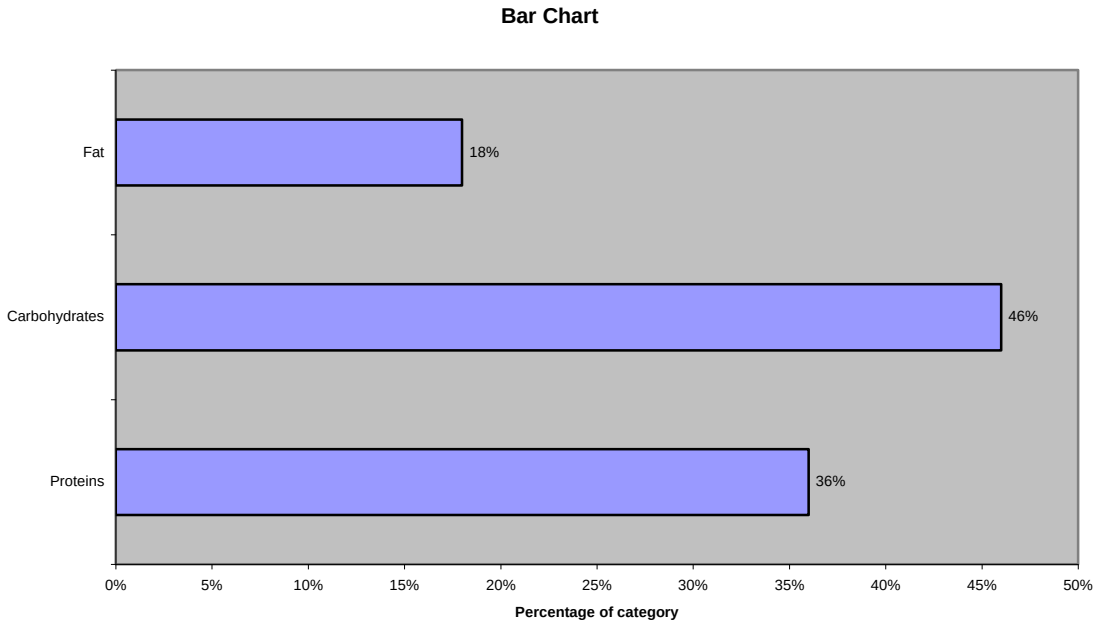
CHAPTER 2

2.1 (a)

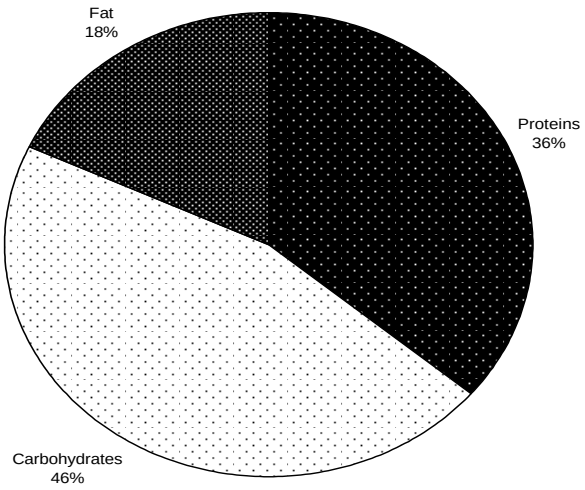
Category	Frequency	Percentage
Proteins	18	36
Carbohydrates	23	46
Fat	9	18

(b) The Bar Chart



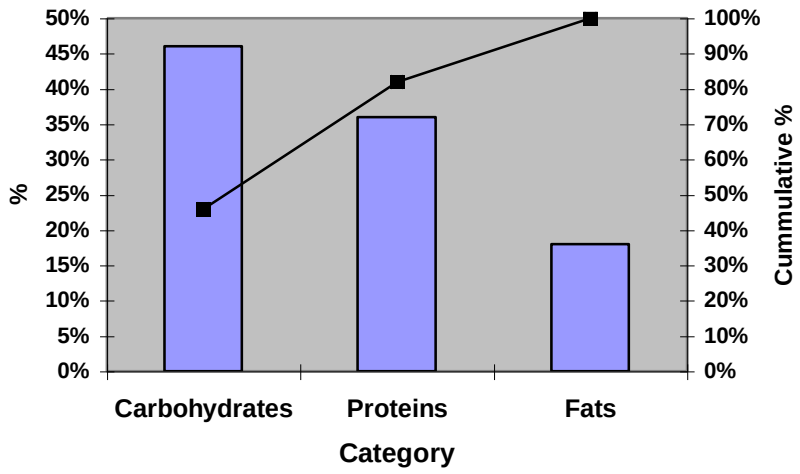
(c)

The Pie Chart

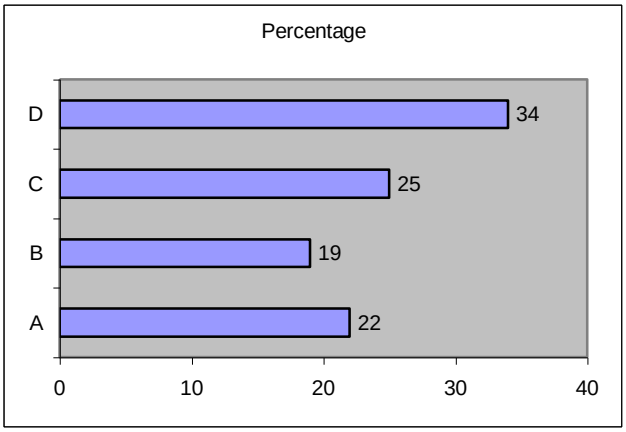


2.1 (d) cont.

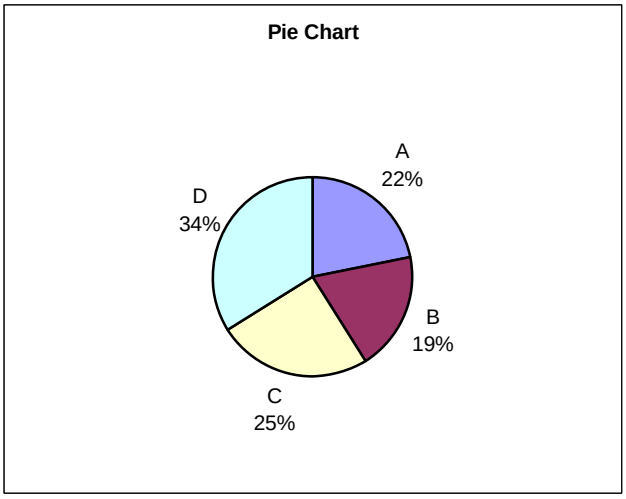
The Pareto Chart



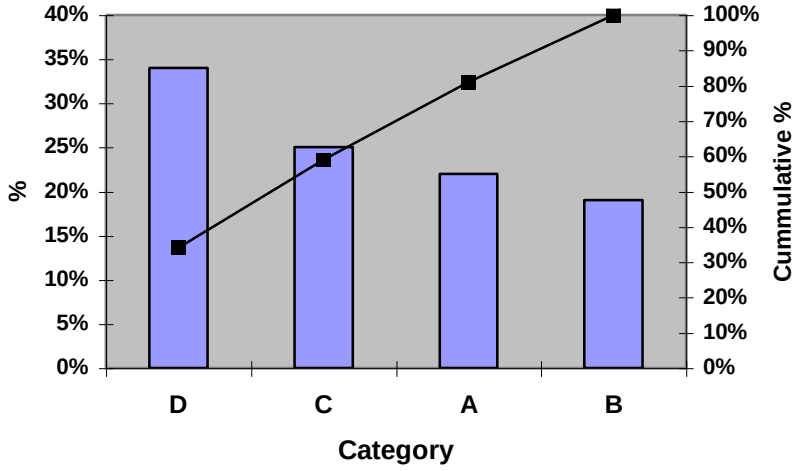
2.2 (a)



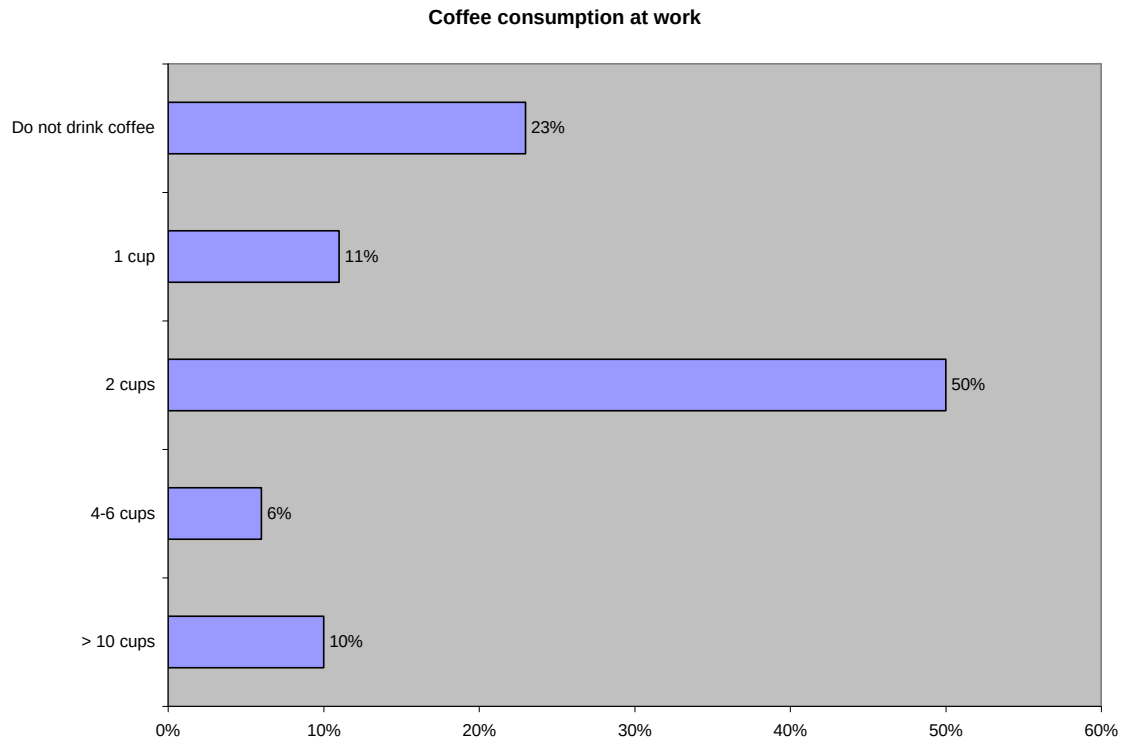
2.2 (b)
cont.



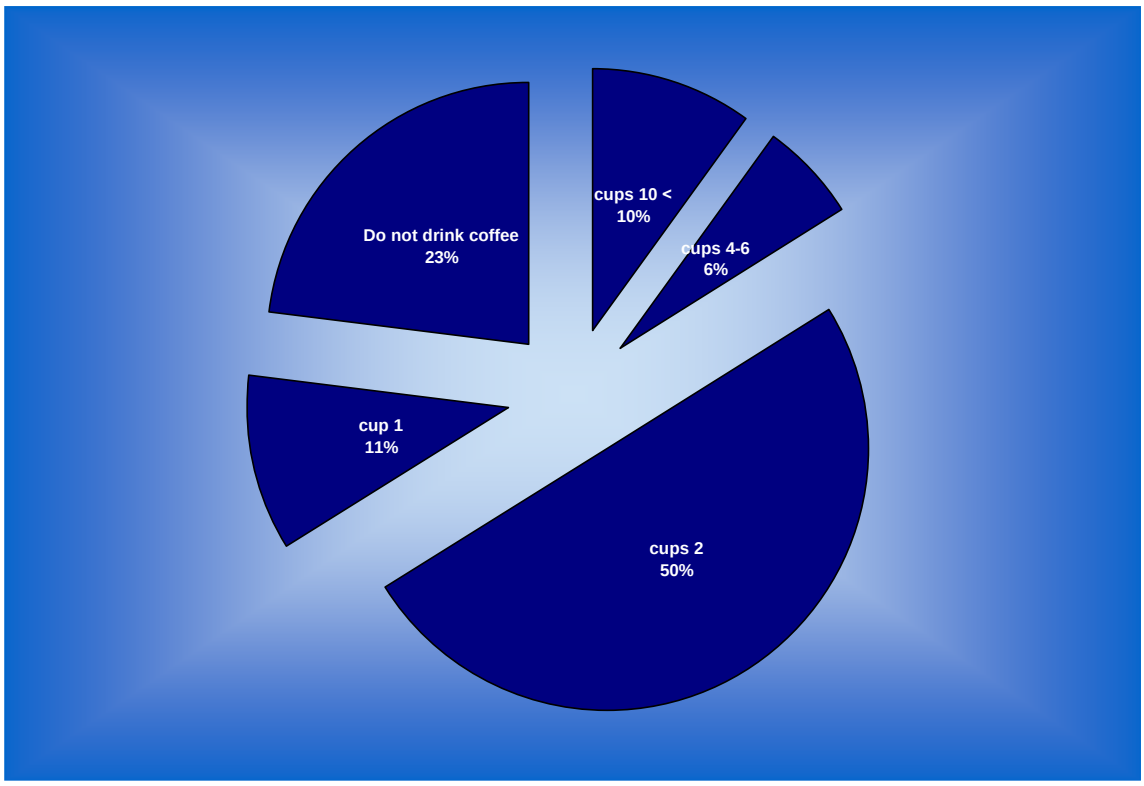
(c)



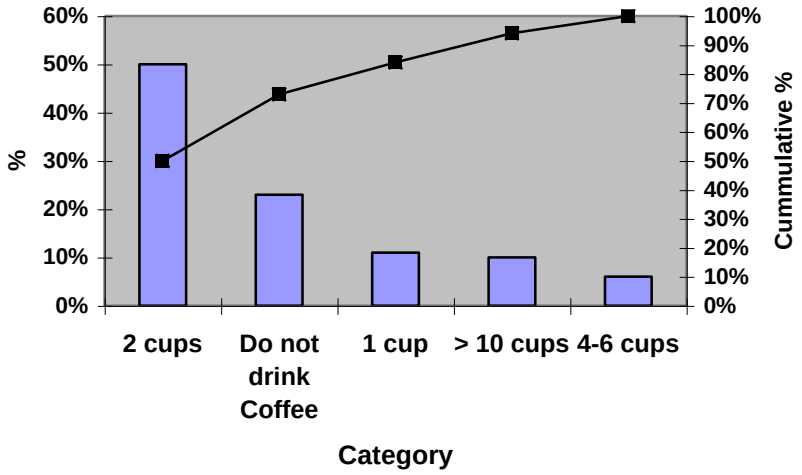
2.3 (a) Bar Chart



Pie Chart



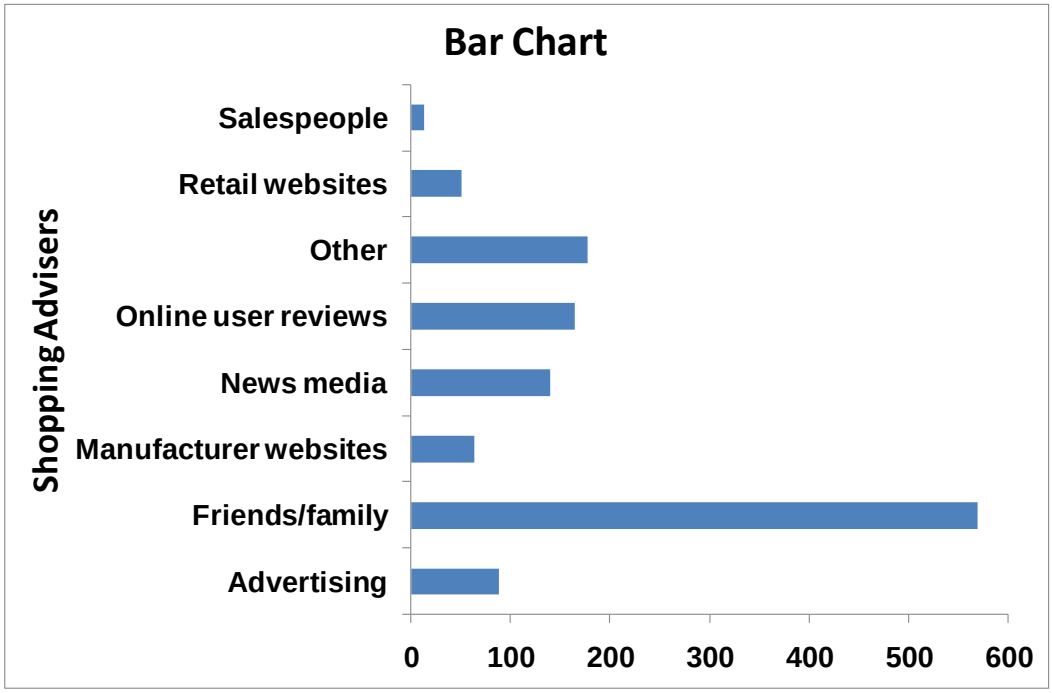
2.3 (a)
cont.
Pareto Diagram



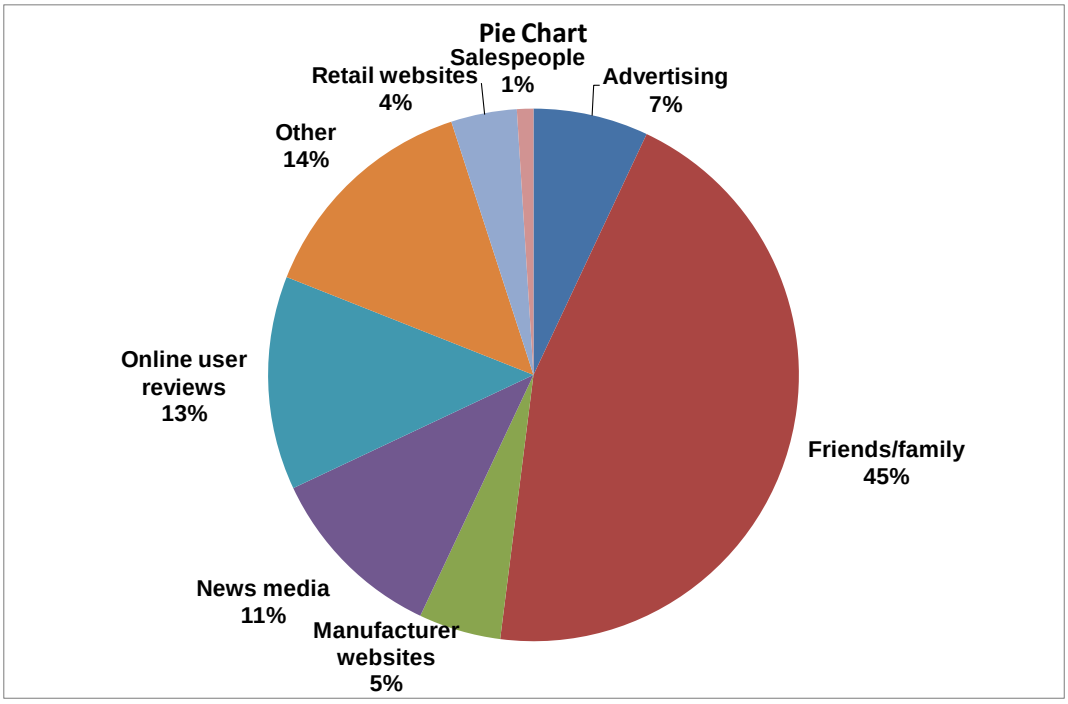
(b) The Pareto diagram is the best to portray these data because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale. From the Pareto diagram, it is obvious more than 70% of the people either consume "2 cups" or "do not drink coffee" at all.
* Note: This is one of the many possible solutions for the question.

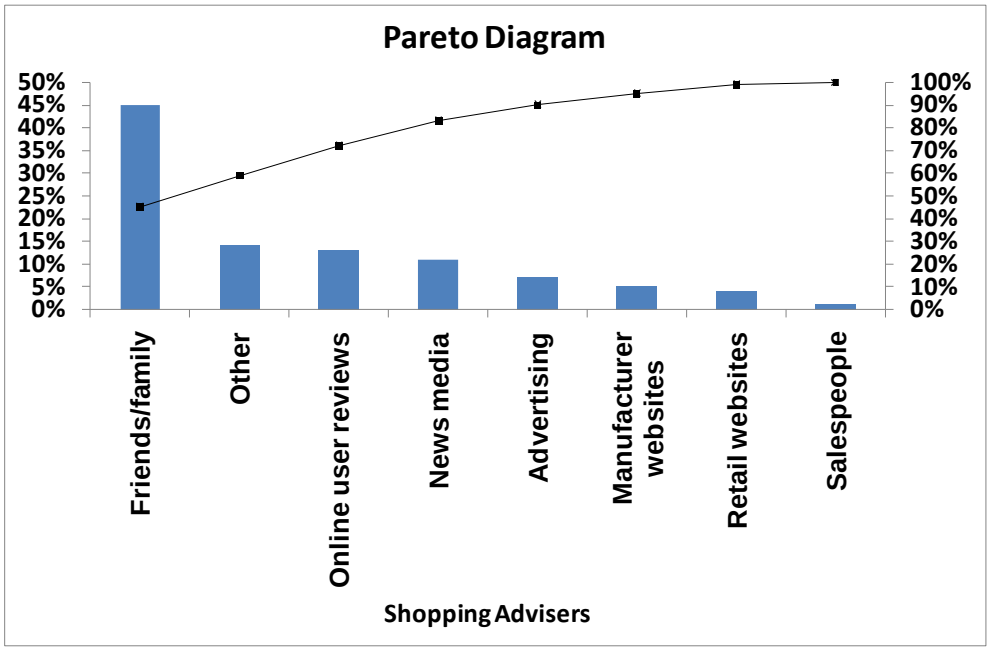
(c) We can conclude that more than 70% of the workers don't have excessive coffee drinking habit at work. A quarter of the people do not even drink coffee and half the population drinks only 2 cups of coffee a day.

2.4 (a)



2.4 cont. (a)

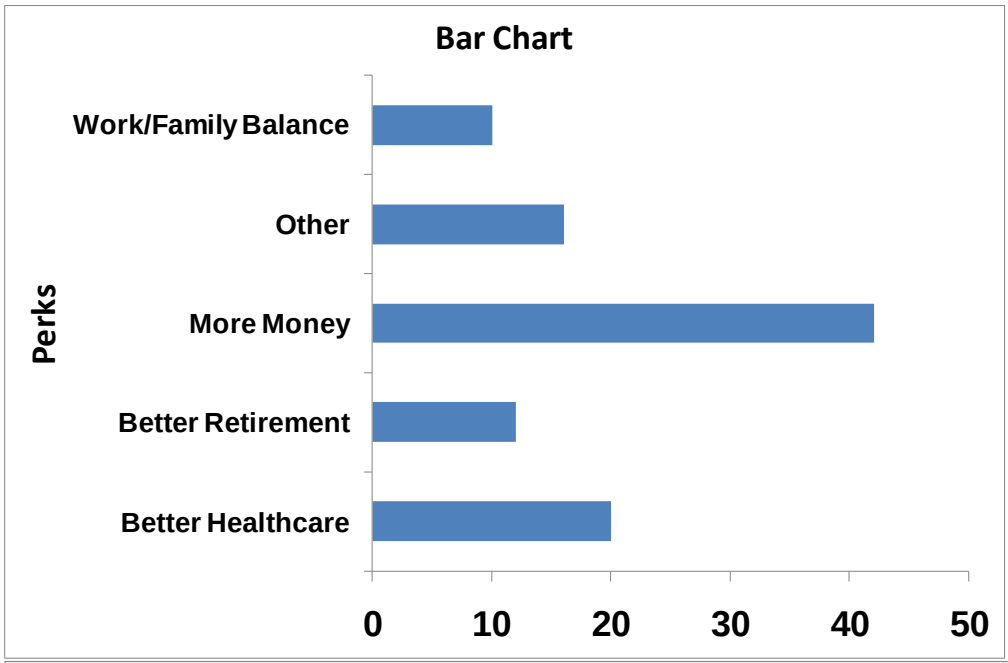


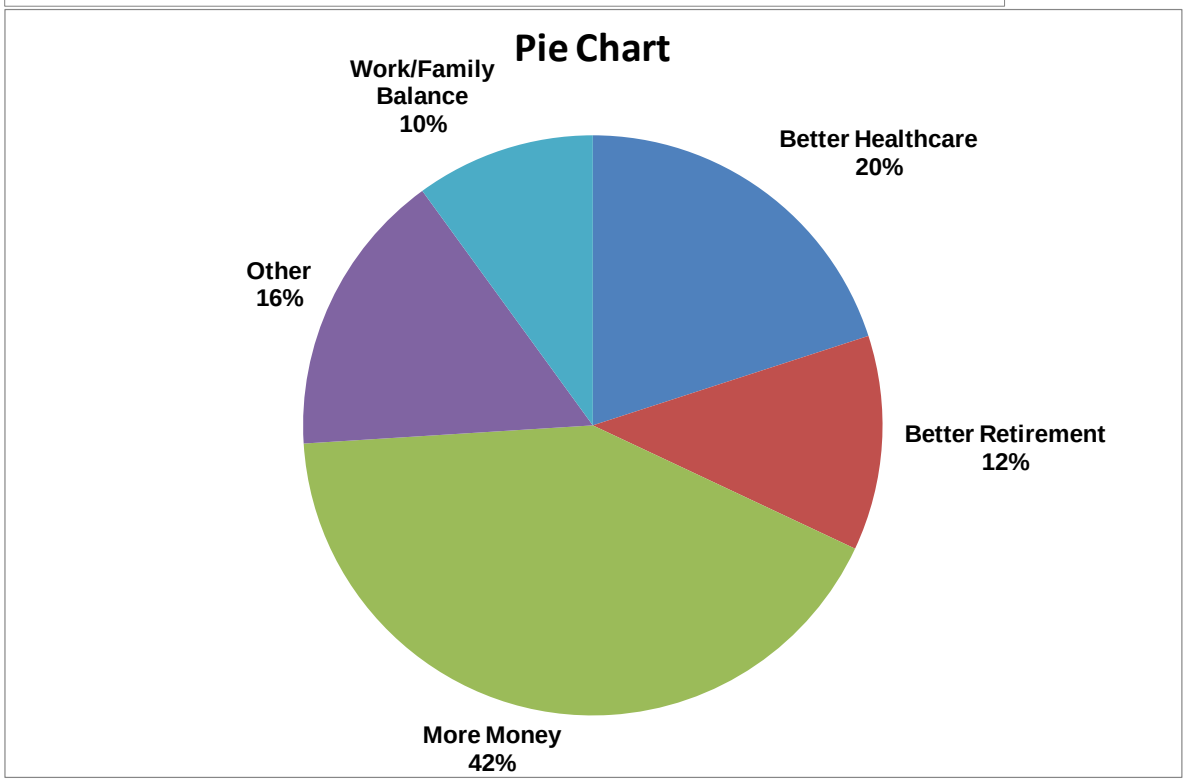


(b) The Pareto diagram is better than the pie chart to portray these data because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

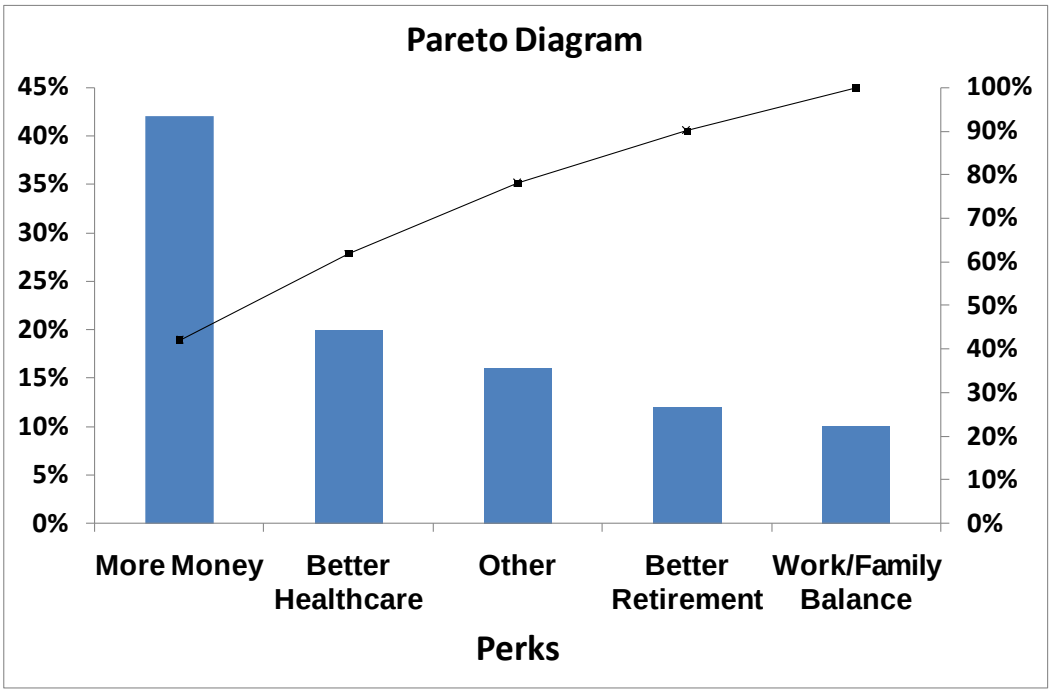
(c) From the Pareto diagram, it is obvious that "Friends/Family" is the women's most trusted source of shopping advisers at 45%.

2.5 (a)





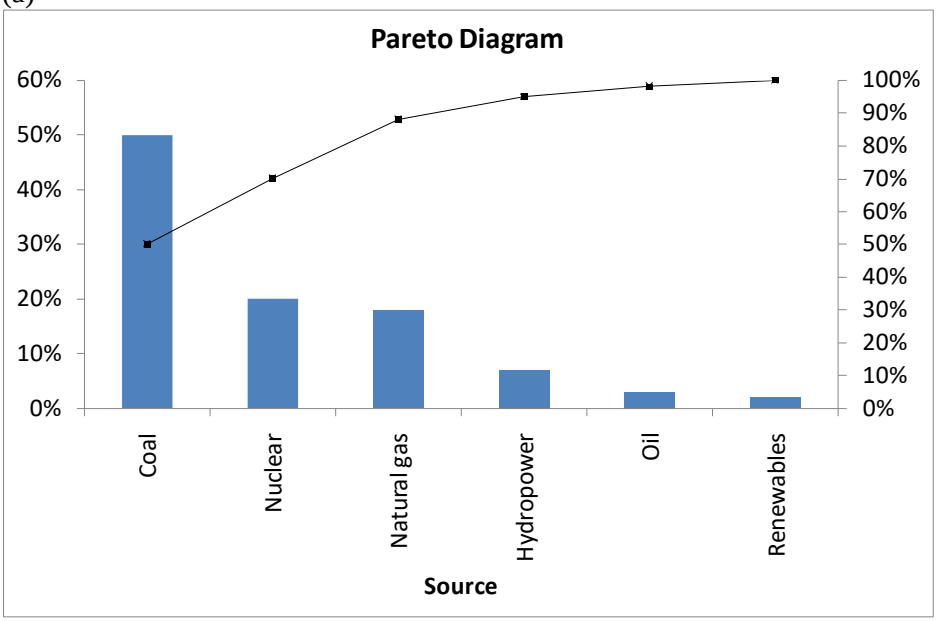
2.5 cont. (a)



(b) The Pareto diagram is better than the pie chart or the bar chart because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

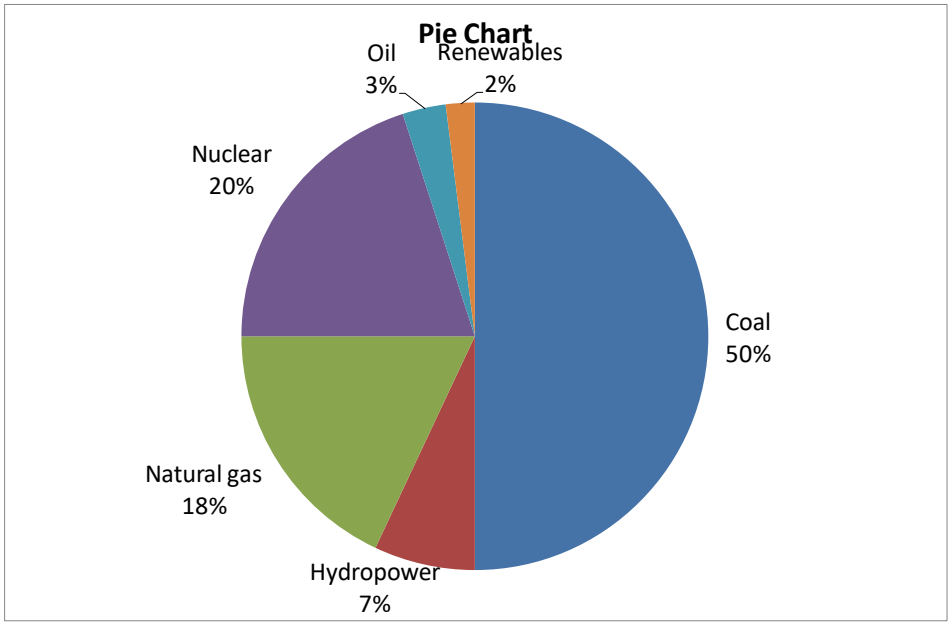
(c) From the Pareto diagram, it is obvious that more than 40% of the workers want "more money". So "more money" is the perk workers want most.

2.6 (a)



(b) Approximately 88% of the electricity is derived from coal, nuclear energy or natural gas.

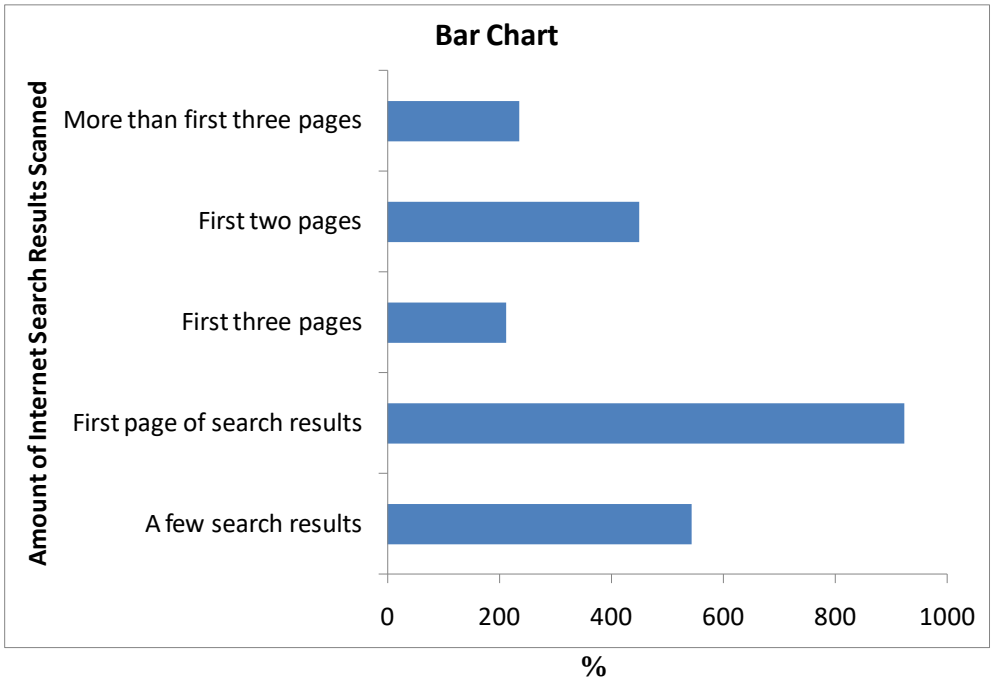
2.6 cont. (c)



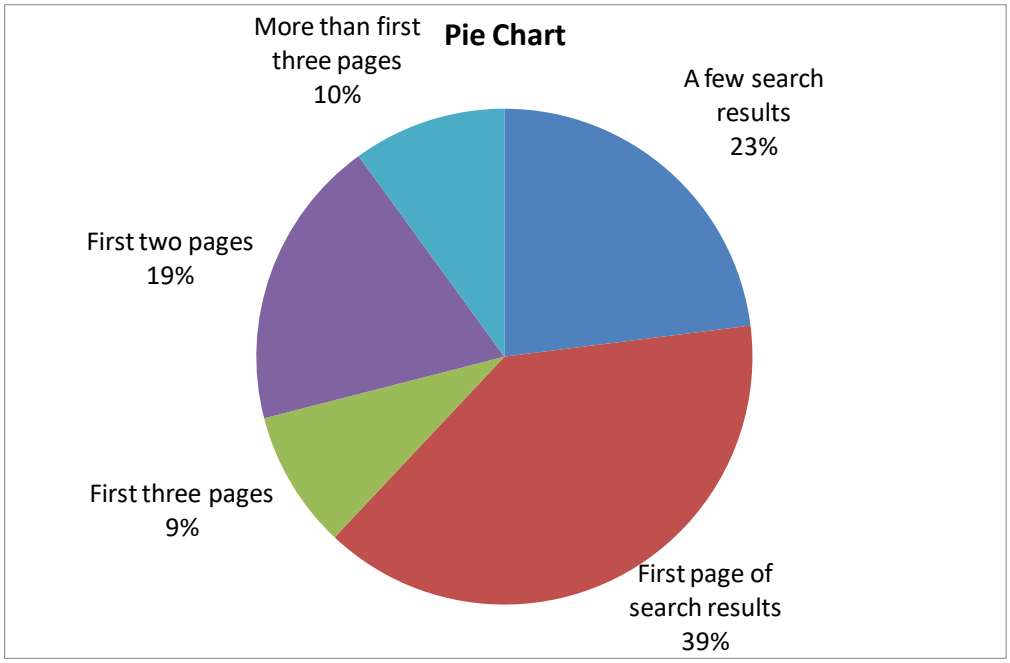
(d) The Pareto diagram is better than the pie chart because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale. From the Pareto diagram, it is obvious that almost 90% of the electricity is derived from coal, nuclear energy or natural gas. *

* Note: This is one of the many possible solutions for the question.

2.7 (a)



2.7 cont. (a)

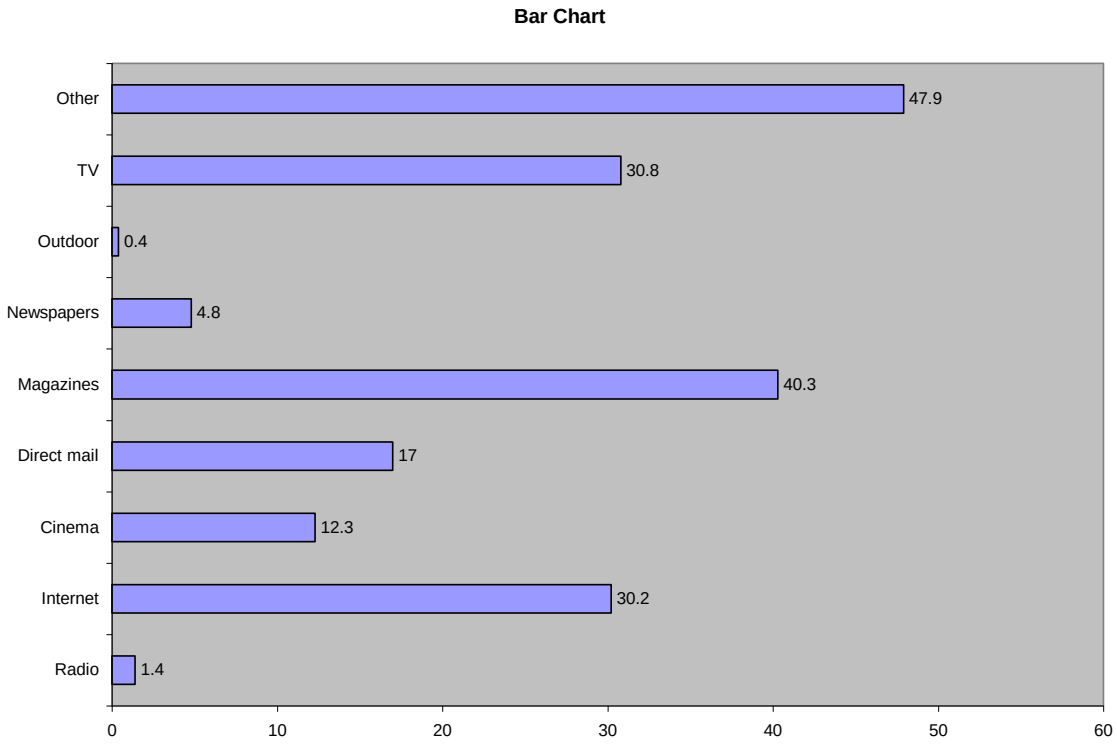


2.7 (b) The bar chart is more suitable if the purpose is to compare the categories. The pie chart is more suitable if the main objective is to investigate the portion of the whole that is in a particular category. *

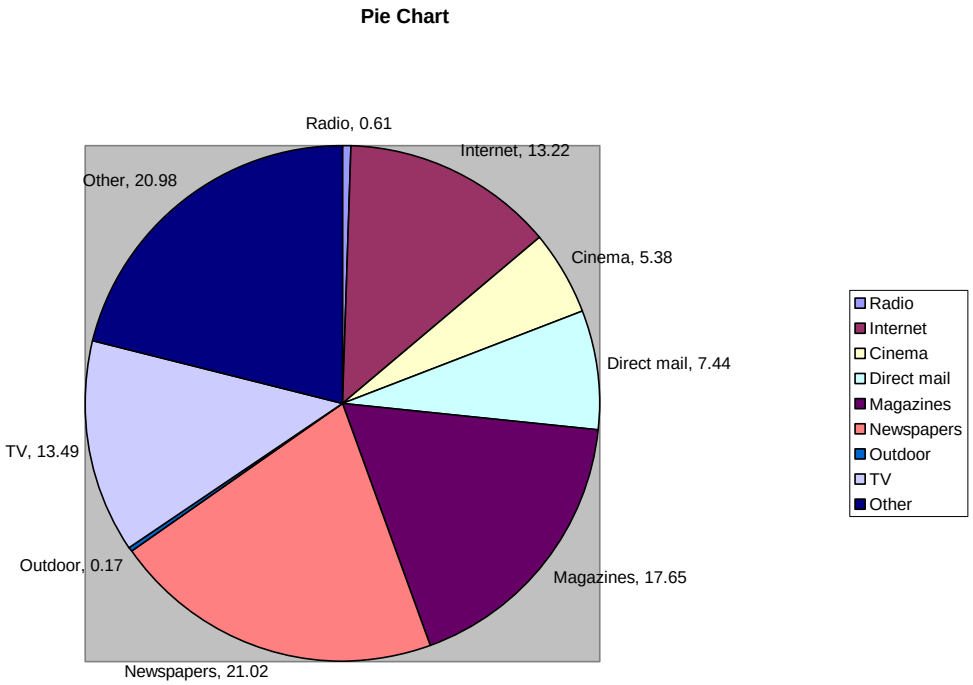
* Note: This is one of the many possible solutions for the question.

