4 Statistical Literacy

Key Words

- percentile
- quartile
- poll
- margin of error
- representative sample
- bias
- valid conclusion
- index
- base value
- inflation

What You'll Learn

To interpret statistical data and to assess whether conclusions based on statistical analysis are valid

And Why

Statistical data are often included in news articles, informational pamphlets, and advertisements. As a media consumer, you need to know how to analyse data to avoid being misled.

Ratios

A ratio compares two or more quantities. Ratios that make the same comparison are **equivalent**.

Example

A marketing company conducted a survey at a mall. Researchers asked 40 teens whether they owned a car. Eight teens owned a car.

a) What is the ratio of teens who owned a car to the total number surveyed?

b) Write an equivalent ratio with each second term:

   i) 20   ii) 100

Solution

a) 8 teens owned cars out of 40 teens surveyed.
   The ratio is 8:40.

b) Divide or multiply each term of the ratio in part a by the same number.

   i) $8:40 = ?:20$ ii) $8:40 = ?:100$
   Since $40 \div 2 = 20$,   Since $40 \times 2.5 = 100$,
   divide each term by 2.   multiply each term by 2.5.
   $8:40 = (8 \div 2):(40 \div 2)$ $8:40 = (8 \times 2.5):(40 \times 2.5)$
   $\quad = 4:20$ $\quad = 20:100$

CHECK ✔

1. Thirty people tasted a new brand of cheese. Twenty of them liked the taste.
   a) What is the ratio of those who liked the taste to the total number who tried it?
   b) Write an equivalent ratio with each second term:
      i) 6   ii) 100

2. There are 700 students enrolled in a high school. The ratio of girls to boys is 4:3.
   a) How many boys and how many girls go to this school?
   b) The average class size is 28 students. Suppose this class is representative of all the students in the school. How many students in the class are girls? How many are boys?

3. In question 1, what is the meaning of the equivalent ratio with a second term of 100?

Measures of Central Tendency and Range

Prior Knowledge for 4.1

The **mean, median,** and **mode** are measures of central tendency for a data set. They represent a typical value for the set. The **range** is a measure of spread. It tells you the difference between the greatest and least numbers in the set.

> Mean, median, and mode are defined in the glossary at the back of this book.

Example

Eight students received these marks on a test: 85, 76, 91, 65, 68, 72, 78, 43

a) Calculate the mean, median, and mode mark.
b) Which measure of central tendency best describes these data? Explain.
c) Determine the range of the data set.

Solution

a) For the mean:
   $(85 + 76 + 91 + 65 + 68 + 72 + 78 + 43) \div 8 = 578 \div 8$, or 72.25
   The mean is 72.25.
   For the median, list the numbers in order.
   43, 65, 68, **72, 76,** 78, 85, 91
   Since there are 2 middle numbers, add them and then divide by 2.
   $(72 + 76) \div 2 = 148 \div 2$, or 74
   The median is 74.
   The mode is the number that occurs most often.
   Each number occurs once. So, there is no mode for this data set.
b) The mark 43 is much less than the other marks. It reduces the mean, but has no effect on the median. There is no mode. So, the median best describes the data.
c) The range is the difference between the greatest and least marks:
   $91 - 43 = 48$

CHECK ✔

1. The heights, in metres, of trees in a woodlot are as follows:

18.0	21.3	17.1	23.5	19.8	17.9	17.0	21.5	19.2	19.0	20.6	19.5

a) Calculate the mean, median, and mode tree height.
b) Determine the range of this data set.

2. Which measure of central tendency best describes the data in question 1? Explain.

Activate Prior Knowledge **193**

Percent Increase and Decrease

Percent increase and **decrease** are often used to describe how quantities have changed over time. They describe the change in a quantity as a percent of its original value.

Example

In August, a sports store priced a shoe at $125.99.
The shoes were so popular that the store increased the price to $134.99 in September. Shoe purchases declined. So, the store reduced the price to $119.99 in October.

a) What was the percent increase in price from August to September?
b) What was the percent decrease in price from September to October?
c) What was the percent decrease in price from August to October?

Solution

For each calculation, write the difference in the prices as a fraction of the earlier price. Then multiply by 100 to determine the percent increase or decrease.

a) $\dfrac{\$134.99 - \$125.99}{\$125.99} \times 100 \doteq 7.1$ The price increased by about 7.1%.

b) $\dfrac{\$119.99 - \$134.99}{\$134.99} \times 100 \doteq -11.1$ The price decreased by about 11.1%.

c) $\dfrac{\$119.99 - \$125.99}{\$125.99} \times 100 \doteq -4.8$ The price decreased by about 4.8%.

CHECK ✓

1. The population of a city was 18 500 last year.
   This year the population is 21 300.
   What is the percent increase in population?

2. Ms. Voisin was trying to sell her house for $325 000.
   When it had not sold after several weeks, she lowered the asking price to $298 000.
   What is the percent decrease in price?
   Use this information to write an advertisement headline for the house.

Ethical Conduct

Colleges and workplaces have strict rules about ethical conduct. Breaking these rules can have serious consequences. College students can be suspended; tradespeople may lose their right to practise; or employees can lose their jobs.

In these scenarios, the student, apprentice, or employee has committed a breach of ethics. Use your Internet skills and general knowledge to investigate rules and penalties for such conduct.

- Troy, a second year fashion student, hands in essays with paragraphs copied from the Internet. He does not reference the sites.
- Nora quotes $1200 to insulate an attic with 10 inches of cellulose fibre. She uses only 8 inches, but says the job is done and collects $1200.
- Joshua uses his 15% discount to purchase appliances for friends.
- Jane got an A on an essay in a business management course. She lends the paper to friends who take similar courses with different professors and at different schools.
- Stefan is an apprentice plumber who does jobs on weekends without permits. On these jobs, he does not make sure everything is "to code."
- Sharon, a receptionist at a doctor's office, discusses some of the patients and their medical problems with friends.
- Jon takes home printer paper from work to use for his evening courses.

1. Describe ethical problems in these scenarios.
2. As you create work of your own and work with others in this chapter, make notes about how you can follow ethical standards.
3. Imagine being an apprentice, employee, or college student. How could you apply what you have learned about ethics?

Search words for student ethics
- ☐ Academic misconduct
- ☐ Academic dishonesty
- ☐ Academic integrity
- ☐ Plagiarism
- ☐ Student resposibilities
- ☐ Student regulations
- ☐ College policies

• Complete questions 2 and 3 after you complete Chapter 4.

Search words for workplace ethics
- ☐ Ethical conduct
- ☐ Ontario employment standards
- ☐ Ontario plumber ethics
- ☐ Ontario paramedic ethics
- ☐ Workplace fraud
- ☐ Workplace racism
- ☐ Corporate bullying
- ☐ Employee code of conduct

A nurse uses growth charts to monitor a child's development. Growth charts are graphs showing what percent of children in the population are at or below particular heights or weights for different ages.

Investigate

Materials
• scientific calculator

Organizing and Describing Data

Work with a partner.

The 16 students in Jesse's math class measured their heights to the nearest centimetre.

> 160, 178, 167, 180, 168, 157, 164, 179,
> 163, 182, 176, 170, 172, 165, 175, 167

■ Determine the measures of central tendency and the range for this set of data.

■ What percent of the class is shorter than each measure of central tendency?

■ Rylan is taller than 65% of the class. How many students are shorter than he is? What is Rylan's height?

■ Why might people want to know how they compare in height to classmates or to other people their age?
■ How might information about typical heights and weights at different ages be of use to a clothing manufacturer?

Connect the Ideas

Television, radio, newspapers, and Web sites often report statistical data. To understand these reports, you need to be familiar with the statistical language they use.

Percentiles

A **percentile** tells approximately what percent of the data are *less than* a particular data value. Percentiles are a good way to rank data when you have a lot of data or want to keep data private.

The World Wise Quiz

Your score:	96/200
Rank among your friends:	3/12
Rank among all quiz-takers:	67th percentile
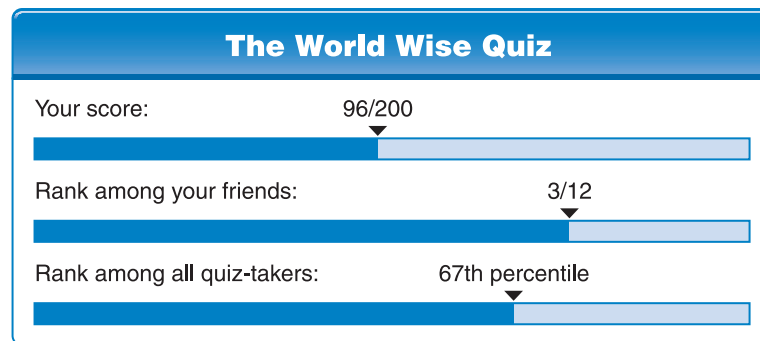
Quartiles

A **quartile** is any of three numbers that separate a sorted data set into four equal parts.
■ The second quartile is the median.
  It cuts the data set in half.
  So, it is the same as the 50th percentile.
■ The first, or lowest, quartile is the median of the data values less than the second quartile.
  It separates the lowest 25% of the data set.
  So, it is the same as the 25th percentile.
■ The third, or upper, quartile is the median of the data values greater than the second quartile.
  It separates the highest 25% of the data set.
  So, it is the same as the 75th percentile.

| Example 1 | **Working with Percentiles** |

Here are the hourly pay rates, in dollars, for 17 high-school students with part-time jobs.

11.50	10.20	8.00	8.25	9.00	9.15
9.75	7.50	8.00	12.50	13.00	11.25
10.75	9.50	9.25	9.45	7.75	

a) What are the quartiles for this data set?
b) Damien's pay is in the 85th percentile for this group. What does the percentile mean? What is Damien's hourly pay rate?

Solution

a) Start by finding the median. Order the data.
7.50, 7.75, 8.00, 8.00, 8.25, 9.00, 9.15, 9.25, **9.45,**
9.50, 9.75, 10.20, 10.75, 11.25, 11.50, 12.50, 13.00
The second quartile is the median: $9.45 per hour

Look at the ordered data that are less than the second quartile.
The lower quartile is the median of these data:
$$\frac{\$8.00 + \$8.25}{2} \doteq \$8.13 \text{ per hour}$$

Look at the ordered data that are greater than the second quartile.
The upper quartile is the median of these data:
$$\frac{\$10.75 + \$11.25}{2} = \$11.00 \text{ per hour}$$

8.13   11.00

7.50, 7.75, 8.00, 8.00, 8.25, 9.00, 9.15, 9.25, 9.45, 9.50, 9.75, 10.20, 10.75, 11.25, 11.50, 12.50, 13.00

Lower quartile   Second quartile   Upper quartile
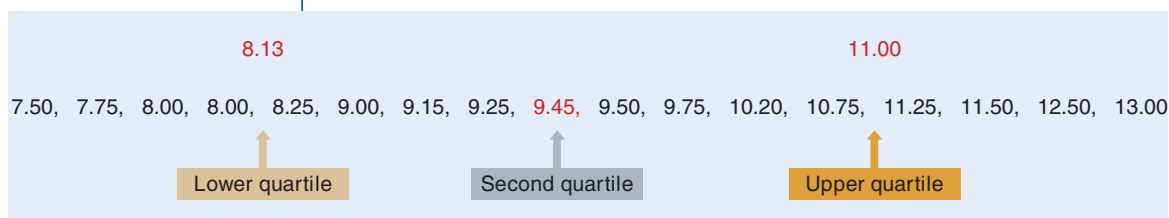
b) The 85th percentile means approximately 85% of the students in the group earn less money per hour than Damien.
$17 \times 0.85 = 14.45$
Round down to the nearest whole number to determine the number of students who earn less money per hour than Damien: 14
So, Damien is the 15th student in the ordered list.
He earns $11.50 per hour.

Data reliability

When you read statistical data, you need to think about the reliability of the source. Data from a government agency are usually more reliable than data from someone who is trying to sell a product or promote a point of view.

| | *Example 2* | | **Comparing Data Sources** |

In each case, a research topic and two sources of information are described. Decide which data source is more likely to provide reliable data.

	Topic	Source 1	Source 2
a)	The benefits or adverse effects of drinking milk	A pamphlet from an animal rights group that opposes dairy farming	Canada's Food Guide produced by Health Canada
b)	The effect of logging on the population of a species of bird	A pamphlet from a wildlife protection organization	A forestry company advertisement
c)	Possible complications of flu shots	A Ministry of Health Web site	A Web site run by a group that opposes immunizations

Solution

a) The animal rights group is promoting a particular point of view and may not be objective. The Food Guide was developed in consultation with thousands of dietitians, scientists, physicians, and public health personnel from across Canada. It will represent a more balanced view.

b) Both the wildlife organization and the forestry company are promoting particular points of view. It would be best to look for additional sources.

c) While the Ministry of Health Web site promotes the use of immunizations, it also describes possible complications. A group opposing immunizations is more likely to present only one side of the issue.

Polls

Polling companies conduct interviews with randomly selected Canadians to determine their opinions about a variety of topics. These surveys are called **polls**.

