Chapter 3 - Data Description

Note: Answers may vary due to rounding, TI 83's, or computer programs.

EXERCISE SET 3-1

1.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{24{,}865}{25} = 994.6$

b. MD: 940

c. Mode: 1180

d. MR: $\frac{590 + 1595}{2} = 1092.5$

2.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{47{,}619}{15} = 3174.6$

b. MD $= 1479$

c. Mode: no mode

d. MR: $\frac{203 + 9822}{2} = 5012.5$

3.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{4982}{37} = 134.6$

b. MD: 41

c. Mode = 2

d. MR $= \frac{1 + 1008}{2} = 504.5$

For the best measure of average, answers will vary.

4.

For Observers:

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{3804}{10} = 380.4$

b. MD: $\frac{352 + 378}{2} = 365$

c. Mode: no mode

d. MR $= \frac{484 + 302}{2} = 393$

For Visits:

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{2769}{10} = 276.9$

b. MD : $\frac{194 + 219}{2} = 206.5$

c. Mode: no mode

d. MR $= \frac{114 + 634}{2} = 374$

The values are higher for observers.

5.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{1136}{10} = 113.6$

b. MD $= \frac{119 + 120}{2} = 119.5$

c. Mode: 120

d. MR $= \frac{74 + 132}{2} = 103$

6.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{247}{13} =$ \$19 million

b. MD: \$10 million

c. Mode: \$7 million

d. MR $= \frac{7 + 50}{2} =$ \$$28.5$ million

The data is positively skewed since the mean is much higher than the median or mode.

7.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{16{,}848}{4} =$ 4212 km (2617.5 mi)

b. MD $= \frac{3643 + 4821}{2} =$ 4232 (2630 mi)

7. continued

c. Mode: none

d. MR $= \frac{3122+5262}{2} = 4192$ km (2605 mi)

The mean, median, and midrange are all very close.

8.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{1221.1}{50} = \$24.42$

b. MD $= \frac{23.2+23.7}{2} = \$23.45$

c. Mode: 16.9, 17.2, 18, 19.1, 24, 25.2, 31.7

d. MR $= \frac{16.5+47.7}{2} = 32.1$

It appears that the mean and median are good measures of average.

9.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{20{,}924}{14} = 1494.6$

b. MD $= \frac{1409+1422}{2} = 1415.5$

c. Mode: none

d. MR $= \frac{1023+2532}{2} = 1777.5$

10.

New England States:

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{14{,}709}{6} = 2451.5$

b. MD $= \frac{1112+1795}{2} = 1453.5$

c. Mode: none

Northwest States:

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{3419}{6} = 569.8$

b. MD $= \frac{172+620}{2} = 396$

c. Mode: none

The measures of central tendency are much larger for New England compared to those for Northwest.

11.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{486.2}{13} = 37.4$

b. MD = 33.7

c. Mode: no mode

d. MR $= \frac{4.4+87.9}{2} = 46.15$

12.

a. $\overline{X} = \frac{\Sigma X}{n} = \frac{209}{21} = 9.952 \approx 10$

b. MD = 9

c. Mode = 8 and 9

d. MR $= \frac{6+14}{2} = 10$

13.

Boundaries	X_m	f	$f \cdot X_m$
0.5 - 3.5	2	11	22
3.5 - 6.5	5	12	60
6.5 - 9.5	8	4	32
9.5 - 12.5	11	2	22
12.5 - 15.5	14	1	14
		30	150

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{150}{30} = 5$

b. modal class: 3.5 - 6.5

14.

Class Limits	*Boundaries*	X_m	f	$f \cdot X_m$
2.48 - 7.48	2.475 - 7.485	4.98	7	34.86
7.49 - 12.49	7.485 - 12.495	9.99	3	29.97
12.50 - 17.50	12.495 - 17.505	15.00	1	15.00
17.51 - 22.51	17.505 - 22.515	20.01	7	140.07
22.52 - 27.52	2.515 - 27.525	25.02	5	125.10
27.53 - 32.53	27.525 - 32.535	30.03	5	150.15
			28	495.15

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{495.15}{28} = 17.68$

b. modal class: $2.48 - 7.48$ and $17.51 - 22.51$ The grouped mean is less.

15.

Boundaries	X_m	f	$f \cdot X_m$
7.5 - 12.5	10	3	30
12.5 - 17.5	15	5	75
17.5 - 22.5	20	15	300
22.5 - 27.5	25	5	125
27.5 - 32.5	30	2	60
		30	590

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{590}{30} = 19.7$

b. modal class: $17.5 - 22.5$

16.

Percentage	*Boundaries*	X_m	f	$f \cdot X_m$
0.8 - 4.4	0.75 - 4.45	2.6	26	67.6
4.5 - 8.1	4.45 - 8.15	6.3	11	69.3
8.2 - 11.8	8.15 - 11.85	10.0	4	40.0
11.9 - 15.5	11.85 - 15.55	13.7	5	68.5
15.6 - 19.2	15.55 - 19.25	17.4	2	34.8
19.3 - 22.9	19.25 - 22.95	21.1	1	21.1
23.0 - 26.6	22.95 - 26.65	24.8	1	24.8
			50	326.1

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{326.1}{50} = 6.5$

b. modal class: $0.8 - 4.4$

The mean is probably not the best measure of central tendency for this data because the data is "top" heavy.

17.

Percentage	*Boundaries*	X_m	f	$f \cdot X_m$
15.2 - 19.6	15.15 - 19.65	17.4	3	52.2
19.7 - 24.1	19.65 - 24.15	21.9	15	328.5
24.2 - 28.6	24.15 - 28.65	26.4	19	501.6
28.7 - 33.1	28.65 - 33.15	30.9	6	185.4
33.2 - 37.6	33.15 - 37.65	35.4	7	247.8
37.7 - 42.1	37.65 - 42.15	39.9	0	0
42.2 - 46.6	42.15 - 46.65	44.4	1	44.4
			51	1359.9

17. continued

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{1359.9}{51} = 26.66$ or 26.7

b. modal class: 24.2 – 28.6

18.

Class Limits	*Boundaries*	X_m	f	$f \cdot X_m$
10 – 20	9.5 – 20.5	15	2	30
21 – 31	20.5 – 31.5	26	8	208
32 – 42	31.5 – 42.5	37	15	555
43 – 53	42.5 – 53.5	48	7	336
54 – 64	53.5 – 64.5	59	10	590
65 – 75	64.5 – 75.5	70	3	210
			45	1929

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{1929}{45} = 42.9$

b. modal class: 32 – 42

19.

Boundaries	X_m	f	$f \cdot X_m$
0.5 - 19.5	10	12	120
19.5 - 38.5	29	7	203
38.5 - 57.5	48	5	240
57.5 - 76.5	67	3	201
76.5 - 95.5	86	3	258
		30	1022

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{1022}{30} = 34.1$

b. modal class: 0.5 – 19.5

20.

Class Limits	*Boundaries*	X_m	f	$f \cdot X_m$
150 – 158	149.5 – 158.5	154	5	770
159 – 167	158.5 – 167.5	163	16	2608
168 – 176	167.5 – 176.5	172	20	3440
177 – 185	176.5 – 185.5	181	21	3801
186 – 194	185.5 – 194.5	190	20	3800
195 – 203	194.5 – 203.5	199	15	2985
204 – 212	203.5 – 212.5	208	3	624
			100	18,028

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{18{,}028}{100} = 180.3$

b. modal class: 177 – 185

21.

Boundaries	X_m	*f*	$f \cdot X_m$
15.5 – 18.5	17	14	238
18.5 – 21.5	20	12	240
21.5 – 24.5	23	18	414
24.5 – 27.5	26	10	260
27.5 – 30.5	29	15	435
30.5 – 33.5	32	6	192
		75	1779

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{1779}{75} = 23.7$

b. modal class: 21.5 – 24.5

22.

