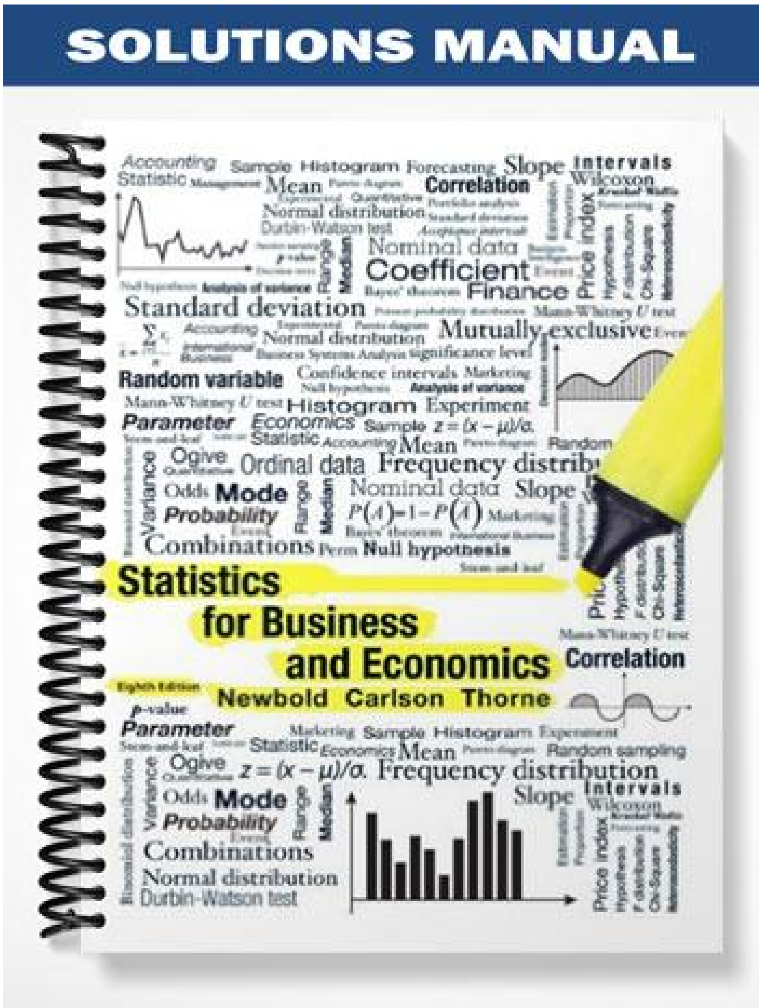

SOLUTIONS MANUAL
Statistics
for Business
and Economics
Eighth Edition
Newbold Carlson Thorne

Chapter 2:

Describing Data: Numerical

2.1

Cruise agency – number of weekly specials to the Caribbean: 20, 73, 75, 80, 82

a. Compute the mean, median and mode

$$\bar{x} = \frac{\sum x_i}{n} = \frac{330}{5} = 66$$

median = middlemost observation = 75

mode = no unique mode exists

b. The median best describes the data due to the presence of the outlier of 20. This skews the distribution to the left. The agency should first check to see if the value '20' is correct.

2.2

Number of complaints: 8, 8, 13, 15, 16

a. Compute the mean number of weekly complaints:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{60}{5} = 12$$

b. Calculate the median = middlemost observation = 13

c. Find the mode = most frequently occurring value = 8

2.3

CPI percentage growth forecasts: 3.0, 3.1, 3.4, 3.4, 3.5, 3.6, 3.7, 3.7, 3.7, 3.9

a. Compute the sample mean: $\bar{x} = \frac{\sum x_i}{n} = \frac{35}{10} = 3.5$

b. Compute the sample median = middlemost observation: $\frac{3.5 + 3.6}{2} = 3.55$

c. Mode = most frequently occurring observation = 3.7

2.4

Department store % increase in dollar sales: 2.9, 3.1, 3.7, 4.3, 5.9, 6.8, 7.0, 7.3, 8.2, 10.2

a. Calculate the mean number of weekly complaints: $\bar{x} = \frac{\sum x_i}{n} = \frac{59.4}{10} = 5.94$

b. Calculate the median = middlemost observation: $\frac{5.9 + 6.8}{2} = 6.35$

2.5 Percentage of total compensation derived from bonus payments: 10.2, 13.1, 15, 15.8, 16.9, 17.3, 18.2, 24.7, 25.3, 28.4, 29.3, 34.7

a. Median % of total compensation from bonus payments =

$$\frac{17.3+18.2}{2}=17.75$$

b. Mean % $\bar{x}=\frac{\sum x_i}{n}=\frac{248.9}{12}=20.7417$

2.6

Daily sales (in hundreds of dollars): 6, 7, 8, 9, 10, 11, 11, 12, 13, 14

a. Find the mean, median, and mode for this store

Mean = $\bar{x}=\frac{\sum x_i}{n}=\frac{101}{10}=10.1$

Median = middlemost observation = $\frac{10+11}{2}=10.5$

Mode = most frequently occurring observation = 11

b. Find the five-number summary

Q1 = the value located in the $0.25(n + 1)^{th}$ ordered position
= the value located in the 2.75^{th} ordered position
= 7 + 0.25(8 –7) = 7.25

Q3 = the value located in the $0.75(n + 1)^{th}$ ordered position
= the value located in the 8.25^{th} ordered position
= 12 + 0.75(13 –12) = 12.75

Minimum = 6

Maximum = 14

Five - number summary:

minimum < Q1 < median < Q3 < maximum

6 < 7.25 < 10.5 < 12.75 < 14

2.7

Find the measures of central tendency for the number of imperfections in a sample of 50 bolts

Mean number of imperfections = $\frac{0(35)+1(10)+2(3)+3(2)}{50}$ = 0.44 imperfections per bolt

Median = 0 (middlemost observation in the ordered array)

Mode = 0 (most frequently occurring observation)

2.8

Ages of 12 students: 18, 19, 21, 22, 22, 22, 23, 27, 28, 33, 36, 36

a. $\bar{x}=\sum\frac{x_i}{n}=\frac{307}{12}=25.58$

b. Median = 22.50

c. Mode = 22

2.9

a. First quartile, Q1 = the value located in the 0.25(n + 1)th ordered position
= the value located in the 39.25th ordered position
= 2.98 + 0.25(2.98 –2.99) = 2.9825