(c) You can conclude that most of the people (39%) scan Internet search results according to the "first page of search results", followed by "a few search results" (23%) and "first two pages" (19%).

2.8 (a)



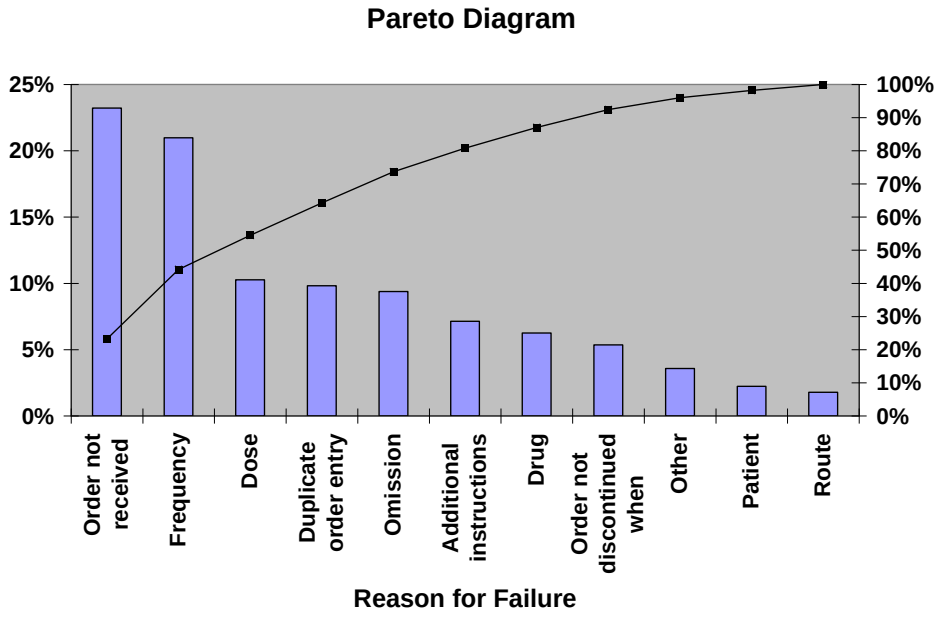
2.8 cont. (a)



(b) The bar chart is more suitable if the purpose is to compare the categories. The pie chart is more suitable if the main objective is to investigate the portion of the whole that is in a particular category. *

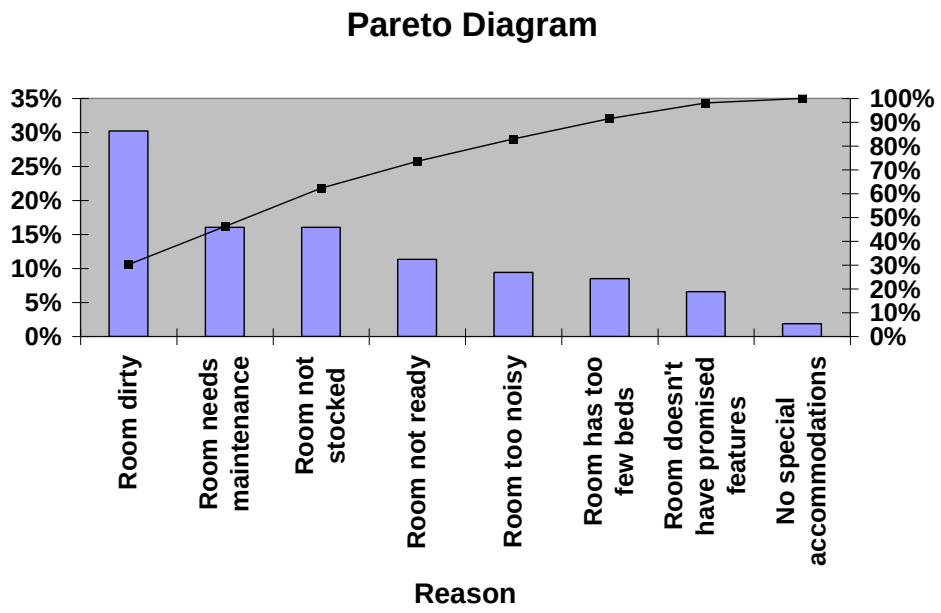
* Note: This is one of the many possible solutions for the question.

2.9 (a)



2.9 cont. (b) The "vital few" are "Order not received" and "Frequency" which accounted for about half of the failures. The remaining reasons constitute the "trivial many" which capture about the remaining half of the failures.

2.10 (a)



(b) The most frequent complain is about "room being dirty" followed by "room needs maintenance" and "room not stocked". The remaining complaints are the "trivial many" reasons.

2.11 Ordered array: 62 63 68 71 72 79 83 99

2.12 Stem-and-leaf of Finance Scores

4	4
5	
6	
7	039
8	348
9	

$n = 7$

2.13 Ordered array: 68 75 78 83 88 91 98

2.14 Ordered array: 52 73 74 75 81 88 94

2.15 (a) Ordered array: 9.1 9.5 9.7 10.0 10.1 10.2 10.3 10.8 11.1 11.2 11.2 11.5 11.5 11.5 11.8 11.8 12.2 12.3 12.3 12.4 12.8 12.9 13.0 13.3

(b) The stem-and-leaf display conveys more information than the ordered array. We can more readily determine the arrangement of the data from the stem-and-leaf display than we can from the ordered array. We can also obtain a sense of the distribution of the data from the stem-and-leaf display.

(c) The most likely petrol purchase is between 11 and 11.9 gallons.

(d) Yes, the third row is the most frequently occurring stem in the display and it is located in the center of the distribution.

2.16 (a) Ordered array: Cost($) 120, 130, 132, 134, 134, 139, 141, 146, 148, 149, 154, 157, 158, 159, 163, 170, 170, 175, 180, 183, 186, 191, 192, 194, 202, 207, 209, 209, 219, 288

(b) PHStat output:

Stem-and-Leaf Display

Stem unit: 10

12	**0**
13	**0 2 4 4 9**
14	**1 6 8 9**
15	**4 7 8 9**
16	**3**
17	**0 0 5**
18	**0 3 6**
19	**1 2 4**
20	**2 7 9 9**
21	**9**
22	
23	
24	
25	
26	
27	
28	**8**

(c) The stem-and-leaf display provides more information because it not only orders observations from the smallest to the largest into stems and leaves, it also conveys information on how the values distribute and cluster over the range of the observations in the data set.

(d) The costs of attending a baseball game do not appear to be concentrating around any particular value. In fact, the costs appear to spread quite evenly between $130 and $210 with the exception of an outlier at $288 for Boston.

2.17 (a) Ordered array: 31, 33.75, 35.05, 36.15, 40.25, 43

(b)

Stem-and-Leaf Display for Price
Stem unit: 1

31	0
32	
33	8
34	
35	1
36	2
37	
38	
39	
40	3
41	
42	
43	0

(c) The stem-and-leaf display conveys more information than the ordered array. We can more readily determine the arrangement of the data from the stem-and-leaf display than we can from the ordered array. We can also obtain a sense of the distribution of the data from the stem-and-leaf display.

(d) The price is not concentrated around any particular value.

* Note that Excel rounds leaves up to the first decimal place.

2.18 (a) Ordered array: 4, 5, 7, 8, 16, 19, 19, 20, 20, 23, 24, 25, 29, 29, 30, 30, 30, 30, 40, 56

(b)

Stem-and-Leaf Display for Fat
Stem unit: 10

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

(c) The stem-and-leaf display conveys more information than the ordered array. We can more readily determine the arrangement of the data from the stem-and-leaf display. We can also obtain a sense of the distribution of the data from the stem-and-leaf display.

(d) The total fat amounts are concentrated around 20 to 29.

2.19 (a) Ordered array: 35, 85, 110, 120, 170, 180, 240, 260, 300, 380, 380, 460

(b)

Stem-and-Leaf Display for Life
Stem unit: 100

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

(c) The stem-and-leaf display conveys more information than the ordered array. We can more readily determine the arrangement of the data from the stem-and-leaf display than we can from the ordered array. We can also obtain a sense of the distribution of the data from the stem-and-leaf display.

(d) The battery life clusters around the high 100s and high 300s.

* Note that Excel rounds leaves up to the tens.

2.20 (a) The class boundaries of the 8 classes can be "40 to less than 50", "50 to less than 60", "60 to less than 70", "70 to less than 80", "80 to less than 90", "90 to less than 100", "100 to less than 110" and "110 to less than 120".

(b) The class-interval width is = $\frac{(113.8 - 41.6)}{8}$ = 9.025 = 10.

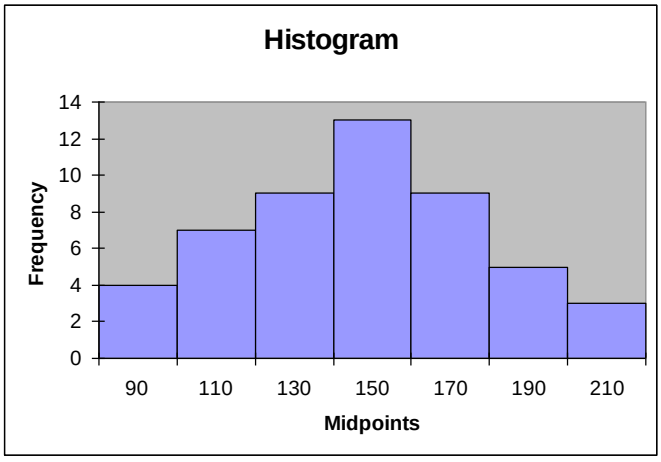
(c) The eight class midpoints are: 45, 55, 65, 75, 85, 95, 105 and 115.

2.21 (a) 4% (b) 32% (c) 36% (d) 100%

2.22 (a)

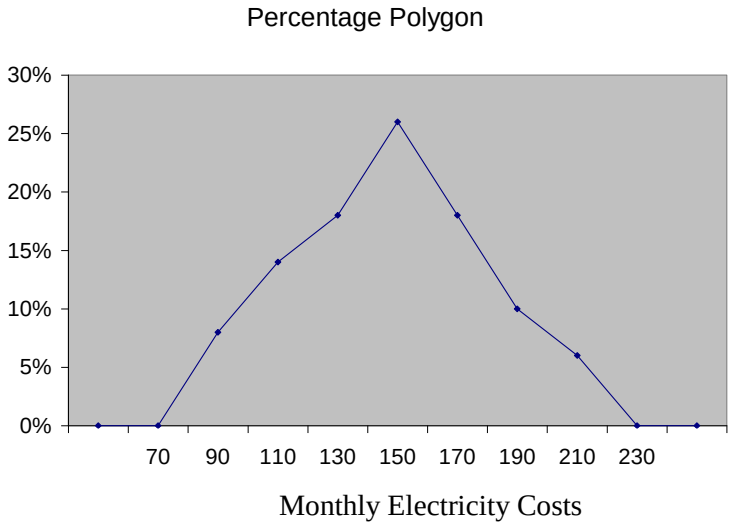
Electricity Costs	Frequency	Percentage
$80 up to $99	4	8%
$100 up to $119	7	14
$120 up to $139	9	18
$140 up to $159	13	26
$160 up to $179	9	18
$180 up to $199	5	10
$200 up to $219	3	6

(b)



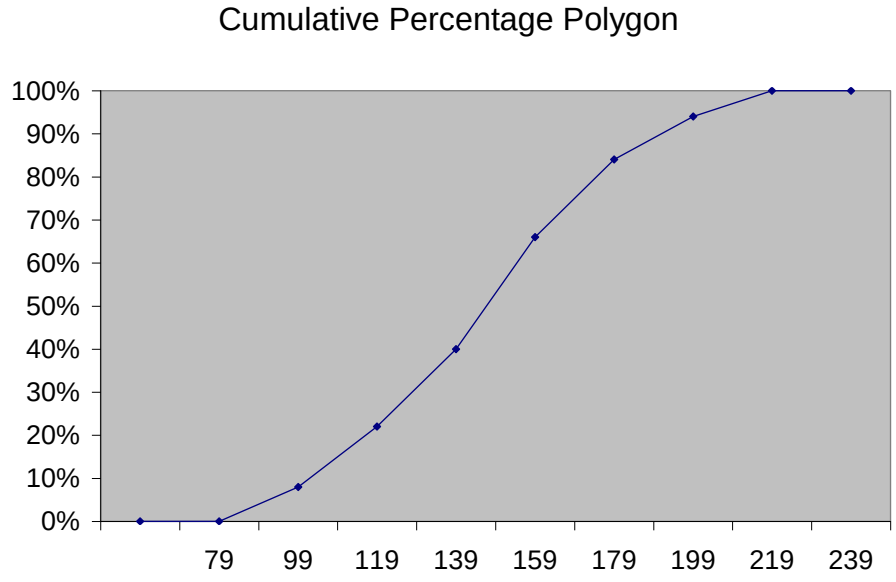
Monthly Electricity Costs

2.22 (b)
cont.



(c)

Electricity Costs	*Frequency*	*Percentage*	*Cumulative %*
$99	4	8%	8%
$119	7	14%	22%
$139	9	18%	40%
$159	13	26%	66%
$179	9	18%	84%
$199	5	10%	94%
$219	3	6%	100%



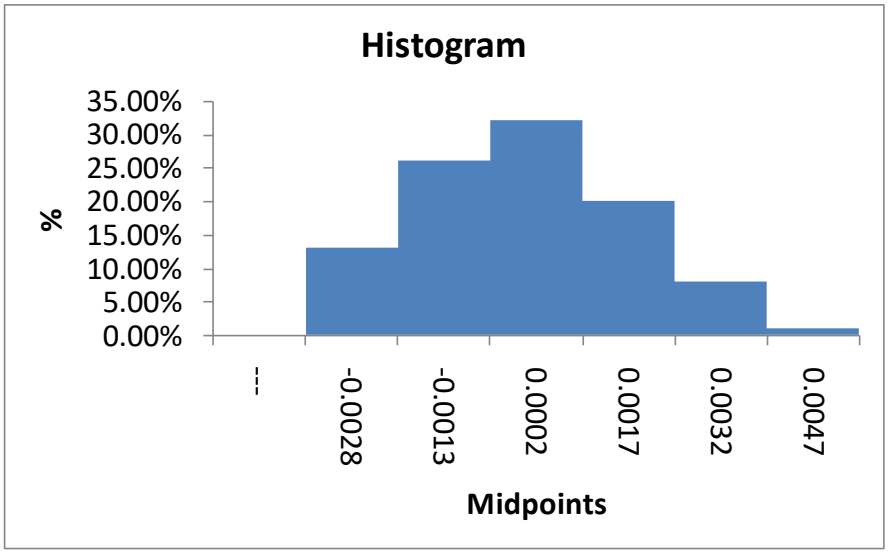
2.22 cont. (d) Monthly electricity costs are most concentrated between $140 and $160 a month, with better than one-fourth of the costs falling in that interval.

2.23 The cost of attending a baseball game is quite evenly distributed between $130 and $210. The cost of attending a game is less than $130 or more than $210 for only a few teams. For about half of the team, it cost less than approximately $170 to attend a baseball game.

2.24 The property tax per capita is rather evenly distributed between $300 to $1,700 with the exception of 16 states with a property tax per capita of somewhere between $900 and $1,100 ,and four states with a property tax per capita of higher than $1,700. Also, half of the states have a property tax per capita of less than $1,000.

2.25 (a)

Error	*Frequency*	*Cumulative %*	*Percentage*
-0.00350 -- -0.00201	13	13%	13%
-0.00200 -- -0.00051	26	39%	26%
-0.00050 -- 0.00099	32	71%	32%
0.00100 -- 0.00249	20	91%	20%
0.00250 -- 0.00399	8	99%	8%
0.00400 -- 0.00549	1	100%	1%

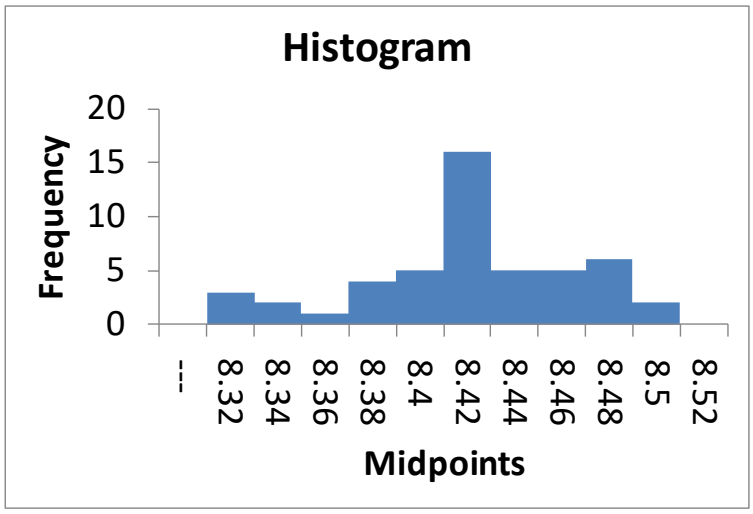


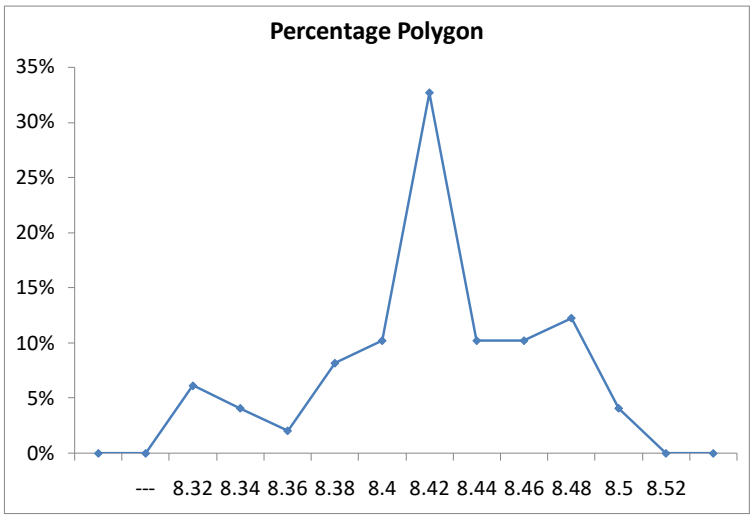
(b) Yes, the steel mill is doing a good job at meeting the requirement as there is only one steel part out of a sample of 100 that is as much as 0.005 inches longer than the specified requirement.

2.26 (a)

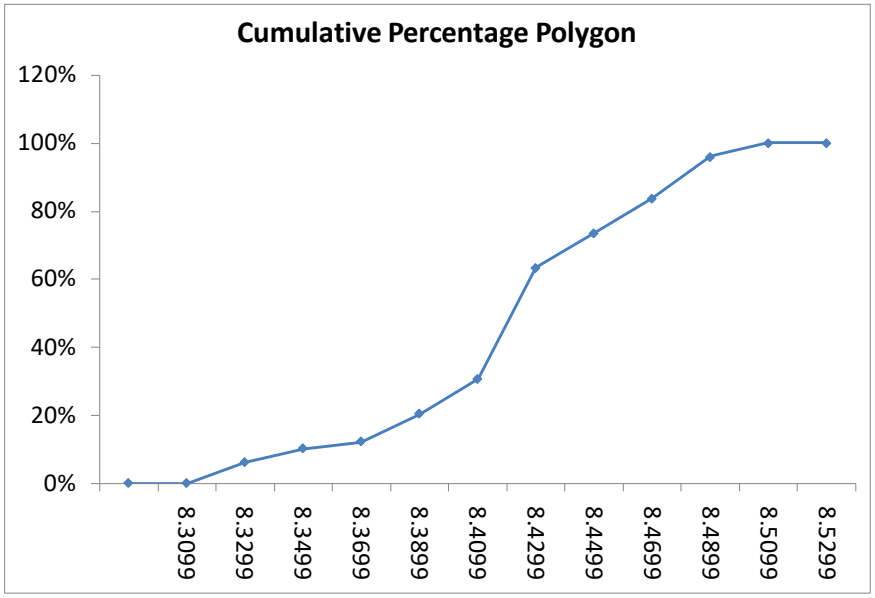
Width	*Frequency*	*Percentage*
8.310 -- 8.329	3	6.12%
8.330 -- 8.349	2	4.08%
8.350 -- 8.369	1	2.04%
8.370 -- 8.389	4	8.16%
8.390 -- 8.409	5	10.20%
8.410 -- 8.429	16	32.65%
8.430 -- 8.449	5	10.20%
8.450 -- 8.469	5	10.20%
8.470 -- 8.489	6	12.24%
8.490 -- 8.509	2	4.08%

(b)





2.26 cont. (c)

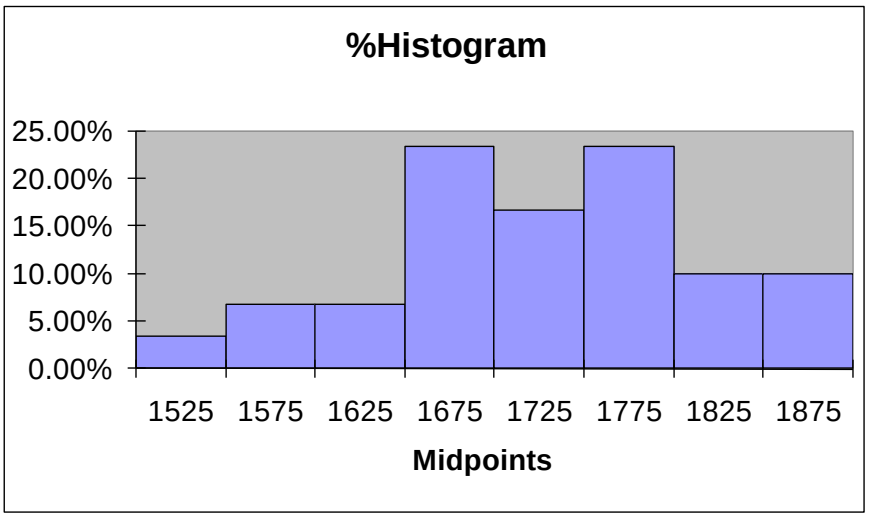


(d) All the troughs will meet the company's requirements of between 8.31 and 8.61 inches wide.

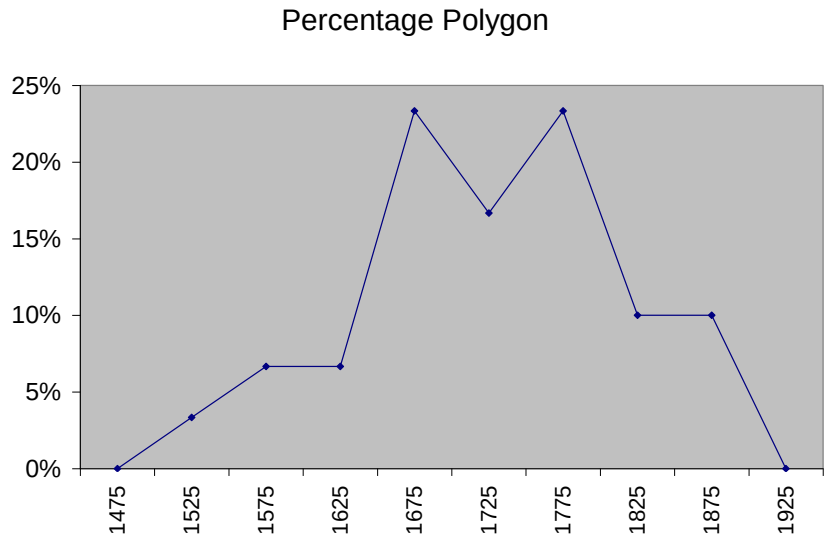
2.27 (a)

Strength	*Frequency*	*Percentage*
1500 -- 1549	1	3.33%
1550 -- 1599	2	6.67%
1600 -- 1649	2	6.67%
1650 -- 1699	7	23.33%
1700 -- 1749	5	16.67%
1750 -- 1799	7	23.33%
1800 -- 1849	3	10.00%
1850 -- 1899	3	10.00%

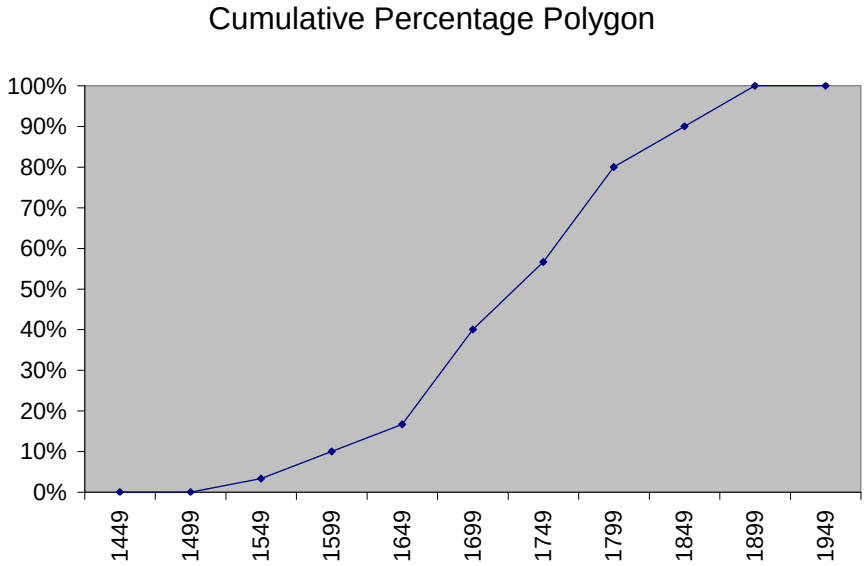
(b)



2.27 cont. (b)



(c)



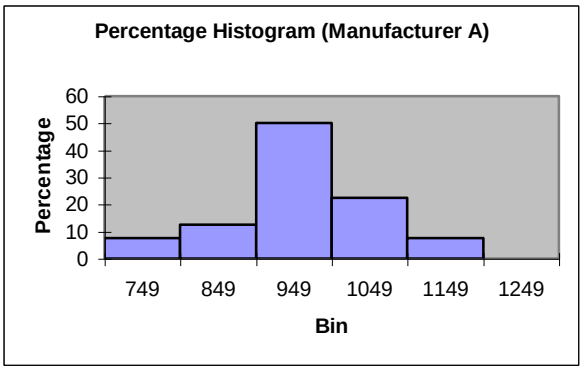
(d) The strength of all the insulators meets the company's requirement of at least 1500.

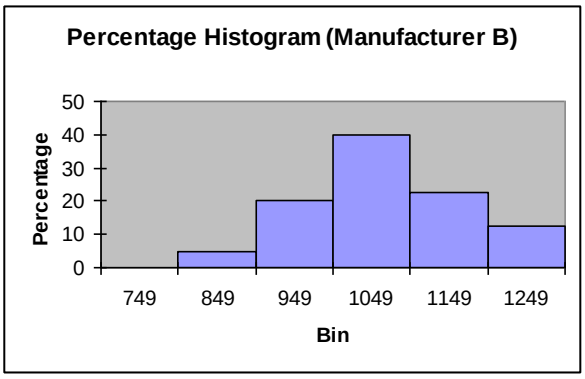
2.28 (a)

Bulb Life (hrs)	*Frequency Manufacturer A*	*Bulb Life (hrs)*	*Frequency Manufacturer B*
650 -- 749	3	750 -- 849	2
750 -- 849	5	850 -- 949	8
850 -- 949	20	950 -- 1049	16
950 -- 1049	9	1050 -- 1149	9
1050 -- 1149	3	1150 -- 1249	5

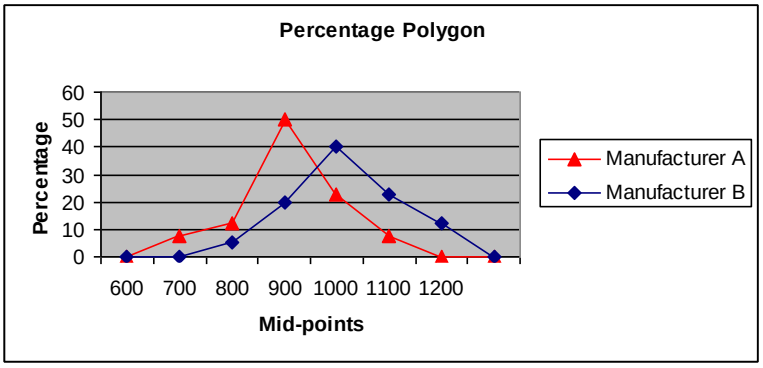
Bulb Life (hrs)	Percentage, Mfgr A	Percentage, Mfgr B
650 – 749	7.5%	0.0%
750 – 849	12.5	5.0
850 – 949	50.0	20.0
950 – 1049	22.5	40.0
1050 – 1149	7.5	22.5
1150 – 1249	0.0	12.5

(b)





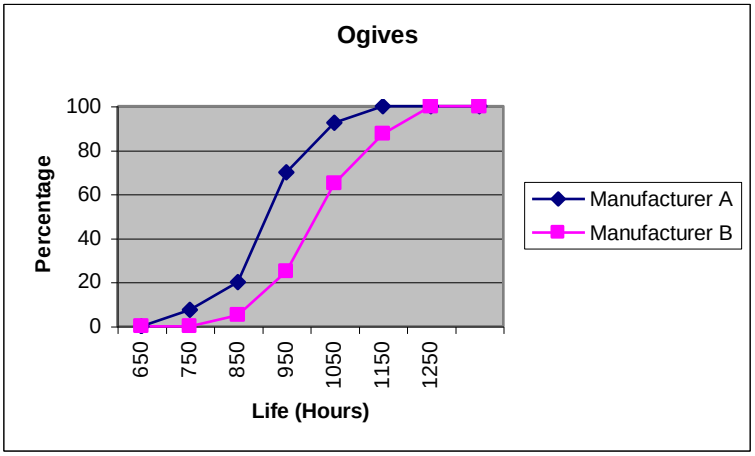
2.28 cont. (b)



(c)

Bulb Life (hrs)	Frequency Less Than, Mfgr A	Frequency Less Than, Mfgr B
650 – 749	3	0
750 – 849	8	2
850 – 949	28	10
950 – 1049	37	26
1050 – 1149	40	35
1150 – 1249	40	40

Bulb Life (hrs)	Percentage Less Than, Mfgr A	Percentage Less Than, Mfgr B
650 – 749	7.5%	0.0%
750 – 849	20.0	5.0
850 – 949	70.0	25.0
950 – 1049	92.5	65.0
1050 – 1149	100.0	87.5
1150 – 1249	100.0	100.0

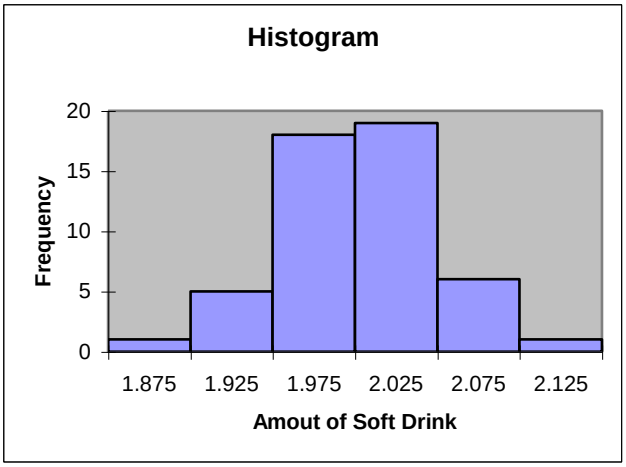


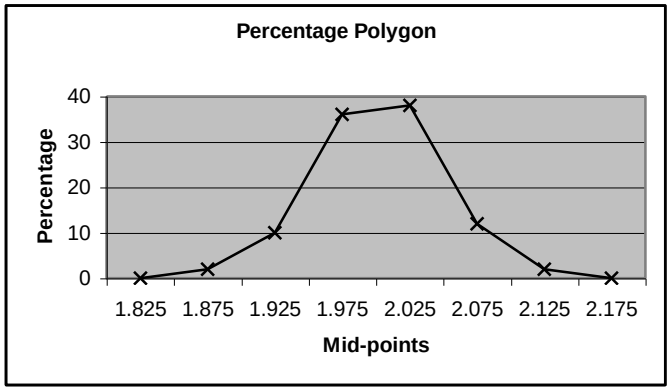
2.28 cont. (d) Manufacturer B produces bulbs with longer lives than Manufacturer A. The cumulative percentage for Manufacturer B shows 65% of their bulbs lasted 1049 hours or less contrasted with 70% of Manufacturer A's bulbs which lasted 949 hours or less. None of Manufacturer A's bulbs lasted more than 1149 hours, but 12.5% of Manufacturer B's bulbs lasted between 1150 and 1249 hours. At the same time, 7.5% of Manufacturer A's bulbs lasted less than 750 hours, while all of Manufacturer B's bulbs lasted at least 750 hours.

2.29 (a)

Amount of Soft Drink	Frequency	Percentage
1.850 – 1.899	1	2%
1.900 – 1.949	5	10
1.950 – 1.999	18	36
2.000 – 2.049	19	38
2.050 – 2.099	6	12
2.100 – 2.149	1	2

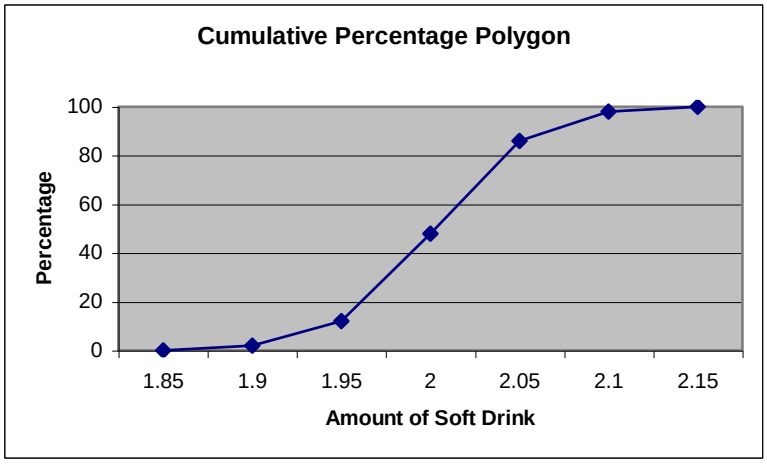
(b)





2.29 cont. (c)

Amount of Soft Drink	Frequency Less Than	Percentage Less Than
1.85 – 1.89	1	2%
1.90 – 1.94	6	12
1.95 – 1.99	24	48
2.00 – 2.04	43	86
2.05 – 2.09	49	98
2.10 – 2.14	50	100



(d) The amount of soft drink filled in the two liter bottles is most concentrated in two intervals on either side of the two-liter mark, from 1.950 to 1.999 and from 2.000 to 2.049 liters. Almost three-fourths of the 50 bottles sampled contained between 1.950 liters and 2.049 liters.

2.30 (a) Table frequencies for all student responses

	Student Major Categories			
Gender	A	C	M	Totals
Male	14	9	2	25
Female	6	6	3	15
Totals	20	15	5	40

(b) Table percentages based on overall student responses

	Student Major Categories			
Gender	A	C	M	Totals
Male	35.0%	22.5%	5.0%	62.5%
Female	15.0%	15.0%	7.5%	37.5%
Totals	50.0%	37.5%	12.5%	100.0%

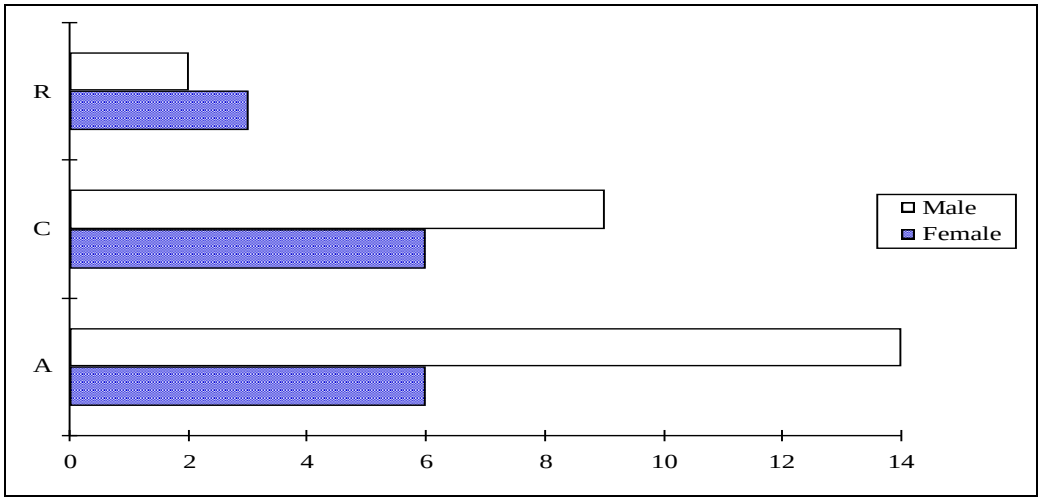
Table based on row percentages

	Student Major Categories			
Gender	A	C	M	Totals
Male	56.0%	36.0%	8.0%	100.0%
Female	40.0%	40.0%	20.0%	100.0%
Totals	50.0%	37.5%	12.5%	100.0%

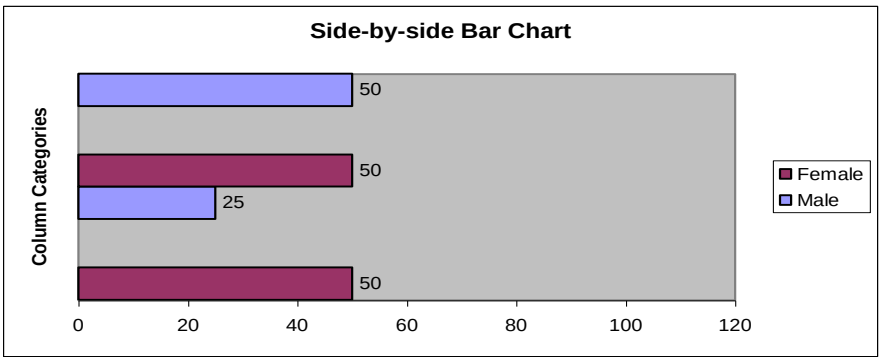
Table based on column percentages

	Student Major Categories			
Gender	A	C	M	Totals
Male	70.0%	60.0%	40.0%	62.5%
Female	30.0%	40.0%	60.0%	37.5%
Totals	100.0%	100.0%	100.0%	100.0%

2.30 (c)
cont.



2.31



2.32 (a) **Contingency Table**

	Condition of Die		
Quality	No Particles	Particles	Totals
Good	320	14	334
Bad	80	36	116
Totals	400	50	450

Table of Total Percentages

	Condition of Die		
Quality	No Particles	Particles	Totals
Good	71%	3%	74%
Bad	18%	8%	26%
Totals	89%	11%	100%

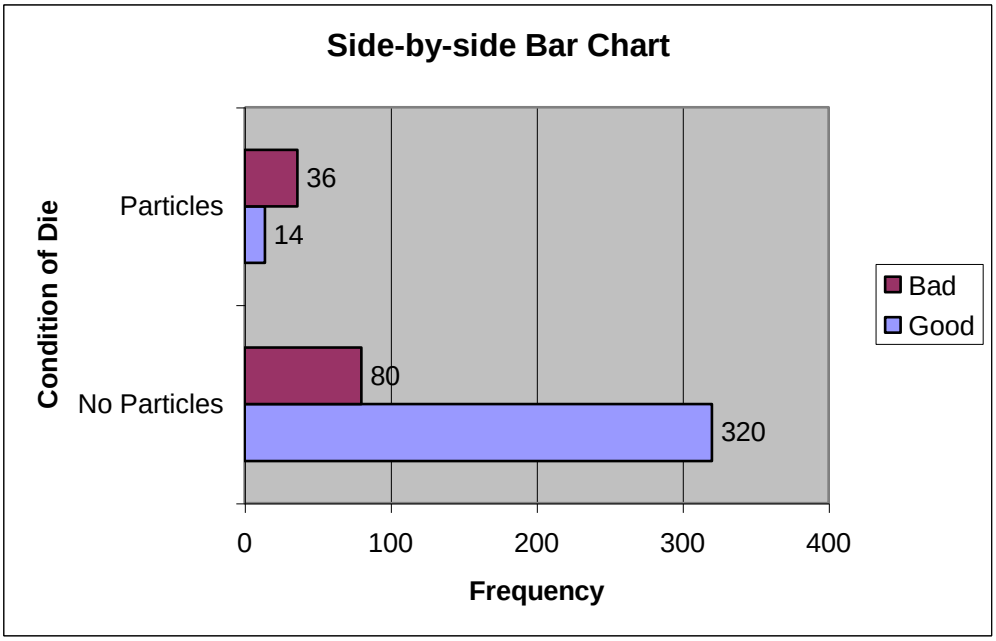
Table of Row Percentages

	Condition of Die		
Quality	No Particles	Particles	Totals
Good	96%	4%	100%
Bad	69%	31%	100%
Totals	89%	11%	100%

Table of Column Percentages

	Condition of Die		
Quality	No Particles	Particles	Totals
Good	80%	28%	74%
Bad	20%	72%	26%
Totals	100%	100%	100%

(b)



(c) The data suggests that there is some association between condition of the die and the quality of wafer because more good wafers are produced when no particles are found in the die and more bad wafers are produced when there are particles found in the die.

