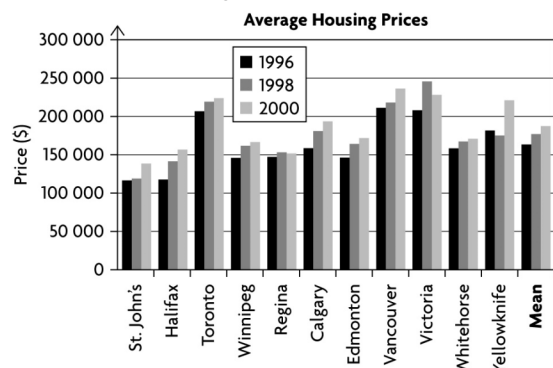


b) e.g., The mean, mode, and median values were the same for each unit test. The smallest range was for Unit 1 test results. Unit 1 test scores were more consistent between students and there were no test scores below 50%. So, Unit 1 test results were better.

c) e.g., The modes were not useful for comparison because they were the same value. The mode indicated the most frequently occurring score on each test but not which unit test the class performed better on.

3. a) Here is the diagram.



For 1996:
 Range = greatest price – least price
 Range = \$212 010 – \$116 443
 Range = \$95 567
 The range in housing prices is \$95 567.

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

$$\text{Mean} = \frac{\$1\,797\,843}{11}$$

$$\text{Mean} = \$163\,440.272\dots$$

The mean housing price is \$163 440.

In order from least to greatest price,
 116 443, 117 990, 144 858, 146 280, 147 889,
 157 677, 157 768, 181 790, 206 738, 208 400,
 212 010
 The median value is the sixth value, or \$157 677.

The mode is the most frequently occurring value. There is no mode.

For 1998:
 Range = greatest price – least price
 Range = \$246 135 – \$118 519
 Range = \$127 616
 The range in housing prices is \$127 616.

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

$$\text{Mean} = \frac{\$1\,946\,310}{11}$$

$$\text{Mean} = \$176\,937.272\dots$$

The mean housing price is \$176 937.

In order from least to greatest price,
 118 519, 141 353, 152 784, 161 337, 164 808,
 167 396, 175 646, 180 258, 218 025, 220 049,
 246 135
 The median value is the sixth value, or \$167 396.

The mode is the most frequently occurring value. There is no mode.

For 2000:
 Range = greatest price – least price
 Range = \$236 617 – \$137 665
 Range = \$98 952
 The range in housing prices is \$98 952.

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

$$\text{Mean} = \frac{\$2\,061\,770}{11}$$

$$\text{Mean} = \$187\,433.636\dots$$

The mean housing price is \$187 434.

In order from least to greatest price,
 137 665, 152 114, 156 988, 166 761, 170 986,
 172 503, 193 275, 221 632, 224 246, 228 983,
 236 617
 The median value is the sixth value, or \$172 503.

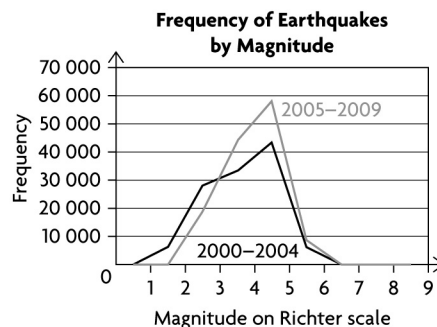
The mode is the most frequently occurring value. There is no mode.

e.g., The price data is not normally distributed about the mean or median values. The values are fairly scattered around the mean or median. The greatest range in prices occurred in 1998. This means some prices were much greater than the mean and some were much lower. Each set of data shows an increase in house price with time, except for Regina and Victoria. The greatest price increase occurred between 1998 and 2000 for Yellowknife. Whitehorse had the smallest price increase from year to year of all the cities.

b) e.g., Comparing the three sets of data is useful because you can determine the trend in housing price with time, and which cities have the greatest price changes. Depending on your needs, you may want to buy a property in a city where housing prices are steadily increasing for a good return on your investment or you may want to live in a city where housing prices are kept relatively low over time.

Lesson 5.2: Frequency Tables, Histograms, and Frequency Polygons, page 221

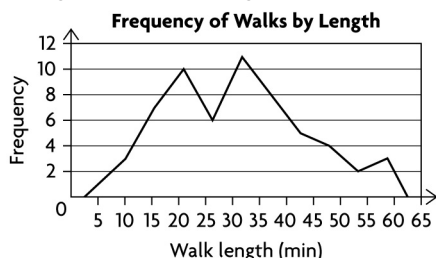
1. a) Here is the diagram.



b) e.g., Between 2005 to 2009, there were more earthquakes measuring 2.0 to 5.9 on the Richter scale relative to 2000 to 2004. The number of earthquakes on the higher end of the scale was approximately the same for both time periods.