The results of polls are often reported in the media, particularly during elections. Poll results usually state a **margin of error** that describes how reliable the data are. As a media consumer, you need to know how to interpret these results.

Example 3

Interpreting Poll Results

The results of a poll conducted by EKOS in 2005 are shown.

"Canada should increase its humanitarian aid to poor countries even if it means less spending in other important areas"

	Disagree (1–3)	Neither (4)	Agree (5–7)
Aug 2005	39	18	43
Jan 2005	46	22	31

0%   20%   40%   60%   80%   100%

■ Disagree (1–3)   ■ Neither (4)   ■ Agree (5–7)
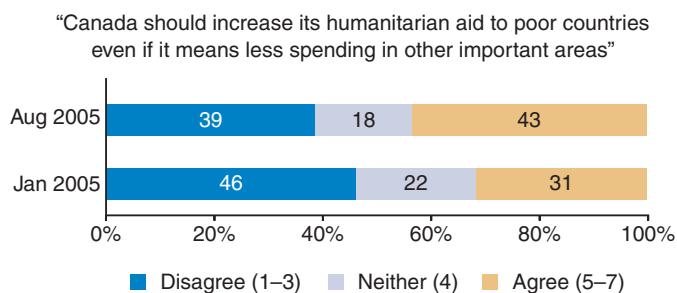
a) What question were people asked?
b) How did the favourable responses compare in January and August?
c) A line below the graph stated "The results are valid within a margin of error of plus or minus 2.5 percentage points, 19 times out of 20." What does this mean?

Solution

a) People were asked whether they agreed with the statement: "Canada should increase its humanitarian aid to poor countries even if it means less spending in other important areas."
b) In January, 31% of the people polled agreed with the statement. In August, the percent agreeing had increased to 43% of those polled.
c) If you had taken another poll, there is a 19 out of 20, or 95% chance that the results would be within 2.5 percent of these results. That is, in August, somewhere between 40.5% and 45.5% of the people polled would agree with the statement.

1. Determine the quartiles for each data set.

   a) 10, 8, 12, 15, 9, 9, 11, 12, 12, 8, 14, 11

   b) 170, 162, 150, 165, 180, 165, 154, 163, 168, 164, 172

2. Identify the statistical information in each excerpt from a Statistics Canada news release about a survey on college and university graduates' debt 5 years after graduation.

a)
Two out of five graduates from the class of 2000 who had left school owing money to government student loans had completely repaid their debt five years after graduation.

b)
Of all graduates from a Canadian college or university in 2000, 56% had no debt from government student loan programs while 44% owed money to such programs.

c)
Slightly less than half of the graduates who still owed money on their student loans reported having difficulty repaying these loans, compared to one out of five among graduates who had paid off their loans by 2005.

3. The Grade 9 students in a high school participated in a national standardized math test. The principal reported on the school Web site that three of the students placed in the upper quartile. Which sentence best describes the meaning of this statement?

   i) Three students received a mark of at least 75%.

   ii) Three students did better than 75% of all those who wrote the test.

4. In each case, a research topic and two sources of information are described. Decide which data source is more likely to provide reliable data. Justify your answers.

	Topic	Source 1	Source 2
a)	The sound quality of a particular stereo	An advertisement in a magazine	A review in a consumer magazine
b)	The possible side effects of a medication	A health information Web site run by a hospital	A blog written by someone who has taken the medication
c)	Job prospects in a particular field	A PDF document written by the Ontario Ministry of Training, Colleges, and Universities	A brochure you receive advertising mail-order courses in this field

5. **Literacy in Math** The headline on the press release for the poll results in *Example 3* read: "Support for Foreign Aid Rises Compared with Other Priorities." Would you say this headline is accurate? Justify your answer.

6. In December 2006, UNU World Institute for Development of Economics Research released a study on the global distribution of personal wealth in 2000. It included these statements:

- The richest 1% of adults alone owned 40% of global assets.
- The richest 10% of adults accounted for 85% of the world total.
- The bottom half of the world adult population owned barely 1% of global wealth.

Demonstrate the meaning of these statements visually by colouring squares on graph paper or using another method of your choice.

7. Sunny works as an assistant to a real estate agent.
She prepares a price comparison for a client thinking of selling a house.
Sunny investigates the list prices and sale prices of comparable houses that
have recently sold in nearby neighbourhoods.

	House 1	House 2	House 3
List price ($)	324 500	379 000	299 900
Sale price ($)	315 000	370 000	295 000

a) Determine the mean and median list and sale prices and the price ranges.
b) Which measure would you use for estimating what the client's house
   might sell for? Explain your choice.

8. Use the growth chart for
girls aged 10 to 20.
  a) Determine the percentile
     ranking for each girl.
     i) Tameika is 14 years old
        and 165 cm tall.
     ii) Audra is 16 years old
        and 152 cm tall.
     iii) Sabrina is 19 years old
        and 174 cm tall.
  b) Asayo is 17 years old
     and 165 cm tall.
     Between what 2 quartiles
     is her percentile ranking?

Stature-for-age Percentiles for Girls
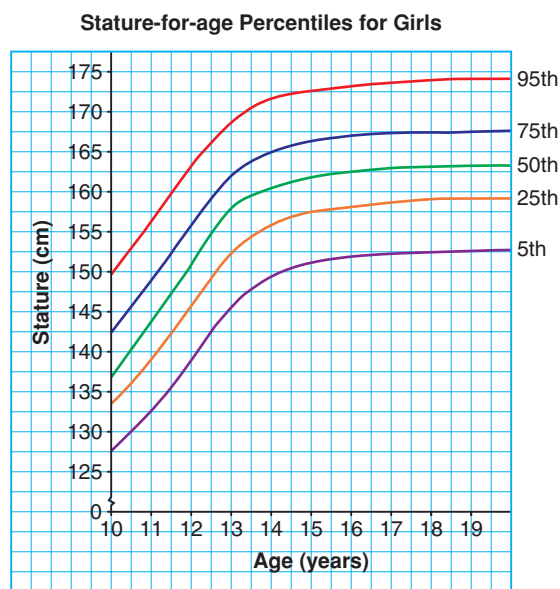
9. Here are the exam marks for a class of 20 math students.

35	72	74	84	90	60	93	48	70	68
75	63	65	75	82	65	54	77	64	59

a) Determine the mean, median, and mode.
   Which measure of central tendency best represents the data?
b) What are the quartiles for this data set?
c) Vince's mark is in the 37th percentile for this group.
   Explain what the percentile means. What is Vince's mark?

10. In November 2007, Ipsos Reid conducted a poll of 1314 randomly selected Ontarians. The report stated:

> …eight in ten (80%) Ontarians support "legislation that would ban smoking in cars and other private vehicles where a child or adolescent under 16 years of age is present". Moreover, a majority of non-smokers (86%) and smokers (66%) would support this legislation…

a) What question were people asked?

b) How did the responses of smokers and non-smokers compare?

c) The poll results are considered accurate to within ±2.7 percentage points, 19 times out of 20. Explain what this statement means.

11. **Assessment Focus** In November 2007, Harris/Decima polled just over 1000 Canadians about their toy-buying habits. The response to one question is shown.

Please say whether you are certain to, likely to, unlikely to or certain not to...
Avoid toys made in China because of concerns about health or safety risks.

Total	30%	25%	25%	13%	6%

0%   20%   40%   60%   80%   100%

☐ Certain to ☐ Likely to ☐ Unlikely to ☐ Certain not to ☐ DK/NR
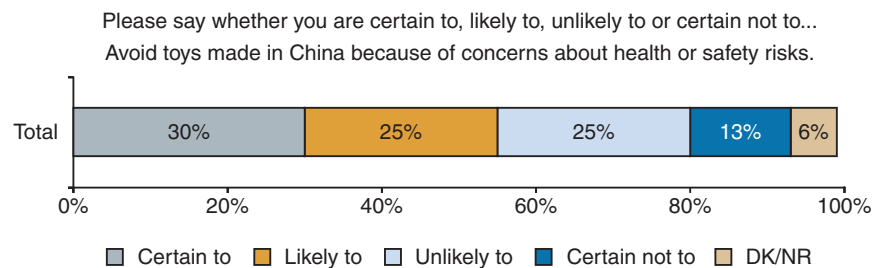
a) What was the question?

b) What percent of respondents said they were certain or likely to avoid toys made in China?

c) The poll's margin of error is 3.1%, 19 times out of 20. Explain what this statement means.

d) Write a headline that could be used in a news story about this poll.

12. Cong reads on a Web site that determining the "85th percentile speed" is part of the process of setting speed limits.

a) Which sentence describes the meaning of an 85th percentile speed? Justify your choice.

  i) The speed below which 85% of motorists are travelling.

  ii) The speed at which 85 out of every 100 cars on the road are driving.

b) What percent of drivers travel faster than the 85th percentile speed?

c) How could an 85% percentile speed be determined for a particular road?

d) Do you think it is reasonable to use this information to help set speed limits? Explain your thinking.

C **13.** The real estate agent in question 7 knows that each of the three houses has upgrades and features that the client's house lacks. She prepares an itemized list for each house and asks Sunny to adjust the prices.

	House 1	House 2	House 3
List price ($)	324 500	379 000	299 900
Sale price ($)	315 000	370 000	295 000
Adjustment ($)	−40 000	−63 000	−23 000

a) Determine the adjusted list and sale price for each house.
b) Determine the mean and median adjusted list and sale prices and the adjusted price ranges. How do they compare to the measures in question 7?
c) Which measure would you use for estimating what the client's house might sell for? Explain your choice.

14. Transportation engineers are considering changing the speed limit on a rural road. Every day for 1 week, a technician records speeds of vehicles using the road in kilometres per hour. Here is a representative sample of data:

76	74	78	75	69	68	87	90	73	70
68	72	85	78	72	70	75	75	76	65

a) What are the quartiles for this data set?
b) Determine the 85th percentile speed for these data.
c) The current speed limit on this road is 70 km/h. Based on these data, would you recommend changing it? What other factors should be considered?

In Your Own Words

Vivian thinks the first quartile of a data set is always a piece of data in the set. Alessandro thinks the first quartile is never in the set. Is either person correct? Justify your answer.

Whether a journalist reports for television, radio, newspapers, magazines, or the Internet, part of her or his job is to describe statistical data in a way that people can understand. How good a job do journalists do?

Inquire

Materials

- newspapers or magazines
- computer with Internet access

If you come across a statistical term you do not know, look it up in a dictionary or the glossary at the back of this book.

Finding Statistics in the Media

Work with a partner.

1. With your partner, create a list of words and phrases related to statistics, based on the materials you have available for this activity.

2. Select two examples from your work in question 1:
  - State where you found the data.
    If a source for the data is given, state that as well.
  - Describe the context in which the data were used.
  - Decide how reliable you think the data are.
    Explain what factors influenced your decision.

3. Find some data that are presented with graphs or other visual presentations. Write to explain how the image helped you understand the data. If you think it did not help, describe the problems with the presentation and suggest how it could be improved.

4. Find an advertisement or an article that is using data to promote a viewpoint.
  a) Decide whether or not you agree with the viewpoint. Explain your decision.
  b) If you agree with the viewpoint, find more data from another source that supports it. If you disagree with it, find data from another source that contradicts it.

5. Many media organizations use their Web sites to pose a "question of the day." Find a few of these informal surveys on the Internet. For each example you find, answer these questions.

a) What was the survey question?

b) How many people responded? What were their opinions?

c) Is the question related to any information presented on the Web site? If so, might the information affect how people answer? Explain your thinking.

d) Are people able to respond to the survey more than once? How can you find out? How might this affect the results?

Do you think your safety is compromised by those who chat on cellphones while driving?

Yes
95%
11444 votes

No
5%
652 votes

Do you think that Canada and the European Union should enter a free-trade deal?

Yes	74%
No	21%
Don't know / don't care	5%

Total votes for this question: 19

Poll

What is your favourite genre of music?

Pop – 17%

Rock – 38%

Country – 24%

Hip-Hop – 16%

Indie – 5%

Total votes 32053

Should politicians be allowed to overrule nuclear safety experts?

Yes   592 **17%**
No   2781 **82%**

VOTE

Which actor would you like to co-star in a flick with?

Jim Carrey
29%

Sandra Oh
14%

Mike Myers
26%

Ellen Page
31%

975 responses,
not scientifically valid, results
updated every minute.

What kind of news do you most prefer to read?

International news	29%
Local news	48%
Political news	12%
Entertainment news	9%
Environmental news	2%

Total votes for this question: 2043

POLL RESULT

What sport are you most likely to follow on TV this weekend?

Basketball	296 votes	(4 %)
Football	3000 votes	(38 %)
Figure skating	1368 votes	(17 %)
Hockey	3192 votes	(41 %)

Reflect

- Share your list from question 1 with the class or other pairs. Create a class list of common statistical terms and expressions.
- Describe three reasons why it is important to be able to recognize and understand common statistical words used in the media.
- Choose one statistic that you think was used effectively in a news article or advertisement. Explain what the statistic means and describe its importance to the article or advertisement in which you found it.

