1.1 Constructing and Interpreting Visual Displays of Data

"More people are born in the spring than in the fall."

In order to draw general conclusions, such as the one above, information must be gathered, organized, and displayed clearly.

When information is gathered from all people in a **population**, the activity is called a census. For example, every five years, Statistics Canada takes a census of the population. The first census of the millennium was on May 15, 2001. The results from that census were released beginning in the spring of 2002.

A poll (or opinion survey) is a method of collecting **data** from a **sample** of a population by asking people to give their answers to a set of questions. Once collected, the data are then organized in a meaningful way so that valid conclusions can be made.

population—refers to the entire group about which data are being collected

data—information providing the basis of a discussion from which conclusions may be drawn; data often take the form of numbers that can be displayed graphically or in a table

sample—part of a population that is selected to gain information about the whole population

frequency—the number of times an event occurs or the number of items in a given category

frequency table—a table listing a variable together with the frequency of each value

Example 1 Organizing Data: Frequency Tables

The members of a Grade 12 class were asked on what day of the week they were born. The results were as follows:

Monday, Tuesday, Wednesday, Thursday, Monday, Friday, Friday, Tuesday, Thursday, Wednesday, Saturday, Friday, Tuesday, Wednesday, Saturday, Monday, Wednesday, Wednesday, Thursday, Thursday, Tuesday, Wednesday, Tuesday, Thursday, Tuesday, Thursday, Saturday, Tuesday, Sunday, Monday

(a) Organize the data in a **frequency table**.

(b) How many students responded to the question?

(c) What percent of the students were born on weekends?

Solution

(a)

Day	Tally	Frequency
Monday	𝍸	4
Tuesday	𝍸 ||	7
Wednesday	𝍸 |	6
Thursday	𝍸 |	6
Friday	|||	3
Saturday	|||	3
Sunday	|	1

(b) Total number of students $= 4 + 7 + 6 + 6 + 3 + 3 + 1$
$= 30$

(c) Percent born on a weekend $= \dfrac{\text{number born on weekend}}{\text{total number of students}}$
$= \dfrac{4}{30} \times 100\%$
$\doteq 13.3\%$

Example 2 Organizing Data: Frequency Tables and Stem-and-Leaf Plots

The heights of the members of two high school classes were measured in centimetres. The results were as follows:

136, 156, 172, 160, 175, 186, 187, 122, 186, 157, 153, 130, 164, 143, 181, 186, 176, 184, 193, 136, 122, 120, 184, 186, 176, 181, 167, 164, 149, 186, 155, 192, 174, 184, 156, 164, 181, 186, 172, 181, 163, 190, 188, 182, 174, 157, 152, 183, 171, 156

stem-and-leaf plot—a way of organizing and presenting a collection of numbers

(a) Display the data using a **stem-and-leaf plot**.

(b) Use your answer to part (a) to set up a frequency table using class intervals.

(c) How is the frequency table in part (b) different from the stem-and-leaf plot?

(d) Explain why class intervals are more useful than individual measurements.

Solution

(a) The set of numbers is rewritten, with each of the hundreds and tens digits becoming a stem and the units digits becoming the leaves. The leaves are entered in numerical order to produce a stem-and-leaf plot.

This branch represents the numbers 143 and 149.

Stem	Leaf
12	0 2 2
13	0 6 6
14	3 9
15	2 3 5 6 6 6 7 7
16	0 3 4 4 4 7
17	1 2 2 4 4 5 6 6
18	1 1 1 1 2 3 4 4 4 6 6 6 6 6 6 7 8
19	0 2 3

The units digits are called the **leaves**.

The hundreds and tens digits are called the **stems**.

class interval—a category or division used for grouping a set of observations

(b) A reasonable **class interval** for this data is a spread of 10 units. Given that the smallest value is 120 and the largest value is 193, the intervals to best display this data are 120–129, 130–139, and so on.

(c) In the stem-and-leaf plot, individual items were listed. In the frequency table in part (b), items were grouped into class intervals.