2.33 (a)

Table of total percentages

	Shift		
	Day	Evening	
Nonconforming	1.6%	2.4%	4%
Conforming	65.4%	30.6%	96%
Total	67%	33%	100%

Table of row percentages

	Shift		
	Day	Evening	
Nonconforming	40%	60%	100%
Conforming	68%	32%	100%
Total	67%	33%	100%

Table of column percentages

	Shift		
	Day	Evening	
Nonconforming	2%	7%	4%
Conforming	98%	93%	96%
Total	100%	100%	100%

(b) The row percentages allow us to block the effect of disproportionate group size and show us that the pattern for day and evening tests among the nonconforming group is very different from the pattern for day and evening tests among the conforming group. Where 40% of the nonconforming group was tested during the day, 68% of the conforming group was tested during the day.

(c) The director of the lab may be able to cut the number of nonconforming tests by reducing the number of tests run in the evening, when there is a higher percent of tests run improperly.

2.34 (a) Table of total percentages

	Gender		
Enjoy Shopping for Clothing	Male	Female	Total
Yes	27%	45%	72%
No	21%	7%	28%
Total	48%	52%	100%

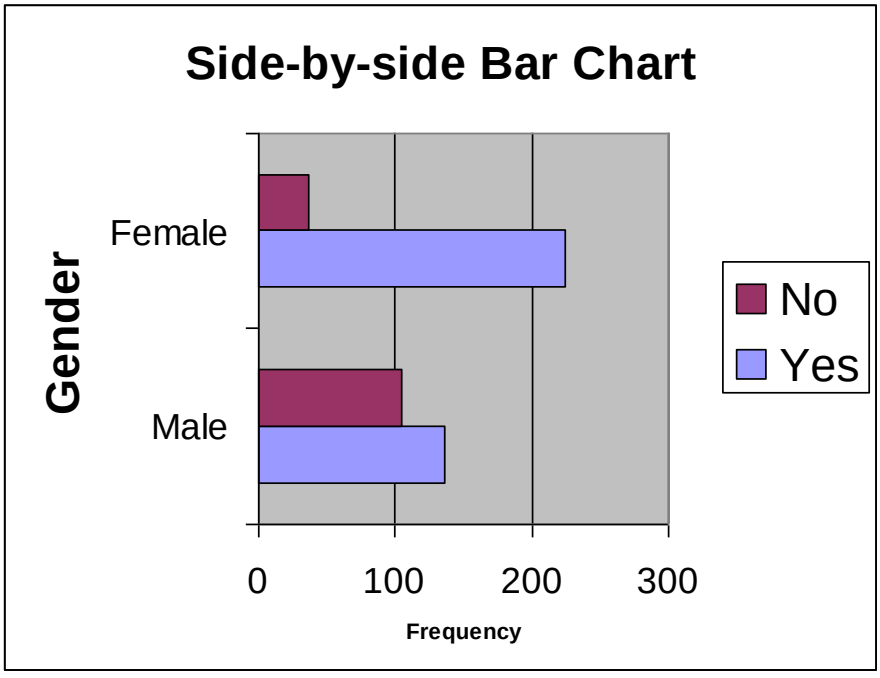
Table of row percentages

	Gender		
Enjoy Shopping for Clothing	Male	Female	Total
Yes	38%	62%	100%
No	74%	26%	100%
Total	48%	52%	100%

Table of column percentages

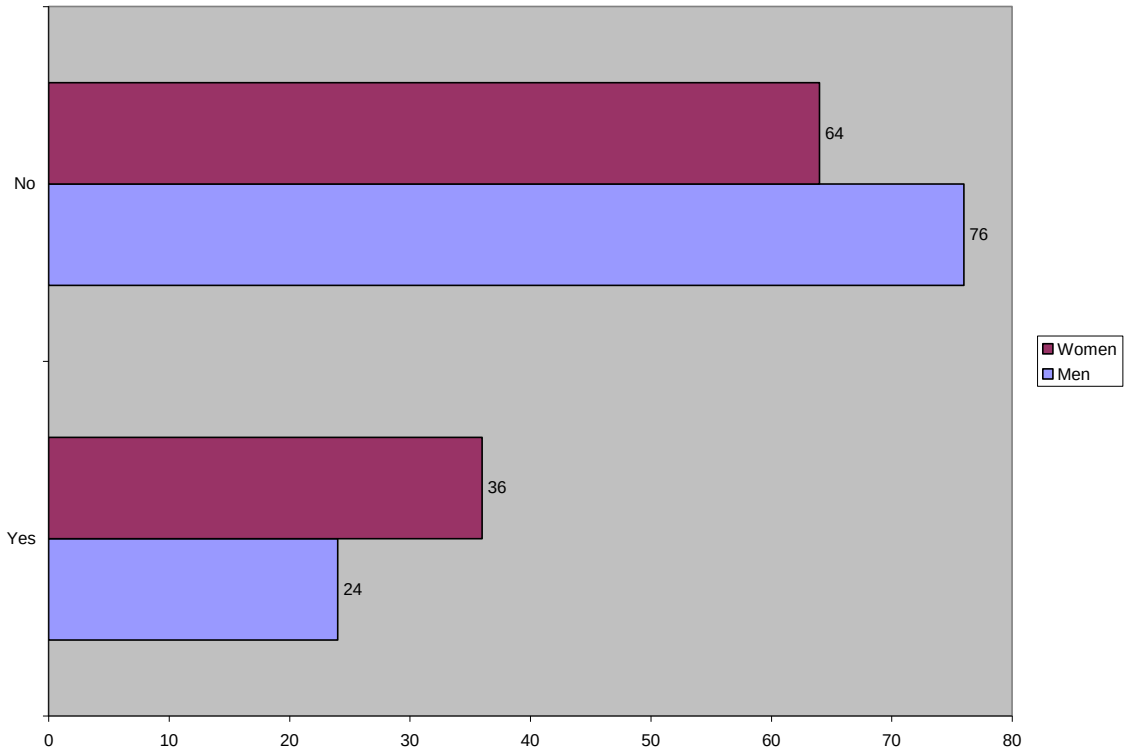
	Gender		
Enjoy Shopping for Clothing	Male	Female	Total
Yes	57%	86%	72%
No	43%	14%	28%
Total	100%	100%	100%

(b)



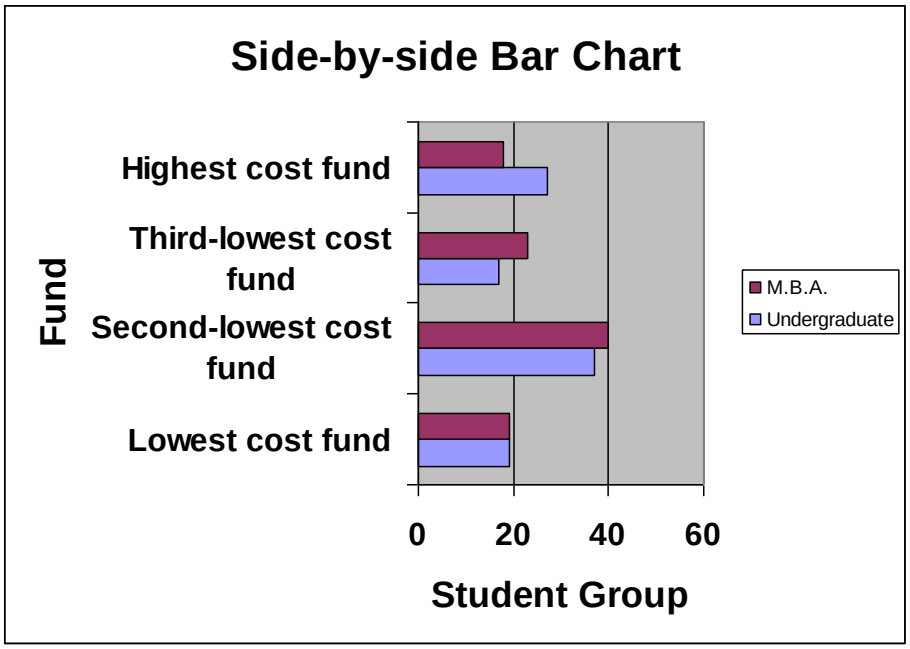
(c) The percentage of shoppers who enjoy shopping for clothing is higher among females than males.

2.35 (a) Side-by-Side Bar Chart



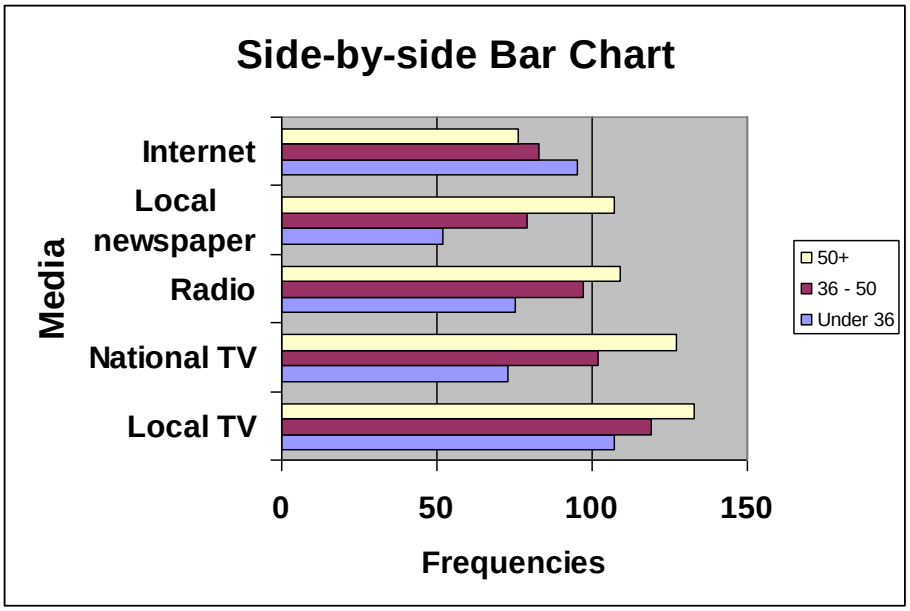
(b) A higher percentage of women feel that it is ok to yawn in public as compared to men.

2.36 (a)



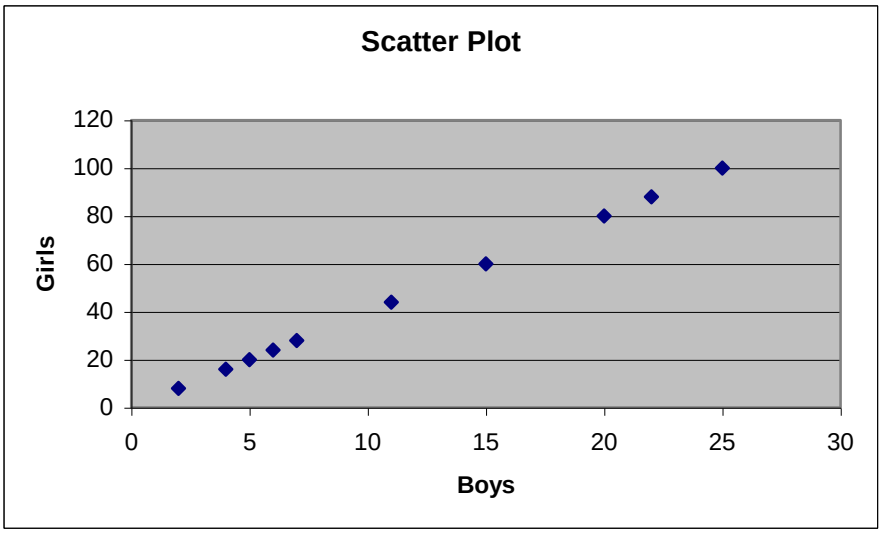
(b) The number of MBA and undergraduate students who choose the lowest cost fund and the second-lowest cost fund is about the same. More MBA students choose the third-lowest cost fund while more undergraduate students choose the highest cost fund.

2.37 (a)



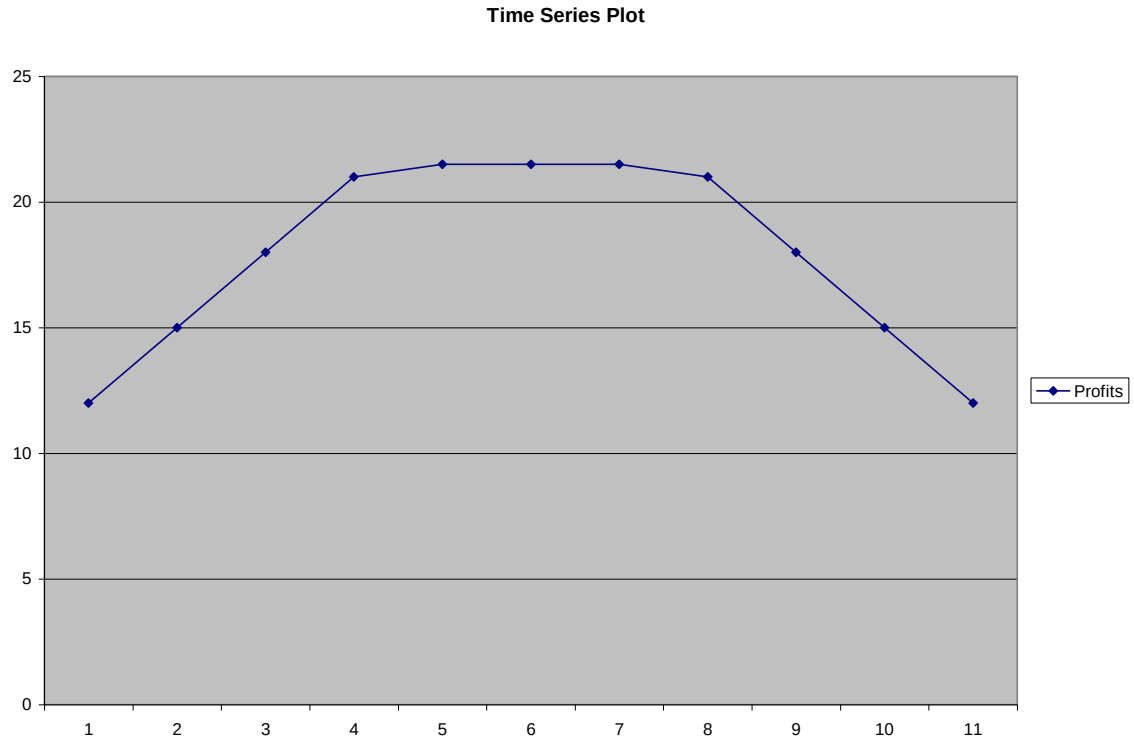
(b) Among the three groups, the "under 36" group has the lowest number of occurrences who get their news from local newspaper and the highest frequency who get their news from the Internet. The "50+" group has the highest frequency who get their news from national TV, local TV, radio, and local newspaper.

2.38 (a)



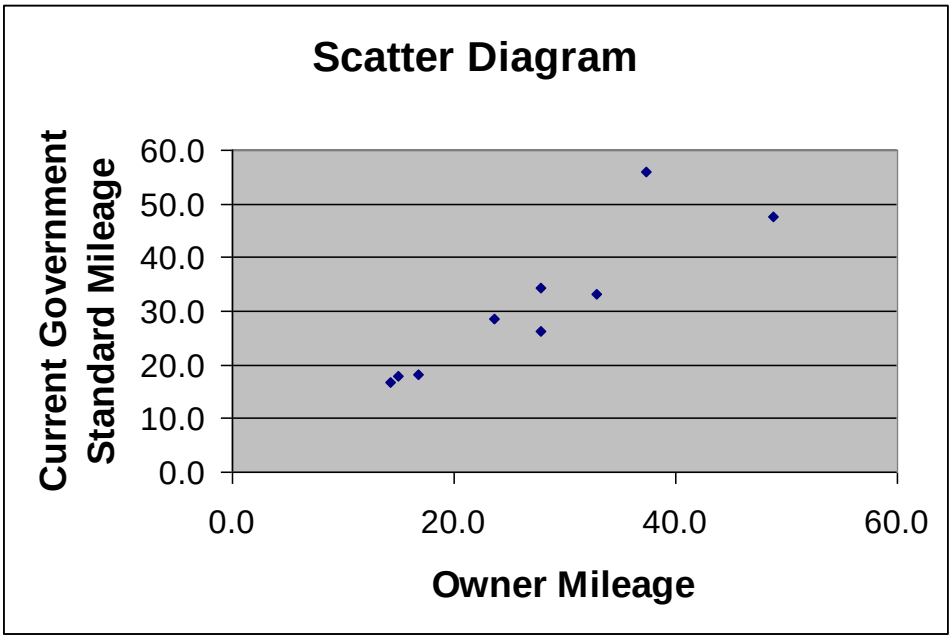
(b) Yes, there appears to be a positive linear relationship between *Boys* and *Girls*

2.39 (a)



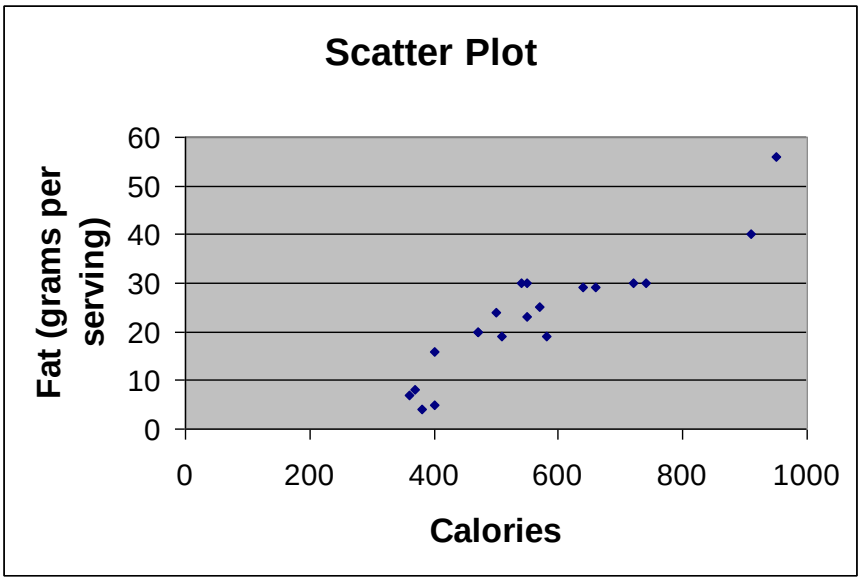
(b) Profit appear to be increasing in the earlier years from 1 to 4, and then it is stable from years 4 to 8 and starts declining thereafter from 9 to 11 years.

2.40 (a)



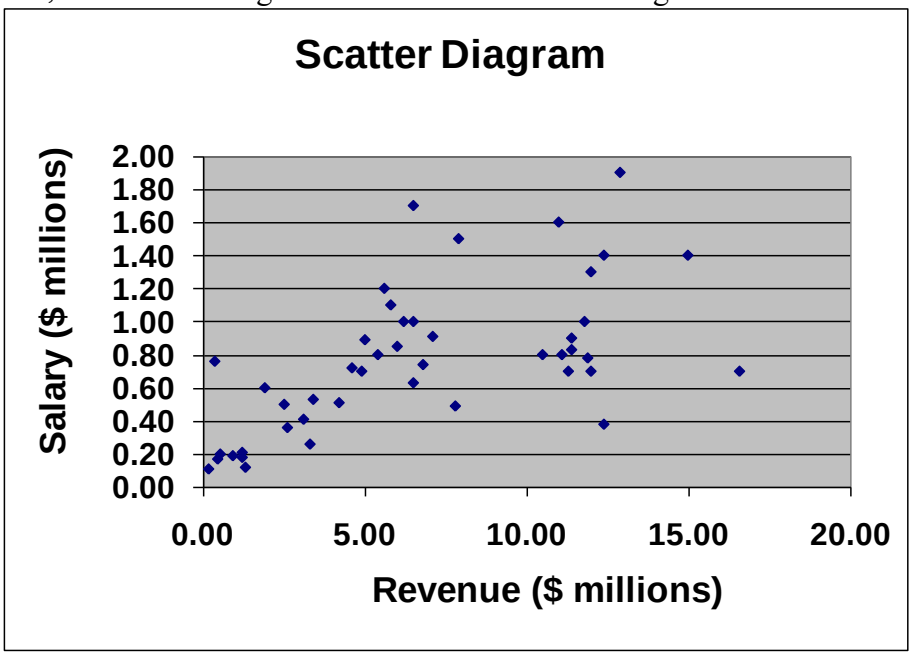
(b) There is a positive relationship between owner mileage and current government standard mileage.

2.41 (a)



(b) There is a positive relationship between calories and total fat in chicken sandwiches.

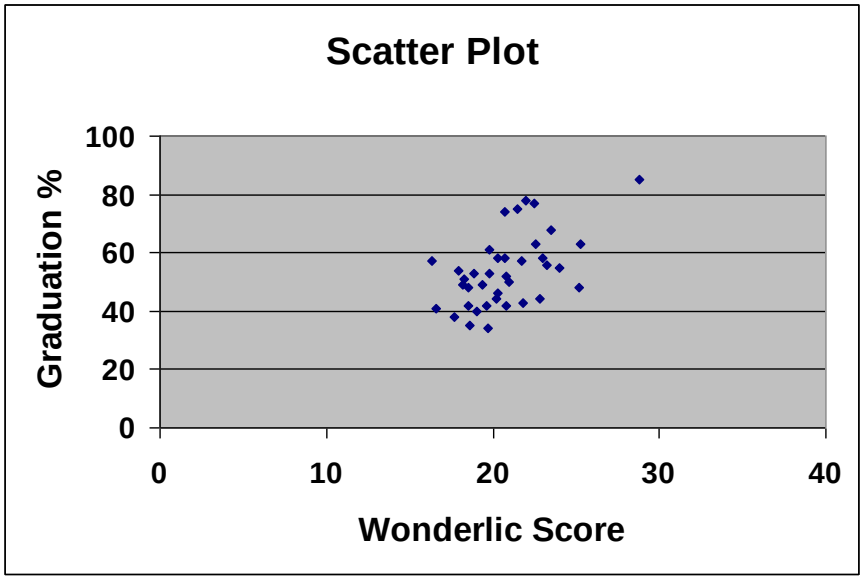
2.42 (a) Yes, schools with higher revenues will also have higher coach's salaries.



(b) There appears to be a positive relationship between coaches' salary and revenue.

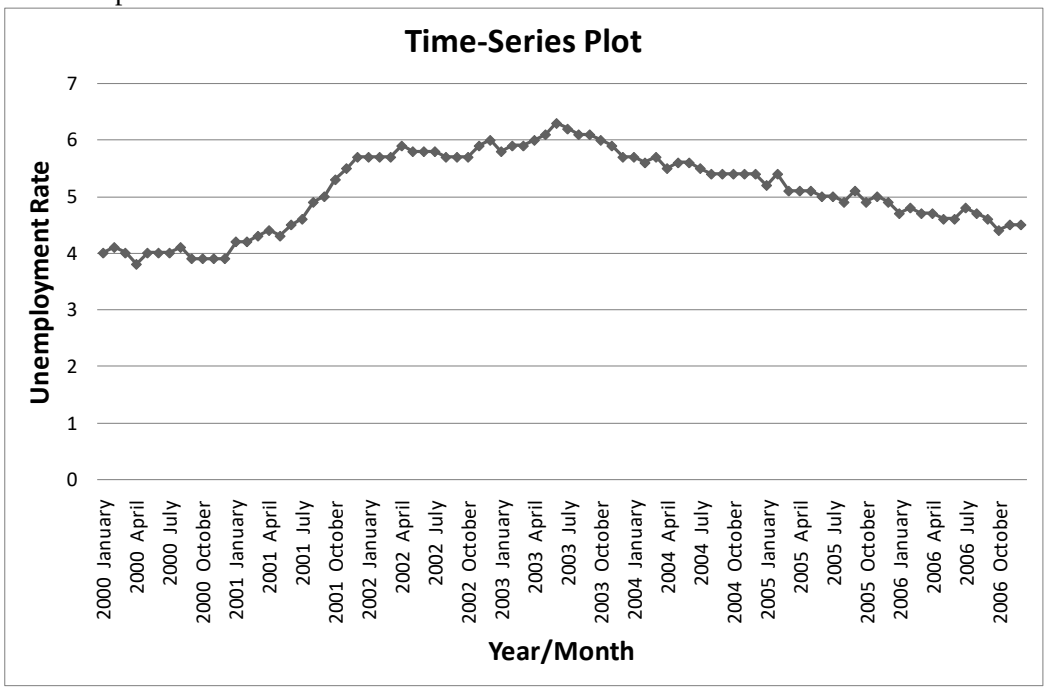
(c) The scatter plot confirms the answer to (a).

2.43 (a)



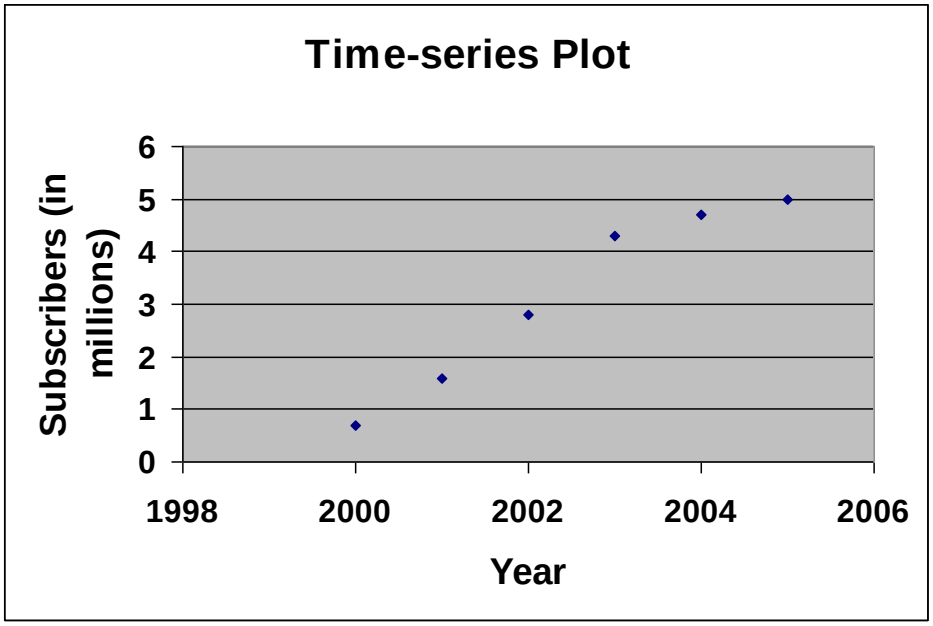
(b) There is a positive relationship between Wonderlic score and graduation rate.

2.44 (a) Excel output:



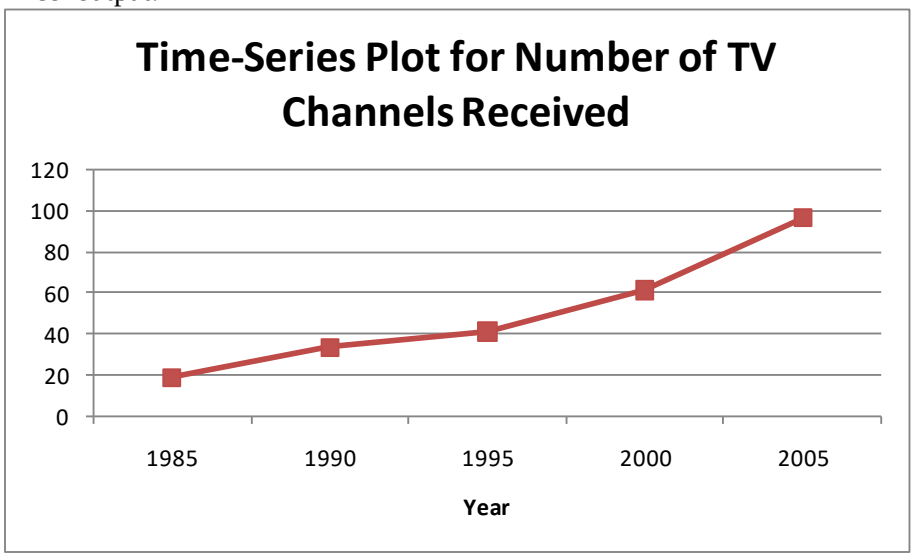
(b) The unemployment rate was quite stable at around 4% from January 2000 to around January 2001. Then it trended upward and leveled off at around 6% by December 2001. Around October 2003, it started to trend downward and reached about 4.5% by December 2006.

2.45 (a)



(b) The number of subscribers is increasing but the rate of growth is decreasing because the slope of the line segments that join the data points is decreasing over the years.

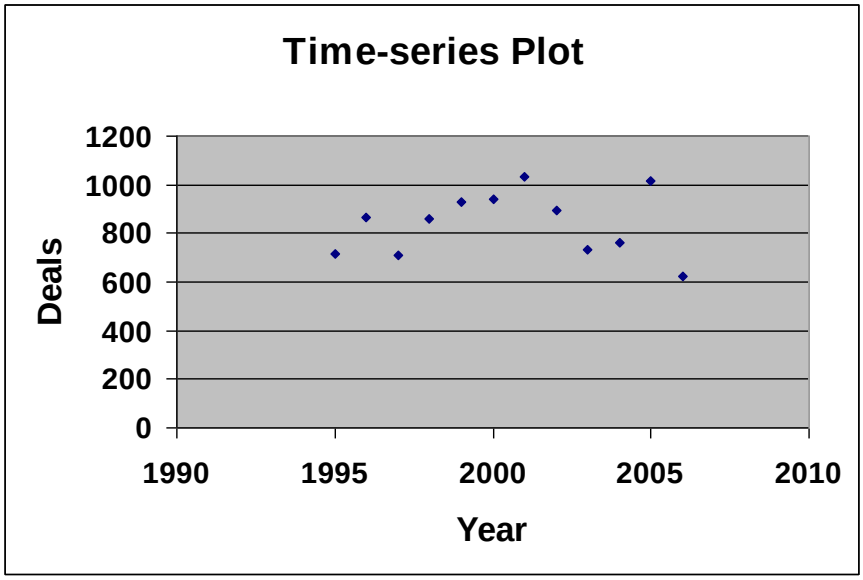
2.46 (a) Excel output:



(b) There is an obvious upward trend in the average number of TV channels that the U. S. home received from 1985 to 2005.

(c) With extrapolation, you would predict the average number of TV channels that the U. S. home will receive in 2010 to be around 140.

2.47 (a)



(b) The data fluctuate around 900 mergers and acquisitions.

(c) You would predict around 800 mergers and acquisitions made during January 1 through January 11 of 2007.

2.48 Student answers will vary.

2.49 Student answers will vary.

2.50 Student answers will vary.

2.51 Student answers will vary.

2.52 Student answers will vary.

2.53 Student answers will vary.

2.54 Student answers will vary.

2.55 Student answers will vary.

2.56 Student answers will vary.

2.57 A histogram uses bars to represent each class while a polygon uses a single point. The histogram should be used for only one group, while several polygons can be plotted on a single graph.

2.58 A summary table allows one to determine the frequency or percentage of occurrences in each category.

2.59 A bar chart is useful for comparing categories. A pie chart is useful when examining the portion of the whole that is in each category. A Pareto diagram is useful in focusing on the categories that make up most of the frequencies or percentages.

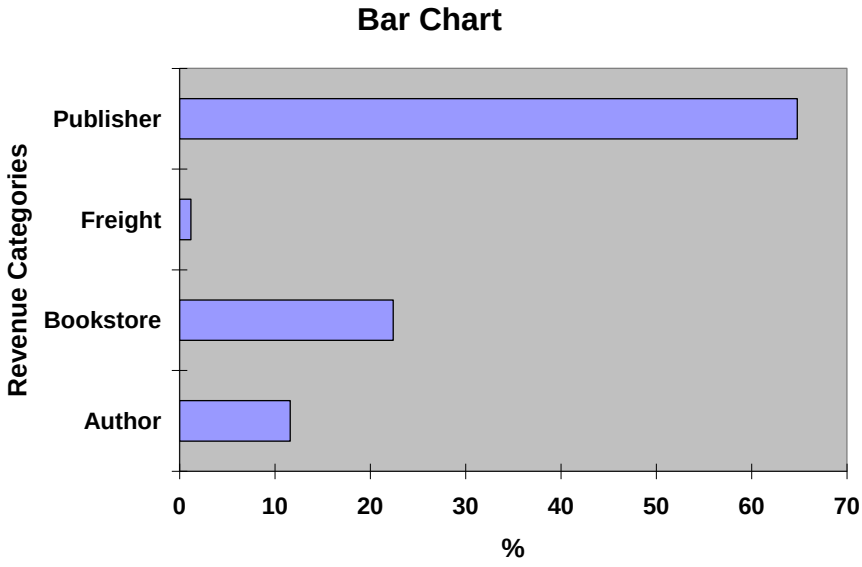
2.60 The bar chart for categorical data is plotted with the categories on the vertical axis and the frequencies or percentages on the horizontal axis. In addition, there is a separation between categories. The histogram is plotted with the class grouping on the horizontal axis and the frequencies or percentages on the vertical axis. This allows one to more easily determine the distribution of the data. In addition, there are no gaps between classes in the histogram.

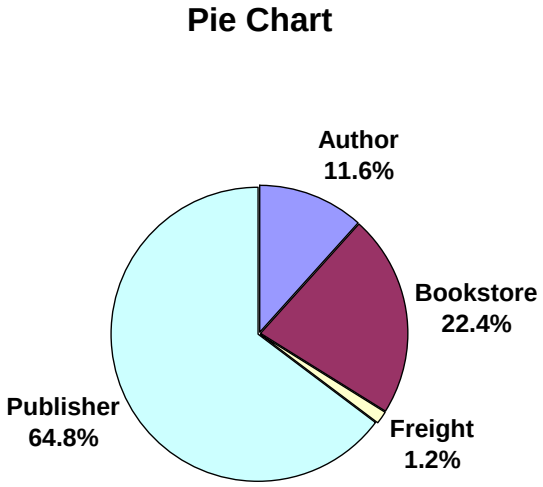
2.61 A time-series plot is a type of scatter diagram with time on the x-axis.

2.62 Because the categories are arranged according to frequency or importance, it allows the user to focus attention on the categories that have the greatest frequency or importance.

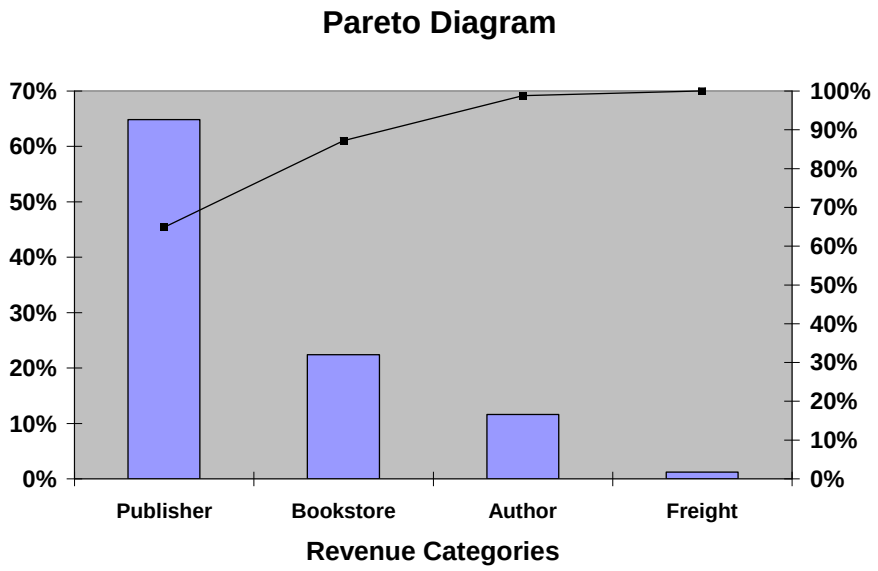
2.63 Percentage breakdowns according to the total percentage, the row percentage, and/or the column percentage allow the interpretation of data in a two-way contingency table from several different perspectives.

2.64 (a)

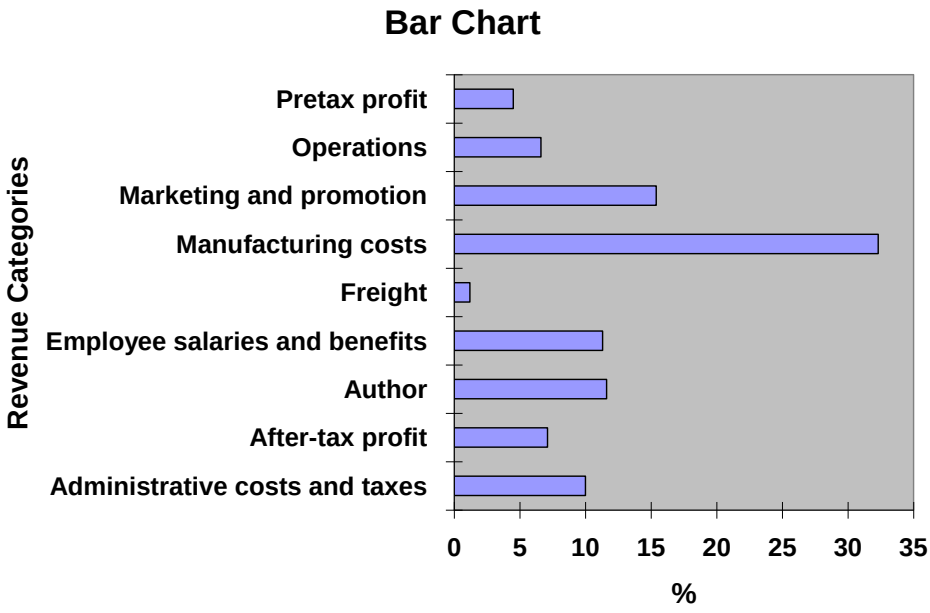




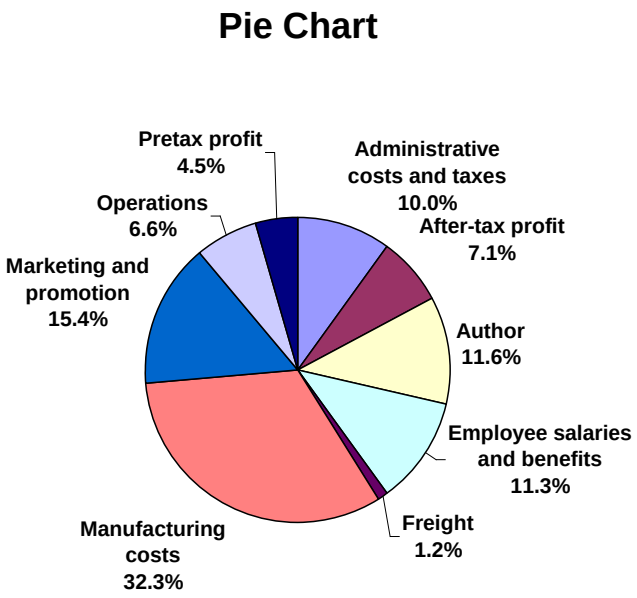
2.64 cont. (a)

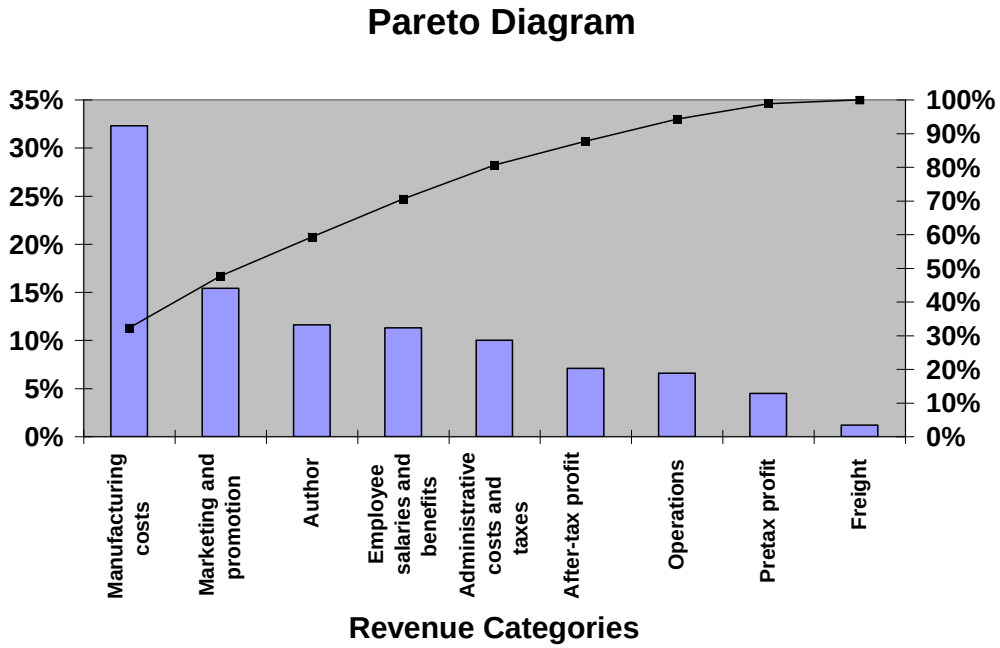


(b)



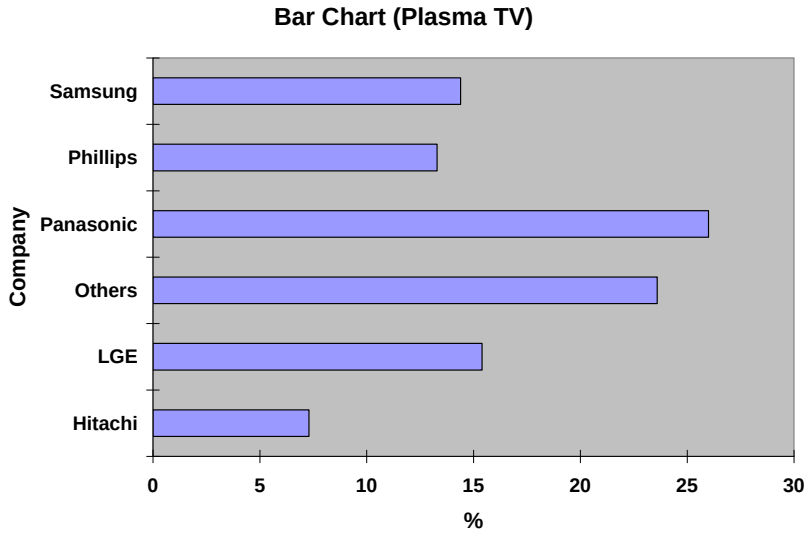
2.64 (b)
cont.

