Chapter 3 Averages and Variation

Section 3.1

1. Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{156+161+152+\cdots+157}{12}$

$$= \frac{1876}{12}$$

$$= 156.33$$

The mean is 156.33.

Organize the data from smallest to largest.

144	148	152	153	156	157
157	157	161	161	162	168

To find the median, add the two middle values and divide by 2 since there is an even number of values.

$$\text{Median} = \frac{157+157}{2} = 157$$

The median is 157.

The mode is 157 because it is the value that occurs most frequently.

A gardener in Colorado should look at seed and plant descriptions to determine if the plant can thrive and mature in the designated number of frost-free days. The mean, median, and mode are all close. About half the locations have 157 or fewer frost-free days.

2. Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{11+29+54+\cdots+46}{12}$

$$= \frac{542}{12}$$

$$= 45.17$$

The mean is 45.17.

Organize the data from smallest to largest.

11	29	41	46	46	46
47	49	54	54	59	60

To find the median, add the two middle values and divide by 2 since there is an even number of values.

$$\text{Median} = \frac{46+47}{2} = 46.5$$

The median is 46.5.

The mode is 46 because it is the value that occurs most frequently.

3. Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{146+152+168+\cdots+144}{14}$

$= \frac{2342}{14}$

$= 167.3$

The mean is 167.3°F.

Organize the data from smallest to largest.

144	146	152	152	165	168	168
174	178	178	178	179	180	180

To find the median, add the two middle values and divide by 2 since there is an even number of values.

$$\text{Median} = \frac{168+174}{2} = 171$$

The median is 171° F.

The mode is 178° F because it is the value that occurs most frequently.

4. Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{13+10+7+\cdots+8}{18}$

$= \frac{111}{18}$

$= 6.2$

The mean is 6.2.

Organize the data from smallest to largest.

2	2	2	3	3	4	4	4	5
7	7	7	7	8	8	10	13	15

To find the median, add the two middle values and divide by 2 since there is an even number of values.

$$\text{Median} = \frac{5+7}{2} = 6$$

The median is 6.

The mode is 7 because it is the value that occurs most frequently.

5. First organize the data from smallest to largest. Then compute the mean, median, and mode.

(a) Upper Canyon

1	1	1	2	3	3	3	3	4	6	9

Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{36}{11} \approx 3.27$

Median = 3 (middle value)

Mode = 3 (occurs most frequently)

(b) Lower Canyon

0	0	1	1	1	1	2	2	3	6	7	8	13	14

$$\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{59}{14} \approx 4.21$$

$$\text{Median} = \frac{2+2}{2} = 2$$

Mode = 1 (occurs most frequently)

(c) The mean for the Lower Canyon is greater than that of the Upper Canyon. However, the median and mode for the Lower Canyon are less than those of the Upper Canyon.

(d) 5% of 14 is 0.7 which rounds to 1. So, eliminate one data value from the bottom of the list and one from the top. Then compute the mean of the remaining 12 values.

$$5\% \text{ trimmed mean} = \frac{\Sigma x}{n} = \frac{45}{12} = 3.75$$

Now this value is closer to the Upper Canyon mean.

6. (a) First arrange the data from smallest to largest. Then compute the mean, median, and mode.

$$\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{1050}{40} \approx 26.3$$

The mean is 26.3 yr.

$$\text{Median} = \frac{25+26}{2} = 25.5$$

The median is 25.5 yr.

Mode = 25

The mode is 25 yr.

(b) The median may represent the age most accurately. The answers are very close.

7. (a) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{93+80+15+\cdots+13}{12}$

$$= \frac{346}{12}$$

$$\approx 28.83$$

The mean is 28.83 thousand dollars.

(b) $\text{Median} = \frac{18+19}{2} = 18.5$

The median is 18.5 thousand dollars.

The median best describes the salary of the majority of employees, since the mean is influenced by the high salaries of the president and vice president.

(c) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{15+25+14+\cdots+13}{10}$

$= \frac{173}{10}$

$= 17.3$

The mean is 17.3 thousand dollars.

$\text{Median} = \frac{16+18}{2} = 17$

The median is 17 thousand dollars.

(d) Without the salaries for the two executives, the mean and the median are closer, and both reflect the salary of most of the other workers more accurately. The mean changed quite a bit, while the median did not, a difference that indicates that the mean is more sensitive to the absence or presence of extreme values.

8. (a) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{9+6+10+\cdots+8}{10} = \frac{74}{10} = 7.4$

$\text{Median} = \frac{8+8}{2} = 8$

Mode = 8 (occurs most frequently)

(b) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{74+36+51+30}{13} = \frac{191}{13} \approx 14.69$

Median = 8 (middle value)

Mode = 8 (occurs most frequently)

(c) The mean is most affected by extreme values.

9. (a) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{15+12+\cdots+15}{7} = \frac{102}{7} \approx 14.57$

Median = 15 (middle value)

Mode = 15 (occurs most frequently)

(b) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{102+57+62}{9} = \frac{221}{9} \approx 24.56$

Median = 15 (middle value)

Mode = 15 (occurs most frequently)

(c) The mean is most affected by extreme values.

10. (a) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{5.2+3.3+\cdots+1.8}{7} = \frac{21.6}{7} \approx 3.09$

Median = 2.9 (middle value)

Mode = 3.3 (occurs most frequently)

(b) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{41.9 + 7.7 + \cdots + 5.1}{8} = \frac{88.2}{8} = 11.025$

$\text{Median} = \frac{6.6 + 6.9}{2} = 6.75$

Mode = 6.6 (occurs most frequently)

Data with first time omitted:

$$\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{46.3}{7} \approx 6.61$$

Median = 6.6 (middle value)

Mode = 6.6 (occurs most frequently)

The mean is very sensitive to extreme values.

