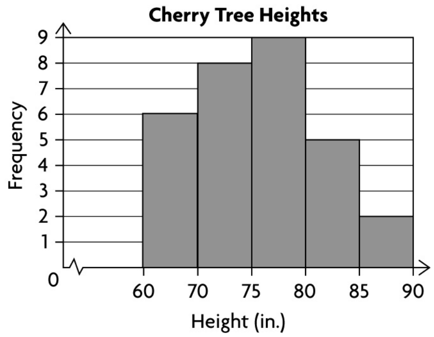


**11. a)** e.g., Grouping large sets of raw data into intervals makes it easier to interpret the data and to draw conclusions about how the data is distributed. It also makes a large set of data more manageable.
**b)** e.g., Frequency polygons and histograms are similar in that they display data that has been grouped into intervals. Histograms are like bar graphs and they display grouped data, whereas frequency polygons are similar to line graphs. For a frequency polygon, the midpoint of each interval is plotted against frequency and the points are joined. Frequency polygons are better to use when comparing multiple sets of data because they can be graphed on top of each other, unlike histograms.
**12. a)** e.g., To determine the mean, I determined the midpoint value for each population interval. Then for each interval, I multiplied the midpoint by the frequency. I added the products for each interval then divided by the sum of the frequencies.

| Midpoint Population | Frequency | Total Population for Interval |
|---|---|---|
| 2 500 | 122 | 305 000 |
| 7 500 | 16 | 120 000 |
| 12 500 | 4 | 50 000 |
| 17 500 | 2 | 35 000 |
| 22 500 | 2 | 45 000 |
| 27 500 | 2 | 55 000 |
| 32 500 | 0 | 0 |
| 37 500 | 0 | 0 |
| 42 500 | 1 | 42 500 |
| 47 500 | 0 | 0 |
| 52 500 | 1 | 52 500 |
| 57 500 | 0 | 0 |
| 62 500 | 0 | 0 |
| 67 500 | 1 | 67 500 |
| **Totals** | 151 | 772 500 |

Mean population = $\frac{\text{total population}}{\text{total number of areas}}$

Mean population = $\frac{772\ 500}{151}$

Mean population = 5115.894…
The mean population per area was 5116 people.
**b)** To determine the median value, I found the middle value of the populations. Since there were 151 areas, the median area is the 76th value. The 76th value is in the first interval, which has an upper limit of 5000 people. The median can be determined by multiplying the upper limit for the interval by the proportion of the frequency that 76 represents.

Median = $\left(\frac{76}{122}\right)(5000)$

Median = 3114.754…
The median population is 3115 people.

**c)** Percent error = $\frac{\text{calulated value} - \text{actual value}}{\text{actual value}}$

Percent error for median = $\frac{3115 - 1700}{1700} \times 100\%$

Percent error for median = 83.235…%
The percent error for the median is 83%.

Percent error for mean = $\frac{5116 - 4275}{4275} \times 100\%$

Percent error for mean = 19.672…%. The percent error for the mean is approximately 19.7%.
The determined mean value is closer to the actual mean than the median but both are higher than their actual values. The discrepancies lie in the use of intervals to group the data. Individual numbers could not be viewed and since the determined values are higher than the actual values, it is assumed that the actual numbers were nearer to the lower limit of each interval.
For the first interval, I assumed that 61 cities would have a population less than 2500, and 61 cities would have a population between 2500 and 5000. Since 75 cities have a population less than 1700, the estimates for both the mean and median are higher than they should be.

## History Connection, page 225

**A.** The floodway was opened 16 times between 1969 and 1999.
**B.** e.g., From 1995 to 1999, there were five consecutive years with flood water levels. Previous to this period, high-level years were usually followed by many low-level years. You would need to look at data for more recent years to determine whether the trend has continued.