Limits	*Boundaries*	X_m	*f*	$f \cdot X_m$
1013 - 1345	1012.5 - 1345.5	1179	11	12969
1346 - 1678	1345.5 - 1678.5	1512	4	6048
1679 - 2011	1678.5 - 2011.5	1845	7	12915
2012 - 2344	2011.5 - 2344.5	2178	3	6534
2345 - 2677	2344.5 - 2677.5	2511	5	12555
2678 - 3010	2677.5 - 3010.5	2844	3	8532
			33	59553

$$\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{59553}{33} = 1804.6$$

modal class: 1013 – 1345

23.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{8(10{,}000) + 10(12{,}000) + 12(8{,}000)}{8 + 10 + 12} = \frac{296{,}000}{8 + 10 + 12} = \frac{296{,}000}{30} = \$\,9866.67$$

24.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{3(3.33) + 3(3.00) + 2(2.5) + 2.5(4.4) + 4(1.75)}{3 + 3 + 2 + 2.5 + 4} = \frac{41.99}{14.5} = 2.896$$

25.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{40(1000) + 30(3000) + 50(800)}{1000 + 3000 + 800} = 35.4\%$$

26.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{9(427000) + 6(365000) + 12(725000)}{9 + 6 + 12} = \frac{14{,}733{,}000}{27} = \$545,666.67$$

27.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{20(83) + 30(72) + 50(90)}{100} = 83.2$$

28.

$$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{1(62) + 1(83) + 1(97) + 1(90) + 2(82)}{6} = \frac{496}{6} = 82.7$$

29.
a. Mode
b. Median
c. Median
d. Mode
e. Mean
f. Median

30.
a. Median
b. Mean
c. Mode
d. Mode
e. Mode
f. Mean

31.
Roman letters, $\overline{\mathrm{X}}$
Greek letters, μ

32.
Both could be true since one could be using the mean for the average salary, and the other could be using the mode for the average.

33.
$5 \cdot 64 = 320$

34.
$5 \cdot 8.2 = 41$
$6 + 10 + 7 + 12 + x = 41$
$x = 6$

35.
The mean of the original data is 30.
The means will be:
a. 40
b. 20
c. 300
d. 3
e. The results will be the same as adding, subtracting, multiplying, and dividing the mean by 10.

36.
a. $\frac{2}{\frac{1}{30}+\frac{1}{45}} = 36$ mph b. $\frac{2}{\frac{1}{40}+\frac{1}{25}} = 30.77$ mph c. $\frac{2}{\frac{1}{50}+\frac{1}{10}} = \$16.67$

37.
a. $\sqrt[3]{(1.35)(1.24)(1.18)} = 1.2547 \approx 1.255$

Average growth rate: $1.255 - 1 = 0.255$ or 25.5%

b. $\sqrt[4]{(1.08)(1.06)(1.04)(1.05)} = 1.057397$

Average growth rate: $1.057 - 1 = 0.057$ or 5.7%

c. $\sqrt[5]{(1.10)(1.08)(1.12)(1.09)(1.03)} = 1.084$

Average growth rate: $1.084 - 1 = 0.084$ or 8.4%

d. $\sqrt[3]{(1.01)(1.03)(1.055)} = \sqrt[3]{1.0975165} = 1.032$

Average growth rate: $1.032 - 1 = 0.032$ or 3.2%

38.
$$\sqrt{\frac{8^2+6^2+3^2+5^2+4^2}{5}} = \sqrt{30} = 5.477$$

39. $\text{MD} = \frac{\frac{50}{2}-0}{26}(3.7)+0.75 = 4.31$

EXERCISE SET 3-2

1.
The square root of the variance is equal to the standard deviation.

2.
One extremely high or low data value would influence the range.

3.
σ^2, σ

4.
s^2, s

5.
When the sample size is less than 30, the formula for the true standard deviation of the sample will underestimate the population standard deviation.

6.
a. s = 4.320
b. s = 5.066
c. s = 6.00
Data set A is least variable and data set C is the most variable.

7.
R = 416 − 190 = 226
$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{15(1{,}222{,}730)-4194^2}{15(15-1)} = \frac{751{,}314}{210} = 3577.7$$

$$s = \sqrt{3577.7} = 59.8$$

8.
R = 70 − 8 = 62
$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{17(30{,}324)-652^2}{17(17-1)} = 332.4$$

$$s = \sqrt{332.4} = 18.2$$

Using the range rule of thumb, $s \approx \frac{70-8}{4} = 15.5$ This is close to the actual standard deviation of 18.2.

9.
Men:
R = 862 − 791 = 71
$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{10(6{,}873{,}116)-8288^2}{10(10-1)} = \frac{40216}{90} = 446.8$$

$$s = \sqrt{446.8} = 21.1$$

9. continued
Women:
$R = 816 - 736 = 80$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{10(5{,}929{,}923) - 7697^2}{10(10-1)} = \frac{55{,}421}{90} = 615.8$$

$$s = \sqrt{615.8} = 24.82$$

The women's scores are more variable.

10.
Eastern states:
$R = 37{,}741 - 20{,}966 = 16{,}775$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{6(5{,}830{,}685{,}308) - 183{,}684^2}{6(6-1)} = 41,476,666.4$$

$$s = \sqrt{41,476,666.4} = 6440.2$$

Western states:
$R = 101{,}510 - 54{,}339 = 47{,}171$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{6(31{,}891{,}035{,}030) - 428{,}362^2}{6(6-1)} = 261,740,237.9$$

$$s = \sqrt{261,740,237.9} = 16,178.4$$

Western states are more variable.

11.
Triplets:
$R = 7110 - 5877 = 1233$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{10(427{,}765{,}643) - 65267^2}{10(10-1)} = \frac{17{,}875{,}141}{90} = 198,612.7$$

$$s = \sqrt{198,612.7} = 445.7$$

Quadruplets:
$R = 512 - 345 = 167$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{10(1{,}925{,}217) - 4347^2}{10(10-1)} = \frac{355{,}761}{90} = 3952.9$$

$$s = \sqrt{3952.9} = 62.9$$

Quintuplets:
$R = 91 - 46 = 45$

$$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{10(56{,}535) - 741^2}{10(10-1)} = \frac{16{,}269}{90} = 180.8$$

$$s = \sqrt{180.8} = 13.4$$

The data for triplets are most variable.

12.
Europe:
$R = \$48{,}704 - \$27{,}789 = \$20{,}915$

12. continued

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{7(8{,}745{,}505{,}887) - 242{,}459^2}{7(7-1)} = 57{,}908{,}917.33$

$s = \sqrt{57,908,917.33} = \$7609.79$

Asia:

R = \$26,852 − \$5862 = \$20,990

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{6(1{,}923{,}668{,}064) - 97{,}958^2}{6(6-1)} = 64{,}874{,}620.67$

$s = \sqrt{64{,}874{,}620.67} = \$8054.48$

The data for Asia are more variable.

13.

R = 46 − 26 = 20

Using the range rule of thumb, $s \approx \frac{20}{4} = 5$

14.

R = 71 − 49 = 22

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{12(38{,}359) - 675^2}{12(12-1)} = 35.48$ or 35.5

$s = \sqrt{35.5} = 5.96 \approx 6$

15.

R = 355 − 22 = 333

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{12(156{,}423) - 877^2}{12(12-1)} = 8393.5$

$s = \sqrt{8393.5} = 91.6$

16.

R = 2786 − 65 = 2721

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = 355{,}427.57$ or 355,427.6

$s = \sqrt{355,427.6} = 596.2$

17.

R = 156 − 26 = 130

$s^2 = \frac{n\sum X^2-(\sum X)^2}{n(n-1)} = \frac{25(271{,}995) - 2471^2}{25(25-1)} = 1156.7$

$s = \sqrt{1156.7} = 34.0$

18.

For unemployment:

$s = 59.8 \qquad \frac{\text{Range}}{4} = \frac{226}{4} = 56.5$

For executions:

$s = 91.6 \qquad \frac{\text{Range}}{4} = \frac{333}{4} = 83.3$

18. continued
For precipitation:
$s = 34.0 \qquad \frac{\text{Range}}{4} = \frac{130}{4} = 32.5$

The closest estimate is for precipitation. The estimate for unemployment is also close.

