

COLLECTING AND ANALYZING DATA

Learning Goals

- analyze data and draw conclusions from statistical measures and graphs



COLLECT & ANALYZE DATA

Now that we know how to choose an **unbiased sample**, we now need to go out and collect the data.

Sampling techniques: Simple Random Sample Cluster Sample Stratified Random Sample
Convenience Sample Systematic Sample Voluntary-Response Sample

WHERE DOES DATA COME FROM?

Primary Source: someone who collects data for their own use

Secondary Source: data that is collected by someone else (i.e., Statistics Canada)

HOW DO YOU CREATE GOOD SURVEY QUESTIONS?

Just like an unrepresentative sample, the wording of questions can also introduce bias into a survey. A survey should be as free from bias as possible.

Reliable: results can be duplicated in another survey

Valid: results represent the whole population

<u>TYPES OF BIAS</u>	
Loaded Questions	contains wording to influence the response
Leading Questions	suggested answers are given ex. multiple choice
Sampling Bias	surveyed people have something in common.
Response Bias	people lie
Non-response Bias	people don't complete the survey

Do these questions have bias? DISCUSS IN YOUR GROUPS

1. "You don't like skinny jeans, do you?"

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____
2. "What do you think of the new driving law, 0.5% alcohol or less?"

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____
3. "What is the best video game series of all time? Legend of Zelda, Super Mario, or Other?"

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____
4. "Are you upset about the construction on Woodbine?" (ask people stuck in traffic on Woodbine)

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____
5. "You do not have to show me, but did you do your homework?"

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____
6. "What will the weather be like in Jamaica for the October break?"

Type of Bias: _____ Source: _____

Revised Question: _____

Sample Technique: _____

Do these questions have bias?

DISCUSS IN YOUR GROUPS

1. "You don't like skinny jeans, do you?"
 Type of Bias: **loaded** Source: **primary**
 Revised Question: **Do you like skinny jeans?**
 Sample Technique: **convenience**
2. "What do you think of the new driving law, 0.5% alcohol or less?"
 Type of Bias: **response** Source: **primary**
 Revised Question: **What do you think the % should be?**
 Sample Technique: **convenience / random / voluntary**
3. "What is the best video game series of all time? Legend of Zelda, Super Mario, or Other?"
 Type of Bias: **leading** Source: **primary**
 Revised Question: **no multiple choice / more choices**
 Sample Technique: **convenience / simple / voluntary**
4. "Are you upset about the construction on Woodbine?" (ask people stuck in traffic on Woodbine)
 Type of Bias: **sampling** Source: **primary**
 Revised Question: **broaden sample**
 Sample Technique: **simple / systematic**
5. "You do not have to show me, but did you do your homework?"
 Type of Bias: **response** Source: **primary**
 Revised Question: **Do you do your homework?**
 Sample Technique: **cluster**
6. "What will the weather be like in Jamaica for the October break?"
 Type of Bias: **NA** Source: **secondary**
 Revised Question: **NA**
 Sample Technique: **weather network / Stats Canada**

DISCUSS IN YOUR GROUPS

1. The members of an elite golf course are asked if they will approve the construction of a subsidized housing complex on the land adjacent to their golf course.

Sampling Bias

*Not representative
of people who live there*

2. A group of professional football players are asked if they have ever taken banned performance enhancing substances.

Response - people lie

3. Melissa asks five of her friends about their favourite type of movie. Four say their favourite is comedy. She reports that 80% of teenagers prefer comedies over all other movie types. Explain two flaws in her survey.

1. Sample size is too small.
2. Only friends
∴ similar taste in movies

Survey Design

- Clearly identify the population you want to survey
- Have a clear purpose
 - » What do you want to find out?
- Decide if you want open or closed questions
 - » open - How should...
 - » closed - multiple choice
- Create clear and precise questions
 - » How many cigarettes do you smoke? per day?
- Avoid Leading questions
 - » Do you think student council has done anything useful yet this year?
- Decide if responders will remain anonymus

In your groups...

Work on handouts

Homework

Finish handouts

1. Identify each as a primary or a secondary source of data.
- a) Rachel has her friend Siobhan survey students at her school about their musical preference.

secondary

- b) Monique asks students in her science class to suggest the most important scientific discovery over the past 100 years.

primary

- c) Rui checks Statistics Canada to find out which Ontario provincial park is the largest.

secondary

- d) Travis' survey indicated that almost 40% of his friends have the same lunch every Monday.

primary

2. a) Explain why information such as the average number of children per family in the year 1910 cannot be found using a primary source.

Very few people would still be alive.

- b) What possible secondary sources could be used to find the information in part a)?

Statistics Canada website

- c) In general, is a primary source more or less accurate than a secondary source?

Both sources can be accurate for some surveys and inaccurate for others.

3. Identify the bias in each sample.

- a) A phone-in radio show asks callers their opinion about the city's baseball team not making the playoffs the day after they are eliminated.
- Callers will be highly critical of their team a day after they missed the playoffs.

- b) To learn which team Canadians feel will win the Stanley Cup, researchers call 100 people from Ottawa.

People from Ottawa will be biased in favour of their team.

- c) To find the age that most Canadians feel should be the legal driving age, researchers ask 3000 high school students.

High school students are likely to choose a lower driving age than most Canadians.

4. Suggest a way to reduce the bias in each sample in question 3.

a. survey people later

survey people randomly

b. choose people from across Canada

c. choose randomly from all age groups

d. use a place where all age groups will go to conduct research

5. Explain why a biased sample may not produce accurate results.

a biased sample may skew the results towards the opinion of the biased group.