Third quartile, Q3 = the value located in the 0.75(n + 1)th ordered position
= the value located in the 117.75th ordered position
= 3.37 + 0.75(3.37 –3.37) = 3.37

b. 30th percentile = the value located in the 0.30(n + 1)th ordered position
= the value located in the 47.1th ordered position
= 3.10 + 0.1(3.10 –3.10) = 3.10

80th percentile = the value located in the 0.80(n + 1)th ordered position
= the value located in the 125.6th ordered position
= 3.39 + 0.6(3.39 –3.39) = 3.39

2.10

a. $\bar{x} = \sum \frac{x_i}{n} = \frac{282}{33} = 8.545$

b. Median = 9.0

c. The distribution is slightly skewed to the left since the mean is less than the median.

d. The five-number summary

Q1 = the value located in the 0.25(n + 1)th ordered position
= the value located in the 8.5th ordered position
= 6 + 0.5(6 – 6) = 6

Q3 = the value located in the 0.75(n + 1)th ordered position
= the value located in the 25.5th ordered position
= 10 + 0.5(11 –10) = 10.5

Minimum = 2

Maximum = 21

Five - number summary:

minimum < Q1 < median < Q3 < maximum

2 < 6 < 9 < 10.5 < 21

2.11

a. $\bar{x} = \sum \frac{x_i}{n} = \frac{23{,}699}{100} = 236.99$. The mean volume of the random sample of 100 bottles (237 mL) of a new suntan lotion was 236.99 mL.

b. Median = 237.00

c. The distribution is symmetric. The mean and median are nearly the same.

d. The five-number summary

Q1 = the value located in the 0.25(n + 1)th ordered position
= the value located in the 25.25th ordered position
= 233 + 0.25(234 – 233) = 233.25

Q3 = the value located in the 0.75(n + 1)th ordered position
= the value located in the 75.75th ordered position
= 241 + 0.75(241 –241) = 241

Minimum = 224
Maximum = 249
Five - number summary:
minimum < Q1 < median < Q3 < maximum
224 < 233.25 < 237 < 241 < 249

2.12

The variance and standard deviation are

x_i	DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)$	SQUARED DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)^2$
6	–1	1
8	1	1
7	0	0
10	3	9
3	–4	16
5	–2	4
9	2	4
8	1	1
$\sum_{i=1}^{8} x_i = 56$	$\sum_{i=1}^{8}\left(x_i - \bar{x}\right) = 0$	$\sum_{i=1}^{8}\left(x_i - \bar{x}\right)^2 = 36$

Sample mean = $\bar{x} = \dfrac{\sum_{i=1}^{8} x_i}{n} = \dfrac{56}{8} = 7$

Sample variance = $s^2 = \dfrac{\sum_{i=1}^{8}\left(x_i - \bar{x}\right)^2}{n-1} = \dfrac{36}{8-1} = 5.143$

Sample standard deviation = s = $\sqrt{s^2}$ = $\sqrt{5.143}$ = 2.268

2.13

The variance and standard deviation are

x_i	DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)$	SQUARED DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)^2$
3	0.5	0.25
0	–2.5	6.25
–2	–4.5	20.25
–1	–3.5	12.25
5	2.5	6.25
10	7.5	56.25
$\sum_{i=1}^{6} x_i = 15$	$\sum_{i=1}^{6}\left(x_i - \bar{x}\right) = 0$	$\sum_{i=1}^{6}\left(x_i - \bar{x}\right)^2 = 101.5$

Sample mean = $\bar{x} = \frac{\sum_{i=1}^{6} x_i}{n} = \frac{15}{6} = 2.5$

Sample variance = $s^2 = \frac{\sum_{i=1}^{6} \left(x_i - \bar{x}\right)^2}{n-1} = \frac{101.5}{5} = 20.3$

Sample standard deviation = s = $\sqrt{s^2}$ = $\sqrt{20.3}$ = 4.5056

2.14

x_i	DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)$	SQUARED DEVIATION ABOUT THE MEAN, $\left(x_i - \bar{x}\right)^2$
10	1	1
8	–1	1
11	2	4
7	–2	4
9	0	0
$\sum_{i=1}^{5} x_i = 45$	$\sum_{i=1}^{5} \left(x_i - \bar{x}\right) = 0$	$\sum_{i=1}^{5} \left(x_i - \bar{x}\right)^2 = 10$

Sample mean = $\bar{x} = \frac{\sum_{i=1}^{5} x_i}{n} = \frac{45}{5} = 9$

Sample variance = $s^2 = \frac{\sum_{i=1}^{5} \left(x_i - \bar{x}\right)^2}{n-1} = \frac{10}{4} = 2.5$

Sample standard deviation = $s = \sqrt{s^2} = \sqrt{2.5} = 1.581$

Coefficient of variation = $CV = \frac{s}{\bar{x}} \times 100\% = \frac{1.581}{9} \times 100\% = 17.57\%$

2.15

Minitab Output:

Descriptive Statistics: Ex2.15

```
Variable   Mean  SE Mean  StDev  Variance  CoefVar  Minimum     Q1  Median
Ex2.15    28.77     2.15  12.70    161.36    44.15    12.00  18.00   27.00

Variable     Q3  Maximum
Ex2.15    38.00    65.00
```

a. Mean = 2.15
b. Standard deviation = 12.70
c. CV = 44.15

2.16

Minitab Output

Stem-and-Leaf Display: Ex2.16

```
Stem-and-leaf of Ex2.16  N  = 35
Leaf Unit = 1.0

3    1  234
10   1  5577889
17   2  0012333
(4)  2  7799
14   3  1
13   3  557788
7    4  002
4    4  59
2    5  3
1    5
1    6
1    6  5
```

IQR = $Q_3 - Q_1$

Q_1 = the value located in the $0.25(35 + 1)^{th}$ ordered position
= the value located in the 9^{th} ordered position
= 18

Q_3 = the value located in the $0.75(35 + 1)^{th}$ ordered position
= the value located in the 27^{th} ordered position
= 38

IQR = $Q_3 - Q_1$ = 38 – 18 = 20 years

2.17

Mean = 75, variance = 25, $\sigma = \sqrt{\sigma^2} = \sqrt{25} = 5$

Using the mean of 75 and the standard deviation of 5, we find the following interval:

$\mu \pm 2\sigma = 75 \pm (2*5) = (65, 85)$, hence we have k = 2

a. According to the Chebyshev's theorem, proportion must be at least $100[1-(1/k^2)]\%$ = $100[1-(1/2^2)]\%$ = 75%. Therefore, approximately 75% of the observations are between 65 and 85

b. According to the empirical rule, approximately 95% of the observations are between 65 and 85

2.18

Mean = 250, σ = 20

a. To determine *k*, use the lower or upper limit of the interval:

Range of observation is 190 to 310.