11. (a) $\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{2723}{20} = 136.15$

The mean is \$136.15.

The median is \$66.50.

The mode is \$60.

(b) 5% of 20 is 1. Eliminate one data value from the bottom and one from the top of the ordered data. In this case eliminate \$40 and \$500.

$$\text{Mean} = \bar{x} = \frac{\Sigma x}{n} = \frac{2183}{18} \approx 121.28$$

The 5% trimmed mean is \$121.28.

Yes, the trimmed mean more accurately reflects the general level of the daily rental cost, but is still higher than the median.

(c) Median. The low and high prices would be helpful also.

12. (a) Since this data is at the ratio level of measurement, the mean, median, and mode (if it exists) can be used to summarize the data.

(b) Since this data is at the nominal level of measurement, only the mode (if it exists) can be used to summarize the data.

(c) Since this data is at the ratio level of measurement, the mean, median, and mode (if it exists) can be used to summarize the data.

13. (a) Since this data is at the nominal level of measurement, only the mode (if it exists) can be used to summarize the data.

(b) Since this data is at the ratio level of measurement, the mean, median, and mode (if it exists) can be used to summarize the data.

(c) The mode can be used (if it exists). If a 24-hour clock is used, then the data is at the ratio level of measurement, so the mean and median may be used as well.

14. Discussion question.

15. (a) If the largest data value is *replaced* by a larger value, the mean will increase because the sum of the data values will increase, but the number of them will remain the same. The median will not change. The same value will still be in the eighth position when the data are ordered.

(b) If the largest value is replaced by a value that is smaller (but still higher than the median), the mean will decrease because the sum of the data values will decrease. The median will not change. The same value will be in the eighth position in increasing order.

(c) If the largest value is replaced by a value that is smaller than the median, the mean will decrease because the sum of the data values will decrease. The median also will decrease because the former value in the eighth position will move to the ninth position in increasing order. The median will be the new value in the eighth position.

16. Answers may vary. Some examples are

(a)	1	2	2	2	3
(b)	1	2	2	2	13
(c)	1	1	5	5	5
(d)	1	2	4	5	5
(e)	−2	−1	0	1	2

17. Answers will vary according to data collected.

Section 3.2

1. (a) Range = largest value − smallest value

$$= 58 - 4 = 54$$

The range is 54 deer/km^2.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{251}{12} \approx 20.9$$

The sample mean is 20.9 deer/km^2.

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{2474.9}{11} \approx 225.0$$

The sample variance is 225.0.

$$s = \sqrt{s^2} = \sqrt{225.0} = 15.0$$

The sample standard deviation is 15.0 deer/km^2.

(b) $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{15.0}{20.9} \cdot 100 \approx 71.8\%$

s is 71.8% of $\bar{x}$.

Since the standard deviation is about 71.8% of the mean, there is considerable variation in the distribution of deer from one part of the park to another.

2. (a) Range = largest value − smallest value

$$= 78.6 - 17.8 = 60.8$$

The range is 60.8%.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{540.8}{10} \approx 54.1$$

The mean is 54.1%.

(b) $s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{3400}{9} \approx 377.78$

The sample variance is 377.78.

$s = \sqrt{s^2} = \sqrt{377.78} \approx 19.44$

The standard deviation is 19.44%.

(c) $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{19.44}{54.1} \cdot 100 \approx 35.9\%$

s is 35.9% of $\bar{x}$.

3. (a) Range = 90.3 − 12.7 = 77.6

The range is 77.6%.

$\bar{x} = \frac{\Sigma x}{n} = \frac{556.7}{10} \approx 55.7$

The mean is 55.7%.

(b) $s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{4833}{9} \approx 537$

The sample variance is 537.

$s = \sqrt{s^2} = \sqrt{537} \approx 23.17$

The standard deviation is 23.17%.

(c) $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{23.17}{55.7} \cdot 100 \approx 41.6\%$

s is 41.6% of $\bar{x}$.

This CV is larger than the CV for geese. So, nesting success rates for ducks have greater relative variability.

4. (a) Range = 14.1 − 6.8 = 7.3

$\bar{x} = \frac{\Sigma x}{n} = \frac{63.7}{7} = 9.1$

$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{53.28}{6} = 8.88$

$s = \sqrt{s^2} = \sqrt{8.88} \approx 2.98$

$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{2.98}{9.1} \cdot 100 = 32.7\%$

(b) Range = 31.0 − 19.1 = 11.9

$$\bar{x} = \frac{\Sigma x}{n} = \frac{182.9}{7} = 26.1$$

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{118.71}{6} \approx 19.79$$

$$s = \sqrt{s^2} = \sqrt{19.79} \approx 4.45$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{4.45}{26.1} \cdot 100 = 17.0\%$$

(c) More relatively consistent productivity at a higher average level.

5. (a) Pax $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{11.56}{11.69} \cdot 100 \approx 98.9\%$

Vanguard $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{12.50}{5.61} \cdot 100 \approx 222.8\%$

Pax World Balanced seems less risky.

(b) Pax: $\bar{x} - 2s = 11.69 - 2(11.56) = -11.43$

$\bar{x} + 2s = 11.69 + 2(11.56) = 34.81$

At least 75% of the data fall in the interval −11.43% to 34.81%.

Vanguard: $\bar{x} - 2s = 5.61 - 2(12.50) = -19.39$

$\bar{x} + 2s = 5.61 + 2(12.50) = 30.61$

At least 75% of the data fall in the interval −19.39% to 30.61%.

The performance range for Pax seems better than for Vanguard (based on these historical data).

6. (a)-(c) Students verify results.

(d) The 3-year moving average has a much lower standard deviation.