2. a) Emmanuella should have placed the data for her first walk of 15 min in the 10 min to 15 min interval.

b) e.g., Here is the diagram.



The data shows a greater number of walks in the 20 min to 25 min interval and 30 min to 35 min interval.

3. a) e.g.,

Range = greatest height – least height

Range = 86 in. – 63 in.

Range = 23 in.

$$\text{Interval width} = \frac{\text{range}}{\text{number of intervals}}$$

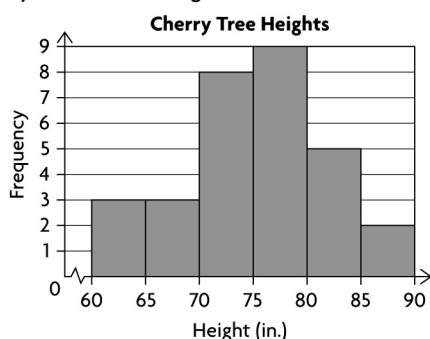
$$\text{Interval width} = \frac{23 \text{ in.}}{6}$$

$$\text{Interval width} = 3.833\dots \text{ in.}$$

I will use an interval size of 5 starting from 60.

Tree Height (in.)	Frequency
60–65	3
65–70	3
70–75	8
75–80	9
80–85	5
85–90	2

b) Here is the diagram.



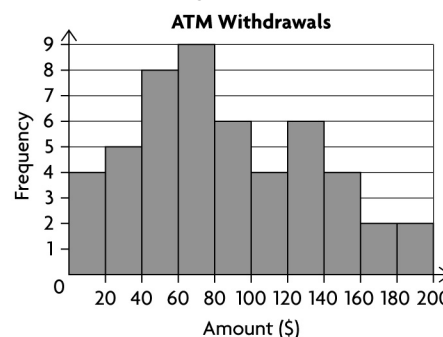
c) The range of heights that occurs most frequently is the 75 in. to 80 in. range. The range of heights that occurs the least is the 85 in. to 90 in. range.

4. a) e.g., Since most of the withdrawal amounts are multiples of \$20, I would like to use an interval width of \$20. This will give me 10 intervals to show how the data is distributed.

b) Create a frequency table.

Withdrawal (\$)	Frequency
0–20	4
20–40	5
40–60	8
60–80	9
80–100	6
100–120	4
120–140	6
140–160	4
160–180	2
180–200	2

c) Here is the diagram.

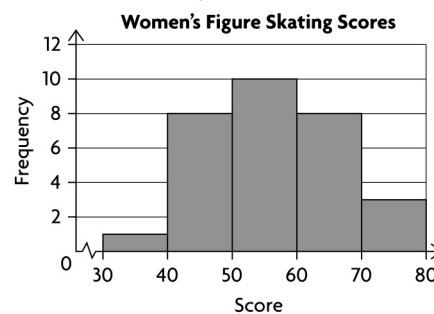


d) e.g., The mean withdrawal amount is \$96.80. The median withdrawal amount is \$80. The mode is between \$60 and \$80. Most of the withdrawal amounts are in the range of \$40 to \$80. Not many withdrawals amounts are greater than \$160.

5. a) e.g.,

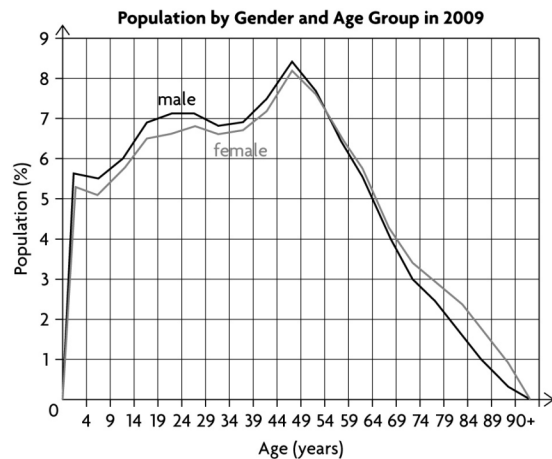
Final Scores	Frequency
30–40	1
40–50	8
50–60	10
60–70	8
70–80	3

b) Here is the diagram.