Government agencies, news organizations, and marketing companies often conduct surveys. The data collected can be factual, such as the number of brothers and sisters a person has, or subjective, such as a person's opinion about the use of lawn pesticides.

Excellent

Good

Satisfactory

Poor

Investigate

Assessing the Validity of Survey Results

Work with a partner.

Choose to analyse one of these case studies:
– Case Study 1: *What Drinks Should Be Sold in the School Vending Machine?*
– Case Study 2: *Are Part-time Jobs Related to Sleep Deprivation?*

■ Decide whether the survey is valid. List reasons for your decision. Think about:
  • Sample size; population size
  • The method of selecting respondents
  • The survey questions

Reflect

■ Find a pair that analysed the other case study.
■ Discuss your findings, inviting the students in the other pair to add any thoughts they have.
■ Listen to their analysis of the other case study and discuss their decision.

Case Study 1
What Drinks Should Be Sold in the School Vending Machine?

A high school vending machine can sell four different bottled drinks. The student council surveys the students to determine what drinks should be sold.
- There are 1200 students in the school.
- The student council prepares this interview script and tally sheet:

Please help us select four bottled drinks for the vending machine by answering these questions.
What grade are you in?
Which drink would you be most likely to buy from a school vending machine? Please choose one.

Vending Machine Survey

Student's grade	Drink choice					
	Cola	Diet cola	Sport drink	Iced tea	Cranberry juice	Apple juice

- Several council members gather outside the cafeteria during lunch one day. They interview students walking by and record the students' choices on tally sheets. They ask each student to participate only once.
- The council members stop when they have 100 responses.
- They tabulate the results and determine the four most popular drinks.

Case Study 2
Are Part-time Jobs Related to Sleep Deprivation?

A reporter wants to see if there is a correlation between the number of hours a high school student spends working at a part-time job and the number of hours of sleep the student gets.
He designs this questionnaire.

1. ☐ Male   ☐ Female

2. ☐ Grade 9   ☐ Grade 10
   ☐ Grade 11   ☐ Grade 12

3. Do you have a part-time job?

   ☐ No   ☐ Yes
                   Number of hours you
                   work in a typical week: _____

4. Doctors say teens require 8.0 h to 9.5 h of sleep each night.
   About how many hours of sleep do you get on a typical night?

   ☐ Less than 6 h
   ☐ Between 6 h and 7 h
   ☐ Between 7 h and 8 h
   ☐ Between 8 h and 9 h
   ☐ Between 9 h and 10 h
   ☐ 10 h or more

He puts the questionnaire in his newspaper and invites students or their parents to e-mail or call in their responses.
He collects 200 responses, then analyses the data.

Connect the Ideas

Sample size

Sample size can affect survey results. If the sample is too small, the survey results may not be reliable. If it is too great, the survey may be costly and difficult to administer.

Representative samples

A sample needs to be typical of the entire population. This is called a **representative sample**. If the sample is not representative, it is **biased** and the survey results are invalid.

Sampling techniques

Some sampling techniques are **random**, which means each member of the population has the same chance of being selected. A non-random technique may not yield a representative sample.

> These sampling techniques are described in the glossary.

Random techniques
- Simple random sampling
- Stratified sampling
- Cluster sampling
- Systematic sampling

Non-random techniques
- Convenience sampling
- Judgement sampling
- Voluntary sampling

Example 1

Assessing the Sample

A town has a population of 20 000 people. The town council conducts a vote at a public meeting about constructing a new ice-hockey rink.
- 50 people attend the meeting.
- 40 of the people at the meeting vote in favour of the hockey rink.
- The council decides to build the hockey rink since 80% of the people support the idea.

a) What percent of the people at the meeting voted for the rink?

b) What percent of the people in the town attended the meeting?

c) Is the sample representative? Justify your answer.

Solution

a) Since $\frac{40}{50} \times 100 = 80$, 80% of the people at the meeting voted for the rink.

b) Since $\frac{50}{20\,000} \times 100 = 0.25$, only 0.25% of the residents of the town attended the meeting.

c) The sample is not representative, for several reasons.
   - The sample size is too small.
   - It is a voluntary sample; only people who attended the meeting could vote.
   - It is probably a biased sample; people who chose to attend the meeting probably have an opinion about the arena.

Biased questions

Biased questions restrict people's choices unnecessarily or use words that could influence people to answer in a certain way.
For results to be valid, survey questions must be unbiased.

Survey techniques

Another factor to consider is how the survey is conducted.
This is particularly important if any of the questions are about sensitive subjects. People may be more likely to answer honestly if they can reply anonymously in writing rather than responding to an interviewer in person or over the phone.

Example 2	**Assessing the Question**

People were asked this survey question in phone interviews:
"We harm the planet when we use pesticides on our lawns.
Should the government ban all residential pesticide use?"
Will the survey results be valid? Justify your answer.
If you feel the survey is not valid, how could it be improved?

Solution

The results will not be valid. The first statement biases the question. Respondents may agree even if they wish to disagree. To determine peoples' true opinions on this issue, the group conducting the survey should omit the first sentence and use a written survey so people can respond anonymously.

| Example 3 | **Assessing the Entire Survey Process** |

About 4000 people visited a large sports equipment store during its annual sale. The store surveyed 100 customers after they paid for their purchases. An employee recorded their answers.

1. Good sports equipment can greatly improve performance. How much do you spend on equipment each year?

___ $200 or less   ___ $200–$400

___ $400–$600   ___ $600–$800

___ $800–$1000   ___ More than $1000

2. How much do you earn per year?

___ Less than $10 000   ___ $10 000–$20 000

___ $20 000–$40 000   ___ $40 000–$60 000

___ $60 000–$80 000   ___ More than $80 000

Why are the survey results invalid? How could the survey be improved?

Solution

To assess the survey, ask yourself these questions.

■ **Is the sample size large enough?**

The store sampled 100 people out of 4000 people.

$\frac{100}{4000} \times 100 = 2.5$

A sample of 2.5% of the customers is too small.

■ **Is the sample representative?**

The store only surveyed people who made a purchase.

The sample does not represent people who visited the store and did not purchase anything.

As people leave the store, every 10th person could be asked to answer the survey questions. This way everyone has a chance of being asked and the store has a greater sample size.

■ **Are the survey questions unbiased?**

The first question contains a statement that may encourage people to exaggerate the amount of money they spend on sports equipment. This sentence should be omitted.

■ **Was the collection method appropriate?**

Having an employee record the answers may be intimidating.

People may not wish to share information about their spending habits or their salaries.

The survey should be conducted as an anonymous written survey.

1. Three schools each survey 300 students about whether they want a longer lunch. What percent of students in each survey want a longer lunch?

School	Number of students who want a longer lunch
a) 1	60
b) 2	270
c) 3	175

2. For each population, determine how many people should be surveyed to include 10% of the population.

a) 350 people   b) 930 people   c) 1180 people   d) 10 360 people

3. The student council at a school surveys 50 students. What percent of each population is this? Choose one population and explain whether you think it is a large enough sample.

a) 450 students   b) 750 students   c) 1200 students

4. Would you conduct a survey on each topic using personal interviews or written forms?

a) Household income and spending on travel

b) Time spent on homework and student marks

c) Preferences for different brands of shampoo

d) Favourite colours and gender

e) Whether people have encountered discrimination in their lives

5. Identify whether each survey question is biased or unbiased.

a)
Old gasoline powered lawn mowers pollute more than cars. People should be forced to replace them with more efficient mowers.
❑ Agree   ❑ Disagree

b)
We will offer yoga classes one weeknight each week. Which night would you prefer?
❑ Thursday
❑ Monday
❑ Wednesday
❑ Tuesday
❑ Friday

c)
Should owners of hybrid vehicles be given an energy efficiency rebate from the government?
❑ Yes   ❑ No

d)
Speed kills! Speed limits on our highways should be reduced to 90 km/h.
❑ Agree   ❑ Disagree

B

6. An Internet survey asks people's opinions about a new software package. Which question is unbiased? How is the other question biased?

i)
> This software is used by some of the biggest names in business. If you have tried this software, what did you think about it?
>
> ___ Excellent ___ Good ___ Fair ___ Poor

ii)
> Have you tried this software? ___ Yes ___ No
> If you have tried this software, what did you think about it?
>
> ___ Excellent ___ Good ___ Fair ___ Poor

7. Choose one part of question 4. Explain how you decided whether to recommend a personal interview or written form.

8. For each part of question 5 that involved a biased question, explain how the question is biased and suggest how it could be improved.

9. An urban music radio station asks its listeners to e-mail or text an answer to this question:
"Do you think students in our city should wear school uniforms?"
95% of respondents say "No". The radio station announces that city schools should not introduce school uniforms since 95% of city residents are against the idea. Is the sample representative? If not, how could it be improved?

10. A newspaper columnist wants to find out what people think of a proposed by-law that would limit the height of fences they can build in their yards. He writes this survey question in his weekly column. Will the survey results be valid? Justify your answer. If you feel the survey is not valid, how could it be improved?

> Once again the government is trying to control us. This time they are interfering with our backyards.
> Do you agree with the proposed law to limit the height of a fence residents can put up in their yards to 2.44 m?
>
> ❑ No ❑ Yes

11. Describe how the town council in *Example 1* could conduct a valid survey to collect people's opinions about the arena.

12. The owner of a coffee shop plans to collect data to see if there is a relationship between the number of cups of coffee a person drinks per day and how happy he or she feels.

> Survey Plans
> - Set up a stand outside the coffee shop on Saturday morning.
> - Hand out a free cup of coffee to each person who participates in the survey.
> - Survey 100 people.
> - Ask these questions and record each person's answers.
>
>   1. How many cups of coffee do you drink in a day?
>
>   ____1-2 cups   ____3-4 cups   ____5-6 cups   ____7-8 cups
>
>   2. How happy are you? Rate yourself using a scale of 1 to 10, with 1 being very unhappy and 10 being very happy. ____

a) Why will the survey results be invalid?

b) What changes could you recommend to improve the survey?

13. Assessment Focus A salesperson for a new cellular phone service provider randomly selected the names of 20 small business owners from an association list of 500 members. She phoned each of them and asked them this question.

> Studies have shown that you can lose business if your cell phone service is not up-to-date. On a scale of 1 to 4, where 1 is low importance and 4 is high importance, how important is it to you to have up-to-date cell phone service for your business?

- All responded with a rating of 3 or 4.
- The salesperson prepared a business proposal for her boss, stating: "100% of the local small business owners I surveyed want up-to-date cell phone service. We should launch a large sales campaign very soon."

a) Are the survey results valid? Justify your answer.

b) What changes would you make to improve the survey?

14. Literacy in Math Create a concept map describing a good survey. Copy and complete this map, adding explanations of each component as well as any additional features you feel are missing, or create your own map.
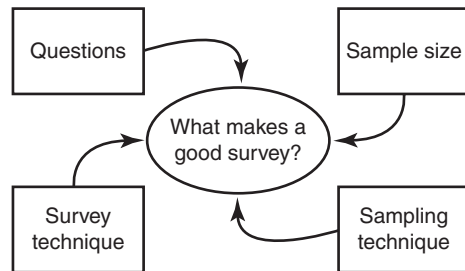
C **15.** An employee with a social service agency in Ottawa wants to know if there is a relationship between the number of years immigrants live in Canada and their sense of "fitting in." He selects the first 25 names from a list of immigrant families in the Ottawa area. He calls each family, states who he works for, and asks these questions:

> • How long have you lived in Canada?
> • On a scale of 1 to 4, with 4 being very well, how well are you fitting in to life in Canada?

Suggest one way to improve each of these components of the survey.
a) Sample size   **b)** Survey questions
c) Sampling technique   **d)** Survey technique

16. Full-time employees of the town of Sunderton belong to a union. Part-time employees who work on 10-month contracts from September to June do not belong to the union. The union held a vote for part-time employees to choose whether to join the union.
- Part-time employees received a letter in early July saying that the vote would be held in the town office on the second Wednesday in July.
- On the day of the vote, there was a power outage. A sign on the door of the town office stated the vote was postponed and told voters to check the office bulletin board for a new vote date.
- The vote was held four weeks later.
- 15% of the part-time employees voted. The majority voted to join the union.
- In September, the union announced that the part-time employees had voted to join the union.

a) In what ways is a vote similar to a survey? How is it different?
b) Based on what you know about planning a survey, describe at least two problems with the way the vote was held. How might these problems have affected the outcome?

In Your Own Words

Explain why it is important to consider respondents' privacy when planning a survey. Describe an example where privacy concerns could affect survey results.

A marketing research assistant helps to prepare and conduct surveys. The data collected are then analysed to help clients make decisions about the design, advertising, and pricing of their products or services.

Inquire

Materials

• computer with statistical software (optional)

For more information about two-variable data and correlations, look back at Chapter 3.

Collecting and Using Data to Answer a Question

In this lesson, you will collect and analyse data about your own research topic.

