24
		% of row	33.3%	37.5%	29.2%	100.0%
		% of column	57.1%	69.2%	53.8%	60.0%
		% of total	20.0%	22.5%	17.5%	60.0%
	Total	Observed	14	13	13	40
		% of row	35.0%	32.5%	32.5%	100.0%
		% of column	100.0%	100.0%	100.0%	100.0%
		% of total	35.0%	32.5%	32.5%	100.0%

a. 17 **b.** 6

c. No relationship.

L02-06

2.47

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

a.   22   b.   4
c.   People who drink more cola are more likely to prefer Rola.
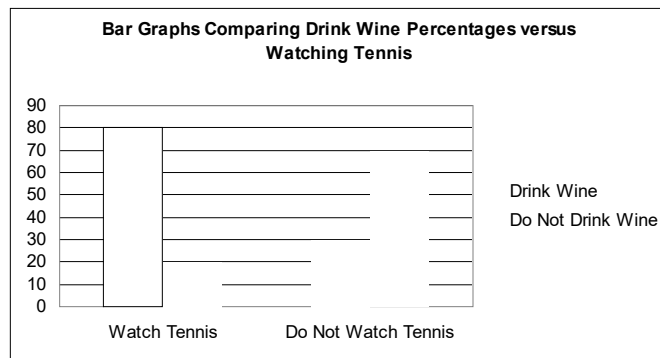
L02-06

2.48 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.

e.

L02-01, L02-06
2.49
a.

	TV Violence Inc.	TV Violence No Inc.	Total
TV Quality Worse	362	92	454
TV Quality Not Worse	359	187	546
Total	721	279	1000

b.

	TV Violence Inc.	TV Violence No Inc.	Total
TV Quality Worse	79.7%	20.3%	100%
TV Quality Not Worse	65.8%	34.2%	100%

c.

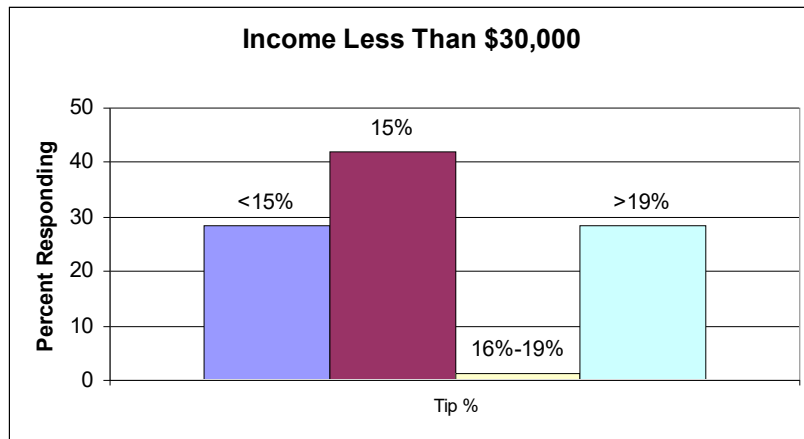|	TV Violence Inc.	TV Violence No Inc.
TV Quality Worse	50.2%	33.0%
TV Quality Not Worse	49.8%	67.0%
Total	100%	100%

