

***Business Statistics - For Contemporary Decision
Making (Canadian) 3rd Edition 3rd edition Black
Solutions Manual***

SKU: 9781119577621-SOLUTIONS

Big Data Case Answers for Chapter 2

1. Cross Tabulation Tables:

- a.) Region and Control – the following crosstabulation table was constructed using the Pivot Table in Excel. In the database, both Region and Control have been converted to categorical values that were then used in the creation of this table. The analyst can see the number of hospitals in the various regions, the number of hospitals under the various types of control, and the intersection of each. Pie charts and other visuals can be used to display the various types of control, for example, in the Mid Atlantic region.

Count of Control		Column Labels				
Row Labels	Gov Fed	Gov	InvestorOwned	NonGov	Other	Grand Total
		NonFed	ForProfit	NonProfit	Nonprofit	
East North Central	7	36	46	38	135	262
East South Central	2	42	34	10	48	136
Mid Atlantic	8	20	18	11	111	168
Mountain	6	28	29	15	82	160
New England	3	6	10	3	78	100
Pacific	2	48	31	38	140	259
South Atlantic	9	44	46	10	143	252
West North Central	6	54	32	27	95	214
West South Central	3	64	233	41	108	449
Grand Total	46	342	479	193	940	2000

b.) Region and Service – This Excel-produced crosstabulation table between region and service was created using the Pivot Table. It shows which types of service dominate the hospital database (General Medical & Surgery, followed by Psychiatric, and Long-term Acute), and how the types of service breakdown by Region.

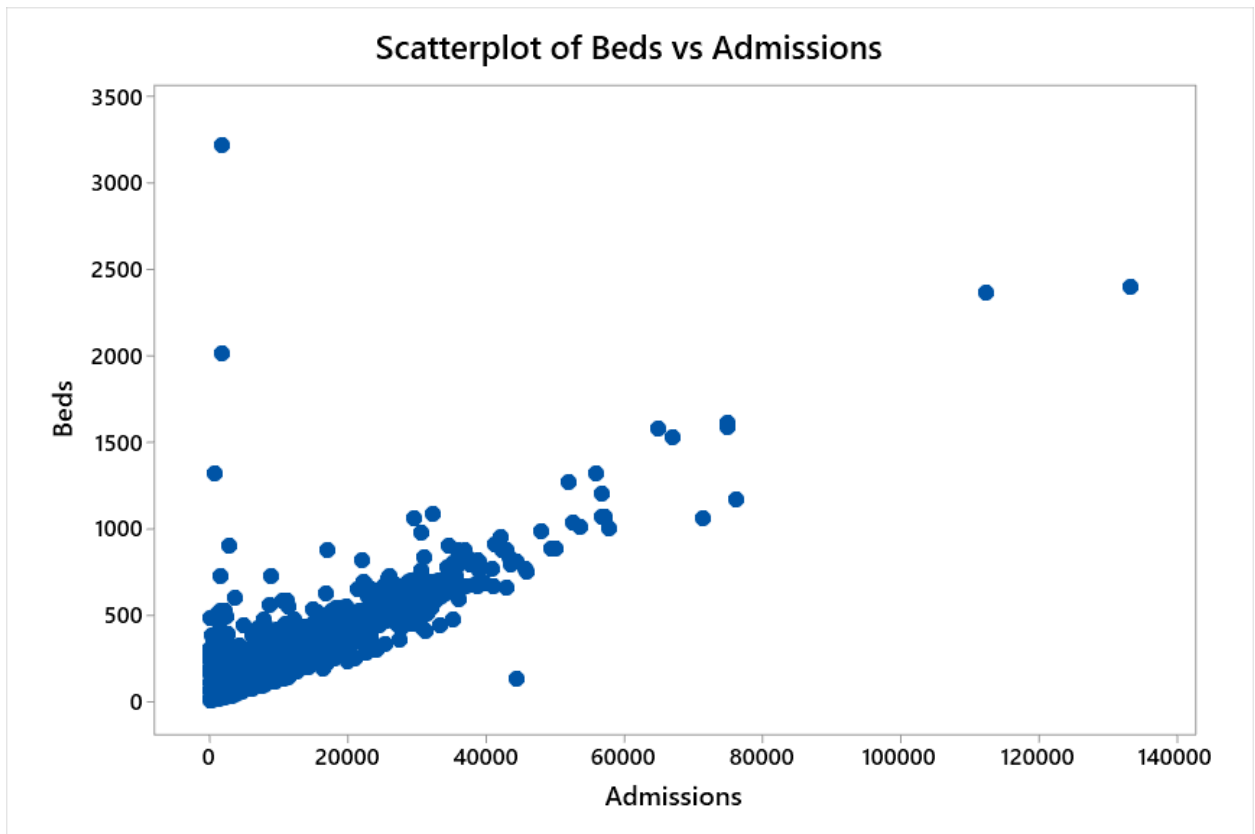
Row Labels	East North Central	East South Central	Mid Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central	Grand Total
Alcoh & Chem Dependency	3						3	1		7
Cancer	2		2			3	1		1	9
Children's	10	3	6	5	5	14	4	12	22	81
Chronic Disease					1		1			2
Gen Med & Surg	170	101	112	127	65	218	179	147	197	1316
Heart	3								4	7
Intellectual Disab								2		2
Long-term Acute	17	6	9	3	5	3	18	17	86	164
Obstetrics & Gyn	1	1	1		1	1			3	8
Orthopedic	2		2	1	1		1	1	4	12
Other Specialty			2	1	1	1	5	2	1	13
Psychiatric	38	14	19	12	13	13	20	18	58	205
Rehab	15	11	14	9	8	4	19	11	58	149
Surgical	1		1	2		2	1	3	15	25
Grand Total	262	136	168	160	100	259	252	214	449	2000

- c.) Service and Control – This Excel-produced crosstabulation table between service and control was created using the Pivot Table. It shows which types of service dominate the hospital database (General Medical & Surgery, followed by Psychiatric, Long-term Acute, and Rehab), and how the types of service breakdown by type of control. For example, Government Federal Control (mostly veteran's hospitals) are mainly General Medical & Surgery hospitals. Whereas, the modal control of rehab hospitals is Owned by Investors and are For Profit.

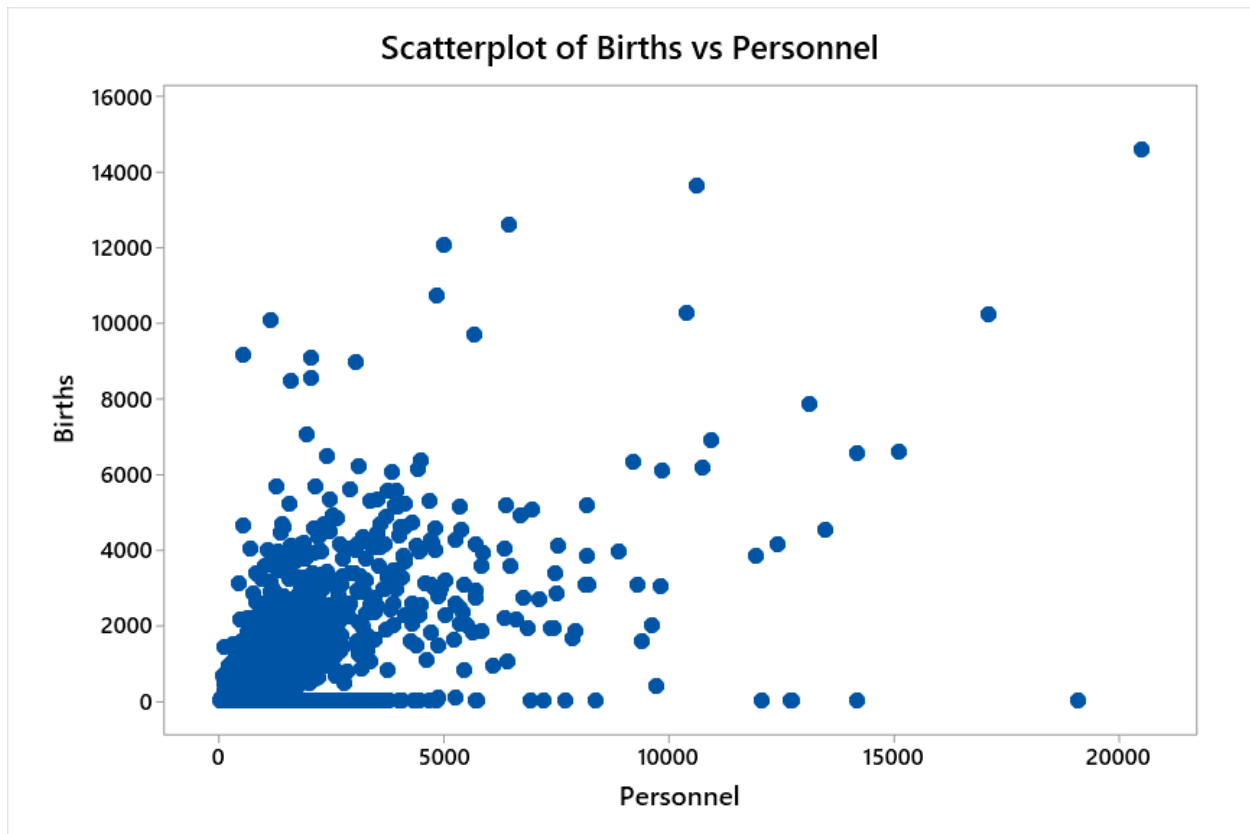
Count of Control Row Labels	Column Labels					
	Gov Fed	Gov NonFed	InvestorOwned ForProfit	NonGov NonProfit	Other Nonprofit	Grand Total
Alcoh & Chem Dependency		1	3		3	7
Cancer		3			6	9
Children's		2	11	2	66	81
Chronic Disease		1			1	2
Gen Med & Surg	43	249	127	167	730	1316
Heart			6		1	7
Intellectual Disab		2				2
Long-term Acute		6	113	13	32	164
Obstetrics & Gynecology			3		5	8
Orthopedic			8		4	12
Other Specialty	1	2			10	13
Psychiatric	2	70	81	4	48	205
Rehab		4	107	6	32	149
Surgical		2	20	1	2	25
Grand Total	46	342	479	193	940	2000

2. Four Different Scatter Plots – As an example, presented here are four different scatter plots from this AHA large database.

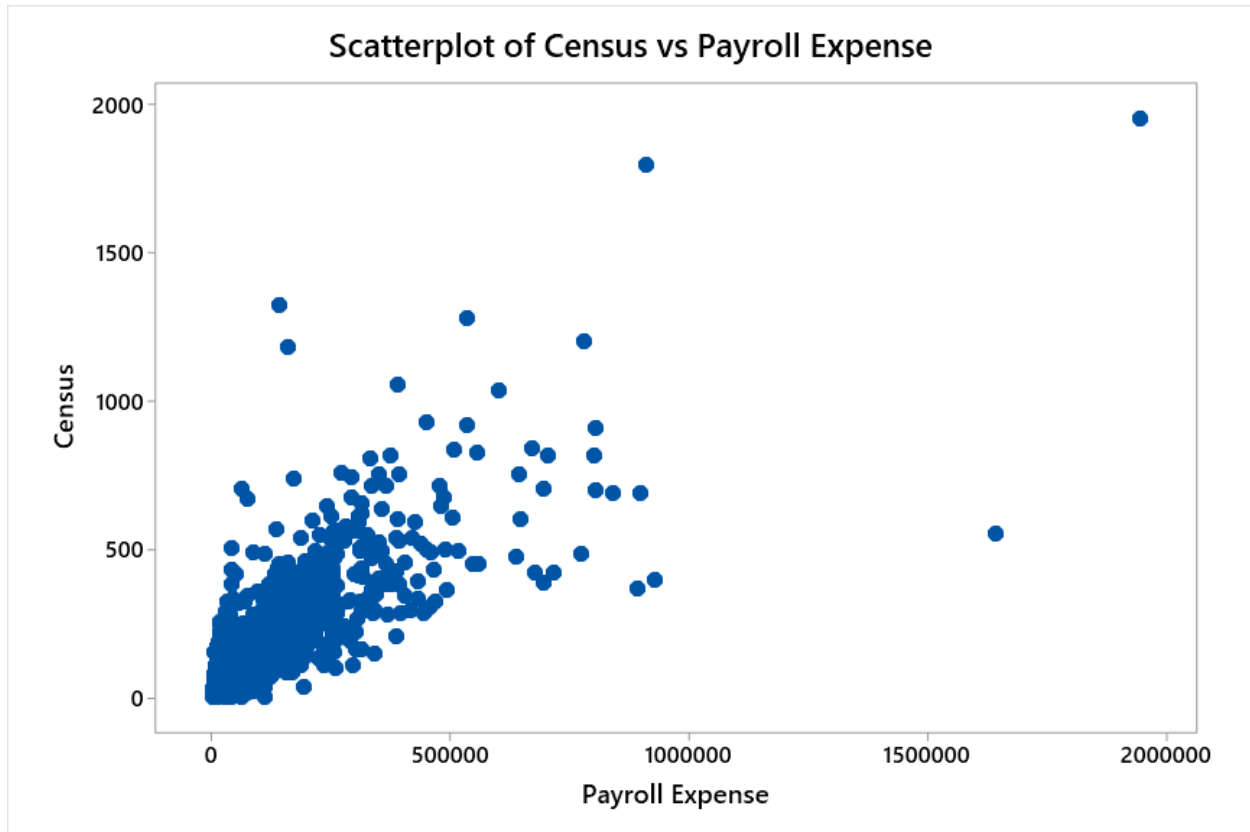
- a.) Shown here is a Minitab-produced scatterplot of Number of Beds vs. Number of Annual Admissions. Of the four scatter plots, this one between beds and admissions, most indicates a potential strong correlation.



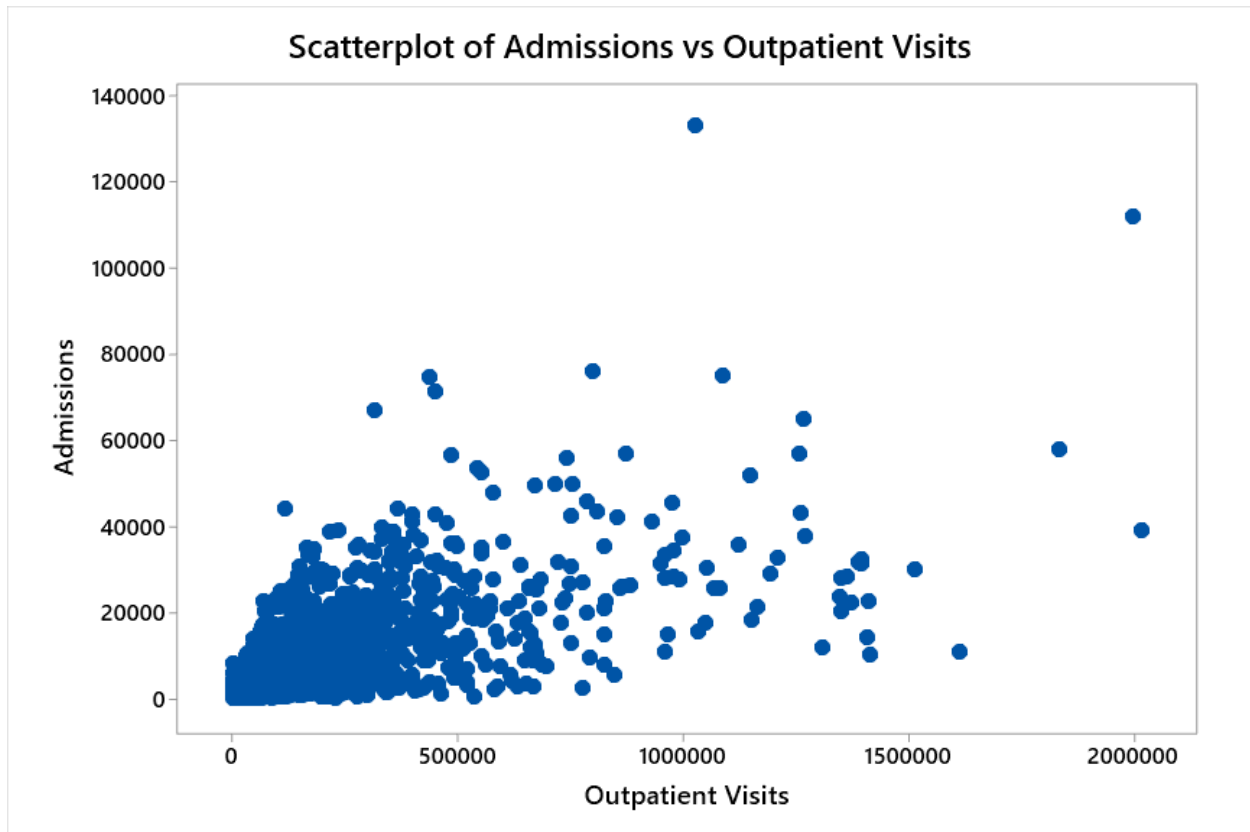
- b.) Shown here is a Minitab-produced scatterplot of Number of Births vs. Number of Personnel. This scatter plot indicates that the relationship between births and personnel is probably the weakest of the four scatter plots.



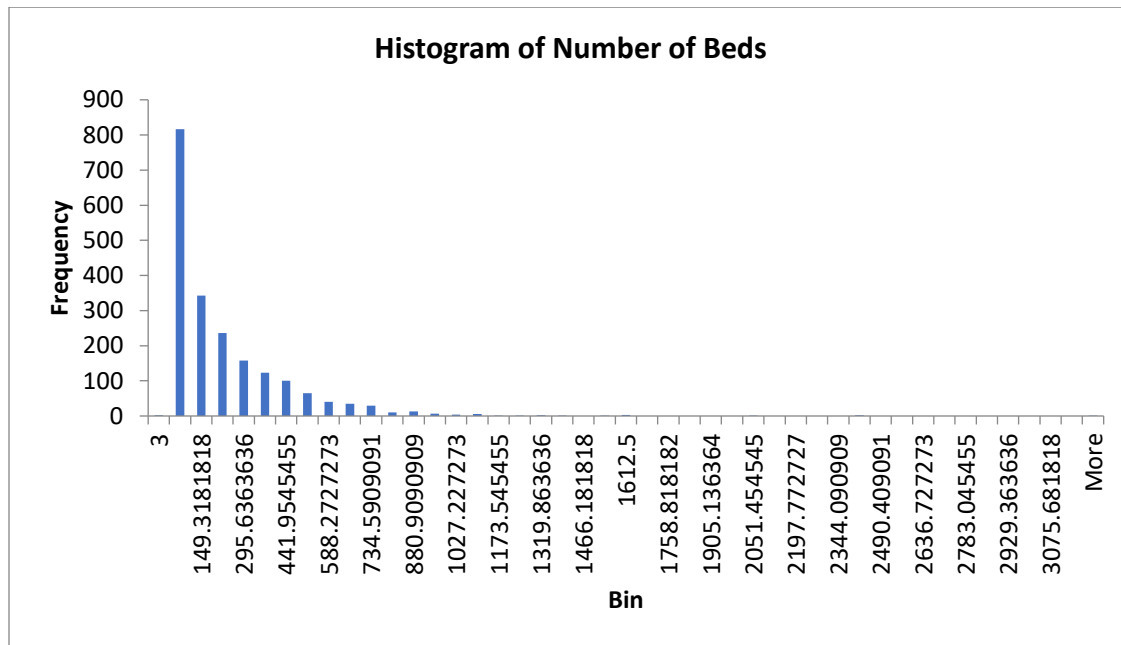
- c.) Shown here is a Minitab-produced scatterplot of Census vs. Payroll Expense. This scatter plot indicates that the relationship between census and payroll expense is stronger than births vs. personnel and weaker than beds vs. admissions.