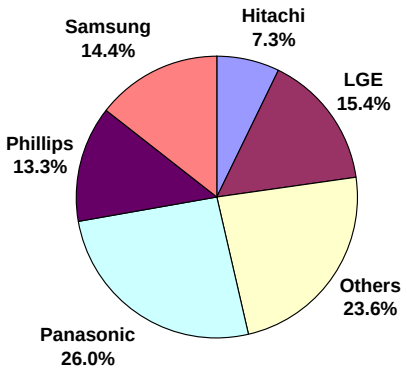


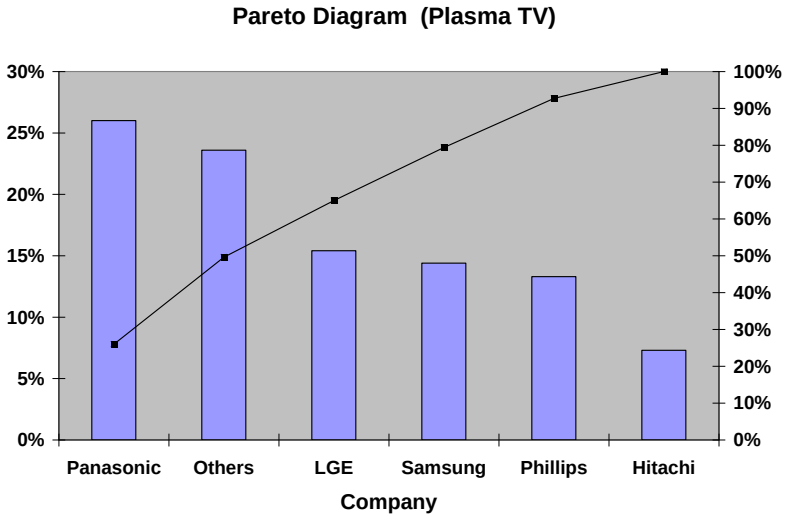


(c) The publisher gets the largest portion (64.8%) of the revenue. About half (32.2%) of the revenue received by the publisher is used for manufacturing costs. Publisher's marketing and promotion account for the next larger share of the revenue at 15.4%. Author, bookstore employee salaries and benefits, and publisher administrative costs and taxes each accounts for around 10% of the revenue while the publisher after-tax profit, bookstore operations, bookstore pretax profit and freight constitute the "trivial few" allocations of the revenue.

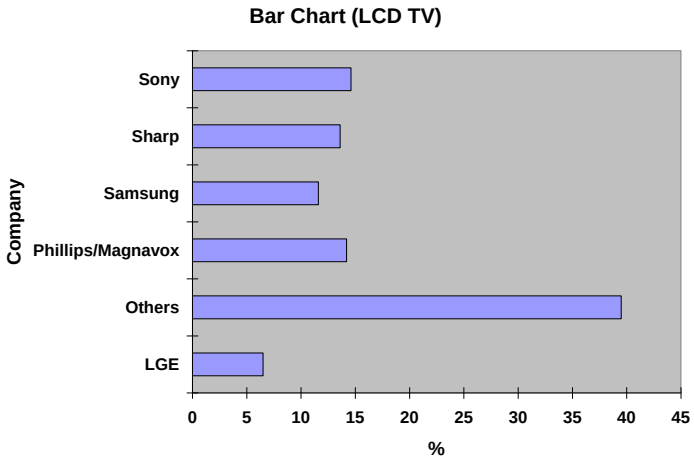
2.65 (a)

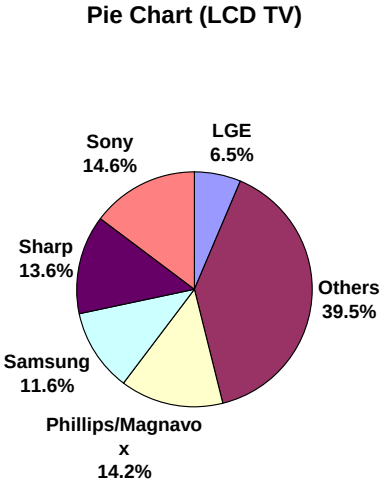


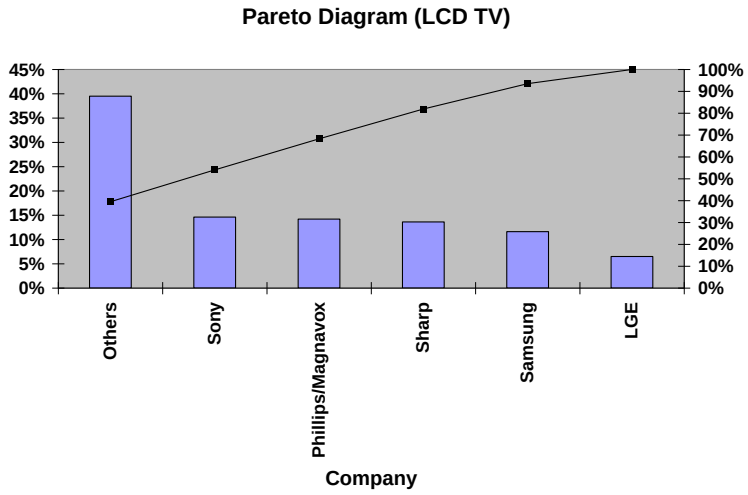




2.65 cont. (a)

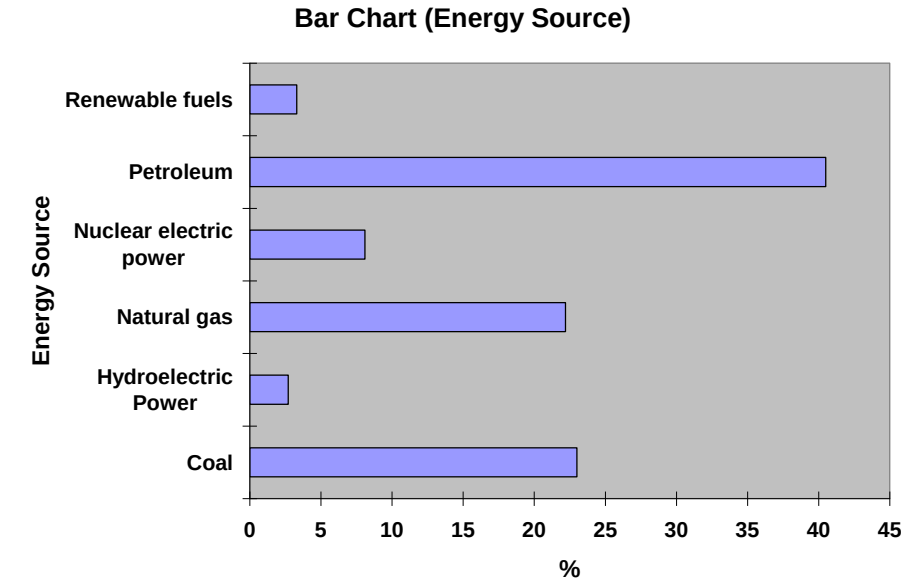


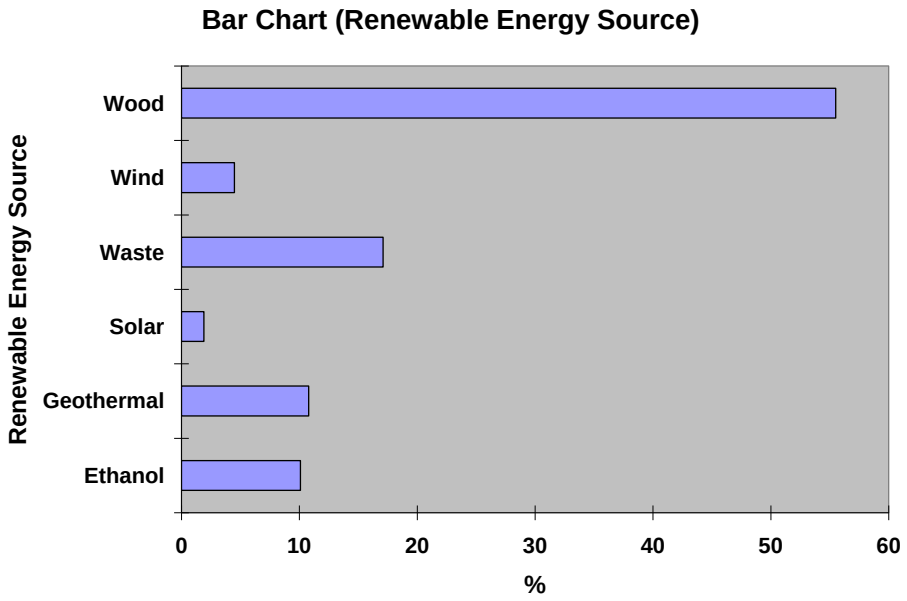




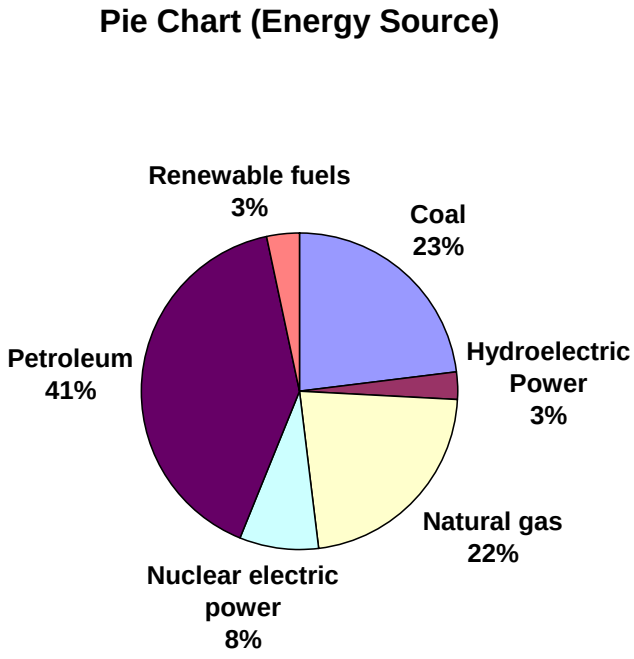
(b) Panasonic has the largest market share at 26% in the plasma TV market while the LCD market is not dominated by any manufacturer.

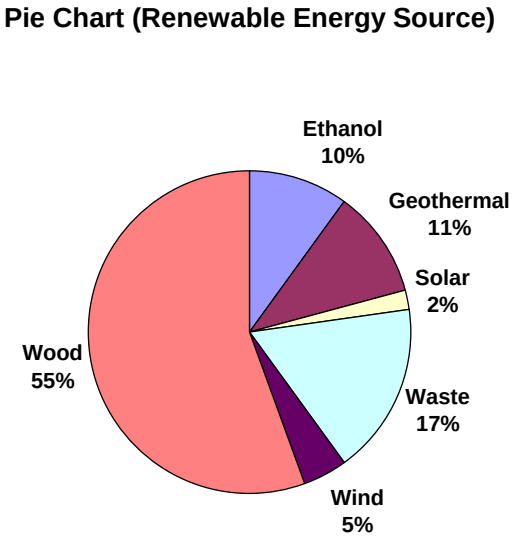
2.66 (a)



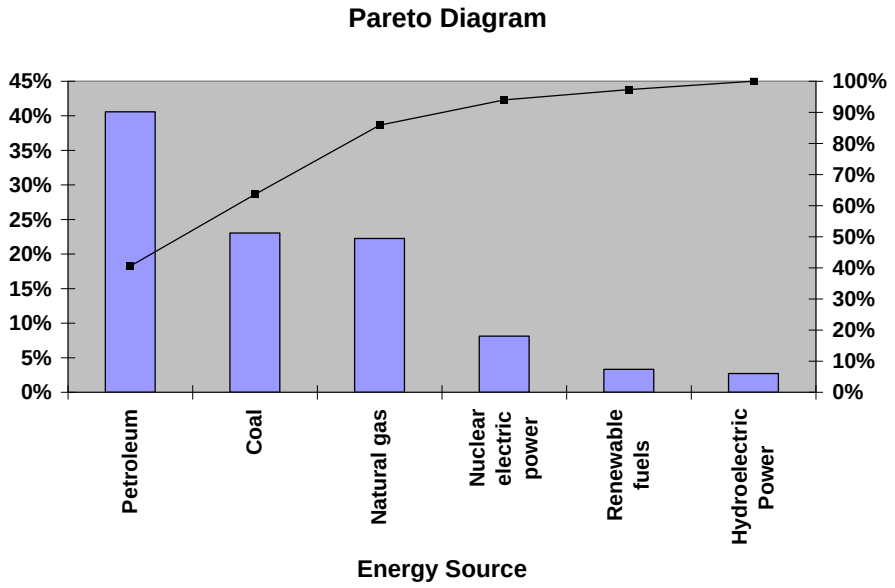


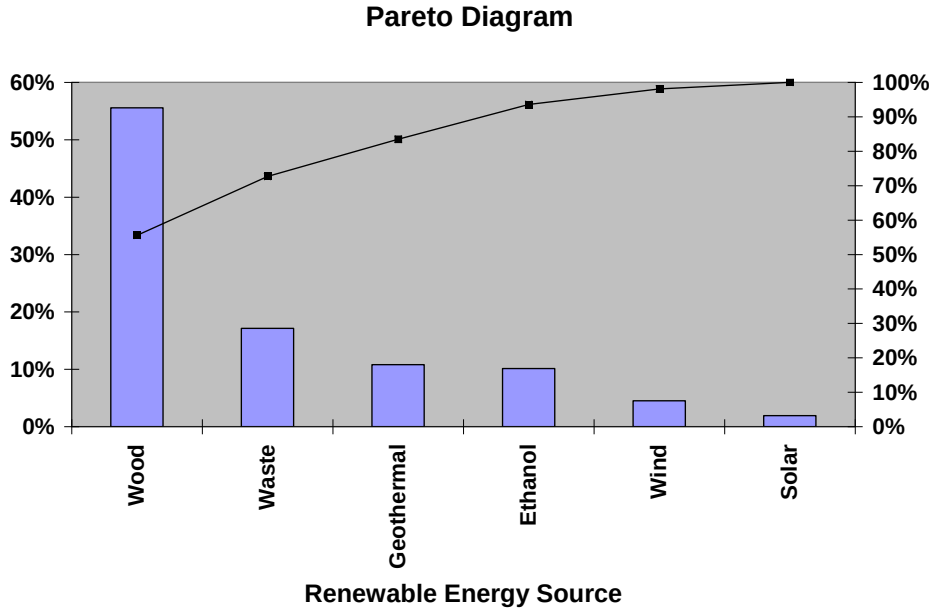
2.66 (a)
cont.





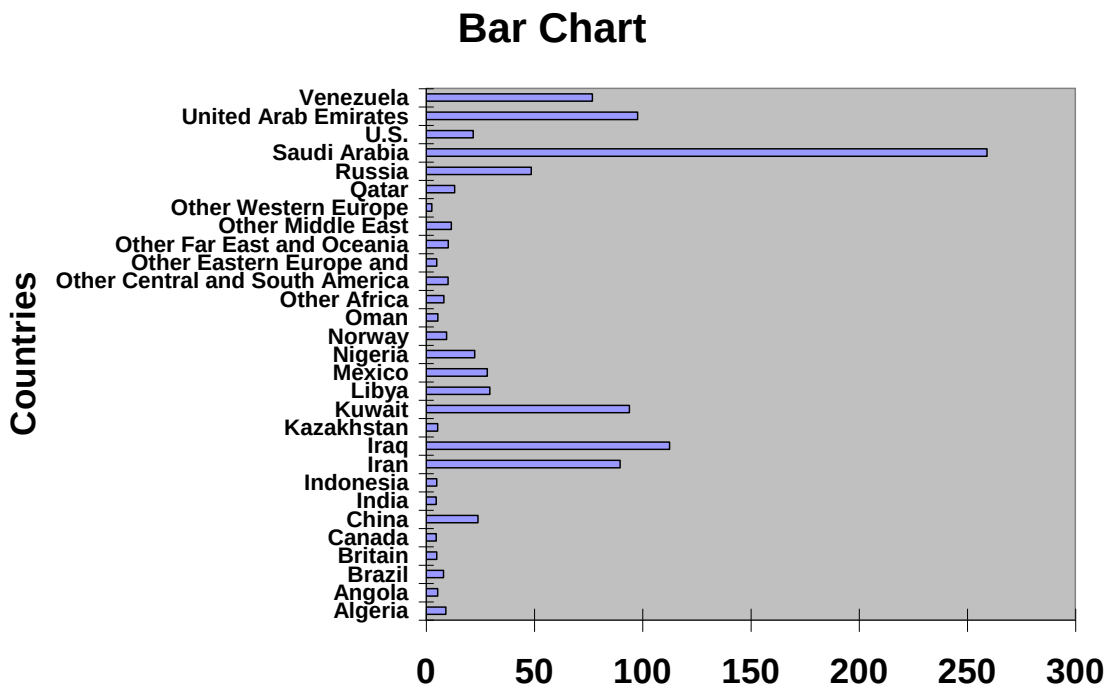
2.66 cont. (a)

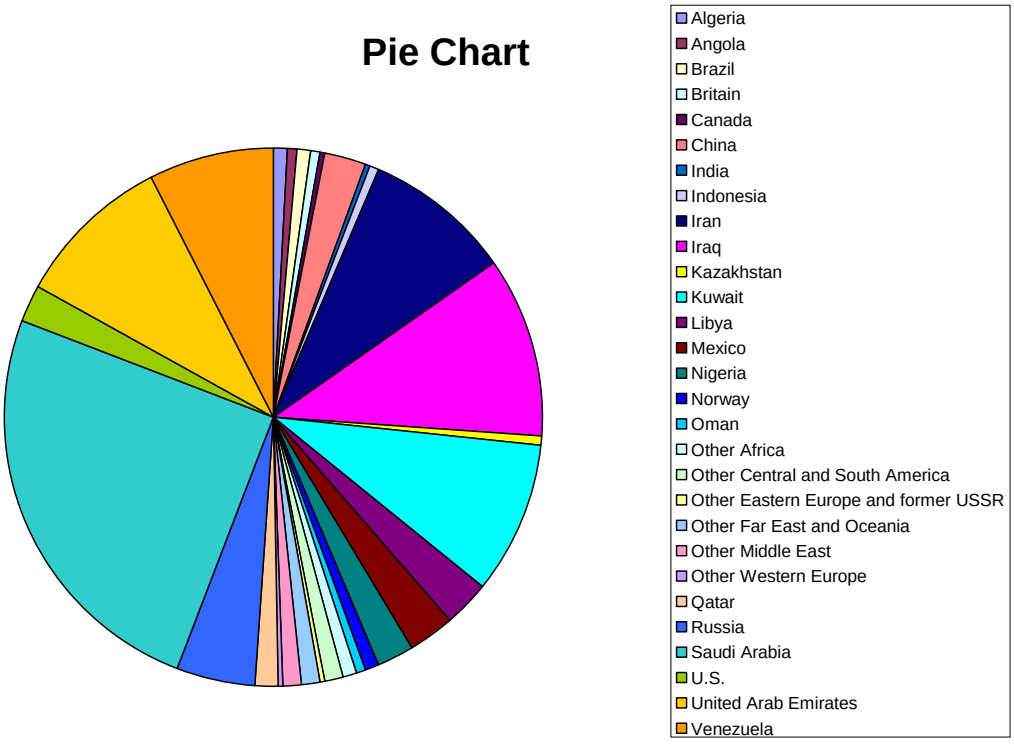




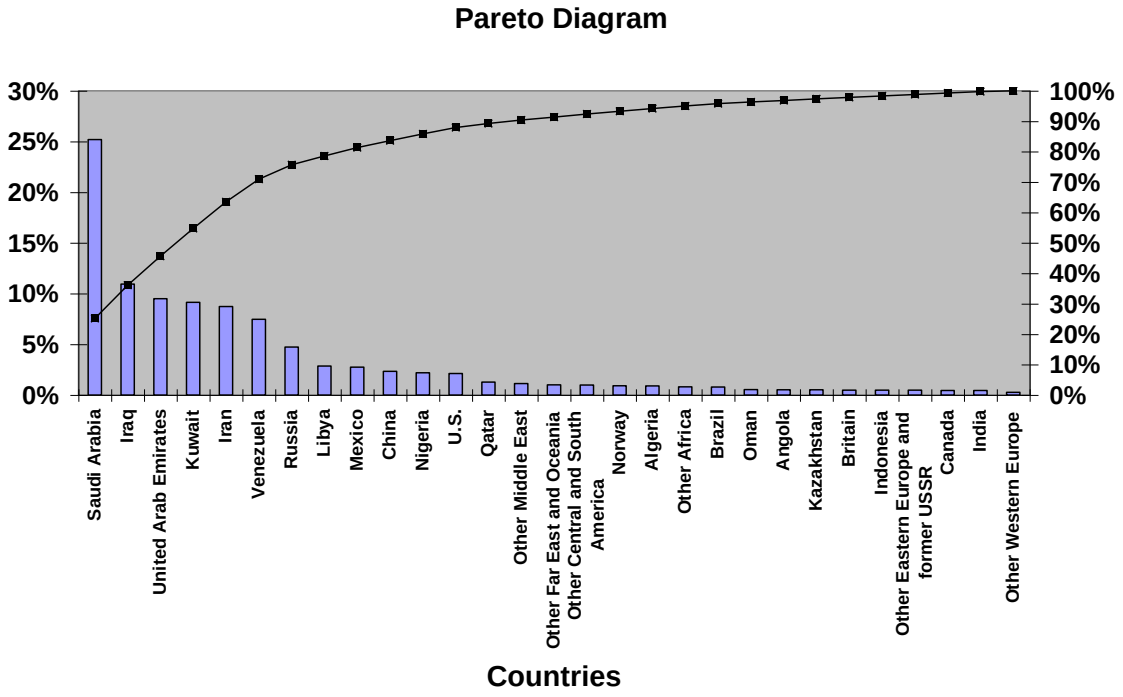
(b) In 2005, U.S. relied on petroleum heavily followed by coal and natural gas as major sources of energy while renewable fuels accounted for less than 4% of the total consumption. Wood accounted for more than half of the renewable energy consumption.

2.67 (a)

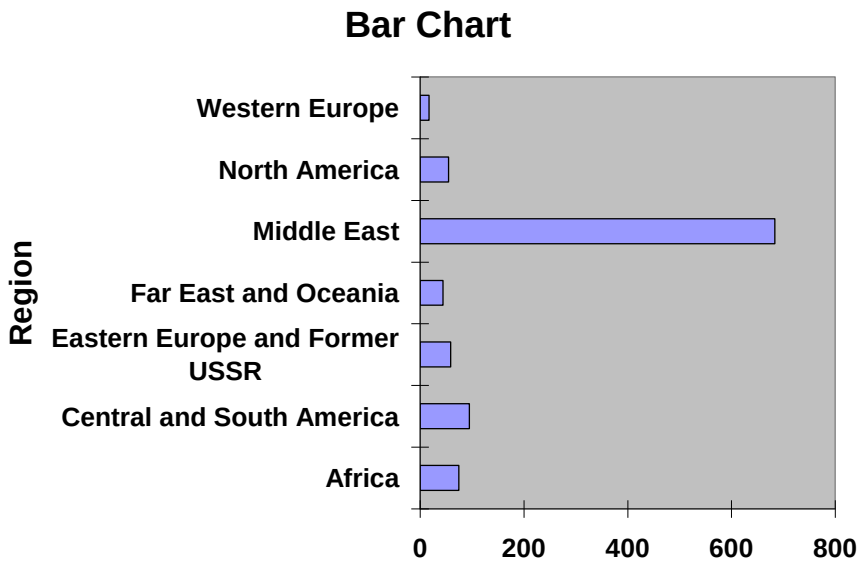




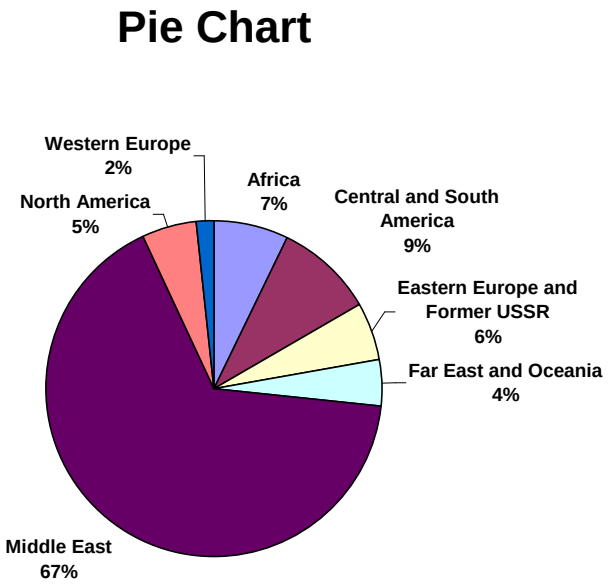
2.67 cont. (a)

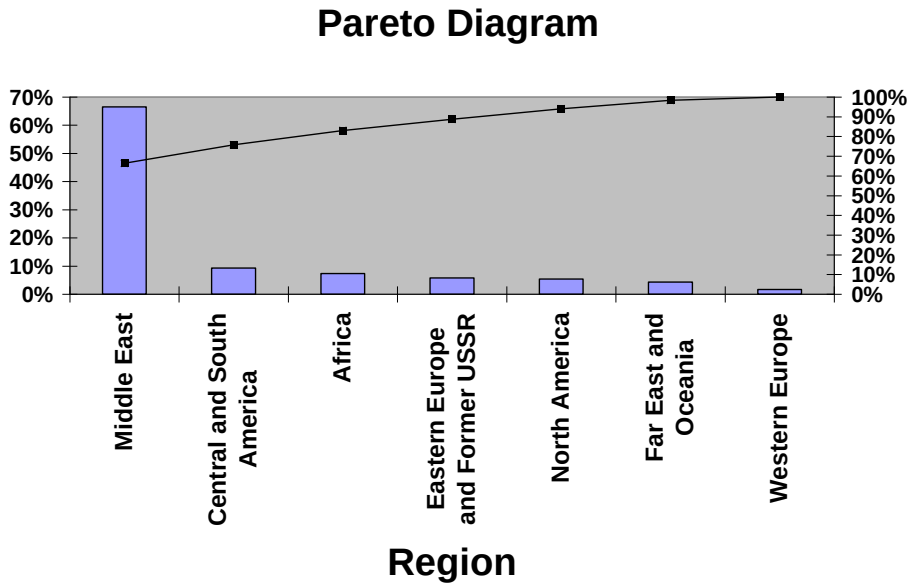


(b)



2.67 (b)
cont.

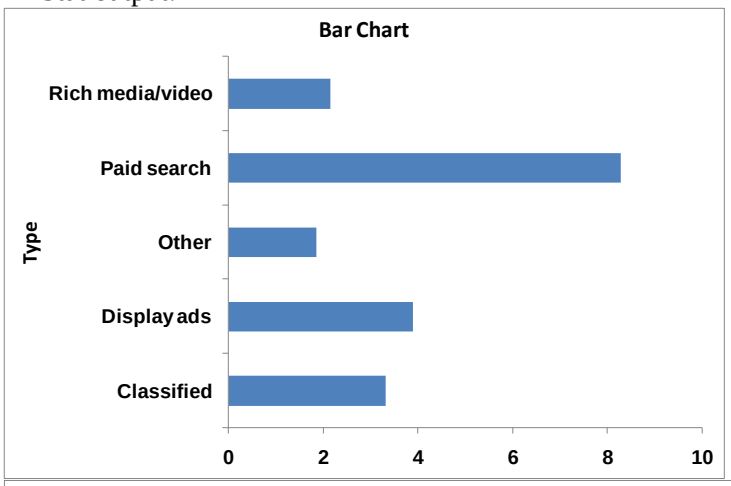


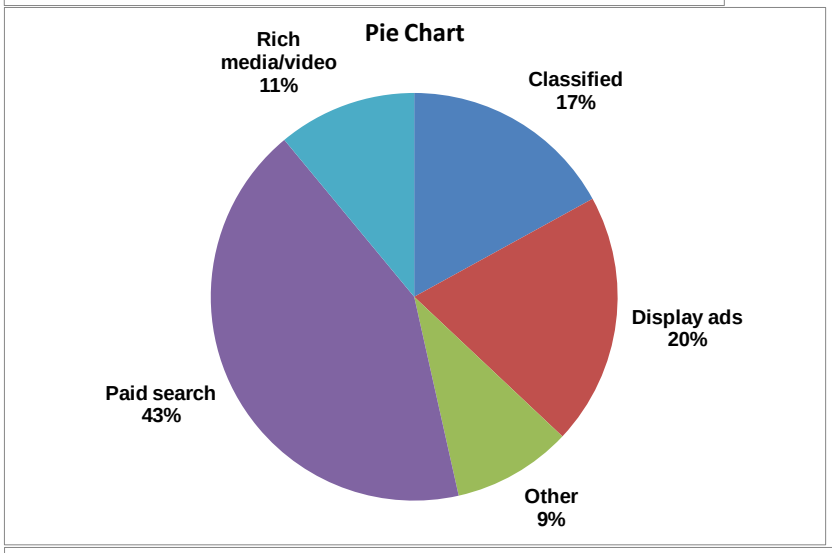


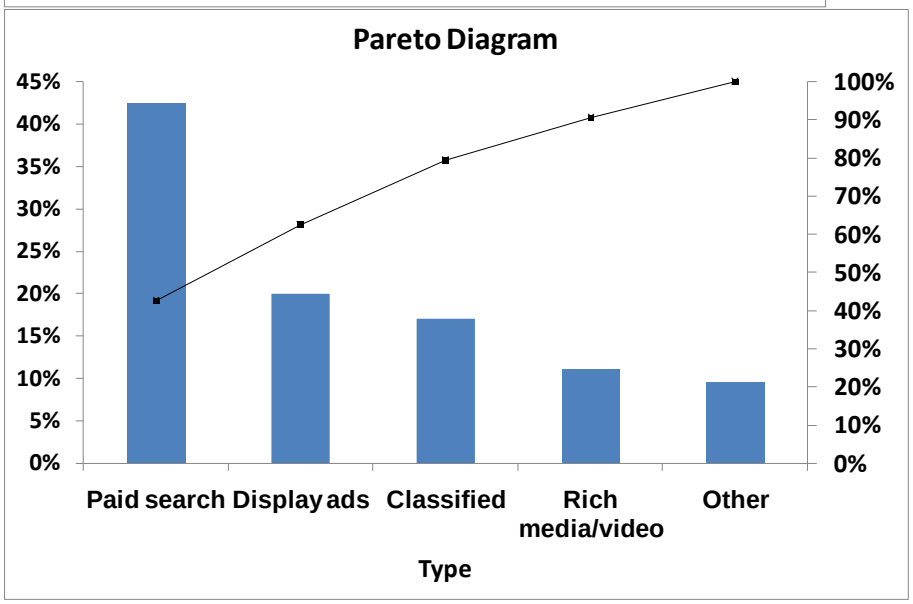
(c) The Pareto diagram is most appropriate because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

(d) The Middle East, with a share of more than 60%, obviously has the largest proven conventional reserves. Among the set of countries, Saudi Arabia has the largest share of proven conventional reserves followed by Iraq, United Arab Emirates and Kuwait. These four countries account for more than half of the reserves among the set of countries.

2.68 (a) PHStat output:

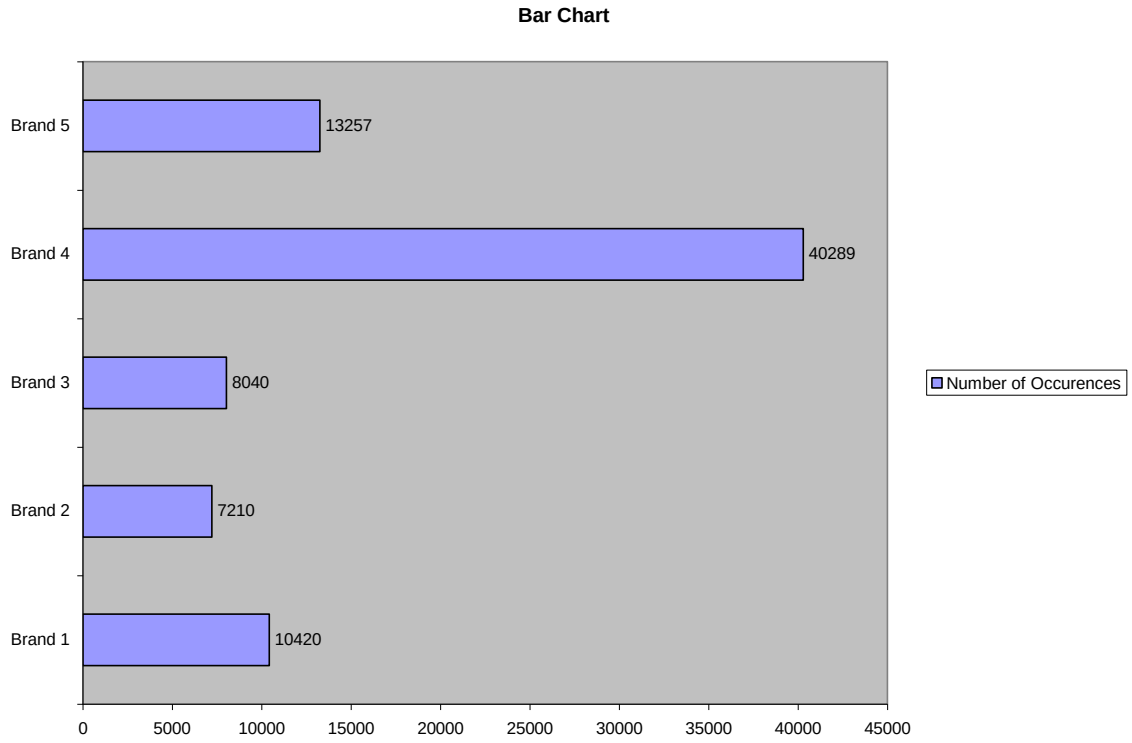


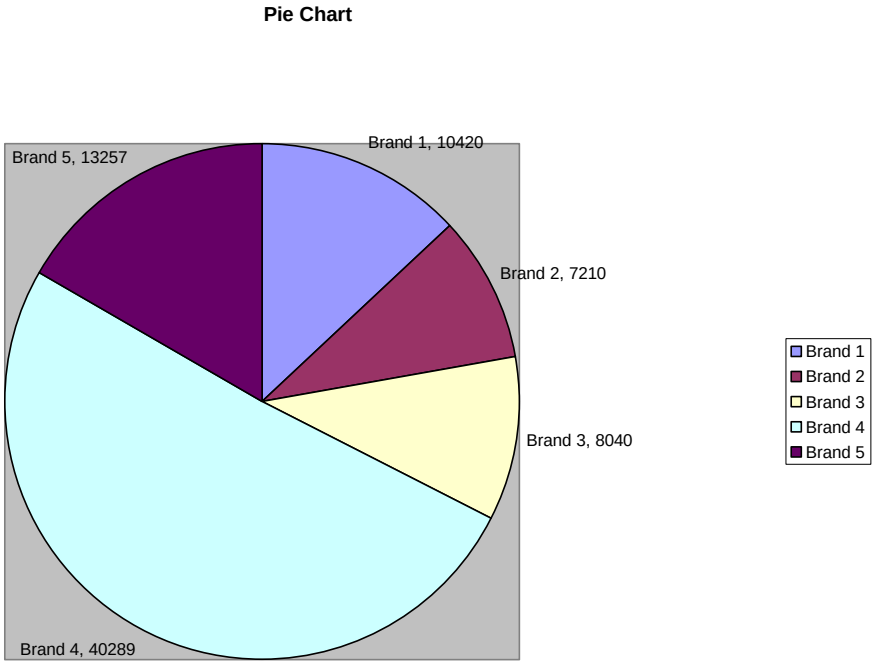




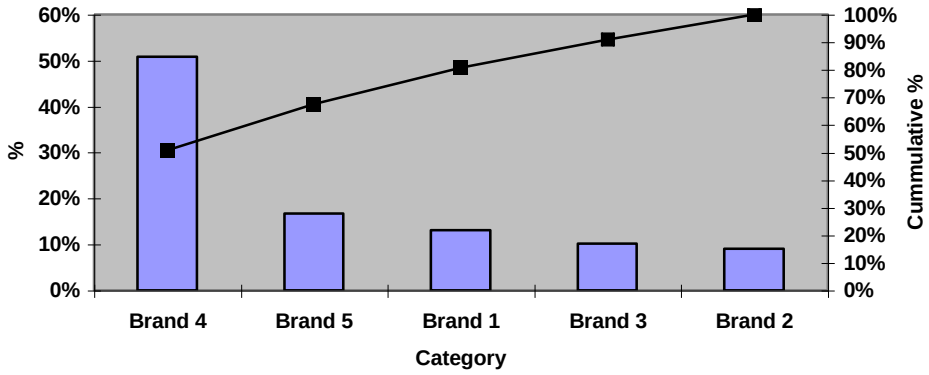
(b) The Pareto diagram is most appropriate because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

(c)
cont.





Pareto Diagram



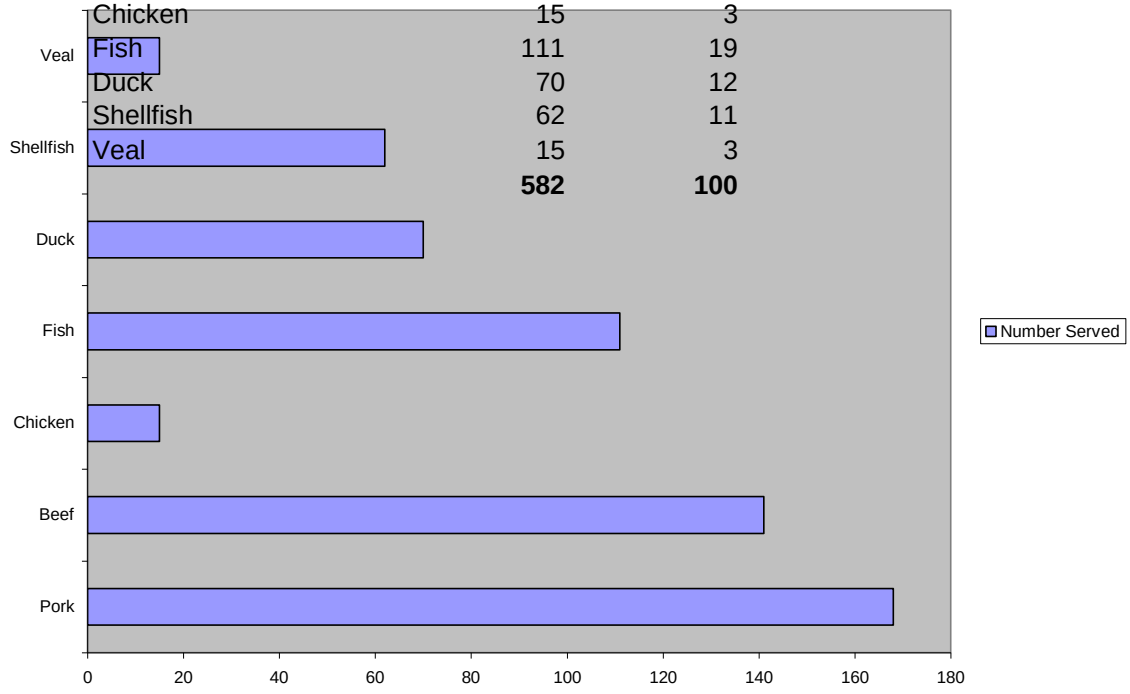
(d) The Pareto diagram is most appropriate because it not only sorts the frequencies in descending order, it also provides the cumulative polygon on the same scale.