7. (a) $\bar{x} - 2s = 11.01 - 2(2.17) = 6.67$

$\bar{x} + 2s = 11.01 + 2(2.17) = 15.35$

We expect at least 75% of the cycles to fall in the interval 6.67 years to 15.35 years.

(b) $\bar{x} - 4s = 11.01 - 4(2.17) = 2.33$

$\bar{x} + 4s = 11.01 + 4(2.17) = 19.69$

We expect at least 93.8% of the cycles to fall in the interval 2.33 years to 19.69 years.

8. (a) Results round to answers given.

(b) $\bar{x} - 2s = 730 - 2(172) = 386$

$\bar{x} + 2s = 730 + 2(172) = 1074$

We expect at least 75% of the years to have between 386 and 1074 tornados.

(c) $\bar{x} - 3s = 730 - 3(172) = 214$

$\bar{x} + 3s = 730 + 3(172) = 1246$

We expect at least 88.9% of the years to have between 214 and 1246 tornados.

9. (a) Range = 956 − 219 = 737

$$\bar{x} = \frac{\sum x}{n} = \frac{3968}{7} \approx 566.9$$

(b) $$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{427{,}213}{6} \approx 71{,}202$$

$$s = \sqrt{s^2} = \sqrt{71{,}202} \approx 266.8$$

(c) $$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{266.8}{566.9} \cdot 100 = 47.1\%$$

s is 47.1% of $\bar{x}$.

(d) $\bar{x} - 2s = 566.9 - 2(266.8) \approx 33$

$\bar{x} + 2s = 566.9 + 2(266.8) \approx 1100$

We expect at least 75% of the artifact counts for all such excavation sites to fall in the interval 33 to 1100.

10. $\bar{x} = 4.0$, $s = 1.2$ for indulgences.

$\bar{x} - 2s = 4.0 - 2(1.2) = 1.6$

$\bar{x} + 2s = 4.0 + 2(1.2) = 6.4$

We would expect at least 75% of the *P*s to fall in the interval 1.6% to 6.4%.

11. $\bar{x} = 9.3$, $s = 2.6$ for domestic routine.

$\bar{x} - 2s = 9.3 - 2(2.6) = 4.1$

$\bar{x} + 2s = 9.3 + 2(2.6) = 14.5$

We would expect at least 75% of the percentages of domestic routine artifacts at different sites to fall in the interval 4.1% to 14.5%.

12. $\bar{x} = 0.4$, $s = 0.1$ for arms.

$\bar{x} - 2s = 0.4 - 2(0.1) = 0.2$

$\bar{x} + 2s = 0.4 + 2(0.1) = 0.6$

We would expect at least 75% of the percentages of arms artifacts at different sites to fall in the interval 0.2% to 0.6%.

13. Construction artifacts have the largest mean and the smallest *CV*. Therefore, construction artifacts have the highest average and lowest relative standard deviation.

14. The mean for arms artifacts is twice that of stable artifacts.

15. (a) Students verify results.

(b) Wal-Mart $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{1.06}{52.03} \cdot 100 \approx 2\%$

Disney $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{0.98}{32.23} \cdot 100 \approx 3\%$

Yes, since the CV's are approximately equal, they appear to be equally attractive.

(c) Wal-Mart:

$\bar{x} - 3s = 52.03 - 3(1.06) = 48.85$

$\bar{x} + 3s = 52.03 + 3(1.06) = 55.21$

Disney:

$\bar{x} - 3s = 32.23 - 3(0.98) = 29.29$

$\bar{x} + 3s = 32.23 + 3(0.98) = 35.17$

The support is \$48.85 and resistance is \$55.21 for Wal-Mart.
The support is \$29.29 and the resistance is \$35.17 for Disney.

16. $CV = \frac{s}{\bar{x}} \cdot 100$

$\frac{\bar{x} \cdot CV}{100} = s$

$s = \frac{\bar{x} \cdot CV}{100}$

$s = \frac{2.2(1.5)}{100}$

$s = 0.033$

17. Answers vary.

Section 3.3

1.

Class	f	x	xf	$x - \bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
21–30	260	25.5	6630	−10.3	106.09	27,583.4
31–40	348	35.5	12,354	−0.3	0.09	31.3
41 and over	287	45.5	13,058.5	9.7	94.09	27,003.8
	$n = \Sigma f = 895$		$\Sigma xf = 32{,}042.5$			$\Sigma(x-\bar{x})^2 f = 54{,}619$

$\bar{x} = \frac{\Sigma xf}{n} = \frac{32{,}042.5}{895} \approx 35.80$

$s^2 = \frac{\Sigma(x-\bar{x})^2 \cdot f}{n-1} = \frac{54.619}{894} \approx 61.1$

$s = \sqrt{61.1} \approx 7.82$

2.

Class	f	x	xf	$x - \bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
1–10	34	5.5	187	−10.6	112.36	3820.24
11–20	18	15.5	279	−0.6	0.36	6.48
21–30	17	25.5	433.5	9.4	88.36	1502.12
31 and over	11	35.5	390.5	19.4	376.36	4139.96
	$n = \Sigma f = 80$		$\Sigma xf = 1290$			$\Sigma(x-\bar{x})^2 f = 9468.8$

$$\bar{x} = \frac{\sum xf}{n} = \frac{1290}{80} \approx 16.1$$

$$s^2 = \frac{\sum (x-\bar{x})^2 f}{n-1} = \frac{9468.8}{79} \approx 119.9$$

$$s = \sqrt{119.9} \approx 10.95$$

3.