Work with a partner or in a small group.

1. **Choosing a topic**

   Choose a topic that relates two variables. Both variables must be measurable. Here are some examples.

   • The number of minutes per day students spend talking to friends on the phone and the number of minutes per day they spend on homework

   • The number of hours per week people spend using a computer and how physically fit they are on a scale of 1 to 10

   • A driver's age and how many kilometres he or she drives in a typical week

   • The number of minutes people spend listening to music per day and how happy they feel on a scale of 1 to 10

   Write your research question in this format:

   Is there a correlation between _____ and _____?

2. Selecting a sample

To help select a sample, answer these questions.

- What is the population for your investigation?
- What is the size of the population?
- What should be the size of your sample? Justify your choice.
- What sampling technique will you use to select your sample?
- How can you ensure your sample will be representative of the population? Explain.

3. Designing your questionnaire

Write your survey questions. Make sure they are unbiased.

As you begin writing, ask yourself these questions.

- Will you conduct your survey by interviewing respondents and recording their answers, or by handing out forms for respondents to complete privately?
- Do you want to include background information about age, gender, or grade on your survey?

Research question:
Is there a correlation between the amount of time students talk on the phone and the amount of time they spend doing homework?

Questionnaire:
Phone Habits and Homework

1. Are you male or female?
   ____M   ____F

2. About how much time each day do you talk on the phone to your friends?
   _____ min

3. About how much time each day do you spend on your homework?
   _____ min

4. Collecting your data

Collect your data using the sampling technique you described in question 2.

5. **Organizing your data**

Organize your data in a table.

You could use a computer with spreadsheet software or a graphing calculator if it is available.

Survey data:

	A	B	C
		Phone	Homework
1	Gender	(min/day)	(min/day)
2	M	5	120
3	F	10	80
4	M	10	90
5	M	14	85
6	F	14	80
7	M	20	75
8	F	20	80
9	F	20	85
10	M	25	90

The background information is shown in the first column.

The answers to the survey questions are in the second and third columns.

6. **Drawing a line of best fit**
   - Create a scatter plot.
   - Does there appear to be a correlation? Justify your answer.
   - If there is a correlation, describe it, and construct a line of best fit. Determine the equation of the line.

Create a scatter plot:
- Independent variable: phone time
- Dependent variable: homework time

Is There a Correlation Between Time Spent on the Phone and Time Spent Doing Homework?

$y = -0.7x + 95.8$
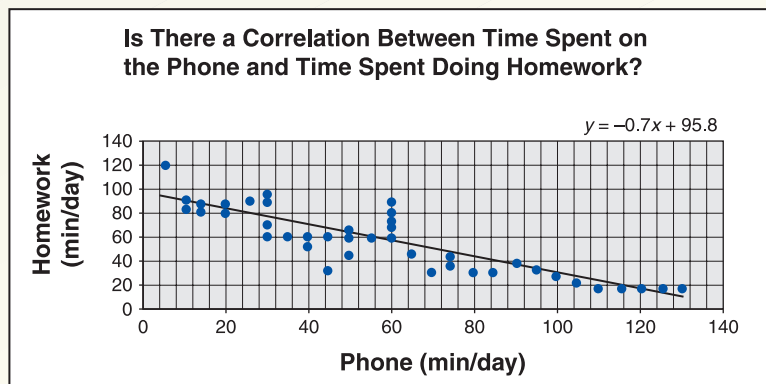
The data lie along a line with negative slope, so there is a strong, negative correlation.

220 CHAPTER 4: Statistical Literacy

7. **Analysing your scatter plot and drawing a conclusion**

■ Write a conclusion that summarizes your results and answers your research question.

Conclusion:

As time students spend on the phone each day increases, the time they spend doing homework decreases.
There is a strong, negative correlation.

While we know it is not the only variable affecting time spent on homework, we believe that spending too much time on the phone will cause students to neglect their homework.

8. **Extending your investigation**

■ If you collected additional data for each subject, you can extend your investigation.

■ If you think there is a cause-and-effect relationship, you might want to do additional research about other variables that might affect the variables you have been examining.

Possibilities for further research:

• Create, then compare separate scatter plots for males and females using the background information already collected.

• Go back and ask students about other factors that may have affected their homework time, such as hours worked at part-time jobs.

Reflect

■ How could you improve your investigation if you were to repeat it? Justify your answer.

■ What related questions could you investigate? Would you have to collect more data? If so, describe your next steps for collecting that data.

4.1

1. Determine the quartiles of these marks.

73	45	79	88	64	70
96	68	72	94	56	81

2. Explain the meaning of each statistic.
  a) The median salary at an advertising firm is $85 000.
  b) In a blind taste test, 95% preferred the new cereal over all the others.
  c) A local athlete scores in the top quartile in a province-wide fitness challenge.
  d) Last year, Ontario's electricity consumption was 11 996 kWh per capita.
  e) 1 in 4 Canadian adolescents is considered overweight.

4.2

3. Four candidates ran for mayor in a town.
  • 35% of the town voted.
  • The distribution of votes was:
    Elson: 30%   Singh: 25%
    Jinah: 23%   Watkins: 22%
  The headline in the paper the next day read: "The People Have Spoken: Elson To Be Mayor"
  Explain why the headline is not appropriate.

4.3

4. For each topic, would you conduct a survey using personal interviews or written forms? Justify your answers.
  a) Age and exercise frequency
  b) Adult literacy
  c) Income and education level
  d) Favourite leisure activities and time spent on them

5. A politician wants to know if seniors in her town would use a seniors' centre. She designs a questionnaire and has her assistant call every 10th phone number in the local phonebook.
  a) What is the population for the survey?
  b) In what way is the sample not representative of the population? How should the sample be changed?

6. A newspaper had the following headline: "People Against New Breed-Specific Dog Ban." The reporter who wrote the story tells you he visited a leash free park and talked to 100 dog owners. His question:

> A new by-law bans owning a Pit Bull, Staffordshire Terrier, or Bull Terrier dog breed, or any hybrid or similar crossbreed. Do you support the by-law?

  a) Why are the survey results invalid?
  b) What changes would you make to improve the survey?

4.4

7. a) Write unbiased survey questions you could use to determine if there is a correlation between the amount of time high-school students spend e-mailing each other and the amount of time they spend talking to each other on the phone.
  b) Describe how you would select your sample and conduct your survey.

4.5

Statistics can both lead and mislead. We are often presented with conclusions that are based on statistical analysis. However, it is our responsibility to determine whether the conclusions are valid.
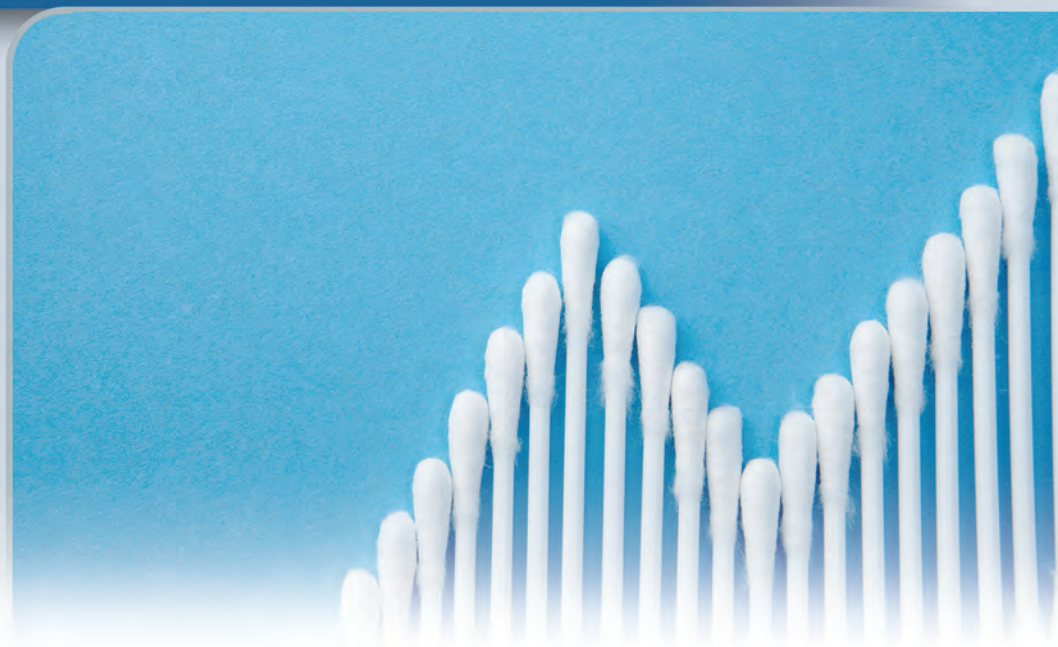
Investigate

Assessing the Validity of Conclusions

Work with a partner.

Refer to Case Study: *Women in the Workforce.*
■ Do you think the conclusion is valid or invalid?
  Include answers to these questions in your response.
  • Are the data reliable?
  • Is there a possible bias in the person analysing the data?
  • Is the sample size reasonable?
  • Is the correlation strong?
  • Is there any evidence to support a cause-and-effect relationship?
  • Does the graph represent the data appropriately?

Reflect

■ List three questions you should ask yourself before you accept a conclusion drawn from statistical data. Choose one question and describe how the answer would affect your decision about whether a conclusion is valid.
■ Why might someone want to mislead others using statistics?

Case Study
Women in the Workforce

The members of a high school debating club are preparing for a debate.

Debate position:
It would be better for families if women stayed at home rather than joined the workforce.

Preparation:
- Analyse data about women in the workforce and single-parent families in Canada.
- Use data collected by Statistics Canada.

Data:

Year	Number of women in workforce (thousands)	Number of single-parent families
2000	6790.4	1 317 760
2001	6910.3	1 406 390
2002	7126.0	1 404 250
2003	7324.2	1 451 150
2004	7466.4	1 444 150

Women in the Workforce and Single-Parent Families

$r = 0.8642$
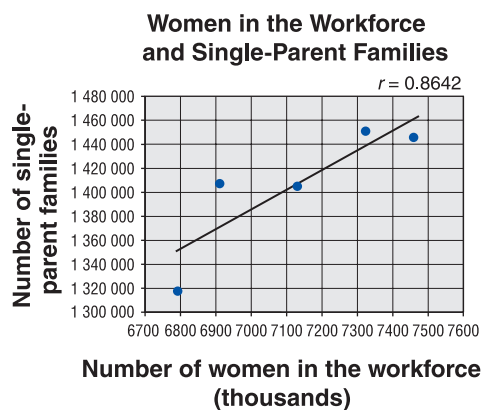
Conclusion:
These data support our position.
There is a strong positive correlation between the two variables.

As the number of women in the workforce increases, the number of single-parent families increases. If we want to reduce the number of single-parent families in Canada, women should stay at home and not go out to work.

Assessing statistical data

A **valid conclusion** is one that is supported by unbiased data that has been interpreted appropriately.

When you read a conclusion someone has made based on statistics, you must decide whether the conclusion is valid. To do this, ask yourself:
- Is there any bias in the data collection—in the way the sample was selected, the questions were phrased, or the survey was conducted?
- If the data involve measurements, were they accurate?
- Are any graphs drawn accurately or do they mislead the viewer?

Example 1

Assessing Graphs

The graphs in each pair show the same data. Choose the graph that displays the data more accurately. Justify your choice.

a) Canada's population by age according to the 2001 census

i) **Ages of Canadians, 2001 Census**

85 and older 6%
65 to 84 11%
0 to 14 19%
45 to 64 20%
15 to 24 20%
25 to 44 24%

ii) **Ages of Canadians, 2001 Census**

85 and older 6%
0 to 14 19%
65 to 84 11%
45 to 64 20%
15 to 24 20%
25 to 44 24%
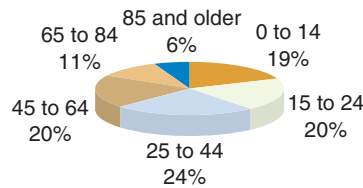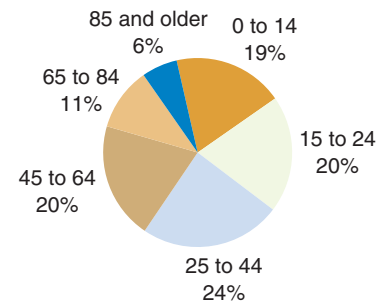
b) A company's profits over a 5-year period

i) **Company Profits**
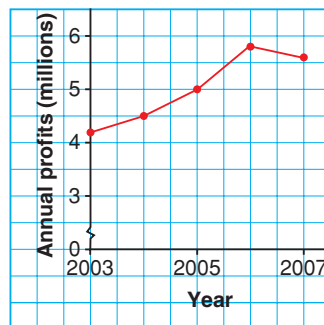
ii) **Company Profits**
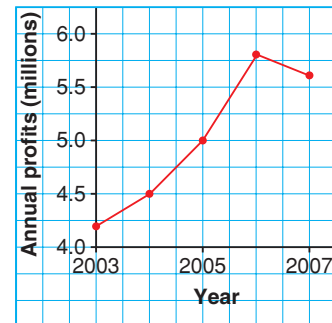
Solution

a) The graph in part ii displays the data more accurately.
   - The graph in part i is a three-dimensional graph.
   - Using three dimensions makes some pieces of the graph appear larger than they should compared to the other pieces.

b) The graph in part i displays the data more accurately.
   - Part of the vertical axis has been omitted in each graph. This makes the differences between the values more striking, particularly in the second graph where more has been deleted.
   - The graph in part i uses a symbol to alert the viewer that part of the axis is missing, while the other graph does not.