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

e.
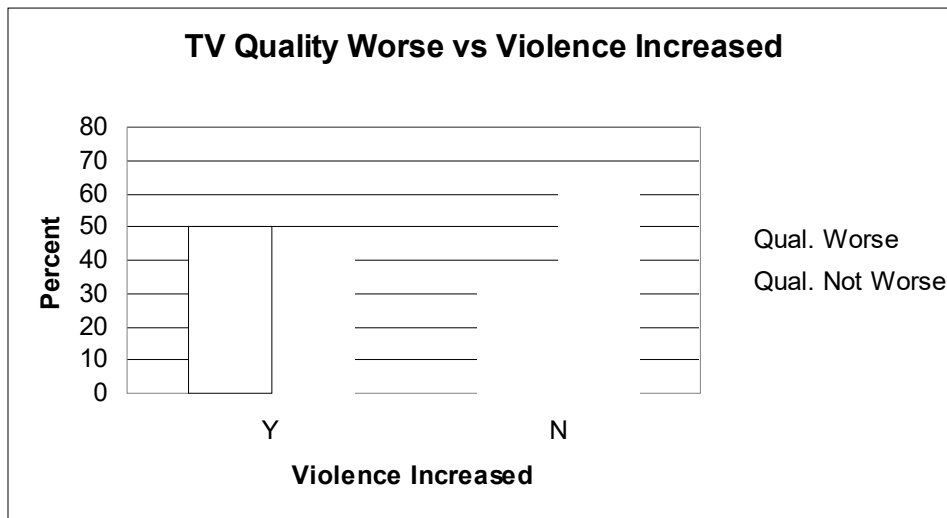
L02-01, L02-06

2.50   a.
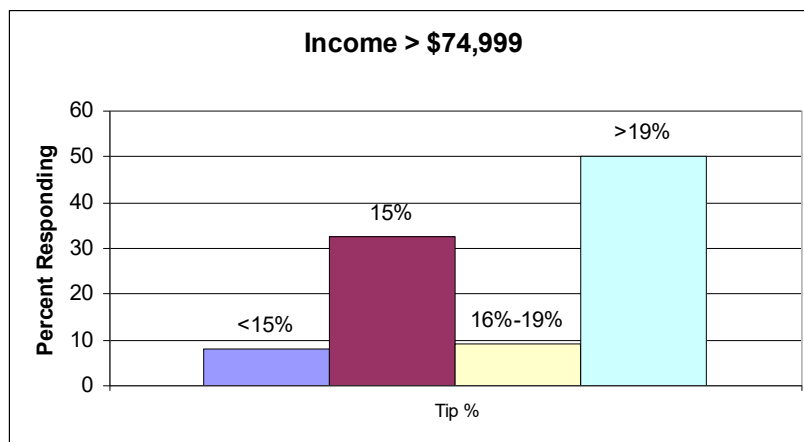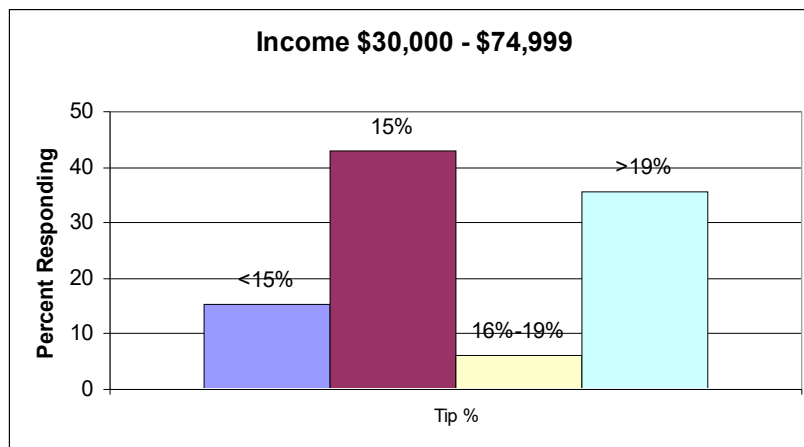
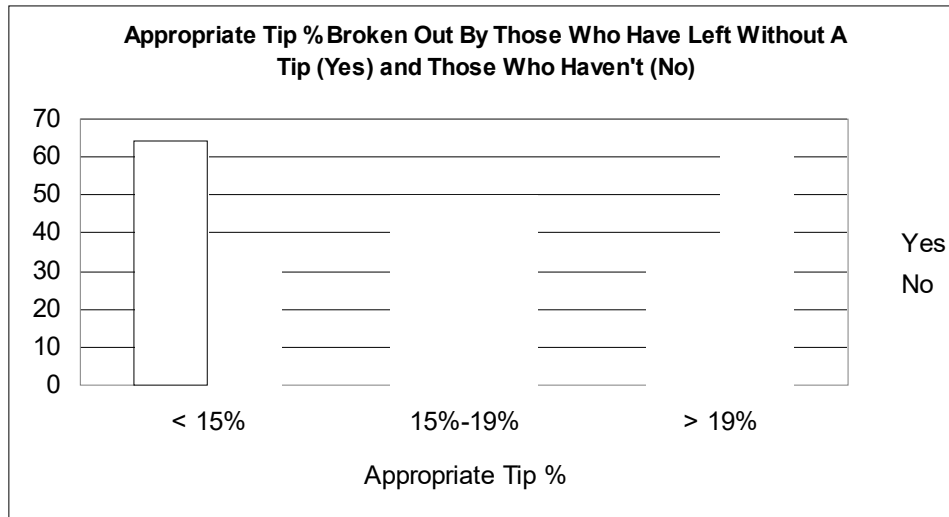**b.** As income rises the percent of people seeing larger tips as appropriate also rises.

L02-01, L02-06

2.51 a.