Class	f	x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
8.6–12.5	15	10.55	158.25	−5.05	25.502	382.537
12.6–16.5	20	14.55	291.00	−1.05	1.102	22.050
16.6–20.5	5	18.55	92.75	2.95	8.703	43.513
20.6–24.5	7	22.55	157.85	6.95	48.303	338.118
24.6–28.5	3	26.55	79.65	10.95	119.903	359.708
	$n = \sum f = 50$		$\sum xf = 779.5$			$\sum (x-\bar{x})^2 f = 1145.9$

$$\bar{x} = \frac{\sum xf}{n} = \frac{779.5}{50} \approx 15.6$$

$$s^2 = \frac{\sum (x-\bar{x})^2 f}{n-1} = \frac{1145.9}{49} \approx 23.4$$

$$s = \sqrt{23.4} \approx 4.8$$

4.

Class	f	x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
12–14	1	13	13	−7.35	54.0225	54.023
15–17	3	16	48	−4.35	18.9225	56.768
18–20	8	19	152	−1.35	1.8225	14.580
21–23	2	22	44	1.65	2.7225	5.445
24–26	6	25	150	4.65	21.6225	129.735
	$n = \sum f = 20$		$\sum xf = 407$			$\sum (x-\bar{x})^2 f = 260.55$

$$\bar{x} = \frac{\sum xf}{n} = \frac{407}{20} = 20.35$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{260.55}{19}} \approx 3.703$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{3.703}{20.35} \cdot 100 = 18.2\%$$

5.

Class	f	x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
18–24	78	21.0	1638.0	−18.12	328.33	25610.1
25–34	75	29.5	2212.5	−9.62	92.54	6940.8
35–44	48	39.5	1896.0	0.38	0.14	6.9
45–54	33	49.5	1633.5	10.38	107.74	3555.6
55–64	33	59.5	1963.5	20.38	415.34	13706.4
65–80	33	72.5	2392.5	33.38	1114.22	36769.4
	$n = \sum f = 300$		$\sum xf = 11{,}736$			$\sum (x-\bar{x})^2 f = 86{,}589$

$$\bar{x} = \frac{\sum xf}{n} = \frac{11{,}736}{300} = 39.12$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{86{,}589}{299}} \approx 17.02$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{17.02}{39.12} \cdot 100 \approx 43.5\%$$

6. Men:

Class	f	x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
0–24	560	12	6720	−15	225	126000
25–49	320	37	11840	10	100	32000
50–74	80	62	4960	35	1225	98000
75–99	40	87	3480	60	3600	144000
	$n = \sum f = 1000$		$\sum xf = 27{,}000$			$\sum (x-\bar{x})^2 f = 400{,}000$

$$\bar{x} = \frac{\sum xf}{n} = \frac{27{,}000}{1000} = 27$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{400{,}000}{999}} \approx 20.01$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{20.01}{27} \cdot 100 \approx 74.1\%$$

Women:

Class	f	x	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
0–24	800	12	9600	−6	36	28800
25–49	170	37	6290	19	361	61370
50–74	20	62	1240	44	1936	38720
75–99	10	87	870	69	4761	47610
	$n = \sum f = 1000$		$\sum xf = 18{,}000$			$\sum (x-\bar{x})^2 f = 176{,}500$

$$\bar{x} = \frac{\sum xf}{n} = \frac{18{,}000}{1000} = 18$$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{176{,}500}{999}} \approx 13.29$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{13.29}{18} \cdot 100 \approx 73.8\%$$

7.

x	f	xf	x^2f
3.5	2	7	24.5
4.5	2	9	40.5
5.5	4	22	121.0
6.5	22	143	929.5
7.5	64	480	3600.0
8.5	90	765	6502.5
9.5	14	133	1263.5
10.5	2	21	220.5
	$\Sigma f = 200$	$\Sigma xf = 1580$	$\Sigma x^2 f = 12,702$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{1580}{200} = 7.9$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 12,702 - \frac{(1580)^2}{200} = 220$$

$$s = \sqrt{\frac{SS_x}{n-1}} = \sqrt{\frac{220}{199}} \approx 1.05$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{1.05}{7.9} \cdot 100 \approx 13.29\%$$

8.

x	f	xf	x^2f
1	15	15	15
2	15	30	60
3	7	21	63
4	9	36	144
5	13	65	325
6	13	78	468
7	9	63	441
8	8	64	512
9	9	81	729
	$\Sigma f = 98$	$\Sigma xf = 453$	$\Sigma x^2 f = 2757$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{453}{98} \approx 4.6$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 2757 - \frac{(453)^2}{98} \approx 663.03$$

$$s^2 = \frac{SS_x}{n-1} = \frac{663.03}{97} \approx 6.8$$

$$s = \sqrt{6.8} \approx 2.6$$

9. (a)

x	f	xf	x^2f
1	3	3	3
2	7	14	28
3	6	18	54
4	5	20	80
5	4	20	100
6	2	12	72
7	0	0	0
8	1	8	64
9	2	18	162
10	1	10	100
	$\Sigma f = 31$	$\Sigma xf = 123$	$\Sigma x^2 f = 663$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{123}{31} \approx 3.97$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 663 - \frac{(123)^2}{31} \approx 175$$

$$s = \sqrt{\frac{SS_x}{n-1}} = \sqrt{\frac{175}{30}} \approx 2.415$$

(b) The results are the same.

10.

x	f	xf	x^2f
35.4	79	2796.6	99000
23.0	565	12995.0	298885
29.4	136	3998.4	117553
38.3	103	3944.9	151090
61.5	400	24600.0	1512900
	$\Sigma f = 1283$	$\Sigma xf = 48{,}335$	$\Sigma x^2 f = 2{,}179{,}427$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{48{,}335}{1283} \approx 37.7$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 2{,}179{,}427 - \frac{(48{,}335)^2}{1283} \approx 358{,}482.2$$

$$s^2 = \frac{SS_x}{n-1} = \frac{358{,}482.2}{1282} \approx 279.6$$

$$s = \sqrt{279.6} \approx 16.7$$

(Calculations may vary slightly due to rounding.)

