Unit 1
Introduction to Data Management

Day 4
Avoiding Bias
Avoiding Bias

Bias occurs when a sample is not representative of the population due to an unintended (or intended) influence in the data gathering.

Data collected with bias distorts the truth.
For example, students in MDM 4U are surveyed to determine attitudes of students in the school.

Bias in the Sample

A. Sample Bias

• when the chosen sample does not accurately represent the characteristics of the population
• Eg. students at the football game are asked whether the school should fund new equipment for the football team or new instruments for the band.
B. Household Bias

• when one type of respondent is overrepresented because of groupings of different sizes are polled equally
• Eg. Fifty students are sampled from each grade despite the fact that they are different sizes

Bias in the Survey

C. Measurement Bias

• when there is an error in the data collection method
• Eg. A highway engineer suggests that to survey traffic speeds police officers could patrol the highway every half hour and record the speed of the traffic
• Eg. You are measuring the mass of student backpacks but did not calibrate the scale, so the data is inaccurate
Bias from the Respondents

D. Non-response Bias
• when the results are influenced because surveys are not returned
• Eg. People do not return mail-in surveys

E. Response Bias
• when participants in a survey deliberately give false or misleading answers
• Eg. poorly written questions, openly biased interviewer, questions that are sensitive or too personal
• Eg. Teachers asks students to raise their hand if they didn’t understand the homework
Example 1
Identify the type(s) of bias that may result from each of the following data collection methods.

a. You wish to find out how many hours teenagers spend playing video games on an average school night, so you survey your school’s computer club.

   **Sampling bias:** The likelihood that members of the computer club play more video games is quite high.

b. You wish to determine how many students will come to an upcoming dance and so you email 100 surveys to students in each grade.

   **Non-response bias** is possible unless you are there to collect the questionnaires.

   May have household or measurement bias if the number of students in each grade is not equal.
c. To collect data on teen shopping habits, you send surveys to every third house on a street.
   Non-response bias: If the survey is sent to every third house on your street, you may not be targeting homes where teens live.
   Sampling bias: homes may be in a very wealthy neighbourhood.

d. You take a random sample of females to survey them about students preference towards various brands of potato chips.
   Sampling bias: Both males and females should be surveyed, as it is possible to skew the results by asking one gender only.

e. You survey the chess club to determine how many hours of television teenagers in your school watch per week.
   Sampling bias: Chess club members aren’t a fair representation of all teens in the school.
**Example 2**

Give an example of a survey with household bias and another example of one with sampling bias. How do they differ? Explain.

- **Household bias:** You send a survey to 1000 Canadian families in each province.

- **Sampling bias:** Surveying only two Grade 10 classes in your school regarding an issue about all teens.

- In sampling bias, you leave out whole groups of people that are represented in the population.

- In household bias, you misrepresent the number of people in certain groups within your population.
**Example 3**
Identify examples of response bias in the following questionnaire. Then, modify the questionnaire to eliminate response bias.

Identifying the person for whom the survey is being conducted may lead to biased responses.

**Election Survey**
*Sponsored by the friends for the election of Celina committee*

Name: **Necessary?**  Gender: M F  Grade: 9 10 11 12

On election day, I intend to vote for

- Brandy
- Nicole
- Jennifer
- Mary
- **Celina**

I want more

- dances
- dress down days
- holiday
- fun

Who wouldn’t want more fun? This question is not likely to generate any useful information.

Bolding Celina’s name may lead to an inordinate number of responses in her favour. Remember that you need to obtain a true reflection of opinion.
Bias

- Bias in the Sample
  - Sample Bias
  - Household Bias
  - Measurement Bias

- Bias in the Survey
  - Response Bias
  - Non-Response Bias

- Bias from Respondents
Homework

• Pg 123 # 1, 2, 3, 6, 8