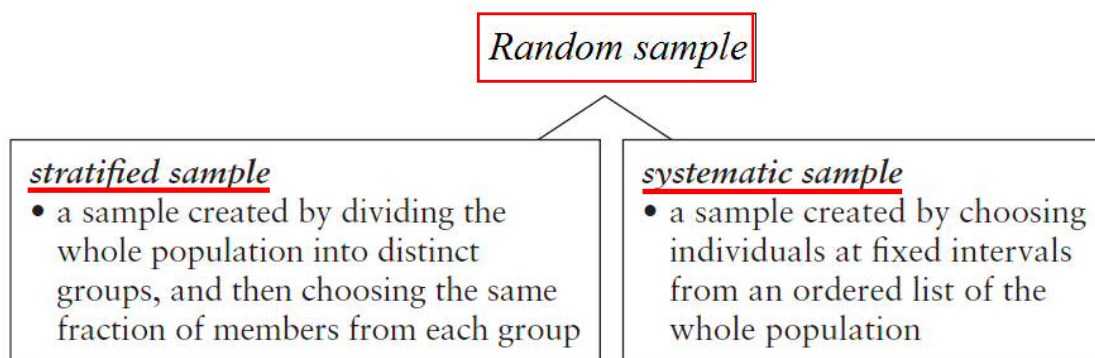# Collecting Data

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.

Random sample

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 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.

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

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

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.

Example 2) 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.

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

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

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) 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

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) 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.

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) 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.

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.

Sampling bias.

- bias that results from the sample not reflecting the characteristics of the population
- it may occur due to the sampling technique or the data collection method

Sampling technique.

- deciding what part of the population will be surveyed to reflect the whole population
- examples include convenience sampling, voluntary surveys, and stratified sampling

Data collection method

- the way in which data are collected or obtained
- examples include direct observation, experiments, and surveys

Example 3) Identify whether any **sampling bias** is in the following survey and state whether it results from the **sampling technique** or from the **data collection method.** If there is sampling bias, how might it affect the validity of the data? How could bias be avoided?

The student council conducted a survey of members of the cheer team. They were asked whether the extra-curricular activity fund should be used to pay for the design of the cheer team's new logo or for new shirts for the football team.

Since the survey is only given to members of the cheer team, it is not representative of all of the students. Sampling bias would occur because of the sampling technique. The results of the survey would be invalid. Asking the whole student body would have resulted in unbiased results.

Homework: p. 426 # 2-10, 14