$\mu + \sigma k = 310$ or $\mu - \sigma k = 190$

$250 + 20k = 310$ or $250 - 20k = 190$

Solving both the equations we arrive at *k* = 3.

According to the Chebyshev's theorem, proportion must be at least

$100[1-(1/k^2)]\%$ = $100[1-(1/3^2)]\%$ = 75%. Therefore, approximately 88.89% of the observations are between 190 and 310.

b. To determine *k*, use the lower or upper limit of the intervals:

Range of observation is 190 to 310.

$\mu + \sigma k = 290$ or $\mu - \sigma k = 210$

$250 + 20k = 290$ or $250 - 20k = 210$

Solving both the equations we arrive at *k* = 2.

According to the Chebyshev's theorem, proportion must be at least

$100[1-(1/k^2)]\%$ = $100[1-(1/2^2)]\%$ = 75%. Therefore, approximately 75% of the observations are between 210 and 290.

2.19

Since the data is Mound shaped with mean of 450 and variance of 625, use the empirical rule .

a. Greater than 425: Since approximately 68% of the observations are within 1 standard deviation from the mean that is 68% of the observations are between (425, 475). Therefore, approximately 84% of the observations will be greater than 425.
b. Less than 500: Approximately 97.5% of the observations will be less than 500.
c. Greater than 525: Since all or almost all of the distribution is within 3 standard deviations from the mean, approximately 0% of the observations will be greater than 525.

2.20

Compare the annual % returns on common stocks vs. U.S. Treasury bills

Minitab Output:

Descriptive Statistics: Stocks_Ex2.20, TBills_Ex2.20

```
Variable        N  N*   Mean  SE Mean  TrMean  StDev  Variance  CoefVar  Minimum
Stocks_Ex2.20   7   0   8.16     8.43       *  22.30    497.39   273.41   -26.50
TBills_Ex2.20   7   0  5.786    0.556       *  1.471     2.165    25.43    3.800

Variable            Q1  Median     Q3  Maximum  Range    IQR
Stocks_Ex2.20   -14.70   14.30  23.80    37.20  63.70  38.50
TBills_Ex2.20    4.400   5.800  6.900    8.000  4.200  2.500
```

a. Compare the means of the populations
Using the Minitab output

$\mu_{stocks} = 8.16$, $\mu_{Tbills} = 5.786$
Therefore, the mean annual % return on stocks is higher than the return for U.S. Treasury bills

b. Compare the standard deviations of the populations
Using the Minitab output,
$\sigma_{stocks} = 22.302$, $\sigma_{Tbills} = 1.471$
Standard deviations are not sufficient for comparision.
We need to compare the coefficient of variation rather than the standard deviations.

$$CV_{Stocks} = \frac{s}{\bar{x}} \times 100 = \frac{8.16}{22.302} \times 100 = 70.93\%$$

$$CV_{Tbills} = \frac{s}{\bar{x}} \times 100 = \frac{5.79}{1.471} \times 100 = 6.60\%$$

Therefore, the variability of the U.S. Treasury bills is much smaller than the return on stocks.

2.21

x_i	x_i^2	DEVIATION ABOUT THE MEAN, $(x_i - \bar{x})$	SQUARED DEVIATION ABOUT THE MEAN, $(x_i - \bar{x})^2$
20	400	–6.8	46.24
35	1225	8.2	67.24
28	784	1.2	1.44
22	484	–4.8	23.04
10	100	–16.8	282.24
40	1600	13.2	174.24
23	529	–3.8	14.44
32	1024	5.2	27.04
28	784	1.2	1.44
30	900	3.2	10.24
$\sum_{i=1}^{10} x_i = 268$	$\sum_{i=1}^{10} x_i^2 = 7830$	$\sum_{i=1}^{10} (x_i - \bar{x}) = -7.1\times10^{-15} \approx 0$	$\sum_{i=1}^{10} (x_i - \bar{x})^2 = 647.6$

a. Sample mean = $\bar{x} = \frac{\sum_{i=1}^{10} x_i}{n} = \frac{268}{10} = 26.8$

b. Using equation 2.13:

$$\text{Sample standard deviation} = s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{10}\left(x_i - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{647.6}{9}} = 8.483$$

c. Using equation 2.14:

$$\text{Sample standard deviation} = s = \sqrt{\frac{\sum_{i=1}^{10} x_i^2 - \frac{\left(\sum x_i\right)^2}{n}}{n-1}} = \sqrt{\frac{7830 - \frac{71824}{10}}{9}} = 8.483$$

d. Using equation 2.15:

$$\text{Sample standard deviation} = s = \sqrt{\frac{\sum_{i=1}^{10} x_i^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{7830-(10)(26.8)^2}{9}} = 8.483$$

e. Coefficient of variation = $CV = \frac{s}{\bar{x}} \times 100 = \frac{8.483}{26.8} \times 100 = 31.65\%$

2.22

Minitab Output:

Descriptive Statistics: Weights

```
Variable   N  N*    Mean  SE Mean   StDev  Variance  CoefVar  Minimum      Q1
Weights   75   0  3.8079   0.0118  0.1024    0.0105     2.69   3.5700  3.7400

Variable  Median      Q3  Maximum   Range
Weights   3.7900  3.8700   4.1100  0.5400
```

a. Using the Minitab output, range = 4.11 – 3.57 = 0.54, standard deviation = 0.1024, variance = 0.010486

b. IQR = Q3 – Q1 = 3.87 – 3.74 = .13. This tells that the range of the middle 50% of the distribution is 0.13

c. Coefficient of variation = $CV = \frac{s}{\bar{x}} \times 100 = \frac{0.1024}{3.8079} \times 100 = 2.689\%$

2.23

Minitab Output:

Descriptive Statistics: Time (in seconds)

```
Variable             Mean  StDev  Variance  CoefVar      Q1  Median      Q3
Time(in seconds)   261.05  17.51    306.44     6.71  251.75  263.00  271.25
```

Using the Minitab output

a. Sample mean = $\bar{x}$ = 261.05

b. Sample variance = $s^2 = 306.44$; $s = \sqrt{306.44} = 17.51$

c. Coefficient of variation = $CV = \frac{s}{\bar{x}} \times 100 = \frac{17.51}{261.05} = 6.708$

2.24

a. Standard deviation (s) of the assessment rates:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{583.75}{39}} = \sqrt{14.974} = 3.8696$$

b. The distribution is approximately mounded. Therefore, the empirical rule applies. Approximately 95% of the distribution is expected to be within +/- 2 standard deviations of the mean.