Example 2 | Assessing How Data Were Collected and Graphed

Four Grade 9 students collected data on school lunch preferences.

Favourite Lunch Meals
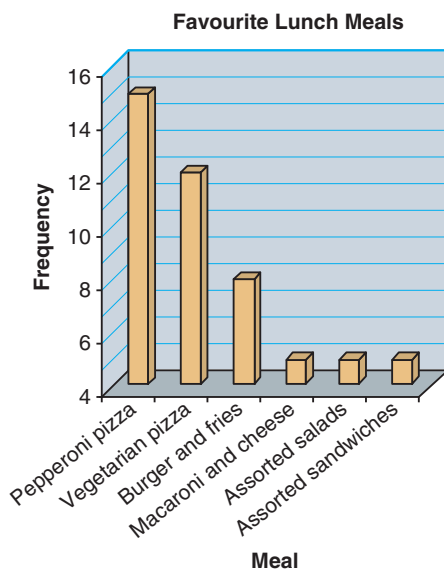
They concluded:

> We asked students to tell us their favourite lunch meals and displayed the results in this bar graph. We conclude that the school cafeteria should serve more pizza since it is clearly the favourite lunch of students.

Is this conclusion valid?

Solution

■ **Was the sample size appropriate?**
By adding the frequencies, you can see that 50 students were surveyed. Depending on the size of the school, this may not be enough data.

■ **Was the sample representative?**
You cannot tell from the information given. Perhaps the researchers surveyed only their friends or only Grade 9 students. You need more information about the sampling technique to judge this.

■ **Was the survey question biased?**
It appears that students were simply asked their favourite lunch meal. This is an unbiased question since it does not try to influence the answer.

■ **How was the survey conducted?**
It appears the survey was conducted orally. This could bias the results because some students might be self-conscious about their eating habits.

■ **Is the graph constructed accurately?**
In general, a three-dimensional bar graph tends to distort the relative quantities being displayed.
Starting the vertical axis at 4 also distorts the relative quantities.

The conclusion is *not* valid, although it may be true. You need more information about the sample selection before recommending any change in lunch choices at the cafeteria.

| Example 3 | **Assessing Assumptions about Cause and Effect** |

A group of Grade 12 students performed a linear regression on data they collected from Statistics Canada about the number of seniors and the number of weapons crimes in Canada.

> The *r*-value is a measure of how strong a correlation is. The closer to 1 or –1, the stronger the correlation.
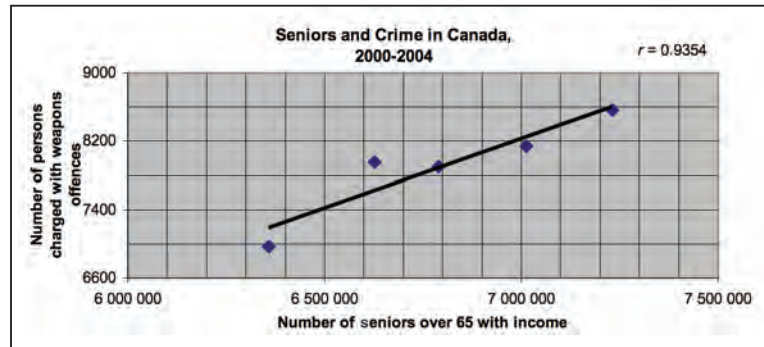
They concluded:

> There is a strong positive correlation between the two variables. As the number of seniors increases, weapons charges increase. Therefore, criminals in Canada are becoming bolder because of our ageing population.

Is this conclusion valid?

Solution

To assess the validity of the conclusion, ask yourself these questions.

■ **Was there bias in the data collection?**
The students gathered the data from Statistics Canada, which is a reliable source of data. However, only 5 years of data were included.

■ **Is the graph constructed accurately?**
Yes. Because the numbers are large, it would be impractical to start the scale on the vertical axis at zero.

■ **Is the correlation strong?**
Yes. The points are close to the regression line.

> The *r*-value 0.9354 is very close to 1. So, the correlation is strong.

■ **Does the analysis support a cause-and-effect relationship?**
Not necessarily. Both variables may be increasing because the population of Canada is increasing.

The conclusion is *not* valid, although it may be true.

You need more data and you need to eliminate any other variables before you can fairly draw the conclusion made by these students.

1. The graphs in each pair show the same data.
Choose the graph that displays the data more accurately.

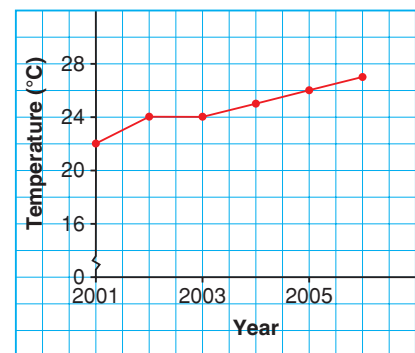**a)** Favourite cola drinks of 95 shoppers in a city mall

i) **Favourite Cola Drinks**

Extra caffeine 3%

Sugar- and caffeine-free 21%

Sugar-free 35%

Regular 25%

Caffeine-free 16%

ii) **Favourite Cola Drinks**

35%
30%
25%
20%
15%
10%
5%
0%

Sugar-free   Caffeine-free   Regular   Sugar- and caffeine-free   Extra caffeine
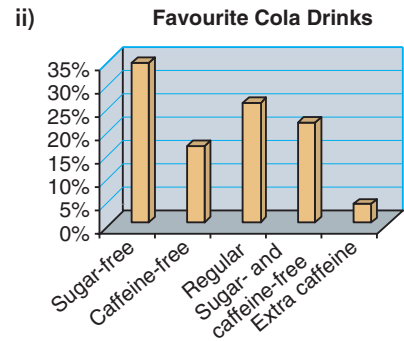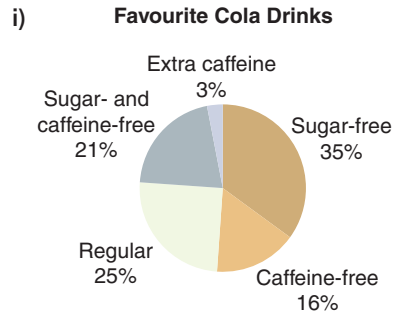
b) Temperature change over time

i) **Average Daily High Temperatures**

Temperature (°C)

28
26
24
22
20

2001   2003   2005
Year

ii) **Average Daily High Temperatures**

Temperature (°C)

28
24
20
16
0

2001   2003   2005
Year
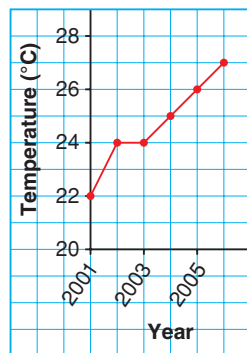
2. For each part in question 1, describe both graphs. What features misrepresent the data in the graph that represents the data *less* accurately?

3. For each survey, who do you think would be less biased in collecting data?
  a) A survey on recycling rates for plastic bottles
    i) A bottled water manufacturer   **ii)** A town's public works department
  b) A survey on people's opinions about health care
    i) A college student doing a project **ii)** A group of doctors

4. A soft-drink company wants to test consumers' reaction to a new soft drink. Which group should collect the data? Why?
  i) The sales and marketing department of the company
  ii) An outside agency specializing in statistical surveys

5. Decide whether you would expect there to be a correlation between each pair of variables.
   a) The numbers of students and teachers in a school
   b) The number of years a person has worked for a company and the number of vacation days he or she receives each year
   c) The population of a town and the amount of precipitation the town receives each year
   d) A person's height and her or his mark in mathematics

6. For each part of question 6 in which you felt there would be a correlation, describe the correlation. Explain whether you think there might be a cause-and-effect relationship between the variables, and why.

7. **Literacy in Math** What additional information would you need before deciding whether each statistical analysis is valid?
   a) The host of a TV infomercial demonstrates a cleaning product. Then a man in a laboratory coat says, "Studies have shown that this product eliminates more bacteria from household surfaces than the leading brands."
   b) You research athletic shoes on the Internet before you purchase a new pair. On one site, a pop-up advertisement displays results from an online survey in a bar graph. The graph shows that people prefer shoes made by Robur to those made by several other brands.

8. Which of the following statements best describes the information in the scatter plot? Justify your choice.
   i) There is no correlation between the number of advertisements shown per month and the monthly cereal sales.
   ii) There is a strong positive correlation between the number of advertisements shown per month and monthly cereal sales.
   iii) As the number of advertisements shown per month increases, cereal sales increase.

Breakfast Cereal Advertisement Effectiveness

$r = 0.8910$

9. A reporter from a TV news show asks 5 people on the street this question: "In light of the many recent home invasions, do you think police are doing all they can to keep us safe?"

Four of those interviewed say the police are not keeping us safe.

On the news that evening, the reporter announces, "4 out of 5 citizens are worried about personal safety," and then shows the interviews. What is wrong with this statistical analysis?

10. The prom committee researched four possible locations for the prom.
- They wanted the graduating students to make the final selection.
- They provided homeroom teachers with copies of a questionnaire to hand out to all graduating students.
- They received completed questionnaires from 85% of the graduating students.

The questionnaire, a graph of the results, and their conclusion are shown. Is the conclusion valid? Justify your answer.

Where would you like to have the prom this year? Please check one.
❑ Crystal Fountain
❑ Empire Hotel
❑ Winston's
❑ Palace Ballroom

Prom Choices

Palace Ballroom 11%
Crystal Fountain 15%
Empire Hotel 19%
Winston's 55%
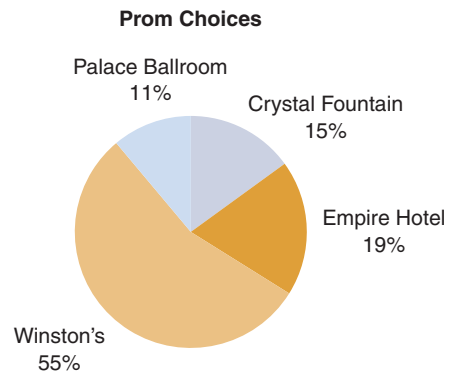
Conclusion:
Most students prefer Winston's, so the prom will be held at Winston's this year.

11. The headline in a newspaper reads:

Hockey Contributes to Increase in Crime

- The article describes a study that compared the number of young people playing on minor hockey teams and the number of arrests over a 5-year period.
- Both variables increased over time.
- A linear regression performed on the data had an r-value of 0.82.

Describe the errors that make this statistical analysis questionable.

12. Assessment Focus A stress management clinic in a city of 250 000 people wanted to find out whether there is a relationship between the number of hours worked in a week and job stress level.
  - They hired an outside agency to collect the data.
  - The agency randomly selected 2500 adults who work in the city.
  - They asked people to tell them the number of hours they work per week and to rate their level of job stress from 1 to 10.

Hours Worked and Stress Level

$r = 0.9176$

Number of hours worked per week
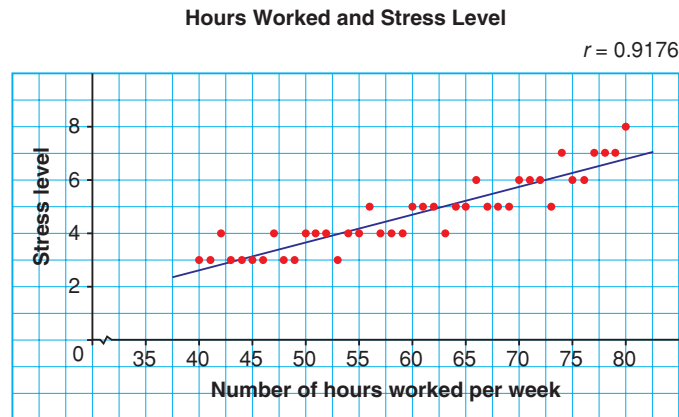
  - The clinic concluded:

Is this conclusion valid?
Justify your answer.

> There is a strong positive correlation between the number of hours people work per week and their stress level on the job. We believe that an increase in working hours is likely to cause an increase in stress level.

C

13. A city plans to widen a road from two lanes to four lanes. Some residents of nearby neighbourhoods are concerned that traffic noise will increase.
They want the city to construct noise barrier walls when they widen the road.
They hire you to do a study. You outline what you will need to do.
  - Design and conduct a survey.
  - Graph and analyse the survey results.
  - Search for data from similar situations in other neighbourhoods or cities.
  - Hire a consultant to test current noise levels and develop projections for the future.
  - Make a recommendation.
  a) Describe how you could design the study so that the recommendation is to construct a wall.
  b) Describe how you could design the study so it is unbiased.