11.

x	f	xf	x^2f
2.8	145	406.0	1136.8
6.3	270	1701.0	10716.3
1.8	224	403.2	725.8
4.8	271	1300.8	6243.8
3.0	67	201.0	603.0
	$\sum f = 977$	$\sum xf = 4012$	$\sum x^2 f = 19{,}426$

$$\bar{x} = \frac{\sum xf}{n} = \frac{4012}{977} \approx 4.11$$

$$SS_x = \sum x^2 f - \frac{(\sum xf)^2}{n} = 19{,}426 - \frac{(4012)^2}{977} \approx 2951$$

$$s^2 = \frac{SS_x}{n-1} = \frac{2951}{976} \approx 3.02$$

$$s = \sqrt{3.02} \approx 1.74$$

12.

x	f	xf	x^2f
1.3	75	97.5	126.8
8.7	190	1653.0	14381.1
11.3	80	904.0	10215.2
5.9	51	300.9	1775.3
3.3	181	597.3	1971.1
	$\sum f = 577$	$\sum xf = 3552.7$	$\sum x^2 f = 28.469$

$$\bar{x} = \frac{\sum xf}{n} = \frac{3552.7}{577} \approx 6.16$$

$$SS_x = \sum x^2 f - \frac{(\sum xf)^2}{n} = 28.469 - \frac{(3552.7)^2}{577} \approx 6594$$

$$s^2 = \frac{SS_x}{n-1} = \frac{6594}{576} \approx 11.45$$

$$s = \sqrt{11.45} \approx 3.38$$

13. Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{92(0.25)+81(0.225)+93(0.225)+85(0.30)}{1}$$

$$= 87.65$$

14. Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{92(0.25)+81(0.25)+93(0.25)+85(0.25)}{1}$$

$$= 87.75$$

The weighted average here is slightly greater than that in Problem 13. Since the weights are the same, we could have computed the mean of the four scores directly.

15. Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{9(2)+7(3)+6(1)+10(4)}{2+3+1+4}$$

$$= \frac{85}{10}$$

$$= 8.5$$

16. (a) Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{9(5)+7(2)+6(1)+8(3)}{5+2+1+3}$$

$$= \frac{89}{11}$$

$$\approx 8.09$$

(b) Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{8(5)+9(2)+5(1)+9(3)}{5+2+1+3}$$

$$= \frac{90}{11}$$

$$\approx 8.18$$

This athlete has the higher average rating.

17. Use the midpoints of the classes for x.

Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{7(0.15)+11(0.39)+14(0.27)+16.5(0.19)}{0.15+0.39+0.27+0.19}$$

$$\approx 12.26$$

The weighted average is 12.26 lb.

18. Use the midpoints of the classes for x.

(a) $$\text{Weighted average} = \frac{\sum xw}{\sum w} = \frac{3.5(0.26)+8(0.40)+13(0.34)}{1} = 8.53$$

(b) $$\text{Weighted average} = \frac{\sum xw}{\sum w} = \frac{3.5(0.26)+8(0.40)+15.5(0.34)}{1} = 9.38$$

(c) Weighted average is sensitive to extreme values.

(d) Answers vary.

Section 3.4

1. 82% or more of the scores were at or below her score. 100% −82% = 18% or less of the scores were above her score. Note: This answer is correct, but it relies on a more precise definition than that given in the text on page 124. An adequate answer, matching the definition in the text would be: 82% of the scores were at or below her score, and (100 − 82)% = 18% of the scores were at or above her score.

2. The upper quartile is the 75th percentile. Therefore, the minimal percentile rank must be the 75th.

3. No, the score 82 might have a percentile rank less than 70.

4. Timothy performed better because a percentile rank of 72 is greater than a percentile rank of 70.

5. Order the data from smallest to largest.

Lowest value = 0.52
Highest value = 1.92

There are 16 data values.

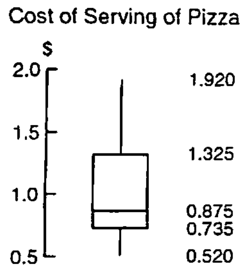
$$\text{Median} = \frac{0.85+0.90}{2} = 0.875$$

There are 8 values less than 0.875 and 8 values greater than 0.875.

$$Q_1 = \frac{0.72+0.75}{2} = 0.735$$

$$Q_3 = \frac{1.15+1.50}{2} = 1.325$$

$$IQR = Q_3 - Q_1 = 1.325 - 0.735 = 0.59$$



6. Order the data from smallest to largest.

Lowest value = 275
Highest value = 393

There are 16 data values.

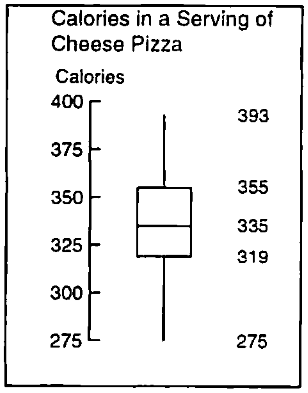
$$\text{Median} = \frac{333+337}{2} = 335$$

There are 8 values less than 335 and 8 values greater than 335.

$$Q_1 = \frac{316+322}{2} = 319$$

$$Q_3 = \frac{353+357}{2} = 355$$

$$IQR = Q_3 - Q_1 = 355 - 319 = 36$$



7. Order the data from smallest to largest.

Lowest value = 2
Highest value = 42

There are 20 data values.

$$\text{Median} = \frac{23+23}{2} = 23$$

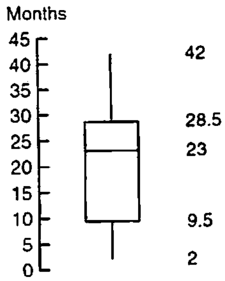
There are 10 values less than the Q_2 position and 10 values greater than the Q_2 position.

$$Q_1 = \frac{8+11}{2} = 9.5$$

$$Q_3 = \frac{28+29}{2} = 28.5$$

$$IQR = Q_3 - Q_1 = 28.5 - 9.5 = 19$$

Nurses' Length of Employment (months)



8. (a) Order the data from smallest to largest.