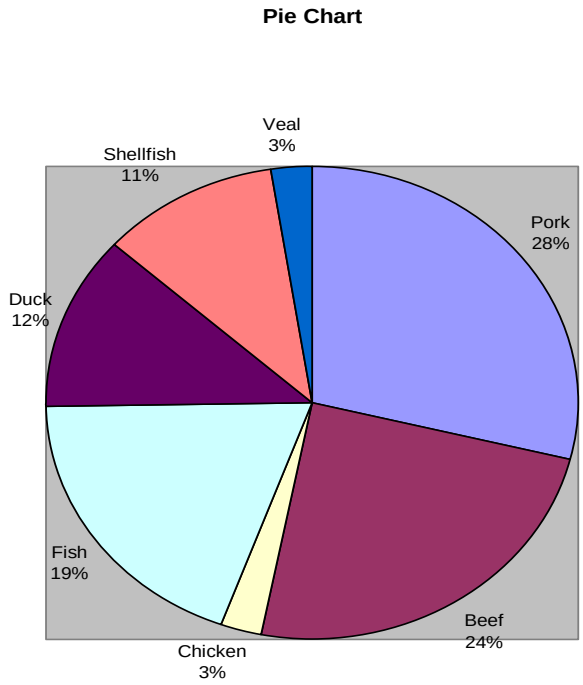
(e) "Paid search" constitutes the largest category on US online ad spending at 43%. Excluding the generic keyword "Cell phones", searches using the keywords "Brand 4" and "Brand 5" make up majority of the search for Cell phones on specific brands.

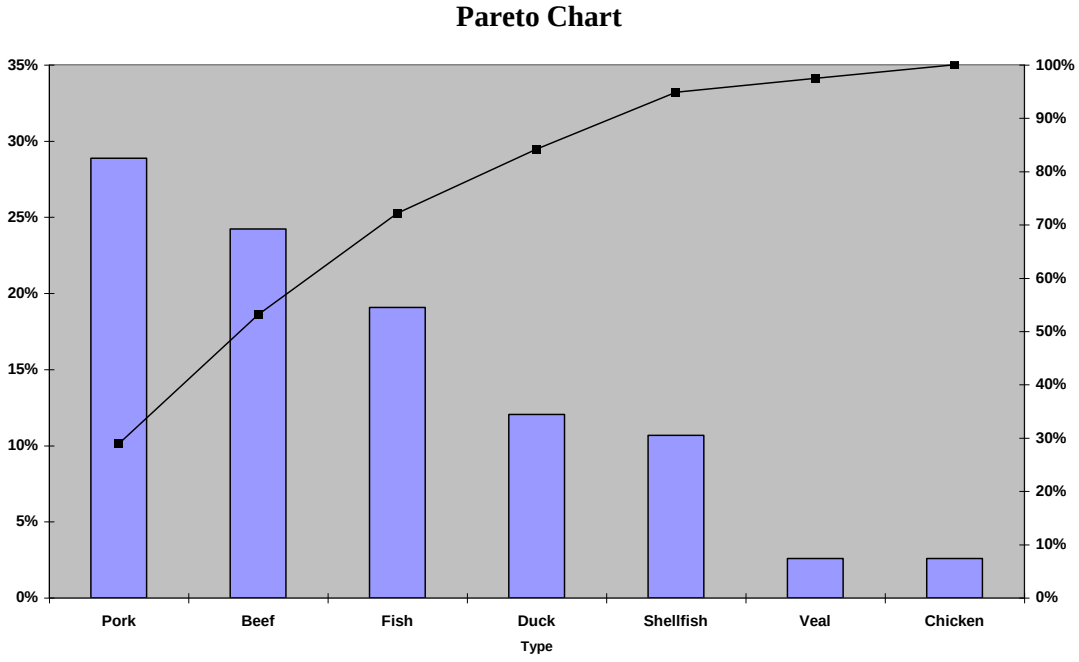
2.69 (a)

(b)

Type of Entrée	Number Served	%
Pork	168	29
Beef	141	24
Chicken	15	3
Fish	111	19
Duck	70	12
Shellfish	62	11
Veal	15	3
	582	**100**







(c) The Pareto diagram has the advantage of offering the cumulative percentage view of the categories and, hence, enables the viewer to separate the "vital few" from the "trivial many".

(d) Pork and Beef account for more than 50% of all entrees ordered by weekend patrons of a continental restaurant. When Fish is included, better than 70% of the entrees are accounted for.

2.70 (a)

	Age group		
Dessert Ordered	20–40	41–60	
Yes	20%	21%	21%
No	80%	79%	79%
Total	100%	100%	100%

	Pork Entrée		
Dessert Ordered	**Yes**	**No**	
Yes	39%	18%	25%
No	61%	82%	75%
Total	100%	100%	100%

	Age group		
Dessert Ordered	**20–40**	**41–60**	**Total**
Yes	51%	49%	100%
No	51%	49%	100%
	51%	49%	100%

	Pork Entrée		
Dessert Ordered	**Yes**	**No**	**Total**
Yes	52%	48%	100%
No	27%	73%	100%
Total	33%	67%	100%

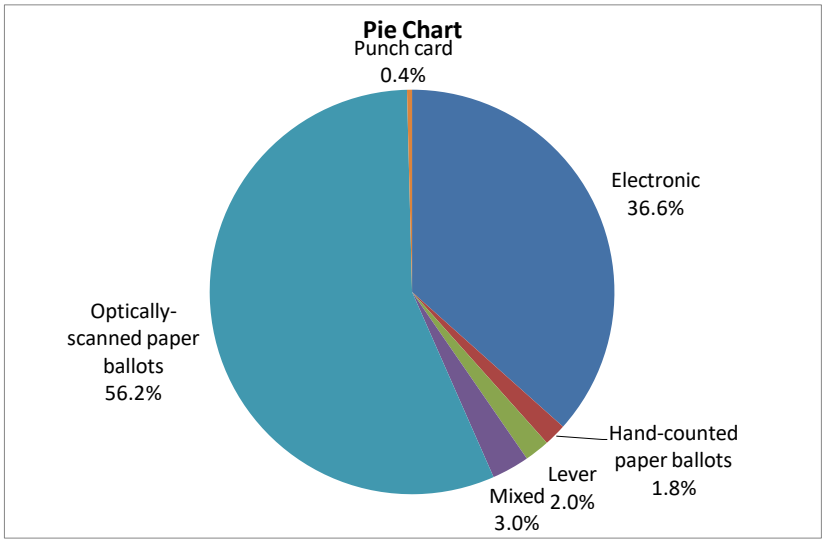
	Age group		
Dessert Ordered	**20–40**	**41–60**	**Total**
Yes	11%	10%	21%
No	41%	39%	79%
Total	51%	49%	100%

	Pork Entrée		
Dessert Ordered	**Yes**	**No**	**Total**
Yes	13%	12%	25%
No	20%	55%	75%
Total	33%	67%	100%

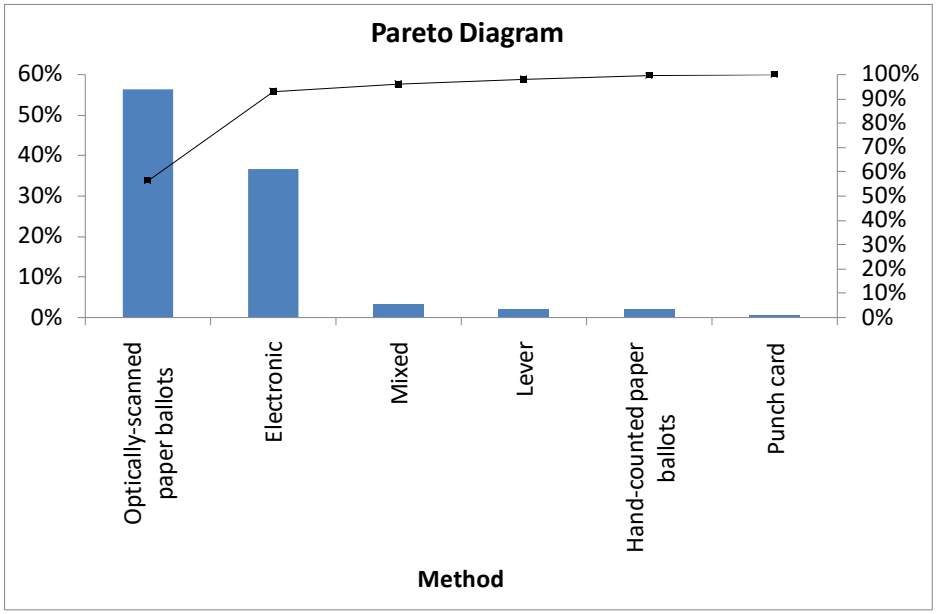
(b) If the owner is interested in finding out the percentage of joint occurrence of age group and ordering of dessert or the percentage of joint occurrence of ordering a pork entrée and a dessert among all patrons, the table of total percentages is most informative. If the owner is interested in the effect of age group on ordering of dessert or the effect of ordering a pork entrée on the ordering of dessert, the table of column percentages will be most informative. Since dessert will usually be ordered after the main entrée and the owner has no direct control over the age group of patrons, the table of row percentages is not very useful here.

(c) There's not a lot of noticeable difference across the age groups in ordering desserts with 20% and 21% ordering in the 2 different age groups. Almost 31% of the patrons ordering a pork entrée ordered dessert compared to 18% of patrons ordering all other entrees. Patrons ordering pork are better than 1.7 times as likely to order dessert as patrons ordering any other entree.

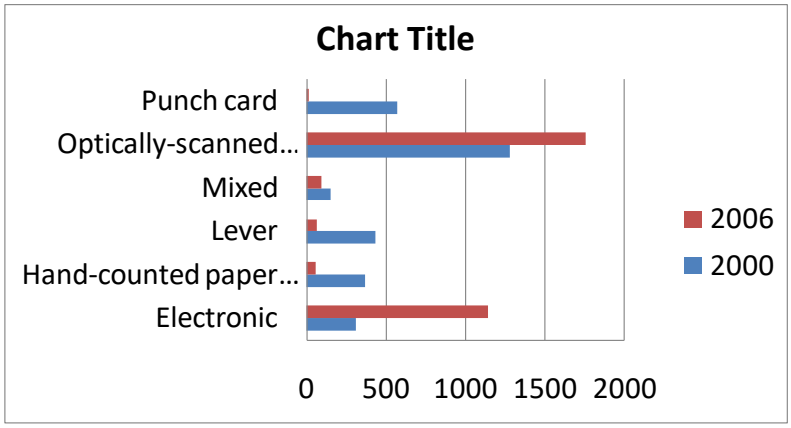
2.71 (a)



2.71 cont. (a)



(b)

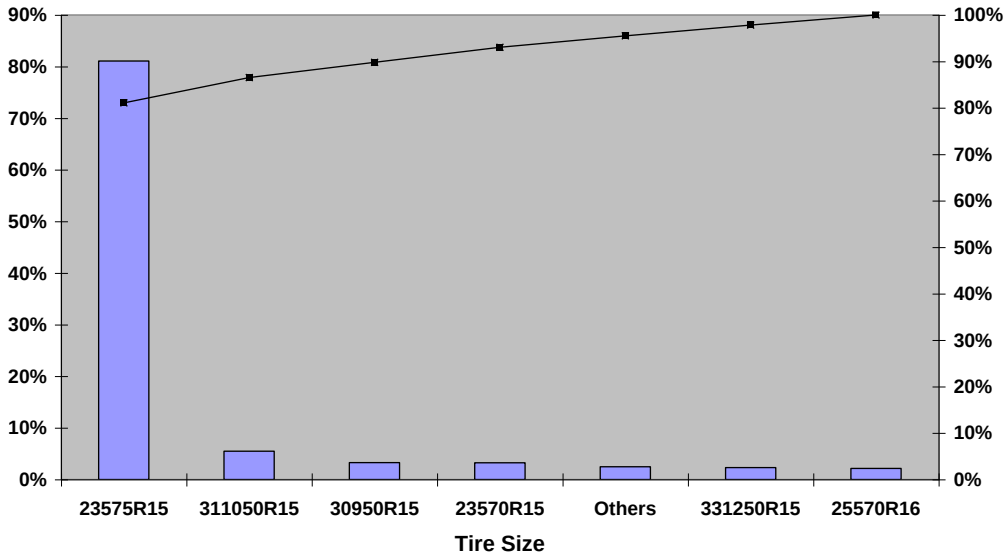


(c) The Pareto diagram has the advantage of offering the cumulative percentage view of the categories and, hence, enables the viewer to separate the "vital few" from the "trivial many". It is better if one is interested in finding out the characteristics of one variable like that in part (a). When performing comparison between two data sets like that in part (b), the side-by-side bar chart is more helpful.

(d) From the Pareto diagram in part (b), one can see that more than 90% of the counties used either the “optically-scanned paper ballots” or “electronic” method in 2006.

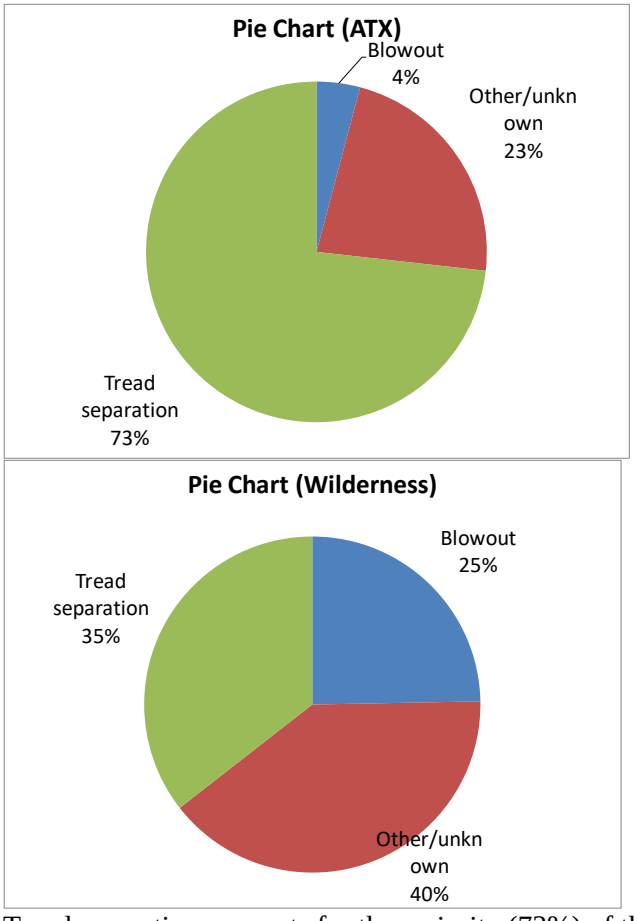
(e) More counties moved from the “punch card”, “mixed”, “level” or “hand-counted paper” methods to using the “optically-scanned paper ballots” or “electronic” methods in 2006 compared to 2000.

2.72 (a)



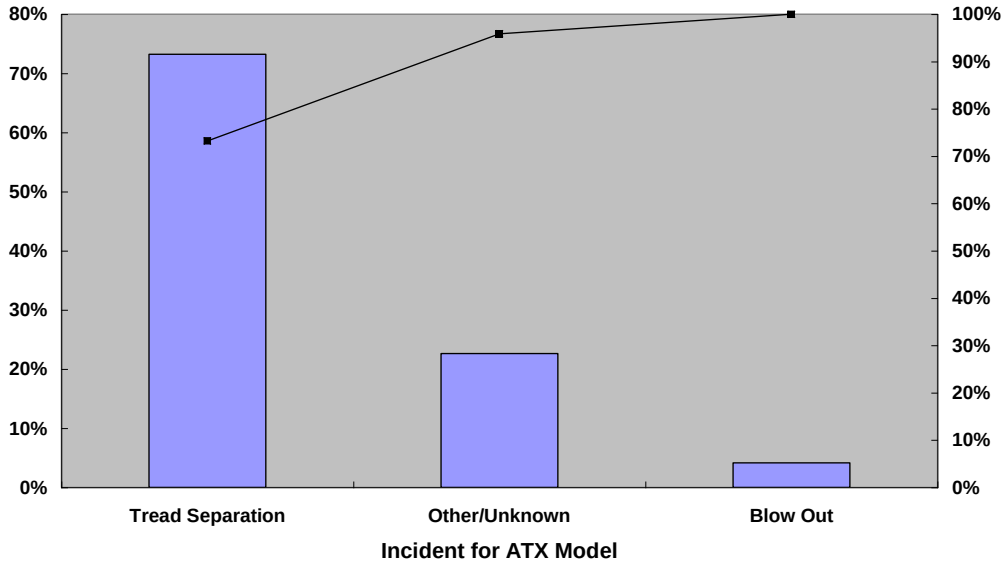
23575R15 accounts for over 80% of the warranty claims.

(b)



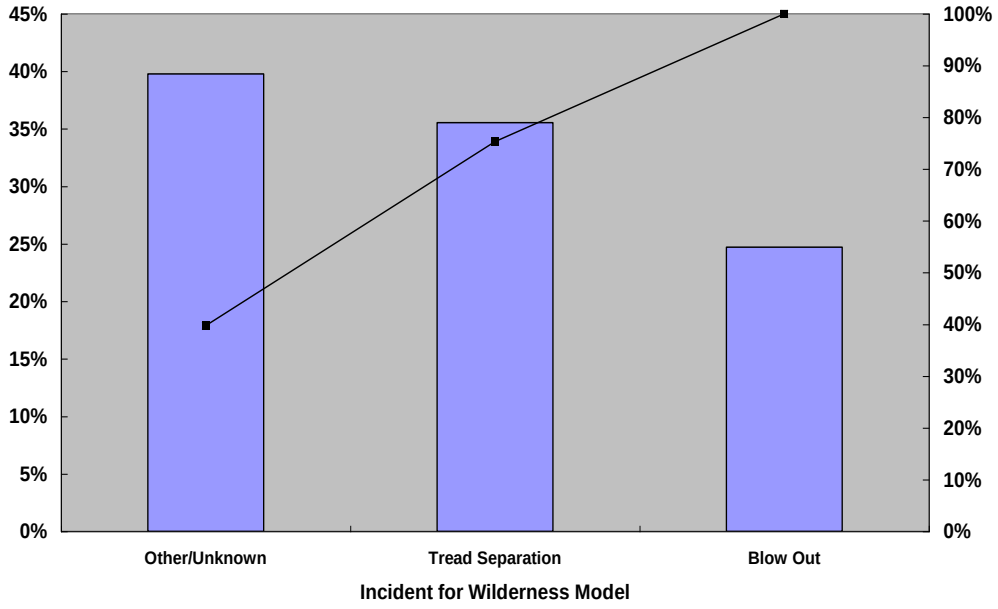
Tread separation accounts for the majority (73%) of the warranty claims for the ATX model but it accounts for only 35% of the warranty claims for the Wilderness model.

2.72 (c)
cont.



Tread separation accounts for more than 70% of the warranty claims among the ATX model.

(d)

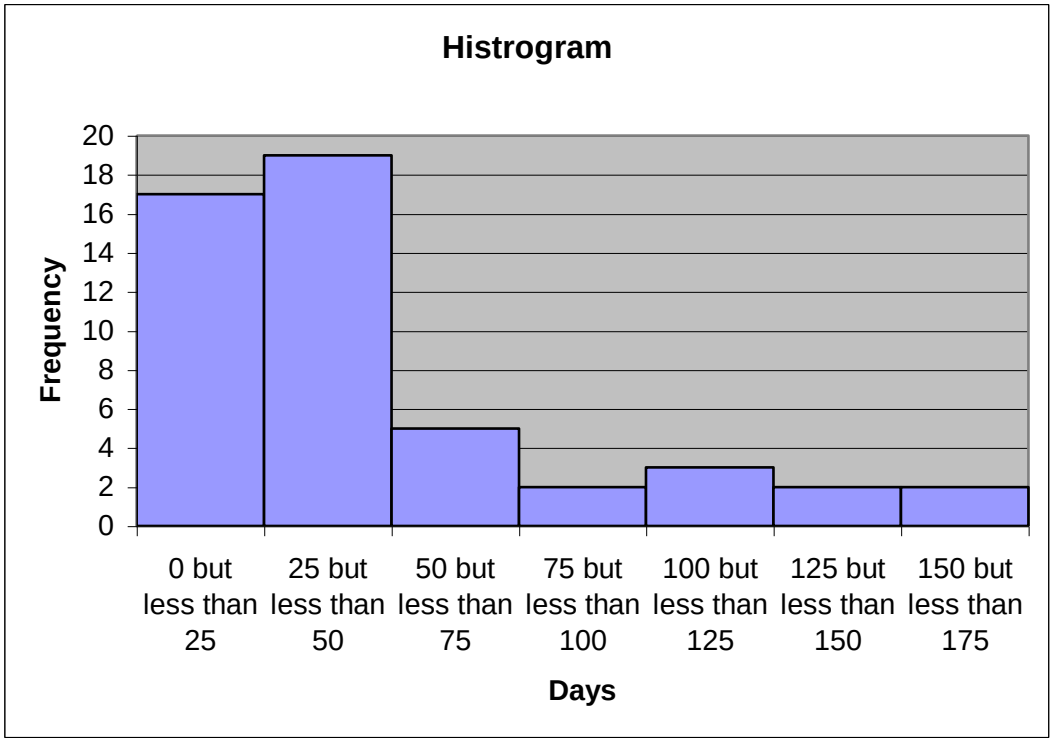


The number of claims is quite evenly distributed among the three incidents. The incident of "other/unknown" accounts for almost 40% of the claims, the incident of "tread separation" accounts for about 35% of the claims while the incident of "blow out" accounts for about 25% of the claims.

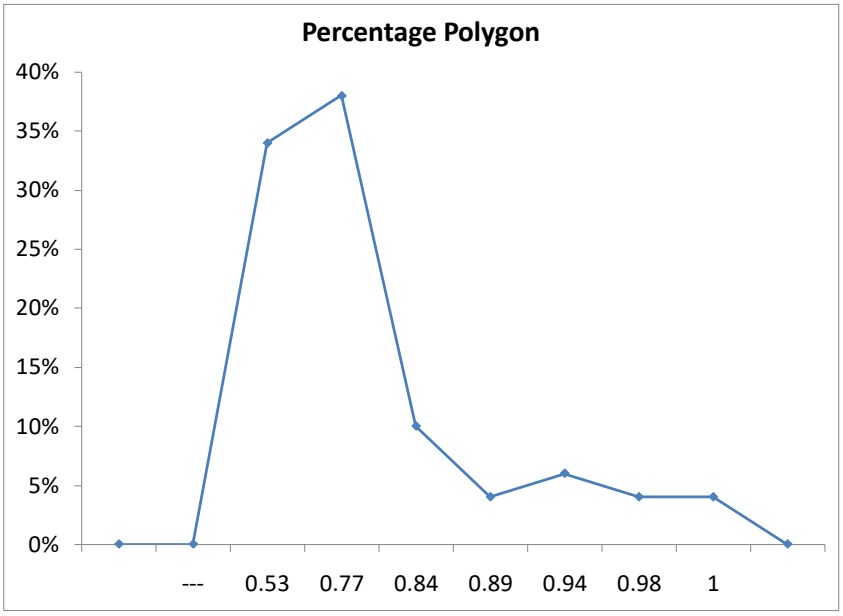
2.73 (a)

Range	*Frequency*	*Percentage*
0 but less than 25	17	34%
25 but less than 50	19	38%
50 but less than 75	5	10%
75 but less than 100	2	4%
100 but less than 125	3	6%
125 but less than 150	2	4%
150 but less than 175	2	4%

(b)

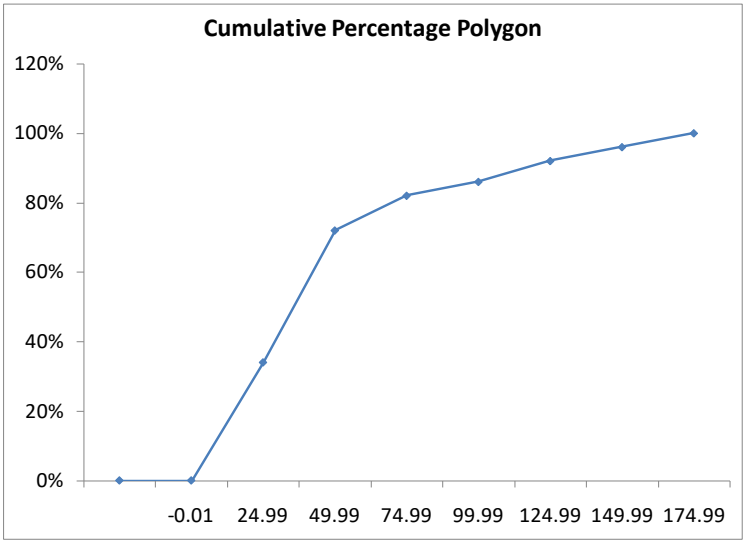


2.73 cont. (b)



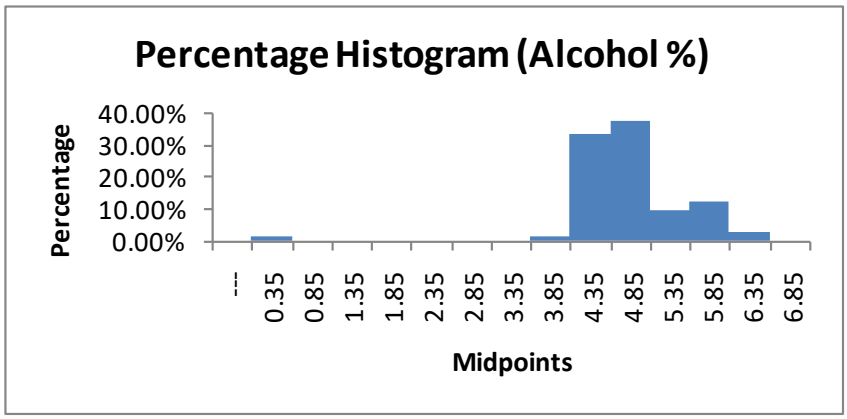
(c)

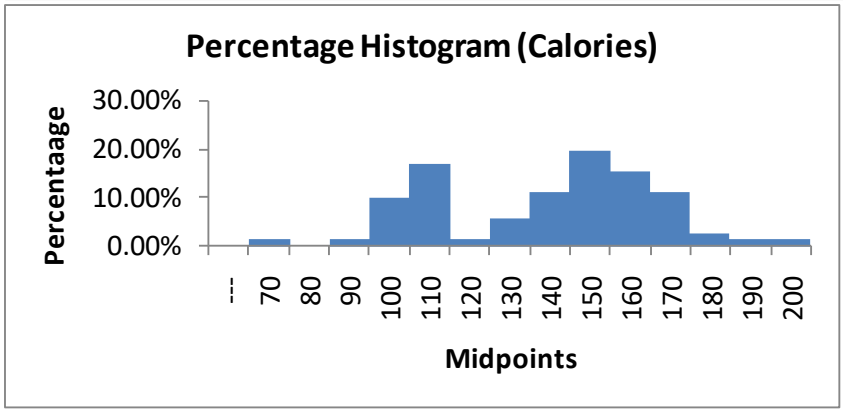
Range	*Cumulative %*
0 but less than 25	34%
25 but less than 50	72%
50 but less than 75	82%
75 but less than 100	86%
100 but less than 125	92%
125 but less than 150	96%
150 but less than 175	100%

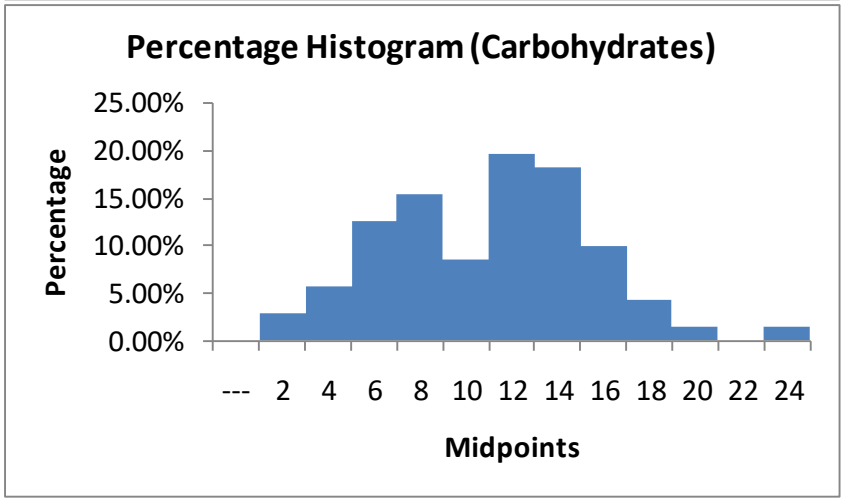


(d) You should tell the president of the company that over half of the complaints are resolved within a month, but point out that some complaints take as long as three or four months to settle.

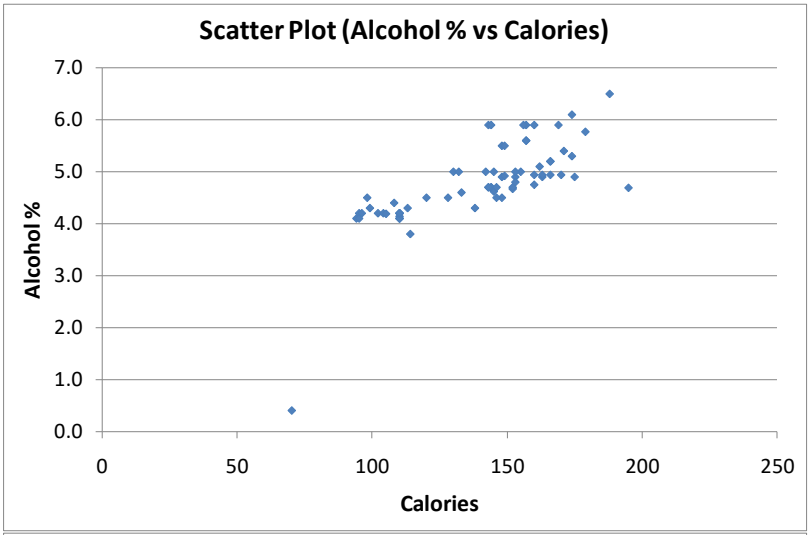
2.74 (a)

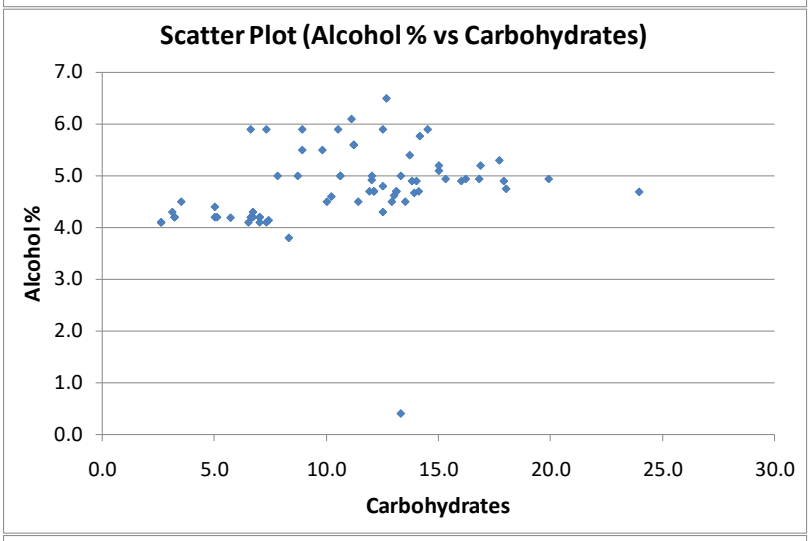


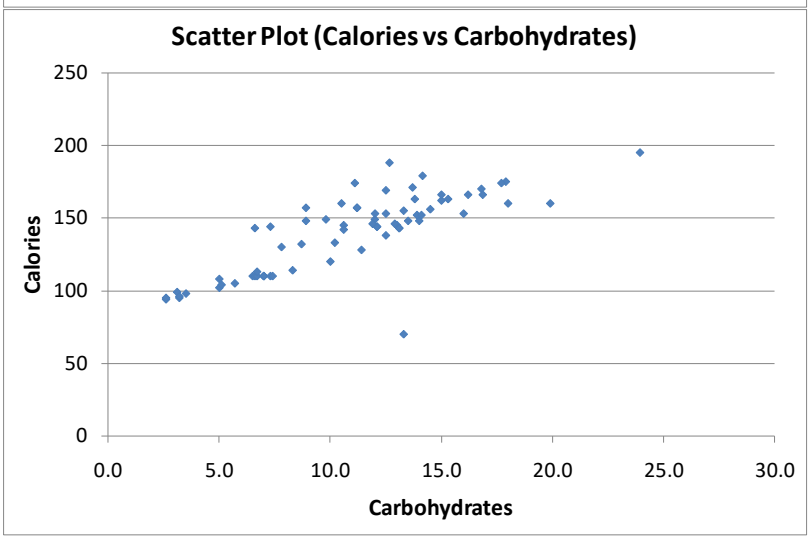




2.74 (b)
cont.







2.74 cont. (c) Majority (about 71%) of the beers have percentage alcohol between 4.1% and 5.1% with one beer (O'Douls) contains only 0.4% alcohol. There are two clusters in calories distribution. About 60% of the beers have calories between 135 and 175 and another cluster of 25% has calories between 95 and 115. The distribution of carbohydrates is slightly right-skewed with Sam Adams Cream Stout having a carbohydrates value of 23.9. There appears to be a positive relationship between percentage alcohol in a beer and its calories content. There is also an obvious positive relationship between calories content and carbohydrates content. Percentage alcohol content and calories content do not appear to be related.

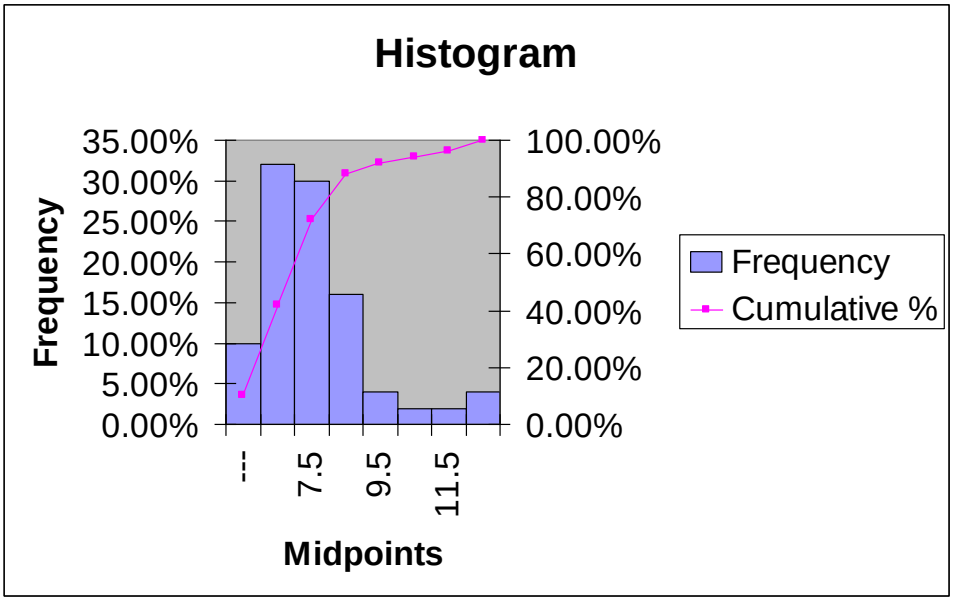
2.75 (a) Ordered array:

5.469	5.644	5.728	5.728	5.981	6.043	6.079	6.084	6.124
6.247	6.308	6.326	6.353	6.388	6.437	6.467	6.474	6.533
6.591	6.751	6.993	7.009	7.080	7.158	7.228	7.298	7.309
7.456	7.484	7.562	7.630	7.649	7.649	7.695	7.701	7.947
8.085	8.248	8.279	8.364	8.564	8.619	8.649	8.673	9.513
9.651	10.437	11.645	12.150	12.885				

(b)

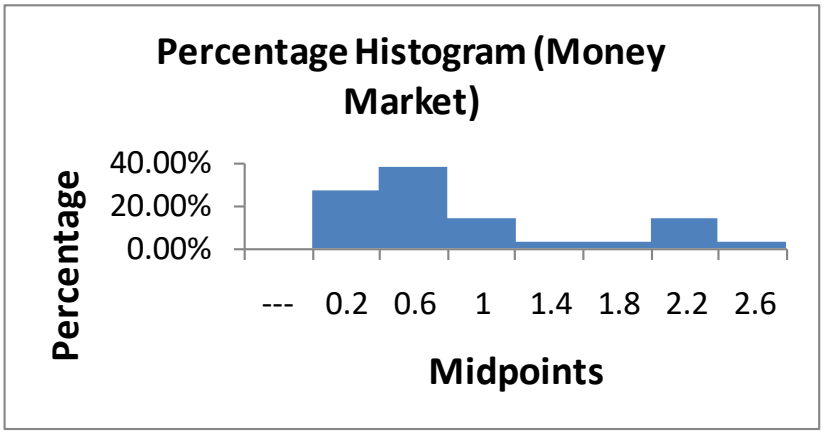
Range	*Frequency*	*Percentage*
5 but less than 6	5	10.00%
6 but less than 7	16	32.00%
7 but less than 8	15	30.00%
8 but less than 9	8	16.00%
9 but less than 10	2	4.00%
10 but less than 11	1	2.00%
11 but less than 12	1	2.00%
12 but less than 13	2	4.00%

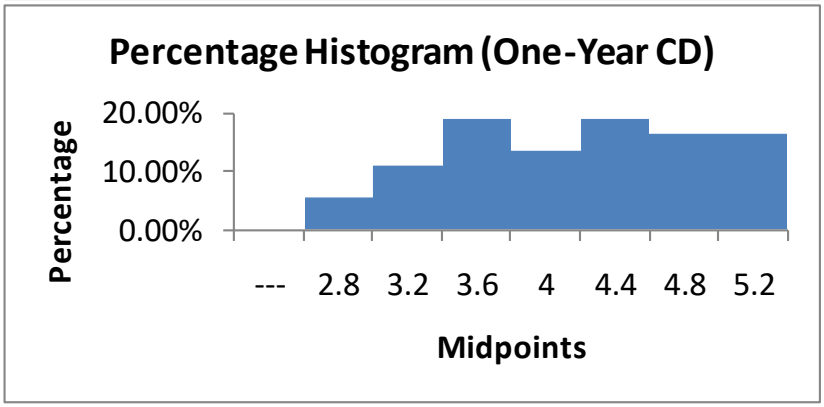
(c)

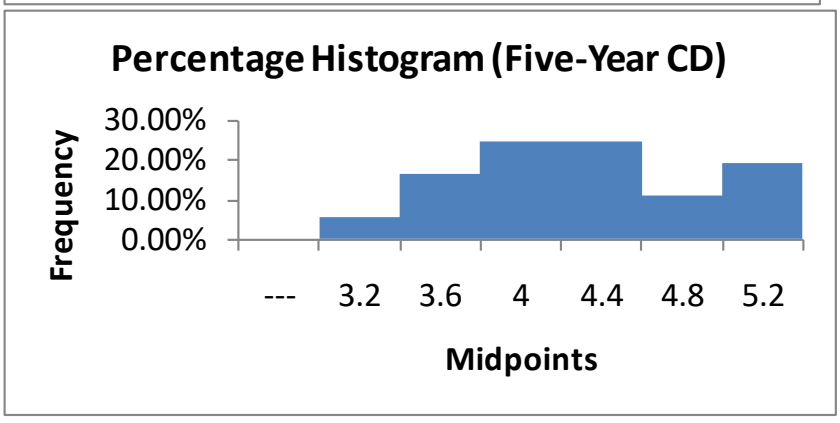


2.75 cont. (d) About 78% of the states spend between 6 and 9 thousand dollars while 10% of the states spend less than 6 thousand dollars and 12% of the states spend more than 9 thousand dollars per capita in 2004.