2.25

Mean dollar amount and standard deviation of the amounts charged to a Visa account at Florin's Flower Shop.

Descriptive Statistics: Cost of Flowers

```
                  Method of
Variable          Payment            N  N*   Mean  StDev  Median
Cost of Flowers   American Express  23   0  52.99  10.68   50.55
                  Cash              16   0  51.34  16.19   50.55
                  Master Card       24   0  54.58  15.25   55.49
                  Other             23   0  53.42  14.33   54.85
                  Visa              39   0  52.65  12.71   50.65
```

Mean dollar amount = \$52.65, standard deviation = \$12.71

2.26

a. mean without the weights $\bar{x} = \sum \frac{x_i}{n} = \frac{21}{5} = 4.2$

b. weighted mean

w_i	x_i	$w_i x_i$
8	4.6	36.8
3	3.2	9.6
6	5.4	32.4
2	2.6	5.2
5	5.2	26.0
24		110.0

$$\bar{x} = \frac{\sum w_i x_i}{\sum w_i} = \frac{110}{24} = 4.583$$

2.27

a. Calculate the sample mean of the frequency distribution for n = 40 observations

Class	m_i	f_i	$f_i m_i$
0-4	2	5	10
5-9	7	8	56
10-14	12	11	132
15-19	17	9	153
20-24	22	7	154
		40	505

$$\bar{x} = \frac{\sum f_i m_i}{n} = \frac{505}{40} = 12.625$$

b. Calculate the sample variance and sample standard deviation

Class	m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0-4	2	5	10	-10.625	112.8906	564.4531
5-9	7	8	56	-5.625	31.64063	253.125
10-14	12	11	132	-0.625	0.390625	4.296875
15-19	17	9	153	4.375	19.14063	172.2656
20-24	22	7	154	9.375	87.89063	615.2344
		40	505			1609.375

$$s^2 = \frac{\sum_{i=1}^{K} f_i \left(m_i - x_i\right)^2}{n-1} = \frac{1609.375}{39} = 41.266$$

$$s = \sqrt{s^2} = \sqrt{41.266} = 6.424$$

2.28

Class	m_i	f_i	$m_i f_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
4 < 10	7	8	56	–8.4	70.56	564.48
10 < 16	13	15	195	–2.4	5.76	86.4
16 < 22	19	10	190	3.6	12.96	129.6
22 < 28	25	7	175	9.6	92.16	645.12
		$\sum f_i = 40$	$\sum m_i f_i = 616$			$\sum f_i(m_i - \bar{x})^2 = 1425.6$

a. Sample mean = $\bar{x} = \frac{\sum m_i f_i}{n} = \frac{616}{40} = 15.4$

b. Sample variance = $s^2 = \frac{\sum_{i=1}^{K} f_i \left(m_i - x_i\right)^2}{n-1} = \frac{1425.6}{39} = 36.554$

Sample standard deviation = $s = \sqrt{s^2} = \sqrt{36.554} = 6.046$

2.29

Calculate the standard deviation for the number of defects per n = 50 radios

m_i # of Defects	f_i # of Radios	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0	12	0	-1.34	1.7956	21.5472
1	15	15	-0.34	0.1156	1.734
2	17	34	0.66	0.4356	7.4052
3	6	18	1.66	2.7556	16.5336
	50	67			47.22

$$s^2 = \frac{\sum f_i(m_i - \bar{x})^2}{n-1} = \frac{47.22}{49} = .96367\,; \quad s = \sqrt{s^2} = .9817$$

2.30

Based on a sample of *n*=50:

m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0	21	0	-1.4	1.96	41.16
1	13	13	-0.4	0.16	2.08
2	5	10	0.6	0.36	1.8
3	4	12	1.6	2.56	10.24
4	2	8	2.6	6.76	13.52
5	3	15	3.6	12.96	38.88
6	2	12	4.6	21.16	42.32
Sum	**50**	**70**			**150**

a. Sample mean number of claims per day = $\bar{X} = \frac{\sum f_i m_i}{n} = \frac{70}{50} = 1.40$

b. Sample variance = $s^2 = \frac{\sum f_i(m_i - \bar{x})^2}{n-1} = \frac{150}{49} = 3.0612$

Sample standard deviation = *s* = $\sqrt{s^2}$ = 1.7496

2.31

Estimate the sample mean and standard deviation

Class	m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0 < 4	2	3	6	-7.36	54.1696	162.5088
4 <8	6	7	42	-3.36	11.2896	79.0272
8 < 12	10	8	80	0.64	0.4096	3.2768
12 < 16	14	5	70	4.64	21.5296	107.648
16 < 20	18	2	36	8.64	74.6496	149.2992
Sum		**25**	**234**			**501.76**

a. Sample mean = $\bar{X} = \frac{\sum f_i m_i}{n} = \frac{234}{25} = 9.36$

b. Sample variance = $s^2 = \dfrac{\sum f_i(m_i - \bar{x})^2}{n-1} = \dfrac{501.76}{24} = 20.9067$