19.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
16	2	32	512
23	7	161	3703
30	12	360	10,800
37	5	185	6845
44	6	264	11,616
51	1	51	2601
58	0	0	0
65	2	130	8450
	35	1183	44527

$s^2 = \frac{n\sum f \cdot X_m^2 - (\sum f \cdot X_m)^2}{n(n-1)} = \frac{35(44{,}527) - 1183^2}{35(35-1)} = 133.58$ or 133.6
$s = \sqrt{133.58} = 11.6$

20.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
10	3	30	300
15	5	75	1125
20	15	300	6000
25	5	125	3125
30	2	60	1800
	30	590	12,350

$s^2 = \frac{30(12{,}350) - 590^2}{30(30-1)} = 25.7$
$s = \sqrt{25.7} = 5.07$ or 5.1

21.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
65	13	845	54,925
128	2	256	32,768
191	0	0	0
254	5	1270	322,580
317	1	317	100,489
380	1	380	144,400
443	0	0	0
506	1	506	256,036
569	2	1138	647,522
	25	4712	1,558,720

$s^2 = \frac{n\sum f \cdot X_m^2 - (\sum f \cdot X_m)^2}{n(n-1)} = \frac{25(1{,}558{,}720) - 4712^2}{25(25-1)} = 27{,}941.76$
$s = \sqrt{27941.76} = 167.16$ or 167.2

22.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
2.4	12	28.8	69.12
3.1	13	40.3	124.93
3.8	7	26.6	101.08
4.5	5	22.5	101.25
5.2	2	10.4	54.08
5.9	1	5.9	34.81
	40	134.5	485.27

$s^2 = \frac{40(485.27) - 134.5^2}{40(40-1)} = 0.85$

$s = \sqrt{0.85} = 0.92$

23.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
5	5	25	125
14	7	98	1372
23	10	230	5290
32	3	96	3072
41	3	123	5043
50	2	100	5000
	30	672	19,902

$s^2 = \frac{n\sum f\cdot X_m^2 - (\sum f\cdot X_m)^2}{n(n-1)} = \frac{30(19{,}902) - 672^2}{30(30-1)} = 167.2$

$s = \sqrt{167.2} = 12.9$

24.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
8	8	64	512
15	5	75	1125
22	7	154	3388
29	1	29	841
36	1	36	1296
43	3	129	5547
	25	487	12709

$s^2 = \frac{25(12{,}709) - 487^2}{25(25-1)} = 134.26$ or 134.3

$s = \sqrt{134.3} = 11.6$

25.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
68	5	340	23,120
79	14	1106	87,374
90	18	1620	145,800
101	25	2525	255,025
112	12	1344	150,528
123	6	738	90,774
	80	7673	752,621

$s^2 = \frac{n\sum f\cdot X_m^2 - (\sum f\cdot X_m)^2}{n(n-1)} = \frac{80(752{,}621) - 7673^2}{80(80-1)} = 211.19$ or 211.2

25. continued
s $= \sqrt{211.2} = 14.5$

No, the variability of the lifetimes of the batteries is quite large.

26.
For National League:

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
0.254	4	1.016	0.258064
0.259	6	1.554	0.402486
0.264	1	0.264	0.069696
0.269	4	1.076	0.289444
0.274	1	0.274	$\underline{0.075076}$
	16	4.184	1.094766

$s^2 = \frac{16(1.094766) - 4.184^2}{16(16-1)} = 0.000043$
$s = \sqrt{0.000043} = 0.0066$

For American League:

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
0.2585	2	0.5170	0.13364450
0.2645	5	1.3225	0.34980125
0.2705	4	1.0820	0.29268100
0.2765	2	0.5530	0.15290450
0.2825	1	0.2825	$\underline{0.07980625}$
	14	3.7570	1.00883750

$s^2 = \frac{14(1.0088375) - 3.757^2}{14(14-1)} = 0.00004767$
$s = \sqrt{0.00004767} = 0.0069$

27.
C. Var = $\frac{s}{\overline{X}}$ = $\frac{2.3}{11}$ = 0.209 = 20.9%
C. Var = $\frac{s}{\overline{X}}$ = $\frac{1.8}{8}$ = 0.225 = 22.5%
The factory workers' data are more variable.

28.
For US: $\overline{X} = 3386.6$, $s = 693.9$; C. Var = $\frac{s}{\overline{X}} = \frac{693.9}{3386.6} = 0.2049$ or 20.49%

For World: $\overline{X} = 4997.8$, $s = 803.2$; C. Var = $\frac{s}{\overline{X}} = \frac{803.2}{4997.8} = 0.1607$ or 16.07%

The data for US is more variable.

29.
C. Var $= \frac{s}{\overline{X}} = \frac{10.5}{80.2} = 0.131 = 13.1\%$

C. Var $= \frac{s}{\overline{X}} = \frac{18.3}{120.6} = 0.152 = 15.2\%$

The waiting time for people who are discharged is more variable.

30.
C. Var $= \frac{\mathrm{s}}{\overline{\mathrm{X}}} = \frac{6}{26} = 0.231 = 23.1\%$
C. Var $= \frac{\mathrm{s}}{\overline{\mathrm{X}}} = \frac{4000}{31{,}000} = 0.129 = 12.9\%$
Age is more variable.

31.
a. $1 - \frac{1}{2^2} = \frac{3}{4}$ or 75%
b. $1 - \frac{1}{1.5^2} = 0.56$ or 56%

32.
a. $1 - \frac{1}{5^2} = 0.96$ or 96%
b. $1 - \frac{1}{4^2} = 0.9375$ or 93.75%

33.
$\frac{120}{160} = 0.75 = 75\%$ so $k = 2$
$72 + 2\mathrm{s} = 77$
$\mathrm{s} = 2.5$
$72 + 2.5k = 82$
$k = 4$
$1 - \frac{1}{4^2} = 0.9375$ or at least 93.75%.

34.
$\overline{\mathrm{X}} = 240$ and $\mathrm{s} = 38$
At least 75% of the data values will fall withing two standard deviations of the mean; hence,
$2(38) = 76$ and $240 - 76 = 164$ and $240 + 76 = 316$. Hence at least 75% of the
data values will fall between 164 and 316 calories.

35.
$1 - \frac{1}{\mathrm{k}^2} = 0.8889$ $\qquad k = 3$
$\overline{\mathrm{X}} = 3$ hours or 180 minutes and $\mathrm{s} = 32$ minutes
$180 - 3(32) = 84$ minutes; $180 + 3(32) = 276$
At least 88.89% of the data values will fall between 84 and 276 minutes.

36.
$1 - \frac{1}{\mathrm{k}^2} = 0.8889$ $\qquad k = 3$
$\overline{\mathrm{X}} = 640$and $\mathrm{s} = 85$
At least 88.89% of the data values will fall within 3 standard deviations of the mean, hence
$640 - 3(85) = 385$ and $640 + 3(85) = 895$. Therefore at least 88.89% of the data values will fall between 385 and 895 pounds.

37.
$1 - \frac{1}{k^2} = 0.75$ $\qquad k = 2$
$\overline{\mathrm{X}} = \$246{,}300$ and $\mathrm{s} = \$48{,}500$
$\$246{,}300 - 2\ (\$48{,}500) = \$149{,}300$ and $\$246{,}300 + 2\ (\$48{,}500) = \$343{,}300$. At least 75% of the homes will fall between \$149,300 and \$343,300.

38.
$\overline{\mathrm{X}} = 12$ and $\mathrm{s} = 3$
$20 - 12 = 8$ and $8 \div 3 = 2.67$
Hence, $1 - \frac{1}{\mathrm{k}^2} = 1 - \frac{1}{2.67^2} = 1 - 0.14 = 0.86 = 86\%$
At least 86% of the data values will fall between 4 and 20.

39.
$\overline{X} = 443$ and $s = 42$
$443 + 42k = 548$ so $k = 2.5$
$1 - \frac{1}{k^2} = 1 - \frac{1}{2.5^2} = 0.84$ or at least 84%

40.
$26.8 + 1(4.2) = 31$
By the Empirical Rule, 68% of consumption is within 1 standard deviation of the mean. Then $\frac{1}{2}$ of 32%, or 16%, of consumption would be more than 31 pounds of citrus fruit per year.

41.
By the Empirical Rule, 68% of scores are within 1 standard deviation of the mean.
Thus, $514 + 1(40) = 554$ and $514 - 1(40) = 474$. Therefore, 68% of the scores would fall between 474 and 554.

To find the percentage of scores above 594, first find k:
$514 + k(40) = 594$
$40k = 80$
$k = 2$
By the Empirical Rule, 95% of the data are within $k = 2$ standard deviations of the mean. This means that $100\% - 95\% = 5\%$ of the scores would be above and below 2 standard deviations of the mean. Thus, $\frac{1}{2}$ of 5%, or 2.5%, of the data are above 594.

42.
(a) $53 + 4.2k = 58.6$
$4.2k = 5.6$
$k = 2$
By Chebyshev's Theorem, $1 - \frac{1}{2^2} = .75$ or 75% of hours worked are within 2 standard deviations of the mean. Then at most $\frac{1}{2}$ of 25%, or 12.5%, work more than 58.6 hours per week.

(b) By the Empirical Rule, $k = 2$ standard deviations of the mean is 95% of hours worked. Then $\frac{1}{2}$ of 5%, or 2.5%, work more than 58.6 hours per week.

43.
$n = 30$ $\overline{X} = 214.97$ $s = 20.76$ At least 75% of the data values will fall between $\overline{X} \pm 2s$.
$\overline{X} - 2(20.76) = 214.97 - 41.52 = 173.45$ and $\overline{X} + 2(20.76) = 214.97 + 41.52 = 256.49$
In this case all 30 values fall within this range.