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

L02-01, L02-06

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

L02-07

2.53 Data are scattered around a straight line with positive slope.

L02-07

2.54 Data are scattered around a straight line with negative slope.

L02-07

2.55 Data are scattered on the plot with the best line to draw through the data being horizontal.

L02-07

2.56   Scatter plot: each value of *y* is plotted against its corresponding value of *x*.
Runs plot: a graph of individual process measurements versus time

L02-07

2.57   As home size increases, sales price increases in a linear fashion. A fairly strong relationship
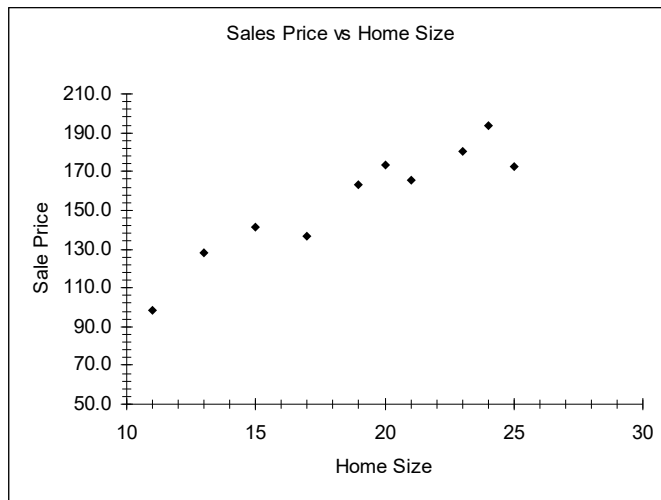
L02-07

2.58   As temperature increases, fuel consumption decreases in a linear fashion. A strong relationship.

L02-07

2.59   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.

L02-07

2.60   Clearly there is a positive linear relationship here. As a brand gets more sales, retailers want to give more shelf space. Also as shelf space increases sales will tend to increase. Its difficult to determine cause and effect here.

L02-07

2.61   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.

The scatterplots below show that average convenience, familiarity, and price are all related in a linear fashion to average preference in a positive, positive, and negative fashion (respectively). These relationships are not as strong as the one between taste and preference.
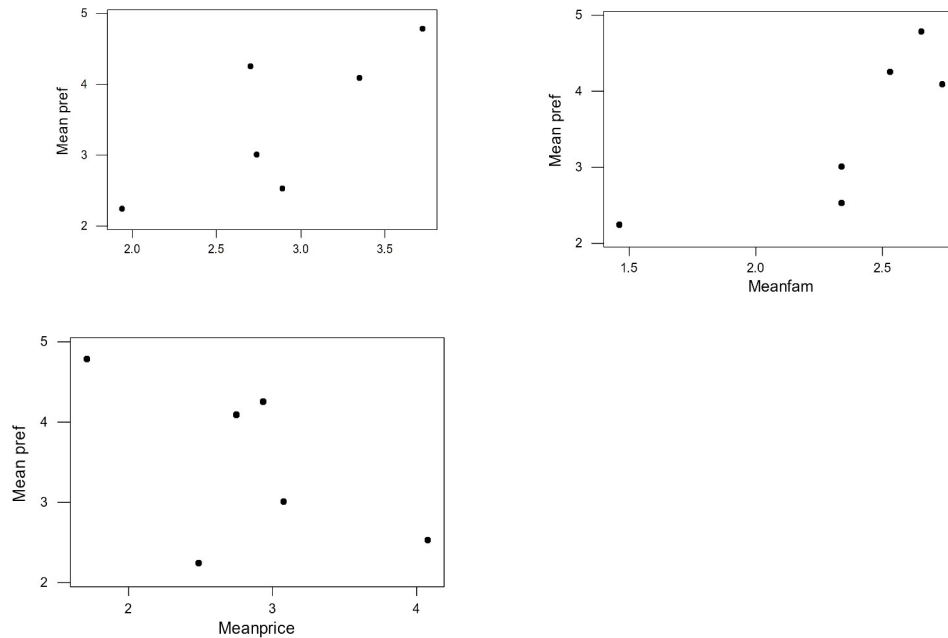
L02-07