- d.) Shown here is a Minitab-produced scatterplot of Admissions vs. Outpatient Visits. This scatter plot indicates that the relationship between admissions and outpatient visits is about as weak as for births vs. personnel.

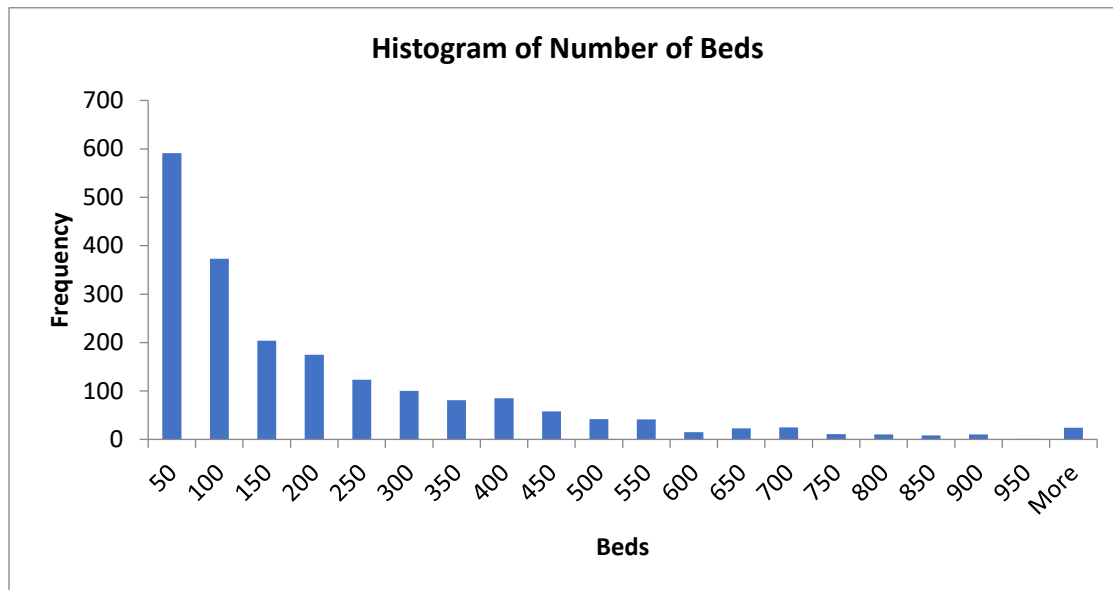


3. Histograms of Number of Beds. Shown first is a histogram of Number of Beds constructed from Excel. For this histogram, Excel created the bins.

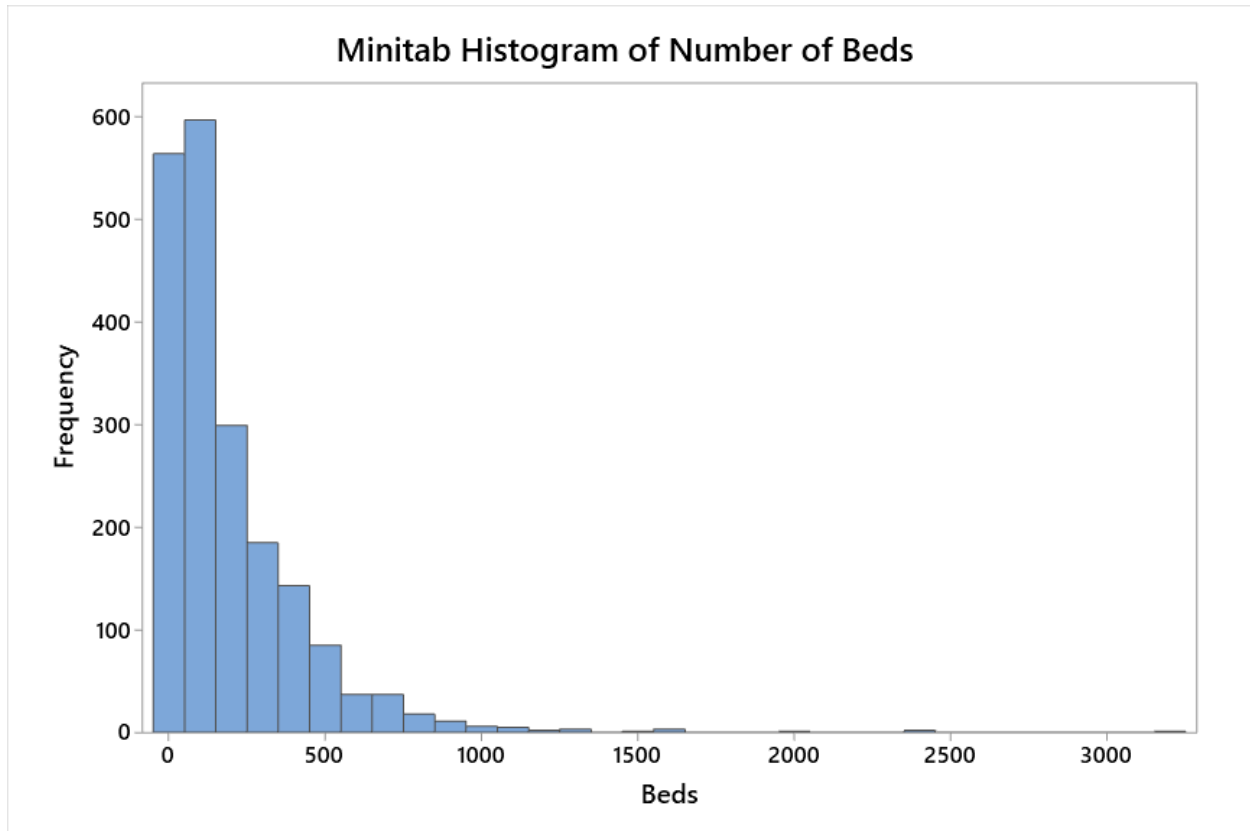


This shows that a few large hospitals are pulling the class intervals way to the right. There is a large “spike” of hospitals probably around 50 beds.

Next, we provided Excel with bins of 50, 100, 150, 200, 250, ..., 950. The resulting histogram makes the distribution a little easier to see the breaks in number of beds classes with many of the hospital having less than 150 beds.



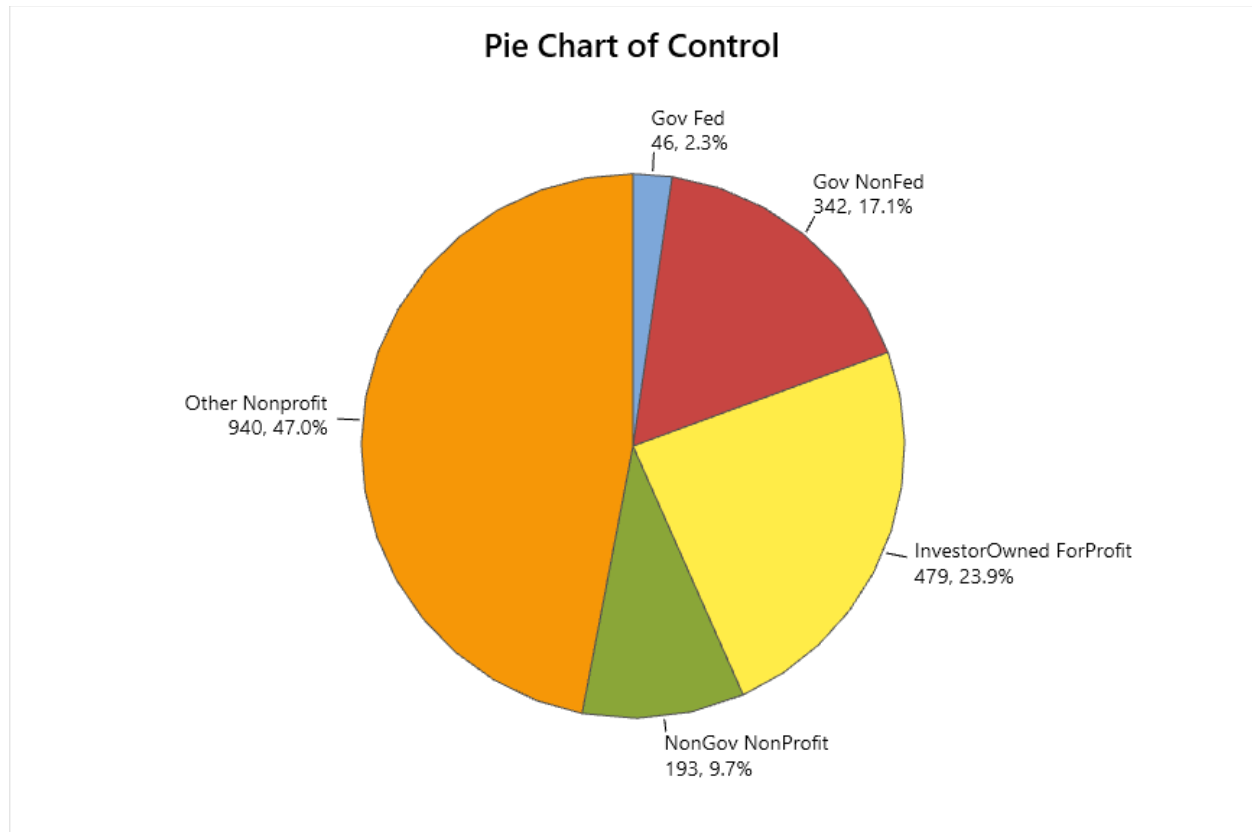
Shown next is a histogram of Number of Beds created in Minitab.



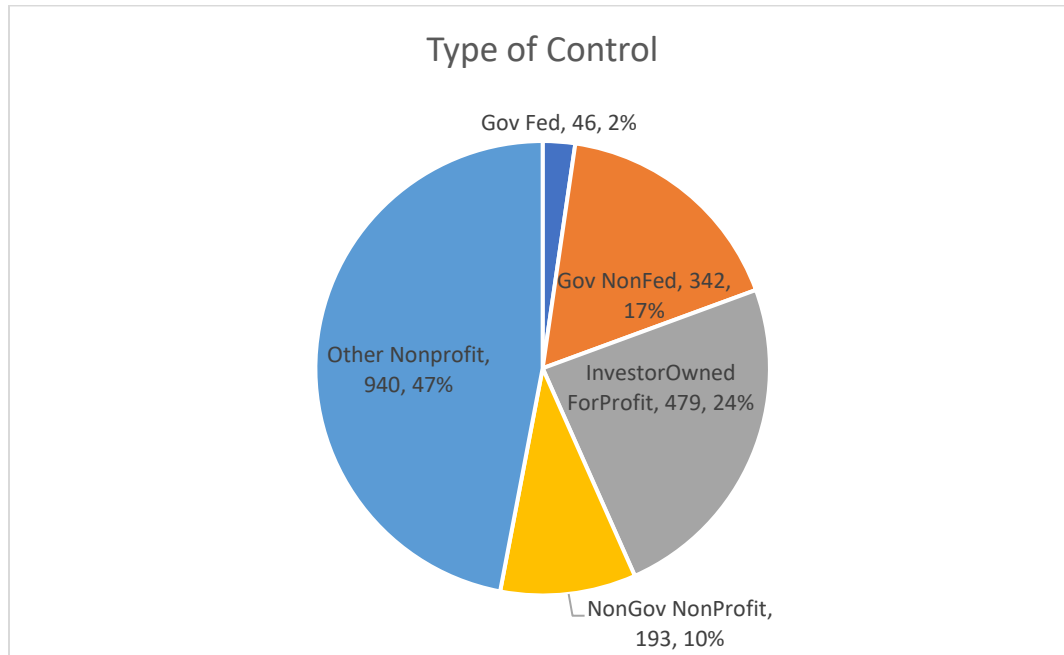
The student should be able to study their histogram and reach some conclusions about the size of the hospitals in this database. Obviously, we can see that they “pile-up” on the left of the graph and have a long thin “tail” to the right which is indicative of a graph that is skewed to the right (an opportunity to introduce a chapter 3 concept visually). Most of the hospitals in the database have less than 500 beds. In addition, one can see from these different histograms that the bin range (class endpoints) makes a difference in how the data are visualized.

- Construct a pie chart of the variable, Control.

From Minitab,



From Excel,



The pie charts show that almost half of all the hospitals in this database are “Other Nonprofit” (47%). Next are investor owned for-profit hospitals at 24% followed by nonfederal government hospitals (state, county, city, etc.) at 17%,. Note that there were 46 federal government hospitals (2%). About 90% of these federal government hospitals were veterans’ hospitals (this information is not contained in the database but is known by the author) and there were a couple of hospitals that were dedicated to native Americans.

Business Statistics, Third Canadian Edition

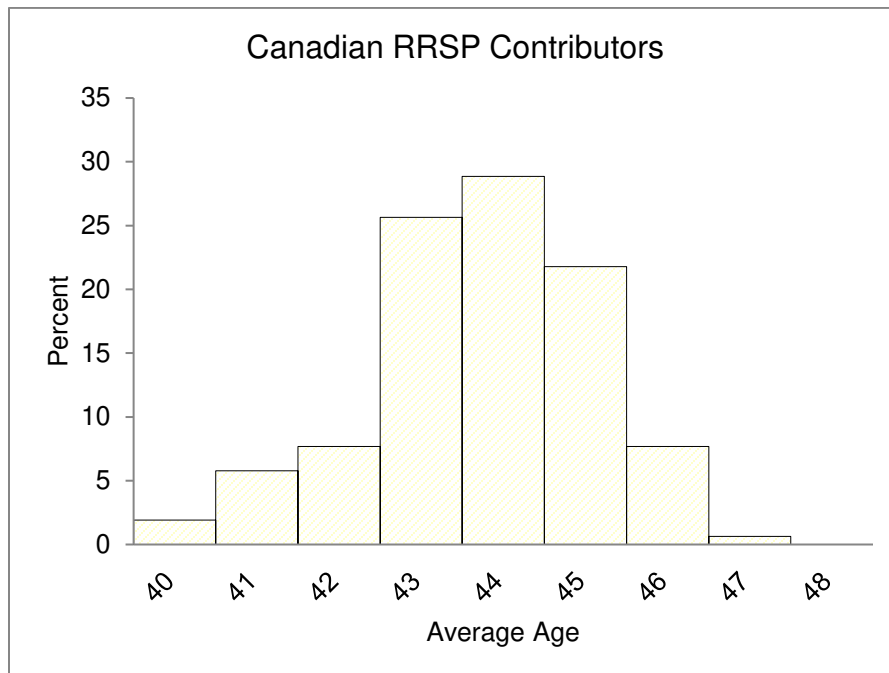
Database Exercise Answers

Chapter 2 Visualizing Data with Charts and Graphs

Question 1

(p.2-33)

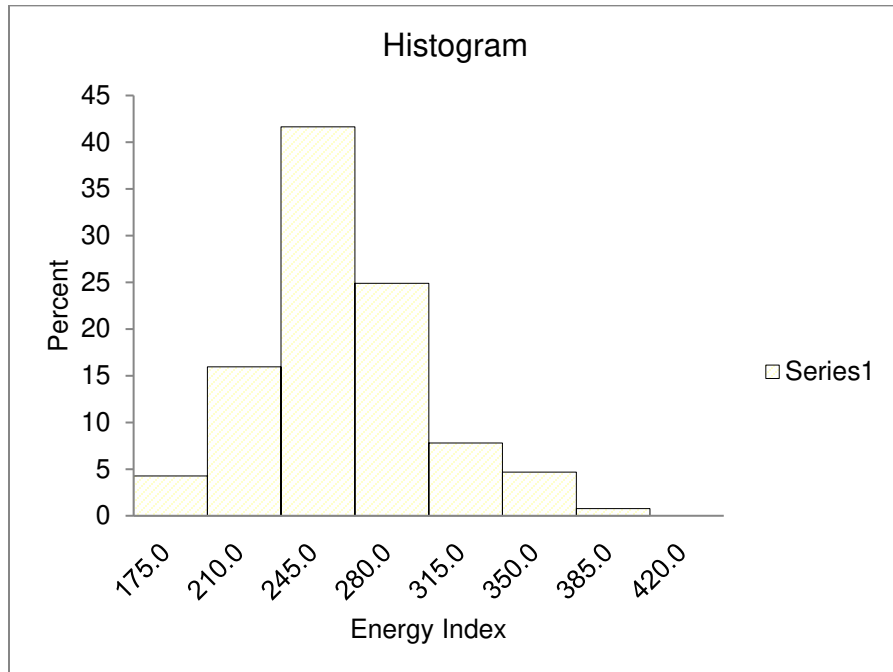
Using the RRSP database, construct a histogram for the variable Average Age of RRSP Contributors. How is the histogram shaped? Is it high in the middle or high near one or both ends of the data? Is it relatively constant in size across the class (uniform), or does it appear to have no shape? Does it appear to be nearly “normally” distributed?



The graph appears to be nearly “normally” distributed with most of the data centered around 43 to 46 years of age.

Question 2:

(p.47)



Frequency Distribution - Quantitative

Energy Index					cumulative			
<i>lower</i>		<i>upper</i>	<i>midpoint</i>	<i>width</i>	<i>frequency</i>	<i>percent</i>	<i>frequency</i>	<i>percent</i>
175.0	<	210.0	192.5	35.0	11	4.3	11	4.3
210.0	<	245.0	227.5	35.0	41	16.0	52	20.2
245.0	<	280.0	262.5	35.0	107	41.6	159	61.9
280.0	<	315.0	297.5	35.0	64	24.9	223	86.8
315.0	<	350.0	332.5	35.0	20	7.8	243	94.6
350.0	<	385.0	367.5	35.0	12	4.7	255	99.2
385.0	<	420.0	402.5	35.0	2	0.8	257	100.0
					257	100.0		

Most of the data center in the middle class and the left endpoint of the histogram. It appears the histogram is not uniform or symmetrical.

Question 3:

(p.47)

Stem	Leaf
4	8 9
5	0 1 2 4 5 5 7
6	0 3
7	0 1 3 6 9
8	0 1 1 6 7 9
9	5
10	0 4 5
11	4 4 5 8

Over the last 30 years the unemployment rate in UK did not drop below 4.8%. The highest rate of unemployment, 11.8%, occurred in Year 2. The lowest unemployment rate occurred in Year 22, after which the rate steadily climbed to 8.0% by Year 30.

SOLUTIONS TO PROBLEMS IN CHAPTER 2

2.1

a)

$$\text{Range} = \text{max} - \text{min} = 29 - (-12) = 41$$

$$\text{Class width} = 41/5 = 8.2 \approx 9$$

The 5 class frequency distribution:

<u>Class Interval</u>	<u>Frequency</u>
-15 - under -6	7
-6 - under 3	12
3 - under 12	13
12 - under 21	9
21 - under 30	<u>9</u>
Totals	50

b) Class width = $41/10 = 4.1 \approx 5$. 10 class frequency distribution:

<u>Class Interval</u>	<u>Frequency</u>
-15 - under -10	2
-10 - under -5	5
-5 - under 0	7
0 - under 5	10
5 - under 10	7
10 - under 15	3
15 - under 20	7
20 - under 25	4
25 - under 30	5
30 - under 35	<u>0</u>
Totals	50

c) The ten class frequency distribution gives a more detailed breakdown of temperatures. It allows locating more accurately the temperatures with the greatest frequency. The class with the highest frequency (modal class), is 0 – under 5 class with a frequency of 10. The five class distribution collapses the intervals into broader classes making it appear that there are nearly equal frequencies in each class.

2.2 The range is 22 (61-39). We will use 8 classes. The class width will be 3. The frequency distribution is as shown below

<u>Class Interval</u>	<u>Frequency</u>
39 - under 42	3
42 - under 45	5
45 - under 48	20
48 - under 51	21
51 - under 54	26
54 - under 57	11
57 - under 60	11
60 - under 63	3
Totals	100

The distribution reveals that 49 of the 100 boxes of raisins contain 51 or less raisin (39 - under 51). It shows that there are three boxes that have 10 or more extra raisins (60 - under 63) and three boxes that have 8-11 less raisins (39 – under 41) than the boxes are supposed to contain.