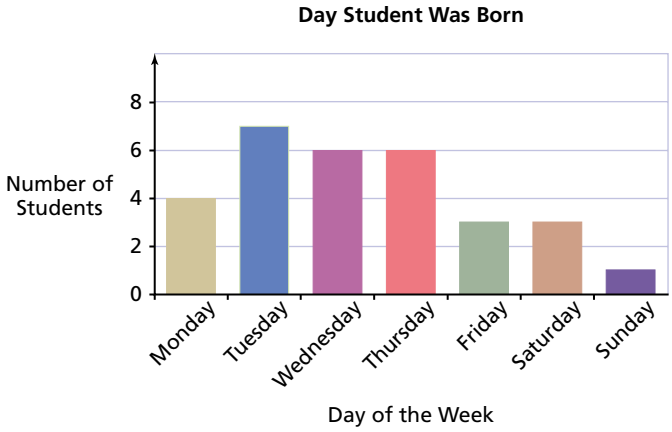
(d) Using individual items would create a table with the data so spread out that it would become difficult to view any trends.

Class Interval	Frequency
120–129	3
130–139	3
140–149	2
150–159	8
160–169	6
170–179	8
180–189	17
190–199	3

DISPLAYING DATA

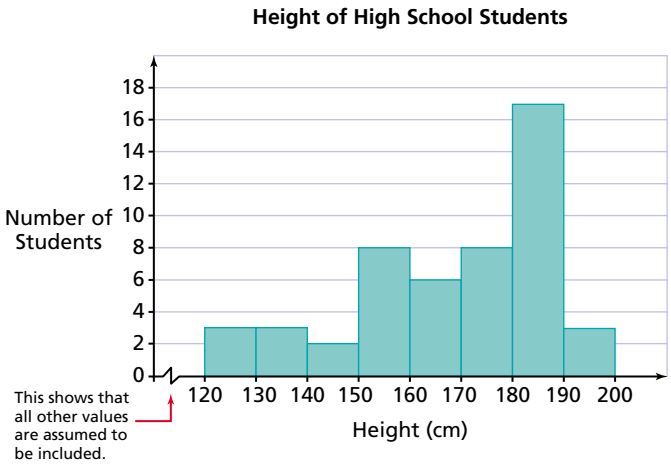
Tables are used to organize data; however, graphs are used to display data in a more meaningful way. A *bar graph* consists of parallel bars of equal widths with lengths proportional to the frequency of the variables they represent. A bar graph is used to represent nominal data, such as days of the week. Typically, bar graphs are used for discrete data. Look at the example on the top of the next page.

Project Connection
Give thoughtful consideration to the type of graph that will best display your data.



frequency distribution—a set of values of a variable, together with the frequency of each value

A *histogram* is a **frequency distribution** where the horizontal axis is divided into equal class intervals into which data have been divided. The heights of the rectangles represent the frequencies associated with the corresponding intervals. A histogram is used to represent data that are organized into class intervals, such as student heights. It is important that each interval have the same width. Typically, histograms are used for continuous data. The following is an example.



A *pictograph* is a symbolic representation of data. The following pictograph displays the number of participants (aged 15 and older) in the five most popular sports activities in Canada.

Sport Activity	Number of Participants
Golf	
Ice Hockey	
Baseball	
Swimming	
Basketball	

Legend: represents 100 000 people

Source: Statistics Canada, General Social Survey, 1998.

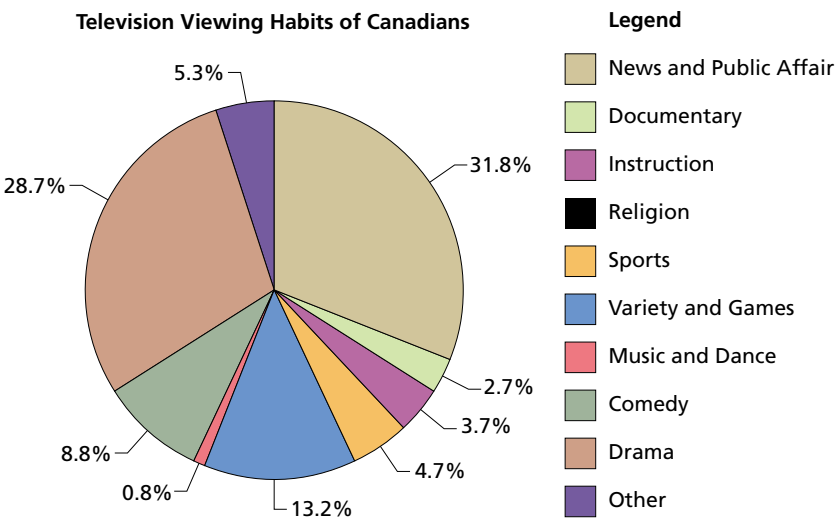
central angle—a measure of proportion in a pie chart. It is calculated as follows: central angle = percent of data for a category × 360°