2.62   The differences in the heights of the bars are more pronounced.

L02-08

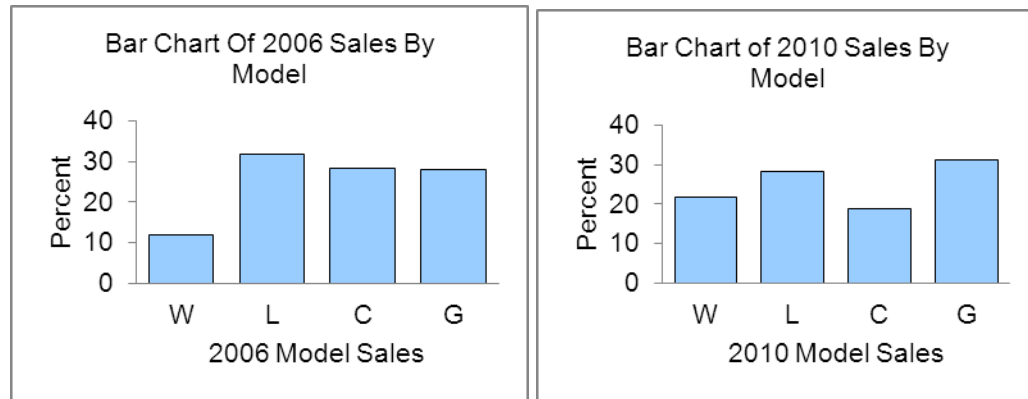**2.63**   Examples and reports will vary.

L02-08

2.64   The administration's plot indicates a steep increase over the four years while the union organizer's plot shows a gradual increase.

L02-08

2.65   **a.**   No, very slight (if any).

   b.   Yes, strong trend.

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

   d.   Probably not. Both distort the data.

L02-08

2.66

Reports will vary but should focus on the Liberty model sales staying around 30% of total sales.

L02-01

2.67   Large portion of manufacturers are rated 3.
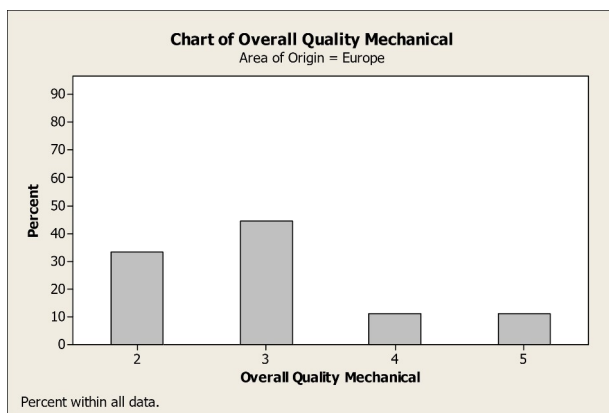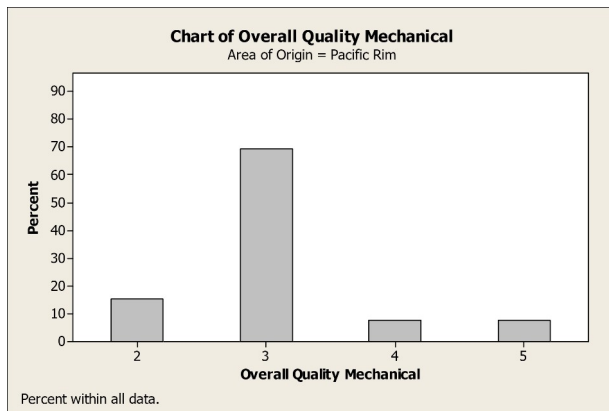
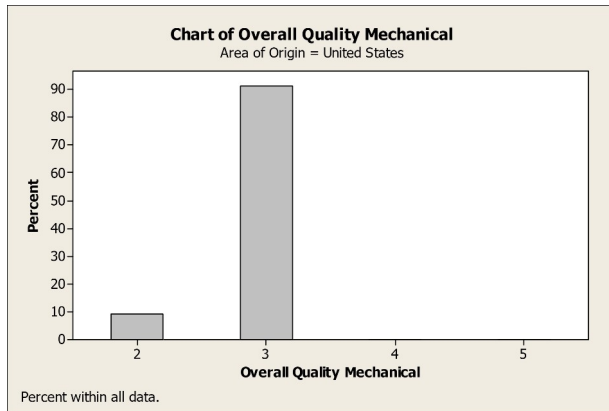| Overall Mechanical Quality	frequency
1	0
2	6
3	23
4	2
5	2
	33

L02-01

2.68   Categories 3 & 4 cover large portion of companies.

Overall Design Quality	frequency	percent
1	0	0.0
2	4	12.1
3	22	66.7
4	6	18.2
5	1	3.0
	33	100.0

L02-01

2.69   Written analysis will vary.

Chart of Overall Quality Mechanical
Area of Origin = United States

Percent within all data.

Chart of Overall Quality Mechanical
Area of Origin = Pacific Rim

Percent within all data.

