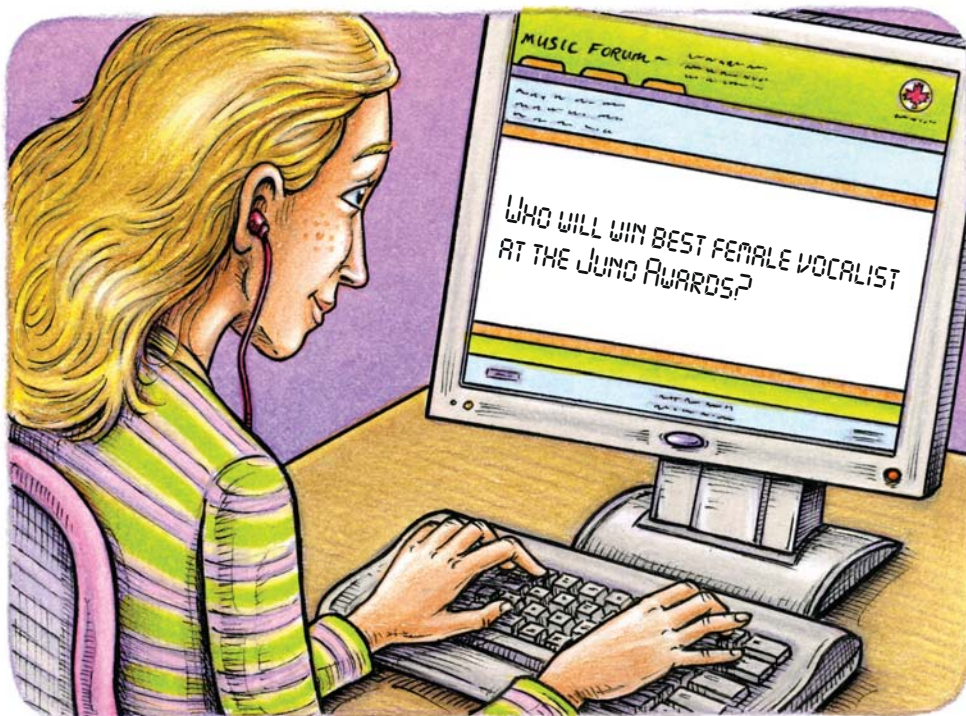# Collecting Data

Focus On ...

After this lesson you will be able to...

- **identify the difference between a population and a sample**
- **identify different types of samples**
- **justify using a population or a sample for given situations**
- **determine whether results from a sample can be applied to a population**

Do Internet forums accurately reflect the opinions of all their readers? Do these people share the same opinions as those who do not read the forums?

How can you reduce the chances of making inaccurate predictions from a survey?

Explore Using Survey Data to Make Predictions

Work as a class. Develop a question to determine the opinions of students at your school about a topic of your choice. For a topic, you might choose favourite foods, sports, actors, or musicians.

1. Write and edit your survey question.

2. Survey everyone in your class.

3. a) Organize the results.
   b) Based on the results of the class survey, predict the whole school's response to your question.

4. Does your prediction accurately reflect the opinions of all students in your school? Explain.

> **Did You Know?**
>
> An Internet forum allows people to participate in online discussions. There are forums for many topics, including sports, fashion, and politics.

Reflect and Check

5. Is your class a population or a sample? Explain.

6. How else might you choose people for your survey to reflect the opinions of all students in your school?

7. How can you reduce the chances of making inaccurate predictions using a survey?

Link the Ideas

There are several different types of samples.

convenience sample
- a sample created by choosing individuals from the population who are easy to access

random sample
- a sample created by choosing a specific number of individuals randomly from the whole population. *Random* means that each individual has an equal chance of being chosen. As a result, a random sample is likely to represent the whole population. Data from a random sample can be used to make predictions about the population. Stratified samples and systematic samples are types of random samples.

stratified sample
- a sample created by dividing the whole population into distinct groups, and then choosing the same fraction of members from each group

systematic sample
- a sample created by choosing individuals at fixed intervals from an ordered list of the whole population

voluntary response sample
- a sample created by inviting the whole population to participate