In Your Own Words

Explain how someone could be misled by a statistical analysis.
Include an example in your explanation.

4.6 Understanding Indices

Prices for everyday items such as gasoline change over time or depending on the geographic location.

Investigate

Creating a Gasoline Price Index

Materials
• scientific calculator

Work with a partner.

- Graph the data in this table. Describe the change in gasoline prices over time.
- Create a new table from this one. Express each price as a percent of the price in January 2006. Graph the data in the table.
- Compare the graphs.

Gasoline Prices (¢/L)		
	January	95.0
	March	93.3
2006	May	104.6
	July	109.7
	September	89.7
	November	86.5
	January	87.1
2007	March	102.4
	May	111.5

Reflect

- For each graph, identify the information it shows that the other graph does not.
- Use the graph of prices as a percent of the price in January 2006. Create a question about gasoline price increase or decrease that can be answered using the graph. Answer the question. Exchange questions with your partner. Check your partner's solution.

Price indices

Price indices help citizens, businesses, and industries follow and predict trends in prices. A price index describes the price of an item compared to a **base value** measured at a particular time or in a particular place.

Statistics Canada tracks price changes using several different indices. The most important is the Consumer Price Index (CPI).

To determine the CPI, Statistics Canada collects thousands of price quotations from across the country for a basket of about 600 popular consumer goods and services. These items range from French fries and bus fares to tuition and Internet service.

Example 1

Reading the Consumer Price Index

Use this CPI graph to answer these questions.

Consumer Price Index (CPI)
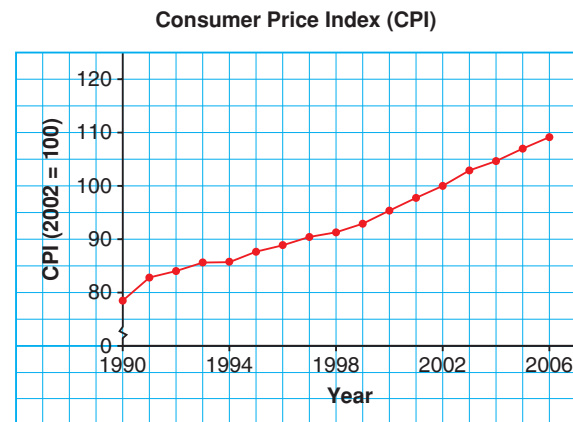
a) What is the base year for the CPI?
b) In what year was the cost of the basket of goods about 90% of the base cost?
c) What was the CPI in 1990? What does this mean?
d) Describe the change in the CPI from 1990 to 1991. What do you notice about the line segment representing this period?
e) Describe the overall trend in the CPI and its significance.

Solution

a) Look for the year with a CPI of 100.
   The base year is 2002.
b) When the cost of a basket of goods is 90% of the cost in the base
   year, the CPI will be 90. The CPI was 90 in 1997.
c) In 1990, the CPI was about 78.
   So, prices in 1990 were about 78% of the prices in 2002.
d) The CPI increased from about 78 to about 83. This is an increase
   of 5% of the base value in one year. This is the greatest one-year
   increase. The line segment representing this increase is the steepest
   on the graph.
e) The CPI increases over the years shown.
   So, Canadians spend more money each year to buy the same basket
   of products and services.

Example 2   **Solving Problems Using an Index**

Use the graph in *Example 1*.
a) Calculate the average annual rate of inflation from 1990 to 2006.
b) Use your answer to part a to predict the CPI for 2010. Justify your
   prediction.

Solution

a) From 1990 to 2006, the CPI rose from about 78 to about 109.
   So, the CPI increased by $109 - 78$, or 31.
   This represents a 31% increase in prices in 16 years.
   $31\% \div 16 \doteq 1.9\%$
   Therefore, the average annual rate of inflation during this time was
   about 1.9%.
b) From 2006 to 2010 is a 4-year period. If the trend observed in part a
   continues, you would expect inflation to increase by 1.9% each year.
   $4 \times 1.9\% = 7.6\%$
   Add this increase to the CPI for 2006: $110 + 7.6 = 117.6$
   So, the CPI for 2010 would be about 118.
   It is reasonable to assume that the trend seen in the 14 years from
   1990 to 2006 would continue for the next 4 years.

Other price indices

Some price indices do not show a change over time. Instead, they compare prices among different geographical regions.

Example 3

Using an Index to Compare Cities

The 2006 UBS *Prices and Earnings* report includes a comparison of clothing prices in 71 cities. The base price is the price in New York.

a) Which cities in this table have index values less than 100? What does this tell you?

b) How do clothing prices in Zurich and Hong Kong compare to clothing prices in New York?

City	Clothing Price Index (New York = 100)
Zurich	115.6
Oslo	114.4
Dublin	97.5
New York	100.0
Toronto	73.8
Tokyo	148.1
Rome	87.5
Hong Kong	75.0
Delhi	43.8

Solution

a) Dublin, Toronto, Rome, Hong Kong, and Delhi have index values less than 100. This means clothing prices in these cities are cheaper than in New York.

b) Zurich's index is 115.6.

$115.6 - 100 = 15.6$

Clothing prices are 15.6% higher in Zurich than in New York.

For every $100 spent on clothing in New York, you would spend $115.60 spent in Zurich for similar items.

Hong Kong's index is 75.0.

$75.0 - 100 = -25.0$

Clothing prices are 25% lower in Hong Kong than in New York.

For every $100 spent on clothing in New York, you would spend only $75 in Hong Kong for similar items.
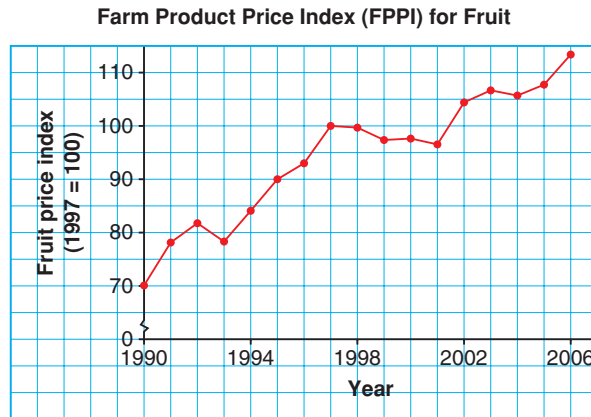

Other types of indices

Some indices do not use a base value. Instead, they use formulas to produce a number that describes something about a person, place, or thing. These numbers can then be compared.

1. a) What is this price index measuring?

b) What is the base year for the index?

c) Estimate the index value for each year.

  i) 1994   **ii)** 2002

Farm Product Price Index (FPPI) for Fruit

2. For each price, calculate the percent price increase from a base value of $124. Round each answer to the nearest percent.

a) $186   **b)** $155   **c)** $248   **d)** $131

3. For each price, calculate the percent price decrease from a base value of $124. Round each answer to the nearest percent.

a) $92   **b)** $62   **c)** $115   **d)** $25

4. Order these top 10 happiest countries from most to least happy.

The world map of happiness is based on a **subjective well-being** (SWB) index. The greater the index value, the happier the population.

Country	SWB Index
Austria	260
The Bahamas	257
Bhutan	253
Brunei	253
Canada	253
Denmark	273
Finland	257
Iceland	260
Sweden	257
Switzerland	273

A Global Projection of Subjective Well-being:
The First Published Map of World Happiness

Map created by Adrian White, Analytic Social Psychologist, University of Leicester (2006)
Map and further analysis incorporates data published by UNESCO, the WHO, the New Economics Foundation, the Veenhoven Database, the Latinbarometer, the Afrobarometer, the CIA, and the UN Human Development Report.

Happy ——— Average ——— Unhappy

Cartographic Unit • University of Leicester

5. Use the Consumer Price Index graph in *Example 1* to answer
these questions.
   a) What was the CPI in January 1996? What does this value mean?
   b) What was the CPI in January 2001? What does this value mean?
   c) Describe the change in the CPI from January 1996 to January 2001.
   d) Calculate the average annual inflation rate from January 1996 to
   January 2001.

B

6. Use the graph of FPPI for fruit in question 1:
   a) Describe the general trend in the graph.
   b) Explain what this trend means.

7. Meteorologists and forestry technicians use the Canadian Forest Fire
Weather Index (FWI) to predict the intensity of potential forest fires.
Severe fires have FWI values greater than 30.
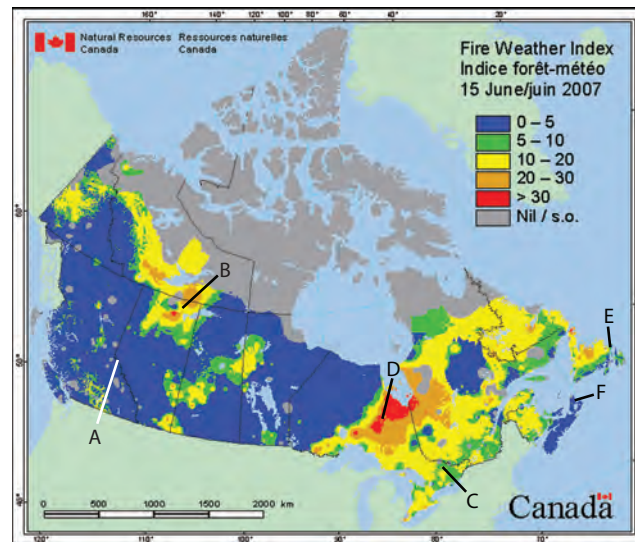
For each pair of forested regions, identify the location that likely had more
intense forest fires on June 15, 2007.
   a) Jasper National Park (A) or Wood Buffalo National Park (B)
   b) Algonquin Provincial Park (C) or Cochrane District, Ontario (D)
   c) Terra Nova National Park (E) or Cape Breton Highlands National
   Park (F)

8. **Literacy in Math** Choose a person that you could talk to about prices from
long ago. Ask what the person remembers about salaries and prices of a few
common items, such as a bottle of pop, a haircut, or a new car. Compare the
salaries and prices then and now. Do you think things are more affordable
today or in the past? Explain. How does this relate to the idea of an index?

Use this graph to answer questions 9, 10, and 11. It shows an index for government spending on education in Canada from 1986 to 2003.

Education Price Index (EPI)
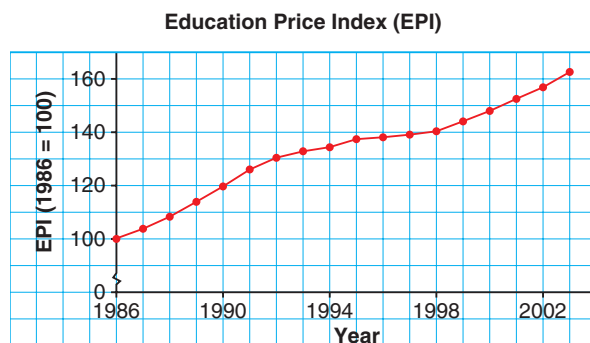
9. a) What is the base year for this index? Explain how you know.
   b) Estimate the EPI for each year.
      i) 1992   **ii)** 1998   **iii)** 2003
   c) By what percent did spending on education rise during each time period?
      i) Base year to 1992   **ii)** 1992 to 1998   **iii)** 1998 to 2003
   d) Compare your answers to part c. Which period had the greatest increase? Which period had the least increase? How does this relate to the line segments on the graph? Explain your thinking.

10. a) Calculate the overall change in the EPI from 1986 to 2003. What was the average rate of change per year for this 17-year period?
   b) Predict the EPI for 2010 if the rate of change you determined in part a continues. Explain your method.

11. Assessment Focus

This graph shows an index for government spending on instructional supplies in Canada from 1986 to 2003. Instructional supplies are part of the education price in Canada.

Education Price Index (EPI) for Instructional Supplies
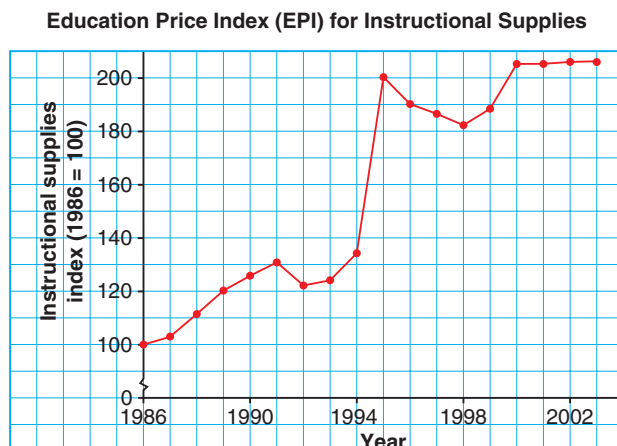
   a) How is this graph the same as the education price index graph?
   b) How is this graph different from the education price index graph?
   c) Calculate the overall change in the instructional supplies index from 1986 to 2003. Then determine the average annual rate of change for the 17-year period.
   d) Compare your answer from part c with your answer in question 10 part a. Give a possible explanation for any differences.