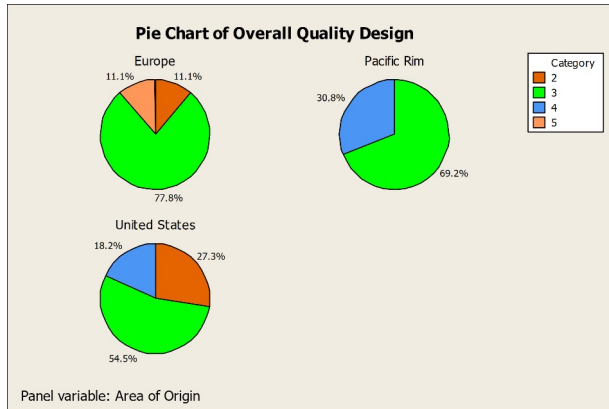Chart of Overall Quality Mechanical
Area of Origin = Europe

Percent within all data.

L02-01

2.70 Written analysis will vary

L02-01

2.71 Europe and the Pacific Rim both have a couple of outliers with ratings of 4 & 5, otherwise there does not seem to be much of a relationship.

Tabulated statistics: Area of Origin, Overall Quality Mechanical

Rows: Area of Origin   Columns: Overall Quality Mechanical

	The Rest	About Average	Better Than Most	Among The Best	All
Europe	3	4	1	1	9
	33.33	44.44	11.11	11.11	100.00
Pacific Rim	2	9	1	1	13
	15.38	69.23	7.69	7.69	100.00
United States	1	10	0	0	11
	9.09	90.91	0.00	0.00	100.00
All	6	23	2	2	33
	18.18	69.70	6.06	6.06	100.00

Cell Contents: Count
                   % of Row

L02-06

2.72 Written reports will vary. See 2.69 for percentage bar charts. See 2.71 for row percentages.

L02-06

2.73 Pacific Rim has a much higher percentage rated 4 or higher than either Europe or United States.

Tabulated statistics: Area of Origin, Overall Quality Design

```
Rows: Area of Origin   Columns: Overall Quality Design

                    2           3           4           5          All

Europe              1           7           0           1            9
                11.11       77.78        0.00       11.11       100.00

Pacific Rim         0           9           4           0           13
                 0.00       69.23       30.77        0.00       100.00

United States       3           6           2           0           11
                27.27       54.55       18.18        0.00       100.00

All                 4          22           6           1           33
                12.12       66.67       18.18        3.03       100.00

Cell Contents:      Count
                    % of Row
```

L02-06

**2.74** Written reports will vary. See 2.70 for pie charts. See 2.73 for row percentages

L02-06

2.75 a. Since there are 50 data points you should use 6 classes.

b.

Frequency Distribution -
Quantitative

ModelAge lower		upper	midpoint	width	frequency	percent	cumulative frequency	percent
17	<	19	18	2	3	6.0	3	6.0
19	<	21	20	2	2	4.0	5	10.0
21	<	23	22	2	3	6.0	8	16.0
23	<	25	24	2	5	10.0	13	26.0
25	<	27	26	2	8	16.0	21	42.0
27	<	29	28	2	15	30.0	36	72.0
29	<	31	30	2	10	20.0	46	92.0
31	<	33	32	2	4	8.0	50	100.0

50   100.0

Histogram
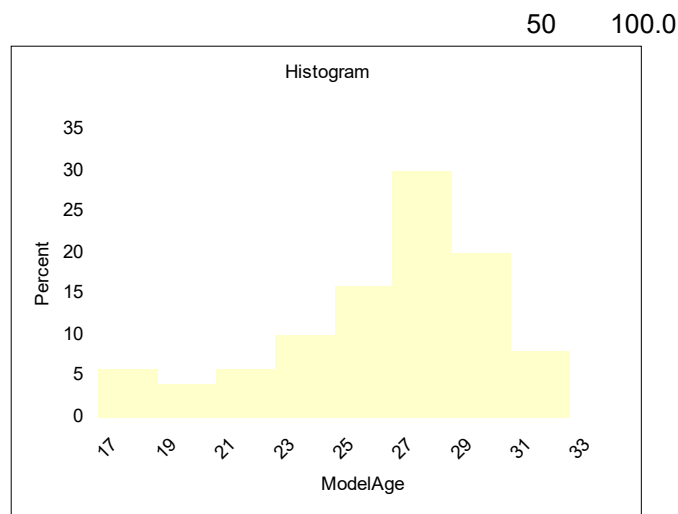
c.

d.   This distribution is skewed to the left.