44.
$n = 30$ $\overline{X} = 34.47$ $s = 13.32$
$\overline{X} - 2s = 34.47 - 2(13.32) = 7.83$ and $\overline{X} + 2s = 34.47 + 2(13.32) = 61.11$
In this case 28 out of 30 data values fall within the range of 7.83 to 61.11. This is 93.3% which is consistent with Chebyshev's Theorem.

45.
For $k = 1.5$, $1 - \frac{1}{1.5^2} = 1 - 0.44 = 0.56$ or 56%
For $k = 2$, $1 - \frac{1}{2^2} = 1 - 0.25 = 0.75$ or 75%
For $k = 2.5$, $1 - \frac{1}{2.5^2} = 1 - 0.16 = 0.84$ or 84%
For $k = 3$, $1 - \frac{1}{3^2} = 1 - 0.1111 = .8889$ or 89%
For $k = 3.5$, $1 - \frac{1}{3.5^2} = 1 - 0.08 = 0.92$ or 92%

46.
a. $s = 15.81$

46. continued
b. s = 15.81
c. s = 15.81
d. s = 79.06
e. s = 3.16
f. The standard deviation is unchanged by adding or subtracting a specific number to each data value. If each data value is multiplied by a number the standard deviation increases by the number times the original standard deviation. For division the standard deviation is divided by the number.
g. When adding or subtracting the same number to each data value the mean will increase or decrease by that number, but the standard deviation will remain unchanged. When multiplying each data value by the same number the mean or standard deviation will be equal to that number times the original mean or standard deviation. When dividing each data value by the same number the mean or standard deviation will be equal to the original mean or standard deviation divided by that number.

47.
$\overline{X} = 13.3$
Mean Dev $= \frac{|5-13.3|+|9-13.3|+|10-13.3|+|11-13.3|+|11-13.3|}{10}$

$+ \frac{|12-13.3|+|15-13.3|+|18-13.3|+|20-13.3|+|22-13.3|}{10} = 4.36$

48.
a. $Sk = \frac{3(10-8)}{3} = 2$ positively skewed

b. $Sk = \frac{3(42-45)}{4} = -2.25$ negatively skewed

c. $Sk = \frac{3(18.6-18.6)}{1.5} = 0$ symmetric

d. $Sk = \frac{3(98-97.6)}{4} = 0.3$ positively skewed

49.
For $n = 25$, $\overline{X} = 50$, and $s = 3$:

$s\sqrt{n-1} = 3\sqrt{25-1} = 14.7$ $\overline{X} + s\sqrt{n-1} = 50 + 14.7 = 64.7$

67 may be an incorrect data value, since it is beyond the range using the formula $s\sqrt{n-1}$.

EXERCISE SET 3-3

1.
A z score tells how many standard deviations the data value is above or below the mean.

2.
A percentile rank indicates the percentage of data values that fall below the specific rank.

3.
A percentile is a relative measure while a percent is an absolute measure of the part to the total.

4.
A quartile is a relative measure of position obtained by dividing the data set into quarters.

5.
$Q_1 = P_{25},\ Q_2 = P_{50},\ Q_3 = P_{75}$

6.
A decile is a relative measure of position obtained by dividing the data set into tenths.

7.
$\mathrm{D}_1 = \mathrm{P}_{10},\ \mathrm{D}_2 = \mathrm{P}_{20},\ \mathrm{D}_3 = \mathrm{P}_{30}$, etc

8.
$\mathrm{P}_{50},\ \mathrm{Q}_2,\ \mathrm{D}_5$

9.
For Canada:
z $= \frac{X - \overline{X}}{s} = \frac{26 - 29.4}{8.6} = -0.40$
For Italy:
z $= \frac{X - \overline{X}}{s} = \frac{42 - 29.4}{8.6} = 1.47$
For US:
z $= \frac{X - \overline{X}}{s} = \frac{13 - 29.4}{8.6} = -1.91$

10.
For Senator Byrd:
z $= \frac{X - \overline{X}}{s} = \frac{86 - 59.5}{11.5} = 2.30$
For Senator Sununu:
z $= \frac{X - \overline{X}}{s} = \frac{40 - 59.5}{8.6} = -1.70$

11.
a. z $= \frac{X - \overline{X}}{s} = \frac{79 - 76}{5} = 0.6$ b. z $= \frac{70 - 76}{5} = -1.2$

c. z $= \frac{88 - 76}{5} = 2.4$ d. z $= \frac{65 - 76}{5} = -2.2$

e. z $= \frac{77 - 76}{5} = 0.2$

12.
If z $= \frac{X - \overline{X}}{s}$ then $X = zs + \overline{X}$

a. $X = 2(10{,}200) + 54{,}166 = \$74{,}566$ b. $X = -1(10{,}200) + 54{,}166 = \$43{,}966$

c. $X = 0(10{,}200) + 54{,}166 = \$54{,}166$ d. $X = 2.5(10{,}200) + 54{,}166 = \$79{,}666$

e. $X = -1.6(10{,}200) + 54{,}166 = \$37{,}846$

13.
For the statistics test: z $= \frac{75 - 60}{10} = 1.5$ For the accounting test: z $= \frac{36 - 30}{\sqrt{16}} = 1.5$

Neither. The scores have the same relative position.

14.
For student #1: z $= \frac{9650 - 8455}{1865} = 0.64$ For student #2: z $= \frac{12360 - 10326}{2143} = 0.95$

The student from the university (student #2) has a higher relative debt.

15.
a. z $= \frac{16{,}000 - 14{,}090}{3500} = 0.55$ b. z $= \frac{10{,}000 - 14{,}090}{3500} = -1.17$

15. continued

c. To find the number of miles, use $X = zs + \overline{X}$

X = 1.6(3500) + 14,090 = 19,690 miles

X = − 0.5(3500) + 14,090 = 12,340 miles

X = 0(3500) + 14,090 = 14,090 miles

16.

a. $z = \frac{3.2-4.6}{1.5} = -0.93$ b. $z = \frac{630-800}{200} = -0.85$ c. $z = \frac{43-50}{5} = -1.4$

The score in part b is the highest.

17.

For 78: $P = \frac{55+0.5}{64} = 0.867$ or the 87th percentile.

For 66: $P = \frac{31+0.5}{64} = 0.492$ or the 49th percentile.

For 59: $P = \frac{12+0.5}{64} = 0.195$ or the 20th percentile.

For the 90th percentile: $c = \frac{(64)(90)}{100} = 57.6$ or 58th value, which is a score of 79 .

For the 80th percentile: $c = \frac{(64)(80)}{100} = 51.2$ or 52nd value, which is a score of 75.

For the 65th percentile: $c = \frac{(64)(65)}{100} = 41.6$ or 42nd value, which is a score of 69.

18.

a. $5,806 b. $6,563 c. $7,566 d. $8,563

e. 24th f. 67th g. 48th h. 88th

19.

a. 6th b. 24th c. 68th d. 76th e. 94th

f. a. 234 g. 251 h. 263 i. 274 j. 284

20.

a. 375 b. 389 c. 433 d. 477 e. 504

f. 13th g. 40th h. 54th i. 76th j. 92nd

21.

Percentile = $\frac{\text{number of values below + 0.5}}{\text{total number of values}} \cdot 100\%$ Data: 228, 489, 524, 597, 623, 659, 736, 777, 804

For 228, $\frac{0+.5}{9} \cdot 100\%$ = 6th percentile For 489, $\frac{1+.5}{9} \cdot 100\%$ = 17th percentile

For 524, $\frac{2+.5}{9} \cdot 100\%$ = 28th percentile For 597, $\frac{3+.5}{9} \cdot 100\%$ = 39th percentile

For 623, $\frac{4+.5}{9} \cdot 100\%$ = 50th percentile For 659, $\frac{5+.5}{9} \cdot 100\%$ = 61st percentile

For 736, $\frac{6+.5}{9} \cdot 100\%$ = 72nd percentile For 777, $\frac{7+.5}{9} \cdot 100\%$ = 83rd percentile

For 804, $\frac{8+.5}{9} \cdot 100\%$ = 94th percentile

$c = \frac{9(40)}{100} = 3.6$ or 4th data value, which is 597

22.

For 12, $\frac{0+.5}{7} \cdot 100\% = 7^{\text{th}}$ percentile

For 28, $\frac{1+.5}{7} \cdot 100\% = 21^{\text{st}}$ percentile

For 35, $\frac{2+.5}{7} \cdot 100\% = 36^{\text{th}}$ percentile

For 42, $\frac{3+.5}{7} \cdot 100\% = 50^{\text{th}}$ percentile

For 47, $\frac{4+.5}{7} \cdot 100\% = 64^{\text{th}}$ percentile

For 49, $\frac{5+.5}{7} \cdot 100\% = 79^{\text{th}}$ percentile

For 50, $\frac{6+.5}{7} \cdot 100\% = 93^{\text{rd}}$ percentile

$c = \frac{n \cdot p}{100} = \frac{7(60)}{100} = 4.2$ or 5 Hence, 47 is the closest value to the 60^{th} percentile.