A *circle graph* (or *pie chart*) is a circle divided into sectors whose areas are proportional to the quantities represented. The size of each piece of a pie chart is determined by the **central angle**. It is calculated by multiplying the percent of data for each category by 360°. For example, in the following table and pie chart, which display the television viewing habits of Canadians, Sports represents 4.7% of the data, or $0.047 \times 360^\circ \doteq 17^\circ$.

Programming Type	Percent of Viewing Time
News and Public Affairs	31.8
Documentary	2.7
Instruction	3.7
Religion	0.3
Sports	4.7
Variety and Games	13.2
Music and Dance	0.8
Comedy	8.8
Drama	28.7
Other	5.3

Source: Statistics Canada

Think about
TV Viewing Habits
What would the results from your class look like? How would you collect these data?



Example 3 Displaying Data

A survey conducted in a math class asked students what mode of transportation they normally use to get to school. The results were as follows:

Bike	10
Walk	15
Bus	9
Car	6

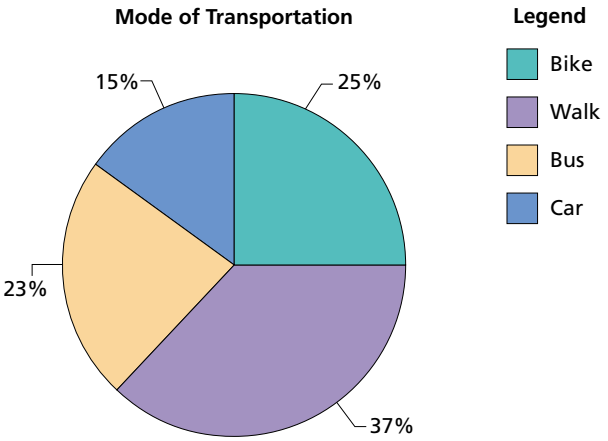
Construct a circle graph to display these data.

Solution

Calculate the angle for each sector, as shown in the following table.

Mode of Transportation	Number of Students	Central Angle
Bike	10	$\frac{10}{40} \times 360° = 90°$
Walk	15	135°
Bus	9	81°
Car	6	54°
Total	40	360°

Use the sector angles and a protractor to construct the circle graph.

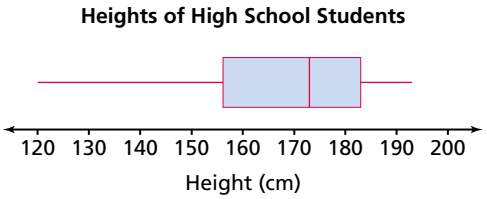


Graphs for Examining the Spread of Data

box-and-whisker plot—a horizontal representation of the spread of a distribution of data

median—the middle value in a frequency distribution

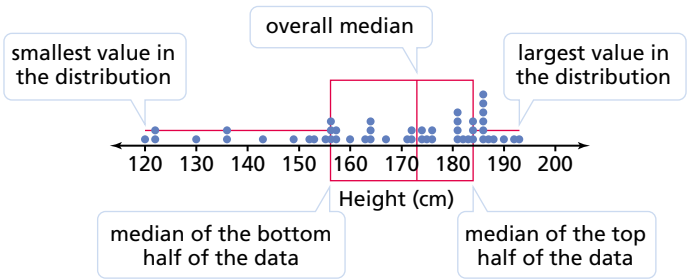
In a **box-and-whisker plot**, the box contains the **median** of the data and its width represents the middle half of the data. From the sides of the box, horizontal lines are drawn extending to the smallest and largest values of the data. The following box-and-whisker plot displays the spread of the heights of a group of high school students.