## Lesson 5.3: Standard Deviation, page 233

**1. a)** Use a table or spreadsheet.

| Test | Class A (%), $x$ | $x - \bar{x}$ | $(x - \bar{x})^2$ |
|---|---|---|---|
| 1 | 94 | 16 | 256 |
| 2 | 56 | –22 | 484 |
| 3 | 89 | 11 | 121 |
| 4 | 67 | –11 | 121 |
| 5 | 84 | 6 | 36 |
| **Total** | 390 | | 1018 |
| **Mean, $\bar{x}$** | $\frac{390}{5} = 78$ | | |
| **Standard deviation, $\sigma$** | | | $\sqrt{\frac{1018}{5}} = 14.268...$ |

The standard deviation for Class A is 14.27%.

| Test | Class B (%), $x$ | $x - \bar{x}$ | $(x - \bar{x})^2$ |
|---|---|---|---|
| 1 | 84 | 5.6 | 31.36 |
| 2 | 77 | –1.4 | 1.96 |
| 3 | 76 | –2.4 | 5.76 |
| 4 | 81 | 2.6 | 6.76 |
| 5 | 74 | –4.4 | 19.36 |
| **Total** | 392 | | 65.20 |
| **Mean, $\bar{x}$** | $\frac{392}{5} = 78.4$ | | |
| **Standard deviation, $\sigma$** | | | $\sqrt{\frac{65.20}{5}} = 3.611$ |

The standard deviation for Class B is 3.61%.

**b)** Using a calculator:

For Class A:

$\bar{x}$ = 78

$\Sigma x$ = 390

$\sigma$ = 14.268…

The standard deviation for Class A is 14.27%.

For Class B:

$\bar{x}$ = 78.4

$\Sigma x$ = 392

$\sigma$ = 3.611…

The standard deviation for Class B is 3.61%.

**c)** The standard deviation for Class B test scores was less than Class A test scores. Class B has the more consistent marks.

**2.** The mean bowling score is 130.42 points.
The standard deviation in bowling scores is 11.51 points.

**3. a)** Create a frequency table.

| Midpoint Bowling Score | Frequency |
|---|---|
| 103 | 1 |
| 108 | 3 |
| 113 | 4 |
| 118 | 7 |
| 123 | 9 |
| 128 | 14 |
| 133 | 11 |
| 138 | 8 |
| 143 | 6 |
| 148 | 5 |
| 153 | 3 |
| 158 | 1 |

The mean bowling score is 130.36 points.
The standard deviation in bowling scores is 12.05 points.

**b)** Ali's mean bowling score is slightly higher than the teams by 0.06 points. Ali's standard deviation is less that the teams by 0.54 points. Ali's data indicates that he is close to the average as a player.

**4. a)** This dispersion of the masses of the bead packages is greater for company B than for company A. For company B, there are more packages that deviate from the mean mass for a package.

**b)** Since company A's packages have a lower standard deviation, Marie should order from company A.

**5.** For group 1:

$\bar{x}$ = 71.9 bpm

$\sigma$ = 6.0 bpm

For group 2:

$\bar{x}$ = 71.0 bpm

$\sigma$ = 4.0 bpm

For group 3:

$\bar{x}$ = 70.4 bpm

$\sigma$ = 5.7 bpm

For group 4:

$\bar{x}$ = 76.9 bpm

$\sigma$ = 1.9 bpm

Group 3 has the lowest mean pulse rate.
Group 4 has the most consistent pulse rate.

**6. a)** For Nazra:

Total stones = 34 + 41 + 40 + 38 + 38 + 45

Total stones = 236

For Diko:

Total stones = 51 + 28 + 36 + 44 + 41 + 46

Total stones = 246

Diko lays more stones during the day.

**b)** For Nazra: $\sigma$ = 3.3 stones

For Diko: $\sigma$ = 7.4 stones

Nazra is more consistent in the number of stones he lays per day.

**7. a**) For 1995 to 2008:

$\bar{x}$ = 10.5 TDs

$\sigma$ = 5.6 TDs

**b)** e.g., The low number of TDs in the first year could be due to lack of practice or fewer games played. The low number of TDs in the last year could be due to injury and fewer games played.

**c)** For 1996 to 2007:

$\bar{x}$ = 11.7 TDs

$\sigma$ = 5.2 TDs

**d)** When the TDs for 1995 and 2008 are eliminated from the data, the mean number of TDs increases and the standard deviation decreases. The number of TDs for each year is now more consistent.

**8. a)** $\bar{x}$ = 1082 yards

$\sigma$ = 428.8 yards

**b)** The standard deviation for Allen Pitts is lower, so Allen Pitts was more consistent in terms of yards gained per year.