L02-03

2.76

Frequency Polygon
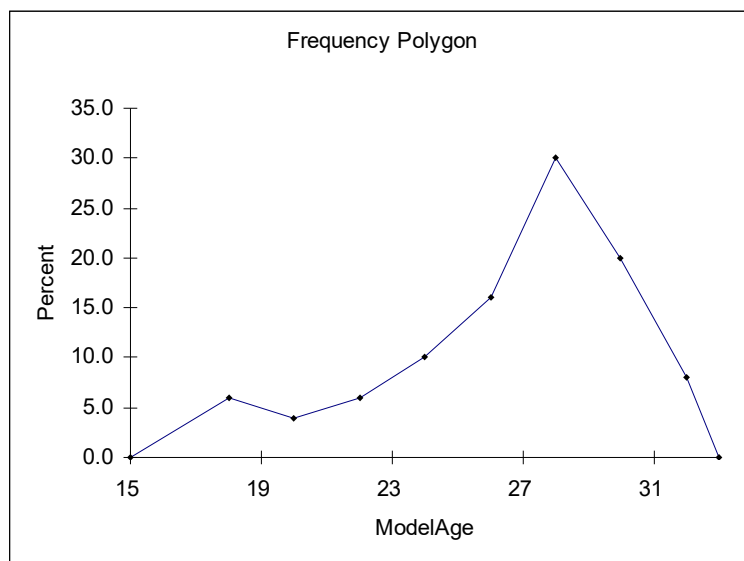
L02-03

2.77   26% of the perceived ages are below 25. Much too high.

DotPlot

ModelAge
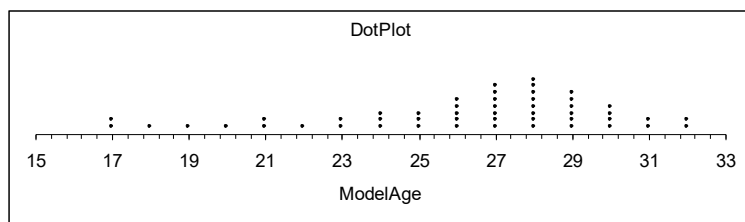
L02-04

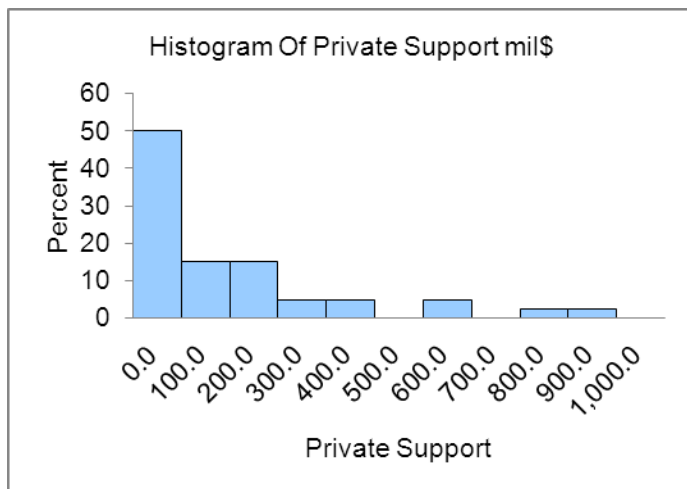**2.78a & b & c.** See table in 2.75

   d.