12. The 2006 UBS *Prices and Earnings*
report compares the cost of a basket
of food in 71 cities. The base cost is
the cost in New York.
Data for 10 cities is given.

a) Which cities have index values greater
than 100? What does this tell you
about food prices in these cities?

b) How do food prices in Oslo and
Delhi compare to food prices in
New York?

c) Name a pair of cities that have
similar food prices. Justify your
answer.

d) Write a question someone could answer using these data.
Answer the question.

City	Food Price Index (New York = 100)
Zurich	115.6
Oslo	112.1
Dublin	86.6
New York	100.0
Copenhagen	99.5
Toronto	80.8
Tokyo	130.3
Rome	87.8
Hong Kong	86.6
Delhi	35.1

C **13.** Use the data from the food index in question 12.

a) Recalculate the index values using Toronto as the base value.
That is, Toronto = 100. Explain your method.

b) Which cities have new index values less than 100? What does this mean?

c) Suppose you were to recalculate the original index using Tokyo as the
base value. How would the index values change? Justify your answer.

14. The S&P/TSX Composite Index compares the current value of certain stocks
traded at the Toronto Stock Exchange relative to their value at a previous
time. Research this index. Prepare a brief report that includes:
- A description of what the index measures
- The base value and the year in which the base value is fixed
- A line graph showing the changes in the index over time using at least
12 pieces of data

In Your Own Words

Explain the difference between a graph showing *Wage Rates* and a
graph showing a *Wage Rate Index*.

Statistics Canada provides many data sets involving indices, including data about housing prices across Canada. In this lesson, you will research indices on the Statistics Canada Web site using E-STAT.

Inquire

Researching Indices

Materials
- computer with Internet access
- E-STAT user name and password

Work with a partner.

■ Go to the Statistics Canada Web site.
Click **English**.
Select **Learning Resources** from the menu on the left.
Click on **E-STAT** in the golden box on the right.
Click on **Accept and enter**.
If you are working from home, you will need to get a user name and password from your teacher.

■ The E-STAT table of contents will be displayed.

1. a) List the categories of data sets.
   b) List the topics under the category *Land and Resources*.
   c) Which category contains the topic *Prices and price indexes*?

■ Select **Prices and price indexes**, and then **Construction price indexes**.

■ You are going to explore the New Housing Price Index.

Click on table **327-0005**.

This table contains data sets for many Canadian cities.

- Scroll through the list of cities in the box labelled *Geography*.
- Hold down the **Ctrl** key and select:

  Montreal, Quebec [**24462**]

  Toronto and Oshawa, Ontario [**35535, 35532**]

  Vancouver, British Columbia [**59933**]
- In the next box, click on **Total** (**house and land**).
- In the next box, click on **1997 = 100**.
- Set the *Reference period* to begin in **Jan 1997** and end in **Jan 2007**.
- Click on **Retrieve as individual Time Series.**
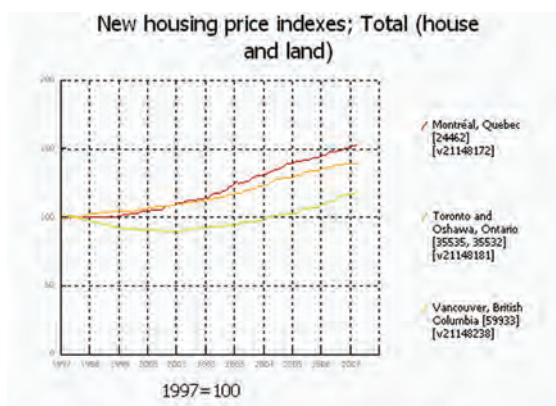
  A summary of the data you have requested is shown.
- Select the output format. Under *Screen output – table*: and *HTML, table*: select **Time as Rows**.
- Click on **Retrieve Now**.

	Statistics Canada	Statistique Canada			Canada
	Français	Contact us	Help	Search	Canada site
	Site map	About us	Privacy	Accessibility	My account

STATISTICS CANADA
CANADA'S NATIONAL STATISTICAL AGENCY

HOME
E-STAT
About E-STAT
What's new in E-STAT
Table of contents
User guides
Search CANSIM
Help with this page
Search Censuses
Search map 2001
Help/Frequently asked questions
Contact E-STAT
Learning resources

Table 327-0005
New housing price indexes

Monthly	v21148172 - Montréal, Quebec [24462]: Total (house and land); 1997=100 (index)	v21148181 - Toronto and Oshawa, Ontario [35535, 35532]: Total (house and land); 1997=100 (index)	v21148238 - Vancouver, British Columbia [59933]: Total (house and land); 1997=100 (index)
Jan 1997	100.4	98.3	101.3
Feb 1997	100.4	98.6	101.4
Mar 1997	100.4	98.7	101.1
Apr 1997	100.1	99.4	101.3
May 1997	99.4	99.8	101.3
Jun 1997	99.7	99.8	100.7
Jul 1997	99.7	100.3	100.1
Aug 1997	100.0	100.2	99.5
Sep 1997	100.1	100.4	99.2
Oct 1997	100.1	100.5	98.2

2. This data set is part of the New Housing Price Index (NHPI).

   a) In which year was the base value set?

   b) How often is the NHPI calculated?

   c) For each city, how many pieces of data are in this set?

   d) Scroll through the data set for one city. What do you notice about the index values?

■ To view the data as a graph, click **Back** on the tool bar.

   • Under *Screen output – Graph*, select **Line graph**.
Then click on **Retrieve now**.

   • Click on **Modify Graphic** if you wish to revise the title for your graph or add gridlines. When you are finished, click **Replot**.

   • If you wish to print the graph, click on **<u>Printer-friendly format</u>** at the bottom of the page.

New housing price indexes; Total (house and land)

Montréal, Quebec [24462] [v21148172]

Toronto and Oshawa, Ontario [35535, 35532] [v21148181]

Vancouver, British Columbia [59933] [v21148238]

1997 = 100

Source: Statistics Canada, CANSIM table 327-0005.

Modify Graphic

Printer-friendly format

3. Use the graph to answer these questions.

   a) How are the lines the same?

   b) How are the lines different?

   c) What most surprises you about the information in the graph?

4. Use the table to answer these questions.

   a) For each city, calculate the percent rate of increase in new housing prices from Jan 1997 to Jan 2007.

   b) Suppose you bought a new house in Toronto in 1997 for $100 000. What would be the price for an equivalent new house in 2007?

c) Repeat part b for Montreal and Vancouver.

d) For each city, suppose the percent rate of increase you calculated in part a remains the same for the next 10-year period. What would be the price of an equivalent new house in each city in 2017?

5. Suppose you were offered the same job in all three cities.

a) If you were to make your decision based on the NHPI, which city would you select? Justify your answer using data from your graph.

b) In addition to the NHPI, what other information on housing would you need to know to help you make your decision?

6. Return to the E-STAT table of contents.
Select another price index and conduct similar research to investigate the index for three cities in Canada.

7. List three things you found interesting about the E-STAT data sets.

Reflect

- What features made the Statistics Canada Web site and E-STAT data sets easy to use? Explain.
- What did you find challenging about using the E-STAT data?
- Describe a real-life question or problem that one of the data sets from E-STAT could help solve.

Data management skills are important in industries such as manufacturing, health sciences, hospitality services, financial services, and resource management. As a result, many college programs include courses that involve organizing, displaying, and analysing data.

Inquire

Materials

- computer with Internet access

Researching College Programs and Occupations

Work with a partner.

■ Record the data management skills you have learned in this course. If you need some hints, look through the *Chapter Reviews* in Chapter 3 and Chapter 4.

> Extrapolate from a graph
>
> Assess validity of a survey
>
> Interpret statistical information

■ On the Internet, search for Web sites of Ontario colleges. Investigate the post-secondary programs offered at a few colleges. Look for courses that involve data management.

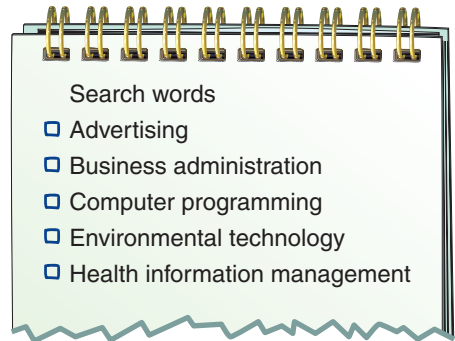You may be able to search for courses using keywords.

Search for Courses

Course Code or Keyword: statistics

[Search for Courses] [Clear Search]

Search Results

Course	Title
BIOT2030	Applied Statistics for Biotechnology
ECOS2029	Landscape Ecology
ENVR1337	Statistics/Quality Control/Data Interpretation
GISC9308	Spatial Analysis and Spatial Statistics
LAW1210	Criminology
MATH1580	Statistics
MATH1760	Introduction to Statistics
MATH1780	Applied Statistics and Research Methods
MATH3010	Quantitative Methods
MATH3020	Mathematical Modelling and Applied Statistics
STAT2060	Statistics for Business Decisions

Search words
- ☐ "Data management"
- ☐ Statistics
- ☐ Data
- ☐ Analysis
- ☐ Quantitative
- ☐ Research
- ☐ Measurement
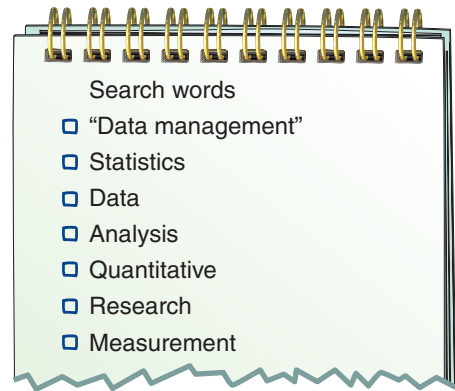
Or, you may have to search for programs likely to involve data management, then look through the course listings for those programs.

Search words
- ☐ Advertising
- ☐ Business administration
- ☐ Computer programming
- ☐ Environmental technology
- ☐ Health information management

- ■ Select three courses that interest you and describe each of them. Include the following information in your description.
  - • Course name
  - • College that offers the course
  - • Programs that require or recommend the course
  - • Brief description of the data management skills used in the course

■ On the Internet, search for occupations that require data management skills. You might try job search Web sites, or career centres on college Web sites.

> Some Occupations Requiring Data Management Skills
>
> • **Business** – legal assistant, marketing consultant, forecasting and inventory planner, customs broker
> • **Sciences** – forensic technologist, automation technician, network security analyst, laboratory assistant
> • **Environment** – earth sciences technician, conservation officer, forestry technician, eco-tourism business operator
> • **Government & Community Services** – property assessor, social service worker, archivist, international trade specialist
> • **Health & Medicine** – health information manager, registered practical nurse, lab technician

■ Select three occupations that interest you and describe each of them. Include the following information in your description.
  • Job title
  • A description of the data management skills used in the job
  • Educational background required for the job
  • Typical salary range for the job
  • Whether there is a demand for people who can perform this job and future prospects for job availability
  • Source of your information

■ How might you use data management skills in the next few years, either at school or in everyday life?

Reflect

■ Describe one thing that surprised you or might surprise someone else about the information you found.
■ What would you say it means to be "statistically literate?"

GAME

Concept Clues

Materials

- 5 *Concept* cards (see *Preparation* section below)
- stopwatch, watch with second hand, or 1-min sand timer

You can use your notes, textbook, and any supplementary material in the classroom to help you create your *Concept* card.

Preparation

Each student should create at least one *Concept* card. Two sample cards are shown. Give your completed card(s) to the teacher.

Concept name →

Four words or short phrases that relate to or describe the concept

Index

Consumer Price Index

Inflation rate

Often measures change over time

Body Mass Index

Sources of Bias

Sample is not representative

Question unfairly restricts choices

Question wording influences response

Personal interview used to collect sensitive information

How to Play

- Play in groups of 3 to 5.
- Each group should get five *Concept* cards from the teacher. Shuffle these and place them face down in a pile.
- The first player reads aloud the concept on a *Concept* card and starts timing. The other group members try to identify each of the other four words or phrases on the card. The player with the card may give clues, but cannot say any of the words or phrases.
- Play ends after 1 min. The group gets 1 point for each word or phrase identified. If all four were identified, the group earns a bonus point.
- After all groups finish their cards, the group with the most points wins.

Reflect

- Explain whether creating a *Concept* card helped you when you played the game.

Study Guide

Assessing the Validity of a Survey

When you read or hear about survey results, ask questions to help you decide whether the survey is valid.

- What is the sample size?
- Is the sample representative?
- Are the survey questions unbiased?
- Is the survey technique suitable?

Assessing the Validity of Data Analysis

When you read or hear a conclusion based on data analysis, ask questions to help you decide whether or not the conclusion is valid.

- What is the source of the data?
- How many pieces of data were used?
- Are graphs constructed accurately?
- Is the correlation strong?
- Is the relationship likely to be cause and effect?