Lowest value = 3
Highest value = 72

There are 20 data values.

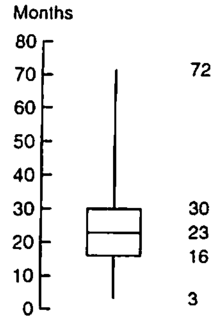
$$\text{Median} = \frac{22+24}{2} = 23$$

There are 10 values less than the median and 10 values greater than the median.

$$Q_1 = \frac{15+17}{2} = 16$$

$$Q_3 = \frac{29+31}{2} = 30$$

$$IQR = Q_3 - Q_1 = 30 - 16 = 14$$



(b) The medians are the same (23) and the IQR's are similar. However, the distances from Q_1 to the minimum value and from Q_3 to the maximum value are greater here than in Problem 7.

9. Order each set of data from smallest to largest.

Suburban

Lowest value = 808
Highest value = 1292

$$\text{Median} = \frac{992+1170}{2} = 1081$$

There are five values above and five values below the median.

$$Q_1 = 972$$
$$Q_3 = 1216$$
$$IQR = 1216 - 972 = 244$$

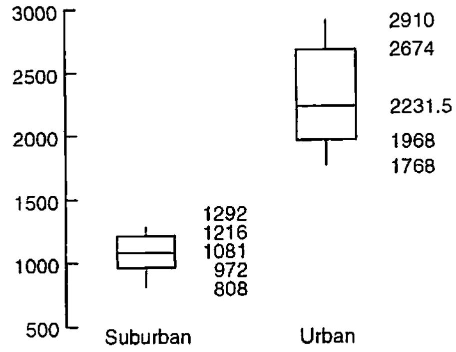
Urban

Lowest value = 1768
Highest value = 2910

$$\text{Median} = \frac{2107+2356}{2} = 2231.5$$

There are five values above and five values below the median.

$$Q_1 = 1968$$
$$Q_3 = 2674$$
$$IQR = 2674 - 1968 = 706$$



The entire box-and-whisker plot for urban is above that for suburban. Even the highest value for suburban is less than the lowest value for urban. The suburban data is less variable than that of urban data.

10. (a) Order the data from smallest to largest.

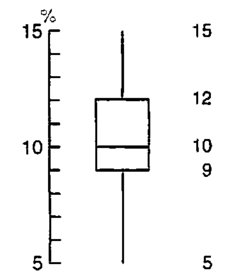
Lowest value = 5
Highest value = 15

There are 50 data values.

$$\text{Median} = \frac{10+10}{2} = 10$$

There are 25 values above and 25 values below the Q_2 position.

$Q_1 = 9$
$Q_3 = 12$
$IQR = 12 - 9 = 3$



(b) 7% is in the 1st quartile, since it is below Q_1.

11. **(a)** Order the data from smallest to largest.

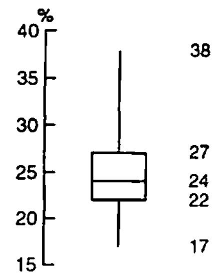
Lowest value = 17
Highest value = 38

There are 50 data values.

$$\text{Median} = \frac{24+24}{2} = 24$$

There are 25 values above and 25 values below the Q_2 position.

$Q_1 = 22$
$Q_3 = 27$
$IQR = 27 - 22 = 5$



(b) 26% is in the 3rd quartile, since it is between the median and Q_3.

12. **(a)** Yes; the data above the median has more spread.

(b) Coca-Cola; the difference between the highest value and lowest value is greatest.

(c) Coca-Cola; the distribution has the most spread.

(d) McDonalds: McDonalds: the line representing the median is at a negative value which means more weekly declines than weekly increases.

(e) Disney: the sizes of the percentage increases are smaller than the others.

(f) Coca-Cola: Coca-Cola; the percentage decreases of this stock are larger than the others.

13. **(a)** California has the lowest premium since its left whisker is farthest to the left. Pennsylvania has the highest premium since its right whisker is farthest to the right.

(b) Pennsylvania has the highest median premium since its line in the middle of the box is farthest to the right.

(c) California has the smallest range of premiums since the distance between the ends of the whiskers is the smallest. Texas has the smallest interquartile range since the distance between the ends of the boxes is the smallest.

(d) Based on the answers to (a)-(c) above, we can determine that part (a) of Figure 3-13 is for Texas. part (b) of Figure 3-13 is for Pennsylvania, and part (c) of Figure 3-13 is for California.

14. **(a)** Order the data from smallest to largest.

