

9.3

Using Samples and Populations to Collect Data**FOCUS**

- Select and defend the choice of using a population or a sample.



To estimate the number of salmon in a river, biologists use a strategy called *mark and recapture*. At one place in the river, biologists capture some fish. Each fish is marked with a tag, then released into the river. At a different place in the river, biologists recapture fish. They track the numbers of marked and unmarked fish caught. They can then estimate the salmon population.

Connect

When collecting data, the **population** is the group about which you are getting information.

A **census** is conducted when data are collected from each member of the population. For example, suppose you test game consoles made in a factory for defects, then *all* the game consoles made in the factory are the population. If you test each game console, then you have conducted a census.

A census can be costly, time consuming, and difficult or impossible to complete. So, a census is only used when an issue is important or when the population is small.



If a census is not feasible or necessary then data are collected from a small portion, or **sample**, of the population. When the sample chosen is representative of the population, the data collection provides **valid conclusions**.

For example, testing 100 game consoles out of 1000 made each day is a sample. If those consoles tested represent the typical quality of consoles made in the factory, the conclusions of the data collection will be valid.

Care must be taken when determining the appropriate size of the sample. If the sample is large, the data collection could be costly or time consuming. If the sample is small, then it may not be representative of the population.

Example 1**Explaining Why Data Are Collected from Populations**

In each case, explain why a population was surveyed instead of a sample.

- a) To determine the average number of siblings of his classmates, Carlos surveyed each person in the class.
- b) Every 5 years, Statistics Canada conducts a census. One question in the survey is used to determine the ages of the people in each household.

▶ A Solution

- a) Carlos knows that surveying the entire population will produce exact results, rather than estimates. So, he chose to survey the entire population, the whole class, because it would not take long or cost him anything.
- b) A census was completed because of the importance of the question. The government requires data about the ages of Canadians so that it can budget for services such as day-care centres, schools, and senior citizens' homes.

Example 2 Reasoning Why and When Samples Should Be Used

The student leadership team is planning a school dance. To attract grade 9 students to the dance, the team decided to collect data about the preferred music of the grade 9 students. The team set up in the hallway to collect the data. By the end of the day it had surveyed 73% of the grade 9 students.

- a) Why do you think the data were collected from a sample instead of the entire population?
- b) Will the opinions of the sample likely reflect those of the population? Explain.

A Solution

- a) There was probably not enough time or people available to ask all grade 9 students. It would also require a lot of effort to find each grade 9 student, especially with absences.
- b) Since the majority of students, 73%, were asked, it is likely that their opinions will reflect those of the entire population.

Example 3 Identifying and Critiquing the Use of Samples

In each case, identify if data were collected from a sample or a population. Wherever a sample was used, explain if you think the conclusion would be valid.

- A province considers banning cell phones in all of its schools. To determine the opinions of students on this issue, you poll each student in your school.
- To determine which politician is expected to win the municipal election, every person over 18 and who is eligible to vote in the election is polled.
- To determine the average lifetime of a type of light bulb, 150 light bulbs were selected randomly from the production line and tested.

**▶ A Solution**

- Sample:** The population is all students of all schools in the province. By asking only the students in your school, your results are based on a sample. If the students in your school do not represent typical students in the province, the conclusion will not be valid. For example, if all students in your school own cell phones, your conclusion would probably be not to ban cell phones. However, not every student in the province owns a cell phone. So, your results would not be representative of the population.
- Population:** All possible voters are polled.
- Sample:** Since not all light bulbs were tested, the results are based on a sample. It would not make sense for the whole population to be tested, since all light bulbs would be destroyed in the process. There would be no light bulbs left to sell. Since a fairly large number of light bulbs were tested, the results will likely give a good estimate of the lifetime of a light bulb. So, the conclusion about the lifetime of a light bulb is likely valid.