To find the median, the heights are listed in ascending order (smallest to largest). The middle value (or the average of the middle two values if there is an even number of items in the data set) is the median. The median for this data set is 173.

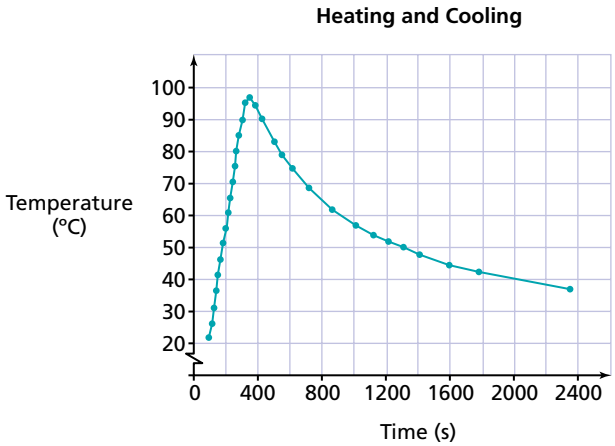
The medians of the upper half and the lower half of the data are calculated to find the upper and lower limits of the box. In this data set, these values are 184 and 156, respectively.

To construct the box-and-whisker plot, the data are plotted on a number line, and the three calculated values are indicated. A box is drawn around the central half of the data, and then lines are drawn extending to the smallest and largest values of the distribution to create the whiskers.



Graphs for Examining Trends

A **broken-line graph** is a graph created by joining data points with line segments. The following is an example. Notice that all graphs must include a title and the axes must be labelled. Pie charts and pictographs must also have legends.



KEY IDEAS

population—refers to the entire group about which data are being collected

data—information providing the basis of a discussion from which conclusions may be drawn; data often take the form of numbers that can be displayed graphically or in a table

sample—part of a population selected so as to gain information about the whole population

frequency—the number of times an event occurs or the number of items in a given category

stem-and-leaf plot—a way of organizing and presenting a collection of numbers

class interval—a category or division used for grouping a set of observations

bar graph—a visual display of data in which quantities are represented by bars of equal width; typically used with discrete data

frequency distribution—a set of values of a variable, together with the number of occurrences (frequency) of each value

histogram—a visual display of data that have been organized into equal class intervals, which appear on the horizontal axis. The heights of the rectangles are proportional to the frequencies they represent. Typically used with continuous data.

pictograph—a graph that uses pictures or symbols to represent variable quantities

circle graph (pie chart)—a circle divided into sectors whose areas are proportional to the variables represented

median—the middle term in a frequency distribution

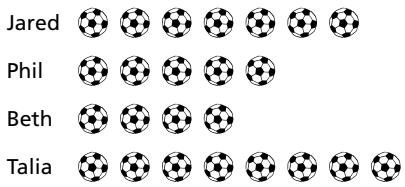
box-and-whisker plot—a horizontal representation of the spread of a distribution

broken-line graph—a graph created by joining data points with line segments

1.1 Exercises

A

1. The number of goals scored by the top four players on the school soccer team are displayed. Jared has 14 goals.

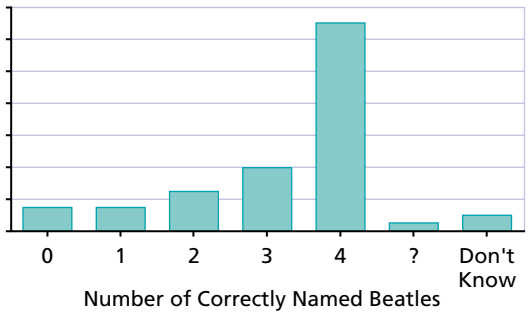


(a) What information is missing from the graph? Provide it.

(b) How many goals does each player have?

(c) What are the advantages and disadvantages of using a pictograph?