2.3

<u>Class Interval</u>	<u>Frequency</u>	<u>Class Midpoint</u>	<u>Relative Frequency</u>	<u>Cumulative Frequency</u>
0 - 5	6	2.5	$6/86 = .0698$	6
5 - 10	8	7.5	.0930	14
10 - 15	17	12.5	.1977	31
15 - 20	23	17.5	.2674	54
20 - 25	18	22.5	.2093	72
25 - 30	10	27.5	.1163	82
30 - 35	<u>4</u>	32.5	<u>.0465</u>	86
TOTAL	86		1.0000	

The relative frequency tells us that it is most probable that a customer is in the 15 - 20 category (.2674). Over two thirds (.6744) of the customers are between 10 and 25 years of age.

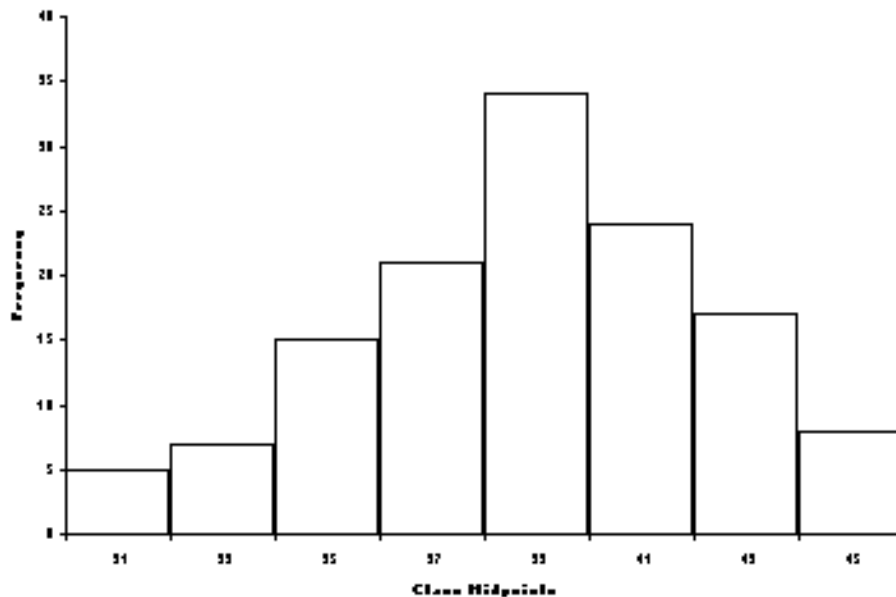
2.4

<u>Class Interval</u>	<u>Frequency</u>	<u>Class Midpoint</u>	<u>Relative Frequency</u>	<u>Cumulative Frequency</u>
0-2	218	1	.436	218
2-4	207	3	.414	425
4-6	56	5	.112	481
6-8	11	7	.022	492
8-10	8	9	.016	500
TOTAL	500		1.000	

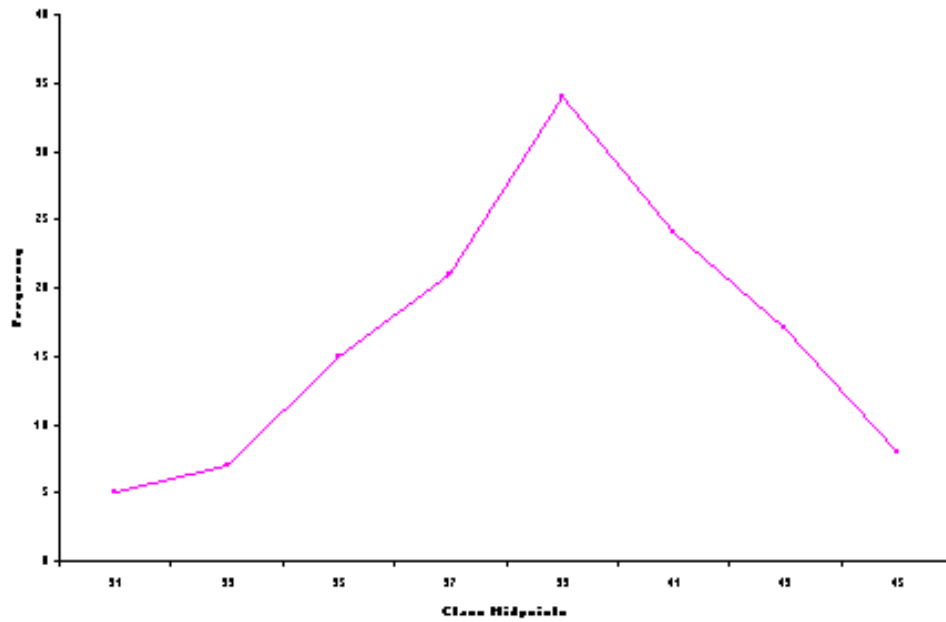
2.5

Some examples of cumulative frequencies in business:
 sales for the fiscal year,
 costs for the fiscal year,
 spending for the fiscal year,
 inventory build-up,
 accumulation of workers during a hiring buildup,
 production output over a time period.

2.6 Histogram:

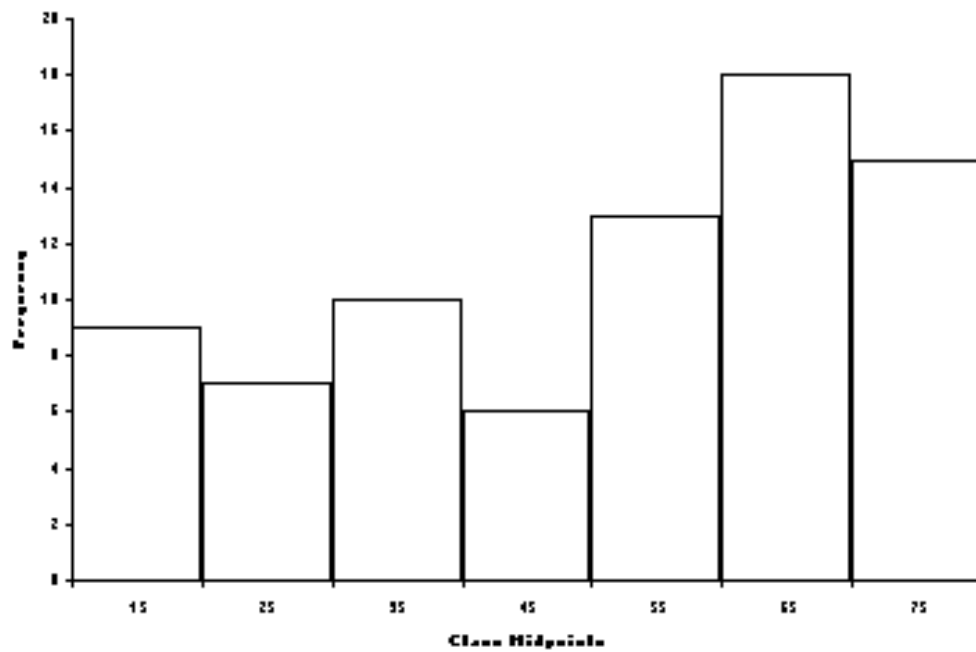


Frequency Polygon:

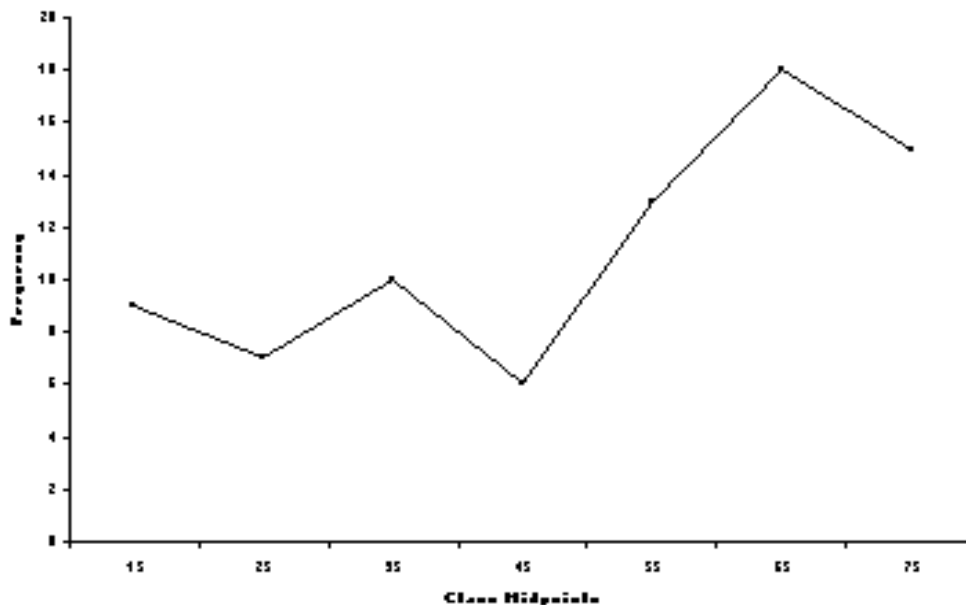


Comment: The assembly times “pile up” near the middle of the graphs indicating that many of the assembly times are between 36 and 42 minutes.

2.7 Histogram:

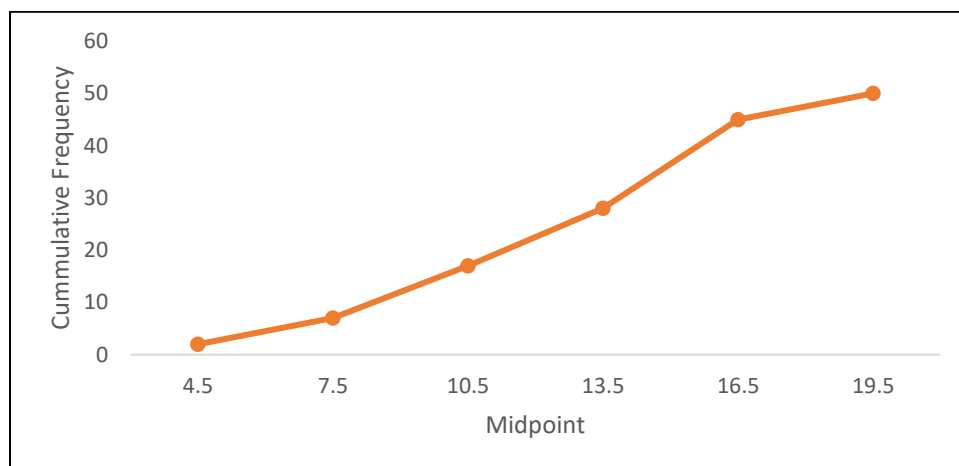


Frequency Polygon:



Comment: The histogram indicates that the number of calls per shift varies widely. However, the heavy numbers of calls per shift fall in the 50 to 80 range. Since these numbers occur quite frequently, staffing planning should be done with these number of calls in mind realizing from the rest of the graph that there may be shifts with as few as 10 to 20 calls.

2.8 Ogive:



2.9 STEM LEAF

21	2 8 8 9
22	0 1 2 4 6 6 7 9 9
23	0 0 4 5 8 8 9 9 9 9

24	0 0 3 6 9 9 9
25	0 3 4 5 5 7 7 8 9
26	0 1 1 2 3 3 5 6
27	0 1 3

The stem and leaf plot indicates that sales prices vary quite a bit within the range of \$212,000 and \$273,000. It is evident from the stem and leaf plot that there is a strong grouping of prices in the five price ranges from the \$220's through the \$260's.

2.10 STEM LEAF

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

The stem and leaf plot shows that the numbers of passengers per flight were relatively evenly distributed between the high teens through the sixties. Rarely was there a flight with at least 70 passengers. The category of 40's contained the most flights (10).

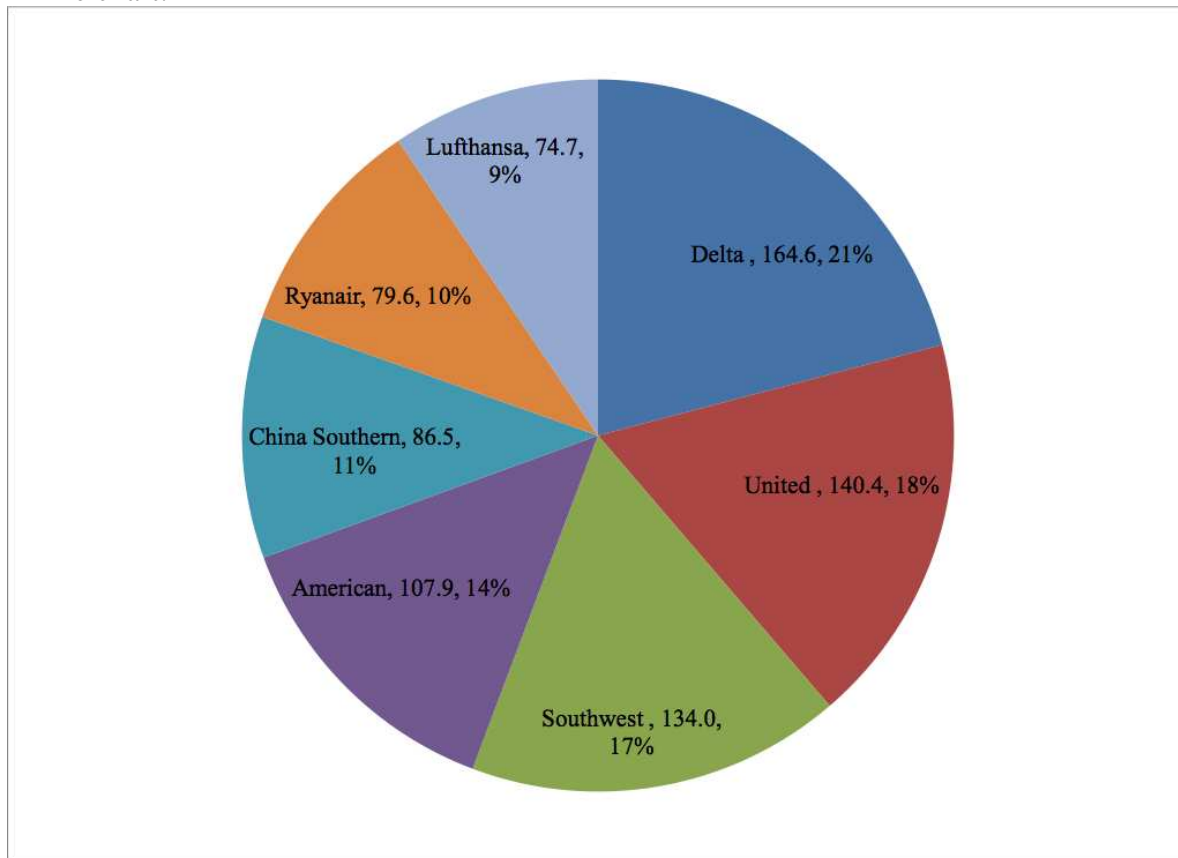
2.11 The histogram shows that there is only one airport with more than 105 million passengers and from the given problem information, we know that that airport is Atlanta's Hartsfield-Jackson International Airport which has more than 105 million passengers. 40% (12) of the top 30 airports have between 65 and 75 million passengers. The next largest grouping is between 55 and 65 million passengers in which there are six airports.

2.12 We assume that the class endpoints (10, 20, 30, ...) are indicated on the horizontal axis and the marks between them represent the class midpoints; the cumulative frequencies are marked on the vertical axis. The ogive shows that out of 200, 50 pots contain less than 10 legal king crabs. About 90 selected pots (45%) have fewer than 20 legal king crabs, and about 120 (60%) contain less than 30 legal king crabs. However, 180 pots (90%) have less than 80 legal king crabs.

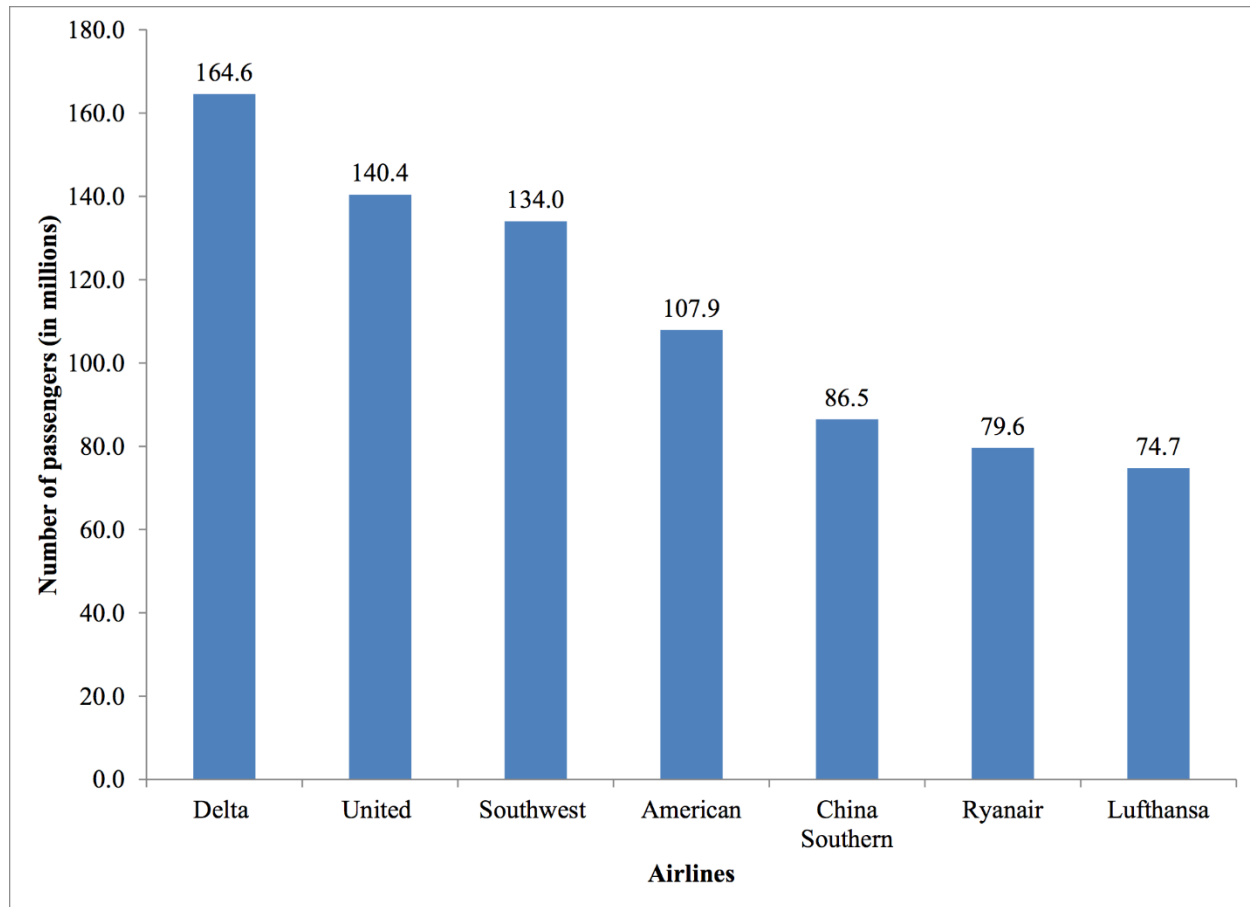
2.13

Airlines	Number of passengers (in millions)	Proportion	Degrees
Delta	164.6	$164.6/787.7 = 0.209$	75
United	140.4	0.178	64
Southwest	134.0	0.170	61
American	107.9	0.137	49
China Southern	86.5	0.110	40
Ryanair	79.6	0.101	36
Lufthansa	74.7	0.095	34
Totals	787.7	1.000	360