Lowest value = 4
Highest value = 80

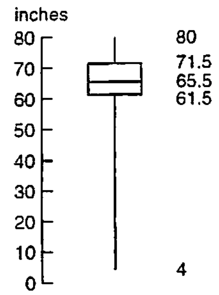
There are 24 data values.

$$\text{Median} = \frac{65+66}{2} = 65.5$$

There are 12 values above and 12 values below the median.

$$Q_1 = \frac{61+62}{2} = 61.5$$

$$Q_3 = \frac{71+72}{2} = 71.5$$



(b) $IQR = Q_3 - Q_1 = 71.5 - 61.5 = 10$

(c) $1.5(10) = 15$
Lower limit: $Q_1 - 1.5(IQR) = 61.5 - 15 = 46.5$
Upper limit: $Q_3 + 1.5(IQR) = 71.5 + 15 = 86.5$

(d) Yes, the value 4 is below the lower limit and so is an outlier; it is probably an error. Our guess is that one of the students is 4 feet tall and listed height in feet instead of inches. There are no values above the upper limit.

15. **(a)** Assistant had the smallest median percentage salary increase since the bar in the middle of the box is the lowest. Associate had the single highest salary increase since it has the highest asterisk.

(b) Instructor had the largest spread between the first and third quartiles since the distance between the ends of the box is greatest.

(c) Assistant had the smallest spread for the lower 50% of the percentage salary increases since the distance between the bar in the box and the maximum value is the smallest.

(d) Professor had the most symmetric percentage salary increases because there are no outliers and the bar representing the median is close to the center of the box.
Yes, if the outliers for the associate professors were omitted, that distribution would appear to be symmetric.

(e) Associate professor:

$$IQR = 5.075 - 2.350 = 2.725$$
$$Q_3 + 1.5(IQR) = 5.075 + 1.5(2.725) \approx 9.16$$

Yes, since 17.7 is greater than 9.16, there is at least one outlier.

Instructor:

$$IQR = 5.800 - 2.850 = 2.950$$
$$Q_3 + 1.5(IQR) = 5.800 + 1.5(2.950) \approx 10.23$$

Yes, since 13.4 is greater than 10.23, there is at least one outlier.

Chapter 3 Review

1. (a) $\bar{x} = \frac{\sum x}{n} = \frac{876}{8} = 109.5$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{7044}{7}} = \sqrt{1006.3} \approx 31.7$$
$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{31.7}{109.5} \cdot 100 \approx 28.9\%$$

range = maximum value − minimum value
= 142 − 73 = 69

(b) $\bar{x} = \frac{\sum x}{n} = \frac{881}{8} = 110.125$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{358.87}{7}} \approx 7.2$$
$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{7.2}{110.125} \cdot 100 \approx 6.5\%$$

range = maximum value − minimum value
= 120 − 100 = 20

(c) The means are about the same. The first distribution has greater spread. The standard deviation, *CV*, and range for the first set of measurements are greater than those for the second set of measurements.

2. (a) Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{1.9 + 2.8 + \cdots + 7.2}{8}$

$$= \frac{36.2}{8}$$
$$= 4.525$$

Order the data from smallest to largest.

1.9 1.9 2.8 3.9 4.2 5.7 7.2 8.6

Median $= \frac{3.9 + 4.2}{2} = 4.05$

The mode is 1.9 because it is the value that occurs most frequently.

(b) $s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{42.395}{7}} \approx 2.46$

$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{2.46}{4.525} \cdot 100 \approx 54.4\%$

Range $= 8.6 - 1.9 = 6.7$

3. (a) Order the data from smallest to largest.

Lowest value = 31
Highest value = 68

There are 60 data values.

$$\text{Median} = \frac{45+45}{2} = 45$$

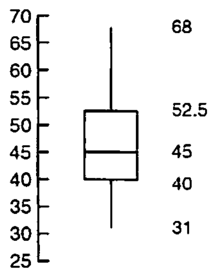
There are 30 values above and 30 values below the Q_2 position.

$$Q_1 = \frac{40+40}{2} = 40$$

$$Q_3 = \frac{52+53}{2} = 52.5$$

$$IQR = 52.5 - 40 = 12.5$$

Percentage of Democratic Vote
by Counties in Georgia



(b) Class width = 8

Class	x Midpoint	f	xf	x^2f
31–38	34.5	11	379.5	13,092.8
39–46	42.5	24	1020	43,350.0
47–54	50.5	15	757.5	38,253.8
55–62	58.5	7	409.5	23,955.8
63–70	66.5	3	199.5	13,266.8
		$n = \Sigma f = 60$	$\Sigma xf = 2766$	$\Sigma x^2 f = 131{,}919$

$$\bar{x} = \frac{\sum xf}{n} = \frac{2766}{60} = 46.1$$

$$SS_x = \sum x^2 f - \frac{(\sum xf)^2}{n} = 131{,}919 - \frac{(2766)^2}{60} = 4406.4$$

$$s = \sqrt{\frac{SS_x}{n-1}} = \sqrt{\frac{4406.4}{59}} \approx 8.64$$

$\bar{x} - 2s = 46.1 - 2(8.64) = 28.82$
$\bar{x} + 2s = 46.1 + 2(8.64) = 63.38$

We expect at least 75% of the data to fall in the interval 28.82 to 63.38.

(c) $\bar{x} = 46.15$, $s \approx 8.63$

4. (a) Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{92(0.05)+73(0.08)+81(0.08)+85(0.15)+87(0.15)+83(0.15)+90(0.34)}{0.05+0.08+0.08+0.15+0.15+0.15+0.34}$$

$$= \frac{85.77}{1}$$

$$= 85.77$$

(b) Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{20(0.05)+73(0.08)+81(0.08)+85(0.15)+87(0.15)+83(0.15)+90(0.34)}{1}$$

$$= 82.17$$

5. Mean weight $= \dfrac{2500}{16} = 156.25$

The mean weight is 156.25 lb.

6.

x	f	xf	$x^2 f$
17	6	102	1734
32	20	640	20,480
47	52	2444	114,868
62	16	992	61,504
	$\sum f = 94$	$\sum xf = 4178$	$\sum x^2 f = 198{,}586$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{4178}{94} \approx 44.4$$

The mean is 44.4 in.