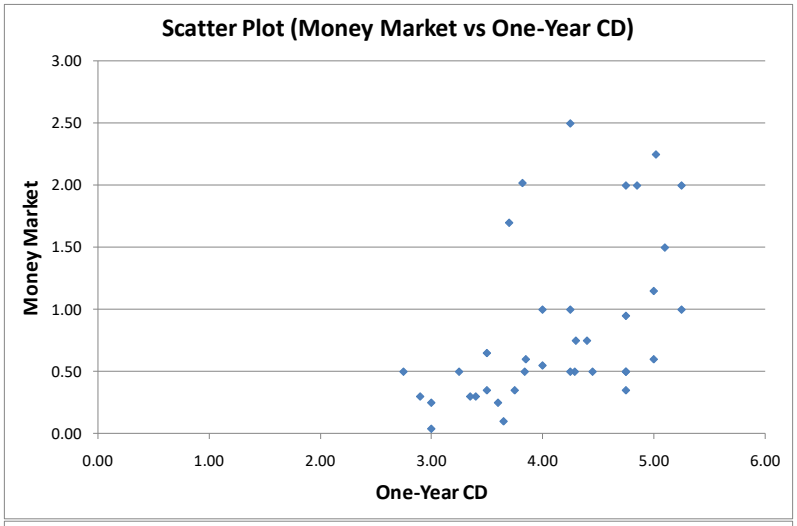
2.76 (a)

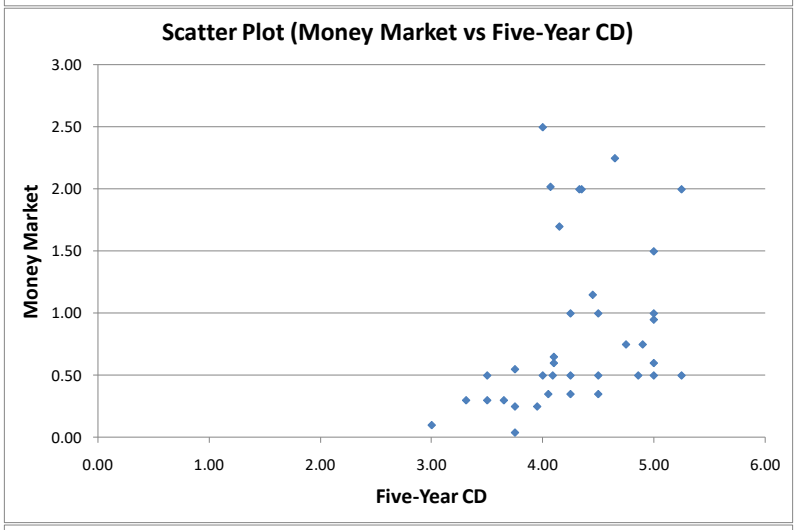


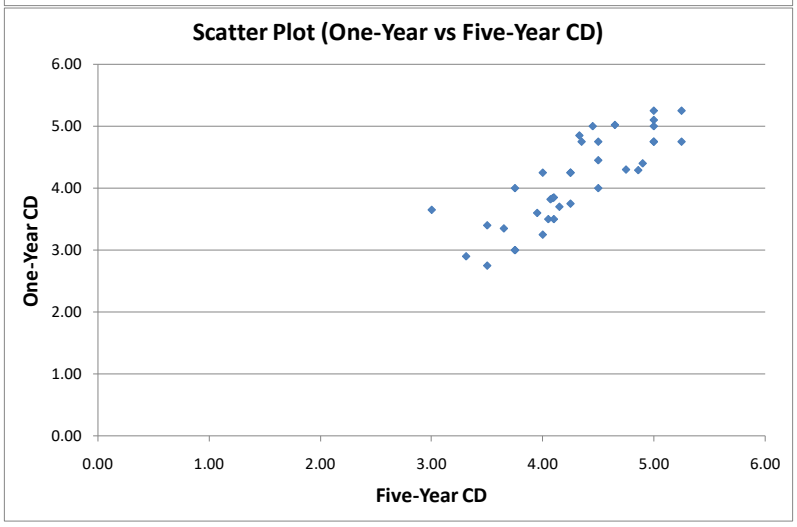




2.76 (b) cont.





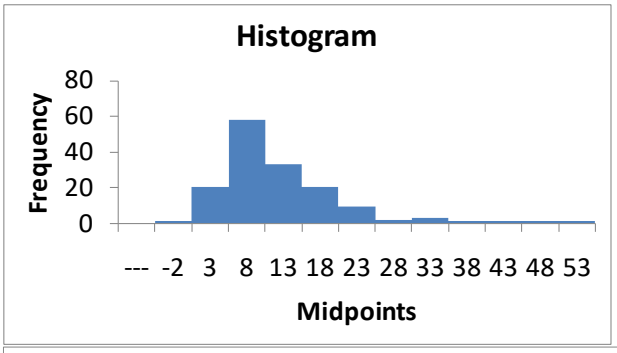


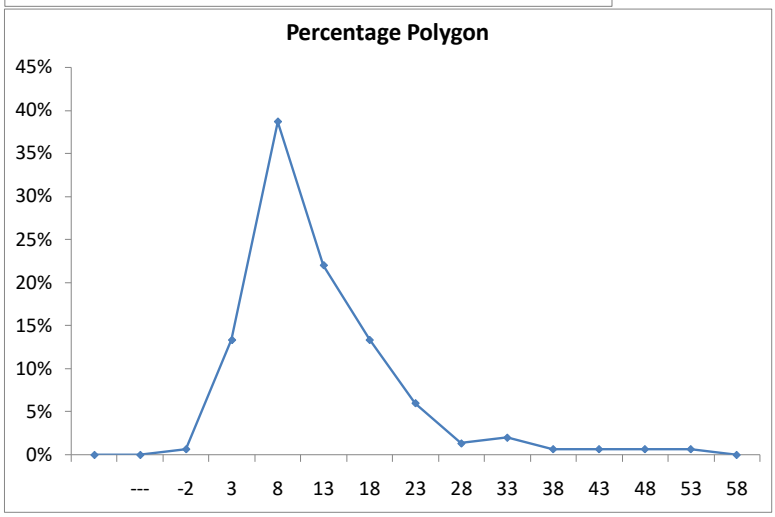
2.76 cont. (c) The distribution of the yields of the money market accounts is right-skewed with almost 80% of them having a return of less than 1.2% while only about 16% of them have yields higher than 2.0%. The distribution of the yields of one-year CD is quite uniform between 3.4% and 5.4% with only about 16% of them having a yield of less than 3.4%. About 65% of the five-year CDs have yields that fall between 3.4% and 4.6%. Only 5% of them have yields that are lower than 3.4% and about 19% of them have yields higher than 5.0%. There appear to be positive relationships between all pairs of yields with one-year CD and five-year CD demonstrating the strongest positive relationship.

2.77 (a), (c)

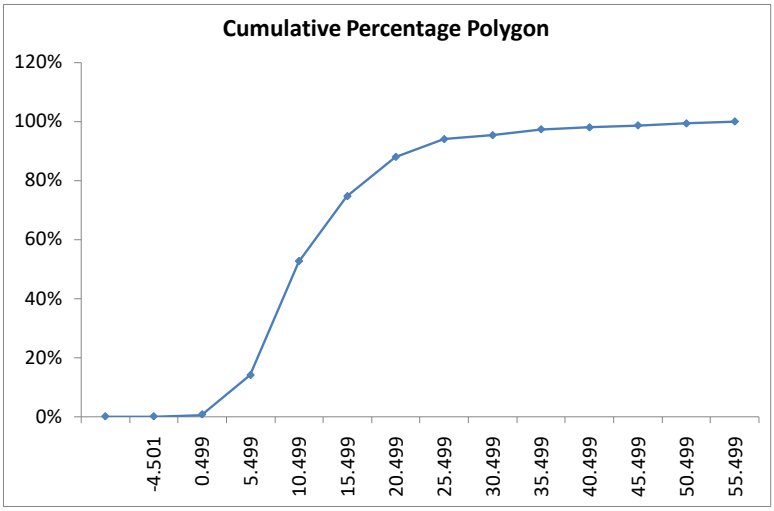
Compensation (in millions$)	*Frequency*	*Percentage*	*Cumulative %*	*Midpts*
Less than -4.5	0	0.00%	0.00%	---
-4.5 but less than 0.5	1	0.67%	0.67%	-2
0.5 but less than 5.5	20	13.33%	14.00%	3
5.5 but less than 10.5	58	38.67%	52.67%	8
10.5 but less than 15.5	33	22.00%	74.67%	13
15.5 but less than 20.5	20	13.33%	88.00%	18
20.5 but less than 25.5	9	6.00%	94.00%	23
25.5 but less than 30.5	2	1.33%	95.33%	28
30.5 but less than 35.5	3	2.00%	97.33%	33
35.5 but less than 40.5	1	0.67%	98.00%	38
40.5 but less than 45.5	1	0.67%	98.67%	43
45.5 but less than 50.5	1	0.67%	99.33%	48
50.5 but less than 55.5	1	0.67%	100.00%	53

2.77 cont. (b)





(c)



(d) The total compensation is right-skewed. Slightly more than half (52.67%) of the CEOs have total compensation below $10.5 millions. Only 6% of the CEOs have total compensation above $25.5 millions.

(e) Yes, Berkshire Hathaway is the only company whose CEO has a total compensation below $500,000. Warren E. Buffet, who is one of the wealthiest persons in the US and a famous philanthropher, is its CEO.

2.78 (a)

Frequencies (Boston)

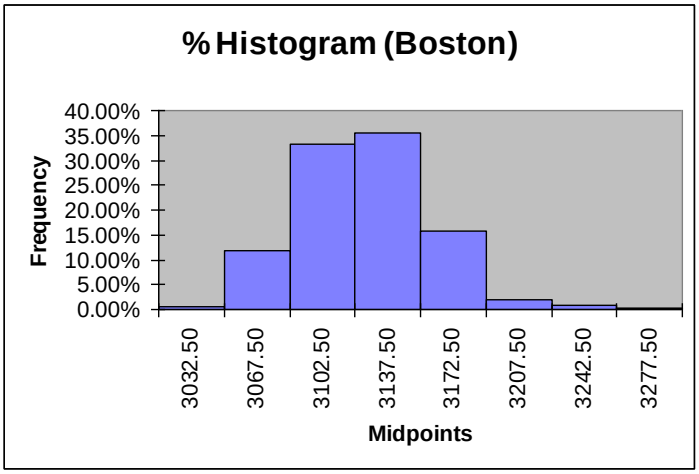
Weight (Boston)	*Frequency*	*Percentage*
3015 but less than 3050	2	0.54%
3050 but less than 3085	44	11.96%
3085 but less than 3120	122	33.15%
3120 but less than 3155	131	35.60%
3155 but less than 3190	58	15.76%
3190 but less than 3225	7	1.90%
3225 but less than 3260	3	0.82%
3260 but less than 3295	1	0.27%

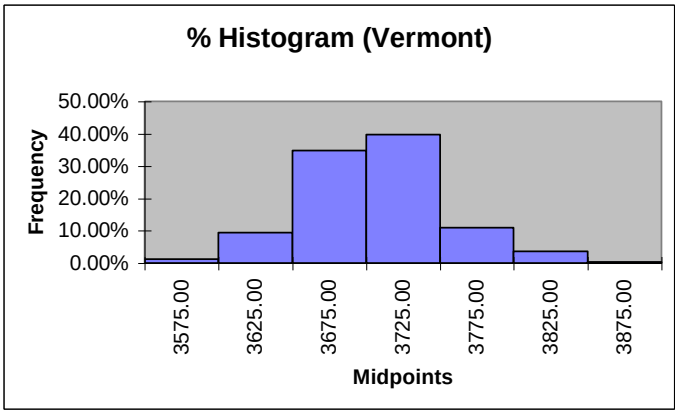
(b)

Frequencies (Vermont)

Weight (Vermont)	*Frequency*	*Percentage*
3550 but less than 3600	4	1.21%
3600 but less than 3650	31	9.39%
3650 but less than 3700	115	34.85%
3700 but less than 3750	131	39.70%
3750 but less than 3800	36	10.91%
3800 but less than 3850	12	3.64%
3850 but less than 3900	1	0.30%

(c)





(d) 0.54% of the "Boston" shingles pallets are underweight while 0.27% are overweight. 1.21% of the "Vermont" shingles pallets are underweight while 3.94% are overweight.

2.79 (a), (c)

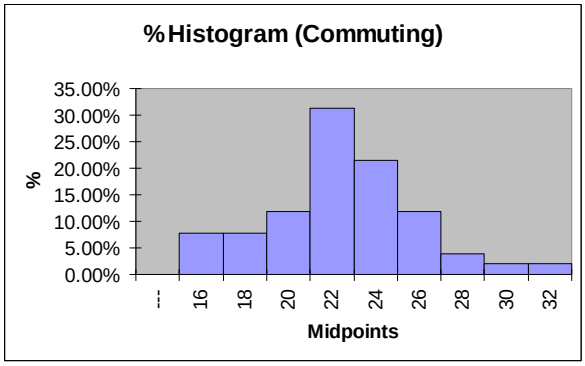
Commuting	*Frequency*	*Percentage*	*Cumulative %*
More than 15 and up to 17	4	7.84%	7.84%
More than 17 and up to 19	4	7.84%	15.69%
More than 19 and up to 21	6	11.76%	27.45%
More than 21 and up to 23	16	31.37%	58.82%
More than 23 and up to 25	11	21.57%	80.39%
More than 25 and up to 27	6	11.76%	92.16%
More than 27 and up to 29	2	3.92%	96.08%
More than 29 and up to 31	1	1.96%	98.04%
More than 31 and up to 33	1	1.96%	100.00%

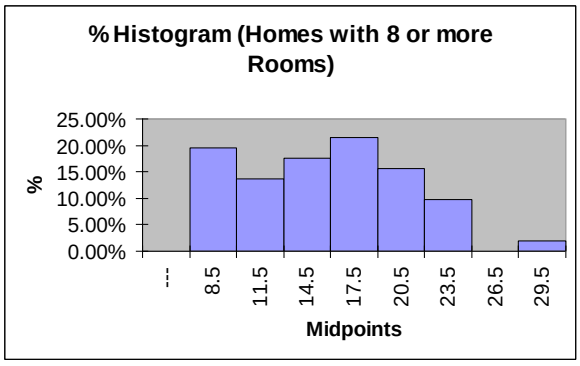
Homes 8 Rooms	*Frequency*	*Percentage*	*Cumulative %*
More than 7 and up to 10	10	19.61%	19.61%
More than 10 and up to 13	7	13.73%	33.33%
More than 13 and up to 16	9	17.65%	50.98%
More than 16 and up to 19	11	21.57%	72.55%
More than 19 and up to 22	8	15.69%	88.24%
More than 22 and up to 25	5	9.80%	98.04%
More than 25 and up to 28	0	0.00%	98.04%
More than 28 and up to 31	1	1.96%	100.00%

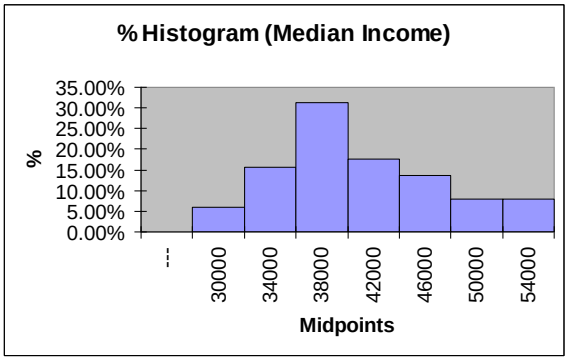
Median Income	*Frequency*	*Percentage*	*Cumulative %*
More than 28000 and up to 32000	3	5.88%	5.88%
More than 32000 and up to 36000	8	15.69%	21.57%
More than 36000 and up to 40000	16	31.37%	52.94%
More than 40000 and up to 44000	9	17.65%	70.59%
More than 44000 and up to 48000	7	13.73%	84.31%
More than 48000 and up to 52000	4	7.84%	92.16%
More than 52000 and up to 56000	4	7.84%	100.00%

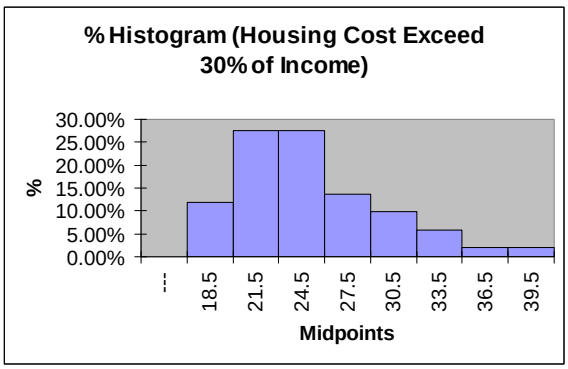
Housing more than 30% Income	*Frequency*	*Percentage*	*Cumulative %*
More than 17 and up to 20	6	11.76%	11.76%
More than 20 and up to 23	14	27.45%	39.22%
More than 23 and up to 26	14	27.45%	66.67%
More than 26 and up to 29	7	13.73%	80.39%
More than 29 and up to 32	5	9.80%	90.20%
More than 32 and up to 35	3	5.88%	96.08%
More than 35 and up to 38	1	1.96%	98.04%
More than 38 and up to 41	1	1.96%	100.00%

2.79 (b)
cont.

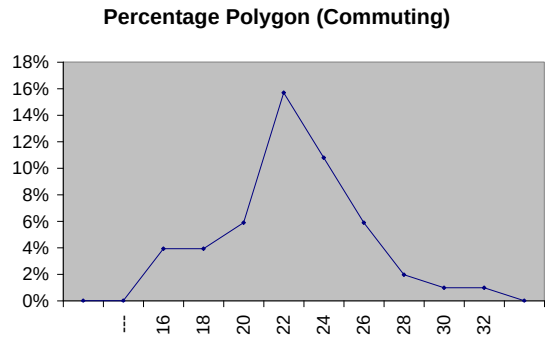


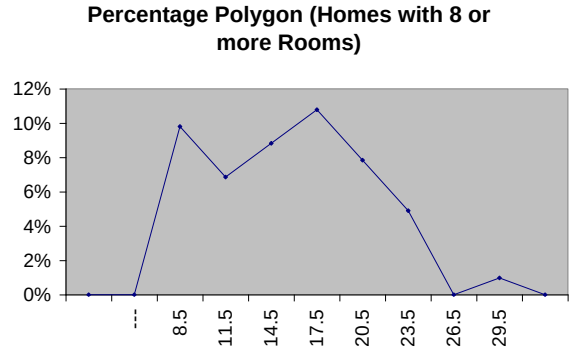


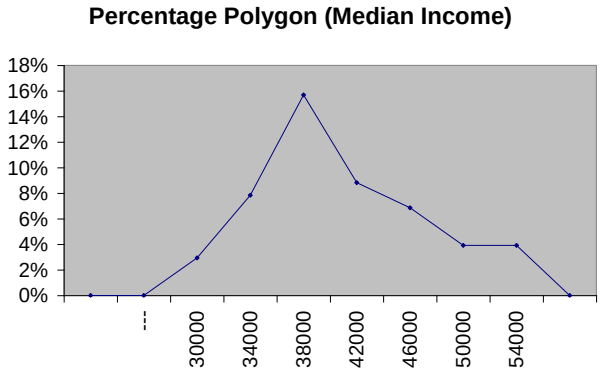


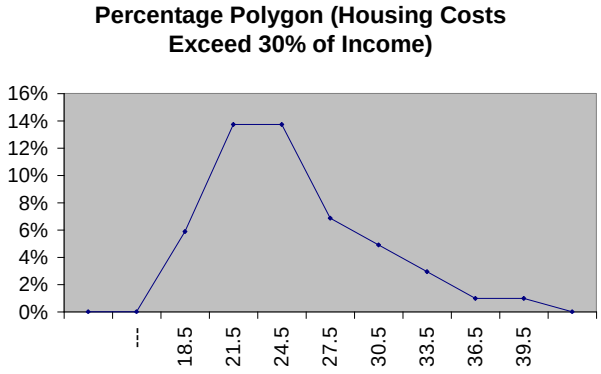


2.79 (b) cont.

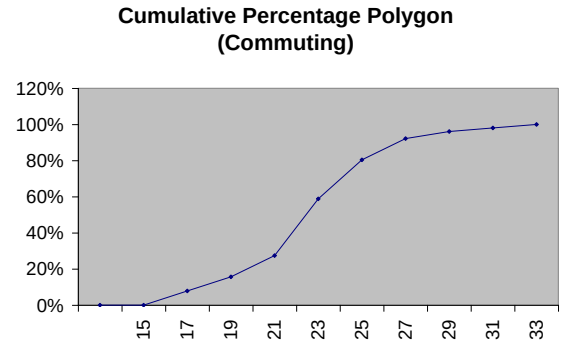


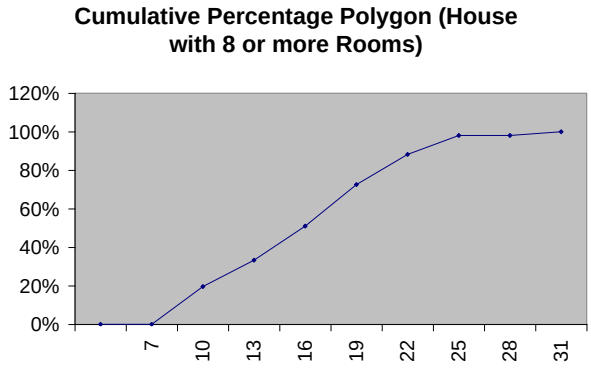


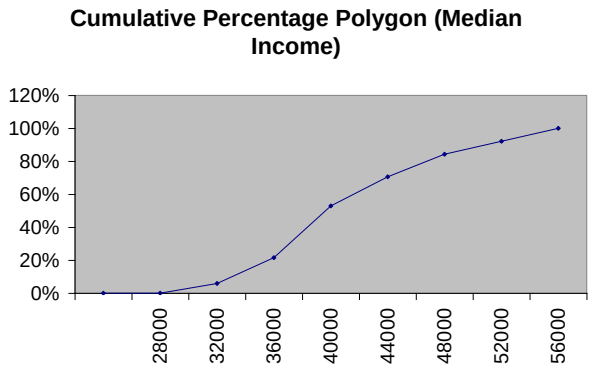


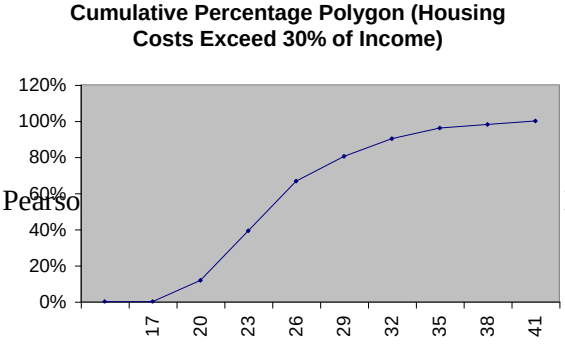


2.79 (c)
cont.



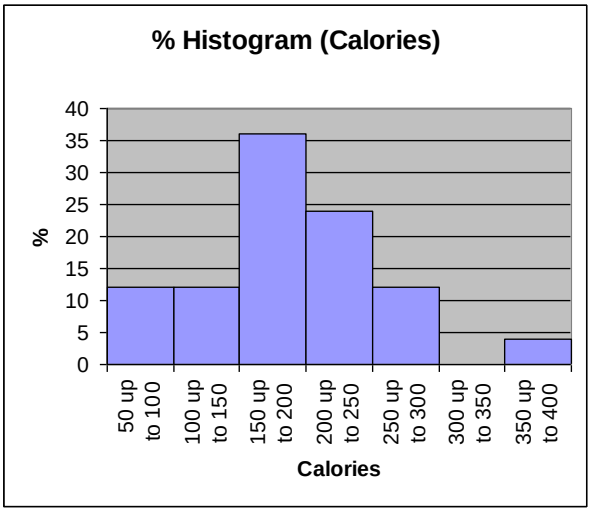






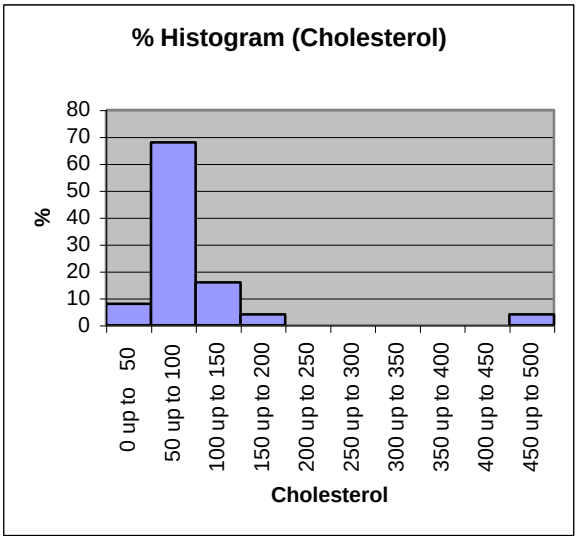
2.80 (a)

Calories	Frequency	Percentage	Percentage Less Than
50 up to 100	3	12%	12%
100 up to 150	3	12	24
150 up to 200	9	36	60
200 up to 250	6	24	84
250 up to 300	3	12	96
300 up to 350	0	0	96
350 up to 400	1	4	100



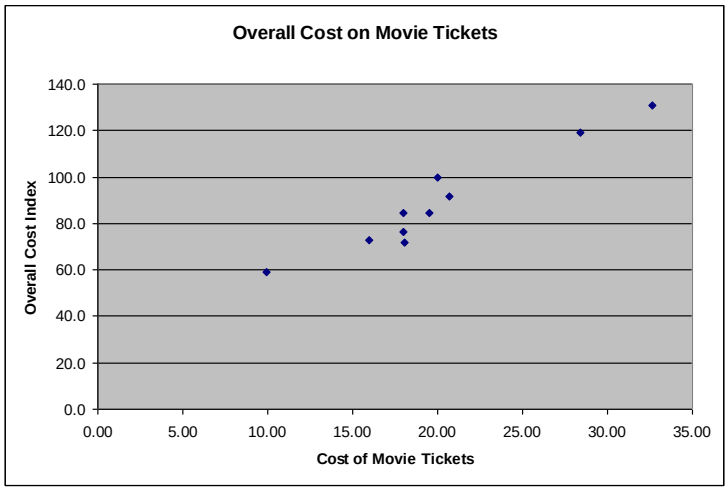
(b)

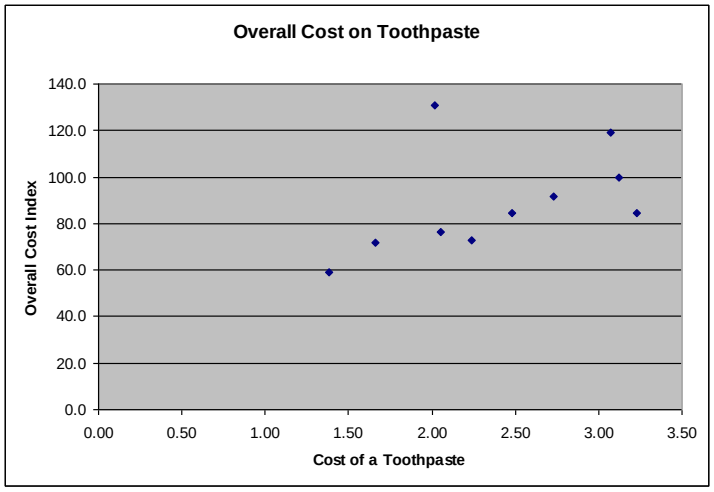
Cholesterol	Frequency	Percentage	Percentage Less Than
0 up to 50	2	8	8%
50 up to 100	17	68	76
100 up to 150	4	16	92
150 up to 200	1	4	96
200 up to 250	0	0	96
250 up to 300	0	0	96
300 up to 350	0	0	96
350 up to 400	0	0	96
400 up to 450	0	0	96
450 up to 500	1	4	100

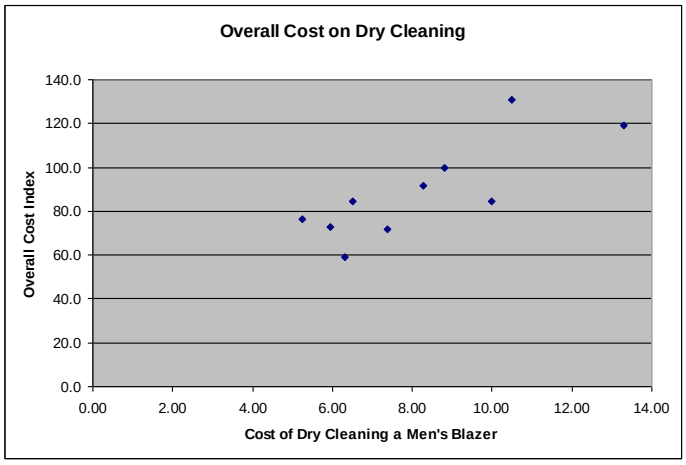


2.80 cont. (c) The sampled fresh red meats, poultry, and fish vary from 98 to 397 calories per serving with the highest concentration between 150 to 200 calories. One protein source, spareribs with 397 calories, was over 100 calories beyond the next highest caloric food. Spareribs and fried liver are both very different from other foods sampled, the former on calories and the latter on cholesterol content.

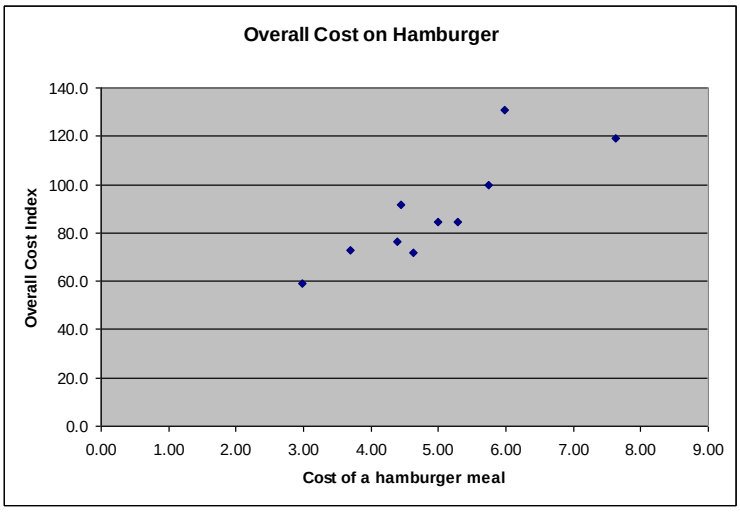
2.81 (a)

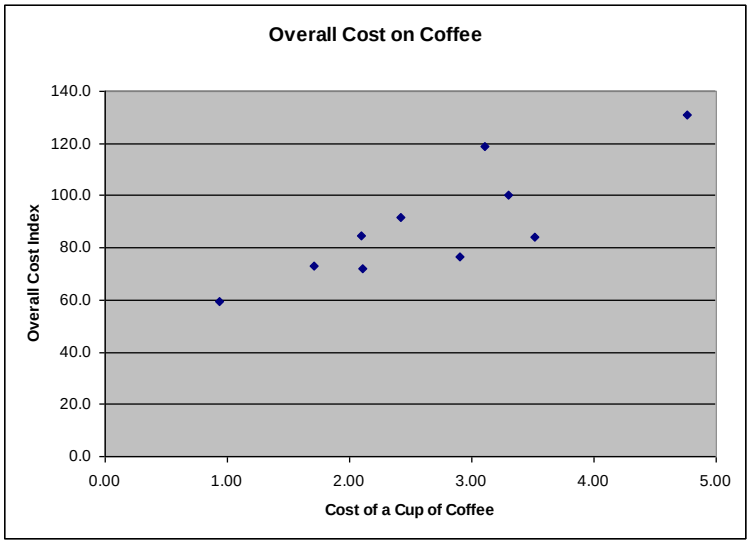


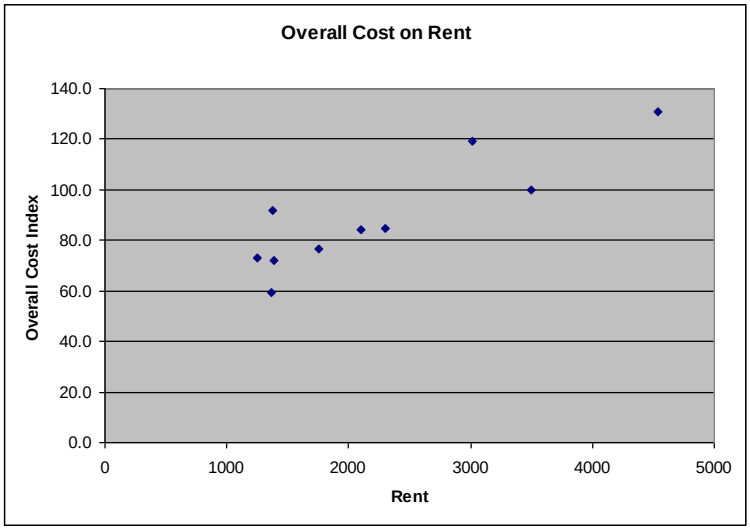




2.81 cont. (a)

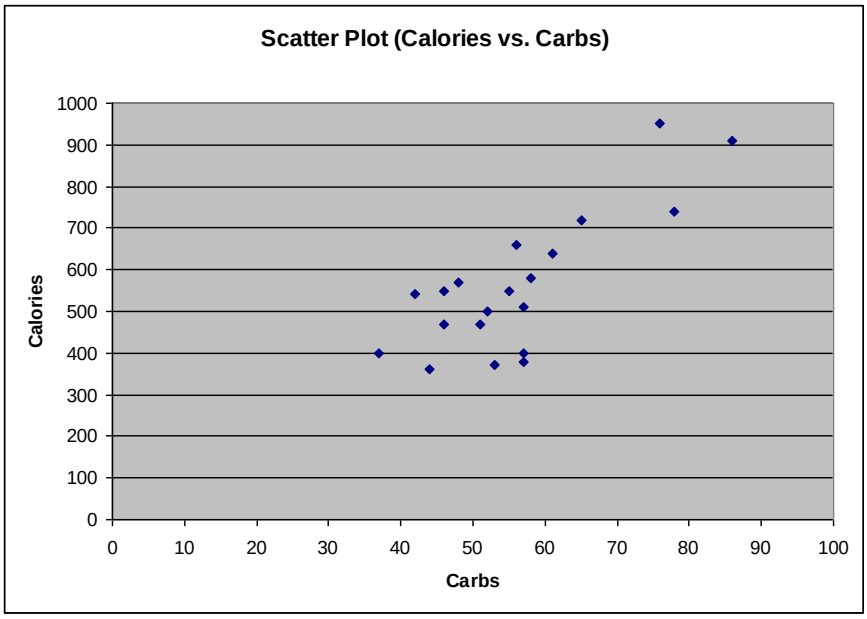




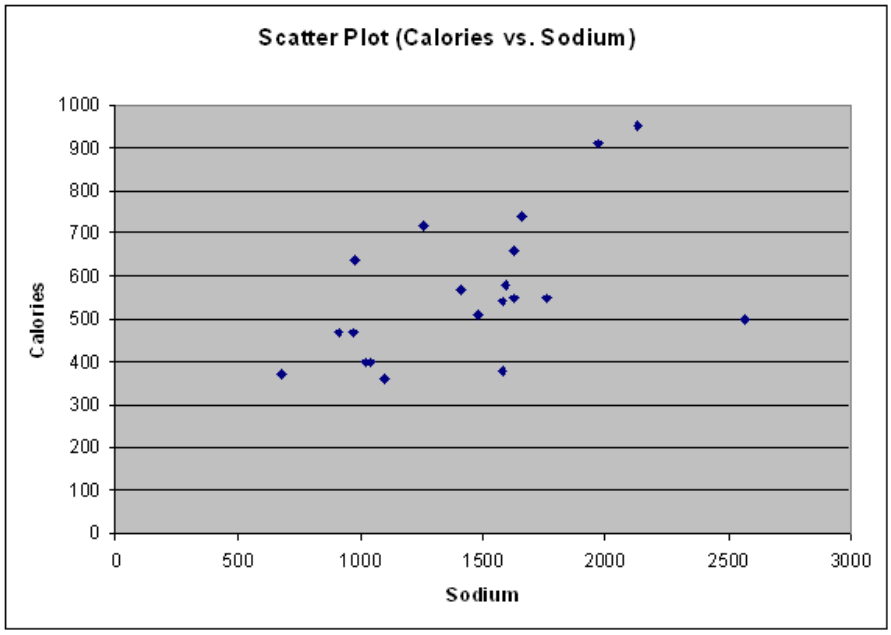


(b) There is a positive relationship between the overall cost index and each of these variables.

2.82 (a)

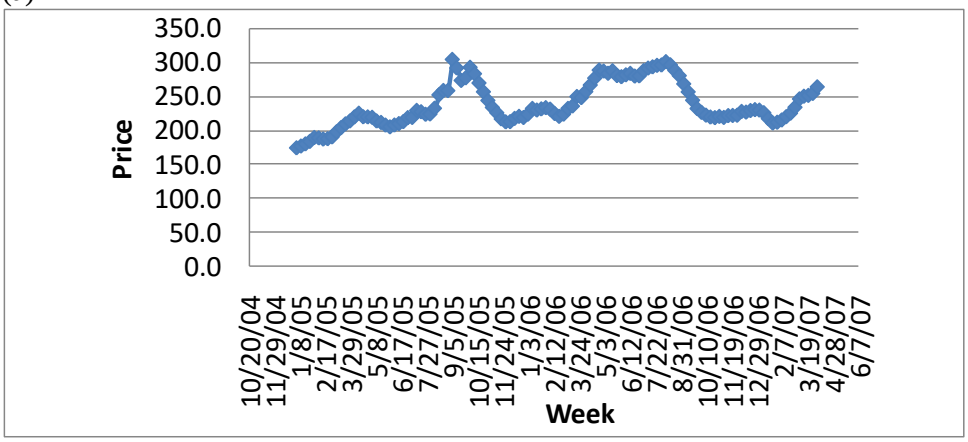


(b)



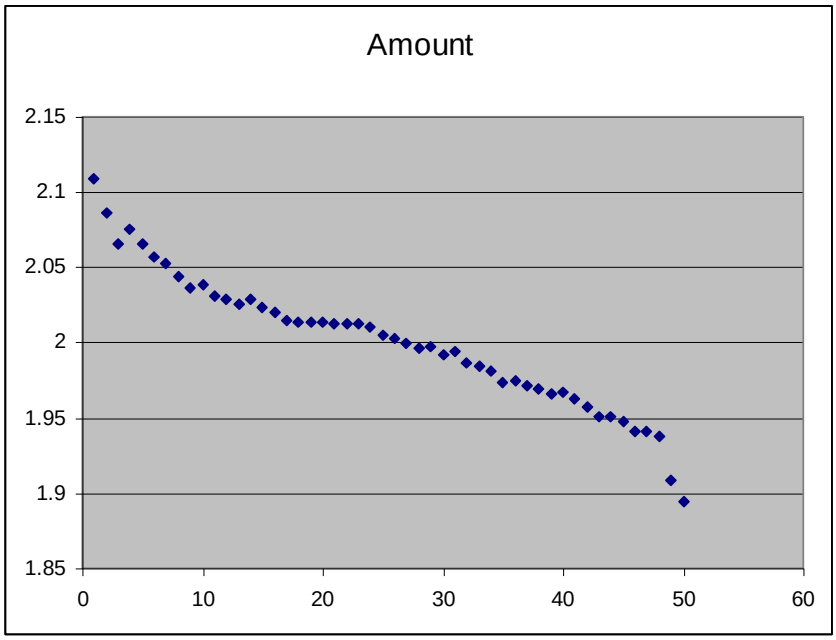
(c) By comparing the scatter plots in (a), (b) and 2.39 (a), total fat seems to be most closely related to calories because the points in the scatter plot are closer to the imaginary line that passes through the data points.

2.83 (a)



(b) There appear to be some cyclical component in the data.

2.84 (a)

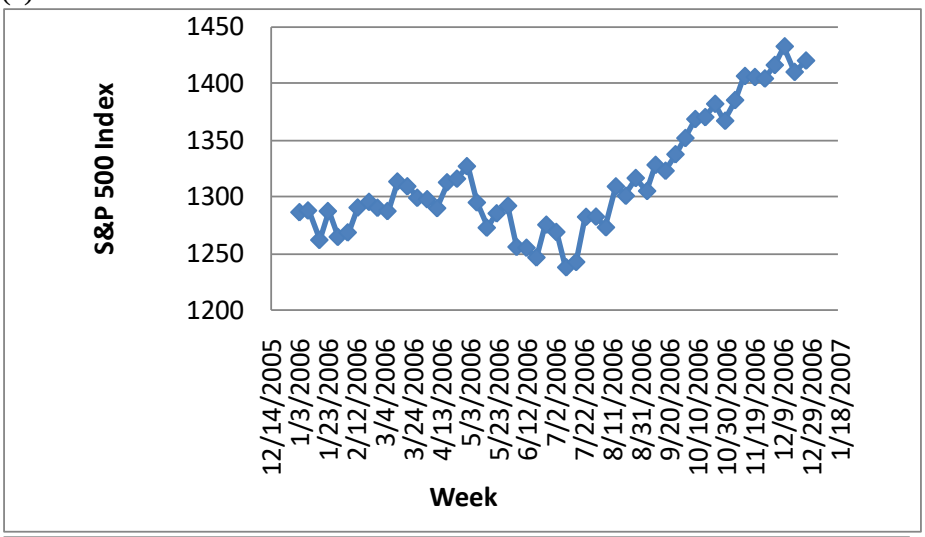


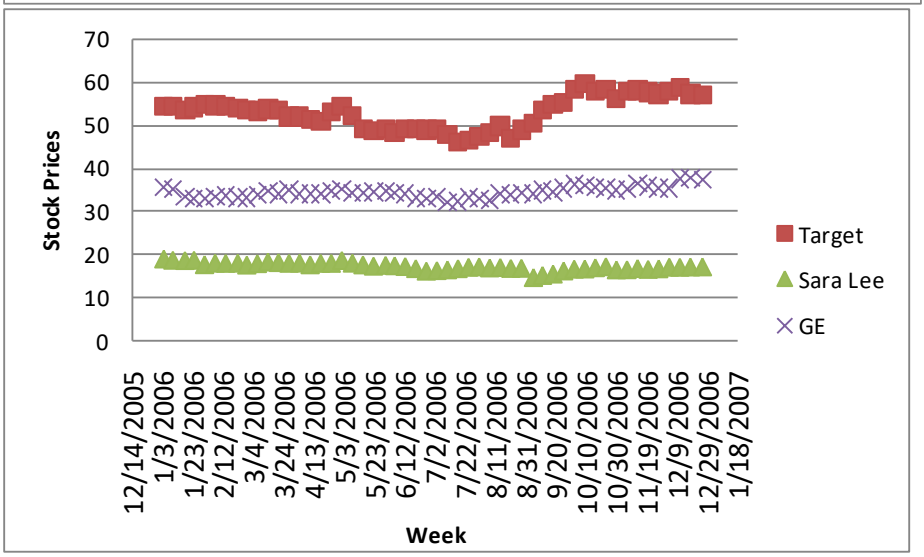
(b) There is a downward trend in the amount filled.

