MTH1W- Unit 9 Statistics

A worksheet on understanding BIAS

Name: _____
Date: _____

Sampling Bias

When the results of a survey do not reflect the population, the survey is said to contain **bias**.

Bias may be introduced:

- **by selecting a sample that is not representative of the population**
- **by the wording of the survey question**
- **by the presentation of the results of the survey.**

Bias is an intentional or unintentional distortion of the data collected in a survey. A survey should be as free of bias as possible.

Common Bias:

- **Response Bias**
  the type of bias which can affect the results of a statistical survey if respondents answer questions in the way they think the questioner wants them to answer rather than according to their true beliefs
- **Measurement Bias**
  the type of bias resulting from a measurement process that favors a particular result, and systematically overstates or understates the true value of the measurement
- **Non-Response Bias**
  the type of bias resulting from poor response rate and limiting the survey analysis to the available data

1. Identify the bias in each survey.

(a) A parent council survey is conducted to learn if an after-school music program should be offered. The survey question reads: *Early musical training helps develop brain areas involved in language and reasoning. Should the school offer an after-school music program?*

(b) Andre wrote this survey question: *Who do you think is the best female tennis player of all time?*

(a) Martina Navratilova
(b) Chris Evert
(c) Serena Williams
(d) Billie Jean King
(e) Others: _____

(c) A Canadian Football League (CFL) team hands out a survey at a home game. The survey asks: *Should the provincial and federal governments help build a new stadium and save the team from being relocated?*

(d) A radio talk show host asks listeners to phone in and express their opinions on an issue.

(e) A market research company mails survey to 1000 households and 200 returned. The company feels that 200 is a good return and that the opinions expressed can represent the sample of 1000.

(f) Shoppers at a mall are asked: *Are you against the poor decision made by a developer to close this mall in order to build a subdivision, making the traffic in the area even more congested?*

(g) The student council is planning events for Spirit Week. They send out this survey question: *Which of the following events would you like to participate in during Spirit Week?*

(a) ***Carnival day***

(b) read-a-thon

(c) music at lunch

(d) after-school sports

2. At a new restaurant, wait staff ask every fifth customer to answer a questionnaire about food quality and service once they have finished their meal. About 30% of customers surveyed fill out the questionnaire. The majority of customers who fill out the questionnaire complain about poor service. The restaurant manager concludes that the wait staff needs more training. Is the conclusion reasonable?

When you conduct a survey or perform an experiment, you are the **primary source** of the data. When you work with data from the Internet, published materials, or Statistics Canada, you are using a **secondary source**.

3. Duran wants to know what the best-selling lunch item in the cafeteria is. She records what she sees on the students' trays as they pass by the cashier after paying for their lunch. Her school has two lunch periods. When she has completed a survey during her lunch, she asks the cafeteria staff to keep track of the items they sell during the other lunch period. She collects the information after school.

(a) Which data are primary data? Justify your answer.

(b) Which data are secondary data? Justify your answer.

4. The student council wants to know the type of music that should be played at the next school dance. Terence and Linda are asked to collect information.

- Terence uses the Internet to find the top 10 songs on 15 radio stations. He assigns 10 points to the number one song on each list, 9 points for the second place song, 8 points for the third and so on. He determines the point rating for each song and uses the ratings to develop his playlist.
- Linda decides to survey the school population. She creates a questionnaire asking students to list their top 5 songs. The homeroom teachers distribute the questionnaires and give students 5 minutes to complete it. Linda collects the questionnaires and ranks the songs according to the results to develop her playlist.

Identify the type of data source each student used.

Answers: **1.** (a) response bias, (b) measurement bias, (c) response bias, (d) non-response bias, (e) non-response bias (f) response bias, (g) measurement bias**;**

2. No. The conclusion is not reasonable because only 30% of those surveyed answered the questionnaire. Customers who responded might have done so because they were unhappy with the service. Those who were happy with the service might have simply ignored the questionnaire because the service did not need to be improved. This non-response bias leads to inaccurate results since only a small number of people in the sample responded**;**

3. (a) Duran's observations are primary data because she collected information for herself, (b) The cafeteria staff's observations are secondary data because they collected the information for Dharma**;**

4. Terence: secondary data; Linda: primary data**.**