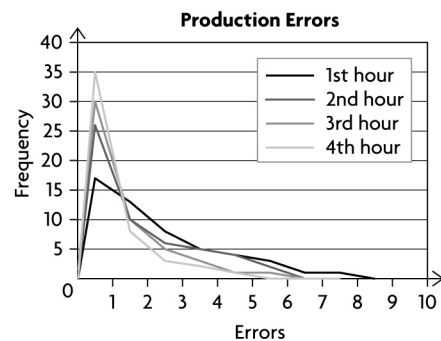
c) e.g., No. It shows that three women scored between 70 and 80, but it does not show the range of scores for a top-five placement.

6. a) Here is the diagram.



b) e.g., There was a slightly greater population of males relative to females aged 0 years to 54 years in 2009. Then the trend changed from the ages of 55 years and greater. The population of females was slightly higher compared to males for these years.

7. a) Here is the diagram.

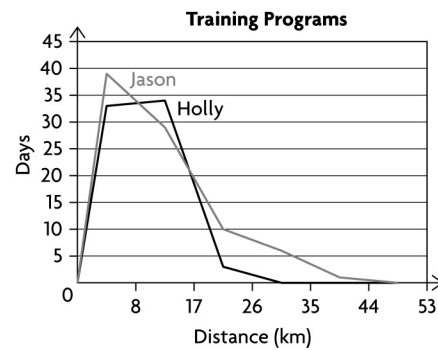


b) All graphs show a downward trend for the frequency of large numbers of errors. This means that as the time increases, the number of errors made decreases. Fewer vehicles have large numbers of errors.

8. a) e.g., I chose interval sizes that created five interval spaces that worked for both tables.

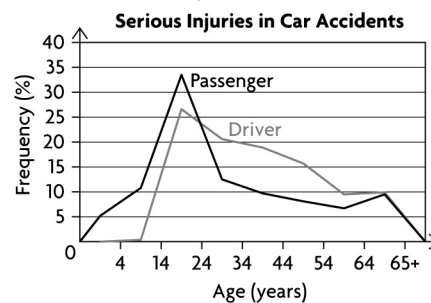
Distance (km)	Holly's Frequency	Jason's Frequency
0-8	33	39
9-17	34	29
18-26	3	9
27-35	0	6
36-44	0	1

b) Here is the diagram.



c) Holly program involves more short distance running, while Jason's program involves more long distance running.

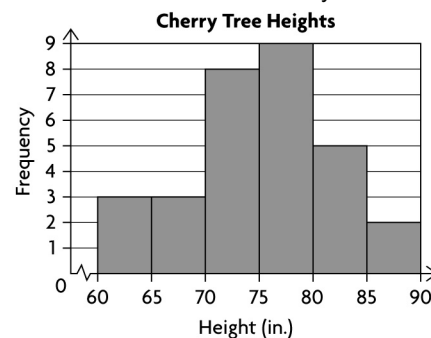
9. Here is the diagram.

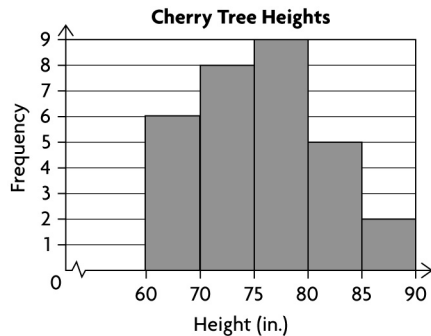


e.g., Younger drivers are involved in accidents where the injuries were serious. The greatest number of injuries for drivers and passengers occurred for the age group between 15 years and 24 years of age. One possible reason is that passengers could have been in the same vehicles as the drivers.

10. e.g., Histograms must use intervals of equal width to display the distribution of data accurately. If unequal widths are used, then the graph will be misleading.

For example, the first histogram shows intervals of equal width. The data looks normally distributed. The second graph uses the same data as the first graph but merges the first two intervals. The data is now skewed or not normally distributed.





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

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

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

$$\text{Mean population} = 5115.894\dots$$

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.

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

$$\text{Median} = 3114.754\dots$$

The median population is 3115 people.

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

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

$$\text{Percent error for median} = 83.235\dots\%$$

The percent error for the median is 83%.

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

$$\text{Percent error for mean} = 19.672\dots\%$$

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, σ			$\sqrt{\frac{1018}{5}} = 14.268\dots$

The standard deviation for Class A is 14.27%.