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 198{,}586 - \frac{(4178)^2}{94} \approx 12{,}887$$

$$s = \sqrt{\frac{SS_x}{n-1}} = \sqrt{\frac{12{,}887}{93}} \approx 11.8$$

The standard deviation is 11.8 in.

7. **(a)** Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{10.1 + 6.2 + \cdots + 5.7}{6} = \frac{47}{6} \approx 7.83$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{26.913}{5}} \approx 2.32$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{2.32}{7.83} \cdot 100 \approx 29.6\%$$

Range = largest value − smallest value
= 10.1 − 5.3 = 4.8

(b) Mean $= \bar{x} = \frac{\Sigma x}{n} = \frac{10.2 + 9.7 + \cdots + 10.1}{6} = \frac{59.7}{6} = 9.95$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{0.415}{5}} \approx 0.29$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{0.29}{9.95} \cdot 100 \approx 2.9\%$$

Range = largest value − smallest value
= 10.3 − 9.6 = 0.7

(c) Second line has more consistent performance as reflected by the smaller standard deviation. *CV*, and range.

8. Order the data from smallest to largest.

Lowest value = 45
Highest value = 109

There are 70 data values.

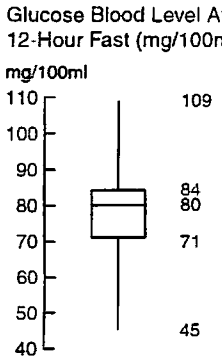
$$\text{Median} = \frac{80 + 80}{2} = 80$$

There are 35 values above and 35 values below the Q_2 position.

$Q_1 = 71$

$Q_3 = 84$

$IQR = 84 - 71 = 13$



9. (a)

x	f	xf	x^2f
50	3	150	7500
61	7	427	26047
72	22	1584	114048
83	26	2158	179114
94	9	846	79524
105	3	315	33075
	$n = \Sigma f = 70$	$\Sigma xf = 5480$	$\Sigma x^2 f = 439{,}308$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{5480}{70} \approx 78.3$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 439{,}308 - \frac{(5480)^2}{70} \approx 10302.3$$

$$s^2 = \frac{SS_x}{n-1} = \frac{10302.3}{69} \approx 149$$

$$s = \sqrt{s^2} = \sqrt{149} \approx 12.2$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = \frac{12.2}{78.3} \cdot 100 \approx 15.6\%$$

(b) $\bar{x} - 2s = 78.3 - 2(12.2) = 53.9$

$\bar{x} + 2s = 78.3 + 2(12.2) = 102.7$

We expect at least 75% of the glucose blood level measurements to fall in the interval 53.9 to 102.7.

10.

x	f	xf	x^2f
4.6	304	1398.4	6432.64
12.0	52	624.0	7488.00
2.6	39	101.4	263.64
4.1	319	1307.9	5362.39
3.9	77	300.3	1171.17
	$\Sigma f = 791$	$\Sigma xf = 3732$	$\Sigma x^2 f = 20{,}718$

$$\bar{x} = \frac{\sum xf}{n} = \frac{3732}{791} \approx 4.7$$

$$SS_x = \sum x^2 f - \frac{(\sum xf)^2}{n} = 20{,}718 - \frac{(3732)^2}{791} \approx 3110$$

$$s^2 = \frac{SS_x}{n-1} = \frac{3110}{790} \approx 3.94$$

$$s = \sqrt{3.94} = 1.98$$

11. Weighted average $= \dfrac{\sum xw}{\sum w}$

$$= \frac{5(2)+8(3)+7(3)+9(5)+7(3)}{2+3+3+5+3}$$

$$= \frac{121}{16}$$

$$\approx 7.56$$

12. (a) Order the data from smallest to largest.

Lowest value = 6
Highest value = 16

There are 50 data values.

$$\text{Median} = \frac{11+11}{2} = 11$$

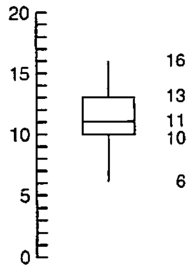
There are 25 values above and 25 values below the Q_2 position.

$Q_1 = 10$

$Q_3 = 13$

$IQR = Q_3 - Q_1 = 13 - 10 = 3$

Soil Water Content



(b)

Class	x Midpoint	f	xf	x^2f
6–8	7	4	28	196
9–11	10	24	240	2400
12–14	13	15	195	2535
15–17	16	7	112	1792
		$n = \Sigma f = 50$	$\Sigma xf = 575$	$\Sigma x^2 f = 6923$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{575}{50} = 11.5$$

$$SS_x = \Sigma x^2 f - \frac{(\Sigma xf)^2}{n} = 6923 - \frac{(575)^2}{50} = 310.5$$

$$s = \sqrt{\frac{SS_x}{n-1}} = \sqrt{\frac{310.5}{49}} \approx 2.52$$

$$\bar{x} - 2s = 11.5 - 2(2.52) = 6.46$$

$$\bar{x} + 2s = 11.5 + 2(2.52) = 16.54$$

We expect at least 75% of the data to fall in the interval 6.46 to 16.54.

(c) $\bar{x} \approx 11.48$; $s \approx 2.44$

13. **(a)** It is possible for the range and the standard deviation to be the same. For instance, for data values that are all the same, such as 1, 1, 1, 1, 1, the range and standard deviation are both 0.

(b) It is possible for the mean, median, and mode to be all the same. For instance, the data set 1, 2, 3, 3, 3, 4, 5 has mean, median, and mode all equal to 3. The averages can all be different, as in the data set 1, 2, 3, 3. In this case, the mean is 2.25, the median is 2.5, and the mode is 3.