Sample standard deviation = $s = \sqrt{s^2} = 4.572$

2.32

Estimate the sample mean and sample standard deviation

Class	m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
9.95-10.45	10.2	2	20.4	-0.825	0.681	1.361
10.45-10.95	10.7	8	85.6	-0.325	0.106	0.845
10.95-11.45	11.2	6	67.2	0.175	0.031	0.184
11.45-11.95	11.7	3	35.1	0.675	0.456	1.367
11.95-12.45	12.2	1	12.2	1.175	1.381	1.381
Sum		**20**	**220.5**			**5.138**

a. sample mean = $\bar{X} = \dfrac{\sum f_i m_i}{n} = \dfrac{220.5}{20} = 11.025$

b. sample variance = $s^2 = \dfrac{\sum f_i(m_i - \bar{x})^2}{n-1} = \dfrac{5.138}{19} = 0.2704$

sample standard deviation = $s = \sqrt{s^2} = 0.520$

2.33

Find the mean and standard deviation of the number of errors per page

m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
0	102	0	-1.654	2.735716	279.043
1	138	138	-0.654	0.427716	59.02481
2	140	280	0.346	0.119716	16.76024
3	79	237	1.346	1.811716	143.1256
4	33	132	2.346	5.503716	181.6226
5	8	40	3.346	11.19572	89.56573
Sum	**500**	**827**			**769.142**

$$\mu = \frac{\sum f_i m_i}{n} = \frac{827}{500} = 1.654$$

$$\sigma^2 = \frac{\sum f_i(m_i - \bar{x})^2}{n} = \frac{769.142}{500} = 1.5383$$

Sample standard deviation = σ = $\sqrt{\sigma^2} = 1.240$

2.34

Using Table 1.7Minutes	m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i(m_i - \bar{x})^2$
220<230	225	5	1125	-36.545	1335.57	6677.851
230<240	235	8	1880	-26.545	704.6612	5637.289
240<250	245	13	3185	-16.545	273.7521	3558.777
250<260	255	22	5610	-6.5455	42.84298	942.5455
260<270	265	32	8480	3.45455	11.93388	381.8843
270<280	275	13	3575	13.4545	181.0248	2353.322
280<290	285	10	2850	23.4545	550.1157	5501.157
290<300	295	7	2065	33.4545	1119.207	7834.446
		110	28770			32887.27

a. Using Equation 2.21, Sample mean, $\bar{x} = \dfrac{\sum f_i m_i}{n} = \dfrac{28770}{110} = 261.54545$

b. Using Equation 2.22, sample variance

$$s^2 = \frac{\sum f_i (m_i - \bar{x})^2}{n-1} = \frac{32887.27}{109} = 301.718\,; \quad s = \sqrt{s^2} = 17.370$$

c. From Exercise 2.23, $\bar{x}$ = 261.05 and s^2 = 306.44. Therefore, the mean value obtained in both the Exercises are almost same, however variance is slightly lower by 4.7219 compared to Exercise 2.23.

2.35

a. Compute the sample covariance

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\,(y_i - \bar{y})$
1	5	-3	9	-1.85714	3.4489796	5.571428571
3	7	-1	1	0.14286	0.0204082	-0.142857143
4	6	0	0	-0.85714	0.7346939	0
5	8	1	1	1.14286	1.3061224	1.142857143
7	9	3	9	2.14286	4.5918367	6.428571429
3	6	-1	1	-0.85714	0.7346939	0.857142857
5	7	1	1	0.14286	0.0204082	0.142857143
28	48	0	22	2.7E-15	10.857143	14
$\bar{x}$ = 4.0	$\bar{y}$ = 6.8571		s_x^2 = 3.667		s_y^2 = 1.8095	Cov(x,y) = 2.333
			s_x =1.9149		s_y =1.3452	

$$Cov(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{14}{6} = 2.3333$$

b. Compute the sample correlation coefficient

$$r_{xy} = \frac{Cov(x, y)}{s_x s_y} = \frac{2.3333}{(1.9149)(1.3452)} = .9059$$

2.36

a. Compute the sample covariance

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\ (y_i - \bar{y})$
12	200	-7	49	-156	24336	1092
30	600	11	121	244	59536	2684
15	270	-4	16	-86	7396	344
24	500	5	25	144	20736	720
14	210	-5	25	-146	21316	730
95	1780	0	236	0	133320	5570
$\bar{x} = 19.00$	$\bar{y} = 356.00$		s_x^2 = 59		s_y^2 = 33330	Cov(x,y) = 1392.5
			s_x =7.681146		s_y =182.5650569	

$$Cov(x,y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{5570}{4} = 1392.5$$

b. Compute the sample correlation coefficient

$$r = \frac{Cov(x,y)}{s_x s_y} = \frac{1392.5}{(7.6811)(182.565)} = 0.9930$$

2.37

a. Compute the sample covariance

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\ (y_i - \bar{y})$
6	80	-2	4	30	900	-60
7	60	-1	1	10	100	-10
8	70	0	0	20	400	0
9	40	1	1	-10	100	-10
10	0	2	4	-50	2500	-100
40	250	0	10	0	4000	-180
$\bar{x}$ = 8.00	$\bar{y}$ = 50.00		s_x^2 = 2.5		s_y^2 =1000	Cov(x,y) = -45
			s_x = 1.5811		s_y =31.623	

$$Cov(x,y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{-180}{4} = -45$$

b. Compute the sample correlation coefficient

$$r_{xy} = \frac{Cov(x,y)}{s_x s_y} = \frac{-45}{(1.58114)(31.6228)} = -.90$$

2.38 Minitab output

Covariances: x_Ex2.38, y_Ex2.38

```
             x_Ex2.38  y_Ex2.38
x_Ex2.38      31.8987
y_Ex2.38       4.2680   34.6536
```

Correlations: x_Ex2.38, y_Ex2.38

```
Pearson correlation of x_Ex2.38 and y_Ex2.38 = 0.128
```

Using Minitab outputa. $Cov(x,y) = 4.268$

b. $r = 0.128$

c. Weak positive association between the number of drug units and the number of days to complete recovery. Recommend low or no dosage units.

2.39 Minitab output

Covariances: x_Ex2.39, y_Ex2.39

```
             x_Ex2.39  y_Ex2.39
x_Ex2.39      9.28571
y_Ex2.39     -5.50000   5.40952
```

Correlations: x_Ex2.39, y_Ex2.39

```
Pearson correlation of x_Ex2.39 and y_Ex2.39 = -0.776
```

Using Minitab output

a. $Cov(x,y) = -5.5$, $r = -.776$

b. Higher prices are associated with fewer days to deliver, i.e., faster delivery time.