23.

$\text{Percentile} = \frac{\text{number of values below + 0.5}}{\text{total number of values}} \cdot 100\%$ Data: 1.1, 1.7, 1.9, 2.1, 2.2, 2.5, 3.3, 6.2, 6.8, 20.3

For 1.1, $\frac{0+.5}{10} \cdot 100\% = 5^{\text{th}}$ percentile

For 1.7, $\frac{1+.5}{10} \cdot 100\% = 15^{\text{th}}$ percentile

For 1.9, $\frac{2+.5}{10} \cdot 100\% = 25^{\text{th}}$ percentile

For 2.1, $\frac{3+.5}{10} \cdot 100\% = 35^{\text{th}}$ percentile

For 2.2, $\frac{4+.5}{10} \cdot 100\% = 45^{\text{th}}$ percentile

For 2.5, $\frac{5+.5}{10} \cdot 100\% = 55^{\text{nd}}$ percentile

For 3.3, $\frac{6+.5}{10} \cdot 100\% = 65^{\text{th}}$ percentile

For 6.2, $\frac{7+.5}{10} \cdot 100\% = 75^{\text{th}}$ percentile

For 6.8, $\frac{8+.5}{10} \cdot 100\% = 85^{\text{th}}$ percentile

For 20.3, $\frac{9+.5}{10} \cdot 100\% = 95^{\text{th}}$ percentile

$c = \frac{10(40)}{100} = 4$ average the 4th and 5th values: $P_{40} = \frac{2.1+2.2}{2} = 2.15$

24.

$\text{Percentile} = \frac{\text{number of values below + 0.5}}{\text{total number of values}} \cdot 100\%$ Data: 5, 12, 15, 16, 20, 21

For 5, $\frac{0+.5}{6} \cdot 100\% = 8^{\text{th}}$ percentile

For 12, $\frac{1+.5}{6} \cdot 100\% = 25^{\text{th}}$ percentile

For 15, $\frac{3+.5}{6} \cdot 100\% = 42^{\text{nd}}$ percentile

For 16, $\frac{4+.5}{6} \cdot 100\% = 58^{\text{th}}$ percentile

For 20, $\frac{5+.5}{6} \cdot 100\% = 75^{\text{th}}$ percentile

For 21, $\frac{5+.5}{6} \cdot 100\% = 92^{\text{nd}}$ percentile

$c = \frac{6(33)}{100} = 1.98$ or 2^{nd} data value, which is 12.

25.

To find Q_1, find P_{25}:

$c = \frac{(14)25)}{100} = 3.5$, round up to 4. Q_1 is at the 4th value, which is 16.

To find Q_3, find P_{75}:

$c = \frac{(14)(75)}{100} = 10.5$, round up to 11. Q_3 is at the 11th value, which is 27.1.

26.

To find Q_1, find P_{25}:

$c = \frac{(12)25)}{100} = 3$, average the 3rd and 4th values. $Q_1 = \frac{120{,}000 + 160{,}000}{2} = 140{,}000$

To find Q_3, find P_{75}:

$c = \frac{(12)(75)}{100} = 9$, average the 9th and 10th values. $Q_3 = \frac{375{,}000 + 375{,}000}{2} = 375{,}000$

27.
To find Q_1, find P_{25}:
$c = \frac{(11)25)}{100} = 2.75$, round up to 3. $Q_1$is at the 3rd value, which is 6.2.

To find Q_3, find P_{75}:
$c = \frac{(11)(75)}{100} = 8.25$, round up to 9. Q_3 is at the 9th value, which is 7.2.

28.
To find Q_1, find P_{25}:
$c = \frac{(9)25)}{100} = 2.25$, round up to 3. $Q_1$is at the 3rd value, which is 6.
Note: TI83 answer is 4.5.

To find Q_3, find P_{75}:
$c = \frac{(9)(75)}{100} = 6.75$, round up to 7. Q_3 is at the 7th value, which is 7.
Note: TI83 answer is 24.

29.
a. 3 16 17 18 19 20 21 22
↑ ↑ ↑
Q_1 MD Q_3

MD = $\frac{18+19}{2}$ = 18.5

For Q_1: $Q_1 = \frac{16+17}{2} = 16.5$ For Q_3: $Q_3 = \frac{20+21}{2} = 20.5$

$Q_3 - Q_1 = 20.5 - 16.5 = 4$ and $4(1.5) = 6$. $16.5 - 6 = 10.5$ and $20.5 + 6 = 26.5$.

Only the value 3 lies outside the range of 10.5 to 26.5 and is a suspected outlier.

b. 14 16 17 18 19 20 24 31 32 54
↑ ↑ ↑
Q_1 MD Q_3

MD = $\frac{19+20}{2}$ = 19.5

For Q_1: Q_1 = 17, the median of 14, 16, 17, 18, and 19.

For Q_3: Q_3 = 31, the median of 20, 24, 31, 32, and 54.

$Q_3 - Q_1$: $31 - 17 = 14$ and $14(1.5) = 21$. $17 - 21 = -4$ and $31 + 21 = 52$. Only the value 54 lies outside the range of -4 to 52 and is a suspected outlier.

c. 200 321 327 343 350
↑ ↑ ↑
Q_1 MD Q_3

MD = 327

For Q_1: $Q_1 = \frac{200+321}{2} = 260.5$ For Q_3: $Q_3 = \frac{343+350}{2} = 346.5$

$Q_3 - Q_1$: $346.5 - 260.5 = 86$ and $86(1.5) = 129$. $260.5 - 129 = 131.5$ and $260.5 + 129 = 475.5$.

Since all the values fall within the range of 131.5 to 475.5, there are no outliers.

30.

a. 72 84 85 86 88 97 100
↑ ↑ ↑
Q_1 MD Q_3

MD = 86
For Q_1: Q_1 = 84 For Q_3: Q_3 = 97.
$Q_3 - Q_1$: $97 - 84 = 13$ and $13(1.5) = 19.5$. $84 - 19.5 = 64.5$ and $97 + 19.5 = 116.5$.

Since all values fall within the range of 64.5 to 116.5, there are no outliers.

b. 116 118 119 122 125 145
↑ ↑ ↑
Q_1 MD Q_3

MD = $\frac{119+122}{2}$ = 120.5
For Q_1: Q_1 = 118. For Q_3: Q_3 = 125.
$Q_3 - Q_1$: $125 - 118 = 7$ and $7(1.5) = 10.5$. $118 - 10.5 = 107.5$ and $125 + 10.5 = 125.5$.

Only the value 145 is outside the range of 107.5 to 135.5 and is a suspected outlier.

c. 13 14 15 16 18 19 20 27 36
↑ ↑ ↑
Q_1 MD Q_3

MD = 18
For Q_1: Q_1 = $\frac{14+15}{2}$ = 14.5 For Q_3: Q_3 = $\frac{20+27}{2}$ = 23.5
$Q_3 - Q_1$: $23.5 - 14.5 = 9$ and $9(1.5) = 13.5$. $14.5 - 13.5 = 1$ and $23.5 + 13.5 = 37$.

Since all values fall within the range of 1 to 37, there are no outliers.

31.

a. 5, 12, 16, 25, 32, 38 Q_1 = 12, Q_2 = 20.5, Q_3 = 32
Midquartile = $\frac{12+32}{2}$ = 22 Interquartile range: $32 - 12 = 20$

b. 53, 62, 78, 94, 96, 99, 103 Q_1 = 62, Q_2 = 94, Q_3 = 99
Midquartile = $\frac{62+99}{2}$ = 80.5 Interquartile range: $99 - 62 = 37$

32.

If $s^2 = 250$, then $s = \sqrt{250} = 15.81$
Using a score of 142:
$142 = -0.5(15.81) + \overline{X}$
$149.9 \approx \overline{X}$

33.

Tom's score is 158. Harry's score can be calculated based on his z score: $X = 2(18) + 125 = 161$.

Since the data are normally distributed, 95% fall within 2 standard deviations of the mean (using the Empirical Rule). Thus, $125 \pm 2(18)$ gives a range of 89 to 161, and a score of 161 is the 95th percentile. Since Dick scored in the 98th percentile, his raw score must be higher than 161.