2. Some students were asked to name the four Beatles. Their responses are displayed below.



(a) The vertical scale is missing. If seven students knew the names of three of the Beatles, what would the scale be?

(b) How can the scale be altered, yet still display the same meaning?

(c) If these data represent the response of 1000 students at a local high school, how many would be able to name all four Beatles?

3. Some Grade 12 students were asked to estimate the number of hours of television they watch each day. These are their responses:

1, 1, 0.5, 1, 1, 0, 3, 2, 1.5, 0.5, 1, 1, 2, 2, 2, 5, 1, 0.5, 0.5, 2, 1, 0.5, 1, 0, 0.5, 3

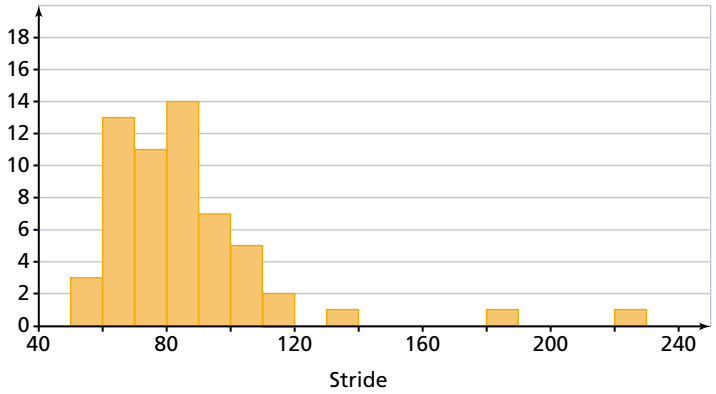
(a) Construct a tally and frequency table.

(b) Use the information to produce a graph.

(c) How many hours of television per day do you watch?

(d) Investigate the average daily television watching time for Canadians. Compare this information with that from the class. Draw conclusion(s) from your comparison. Give reasons for your answer.

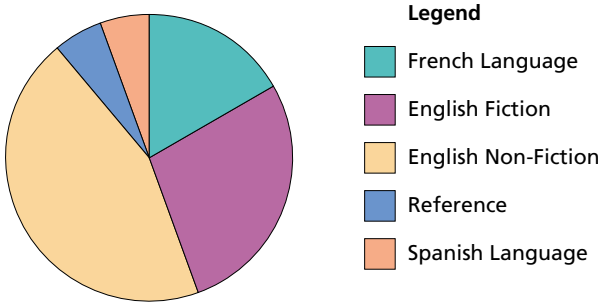
4. The stride lengths, in centimetres, of a group of Grade 12 students are displayed in the following histogram.



(a) What information is missing?
(b) In what ways are bar graphs and histograms similar?
(c) In what ways are they different?
(d) When is a histogram more appropriate than a bar graph? Give reasons for your answer.

5. The school library has received a donation and will purchase books based on the borrowing patterns of the students.

Library Borrowing Patterns



(a) Based on the graph, estimate what percent of the money should be spent on each book type.
(b) If the library received a donation of $125 000, estimate how much money should be spent on each book type.

6. Thirty people were asked to state their favourite sport. The responses are listed below.

Tennis	6
Football	8
Swimming	10
Badminton	3
Volleyball	3

Construct a circle graph to display this information.

7. **Knowledge and Understanding** The lengths of the songs on a CD compilation set are given in minutes and seconds below.

3:44	3:40	4:57	4:08	4:02	4:47
4:08	2:39	4:04	4:06	3:24	4:48
4:27	3:03	4:13	4:05	4:16	4:01
3:40	3:53	3:02	3:43	5:13	3:49
4:38	5:21	4:01	4:32	3:45	3:59
4:00	3:29	4:05	4:06		

What percent of the songs are less than 5 min in length?

(a) First, use a 30-s class interval to construct a frequency table.

(b) Then, use the table constructed in part (a) to answer the question.

(c) Explain why the table was more helpful than the list in answering the question.

(d) Which type of graph should be used to display these data? Construct it.