Pie chart:



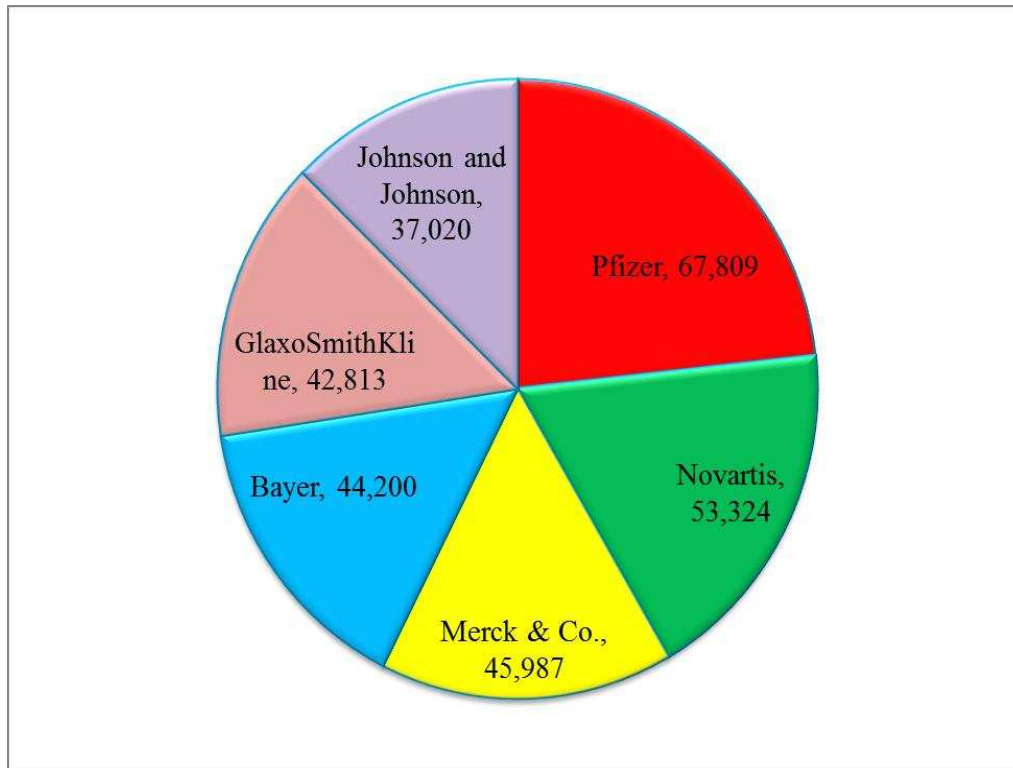
Bar chart:



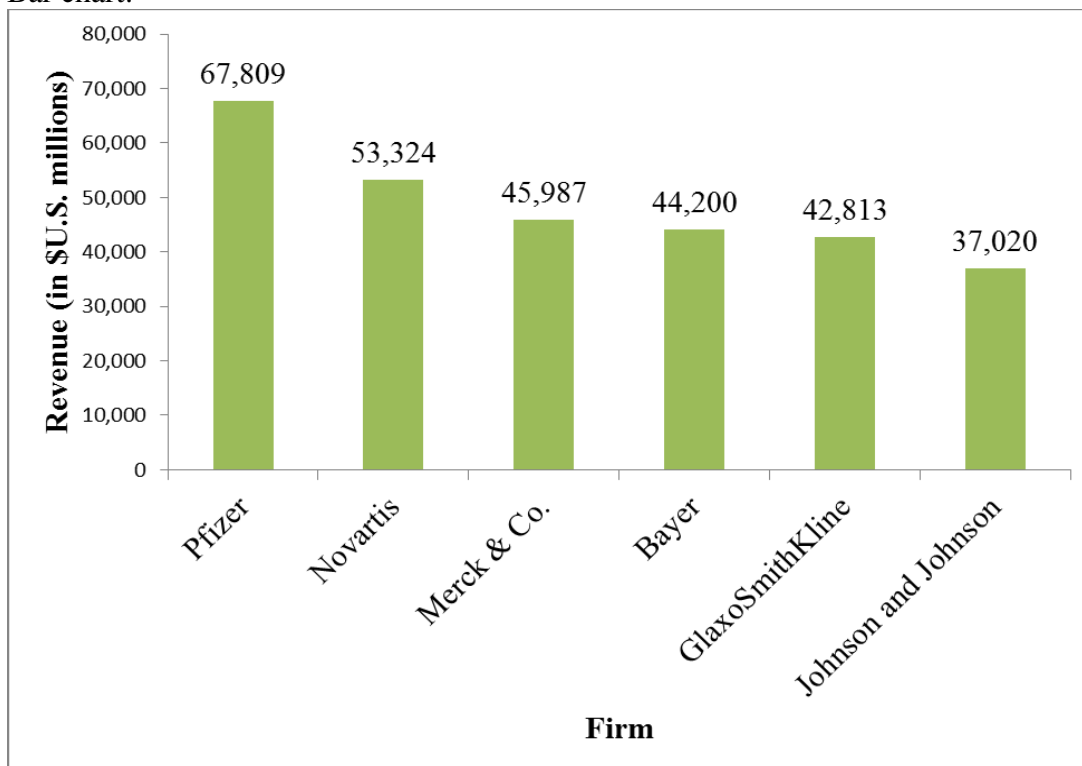
2.14

Firm	Revenue (\$U.S. millions)	Proportion	Degrees
Pfizer	67,809	$67,809/291,153=0.233$	84
Novartis	53,324	0.183	66
Merck & Co.	45,987	0.158	57
Bayer	44,200	0.152	55
GlaxoSmithKline	42,813	0.147	53
Johnson and Johnson	37,020	0.127	46
Totals	291,153	1.000	360

Pie chart:



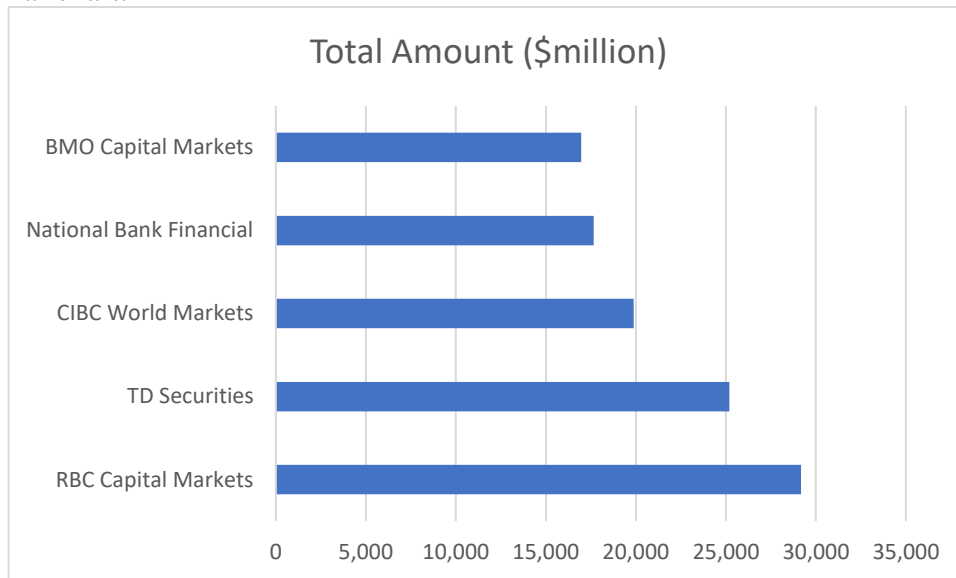
Bar chart:



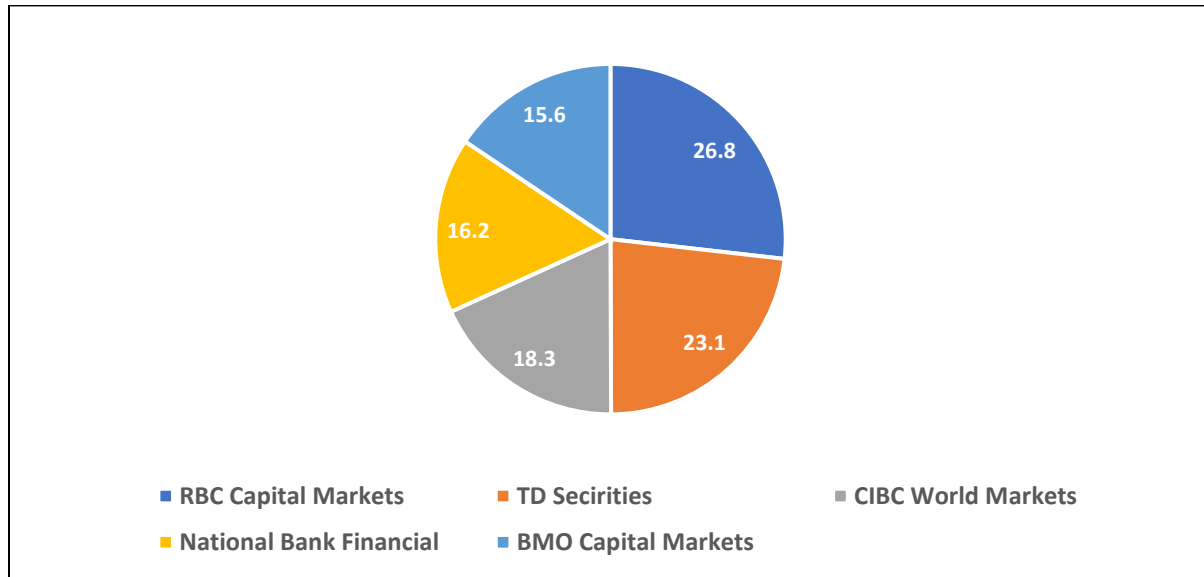
2.15

Underwriting Firm	Total Amount (\$million)	Percent
RBC Capital Markets	29,172	26.8
TD Securities	25,193	23.1
CIBC World Markets	19,875	18.3
National Bank Financial	17,653	16.2
BMO Capital Markets	16,958	15.6
Total	108,851	100.0

Bar chart:



Pie chart:

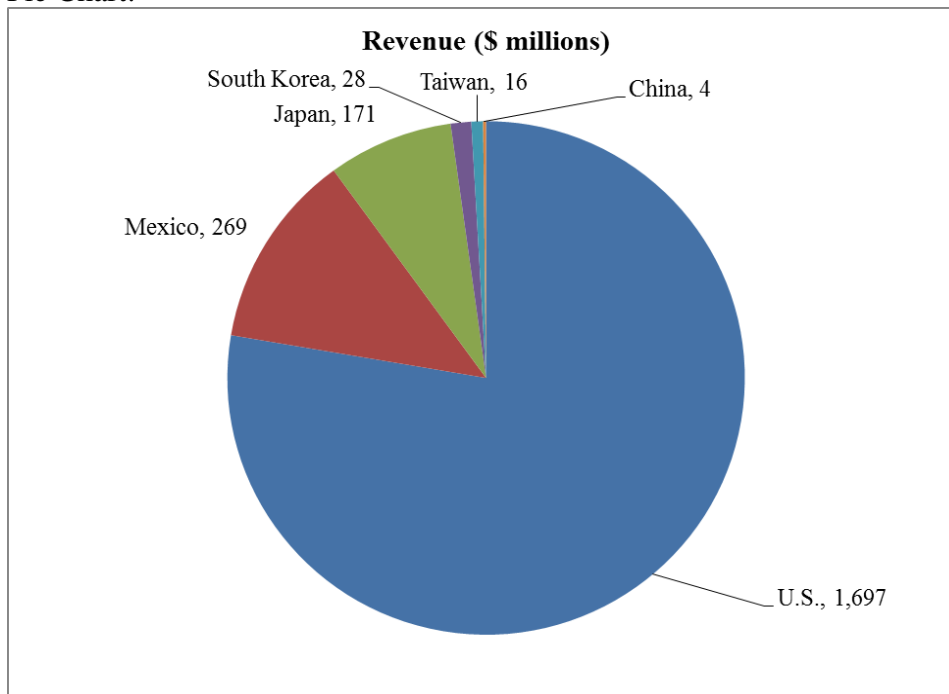


The proportion, sizes and color of the pie slices clearly shows that RBC Capital Markets raised the highest amount (\$29,172 million, 26.8%) and BMO Capital Markets raised the lowest amount (\$16,958 million, 15.8%).

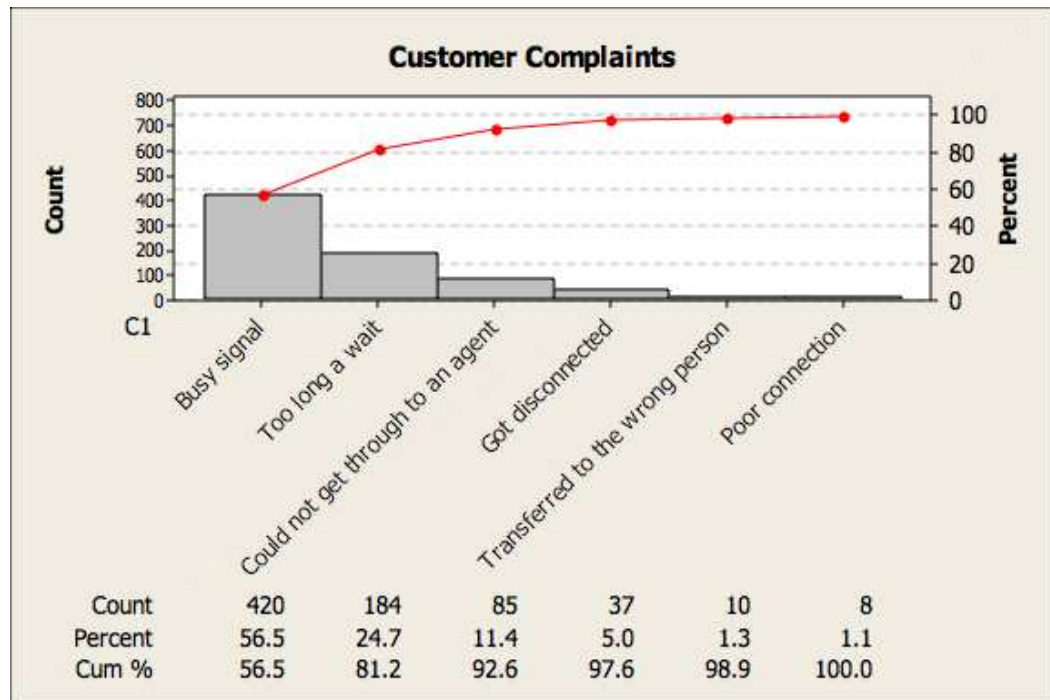
2. 16

Destination	Revenue (\$ millions)	Proportion	Degrees
U.S.	1,697	0.7767	279.61
Mexico	269	0.1231	44.32
Japan	171	0.0783	28.19
South Korea	28	0.0128	4.61
Taiwan	16	0.0073	2.63
China	4	0.0018	0.65
Totals	2,185	1.0000	360.0

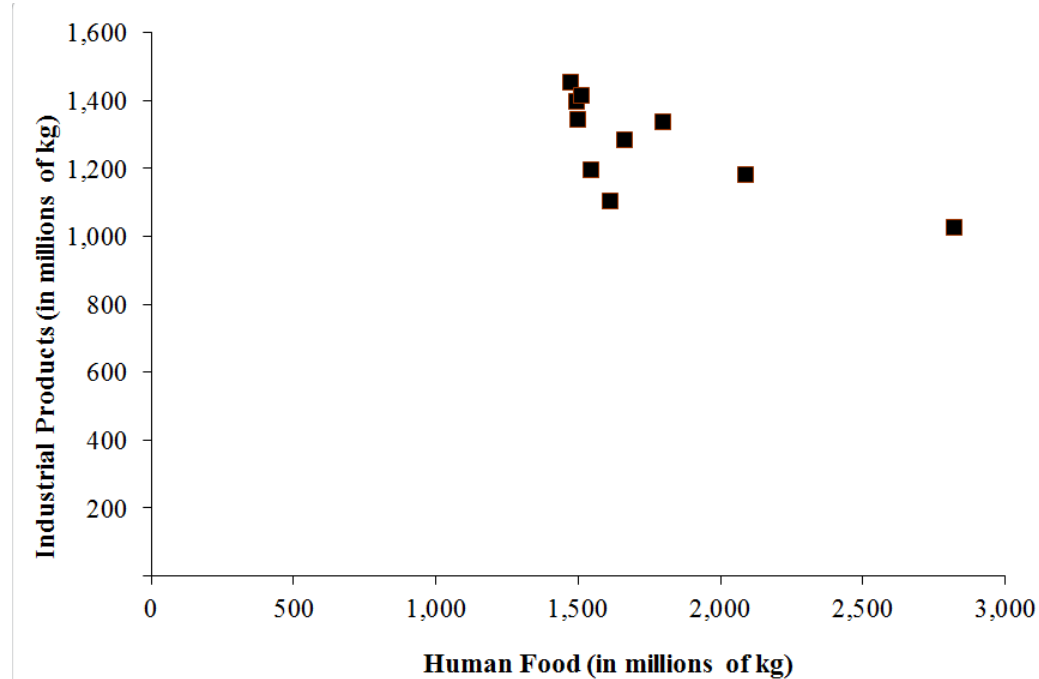
Pie Chart:



2.17	<u>Complaint</u>	<u>Number</u>	<u>% of Total</u>
	Busy Signal	420	56.45
	Too long a Wait	184	24.73
	Could not get through	85	11.42
	Got Disconnected	37	4.97
	Transferred to the Wrong Person	10	1.34
	Poor Connection	8	1.08
	Total	744	99.99

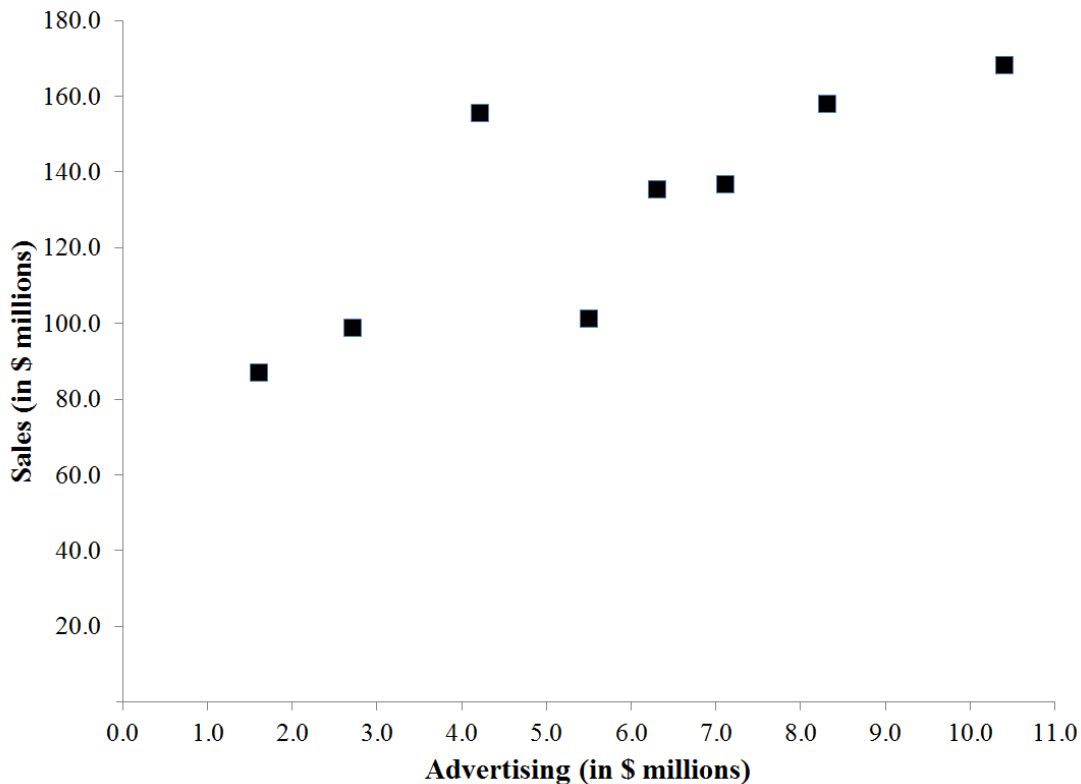


2.18



Generally speaking, the tendency is that less fish caught and used for industrial products when more fish is caught and used for human food.