(c) The amount filled in next bottle will most likely be below 1.894 liter.

(d) The scatter plot of the amount of soft drink filled against time reveals the trend of the data while a histogram only provides information on the distribution of the data.

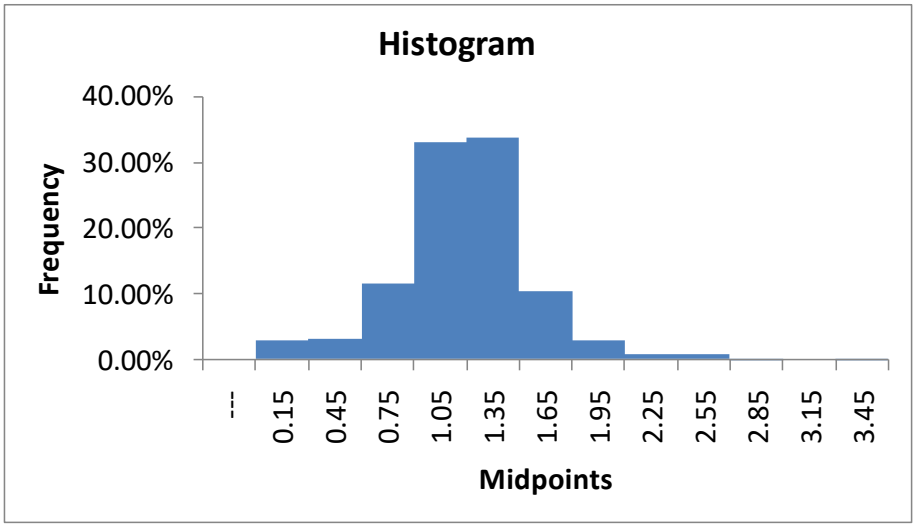
2.85 (a)



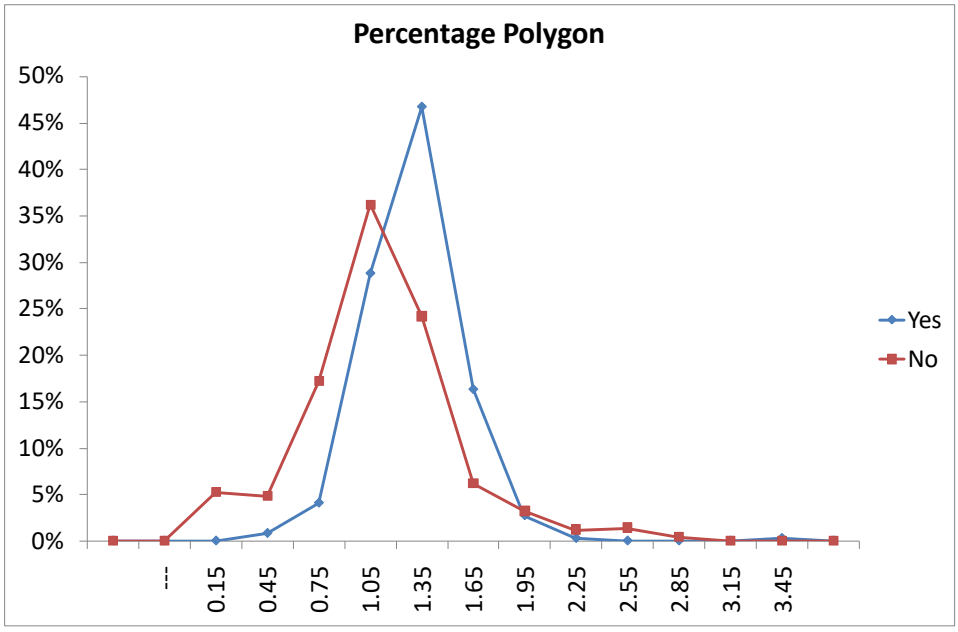


(b) Even though there appears to be cyclical pattern in the S&P index, there was a general upward trend after June 2, 2006. The stock price of Target fluctuated slightly around $50 before August 11, 2006 and then trended upward until it leveled off at around $60 beginning October 2, 2006. The stock price of GE and Sara Lee is quite stable hovering at around $35 and $15 respectively.

2.86 Student answers will vary.
2.87 Student answers will vary.
2.88 Student answers will vary.
2.89 Student answers will vary.
2.90 (a)

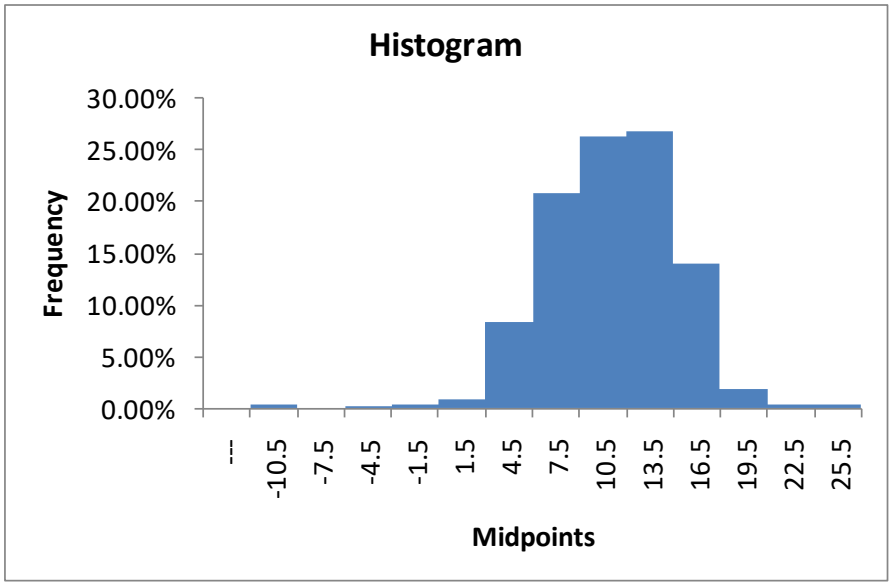


(b)

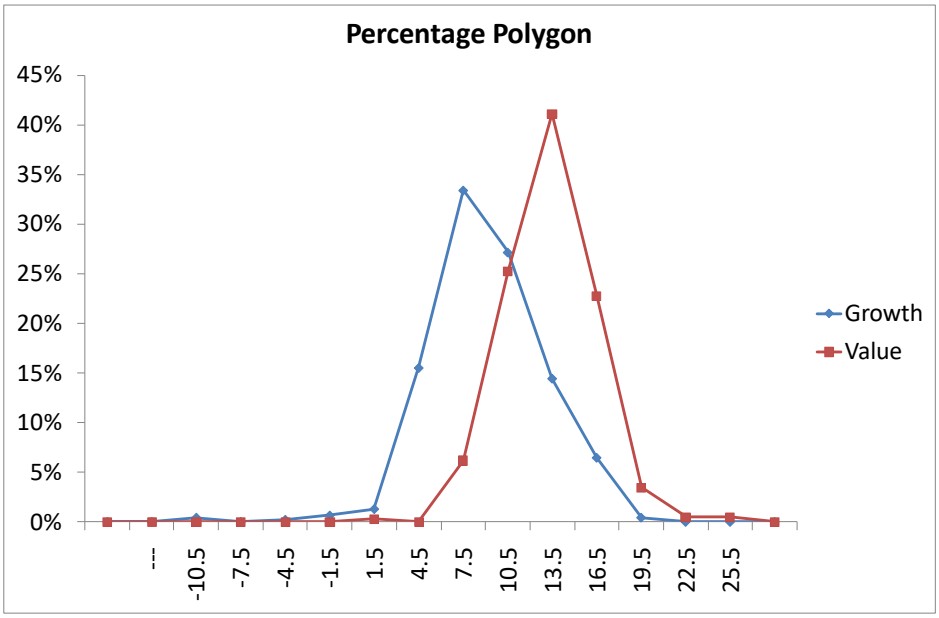


(c) The expense ratio of all mutual funds is scattered around 1.2. Mutual funds with fees have expense ratio that is more symmetrically distributed around 1.35 while mutual funds without fees have expense ratio that is right-skewed with majority of them scattered between 0.75 and 1.35.

2.91 (a)

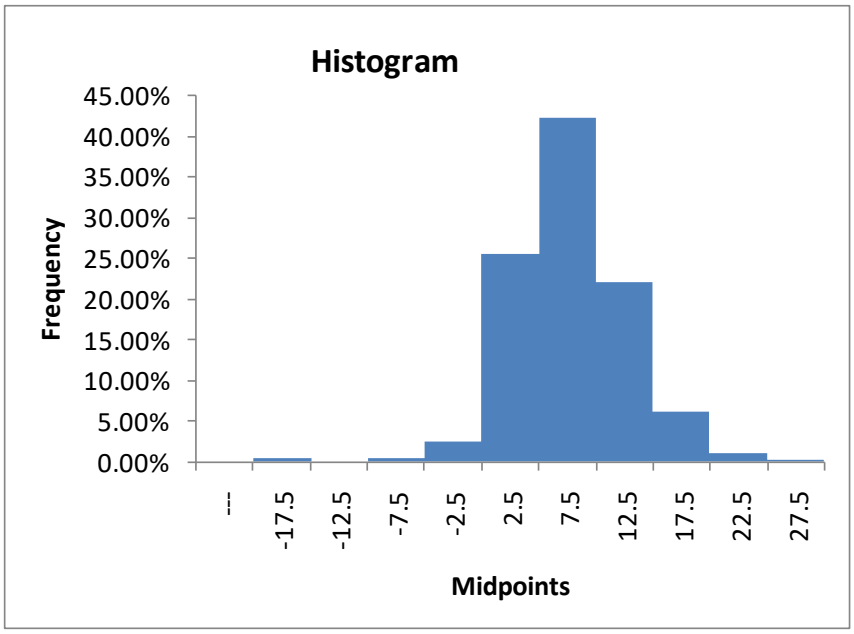


(b)

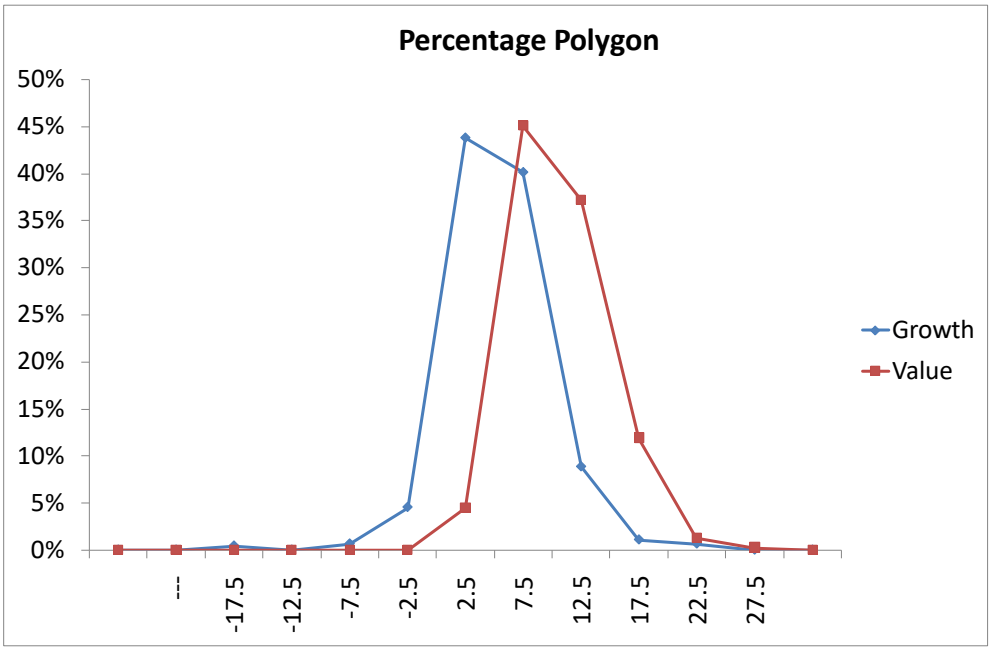


(c) The three-year annualized return of the 868 mutual funds is left-skewed with majority of them (about 96.2%) scattered between 3% and 18%. About 0.69% of the mutual funds have a negative three-year annualized return while about 2.3% of them have return higher than 18%. In general, the value mutual funds have higher three-year annualized return than growth mutual funds. Both types of mutual funds have three-year annualized return skewed to the left.

2.92 (a)

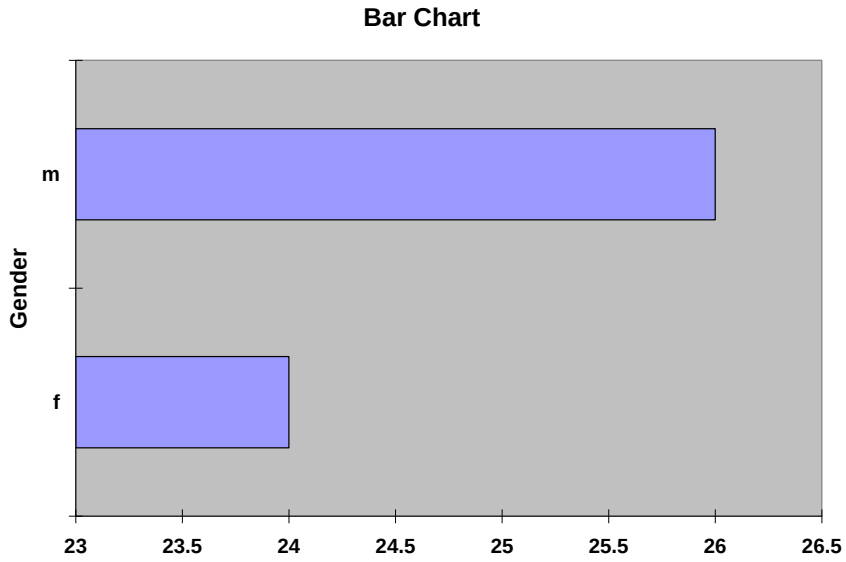


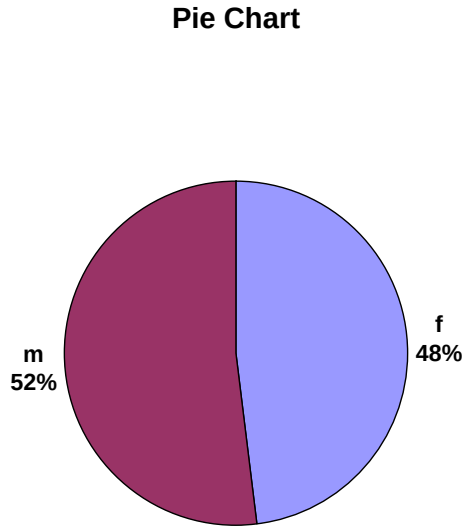
(b)



(c) The five-year annualized return of the 868 mutual funds is left-skewed with majority of them (about 89.86%) scattered between 0% and 15%. About 3% of the mutual funds have a negative five-year annualized return while about 7.14% of them have return higher than 20%. In general, the value mutual funds have higher five-year annualized return than growth mutual funds. Both types of mutual funds have five-year annualized return skewed to the left.

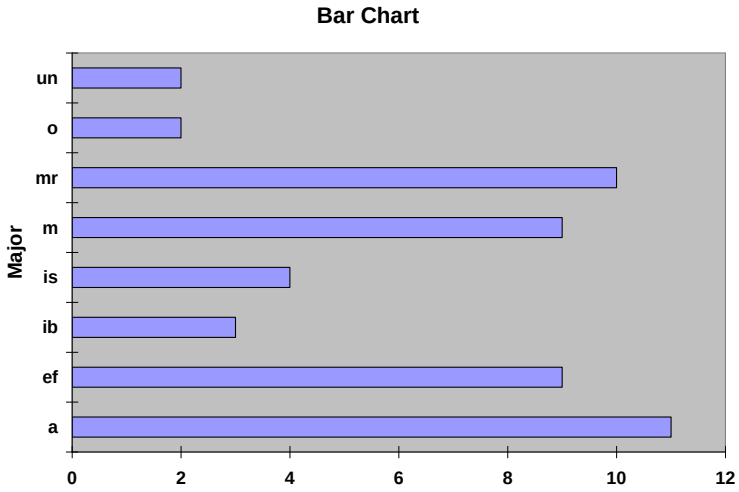
2.93 **Gender:**

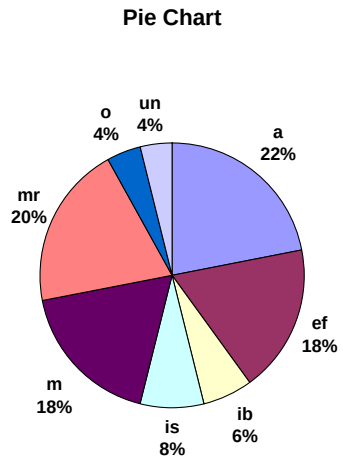


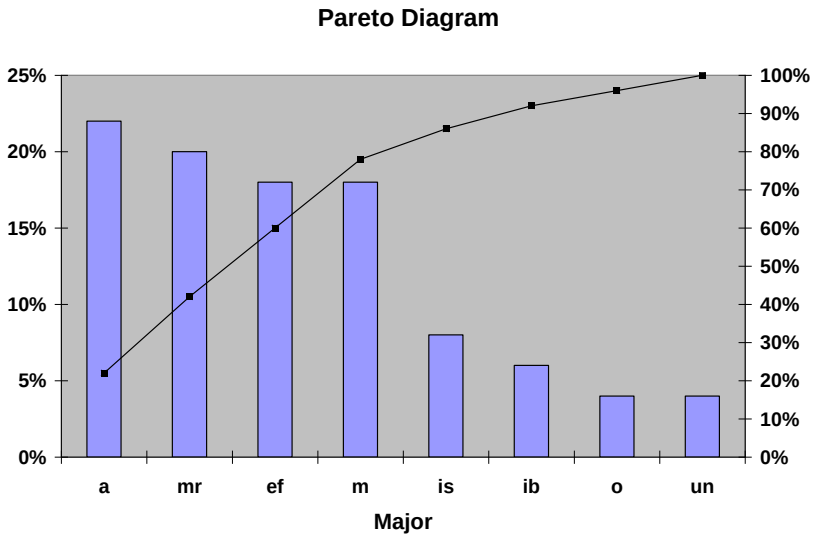


There are more males than females in the survey.

2.93
cont. **Major:**

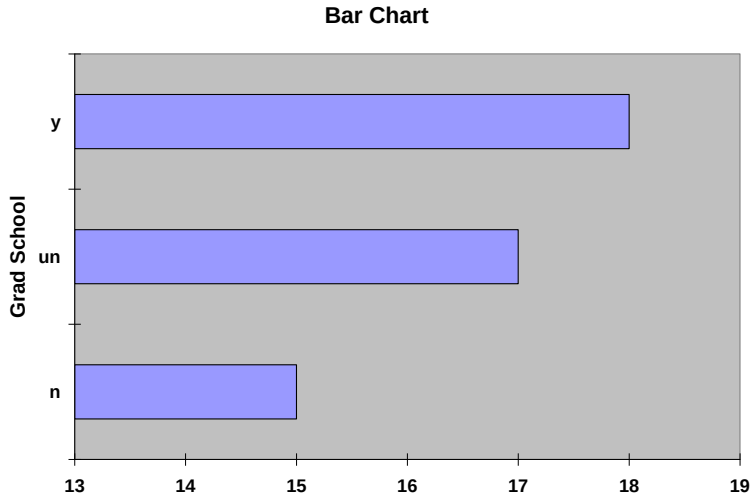


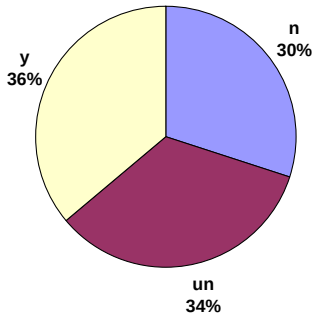


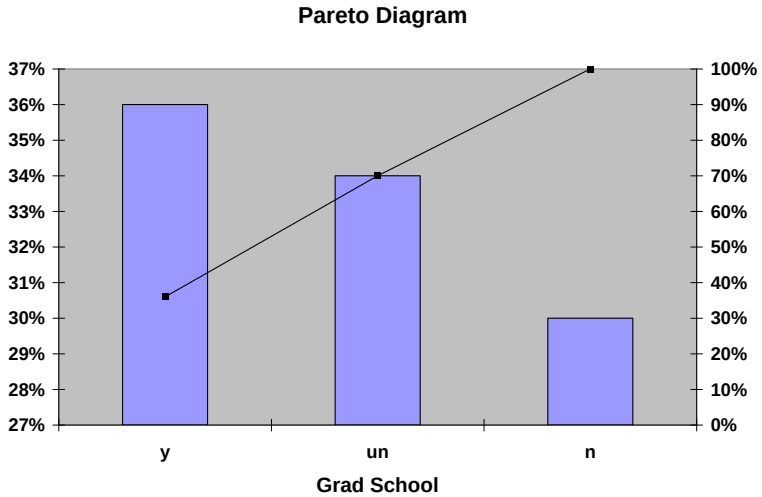


2.93 cont. Accounting, marketing/retailing, economics/finance and management constituted the "vital few" while the rest of the majors make up the "trivial many".

Grad School

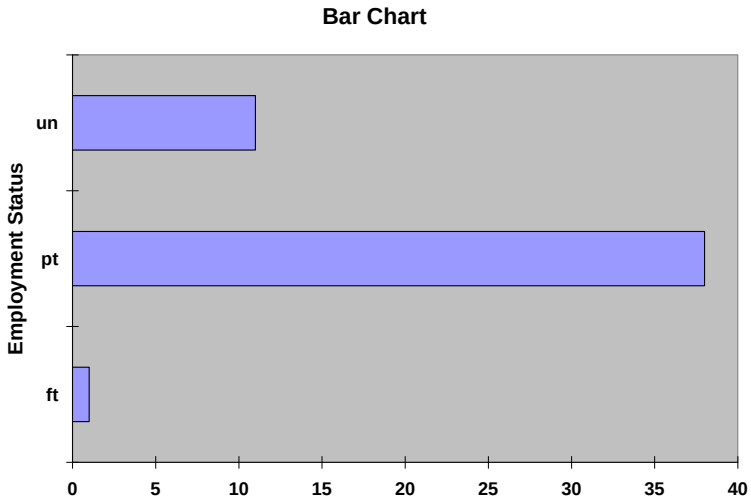


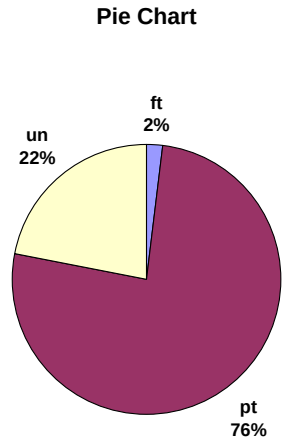


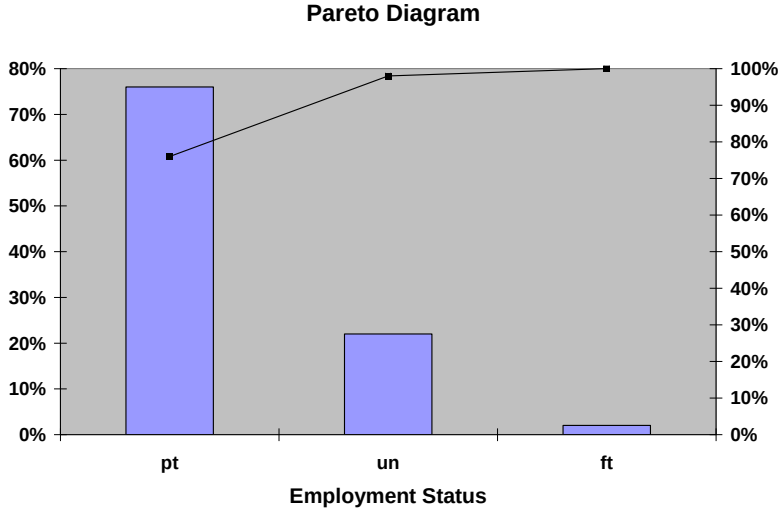


The percentages of students planning to attend graduate school are roughly evenly distributed among "Yes", "No" and "Not Sure".

2.93
cont. **Employment Status:**

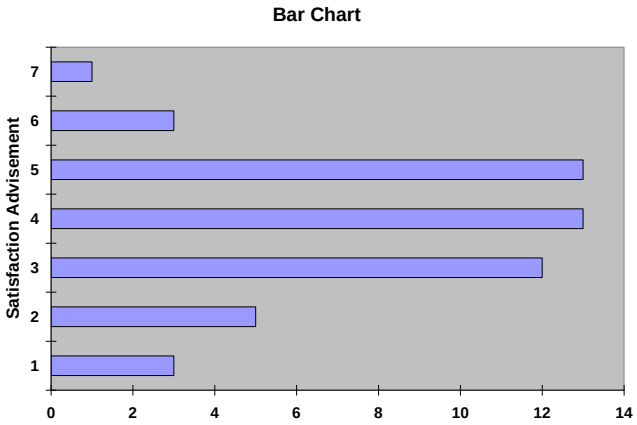


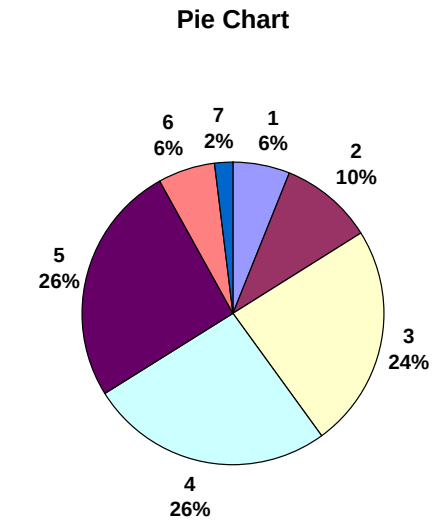


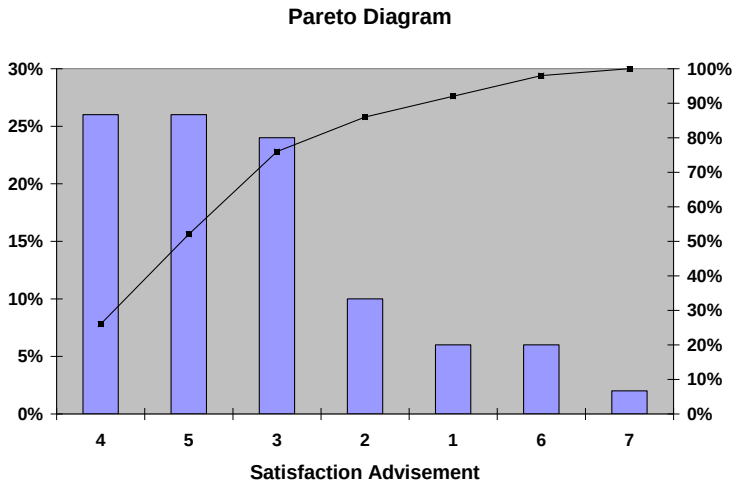


Part-time students constitute the "vital few" while full-time and unemployed students make up the "trivial many".

2.93
cont. **Satisfaction Advisement:**







About 80% of the students rated satisfaction with advisement in the range between 3 to 5.

2.93
cont. **Number Affiliations:**

Stem-and-Leaf Display
for Number of Affiliations
Stem unit: 1

0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
	0 0 0 0 0
1	0 0 0 0 0 0 0 0 0 0 0
2	0 0 0 0 0
3	0 0

Majority of the students (64%) have no affiliation with clubs, groups, organizations or teams currently.

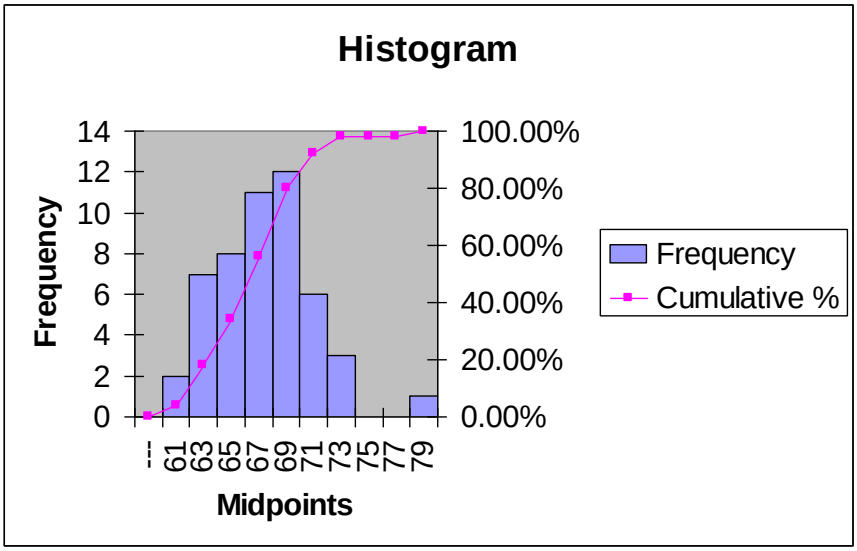
Age:

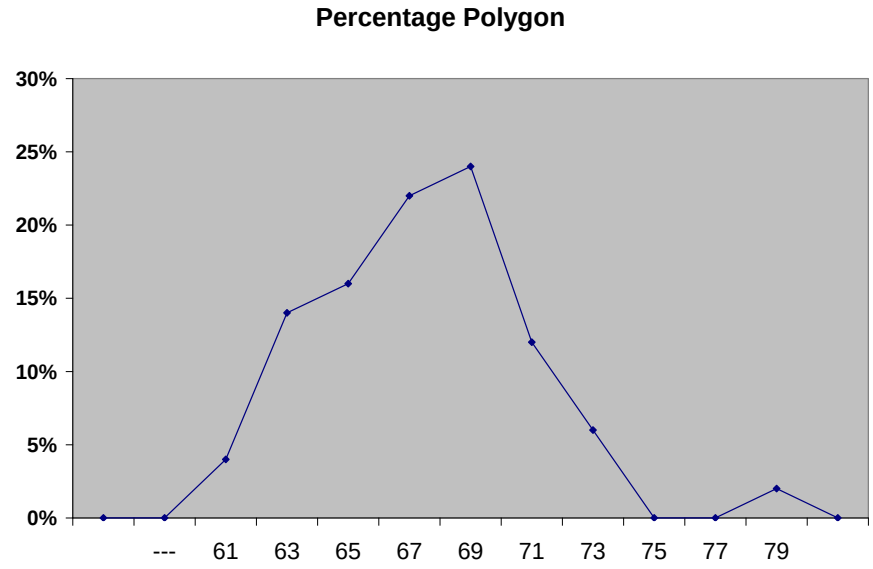
Stem-and-Leaf Display
for Age
Stem unit: 1

18	0
19	0 0 0 0 0 0 0 0 0 0 0 0 0 0
20	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
21	0 0 0 0 0 0 0 0 0 0
22	0 0 0 0
23	0
24	0
25	
26	
27	
28	
29	
30	0
31	
32	
33	0
34	
35	
36	0

Majority (92%) of the students surveyed are between 19 and 22 year old.

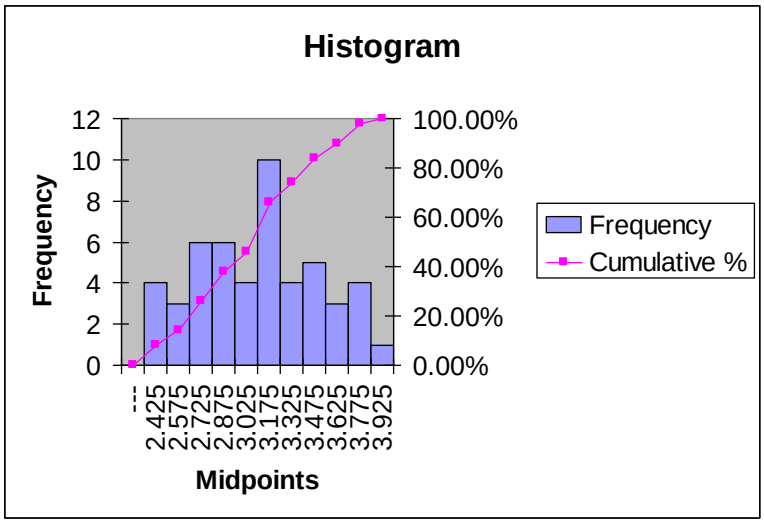
2.93
cont. **Height:**



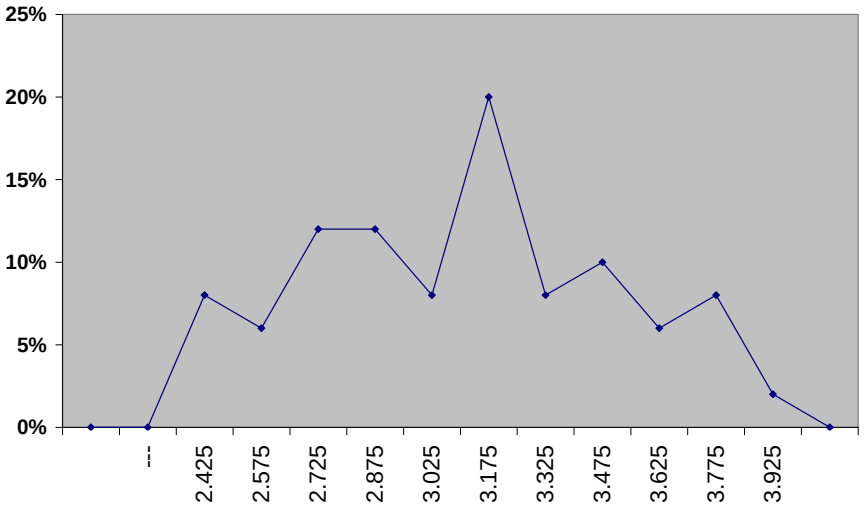


Height is right-skewed.

2.93
cont. **GPA:**



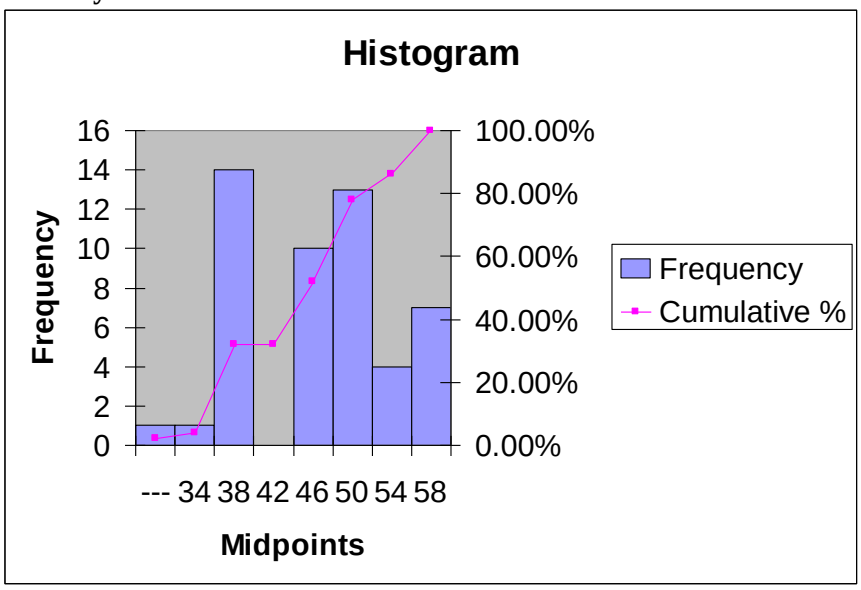
Percentage Polygon

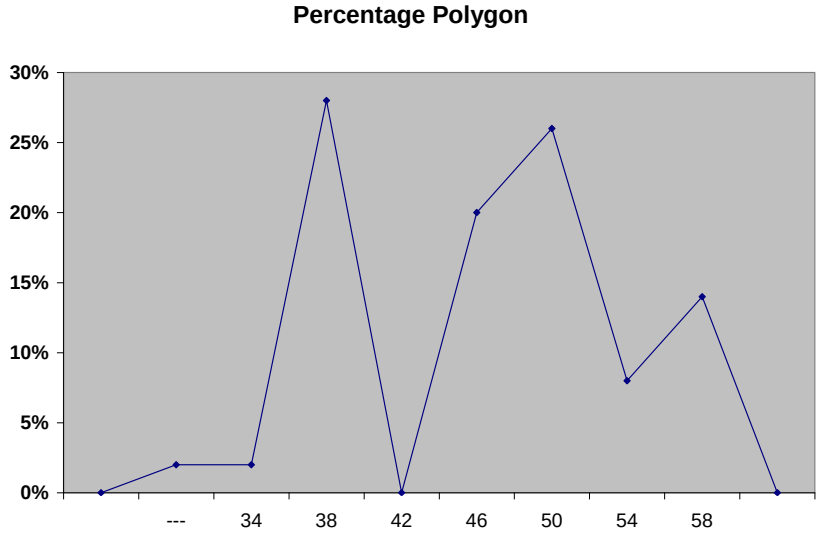


GPA is quite symmetrically distributed around 3.0.

2.93
cont.

Expected Salary:

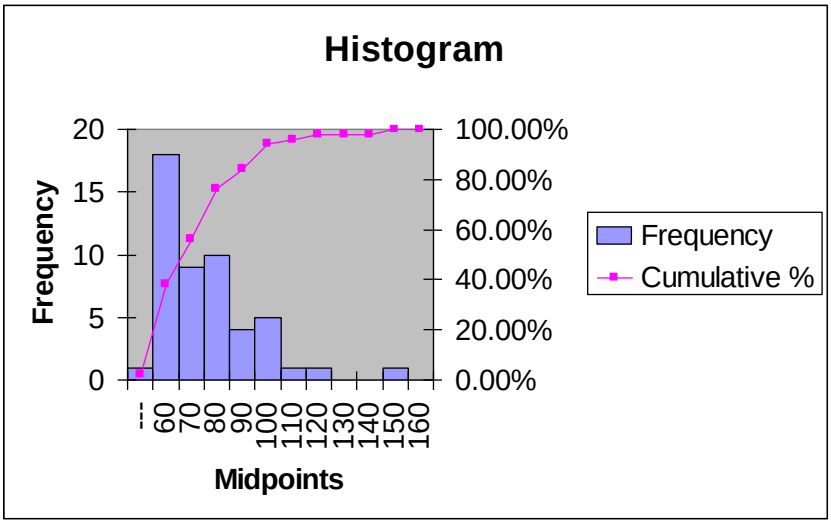


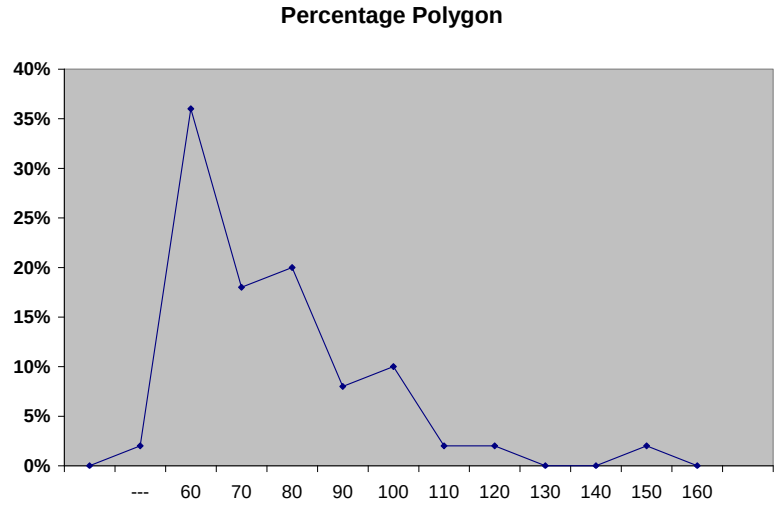


Expected salary is somewhat left-skewed.

2.93
cont.