Example 1: Identify the Population

Identify the population for each situation. Then, state whether you would survey the population or a sample of the population. Explain your reasoning.

a) A bicycle store owner wants to know which brand of mountain bike her customers prefer.

b) The school board wants to know how many hours of homework students do each day.

c) A candle manufacturer wants to know how many of its candles are made with flaws.

population
- all of the individuals in the group being studied
- for example, the population in a federal election is all eligible voters

sample
- any group of individuals selected from the population
- for example, a sample of the population in a federal election might be 100 individuals chosen from each province or territory

Solution

a) The population is the store's customers. It depends on the size of the store. A small store might survey all of its customers. A large, busy store would likely survey a sample of customers. For them it would be time-consuming and costly to survey all of the customers.

b) The population is students in schools within the school board. The school board would likely survey a stratified sample of its students. They would want to include the same fraction of students at various grade levels because the amount of homework done varies from grade to grade.

c) The population is all candles made by the company. A small company might check each candle made. A large company might check every tenth candle. It would be costly and time-consuming to check every item.

Show You Know

For each scenario, identify the population. Then, indicate whether you would survey the population or a sample. Explain your reasoning.

a) The Royal Garden restaurant needs to know which main dish its customers favour.

b) Stephan wants to find out if teachers in Canada prefer to wear glasses or contact lenses.

c) A junior hockey team wants to find out why some people who bought season tickets last year are not buying them this year.

Example 2: Identify a Sample

For each situation, describe how the sample could be selected. Identify the type of sample.

a) A teacher wishes to get feedback from her class about the school dance. She plans to survey 5 students out of a class of 30.

b) A telephone company wants to determine whether a fitness centre would be well used by its 3000 employees. The company plans to survey 300 employees.

c) A chain store is trying to decide whether to open a store in Camrose, Alberta. The company decides to survey people in Camrose and three nearby towns. The population of each location is shown in the table.

Location	Population
Camrose	16 000
Bashaw	825
Tofield	1 876
Daysland	876

d) A marketing research company mails surveys to all of the adult residents in a town. The survey asks about brands of consumer products. The residents are asked to mail their responses in a prepaid envelope.

e) A restaurant owner wants to know the favourite pizza topping of his customers. He plans to survey every customer who orders a pizza in his restaurant between 5:00 p.m. and 10:00 p.m. one evening.

Solution

You can use different types of samples in a survey. These are some possible solutions.

a) The teacher could put all the students' names in a box and draw five names. This is a *random sample*.

b) The company plans to survey 300 out of their 3000 employees. To ensure that the sample fairly represents the population, the company might interview every tenth person on the payroll list. This is a *systematic sample*. This type of sample is time and cost efficient.

c) Since the city has more people who use the company's products than the nearby towns, the company could survey 25% of the population in each location. This is a *stratified sample*. Since 25% of each group is surveyed, the same proportion of each town is represented in the sample. In this case, the company would survey 4000 people from Camrose, 206 people from Bashaw, 469 people from Tofield, and 219 people from Daysland.

d) The marketing research company is inviting all residents to participate. This is a *voluntary response sample*. This sample may not represent the population because only those who are interested will respond.

e) This is a *convenience sample*. It is not random since only customers who order pizza are surveyed. However, the sample does target customers who will provide useful input. These customers are easily accessible. The sample provides the owner with information right away and costs no extra money.
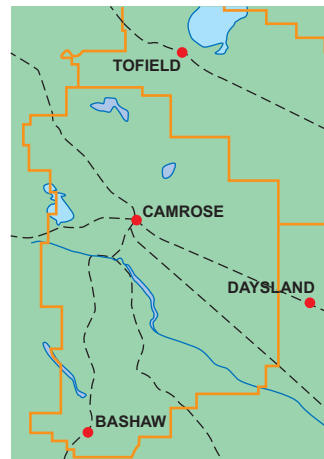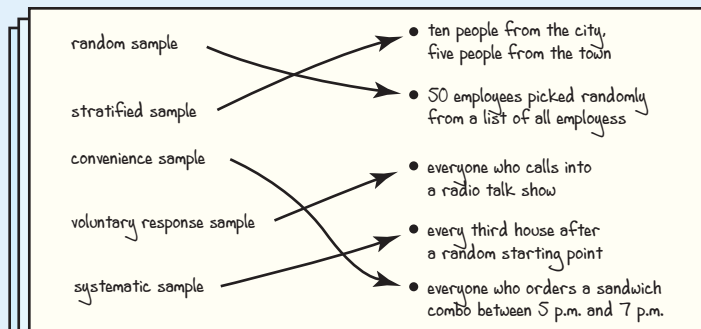
> For each situation, is there a different type of sample that might be used? If so, explain what type of sample and how you would organize it.