2.19



Generally speaking, the tendency is that sales are higher when more money is spent on advertising.

2.20 It appears from the graph that as job satisfaction decreases, there is an increase in tardiness. Thus, there appears to be an inverse relationship between job satisfaction and tardiness. The scatter plot also shows that when employees are highly satisfied, the level of tardiness is low.

2.21

		One-Way Commute Distance (in km)			Total
		0 - 3	4 - 10	More than 10	
Number of Annual	0 - 2	95	184	117	396
Non-vacation -Day	3 - 5	21	40	53	114
Absences	More than 5	3	7	12	22
Total		119	231	182	532

There is a slight tendency for there to be a few more absences as plant workers commute further distances. Say, 6.6% of those who commute more than 10 km had more than 5 non-vacation absent days, as compared to 2.5% and 3% for those who commute 0-3 km and 4-10 km respectively. Comparing workers who travel 4-10 km to those who travel 0-3 km, there is about a 2:1 ratio in all three cells (0-2, 3-5, more than 5 non-vacation day absences) indicating that for these two categories (0-3 and 4-10), number of absences is essentially independent of commute distance.

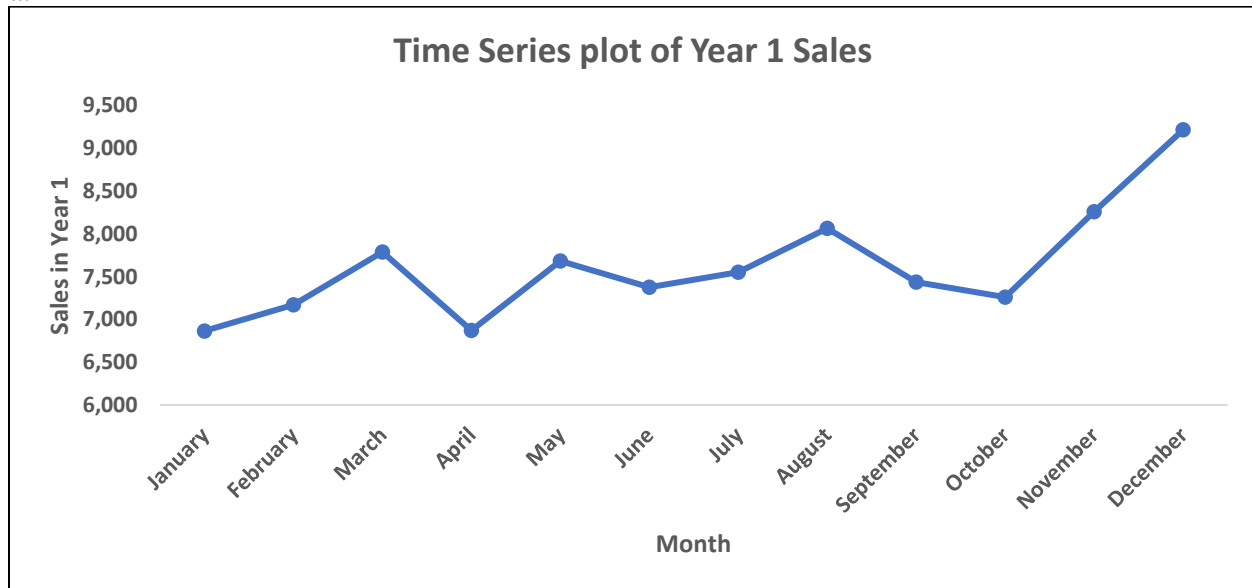
2.22

		Level of Education		Total
		High School only	University Degree	
Rating of Service	Acceptable	9	6	15
	Unacceptable	2	8	10
Total		11	14	25

There is a tendency that the lower the level of education, the more acceptable is the service. First of all, according to the table, the numbers of customers with different level of education are almost the same (11 with “high school only” and 14 with “university degree”). It appears that a much higher proportion of high school level education customers rate the service as “acceptable” than as “unacceptable” (9 to 2 ratio or about 4.5 times as many). On the other hand, more of the university degree educated customers rated the service as “unacceptable” than as “acceptable” (8 to 6 ratio). In addition, 60% or 9 customers out of 15 who rate the service as “acceptable” have only high school education, and 80% or 8 customers out of 10 who rate the service as “unacceptable” are university degree level educated.

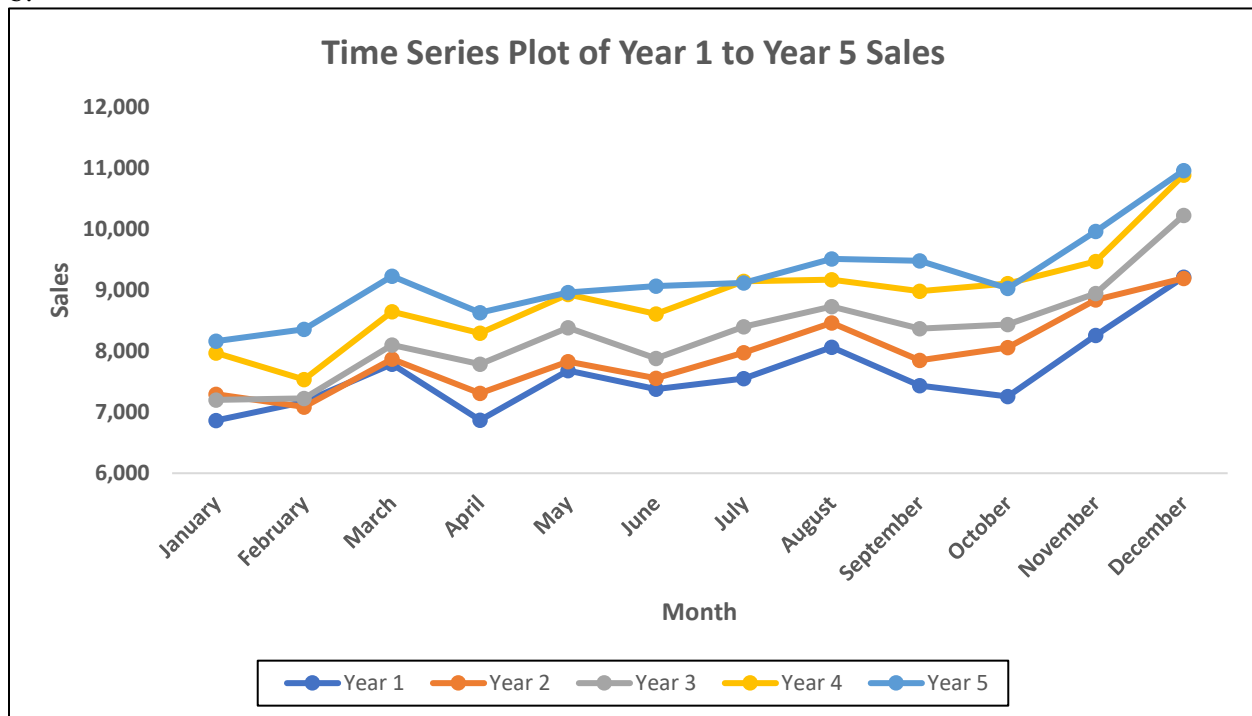
2.23

a.



This time-series plot for the year 1 shows up and down swings of sales until October with sales strongly increasing for November and December due to perhaps the holiday season. If true, we could be seeing possible seasonal effects of the holiday season followed by low sales in January.

b.



When placed together on one time series plot, the graphs tend to show the same shape virtually every year. This underscores the notion that there is some seasonality in the data. Sales are down

in January, April, June, and October. Sales are highest in December, increases every year from October through December, and then drops dramatically in January.

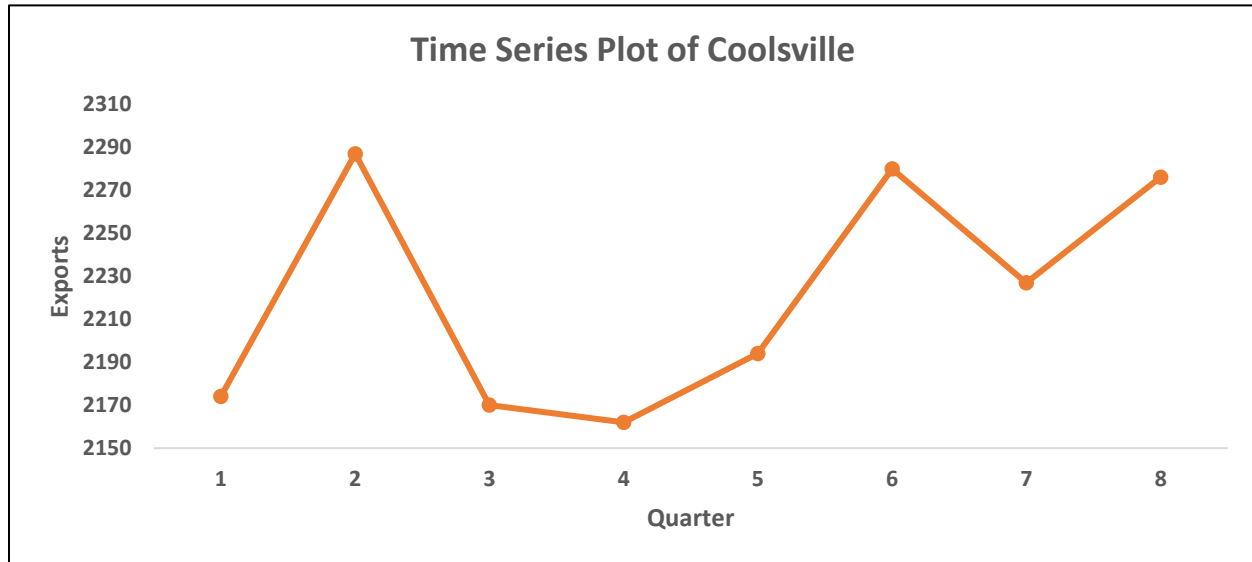
c.



With all five years of data shown chronologically, it is apparent that there is a repeating seasonal cycle within each year. In addition, however, there is a general trend of increasing sales over the five years.

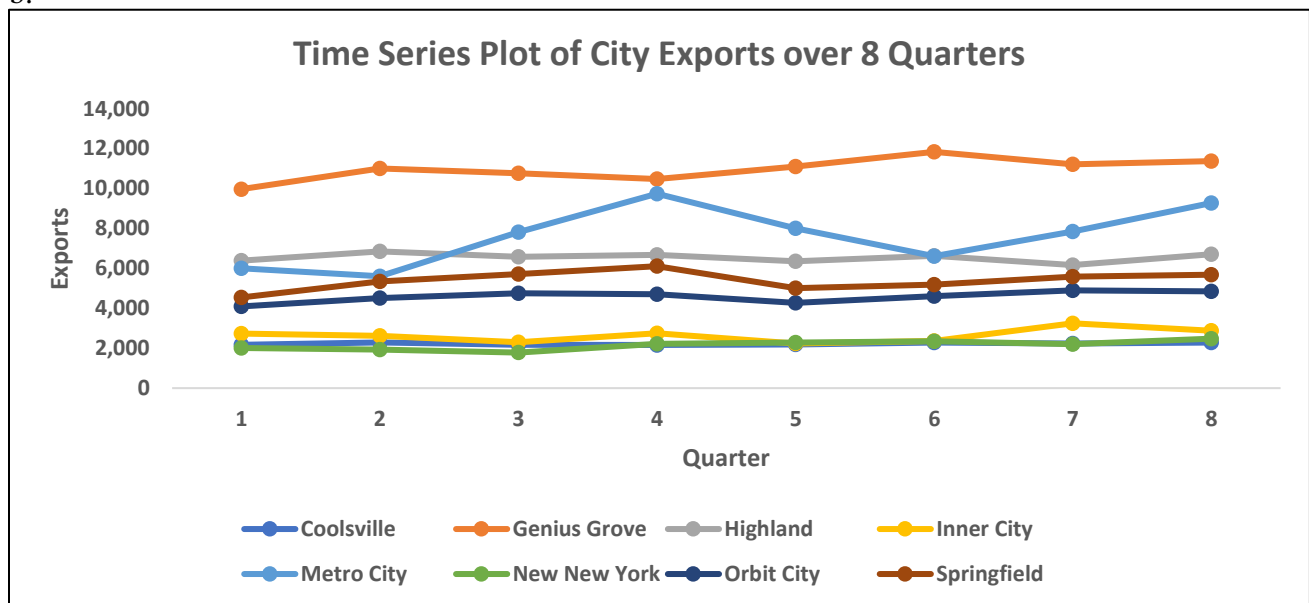
2.24

a.



Examining the exports for Coolsville over these two years, there was an increase in exports from quarter 1 to quarter 2, followed by a decline in quarter 3 and 4 in year 1. In year 2, there was an increase in exports from quarter 1 to 2 followed by a decline in quarter 3 and other increase in quarter 4. Overall, there seems to be little or no trend in exports over time just a pattern of ups and downs.

b.



Many of the cities actually have relatively constant exports over these 8 quarters with few significant ups and downs. Metro City (second highest at the end) seems to be the most volatile of the cities in terms of exports. Genius Grove consistently has the highest exports.

2.25

$$\text{Range} = \text{max} - \text{min} = 57 - 18 = 39$$

$$\text{Class width} = 39/6 = 6.5 \approx 7$$

<u>Class Interval</u>	<u>Frequencies</u>
16 - under 23	6
23 - under 30	9
30 - under 37	4
37 - under 44	4
44 - under 51	4
51 - under 58	<u>3</u>
TOTAL	30

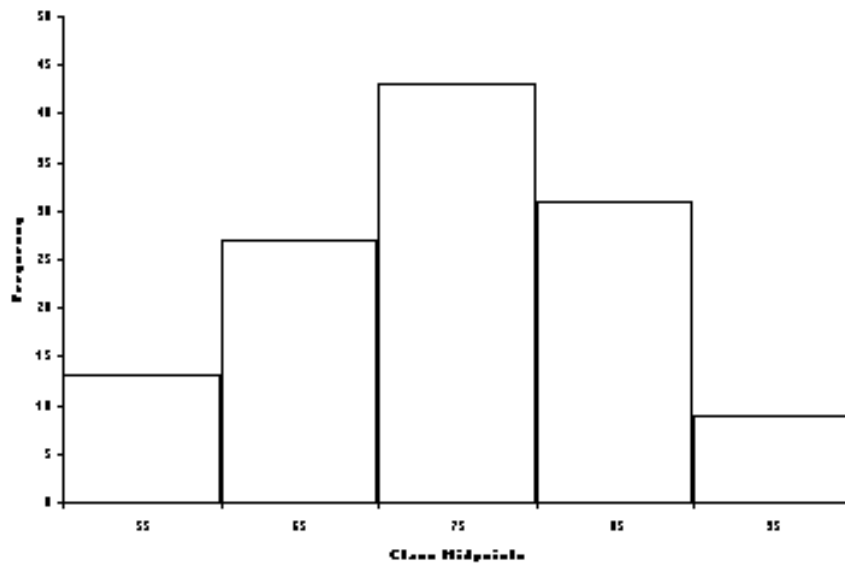
2.26

<u>Class Interval</u>	<u>Frequency</u>	<u>Midpoint</u>	<u>Rel. Freq.</u>	<u>Cum. Freq.</u>
20 - under 25	17	22.5	.207	17
25 - under 30	20	27.5	.244	37
30 - under 35	16	32.5	.195	53
35 - under 40	15	37.5	.183	68
40 - under 45	8	42.5	.098	76
45 - under 50	6	47.5	.073	82

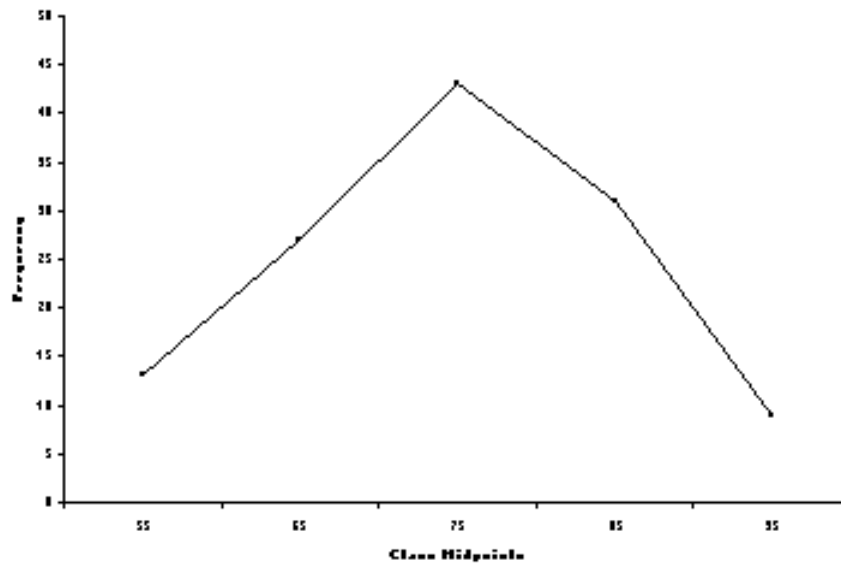
2.27

<u>Class Interval</u>	<u>Frequencies</u>
50 - under 60	13
60 - under 70	27
70 - under 80	43
80 - under 90	31
90 - under 100	<u>9</u>
TOTAL	123

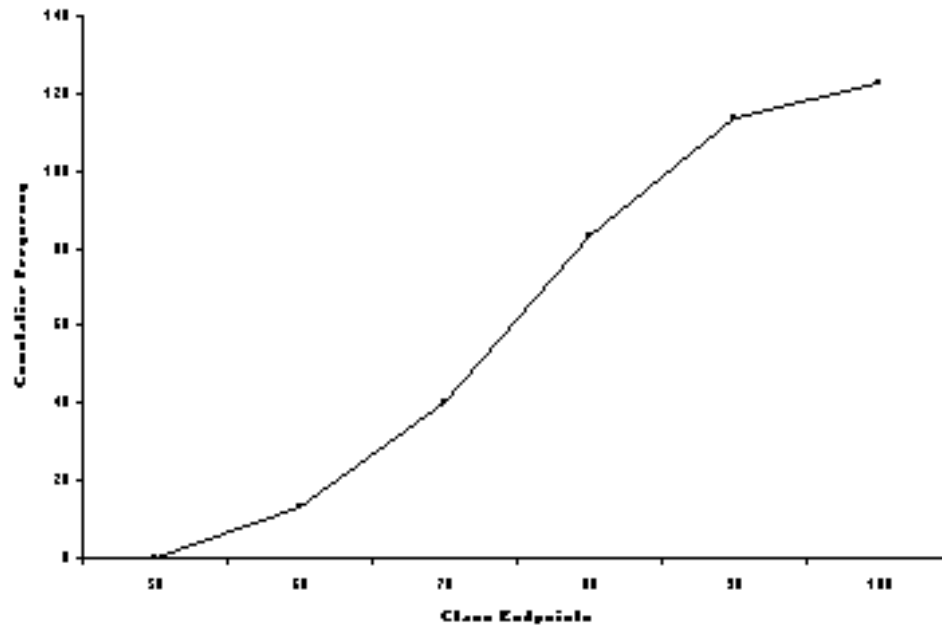
Histogram:



Frequency Polygon:



Ogive:



2.28

STEM	LEAF
28	4, 6, 9
29	0, 4, 8
30	1, 6, 8, 9
31	1, 2, 4, 6, 7, 7
32	4, 4, 6
33	5

2.29

<u>Label</u>	<u>Value</u>	<u>Proportion</u>	<u>Degrees</u>
A	55	.180	65
B	121	.397	143
C	83	.272	98
D	<u>46</u>	<u>.151</u>	<u>54</u>
TOTAL	305	1.000	360

Pie

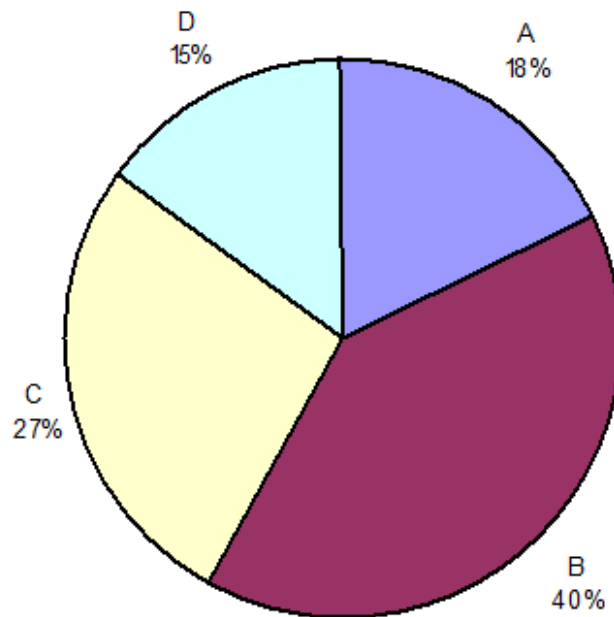
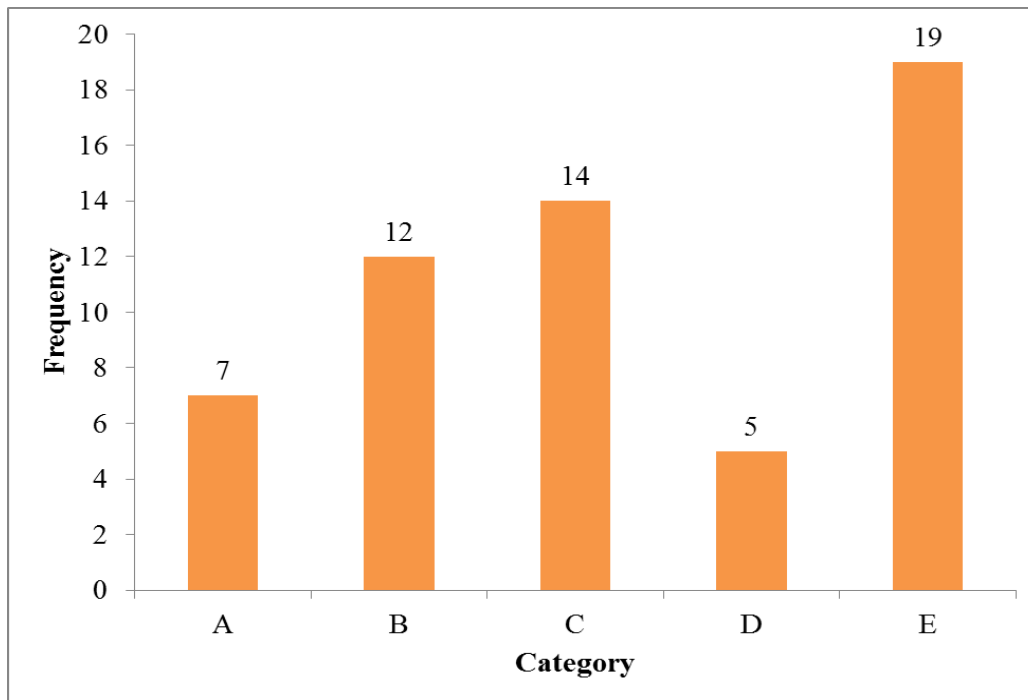


Chart:

2.30

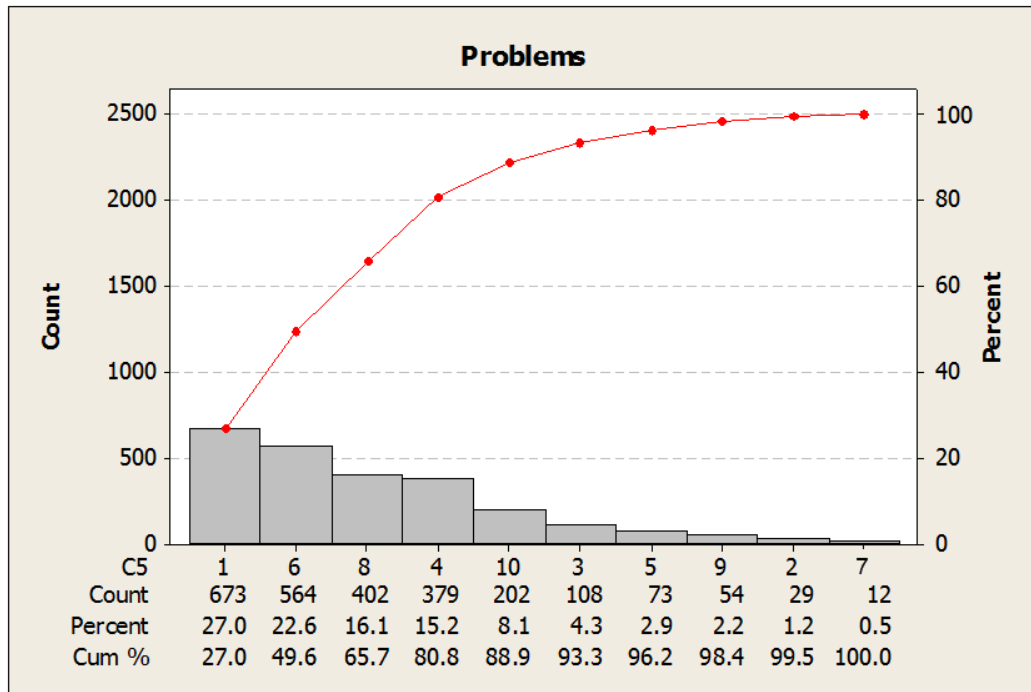
Bar Graph:

<u>Category</u>	<u>Frequency</u>
A	7
B	12
C	14
D	5
E	19

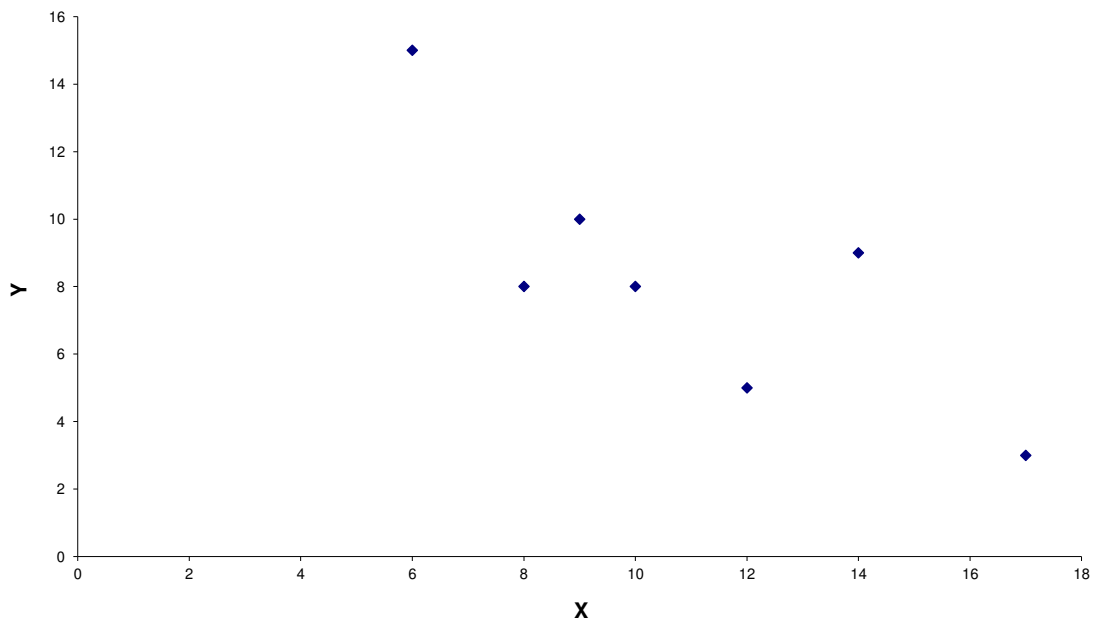


2.31	<u>Problem</u>	<u>Frequency</u>	<u>Percent of Total</u>
	1	673	26.96
	2	29	1.16
	3	108	4.33
	4	379	15.18
	5	73	2.92
	6	564	22.60
	7	12	0.48
	8	402	16.11
	9	54	2.16
	10	<u>202</u>	8.09
		2496	

Pareto Chart:



2.32



2.33 Yellowknife Steel Company

$$\text{Range} = \text{max} - \text{min} = 69 - 36 = 33$$

$$\text{Class width} = 33/8 = 4.125 \approx 5$$

<u>Class Interval</u>	<u>Frequency</u>
32 - under 37	1
37 - under 42	4
42 - under 47	12
47 - under 52	11
52 - under 57	14
57 - under 62	5
62 - under 67	2
67 - under 72	<u>1</u>
TOTAL	50

The highest frequencies are between 42 and 57.

2.34

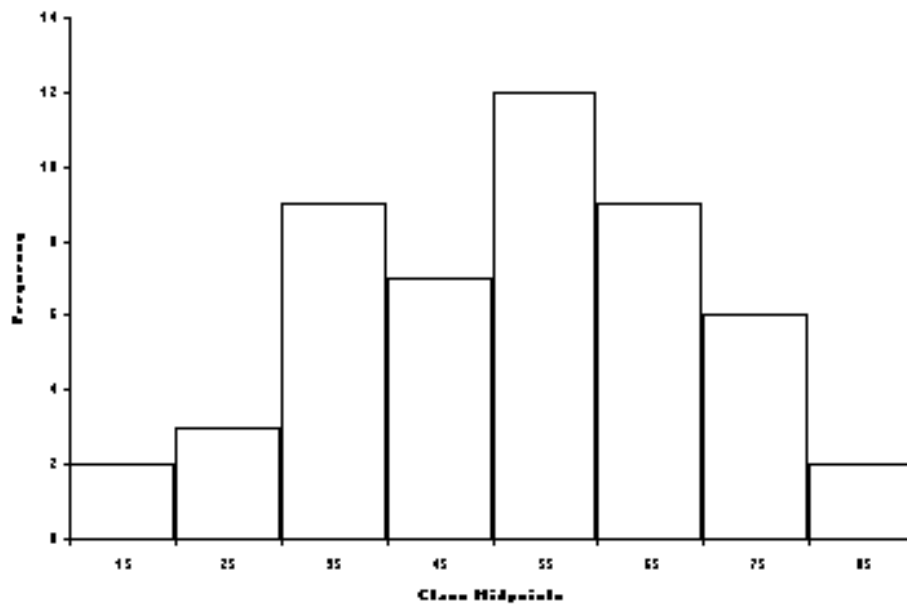
<u>Class Interval</u>	<u>Frequency</u>	<u>Class Midpoint</u>	<u>Relative Frequency</u>	<u>Cumulative Frequency</u>
20 – 25	8	22.5	$8/53 = .1509$	8
25 – 30	6	27.5	.1132	14
30 – 35	5	32.5	.0943	19
35 – 40	12	37.5	.2264	31
40 – 45	15	42.5	.2830	46
45 – 50	<u>7</u>	47.5	<u>.1321</u>	53
TOTAL	53		.9999 ≈ 1	

2.35 Frequency Distribution:

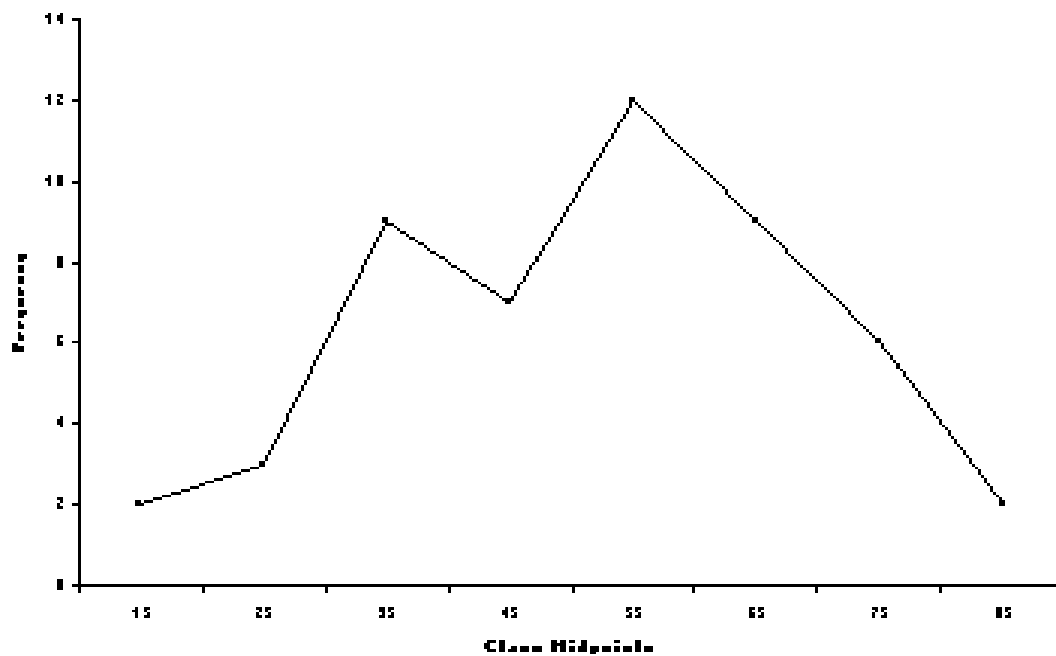
<u>Class Interval</u>	<u>Frequency</u>
10 - under 20	2
20 - under 30	3
30 - under 40	9
40 - under 50	7
50 - under 60	12
60 - under 70	9
70 - under 80	6
80 - under 90	<u>2</u>

50

Histogram:



Frequency Polygon:

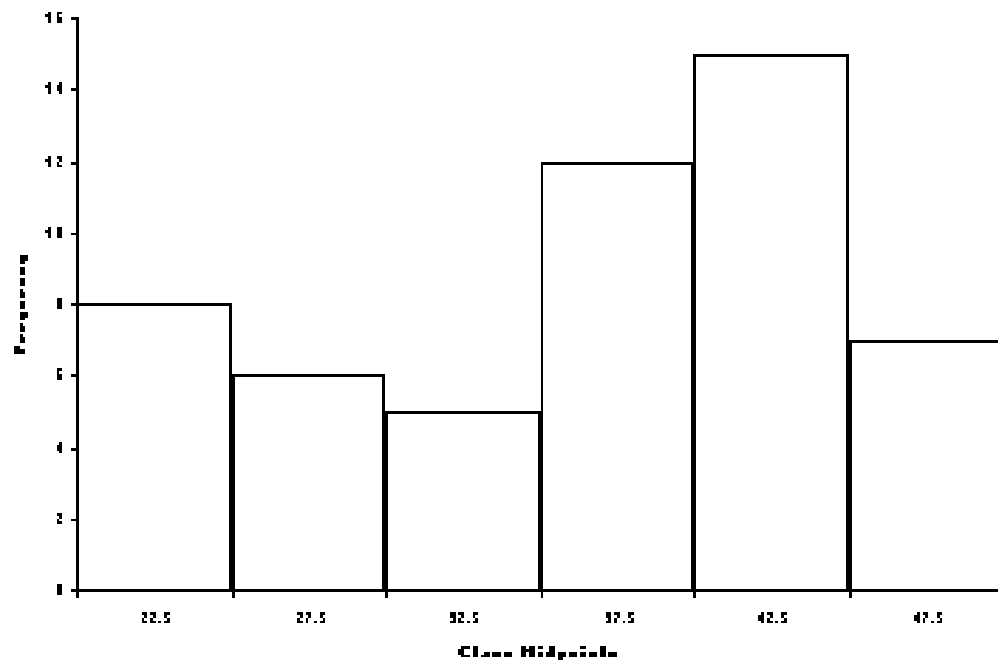


The normal distribution appears to peak near the center and diminish towards the end intervals.

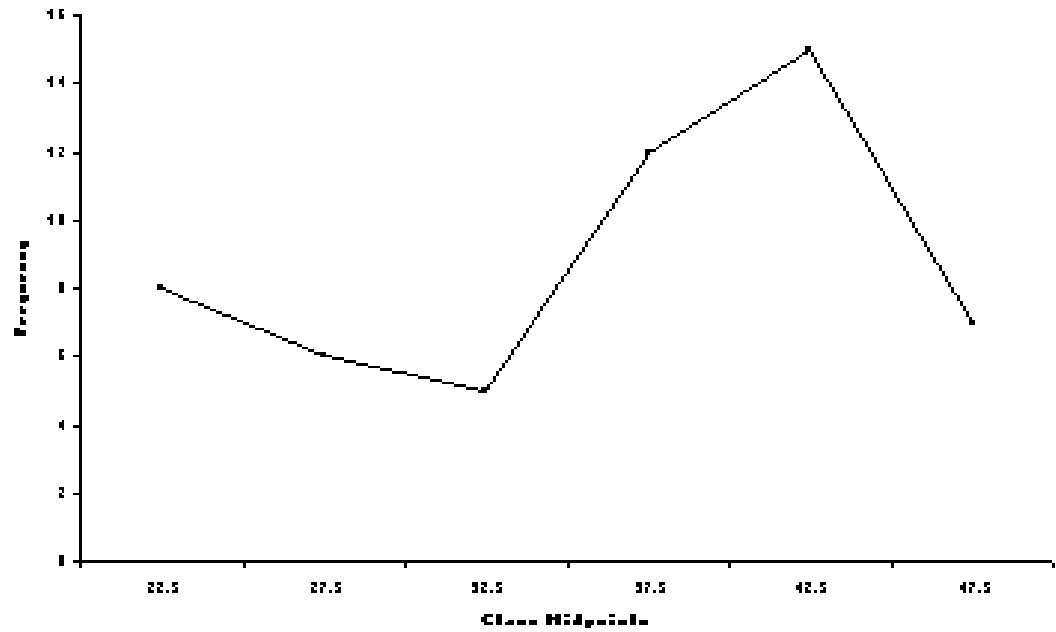
2.36 a. Histogram and a Frequency Polygon for Problem 2.34

<u>Class Interval</u>	<u>Frequency</u>	<u>Cumulative Frequency</u>
20 – 25	8	8
25 – 30	6	14
30 – 35	5	19
35 – 40	12	31
40 – 45	15	46
45 – 50	7	53
TOTAL	53	