Ogive
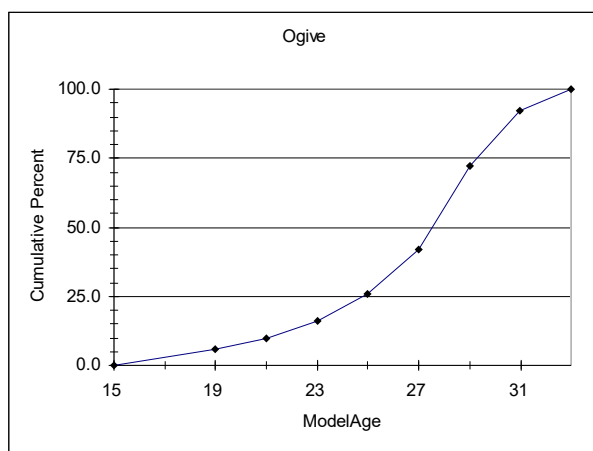
   e.   36 out of 50 = 72%

   f.   8 out of 50 = 16%

L02-03

2.79
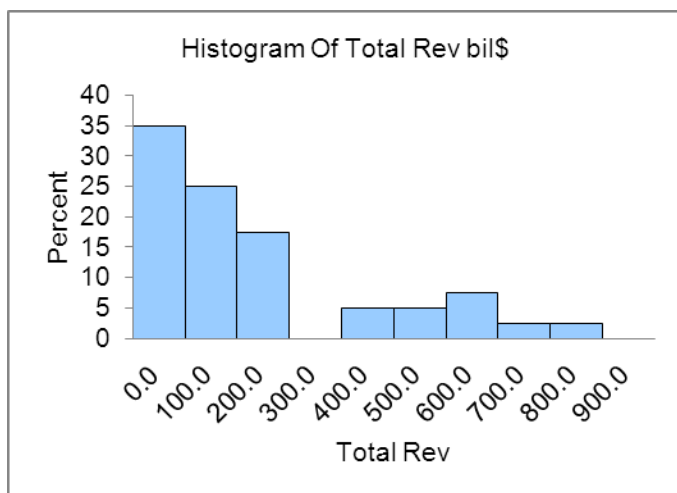
Distribution is skewed to the right

Distribution is skewed to the right

Histogram of Efficiency
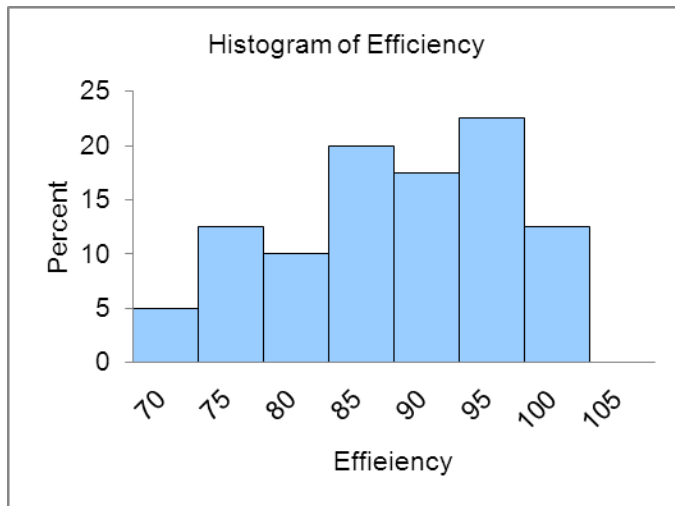
Distribution is skewed to the left
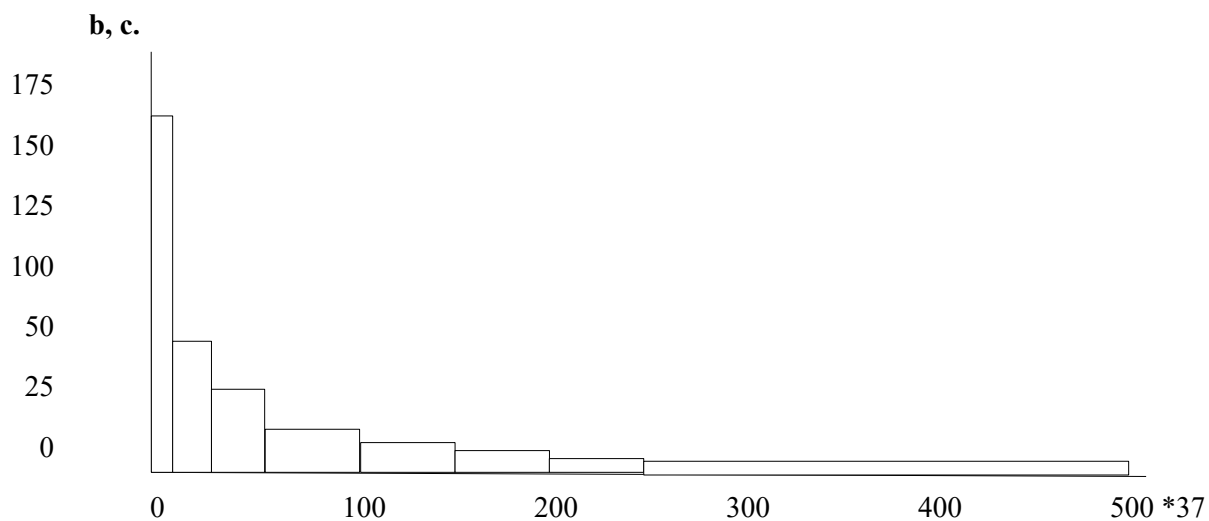
L02-03

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

L02-04

2.81 a.