2.40

a. Compute the covariance

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\,(y_i - \bar{y})$
5	55	-2	4	12.4	153.76	-24.8
6	53	-1	1	10.4	108.16	-10.4
7	45	0	0	2.4	5.76	0
8	40	1	1	-2.6	6.76	-2.6
9	20	2	4	-22.6	510.76	-45.2
35	213	0	10	0	785.2	-83
$\mu_x = 7.00$	$\mu_y = 42.60$		$\sigma_x^2 = 2.0$		$\sigma_y^2 = 157.04$	$Cov(x,y) = -16.6$
			$\sigma_x = 1.4142$		$\sigma_y = 12.532$	

=

$$Cov(x,y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N} = \frac{-83}{5} = -16.6$$

b. Compute the correlation coefficient

$$r_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{-16.6}{(1.4142)(12.5316)} = -.937$$

2.41

Minitab output

Covariances: Temperature (F), Time(hours)

```
                  Temperature (F)      Time(hours)
Temperature (F)         145.67273
Time(hours)               2.80136          0.05718
```

Correlations: Temperature (F), Time(hours)

```
Pearson correlation of Temperature (F) and Time(hours) = 0.971
```

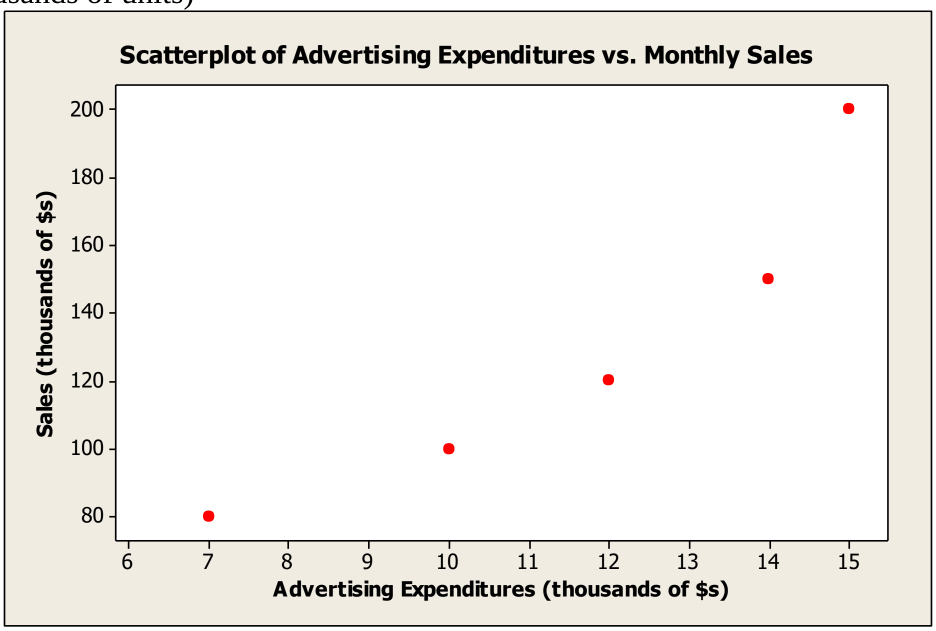
Using Minitab output

a. Covariance = 2.80136

b. Correlation coefficient = 0.971

2.42

Scatter plot – Advertising expenditures (thousands of $s) vs. Monthly Sales (thousands of units)



x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\ (y_i - \bar{y})$
10	100	-1.6	2.56	-30	900	48
15	200	3.4	11.56	70	4900	238
7	80	-4.6	21.16	-50	2500	230
12	120	0.4	0.16	-10	100	-4
14	150	2.4	5.76	20	400	48
58	650		41.2		8800	560
$\bar{x}$ = 11.60	$\bar{y}$ = 130.00		s_x^2 = 10.3		s_y^2 =2200	Cov(x,y) = 140
			s_x = 3.2094		s_y = 46.9042	

Covariance = $Cov(x,y) = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$ = 560 / 4 = 140

Correlation = $\dfrac{Cov(x,y)}{s_x s_y} = \dfrac{140}{(3.2094)(46.9042)}$ = .93002

2.43

Compute covariance and correlation between retail experience (years) and weekly sales (hundreds of dollars)

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\ (y_i - \bar{y})$
2	5	-1.875	3.515625	-5.75	33.0625	10.78125
4	10	0.125	0.015625	-0.75	0.5625	-0.09375
3	8	-0.875	0.765625	-2.75	7.5625	2.40625
6	18	2.125	4.515625	7.25	52.5625	15.40625
3	6	-0.875	0.765625	-4.75	22.5625	4.15625
5	15	1.125	1.265625	4.25	18.0625	4.78125
6	20	2.125	4.515625	9.25	85.5625	19.65625
2	4	-1.875	3.515625	-6.75	45.5625	12.65625
31	86		18.875		265.5	69.75
$\bar{x}$ = 3.875	$\bar{y}$ = 10.75		s_x^2 = 2.6964		s_y^2 = 37.9286	Cov(x,y) = 9.964286
			s_x = 1.64208		s_y = 6.15862	

Covariance = $Cov(x,y) = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$ =69.75 / 7 = 9.964286

Correlation = $\dfrac{Cov(x,y)}{s_x s_y} = \dfrac{9.964286}{(1.64208)(6.15862)}$ = .9853

2.44

Air Traffic Delays (Number of Minutes Late)

m_i	f_i	$f_i m_i$	$(m_i - \bar{x})$	$(m_i - \bar{x})^2$	$f_i (m_i - \bar{x})^2$
5	30	150	-13.133	172.46	5173.90
15	25	375	-3.133	9.81	245.32
25	13	325	6.867	47.16	613.11
35	6	210	16.867	284.51	1707.07
45	5	225	26.867	721.86	3609.30
55	4	220	36.867	1359.21	5436.84
	83	1505			16785.54
$\bar{x}$	18.13			variance =	204.7017

a. Sample mean number of minutes late = 1505 / 83 = 18.1325

b. Sample variance = 16785.54/82 = 204.7017

Sample standard deviation = s = 14.307

2.45

Minitab Output

Descriptive Statistics: Cost ($)

```
           Total
Variable   Count   Mean  StDev  Variance  Minimum     Q1  Median     Q3  Maximum
Cost ($)      50  43.10  10.16    103.32    20.00  35.75   45.00  50.25    60.00
```

Using the Minitab output

a. Mean charge = $43.10
b. Standard deviation = $10.16
c. Five - number summary:

minimum < Q1 < median < Q3 < maximum

20 < 35.75 < 45 < 50.25 < 60

2.46

For Location 2:

x_i	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$
1	-9.2	84.64
19	8.8	77.44
2	-8.2	67.24
18	7.8	60.84
11	0.8	0.64
10	-0.2	0.04
3	-7.2	51.84
17	6.8	46.24
4	-6.2	38.44
17	6.8	46.24
102		473.6