Indices

An index is a single number calculated from several pieces of data.
Many indices describe a cost or quantity in comparison to a cost or quantity at a particular time or in a particular place. This is called the base for the index.

Price indices are often presented as broken line graphs with these features:

- The horizontal axis is time.
- The vertical axis values are relative to a base value.
- The graph shows changes in value over time.
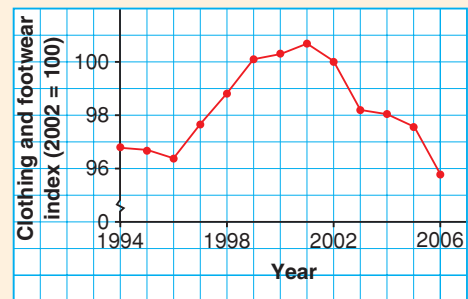- The index may be used to predict future trends.

Some other indices, such as the Body Mass Index, do not have a base value.
They use a formula to determine a number that describes something about a person, place, or thing.
This number allows comparisons to be made.
These indices are often presented as tables or bar graphs.

Consumer Product Index for Clothing and Footwear

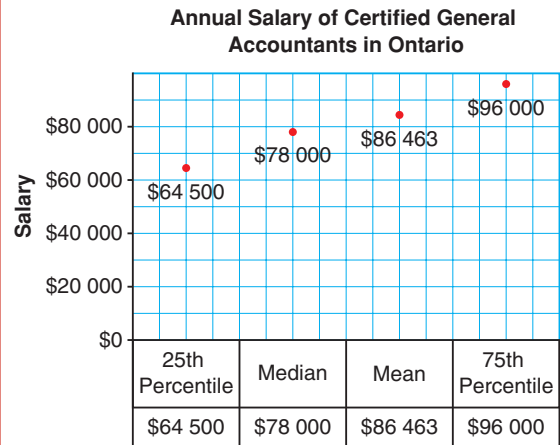$$BMI = \frac{mass \text{ in kilograms}}{(height \text{ in metres})^2}$$

4.1

1. A group of people were discussing marks of high-school students. Irfon overheard this statement: "All of our students should have marks above average."
a) Explain the error in the statement.
b) What do you think the speaker actually intended by the statement?

2. Here are the batting averages of 18 Blue Jays players for the 2007 season.

.297	.178	.245	.233	.240	.262
.289	.242	.251	.238	.236	.208
.238	.277	.216	.204	.291	.167

a) Order the data from least to greatest.
b) What are the quartiles for this data set?
c) Frank Thomas's batting average is in the 80th percentile for this group. Explain what this percentile means. What was Frank Thomas's batting average for the 2007 season?

3. Rahim researches whether it is better to drink bottled water or tap water. Explain which of these organizations' Web sites are likely to have reliable information.
a) A think tank studying environmental issues
b) An association of water bottling companies
c) A company selling water purifying equipment for the home
d) The city of Kingston
e) A scientific organization concerned with water management issues

4.2

4. Livia found this graph on the Web site of the Certified General Accountants of Ontario. It is based on data collected in an e-mail survey in 2006. Almost 6000 people responded to the survey—a response rate of 38%.

Annual Salary of Certified General Accountants in Ontario

	25th Percentile	Median	Mean	75th Percentile
	$64 500	$78 000	$86 463	$96 000

a) What was the 25th percentile salary? What does this value mean?
b) What was the 75th percentile salary? What does this value mean?
c) How confident are you in the results of the survey? Justify your answer.

4.3

5. A school has 1240 students. How many should be surveyed to sample each portion of the population?
a) 5%   b) 10%   c) 15%   d) 30%

6. A TV station wants to know if the community would like more local news included in the evening news. Viewers are asked to e-mail or text their opinions.
a) Why is this sample not representative of the community?
b) Describe how you could make the sample more representative.

7. A town's recreation director conducts a survey about building a swimming pool in the community centre.

- He designs a questionnaire and places it at the community centre.
- People using the centre pick up a copy, fill it in, and drop it in a box.

a) The sample is not representative of the people who live in the town. Explain why not.

b) Suggest a way to make the sample more representative.

4.4

8. The mayor of a town of 15 000 wants to find out if there is a relationship between the amount of time people spend commuting to work and the amount of time they spend with family and friends. She hires you to conduct a survey.

a) What sample size would you use?

b) How would you select your sample?

c) How would you make sure your sample is representative?

4.5

9. Is each statement true or false? Justify your answers.

a) Three-dimensional graphs display data more accurately than two-dimensional graphs.

b) When the population of a survey is the students at a school, the sample should include students from each grade.

c) People may be reluctant to provide salary information or other sensitive information in a personal interview.

d) A graph of two-variable data that has a strong positive correlation means there is a cause-and-effect relationship between the variables.

10. For each survey, decide who you think would be more likely to collect reliable data.

a) A survey on pesticide restrictions
  i) A company that makes pesticides
  ii) A college student doing a project

b) A survey on college students' opinion of a career in the entertainment industry
  i) A radio show
  ii) A college's student life office

11. A PhD student in health economics studied U.S. fuel prices and health data. In a research paper, he wrote:

> A $1 (U.S.) increase in gas prices would, after 7 years, reduce U.S. obesity by approximately 9%, saving 11 000 lives and $11 billion per year.

a) List at least two questions you could ask him about his research before deciding whether you agree with his conclusions.

b) Why might gas prices have an effect on obesity?

12. A flyer advertising a children's soccer camp is delivered to Jan's home. It says: "Parents prefer soccer to any other sport!" Jan asks about the statement. He is told that at a soccer tournament, 50 parents were asked: "Which sport do you prefer your children to play: soccer, tennis, or judo?" All of the parents chose soccer.

a) Record as many reasons as you can that would cause you to question the above conclusion.

b) Compare your list with a partner.

c) Select the most convincing reasons and write a response to the soccer camp, explaining why you think their statement is misleading.

13. Here are the 2008 Environmental Performance Index (EPI) results for North and Central American countries. The EPI was designed to help countries evaluate their environmental policies. The index uses a formula to calculate a score out of 100 for each country, with 100 being the best possible score.

Country	EPI
Canada	86.6
United States	81.0
Mexico	79.8
Belize	71.7
Guatemala	76.7
Honduras	75.4
El Salvador	77.2
Nicaragua	73.4
Costa Rica	90.5
Panama	83.1

a) Which country has a higher EPI score than Canada's?
b) Which countries have a lower EPI score than Canada's?

Use this graph for questions 14 and 15.

Computer Price Index for Consumers, Jul 2001 to Jan 2007
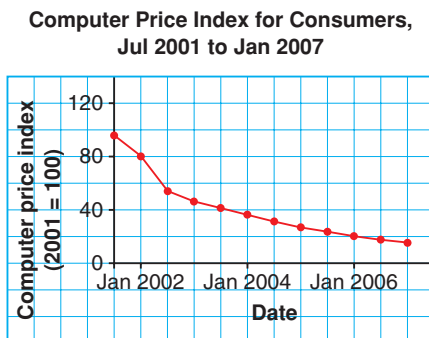
14. a) What is the price index measuring?
b) What was the base year? What does this mean?
c) What was the index value in January 2002? What does this value mean?
d) Describe the general trend in the data.

15. a) When was the computer cost about 40% of the cost in 2001? How can you tell this from the graph?
b) Determine the change in the Computer Price Index between January 2002 and January 2007. Calculate the average annual rate of decrease in the Computer Price Index during this 5-year period.
c) What does this information mean to consumers?

16. Go to the Statistics Canada Web site. Follow the steps from Lesson 4.7 to use data from E-STAT, but select **Agriculture price indexes** rather than **Construction price indexes**. Use the annual Farm Product Price Index from table **002-0022**. Retrieve data about the prices of all the listed commodities from 1997 to 2006.
a) Which two commodities showed the greatest increase in price from 1997 to 2006?
b) Which two commodities showed the least increase in price from 1997 to 2006?

17. Name an occupation that involves data management. Describe the educational background required for the job and how data management skills are used.

Choose the best answer for questions 1 and 2. Justify your choice.

1. "An extremely cold day occurs when the temperature is below the 10th percentile of historical temperatures for that day."
What does this phrase mean?
A. The temperature is below −10°C.
B. 10% of the days in a year are extremely cold.
C. February 2 is extremely cold about 10% of the time.
D. In 100 years, February 2 will be extremely cold 10 times.

2. Which of the following statements is false?
A. A price index can be used to compare changes in prices over time.
B. A graph of the average monthly price of gas is an example of an index.
C. Price index values are often calculated relative to a base value at a certain point in time or in a certain place.
D. Index values can be greater than 100.

Show your work for questions 3 to 6.

3. **Knowledge** These tables show the number of paid admissions to movie theatres in Canada and the population of Canada for 5 years.

Paid Admissions to Movie Theatres in Canada	
Year	**Paid Admissions**
1996	89 024 000
1997	96 805 000
1998	109 688 000
1999	117 352 000
2000	117 574 000

Population of Canada	
Year	**Population**
1996	29 610 757
1997	29 907 172
1998	30 157 082
1999	30 403 878
2000	30 689 035

a) In 1996, the number of movie admissions per capita was
89 024 000 ÷ 29 610 757, or about 3.0. What does this mean?
b) For each year, determine the per capita rate of movie theatre admissions.
c) From your results in part b, can you say that Canadians are going to the movies more often? Explain.
d) Why is the per capita rate more meaningful than the raw numbers?

4. Application A high-school soccer coach wonders if there is a relationship between the number of hours students work at part-time jobs each week and the number of hours they spend on school sports and clubs.

a) Create unbiased survey questions that could be used to research this topic. Explain why your survey questions are unbiased.

b) The school has 1000 students in Grades 9 through 12. Suggest a sample size and sampling technique that could be used for this study.

c) Would you recommend a written questionnaire or interview for this study? Justify your answer.

5. Thinking Ari and Bianca plan to open a computer store in their town of 25 000 people. They need a bank loan to begin the business. As part of their loan application, they prepare a business plan. It includes this information.

- We surveyed 60 adults in town to determine the level of support for our business venture.
- We selected our sample by randomly choosing numbers out of the town phone book.
- We called, asked to speak to an adult, then posed our question.

> It is nearly impossible to function today without a computer. Our town has no computer store. Do you agree that it would be a good idea to have a computer store in town to serve all your computer needs?
> ❑ Strongly agree
> ❑ Somewhat agree
> ❑ Somewhat disagree
> ❑ Strongly disagree

- Our conclusion: Since 70% of the people would like a computer store in town, we will have a good customer base for our business.

If you were the bank manager, would you find their survey persuasive? Justify your answer.
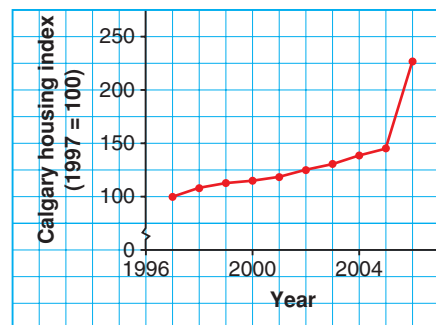
It Would Be a Good Idea to Have a Computer Store in Town

Strongly disagree 10%
Somewhat disagree 20%
Somewhat agree 15%
Strongly agree 55%
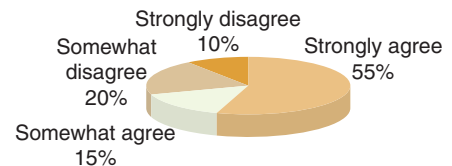
6. Communication Use this graph.

a) Your friend has never heard of a price index. Explain to your friend what the graph is showing and how it can be used.

b) In which year did the price of new houses in Calgary have the greatest increase? Explain how you know.

c) What was the average annual percent price change for new houses in Calgary from 1997 to 2006?

New Housing Price Index for Calgary

(graph: Calgary housing index (1997 = 100) vs Year, values rising from 100 in 1996 to about 225 in 2006)

Materials

- *Microsoft Excel*
- frenchfries.xls

Suppose your school was considering banning French fries from its cafeteria. Do you think students would agree or disagree with this plan?

You will prepare a report on this issue for the school administration. For the report, you will collect, analyse, and display data. Your report should be fair and unbiased.

1. Prepare a survey about students' opinions on banning French fries from the cafeteria. Include questions about grade, gender, and how often students purchase food from the cafeteria. Decide on a sample size. Plan a sampling technique.

2. Conduct your survey or use the sample data provided in the file *frenchfries.xls*.

3. Analyse the data, using a computer if one is available. Create tables and graphs.

4. Find supporting data in the media or on the Internet.

5. Prepare your report.

	A	B	C	D	E	F	G
1	Collected in Gym				Collected in Cafeteria		
2	Grade	Gender	French Fries should be banned from the school cafeteria.		Grade	Gender	French Fries should be banned from the school cafeteria.
3	9	M	Agree strongly		9	M	Disagree strongly
4	9	M	Agree strongly		9	M	Disagree strongly
5	9	F	Disagree strongly		11	F	Disagree