Annual Salary in 5 Years:

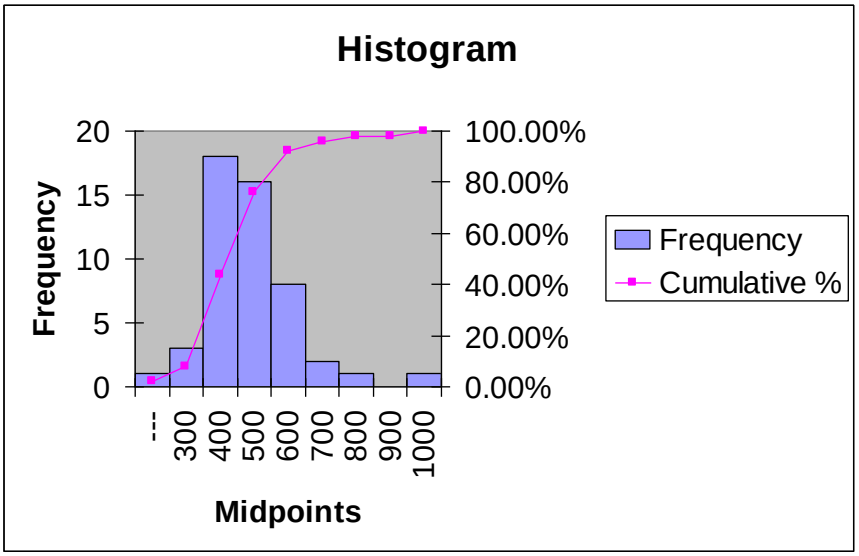


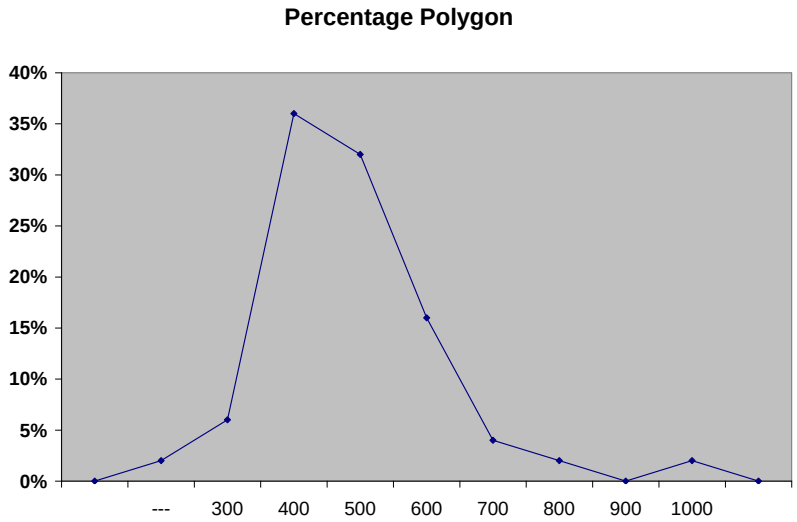


Annual salary in five years is right-skewed.

2.93
cont.

Spending:

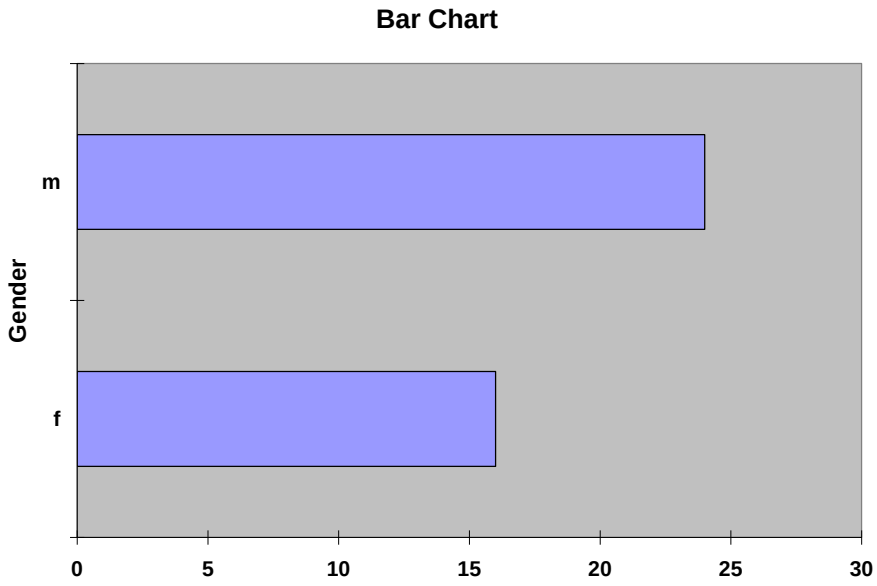


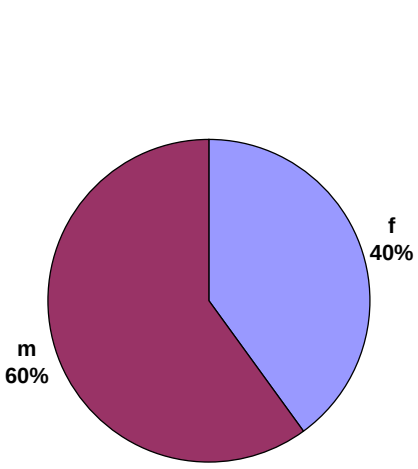


Spending is also right-skewed.

2.94 Student answers will vary.

2.95 **Gender:**

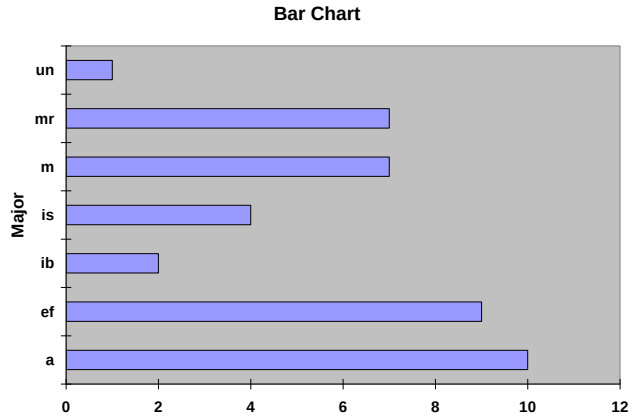


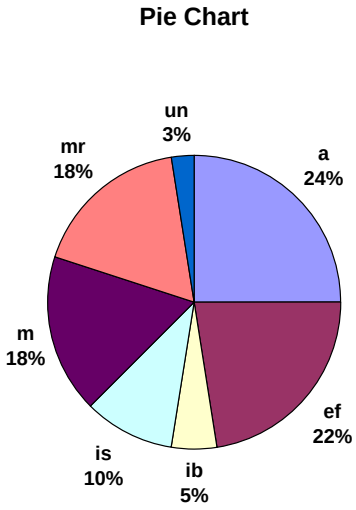


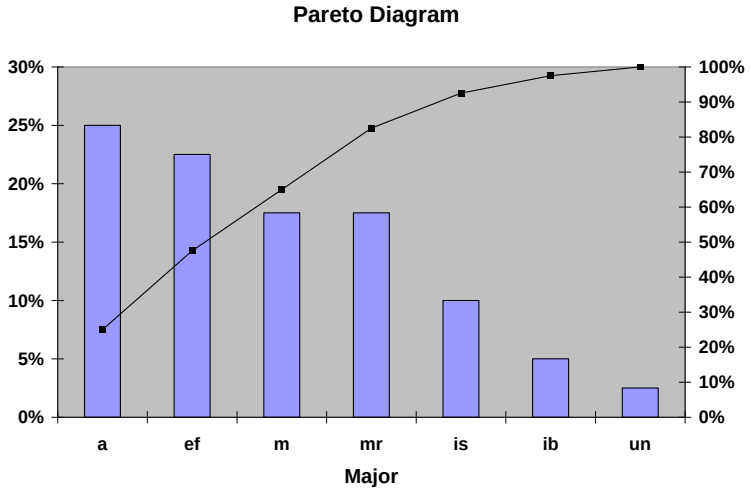
There are more males than females in the survey.

2.95
cont.

Major:

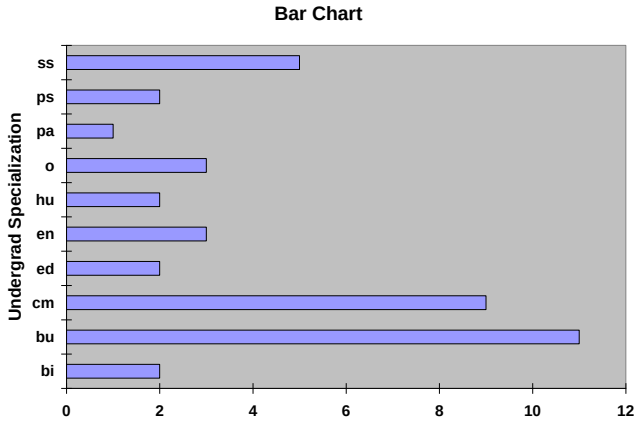


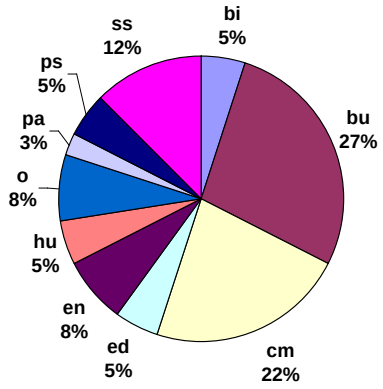


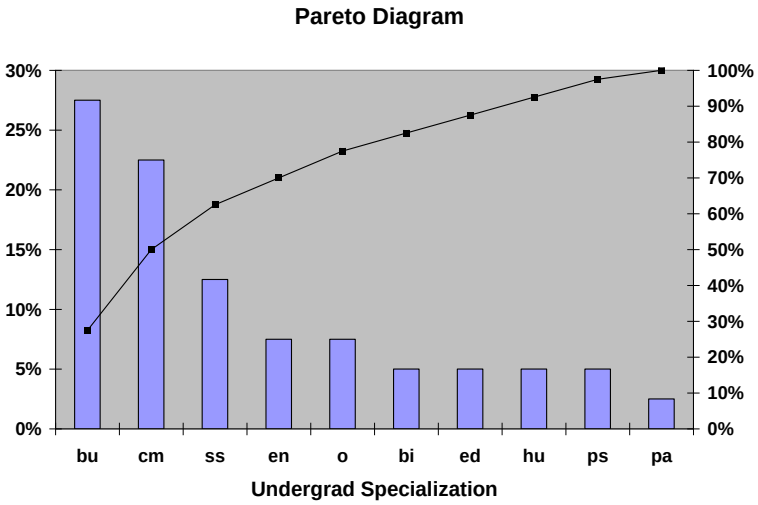


The "vital few" of accounting, economics/finance, management and marketing/retailing account for more than 80% of the majors.

2.95
cont. **Undergraduate Specialization:**

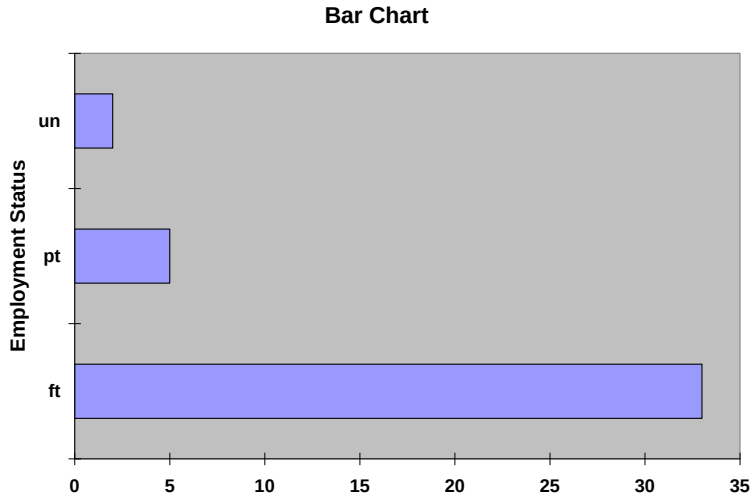


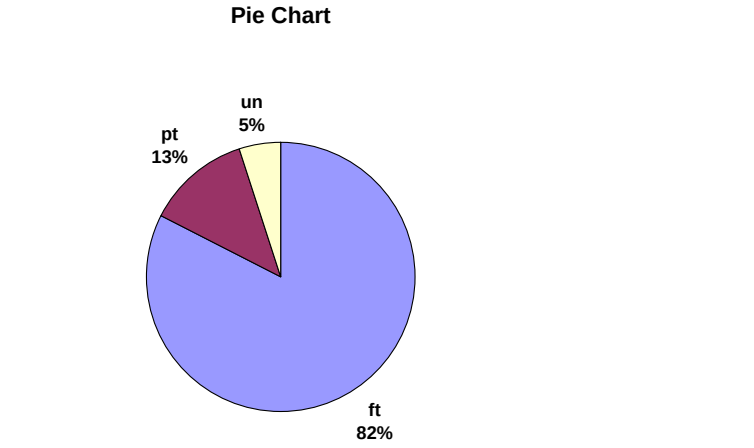


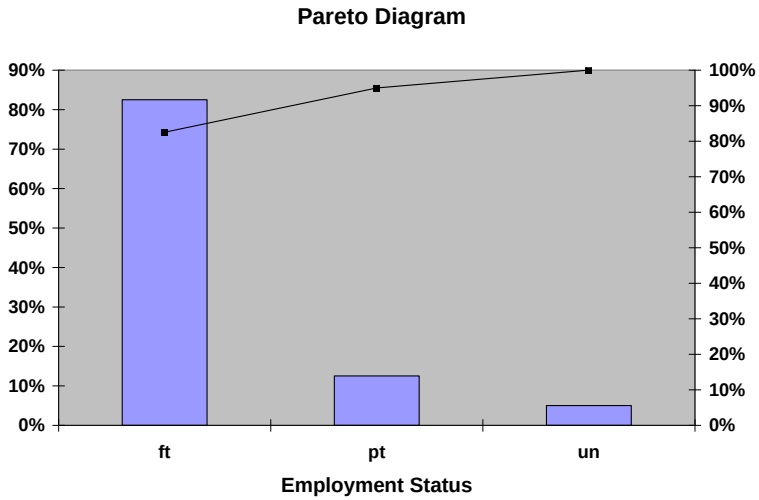


The "vital few" of business administration and computer/math account for half of the undergraduate specialization.

2.95
cont. **Employment Status:**

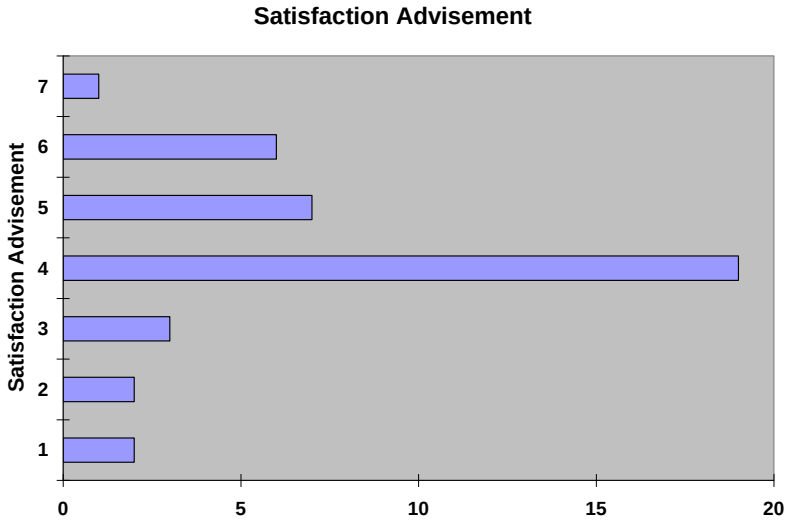


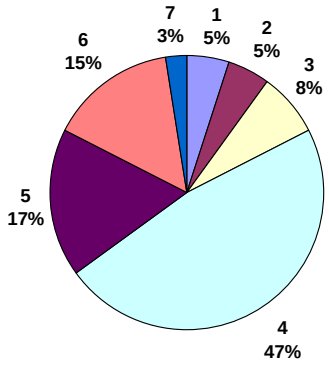


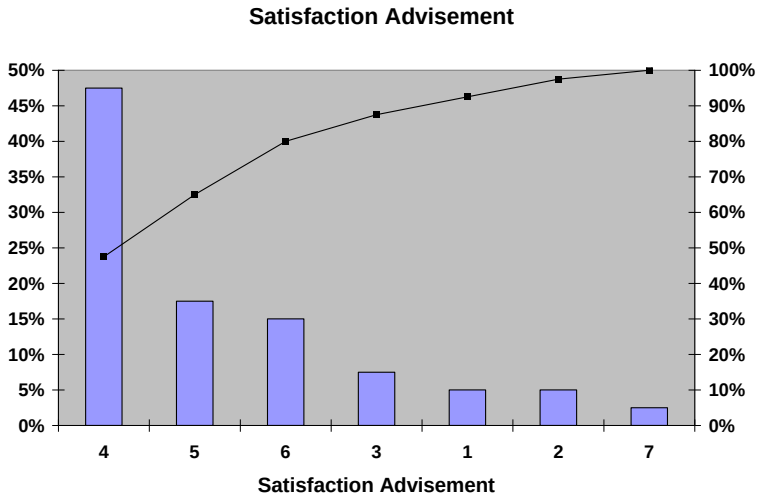


"Full-time" employment status accounts for more than 80% of the students.

2.95
cont. **Satisfaction Advisement:**







80% of the students rated their satisfaction advisement at between 4 and 5.

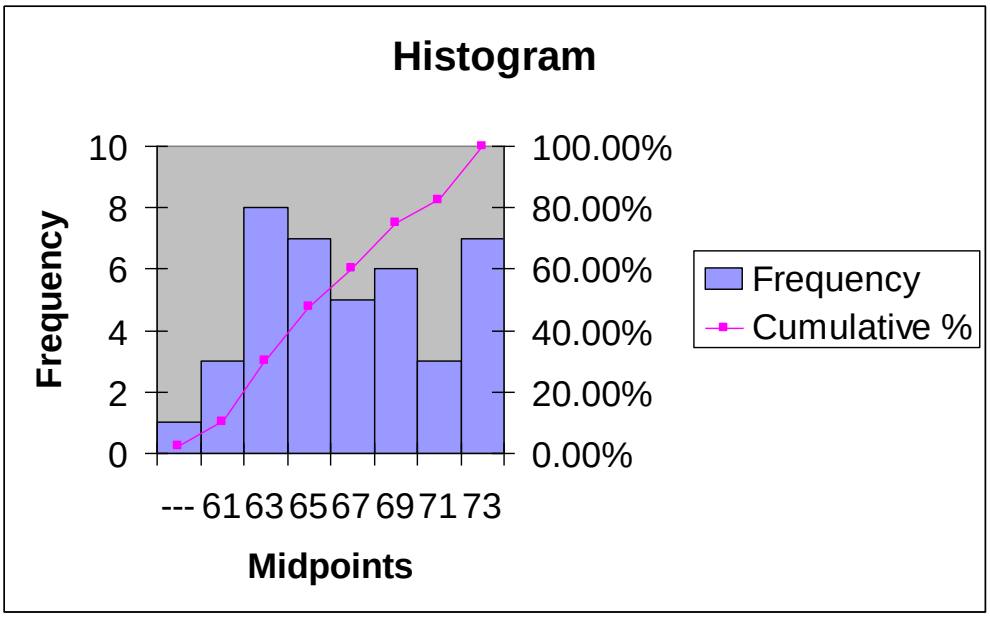
2.95
cont. **Age:**

Stem-and-Leaf Display for Age
Stem unit: 1

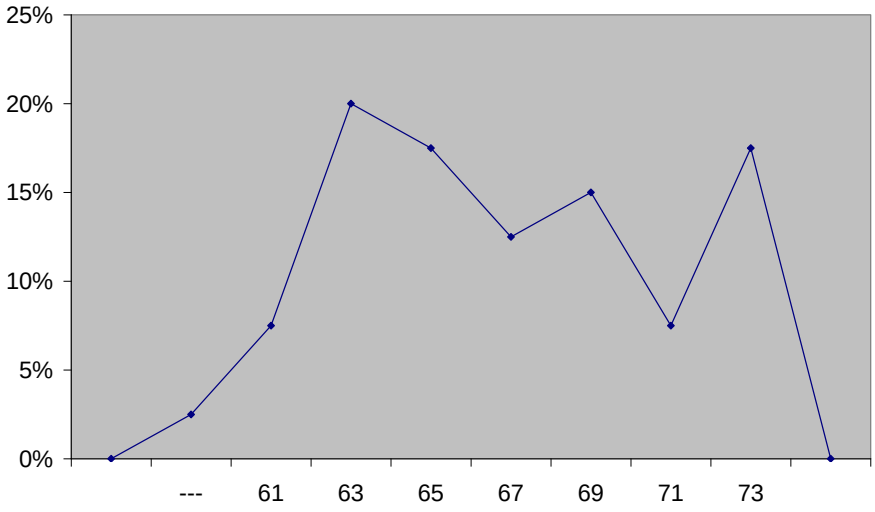
22	0
23	0
24	0 0
25	0 0 0 0 0
26	0 0 0
27	0 0 0
28	0 0 0
29	0 0
30	0 0 0 0
31	0 0 0
32	0 0 0
33	0 0
34	0
35	0 0
36	0
37	0
38	0
39	0
40	
41	0

Age is right-skewed.

2.95 cont. **Height:**



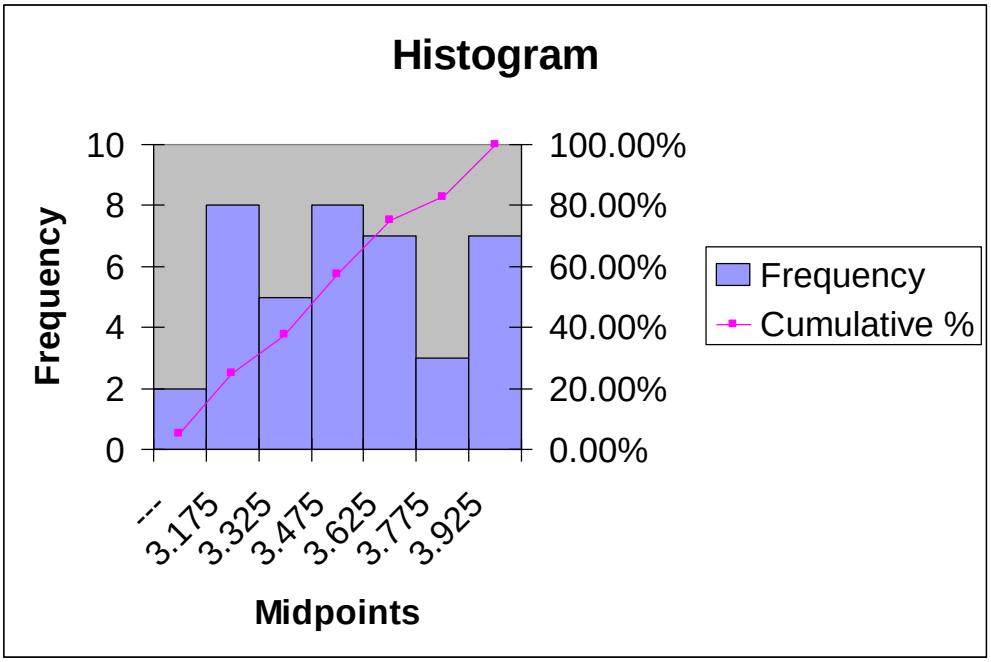
Percentage Polygon



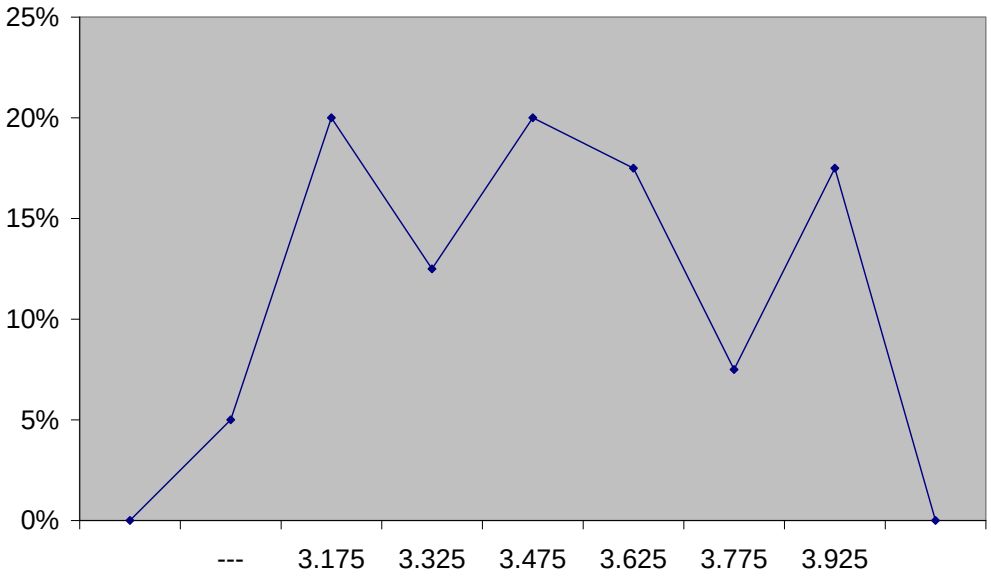
Height is left-skewed.

2.95
cont.

Grad GPA:



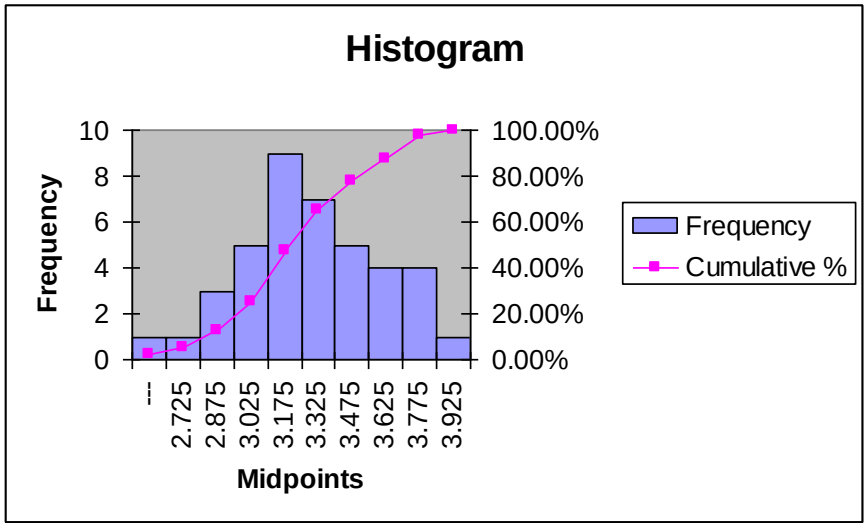
Percentage Polygon

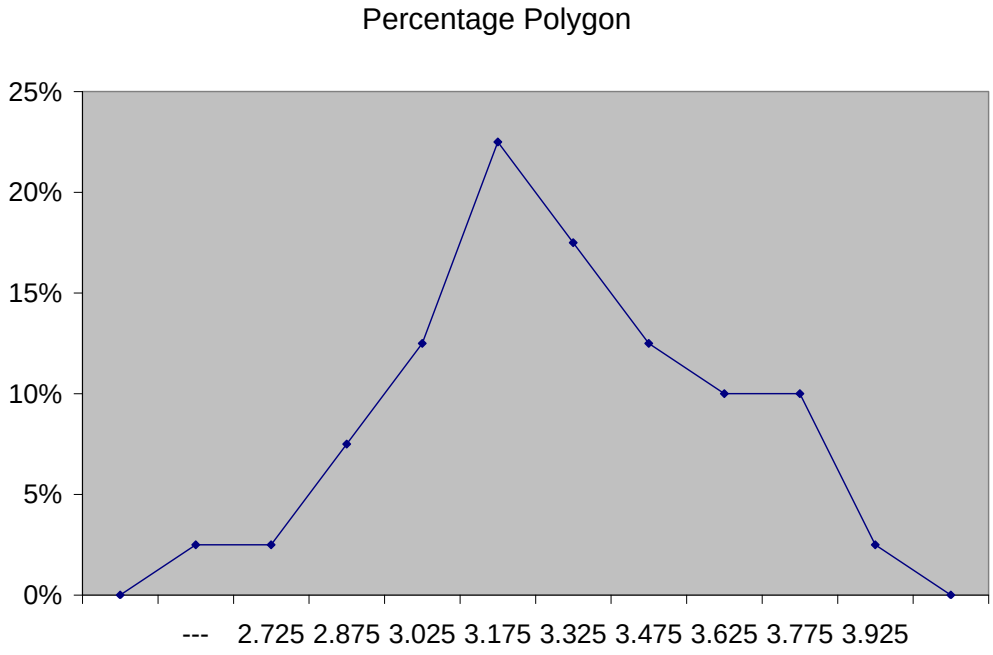


Graduate GPA is left-skewed.

2.95
cont.

Undergraduate GPA:

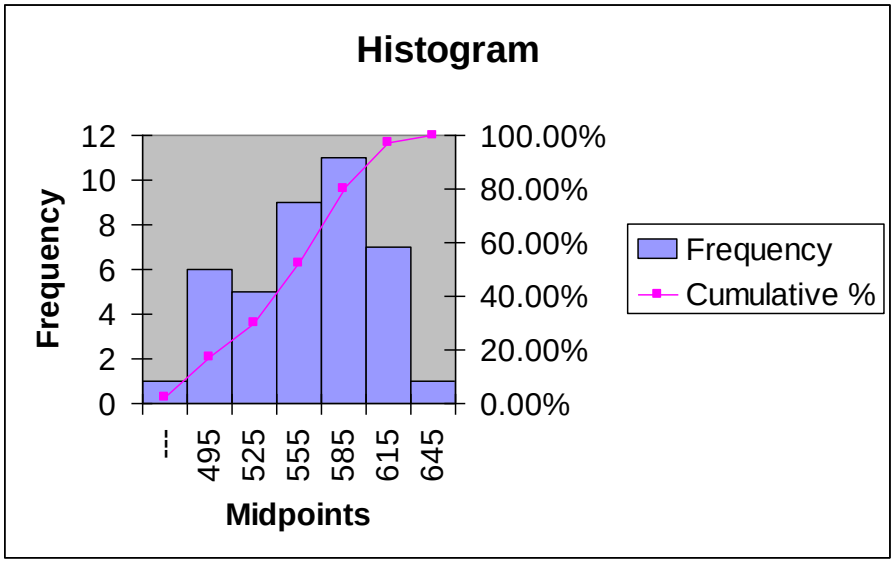




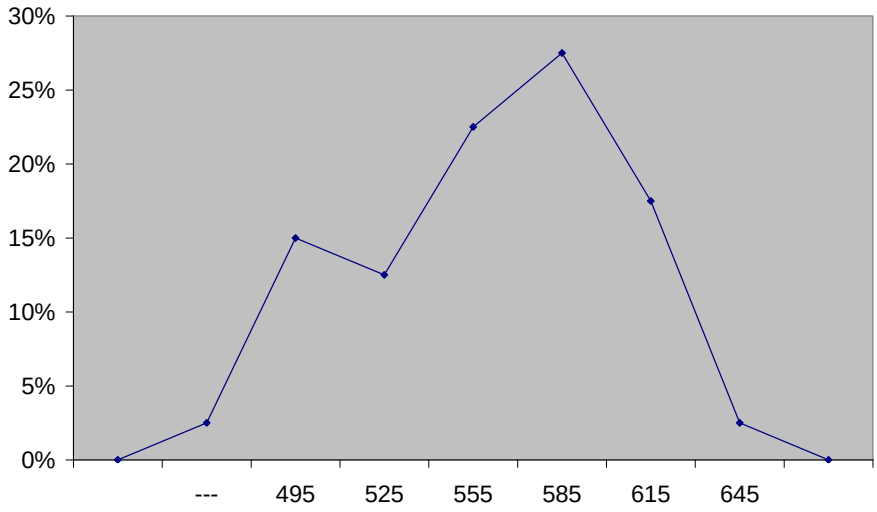
Undergraduate GPA is left-skewed.

2.95
cont.

GMAT:



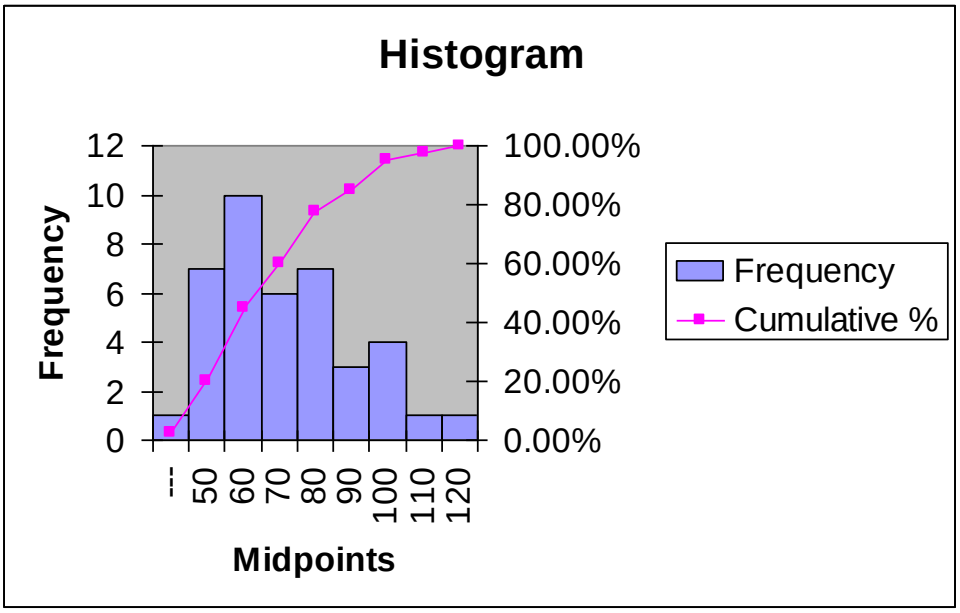
Percentage Polygon



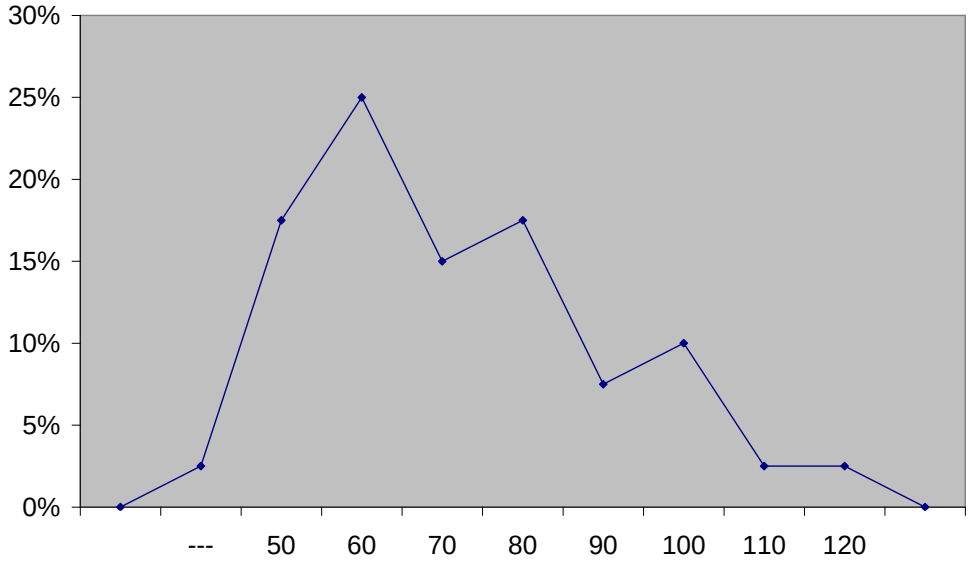
GMAT score is left-skewed.

2.95
cont.

Expected Salary:



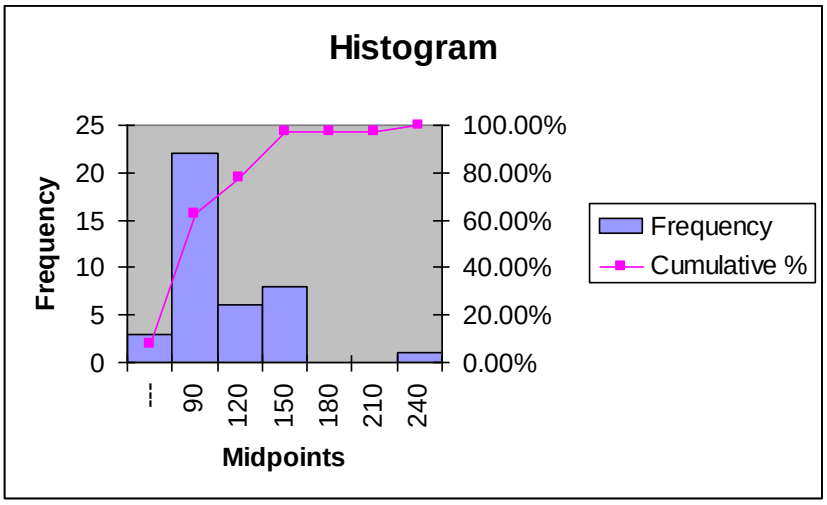
Percentage Polygon



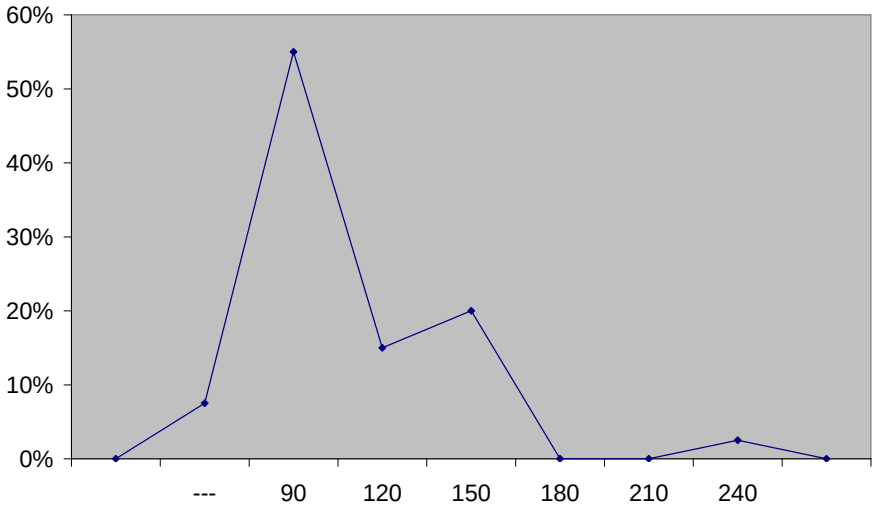
Expected salary is right-skewed.

2.95
cont.

Anticipated Salary in 5 Years:



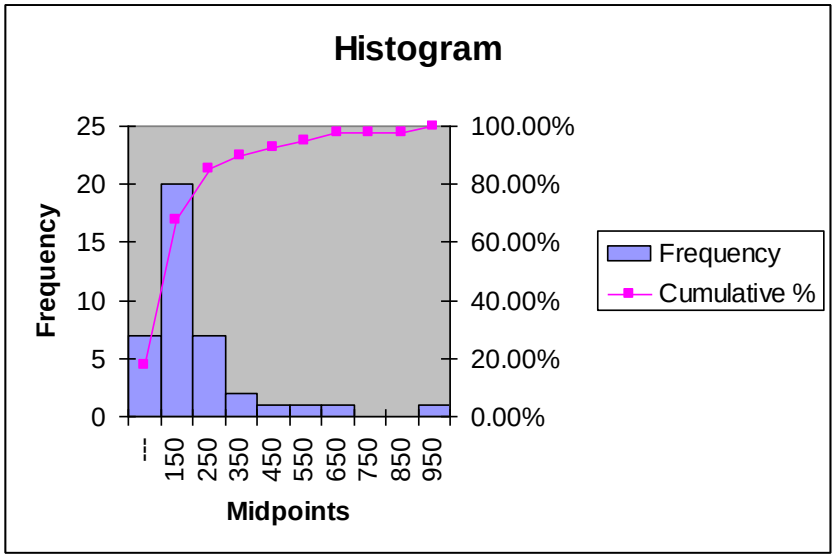
Percentage Polygon



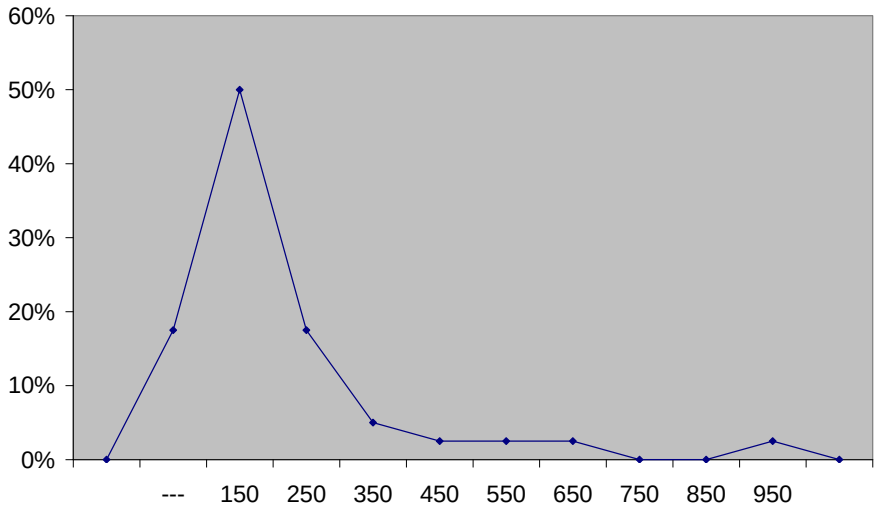
Anticipated salary is right-skewed.

2.95
cont.

Spending:



Percentage Polygon



Spending is right-skewed.

2.96 Student answers will vary.