Therefore, Tom's score is the lowest followed by Harry, with Dick's score being the highest.

EXERCISE SET 3-4

1. Data arranged in order: 6, 8, 12, 19, 27, 32, 54
Minimum: 6
Q_1: 8
Median: 19
Q_3: 32
Maximum: 54
Interquartile Range: $32 - 8 = 24$

2. Data arranged in order: 7, 16, 19, 22, 48
Minimum: 7
Q_1: $\frac{7+16}{2} = 11.5$
Median: 19
Q_3: $\frac{22+48}{2} = 35$
Maximum: 48
Interquartile Range: $35 - 11.5 = 23.5$

3. Data arranged in order: 188, 192, 316, 362, 437, 589
Minimum: 188
Q_1: 192
Median: $\frac{316+362}{2} = 339$
Q_3: 437
Maximum: 589
Interquartile Range: $437 - 192 = 245$

4. Data arranged in order: 147, 156, 243, 303, 543, 632
Minimum: 147
Q_1: 156
Median: $\frac{243+303}{2} = 273$
Q_3: 543
Maximum: 632
Interquartile Range: $543 - 156 = 387$

5. Data arranged in order: 14.6, 15.5, 16.3, 18.2, 19.8
Minimum: 14.6
Q_1: $\frac{14.6+15.5}{2} = 15.05$
Median: 16.3
Q_3: $\frac{18.2+19.8}{2} = 19.0$
Maximum: 19.8
Interquartile Range: $19.0 - 15.05 = 3.95$

6. Data arranged in order: 2.2, 3.7, 3.8, 4.6, 6.2, 9.4, 9.7
Minimum: 2.2
Q_1: 3.7
Median: 4.6
Q_3: 9.4
Maximum: 9.7
Interquartile Range: $9.4 - 3.7 = 5.7$

7. Minimum: 3
Q_1: 5
Median: 8
Q_3: 9

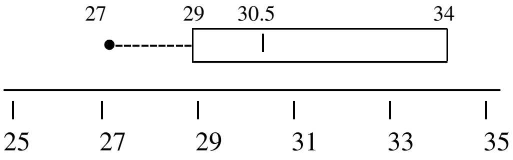
7. continued
Maximum: 11
Interquartile Range: $9 - 5 = 4$

8. Minimum: 200
Q_1: 225
Median: 275
Q_3: 300
Maximum: 325
Interquartile Range: $300 - 225 = 75$

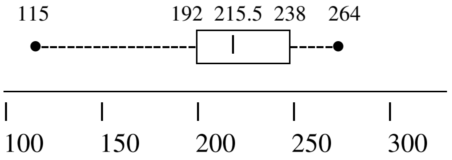
9. Minimum: 55
Q_1: 65
Median: 70
Q_3: 90
Maximum: 95
Interquartile Range: $90 - 65 = 25$

10. Minimum: 2000
Q_1: 3000
Median: 4000
Q_3: 5000
Maximum: 6000
Interquartile Range: $5000 - 3000 = 2000$

11.
MD $= \frac{30+31}{2} = 30.5$ $\quad Q_1 = 29$ $\quad Q_3 = 34$

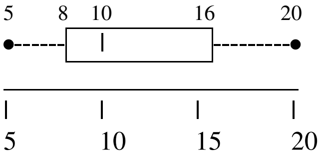


12.
MD $= \frac{214+217}{2} = 215.5$ $\quad Q_1 = 192$ $\quad Q_3 = 238$



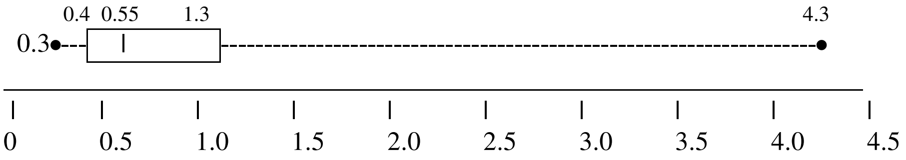
The distribution is slightly left-skewed.

13.
MD $= 10$ $\quad Q_1 = 8$ $\quad Q_3 = 16$



13. continued
The box plot of the data is somewhat positively skewed.

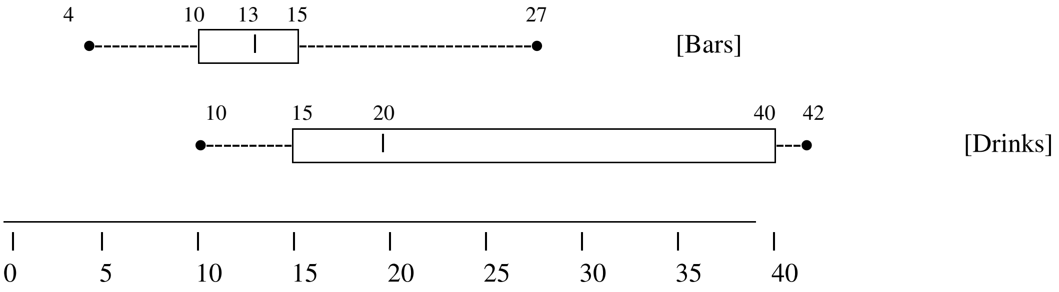
14.
MD $= \frac{0.5+0.6}{2} = 0.55$ $\quad Q_1 = 0.4$ $\quad Q_3 = 1.3$



The box plot of the data is somewhat positively skewed.

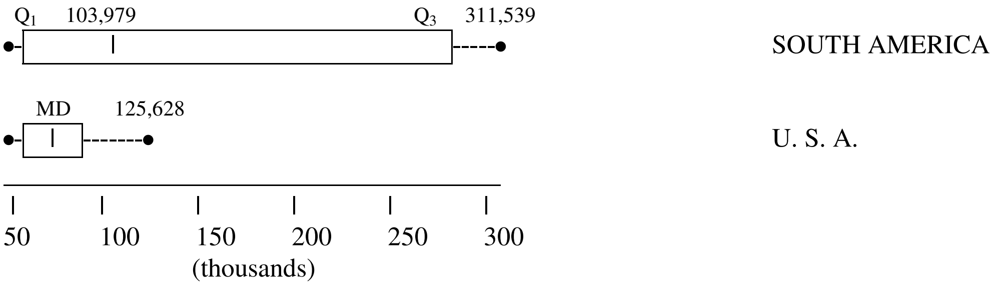
15.
For bars:
MD $= \frac{12+14}{2} = 13$ $\quad Q_1 = 10$ $\quad Q_3 = 15$

For drinks:
MD $= 20$ $\quad Q_1 = 15$ $\quad Q_3 = 40$



The data for protein drinks have a higher median amount of grams and are more variable.

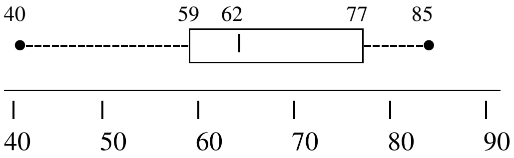
16.
For USA: min = 50,000, max = 125,628, MD = 72,100, $Q_1 = 57{,}642.5$, and $Q_3 = 85{,}004$
For South America: min = 46,563, max = 311,539, MD = 103,979, $Q_1 = 56{,}242$, and $Q_3 = 274{,}026$



The range and variation of the capacity of the dams in South America is considerably larger than those of the United States.

17.
MD $= \frac{60+64}{2} = 62$ $\quad Q_1 = 59$ $\quad Q_3 = 77$

17. continued



18.

(a)

For April: $\overline{X} = 138$

For May: $\overline{X} = 391.7$

For June: $\overline{X} = 292$

For July: $\overline{X} = 143$

The month with the highest mean number of tornadoes is May.

(b)

For 2005: $\overline{X} = 177.3$

For 2004: $\overline{X} = 256.5$

For 2003: $\overline{X} = 289.8$

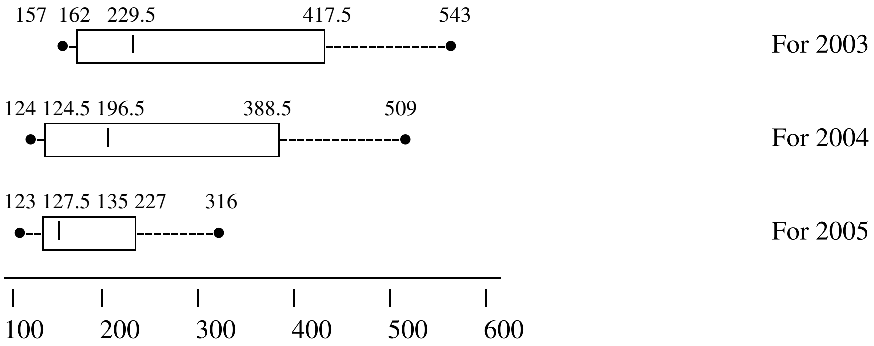
The year with the highest mean number of tornadoes is 2003.