Mean = $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{102}{10} = 10.2$

Variance = $s^2 = \dfrac{\sum\left(x_i - \bar{x}\right)^2}{n-1} = \dfrac{473.6}{9} = 52.622$

Standard deviation = $s = \sqrt{s^2} = 7.254$

For Location 3:

x_i	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$
2	-16.4	268.96
3	-15.4	237.16
25	6.6	43.56
20	1.6	2.56
22	3.6	12.96
19	0.6	0.36
25	6.6	43.56
20	1.6	2.56
22	3.6	12.96
26	7.6	57.76
184		682.4
18.4		

Mean = $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{184}{10} = 18.4$

Variance = $s^2 = \dfrac{\sum\left(x_i - \bar{x}\right)^2}{n-1} = \dfrac{682.4}{9} = 75.822$

Standard deviation = $s = \sqrt{s^2} = 8.708$

For Location 4:

x_i	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$
22	9.5	90.25
20	7.5	56.25
10	-2.5	6.25
13	0.5	0.25
12	-0.5	0.25
10	-2.5	6.25
11	-1.5	2.25
9	-3.5	12.25
10	-2.5	6.25
8	-4.5	20.25
125		200.5

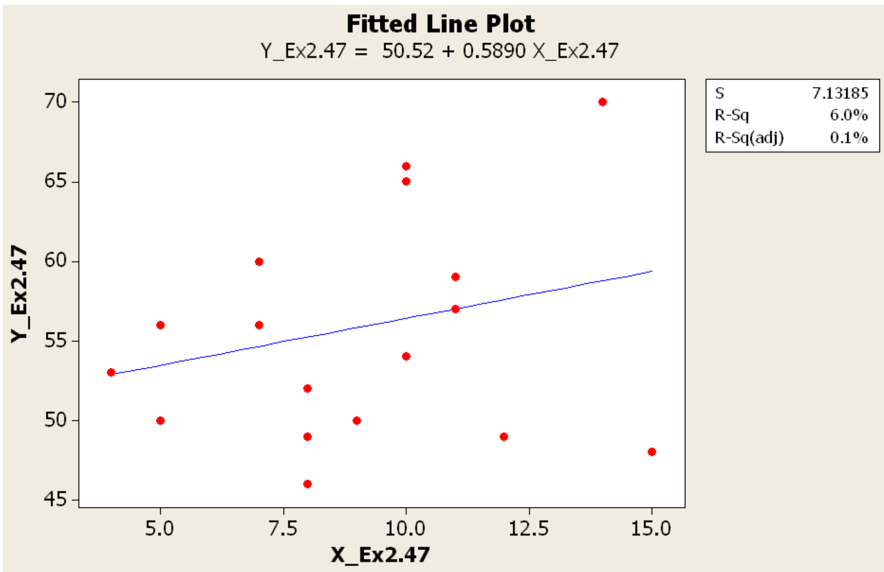
Mean = $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{125}{10} = 12.5$

$$\text{Variance} = s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{200.5}{9} = 22.278$$

$$\text{Standard deviation} = s = \sqrt{s^2} = 4.720$$

2.47

Describe the data numerically



Covariances: X_Ex2.47, Y_Ex2.47

	X_Ex2.47	Y_Ex2.47
X_Ex2.47	8.81046	
Y_Ex2.47	5.18954	50.92810

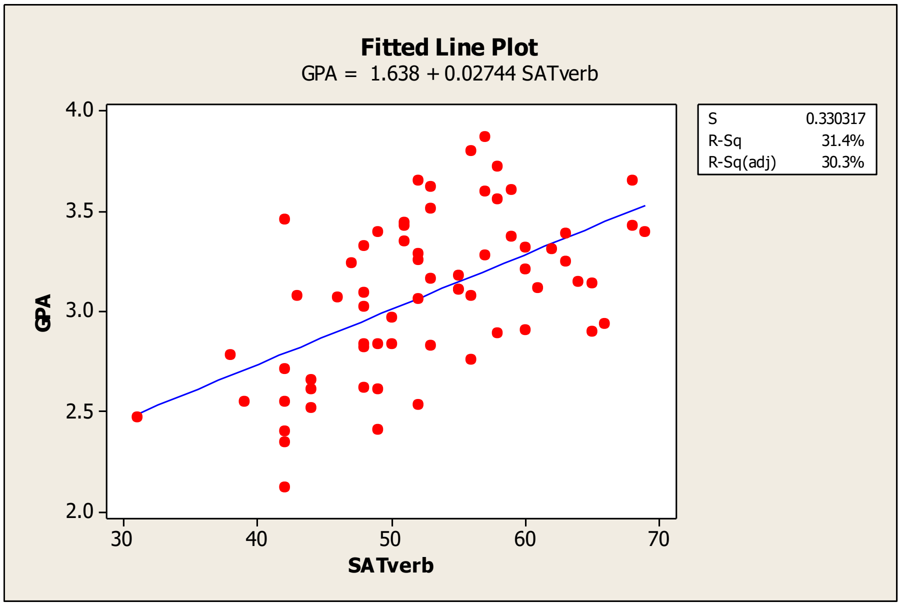
Correlations: X_Ex2.47, Y_Ex2.47

```
Pearson correlation of X_Ex2.47 and Y_Ex2.47 = 0.245
P-Value = 0.327
```

There is a very weak positive relationship between the variables.

2.48

a. Describe the data graphically between graduating GPA vs. entering SAT Verbal scores



b.