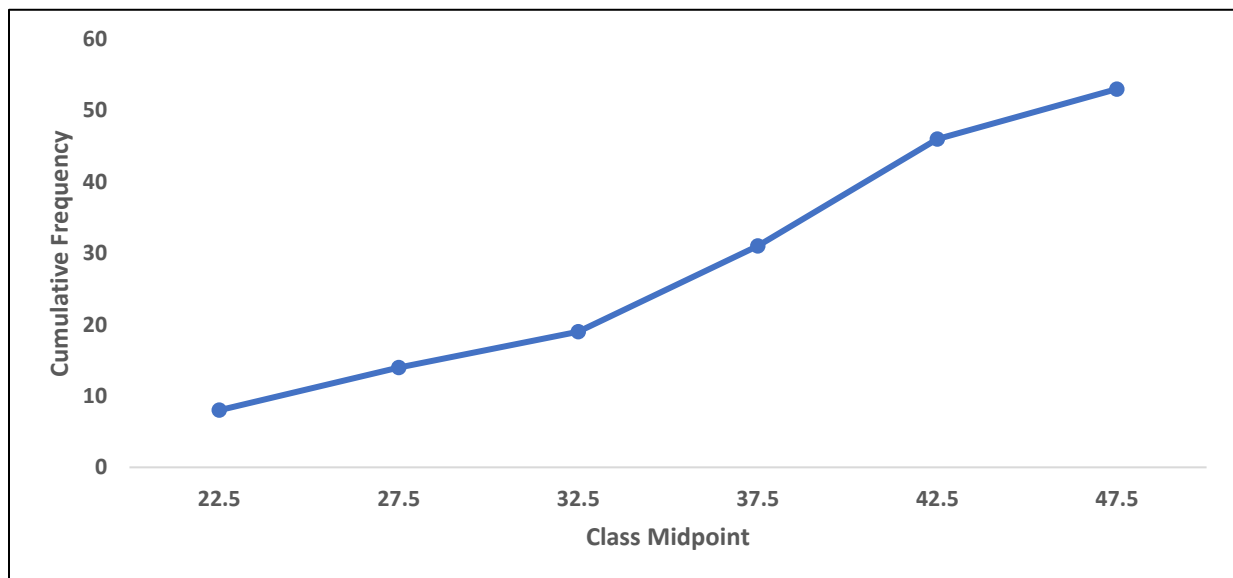
Histogram:



Frequency Polygon:



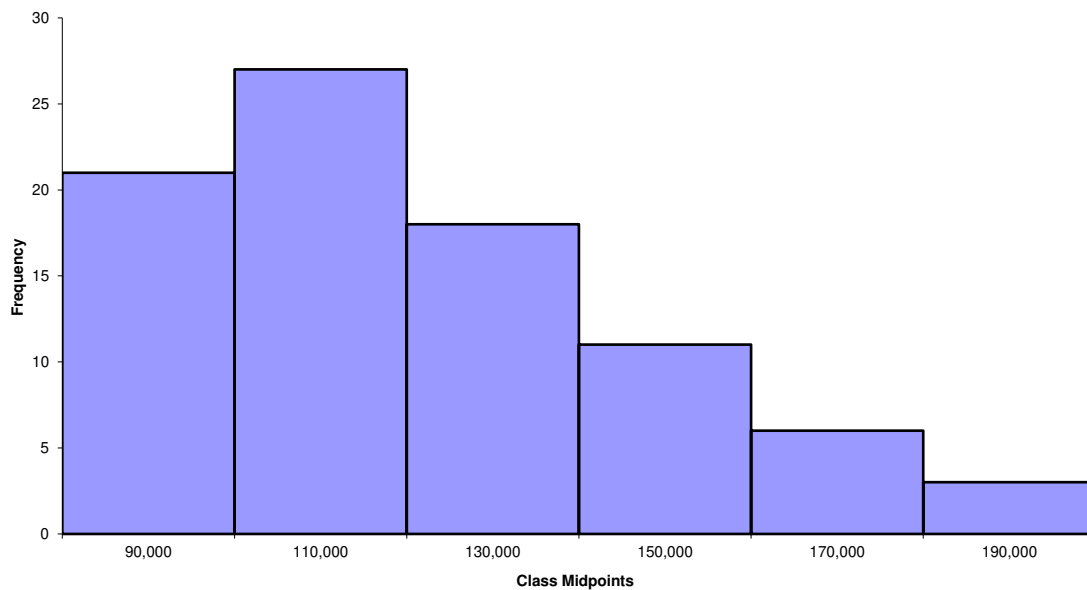
b. Ogive:



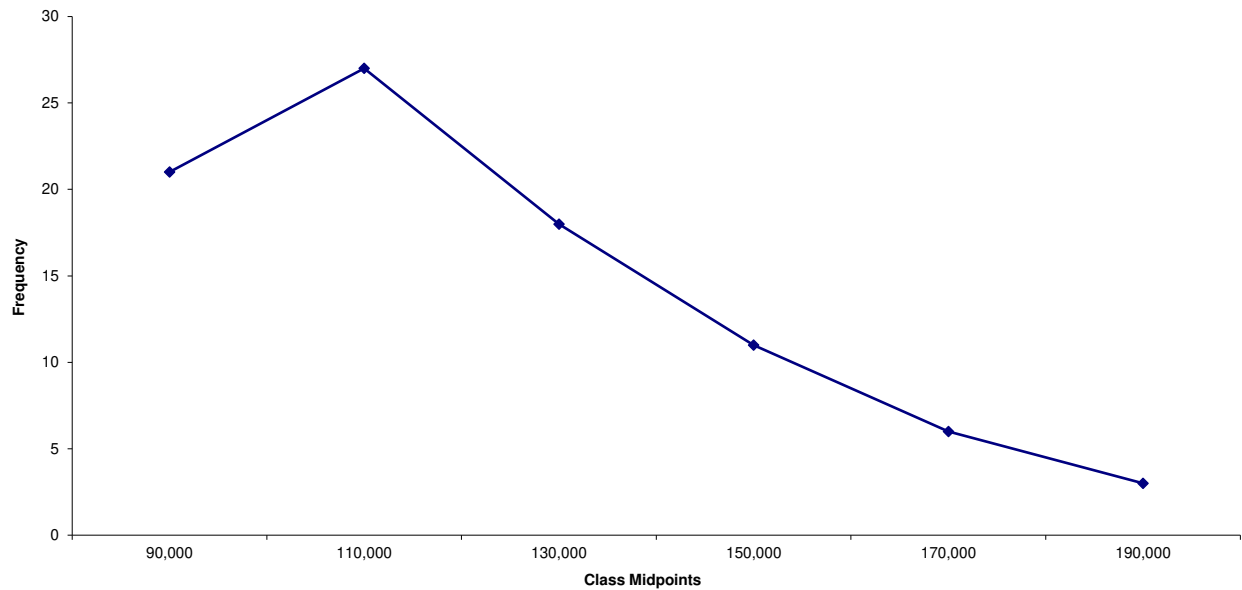
2.37

<u>Asking Price</u>	<u>Frequency</u>	<u>Cumulative Frequency</u>
\$ 80,000 - under \$ 100,000	21	21
\$ 100,000 - under \$ 120,000	27	48
\$ 120,000 - under \$ 140,000	18	66
\$ 140,000 - under \$ 160,000	11	77
\$ 160,000 - under \$ 180,000	6	83
\$ 180,000 - under \$ 200,000	<u>3</u>	86
	86	

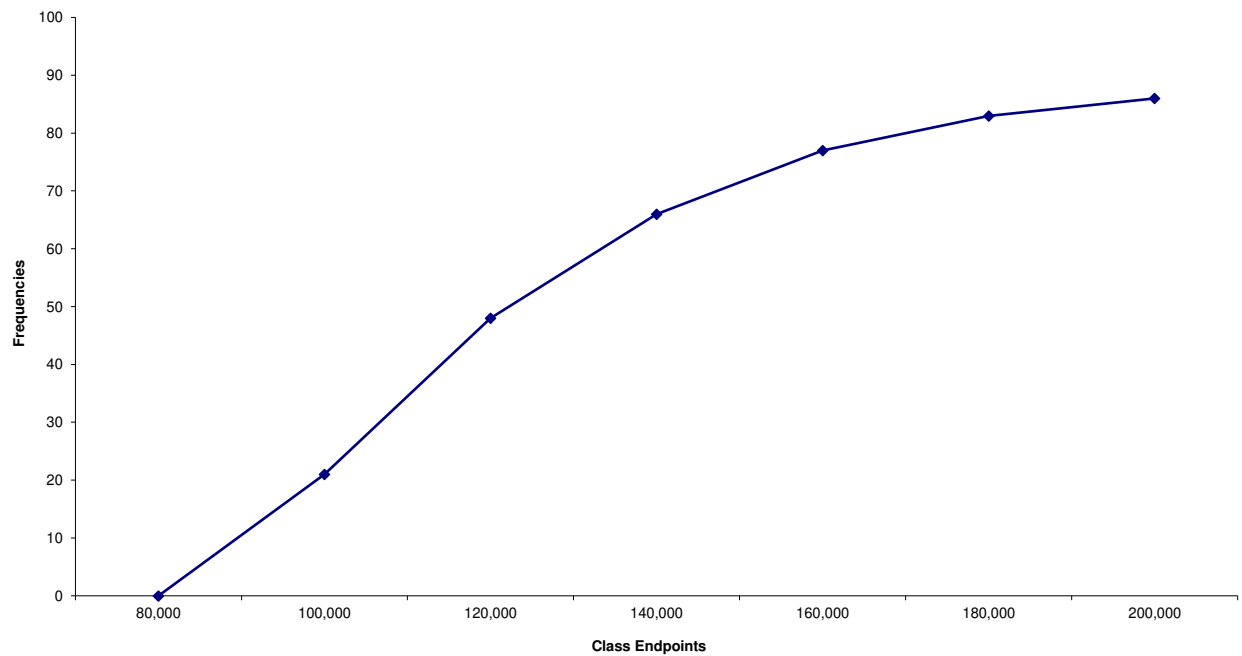
Histogram:



Frequency Polygon:



Ogive:



2.38 Stem and Leaf Plot:

STEM	LEAF
1	2, 3, 6, 7, 8, 8, 8, 9, 9
2	0, 3, 4, 5, 6, 7, 8

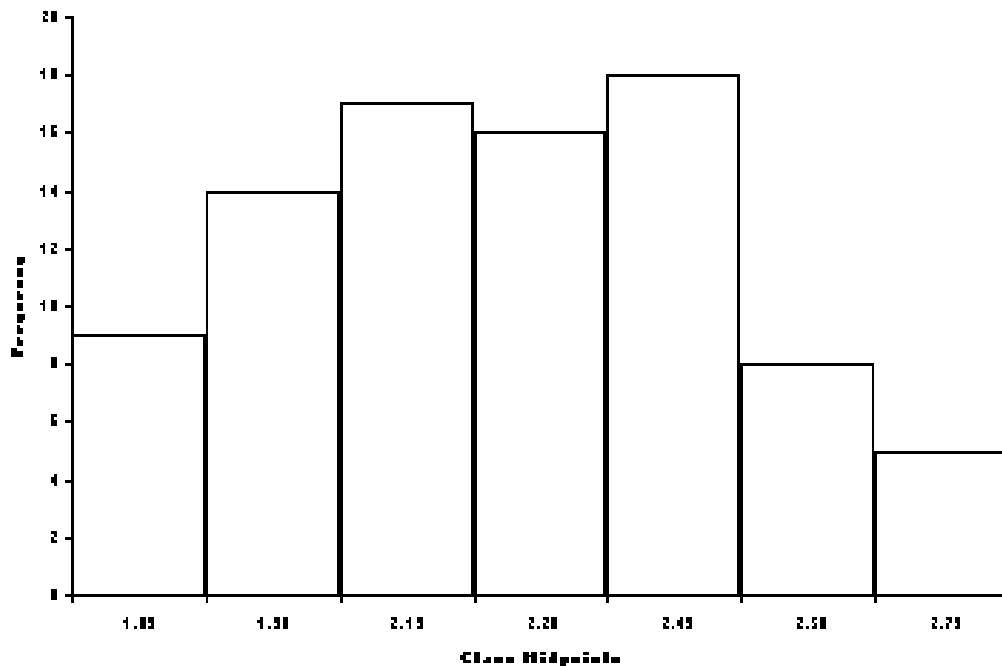
3 0, 1, 2, 2

Comments: The stem and leaf plot shows that the travel times are relatively evenly spread out between 12 days and 32 days. It also shows that the most travel times fall in the 12 to 19 day interval followed by the 20 to 28 day interval. Only four of the travel times were thirty or more days; 18 days is the most frequently occurring travel time (occurred three times).

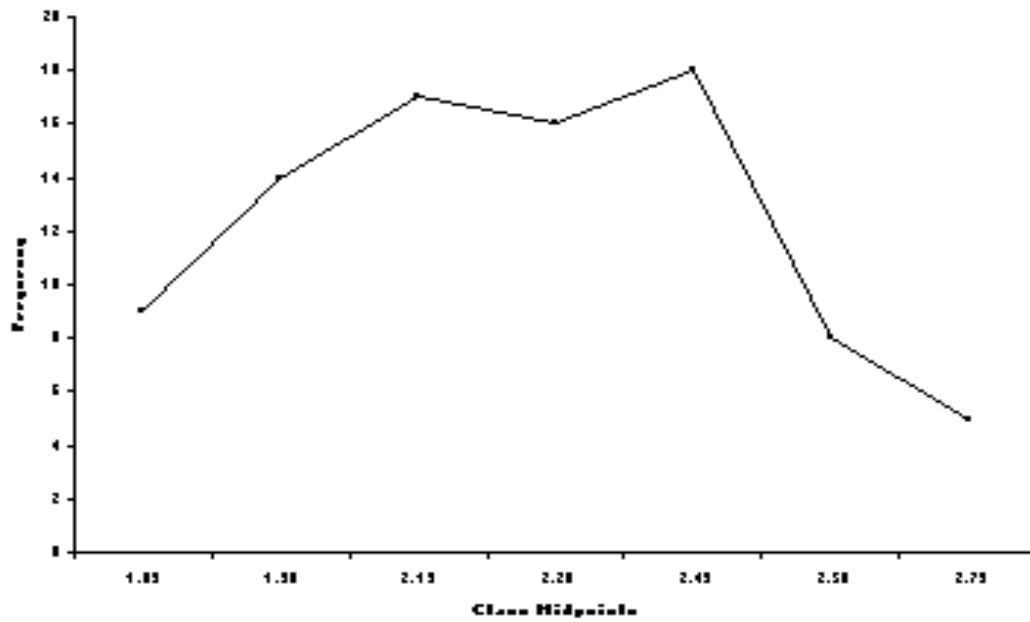
2.39

<u>Price</u>	<u>Frequency</u>	<u>Cumulative Frequency</u>
\$1.75 - under \$1.90	9	9
\$1.90 - under \$2.05	14	23
\$2.05 - under \$2.20	17	40
\$2.20 - under \$2.35	16	56
\$2.35 - under \$2.50	18	74
\$2.50 - under \$2.65	8	82
\$2.65 - under \$2.80	<u>5</u>	87
	87	

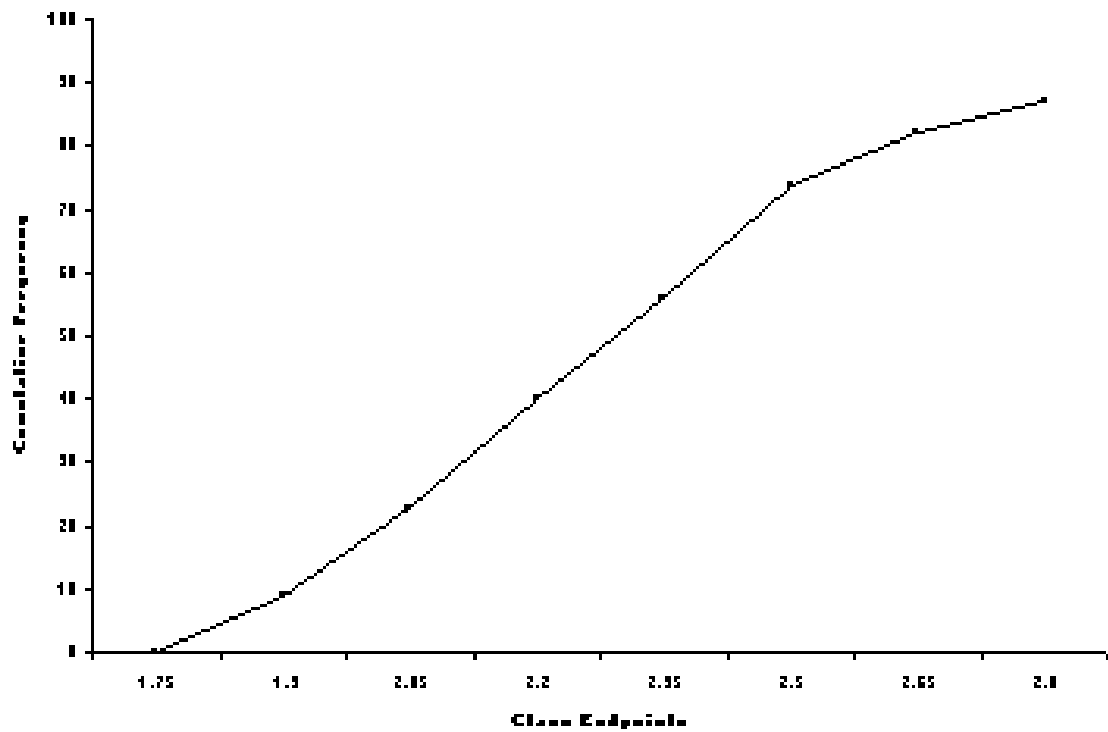
Histogram:



Frequency Polygon:

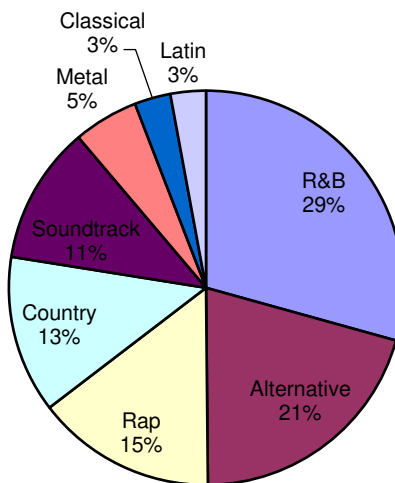


Ogive:

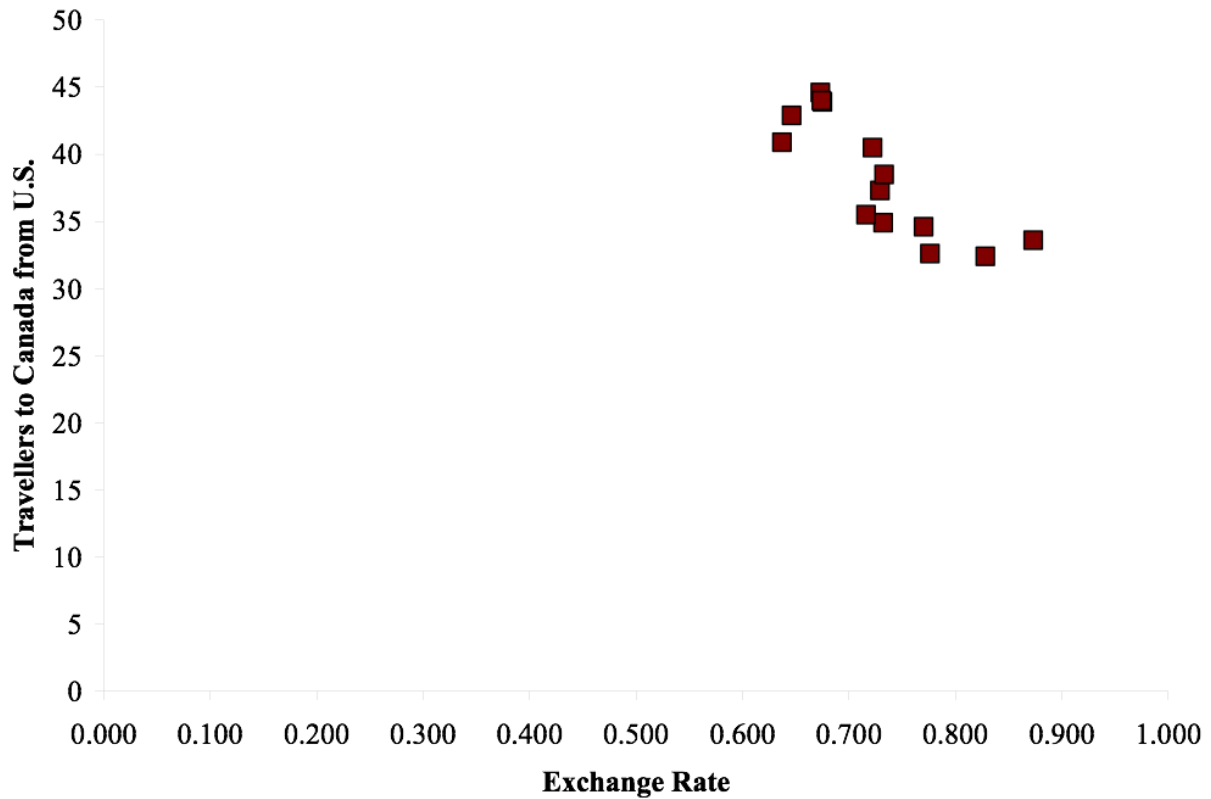
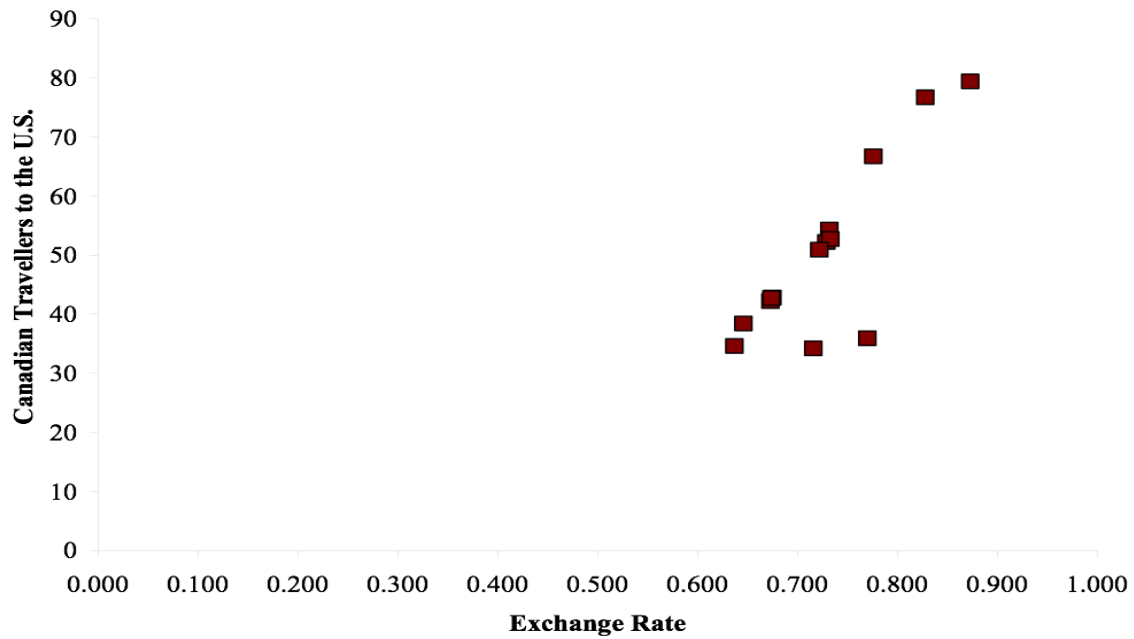


2.40	<u>Genre</u>	<u>Albums Sold</u>	<u>Proportion</u>	<u>Degrees</u>
	R&B	146.4	.29	104.4
	Alternative	102.6	.21	75.6
	Rap	73.7	.15	54.0
	Country	64.5	.13	46.8
	Soundtrack	56.4	.11	39.6
	Metal	26.6	.05	18.0
	Classical	14.8	.03	10.8
	Latin	14.5	.03	10.8
	TOTAL		1.00	360.0

Pie Chart:



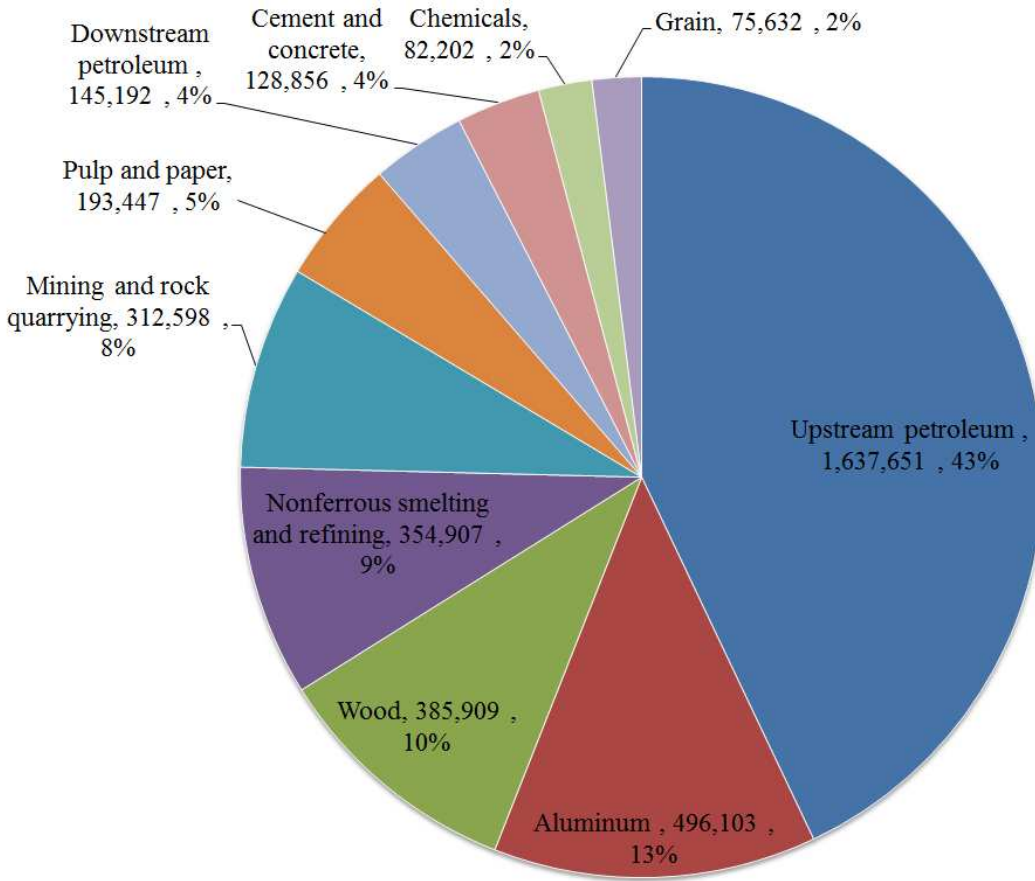
2.41



The higher is the exchange rate more Canadian travellers go to the U.S. and less the U.S. travellers do to Canada.

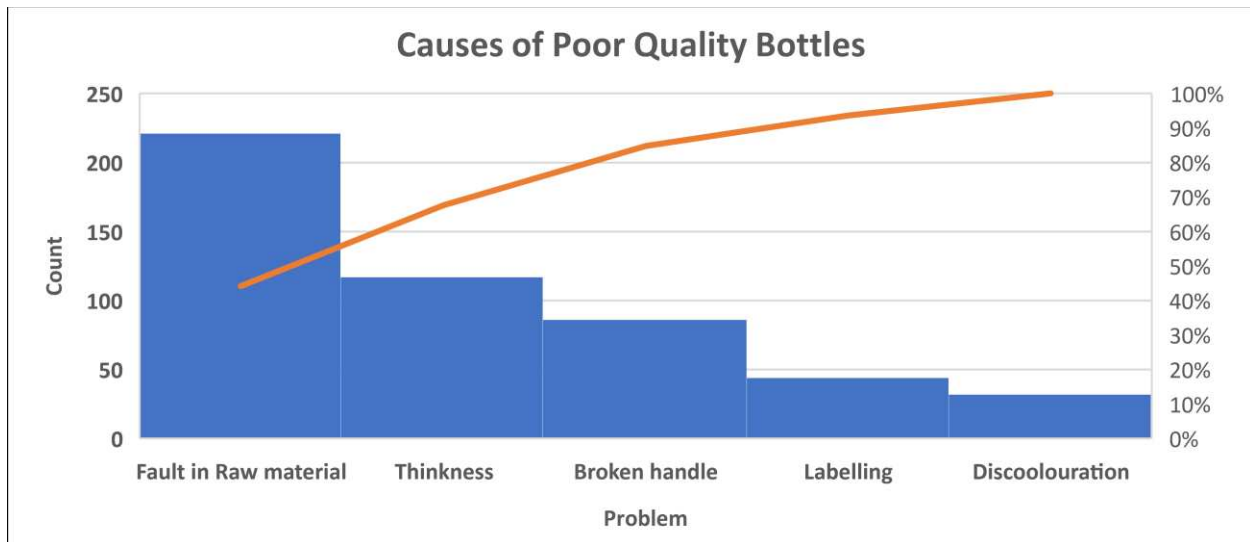
2.42

Industrial Source	Total Emissions (tonnes)	Proportion	Degrees
Upstream petroleum	1,637,651	0.430	154.6
Aluminum	496,103	0.130	46.8
Wood	385,909	0.101	36.4
Nonferrous smelting and refining	354,907	0.093	33.5
Mining and rock quarrying	312,598	0.082	29.5
Pulp and paper	193,447	0.051	18.3
Downstream petroleum	145,192	0.038	13.7
Cement and concrete	128,856	0.034	12.2
Chemicals	82,202	0.022	7.8
Grain	75,632	0.020	7.1
Totals	3,812,497	1.0000	360.0



2.43

The Pareto chart indicates that fault in raw material causes 44.2% of the defects and becomes the major problem. According to the chart, 23.4% of the bottles were rejected because of incorrect thickness which can be identified as the second severe problem. The steepest slopes correspond to “fault in raw material”, “thickness”, and “broken handle” categories. They represent 84.8% causes of poor-quality bottles.



2.44	STEM	LEAF
	42	12, 16, 24, 32, 99, 99
	43	04, 28, 39, 46, 61, 88
	44	20, 40, 59
	45	12
	46	53, 54
	47	30, 34, 58
	48	22, 34, 66, 78
	49	63
	50	48, 49, 90
	51	66
	52	21, 54, 57, 63, 91
	53	38, 66, 66
	54	31, 78
	55	56
	56	69
	57	37, 50
	58	31, 32, 58, 73
	59	19, 23

2.45

STEM	LEAF
92	00, 68

93	01, 37, 44, 75
94	05, 37, 48, 60, 68
95	24, 55
96	02, 56, 70, 77
97	42, 60, 64
98	14, 30
99	22, 61, 75, 76, 90, 96
100	02, 10

2.46 The histogram shows that all household incomes fall between \$30,000 and 140,000. Since the distribution is almost bell-shaped, the data are approximately normally distributed. The centre of the histogram is located left of \$100,000 and indicates that the average household income of mall shoppers is about \$90,000. The heights of the tallest three rectangles (around 30 observations each) show most of the mall shoppers have household income between \$75,000 and \$105,000. There are no outliers.

2.47 Family practice is most prevalent - followed by pediatrics
 . A virtual tie exists between ob/gyn, general surgery, anesthesiology, and psychiatry.

2.48 Stem-and-leaf plots can be thought of as histograms that are produced by the observations themselves, rather than by columns or bars representing the values. Since we have the observations themselves, stem-and-leaf plots often illustrate more detailed information than equivalent histograms.

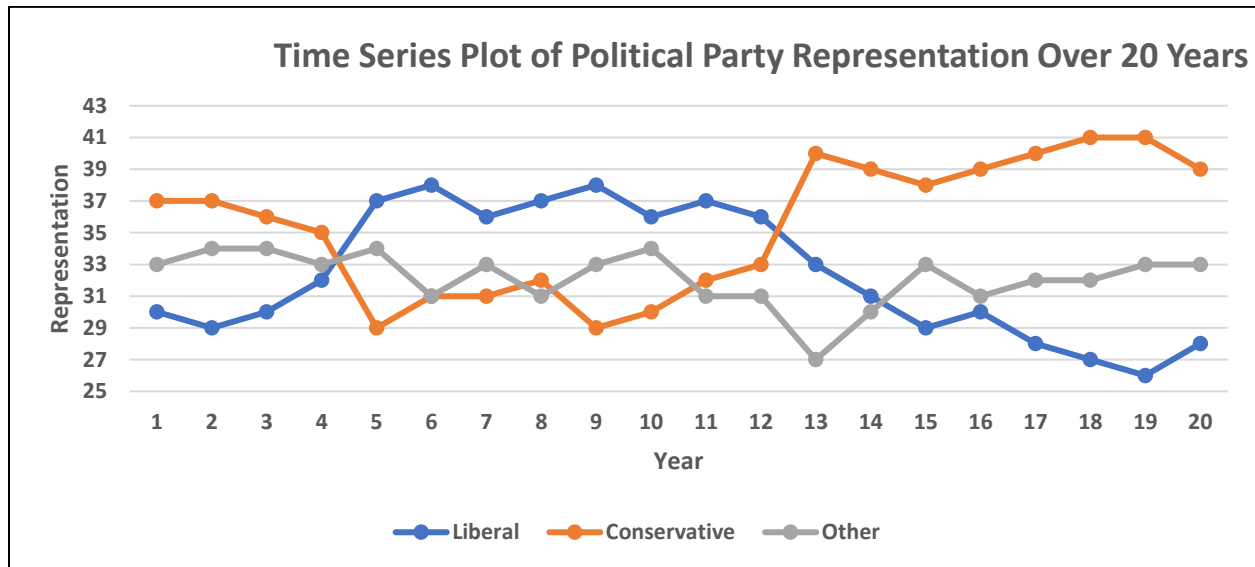
In this stem-and-leaf plot of the audits for 100 CPA firms, we can observe that the range is 30 audits (the largest value of 42 minus the smallest value of 12). The observations are concentrated on the top stems between 12-19 audits, with other concentrations for 26-29 audits and also for 34-35 audits. The stem-and-leaf plot does not show a clear central tendency, indeed, one could say that the plots is skewed since the density of observations is higher on the top stems.

2.49 Ogives are useful when the analyst wants to see running totals (or cumulative frequencies). Thus, ogives are useful when identifying steep increases in frequencies. In this ogive, particularly steep increases occur in November and December, indicating large jumps (the largest being in

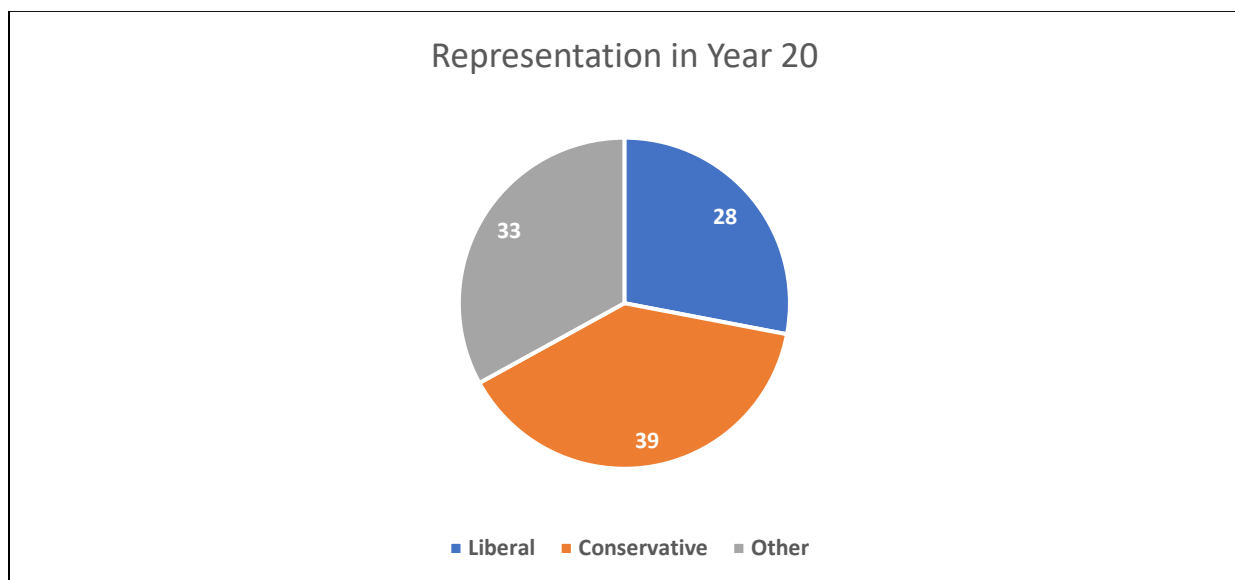
December) in monthly frequency totals, likely due to the holidays season. Notice also a steep increase in April, can you explain why this might be?

2.50

This is a time series, cross-sectional data. One can examine trends in representation using time series plots. Pie chart or bar chart can also be used to visualize representation for each year.

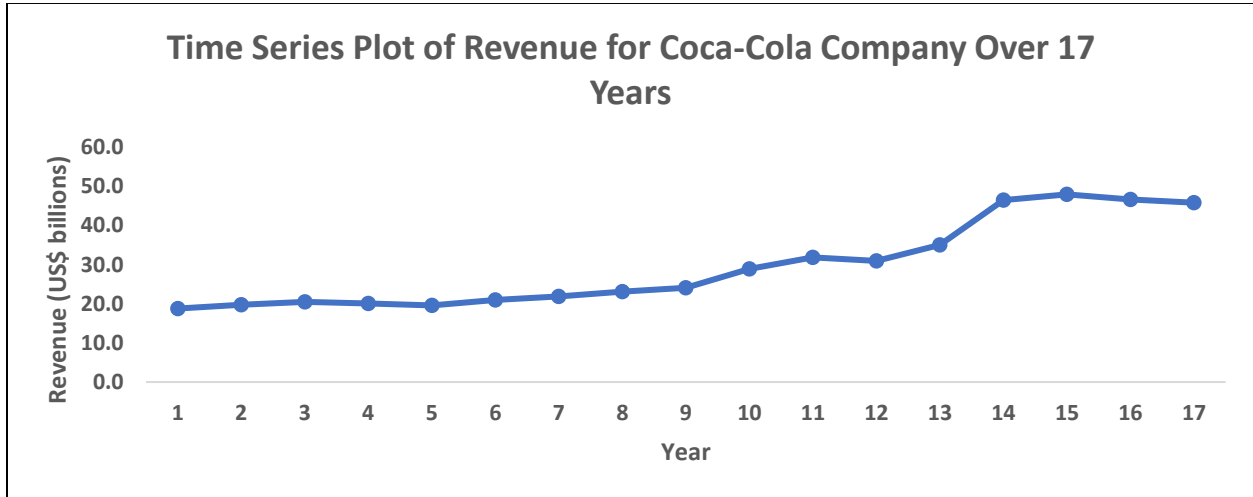


Conservative had greater representation in the first 4 year while the Liberal Party had greater representation from year 5 through year 12. Over the last 8 years the Conservative Party had the majority representation.



The sizes and color of the pie slices clearly shows that the Conservative Party had majority representation in year 20 followed by the Liberal Party.

2.51



This time-series plot shows that the company enjoyed revenue growth from year 1 to year 15. Revenue appears to be declining after year 15.

Legal Notice

Copyright

WILEY

Copyright © 2020 by John Wiley & Sons Canada, Ltd. or related companies. All rights reserved.

The data contained in these files are protected by copyright. This manual is furnished under licence and may be used only in accordance with the terms of such licence.

The material provided herein may not be downloaded, reproduced, stored in a retrieval system, modified, made available on a network, used to create derivative works, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise without the prior written permission of John Wiley & Sons Canada, Ltd.