(c) The 5-number summaries for each year are:

For 2005: 123, 127.5, 135, 227, 316

For 2004: 124, 124.5, 196.5, 388.5, 509

For 2003: 157, 162, 229.5, 417.5, 543



The distribution for 2003, 2004, and 2005 are positively skewed. The data for 2005 appears to be the least variable.

19. Data arranged in order: 39, 39, 42, 43, 43, 53, 54, 66, 91, 97

Minimum: 39

Q_1: 42

Median: $\frac{43+53}{2} = 48$

Q_3: 66

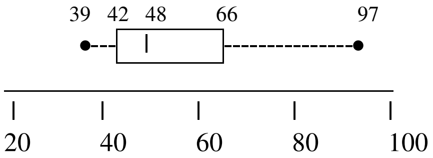
Maximum: 97

Interquartile Range: $66 - 42 = 24$

$1.5(24) = 36$ for mild outliers; $3(24) = 72$ for extreme outliers

There are no outliers.

19. continued



REVIEW EXERCISES - CHAPTER 3

1.

$\overline{\text{X}} = \frac{\sum X}{n} = \frac{272}{10} = 27.2$

Data arranged in order: 17, 17, 17, 18, **19, 19**, 26, 28, 52, 59

MD $= \frac{19+19}{2} = 19$

Mode $= 17$

MR $= \frac{17+59}{2} = 38$

2.

Attacks:

$\overline{\text{X}} = \frac{318}{5} = 63.6$

Data arranged in order: 57, 61, **64**, 65, 71

MD $= 64$

no mode

MR $= \frac{57+71}{2} = 64$

Deaths:

$\overline{\text{X}} = \frac{20}{5} = 4$

Data arranged in order: 1, 4, **4**, 4, 7

MD $= 4$

Mode $= 4$

MR $= \frac{1+7}{2} = 4$

3.

Class	X_m	f	$f \cdot X_m$	$f \cdot X_m^2$	cf
1 - 3	2	1	2	4	1
4 - 6	5	4	20	100	5
7 - 9	8	5	40	320	10
10 - 12	11	1	11	121	11
13 - 15	14	1	14	196	12
		12	87	741	

$\overline{\text{X}} = \frac{\sum f \cdot X_m}{n} = \frac{87}{12} = 7.3$

3. continued
Modal Class $= 7 - 9$ or $6.5 - 9.5$

4.

X_m	f	$f \cdot X_m$	$f \cdot X_m^2$
491	4	1964	964,324
518	6	3108	1,609,944
545	2	1090	594,050
572	2	1144	654,368
599	2	1198	717,602
	16	8504	4,540,288

$\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{8504}{16} = 531.5$
Modal Class $= 505 - 531$

5.
$\overline{X} = \frac{\sum w \cdot X}{\sum w} = \frac{1.6(1.4) + 0.8(0.8) + 0.4(0.3) + 1.8(1.6)}{1.4 + 0.8 + 0.3 + 1.6} = 1.43$ viewers per household

6.
$\overline{X} = \frac{0.3(10{,}000) + 0.5(3000) + 0.2(1000)}{0.3 + 0.5 + 0.2} = \$4{,}700.00$

7.
Range $= 316 - 10 = 306$
$s^2 = \frac{n \sum f \cdot X_m^2 - (\sum f \cdot X_m)^2}{n(n-1)} = \frac{12(212{,}804) - 1264^2}{12(12-1)} = 7242.06$

$s = \sqrt{7242.06} = 85.1$

8.
Range $= 75 - 47 = 28$
$s^2 = \frac{n \sum f \cdot X_m^2 - (\sum f \cdot X_m)^2}{n(n-1)} = \frac{13(41{,}379) - 725^2}{13(13-1)} = 78.85$

$s = \sqrt{78.9} = 8.9$

9.

Class Boundaries	X_m	f	$f \cdot X_m$	$f \cdot X_m^2$	cf
12.5 - 27.5	20	6	120	2400	6
27.5 - 42.5	35	3	105	3675	9
42.5 - 57.5	50	5	250	12,500	14
57.5 - 72.5	65	8	520	33,800	22
72.5 - 87.5	80	6	480	38,400	28
87.5 - 102.5	95	$\underline{2}$	$\underline{190}$	$\underline{18{,}050}$	30
		30	1665	108,825	

a. $\overline{X} = \frac{\sum f \cdot X_m}{n} = \frac{1665}{30} = 55.5$

b. Modal class $= 57.5 - 72.5$

c. $s^2 = \frac{n \sum f \cdot X_m^2 - (\sum f \cdot X_m)^2}{n(n-1)} = \frac{30(108{,}825) - 1665^2}{30(30-1)} = 566.1$

9. continued

d. $s = \sqrt{566.1} = 23.8$

10.

Class	X_m	f	$f \cdot X_m$	$f \cdot X_m^2$	cf
10 - 12	11	6	66	726	6
13 - 15	14	4	56	784	10
16 - 18	17	14	238	4046	24
19 - 21	20	15	300	6000	39
22 - 24	23	8	184	4232	47
25 - 27	26	2	52	1352	49
28 - 30	29	1	29	841	50
		50	925	17981	

a. $\overline{X} = \frac{925}{50} = 18.5$

b. Modal Class $= 19 - 21$

c. $s^2 = \frac{50(17{,}981) - 925^2}{50(50-1)} = 17.7$

d. $s = \sqrt{17.7} = 4.2$

11.

$s \approx \frac{24}{4} = 6$

12.

$s \approx \frac{56}{4} = 14$

13.

Textbooks: C. Var $= \frac{5}{16} = 0.3125$ or 31.25%

Ages: C. Var $= \frac{8}{43} = 0.186$ or 18.6%

The number of books is more variable.

14.

Magazines: C. Var $= \frac{s}{\overline{X}} = \frac{12}{56} = 0.214$ or 21.4%

Year: C. Var $= \frac{s}{\overline{X}} = \frac{2.5}{6} = 0.417$ or 41.7%

The number of years is more variable.

15.

$\overline{X} = 0.32 \quad s = 0.03 \quad k = 2$

$0.32 - 2(0.03) = 0.26$ and $0.32 + 2(0.03) = 0.38$

At least 75% of the values will fall between \$0.26 and \$0.38.

16.

$\overline{X} = \$58{,}500 \quad s = \$11{,}200$

a. $58{,}500 + 11{,}200k = 69{,}700$

$k = 1$

Since Chebyshev's Theorem is appropriate only for $k > 1$, no information can be obtained about the percentage of workers earning between \$47,300 and \$69,700.

b. $58{,}500 + 11{,}200k = 80{,}900$

$k = 2$

$1 - \frac{1}{2^2} = 0.75$ or 75%

Hence at most $100\% - 75\% = 25\%$ earn more than \$80,900.

16. continued

c. $58{,}500 + 11{,}200k = 100{,}000$

$k = 3.7054$

$1 - \frac{1}{3.7054^2} = 0.927$ or 92.7%

Hence at most $100\% - 92.7\% = 7.3\%$ earn more than \$100,000.

17.

$\overline{\text{X}} = 54 \quad \text{s} = 4 \quad 60 - 54 = 6 \quad \text{k} = \frac{6}{4} = 1.5 \quad 1 - \frac{1}{1.5^2} = 1 - 0.44 = 0.56$ or 56%

18.

$\overline{\text{X}} = 231 \quad \text{s} = 5 \quad 243 - 231 = 12 \quad \text{k} = \frac{12}{5} = 2.4 \quad 1 - \frac{1}{2.4^2} = 0.83$ or 83%

19. By the Empirical Rule, 68% of the scores will be within 1 standard deviation of the mean.

$29.7 + 1(6) = 35.7$

$29.7 - 1(6) = 23.7$

Then 68% of commuters will get to work between 23.7 and 35.7 minutes.

20.

Since the data are normally distributed, the Empirical Rule can be used. For 95%, use $k = 2$ standard deviations.

$\overline{X} \pm 2s = 44 \pm 2(9)$

Thus, 95% of the times are between 26 and 62 minutes.

21.

$\overline{X} = 68.09$

$s = 13.95$

a. $z = \frac{80^\circ - 68.09^\circ}{13.95} = 0.86$

b. $z = \frac{56^\circ - 68.09^\circ}{13.95} = -0.87$

22.