| Class	Factor	Height
$25K to 50K	$\dfrac{50-25}{10-0} = \dfrac{25}{10} = 2.5$	$\left(\dfrac{1}{2.5}\right)(62) = 24\dfrac{4}{5}$
$50K to 100K	$\dfrac{100-50}{10-0} = \dfrac{50}{10} = 5$	$\left(\dfrac{1}{5}\right)(60) = 12$
$100K to 150K	$\dfrac{150-100}{10-0} = \dfrac{50}{10} = 5$	$\left(\dfrac{1}{5}\right)(24) = 4\dfrac{4}{5}$
$150K to 200K	$\dfrac{200-150}{10-0} = \dfrac{50}{10} = 5$	$\left(\dfrac{1}{5}\right)(19) = 3\dfrac{4}{5}$
$200K to 250K	$\dfrac{250-200}{10-0} = \dfrac{50}{10} = 5$	$\left(\dfrac{1}{5}\right)(22) = 4\dfrac{2}{5}$
$250K to 500K	$\dfrac{500-250}{10-0} = \dfrac{250}{10} = 25$	$\left(\dfrac{1}{25}\right)(21) = \dfrac{21}{25}$

b, c.

L02-03

2.82   Since the runs plot is not in control, the stem & leaf is not representative of the number of missed shots.

```
            Stem-and-leaf of Shots Missed     N = 30
            Leaf Unit = 0.10

                1      5 0
                2      6 0
                4      7 00
                9      8 00000
               15      9 000000
               15     10 00000
               10     11 00
                8     12 0
                7     13 0
                6     14 0
                5     15 00
                3     16 0
                2     17 0
                1     18 0
```
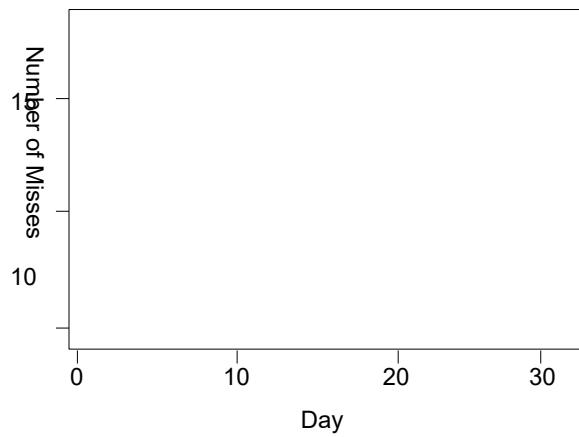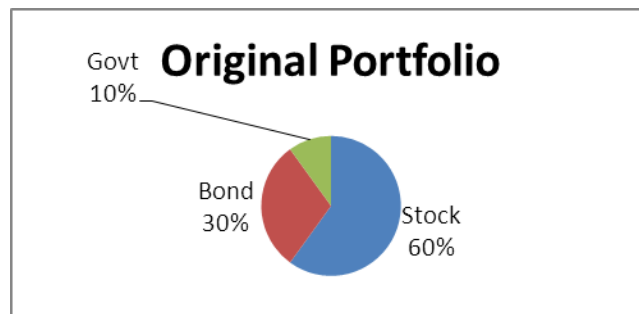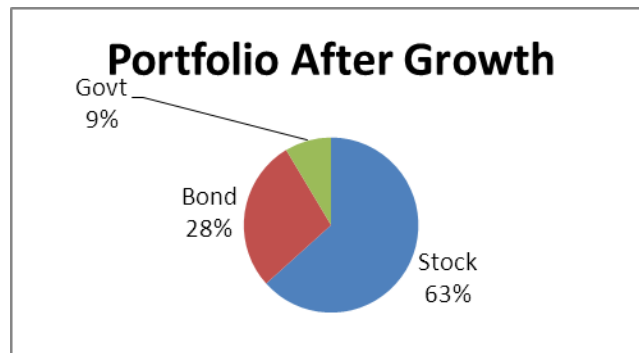
L02-05

2.83 The graph indicates that Chevy trucks far exceed Ford and Dodge in terms of resale value, but the y-axis scale is misleading.
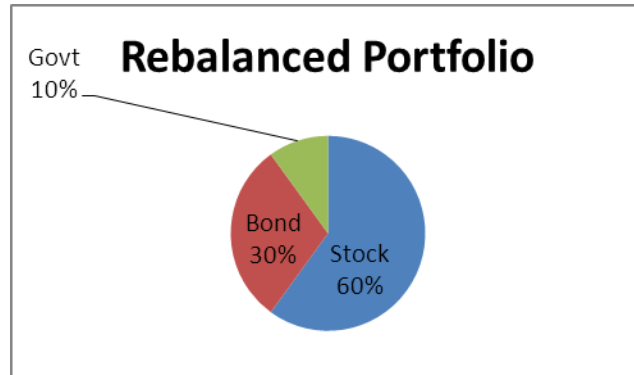
L02-08

2.84 **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

L02-01

Internet Exercises

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

L02-01 – L02-08