

Worksheet 8-1: Sampling Techniques

What is Statistics?

Even though you may not have realized it, you probably have made some statistical statements in your everyday conversation or thinking. Statements like "I sleep for about eight hours per night on average" and "You are more likely to pass the exam if you start preparing earlier" are actually statistical in nature.

Statistics involve gathering data from real-life events in order to make predictions about future events. Surveys are usually used to collect statistical data. Since it is not practical or possible to survey an entire population in many situations, a representative sample of the population is selected. There are a variety of sampling techniques, each with its strengths and weaknesses. The choice of sampling technique depends on the population and the type of survey being conducted.

Let's use a statistical problem to learn about specific terms related to statistics.

Problem: There are 300 students taking mathematics this semester at Royal Secondary School. Mrs. Barron, the principal, wishes to survey a total of 60 mathematics students. She has chosen 20 grade 9s, 16 grade 10s, 15 grade 11s, and 9 grade 12s.

How many mathematics students are failing their math courses at mid-term?
 How many hours on average are spent on doing math homework in a week?

Definition	Example
A population is a group of individuals or items that are being studied, and from which samples are taken for statistical measurement.	In the problem above, the population of the survey is the 300 students taking mathematics this semester at Royal Secondary School.
A sample is a group of individuals or items that are representative of the population from which they are selected.	In the problem above, the 60 mathematics students chosen by the principal from each grade are representative of the population (300 students taking mathematics at the school).
A bias is a statistical sampling or testing error caused by systematically favouring some outcomes over others.	In the problem above, if the 60 mathematics students are not chosen randomly, and only those who are passing their math courses are selected to complete the survey, the statistical data collected by the survey would be biased, and the sample would not be representative of the population.

1. In each situation, identify the sample and the population.

- (a) Mary surveyed her science class to identify the rock group most listened to by the students in her grade.

(b) Tony works for a political party. He stands at a busy street corner asking people whom they plan to vote for in the upcoming election.

(c) Karen noticed that most teenagers at the local library were reading murder mysteries. She suggested to the student council at her school that they organize a school murder mysteries read-a-thon.

(d) Angelo is the manager of a gym. He asks members who are at the gym on Wednesday night if he should have a party for gym members.

2. Meena wants to know which band Ontario high school students think is the best. Meena's friend, Cindy, goes to a different school, so they each survey students at their own schools. Meena uses the completed surveys from both schools to draw conclusions.

(a) Identify the population and the sample.

(b) Is the sample representative of the population? If it is representative, explain why. If not, suggest how the sample could be improved.

Sampling Techniques:

There are a number of different ways to choose a sample.

- **Simple Random Sample**
Each member of the population has an equal chance of being selected.
- **Stratified Random Sample**
The population is divided into subgroups (for example, by gender, age or nationality) and a random sample is selected from each subgroup in proportion to its size in the population.
- **Voluntary-Response Sample**
The sample contains those members of the population who have chosen to respond to the survey.
- **Cluster Sample**
The population is divided into clusters and a certain number of clusters are chosen. Every member of these clusters is part of the sample.
- **Convenience Sample**
The sample contains those members of the population from which data are most easily collected.
- **Systematic Sample**
Every n^{th} member of the population is selected.

3. Identify the sampling technique for each scenario.

(a) A mineral water company wants to find out why people choose a particular brand of mineral water.

(i) a mail-in response card is attached to the neck of every bottle of mineral water the company sells

(ii) phone every 200th person who mails in a coupon as part of a money-back offer on the purchase of mineral water

(iii) phone 400 numbers chosen randomly from all the residential phone numbers in a province

(iv) mail questionnaires to 200 men and 200 women chosen randomly from a list of men and a list of women compiled from a market research company's mailing list

(v) ask the first 20 people seen carrying mineral water in a park on a summer day:

(b) You intend to survey your school population to determine whether the students would like a dance in June.

(i) ask every third person walking in the hall

(ii) distribute a questionnaire to each student and then wait for the responses

(iii)ask your friends

(iv)obtain all the class lists and generate random numbers to identify who should be in your sample

(v) use random numbers to select one class from each grade, then use random numbers again to pick names from each class

4. Would the following produce a random sample of students from the class? Yes or No.

(a) One student lists the names of six students from memory.

(b) Five letters of the alphabet are chosen randomly. Students whose surnames begin with these letters form the sample.

(c) Choose the five oldest students in the class.

(d) The name of each student is written on a piece of paper, and all pieces are placed in a hat. The names of eight students are drawn.

(e) The sample consists of the first six students to arrive in the class.

5. Determine the best sampling method for each survey.

- (a) The school newspaper wants to determine which presidential candidate in the upcoming student council elections is supported by the majority of students.
- (b) A light bulb manufacturer wants to determine the lifespan of a certain type of light bulbs, in hours.
- (c) The Parent-Teacher Association wants to determine the average number of hours per week that students spend on homework.
- (d) The producers of “Rock Idol” want to determine which of the two remaining candidates should be the next rock idol.

- Answers:**
- (a) sample: students in Mary’s science class; population: all students in Mary’s grade,
(b) sample: people who pass Tony; population: all voters,
(c) sample: teenagers at local library; population: all students at Karen’s school,
(d) sample: people at the gym on Wednesday night; population: all members of Angelo’s gym;
 - (a) sample: students surveyed in both schools; population: all Ontario high school students,
(b) No. Only two schools are surveyed out of about 5000 high schools in Ontario. Sample size is too small to be representative of the population. Should include more schools from different areas in Ontario;
 - (a) (i) voluntary-response, (ii) systematic, (iii) simple random, (iv) stratified random, (v) convenience
(b) (i) systematic, (ii) volunteer, (iii) convenience, (iv) simple random, (v) stratified random;
 - (a) No, (b) No, (c) No, (d) Yes, (e) No;
 - (a) simple random, (b) systematic, (c) stratified random, (d) voluntary-response