Show You Know

a) For each scenario, what type of sample would you use? How would you select the sample?
- A marketing firm plans to conduct a telephone survey in a city of 800 000 people. The survey asks whether there is interest in a new art gallery.
- A student wants to know the most popular cell phone provider that grade 9 students use.

b) In each case, identify the type of sample.
- A coach puts the names of all of the basketball players into a hat and draws one name for a free basketball.
- A questionnaire is sent to every ninth person on an alphabetical list of a store's credit card customers.
- The student council invites all students to provide ideas for activities.

Key Ideas

- A population is the whole group of individuals being studied. It is not always practical or cost effective to survey everyone in a large population. You might survey a sample of the population.
- A sample is any part of the population.
- A random sample ensures that all people have an equal chance of being selected for a study. You can use data from a random sample to make predictions about the population. Systematic samples and stratified samples are types of random samples.
- Voluntary response samples and convenience samples are types of non-random samples.

random sample	→	• ten people from the city, five people from the town
stratified sample		• 50 employees picked randomly from a list of all employess
convenience sample		• everyone who calls into a radio talk show
voluntary response sample		• every third house after a random starting point
systematic sample		• everyone who orders a sandwich combo between 5 p.m. and 7 p.m.

Check Your Understanding

Communicate the Ideas

1. Your friend sends you this e-mail. Write your response.

FROM: Nick
TO: Jonah
I am confused about populations and samples. Would you please explain the difference? It would help if you gave an example of each.
Nick

2. A group can be the population or a sample, depending on the survey question. Do you agree? Give examples to support your opinion.

3. Kim and Scott want to know how many people took public transit to come to the sold-out concert. The seating capacity at the venue is 18 000 people.

Kim: *I plan to ask the first 20 people who arrive.*

Scott: *I plan to ask the first 200 people.*

a) Whose sample would provide a more accurate representation of the population? Explain why.

b) Recommend a different type of sample that would give more accurate results. Explain your reasoning.

Practise

4. Identify the population for each survey question. Indicate whether the population or a sample should be used for the survey. Justify your choice.

a)
> Which band rocks like no other?

b)
> Who will be next year's grade 10 representative on student council?

c)
> Which styles of soccer shirts should be sold at Sports R Us?

d)
> What brand of shampoo do you prefer?

5. In each case, identify the population. State whether you would survey the population or a sample of the population. Justify your choice.

a) A newspaper wants to know the online activities of Internet users at home.

b) The principal wants to know if people associated with the school are in favour of school uniforms.

c) An electronics store needs to find out whether its customers are satisfied with the repairs and services department.

d) The municipal government wants to determine if bus transit is needed for people with special needs.

6. For each situation in #5, what type of sample would you use? How would you select the sample?

7. For each context, identify and describe the sample you would select for a survey.

a) A radio talk-show host wants listeners' views on a proposed by-law about watering lawns and gardens.

b) A province wants to select 25 schools to participate in a new physical education program.

c) A marketing firm wants to know teens' favourite magazines.

d) A reporter wants to ask people downtown about their plans to participate in the Big Valley Jamboree.

> **Did You Know?**
>
> The Big Valley Jamboree, which is held in Camrose, Alberta, is Canada's biggest country music show. It was voted the Canadian Country Music Association's Country Music Event of the Year in 2006.

Apply

8. For each context, would you recommend surveying the population or a sample? Justify your choice.

a) You want to determine the air quality in hospitals in Edmonton, Alberta.

b) You want to know post-secondary plans of grade 9 students.

c) You want to test the quality of parachutes.

d) You want to test the quality of bike tires.