**9. a)** For Fitness Express:

| Midpoint Hours | Frequency |
|---|---|
| 8 | 9 |
| 11 | 18 |
| 13 | 23 |
| 15 | 32 |
| 17 | 39 |
| 19 | 42 |
| 21 | 31 |
| 23 | 22 |
| 25 | 16 |
| 27 | 11 |
| 29 | 7 |

$\bar{x}$ = 18.3 h
$\sigma$ = 4.9 h

For Fit for Life:

| Midpoint Hours | Frequency |
|---|---|
| 7.5 | 8 |
| 10.5 | 13 |
| 13.5 | 32 |
| 16.5 | 47 |
| 19.5 | 52 |
| 22.5 | 42 |
| 25.5 | 27 |
| 28.5 | 19 |

$\bar{x}$ = 19.1 h
$\sigma$ = 5.3 h

**b)** Fitness Express has a lower standard deviation, so it is the better club for encouraging its members to workout more consistently.

**10.** $\bar{x}$ = 21.2 min
$\sigma$ = 3.5 min
e.g., Jaime's mean arrival time is 21.2 min.
On average, Jaime is late by 1.2 min. The standard deviation is 3.5 min. This means there is a wide dispersion of arrival times from the mean. Based on the average, Jaime will more likely be late than early. I think Jaime will lose her job.

**11.** Create a frequency table.

| Midpoint Calls | Frequency |
|---|---|
| 28 | 2 |
| 33 | 13 |
| 38 | 42 |
| 43 | 53 |
| 48 | 42 |
| 53 | 36 |
| 58 | 8 |
| 63 | 4 |

$\bar{x}$ = 45.0 calls
$\sigma$ = 7.1 calls
The standard deviation is greater than the maximum standard deviation. So the manager should hire more employees.

**12. a)** Games played $\bar{x}$ = 57.5 games
Games played $\sigma$ = 11.4 games
Goals $\bar{x}$ = 12.5 goals
Goals $\sigma$ = 11.6 goals
Assists $\bar{x}$ = 16.1 assists
Assists $\sigma$ = 13.1 assists
Points $\bar{x}$ = 28.6 points
Points $\sigma$ = 24.5 points

**b)** e.g., If Jordin's statistics for the 2005-2006 season were eliminated, then his statistic for each column would rise and the standard deviation for each column should decrease.

**c)** Games played $\bar{x}$ = 60.1 games
Games played $\sigma$ = 8.7 games
Goals $\bar{x}$ = 13.4 goals
Goals $\sigma$ = 11.8 goals
Assists $\bar{x}$ = 17.2 assists
Assists $\sigma$ = 13.3 assists
Points $\bar{x}$ = 30.7 points
Points $\sigma$ = 24.9 points

**d)** For the number of games played, the mean number of games decreased and the standard deviation decreased. For the number of goals, the mean number of goals decreased and the standard deviation increased. For the number of assists, the mean number of assists decreased and the standard deviation increased. For the number of points, the mean number of points decreased and the standard deviation decreased. The means and standard deviations increased and decreased as I predicted.

**e)** e.g., When using the raw data, the statement "Goals + assists = points" is true.
For example,
6 goals + 10 assists = 16 points.
When using the mean, mean number of goals + mean number of assists = mean number of points.
For example,
12.5 goals + 16.1 goals = 28.6 points.
When using the standard deviation, however, the statement is not true.
For example,
11.6 goals + 13.1 assists = 24.7 points.

**13.** To have the same mean, Jordana for example, could have been more consistent in her scores while Jane had inconsistent scores but the sum of her test scores could have been the same as Jordana's. In this case, Jane's standard deviation would have been higher than Jordana's.
Example:
Jane's scores: 80%, 85%, 82%, 87%, 86%, 84%, 87%, 85%, 85%, 89%
mean = 85.0%
standard deviation = 2.6%

Jordana's scores: 78%, 92%, 99%, 64%, 72%, 82%, 77%, 95%, 98%, 93%
mean = 85.0%
standard deviation = 12.0%

**14. a)** Group A $\bar{x}$ = 8.56 s
Group A $\sigma$ = 7.99 s
Group B $\bar{x}$ = 5.55 s
Group B $\sigma$ = 4.73 s

**b)** Group B, the group given information, was able to recognize the pattern more quickly than Group A. However, Group A was more consistent in their image recognition times.

## Mid-Chapter Review, page 239

**1.** For Paris:

$$\text{Mean} = \frac{\text{sum of temperatures}}{\text{number of temperatures}}$$

$$\text{Mean} = \frac{187\ ^\circ\text{C}}{12}$$

Mean = 15.583… °C
The mean average daily temperature for Paris, France is 15.6 °C.