8. The number of hot dogs sold by a street vendor for each day in the month of June is recorded below.

112	98	108	128	24	30	89
106	48	34	16	71	122	71
102	118	53	76	76	25	72
52	33	122	33	109	109	110
116	21					

(a) Construct a stem-and-leaf plot to display the data.

(b) Use the plot to determine the number of days when fewer than 80 hot dogs were sold.

(c) On what percent of the days were more than 100 hot dogs sold?

9. The members of a Grade 9 class and a Grade 12 class were given a list of five animals: cat, dog, bird, fish, and iguana. From the list, they were asked to pick the animal they would want as a pet. The results were as follows:

Pet	Grade 9	Grade 12
Cat	2	7
Dog	4	6
Bird	9	3
Fish	10	6
Iguana	5	8

(a) Construct a pictograph for both classes.

(b) What percent of students did not choose a cat or dog?

(c) Compare your answer in part (b) to the percent of Grade 9 students who did not choose a cat or dog. Show your work.

10. **Application** For each of the following, determine the type of graph that would best represent the data collected. Explain.

(a) What portion of the general population reads a newspaper on a daily basis? On Saturday only?

(b) How does your heart rate change when doing each of the following activities: sleeping, walking slowly, jogging, running?

(c) Which type of vacation do people prefer: camping, tropical holiday, cruise, cottage, or visiting relatives?

(d) How do high school students spend a typical Saturday?

(e) How do students spend a typical weekday during the school year?

B

11. **Communication** For Question 12, below, write a plan explaining how you would carry out the investigation. List all the steps clearly and reference specific resources. For example, if you plan to use the Internet, list specific Web-site addresses.

12. Investigate the breakdown of blood types in Canada.

(a) Use your findings to construct a circle graph.

(b) If 750 people attended a high school basketball game, calculate the number that you would expect to have blood type A.

C

13. Choose five pop stars. Conduct a survey of students to determine their favourite pop star.

(a) Organize the data in a frequency table.

(b) Construct a bar graph to display the data.

(c) Create a question about your graph. Have someone in the class answer your question and provide an explanation for his or her answer.

14. Investigate the breakdown of enrolment in each of the Grade 12 math courses in your school.

(a) Construct an appropriate display of the data.

(b) Give reasons for your choice of display.

(c) Draw conclusions from your data. Give reasons for your answer.

15. **Thinking, Inquiry, Problem Solving**

(a) Conduct a survey of students that asks one or more of the questions that follow.

(i) How many hours per week do you spend at your part-time job?

(ii) How much money per month do you spend on entertainment (e.g., CDs, concerts, movies, etc.)?

(iii) How tall are you in centimetres?

(iv) How many hours of television do you watch each week?

(v) List all the sports that you have participated in during the last four weeks.

(b) Display your results in graphical form.

(c) Draw any conclusions that arise from the data. Explain.

(d) Reflect on how you conducted the survey. Explain any improvements that you would make if you were to conduct the survey with a different group of students.

ADDITIONAL ACHIEVEMENT CHART QUESTIONS

16. **Knowledge and Understanding** The 1996 Census of Canada reports the following composition of Canadian families. Create two different displays of these data.

Number of Children	Number of Families (in thousands)
Without children at home	2730
With one child	2106
With two children	2047
With three children	729
With four children	175
With five or more children	51
Total families	7838

Source: Statistics Canada

17. **Application**

(a) Choose a topic that interests you and survey your classmates to find their responses.
(b) Organize your data in a frequency table.
(c) Create the most appropriate graph to display your data.
(d) Use the table and your graph to draw a conclusion about your class.

18. **Thinking, Inquiry, Problem Solving**

(a) Use the Internet to find data that show a trend over time.
(b) Use the most appropriate graph to display your data.
(c) Use the graph to make a prediction. Give reasons for your answer.

19. **Communication** Explain the difference between histograms and bar graphs, and provide an example of a set of data that is best suited for each of these types of graphs.

Chapter Problem

Trends in Canada's Population

Use the data given in the chapter problem on page 2 to answer these questions.

CP1. Create three different types of graphs that can be used to compare the structure of Canada's population in 1996.
CP2. Of the three graphs you created, which graph best displays the similarities and differences in the structure of the population? Explain.