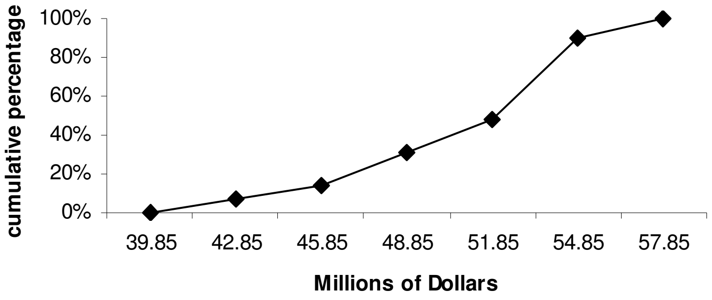
a. $\text{z} = \frac{82 - 85}{6} = -0.5$

b. $\text{z} = \frac{56 - 60}{5} = -0.8$

The exam in part a has a better relative position.

23.

a.

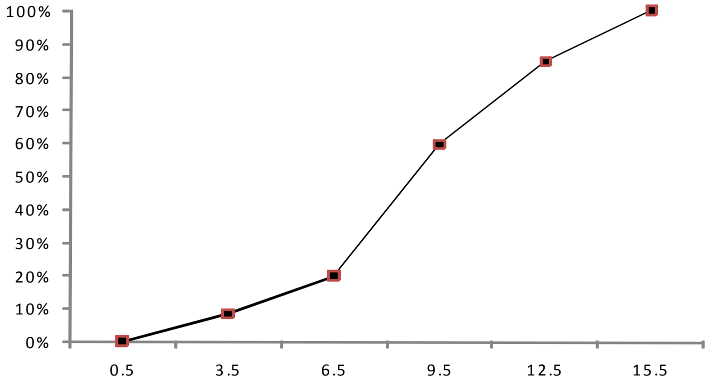


b. $\text{P}_{35} = 50$; $\text{P}_{65} = 53$; $\text{P}_{85} = 55$ (answers are approximate)

c. 44 = 10th percentile; 48 = 26th percentile; 54 = 78th percentile (answers are approximate)

24.

a.



b. $P_{20} = 6.5$; $P_{50} = 9.4$; $P_{70} = 14.3$ (answers are approximate)

c. 5 = 13th percentile; 10 = 67th percentile; 14 = 93rd percentile (answers are approximate)

25.

a. 400 506 511 514 517 521
↑ ↑
Q_1 Q_3

For Q_1: $c = \frac{np}{100} = \frac{6(25)}{100} = 1.5$ round up to 2 $Q_1 = 506$

For Q_3: $c = \frac{np}{100} = \frac{6(75)}{100} = 4.5$ round up to 5 $Q_3 = 517$

$Q_3 - Q_1 = 517 - 506 = 11$; $11(1.5) = 16.5$; $506 - 16.5 = 489.5$ and $517 + 16.5 = 533.5$
Therefore, only the value 400 lies outside the range of 489.5 to 533.5 and is a suspected outlier.

b. 3 6 7 8 9 10 12 14 16 20
↑ ↑
Q_1 Q_3

For Q_1: $c = \frac{np}{100} = \frac{10(25)}{100} = 2.5$ round up to 3 $Q_1 = 7$

For Q_3: $c = \frac{np}{100} = \frac{10(75)}{100} = 7.5$ round up to 8 $Q_3 = 14$

$Q_3 - Q_1 = 14 - 7 = 7$; $7(1.5) = 10.5$; $7 - 10.5 = -3.5$ and $14 + 10.5 = 24.5$
Since all values fall within the range of -3.5 to 24.5, there are no outliers.

26.

a. 5 13 14 18 19 25 26 27
↑ ↑
$Q_1 = 13.5$ $Q_3 = 25.5$

For Q_1: $c = \frac{np}{100} = \frac{8(25)}{100} = 2.0$ Use the value between the 2nd and 3rd position: $Q_1 = \frac{13+14}{2} = 13.5$

For Q_3: $c = \frac{np}{100} = \frac{8(75)}{100} = 6.0$ Use the value between the 6th and 7th position: $Q_3 = \frac{25+26}{2} = 25.5$

$Q_3 - Q_1 = 25.5 - 13.5 = 12$; $12(1.5) = 18$: $13.5 - 18 = -4.5$ and $25.5 + 18 = 43.5$
Since all values fall within the range of -4.5 to 43.5, there are no outliers.

26. continued

b. 112 116 129 131 153 157 192

$$\begin{array}{ccccccc} 112 & 116 & 129 & 131 & 153 & 157 & 192 \\ & \uparrow & & & & \uparrow & \\ & Q_1 & & & & Q_3 & \end{array}$$

For Q_1: $c = \frac{np}{100} = \frac{7(25)}{100} = 1.75$ Round up to 2. $Q_1 = 116$

For Q_3: $c = \frac{np}{100} = \frac{7(75)}{100} = 5.25$ Round up to 6. $Q_3 = 157$

$Q_3 - Q_1 = 157 - 116 = 41$; $41(1.5) = 61.5$: $116 - 61.5 = 54.5$ and $157 + 61.5 = 218.5$

Since all values fall within the range of 54.5 to 218.5, there are no outliers.

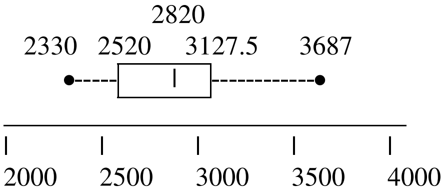
27.

Top Nine Movies:

MD = 2820 $Q_1 = 2520.5$ $Q_3 = 3127.5$

Top Ten Movies:

MD = 2699.5 $Q_1 = 2516$ $Q_3 = 3044$



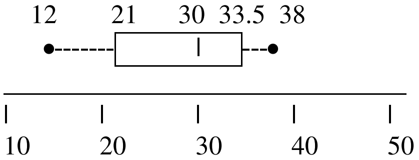
Top Nine

The range is much larger.

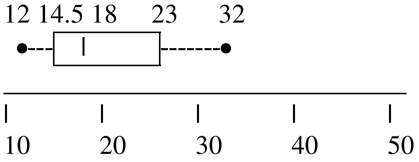
28. The five-number summaries are:

Before: 12, 21, 30, 33.5, 38

After: 12, 14.5, 18, 23, 32



Before

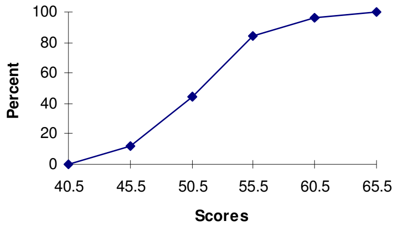


After

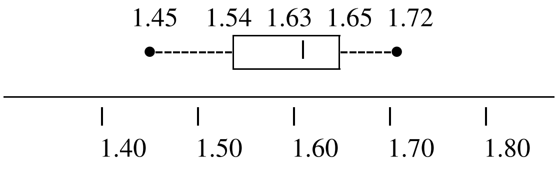
CHAPTER 3 QUIZ

1. True
2. True
3. False
4. False
5. False
6. False
7. False
8. False
9. False

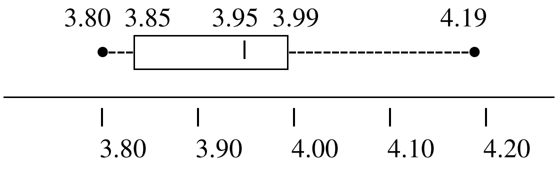
10. c
11. c
12. a and b
13. b
14. d
15. b
16. Statistic
17. Parameters, statistics
18. Standard deviation
19. σ
20. Midrange
21. Positively
22. Outlier
23. a. 15.3 b. 15.5 c. 15, 16, 17 d. 15 e. 6 f. 3.57 g. 1.9
24. a. 6.4 b. $6-8$ c. 11.6 d. 3.4
25. 4.46 or 4.5
26. 0.107 or 10.7%, 0.114 or 11.4%; newspapers sold in a convenience store are more variable
27. 88.89%
28. For above 1129: 16%; For above 799: 97.5%
29. $s \approx \frac{18}{4} = 4.5$
30. -0.75; -1.67; science
31.
a.



b. 47; 55; 64 c. 56th percentile; 6th percentile; 99th percentile
32.
For Pre-buy:
MD = 1.625 or 1.63 $Q_1 = 1.54$ $Q_3 = 1.65$
For No Pre-buy:
MD = 3.95 $Q_1 = 3.85$ $Q_3 = 3.99$



Pre-Buy Cost



No Pre-buy Cost

The cost of pre-buy gas is much less than to return the car without filling it with gas. The variability of the return without filling with gas is larger than the variability of the pre-buy gas.