9. Kristi wants to create a menu for a family picnic that will appeal to adults, teens, and children. Her family includes 20 adults, 8 teens, and 12 children. If she has time to talk to only 10 people, how should she choose her sample? Explain.

10. Jason, a member of the Graduation Committee, plans to ask each student who enters the cafeteria the following questions.

> What is your favourite paint colour for the cafeteria walls? _____
>
> Should the cafeteria be used for graduation?   YES   NO

a) Identify the population.

b) Identify the sample.

c) Will the results of his survey accurately represent the population? Explain.

d) Is Jason correct in using the same sample for both questions? Explain your thinking.

11. The student council plans to survey students about how best to spend the budget for activities. Enzo prefers to spend the money on baseball equipment. He decides to randomly survey students at a baseball game.

a) Is there a bias in Enzo's sample? If so, what is the bias?

b) Describe a sample that would reflect the overall opinion of students. Explain your thinking.

12. Anita and Cindi are asked to find out what type of mural to paint in the entrance of their office tower. There are 1400 employees. Cindi proposes using a random sample of 20 employees. Anita suggests using a stratified sample to get input from every department. Whose sampling method is better? Explain your reasoning.

13. Erin plans to survey her friends to determine the average number of children per household in Canada. Is this a random sample? Explain your reasoning.

14. Ben asked 50 people at random in a mall the following question.

> Are you allergic to any of these animals?
> A   cats
> B   dogs
> C   birds
> D   gerbils
> E   hedgehogs

His results are displayed.

Animal Allergy	Frequency	Total
Cats	卌 卌 卌 卌 卌 |	26
Dogs	卌 卌 ||	12
Birds	卌 |||	8
Gerbils	||||	4
Hedgehogs		0

a) Ben made the following predictions based on his data. Do you agree with each prediction? Explain.
   - Almost 25% of the population is allergic to dogs.
   - Hedgehogs do not cause allergies.

b) Improve the survey question. Explain your reasoning.

Extend

15. Search various media for information about a recent survey. Use sources such as magazine and newspaper articles, or radio, television, and Internet reports.

a) Identify and comment on the population and the sample.

b) Are the predictions valid for the population? Explain your reasoning.

c) Was the survey well conducted? What improvements, if any, would you recommend?

16. Five firefighters conducted a survey to assess how well the fire department is performing. They asked local residents the following question.
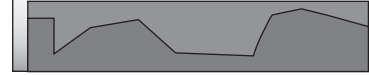
> Which of the following choices best describes your opinion of how well the fire department is doing?
> A   Excellent
> B   Very Good
> C   Good
> D   Poor

A week later, the local paper has the headline as shown.

THE DAILY NEWS

Survey Reveals Majority Very Satisfied With Fire Department

a) How might the wording of the question affect the collection of data? Rewrite the question to produce more accurate responses.

b) What else may have influenced the collection of data? Describe how to conduct a survey that would reflect opinions more accurately.

Math Link

For your research project, choose one of the research questions you wrote for the Math Link on page 421.

a) What is your question?

b) Write a hypothesis that clearly states what you want to prove or disprove.

c) Identify and describe the population for your question.

d) To answer your question, you will need to find data from studies and surveys that have already been done. You will do the research in the Math Link on page 439. What sampling methods do you think would be used to collect data related to your question? Explain your reasoning.

An example is shown.

Literacy Link

A *hypothesis* is a statement put forward to guide an investigation.

Research Question: What are the population trends for beluga whales in Canada?

We predict that all beluga whale populations in Canada are threatened.

We hypothesize that there are few differences in the population trends for the whale populations in Canada.

Beluga whale, Somerset Island, Nunavut

The beluga whale populations in Canada can be organized into seven groups.

1. St. Lawrence
2. Ungava Bay
3. Eastern Hudson Bay
4. Western Hudson Bay
5. Cumberland Sound
6. High Arctic
7. Beaufort Sea

We expect that studies will use samples of beluga whale populations since there are large numbers of whales spread out over large areas.