Correlations: GPA, SATverb

```
Pearson correlation of GPA and SATverb = 0.560
P-Value = 0.000
```

2.49

Arrange the populations according to their variances and calculate the variances manually (a) has the least variability, then population (c), followed by (b) and then (d)

Population standard deviation $\sigma^2 = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}}$

a	b	c	d	$(a-\bar{a})^2$	$(b-\bar{b})^2$	$(c-\bar{c})^2$	$(d-\bar{d})^2$
1	1	1	-6	12.25	12.25	12.25	110.25
2	1	1	-3	6.25	12.25	12.25	56.25
3	1	4	0	2.25	12.25	0.25	20.25
4	1	4	3	0.25	12.25	0.25	2.25
5	8	5	6	0.25	12.25	0.25	2.25
6	8	5	9	2.25	12.25	0.25	20.25
7	8	8	12	6.25	12.25	12.25	56.25
8	8	8	15	12.25	12.25	12.25	110.25
36	36	36	36	42	98	50	378
$\bar{x} = 4.5$	$\bar{x} = 4.5$	$\bar{x} = 4.5$	$\bar{x} = 4.5$	σ^2_a = 5.25	σ^2_b = 12.25	σ^2_c = 6.25	σ^2_d = 47.25

2.50

Mean of \$295 and standard deviation of \$63.

a. Find a range in which it can be guaranteed that 60% of the values lie.
   Use Chebyshev's theorem: at least 60% = [1-(1/k^2)]. Solving for k, k = 1.58. The interval will range from 295 +/- (1.58)(63) = 295 +/- 99.54. 195.46 up to 394.54 will contain at least 60% of the observations.

b. Find the range in which it can be guaranteed that 84% of the growth figures lie
   Use Chebyshev's theorem: at least 84% = [1-(1/k^2)]. Solving for k, k = 2.5. The interval will range from 295 +/- (2.50)(63) = 295 +/- 157.5. 137.50 up to 452.50 will contain at least 84% of the observations.

2.51

Growth of 500 largest U.S. corporations had a mean of 9.2%, standard deviation of 3.5%.

a. Find the range in which it can be guaranteed that 84% of the growth figures lie.
   Use Chebyshev's theorem: at least 84% = [1-(1/k^2)]. Solving for k, k = 2.5. The interval will range from 9.2 +/- (2.50)(3.5) = 9.2 +/- 8.75. 0.45% up to 17.95% will contain at least 84% of the observations.

b. Using the empirical rule, approximately 68% of the earnings growth figures lie within 9.2 +/- (1)(3.5). 5.7% up to 12.7% will contain at least 68% of the observations.

2.52

Tires have a lifetime mean of 29,000 miles and a standard deviation of 3,000 miles.

a. Find a range in which it can be guaranteed that 75% of the lifetimes of tires lies
   Use Chebyshev's theorem: at least 75% = $[1-(1/k^2)]$. Solving for $k = 2.0$. The interval will range from 29,000 ± (2.0)(3,000) = 29,000 ± 6,000 23,000 to 35,000 will contain at least 75% of the observations .

b. 95%: solve for $k = 4.47$. The interval will range from 29,000 ± (4.47)(3000) = 29,000 ± 13,416.41. 15,583.59 to 42,416.41 will contain at least 95% of the observations.

2.53

Minitab Output:

Descriptive Statistics: Time (in seconds)

```
                      Total
Variable              Count    Mean  StDev  Variance  Minimum      Q1
Time (in seconds)       110  261.05  17.51    306.44   222.00  251.75


Variable               Median      Q3  Maximum    IQR
Time (in seconds)      263.00  271.25   299.00  19.50
```

Using the Minitab output

a. Interquartile Range = 19.50. This tells that the range of the middle 50% of the distribution is 19.50.

b. Five - number summary:

   minimum < Q1 < median < Q3 < maximum

   222 < 251.75 < 263 < 271.25 < 299

2.54

Minitab Output:

Descriptive Statistics: Time

```
           Total
Variable   Count   Mean  StDev  Variance  CoefVar  Minimum     Q1  Median
Time         104  41.68  16.86    284.35    40.46    18.00  28.50   39.00


Variable      Q3     Maximum
Time       56.50       73.00
```

Using the Minitab output

a. Mean shopping time = 41.68

b. Variance = 284.35
   Standard deviation = 16.86

c. 95^{th} percentile = the value located in the $0.95(n + 1)^{th}$ ordered position
   = the value located in the 99.75^{th} ordered position
   = 70 + 0.75(70 –70) = 70.

d. Five - number summary:

   minimum < Q1 < median < Q3 < maximum

   18 < 28.50 < 39 < 56.50 < 73

e. Coefficient of variation = 40.46

f. Find the range in which ninety percent of the shoppers complete their shopping. Use Chebyshev's theorem: at least 90% = [1-(1/k^2)]. Solving for k, k = 3.16. The interval will range from 41.68 +/- (3.16)(16.86) = 41.88 +/- 53.28. -11.60 up to 94.96 will contain at least 90% of the observations.

2.55

x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})\ (y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
3.5	88	0.3	7.8	2.34	0.09	60.84
2.4	76	-0.8	-4.2	3.36	0.64	17.64
4	92	0.8	11.8	9.44	0.64	139.24
5	85	1.8	4.8	8.64	3.24	23.04
1.1	60	-2.1	-20.2	42.42	4.41	408.04
16	401			66.2	9.02	648.8
$\bar{x}$ = 3.2	$\bar{y}$ = 80.2				s_x^2 = 2.255	s_y^2 = 162.2
					s_x = 1.5017	s_y = 12.7358

Covariance = $Cov(x,y) = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \dfrac{66.2}{4} = 16.55$

Correlation coefficient = $\dfrac{Cov(x,y)}{s_x s_y} = \dfrac{16.55}{(1.5017)(12.7358)} = 0.8654$

2.56

x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})\ (y_i - \bar{y})$
12	20	-9.3	86.49	-21.20	449.44	197.16
30	60	8.7	75.69	18.80	353.44	163.56
15	27	-6.3	39.69	-14.20	201.64	89.46
24	50	2.7	7.29	8.80	77.44	23.76
14	21	-7.3	53.29	-20.20	408.04	147.46
18	30	-3.3	10.89	-11.20	125.44	36.96
28	61	6.7	44.89	19.80	392.04	132.66
26	54	4.7	22.09	12.80	163.84	60.16
19	32	-2.3	5.29	-9.20	84.64	21.16
27	57	5.7	32.49	15.80	249.64	90.06
213	412		378.1		2505.6	962.4
$\bar{x} = 21.3$	$\bar{y} = 41.2$		s_x^2 = 42.01		s_y^2 = 278.4	
			s_x = 6.4816		s_y =16.6853	

Covariance = $Cov(x, y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{962.4}{9} = 106.9333$

Correlation coefficient = $\frac{Cov(x,y)}{s_x s_y} = \frac{106.9333}{(6.4816)(16.6853)}$ = 0.9888