Forms of Bias

1. Sampling Bias:

Sampling bias occurs when the sampling frame does not reflect the characteristics of the population. Biased samples can result from problems with either the sampling technique or the data-collection method.

General Examples:

- A survey of high school students to measure teenage use of illegal drugs will be a biased sample because it does not include home schooled students or dropouts.
- A sample is also biased if certain members are underrepresented or overrepresented relative to others in the population. For example, distributing a questionnaire at the end of a 3-day conference is likely to include more people who are committed to the conference so their views would be overrepresented.
- Interviews with people who walk by a certain location is going to over-represent healthy individuals or those who live near the location.
- Selecting a sample using a telephone book will underrepresented people who cannot afford a telephone, do not have a telephone, or do not list their telephone numbers.

1. Sampling Bias:

Example 1:

Identify the bias of the survey and suggest how it could be avoided.

A survey asked students at a high school football game whether a fund for extra-curricular activities should be used to buy a new equipment for the football team or instruments for the school band.

Solution:

Since the sample includes only football fans, it is not representative of the whole student body. A poor choice of sampling technique makes the results of the survey invalid. A random sample selected from the entire student body would give unbiased results.

1. Sampling Bias:

Example 2:

Identify the bias of the survey and suggest how it could be avoided.

An aid agency in a developing country wants to know what proportions of households have at least one personal computer. One of the agency's staff members conducts a survey by calling households randomly selected from the telephone directory.

Solution:

There could be a significant number of households without telephones. Such households are unlikely to have computers. Since the telephone survey excludes these households, it will overestimate the proportion of households that have computers. By using telephone survey as the data-collection method, the researcher has inadvertently biased the sample. Visiting randomly selected households would give a more accurate estimate of the proportion that have computers. However, this method of data collection would be more time-consuming and more costly than a telephone survey.

1. Sampling Bias:

How can we know if the sample is biased?

Sometimes you can identify sampling bias just by being very thoughtful and comparing the characteristics of respondents in your sample to what you know about the population in general. Think about the demographic characteristics that might have an important relationship to their answers.

For example, if you know that gender is an important variable, and you know that the population includes 50% males and 50% females, then the sample needs to include the same proportions. If the sample includes 20% males, your results are likely to be biased because you don't have enough responses from men.

2. Non-response Bias:

Non-response bias occur when particular groups are under-represented in a survey because they choose not to participate. Thus, non-response bias is a form of sampling bias.

General example:

- A restaurant may give each table a “customer satisfaction” survey with their bill. Generally, there are two types of people that respond to these surveys: those that really enjoy the service, or those that hate it. A large portion of diners – those whose opinions are somewhere in between – are not well represented. When surveying, one method to ensure that a variety of groups are represented is to include questions that identify the group itself.

2. Non-response Bias:

Example:

A science class asks every fifth student entering the cafeteria to answer a survey on environmental issues. Less than half agree to complete the questionnaire. The completed questionnaires show that a high proportion of the respondents are concerned about the environment and well-informed about environmental issues. What bias could affect these results?

Solution:

The student who chose not to participate in the survey are likely to be those least interested in environmental issues. At a result, the sample may not be representative of all the students at the school.

2. Non-response Bias:

How to avoid it?:

To avoid non-response bias, researchers must ensure that the sampling process is truly random. For example, they could include questions that identify members of particular groups to verify that they are properly represented in the sample.

3. Measurement Bias:

Measurement bias occurs when the data-collection method consistently either under- or overestimates a characteristic of the population. While random errors tend to cancel out, a consistent measurement error will skew the results of a survey. Often, measurement bias results from a data-collection process that affects the variable it is measuring.

Measurement bias is a consistent measurement error which skews the results of the survey. Often the data collection process affects the variable that is being measured.

General examples:

- When measurement instruments are poorly calibrated or affect the result (e.g. airflow) in data collection where the data collection method affects the result (e.g.. test anxiety).

- A principal might want to get an estimate of the number of students who use their cell phones in the hallways. She might decide to have teachers keep a tally of how many students they see using their cell phones, as they patrol the halls. Since many students hide their phones when they see a teacher coming, this may not be an accurate data-collection method.

3. Measurement Bias:

Example 1:

Identify the bias in the following survey and suggest how it could be avoided.

A highway engineer suggests that an economical way to survey traffic speeds on an expressway would be to have the police officers who patrol the highway record the speed of the traffic around them every half hour.

Solution:

Most drivers who are speeding will slow down when they see a police cruiser. A survey by police cruisers would underestimate the average traffic speed. Here, the data-collection method would systematically decrease the variable it is measuring. A survey by unmarked cars or hidden speed sensors would give more accurate results.

3. Measurement Bias:

Example 2:

Identify the bias in the following survey and suggest how it could be avoided.

As part of a survey of the "Greatest Hits of All Time," a radio station asks its listeners: Which was the best song by the Beatles?

- *Help!*
- *Nowhere Man*
- *In My Life*
- *Other*

Solution:

The question was intended to remind listeners of some of the Beatles' early recordings that might have been overshadowed by their later hits. However, some people will choose one of the suggested songs as their answer even though they would not have thought of these songs without prompting. The survey would more accurately determine listeners' opinions if the question did not include any suggested answers.

3. Measurement Bias:

Example 3:

Identify the bias in the following survey and suggest how it could be avoided.

A poll by a tabloid newspaper includes the question: "Do you favour the proposed bylaw in which the government will dictate whether you have the right to smoke in a restaurant?"

Solution:

This question distracts attention from the real issue, namely smoking in restaurants, by suggesting that the government will infringe on the respondents' rights.

A question with straightforward neutral language will produce more accurate data. For example, the question could read simply: "Should smoking in restaurants be banned?"

4. Response Bias:

Response bias occurs when participants in a survey deliberately give false or misleading answers. The respondents might want to influence the results unduly, or they may simply be afraid or embarrassed to answer sensitive questions honestly.

General examples:

- The results of a survey of college fraternity students indicate that each student drinks approximately 50 alcoholic beverages per week. While possible, the data are probably skewed by those students who are exaggerating the number of drinks they consume. When analyzing data, it is important to ask how realistic the data are.
- Often in face to face interview where the true answer may be illegal, embarrassing, or unpopular. e.g. survey on illegal drug use or number of speeding tickets.

4. Response Bias:

Example 1:

A teacher has just explained a particularly difficult concept to her class and wants to check that all the students have grasped this concept. She realizes that if she asks those who did not understand to put up their hands, these students may be too embarrassed to admit that they could not follow the lesson. How could the teacher eliminate this response bias?

Solution:

The teacher could say: “This material is very difficult. Does anyone want me to go over it again?” This question is much less embarrassing for students to answer honestly, since it suggests that it is normal to have difficulty with the material. Better still, she could conduct a survey that lets the students answer anonymously. The teacher could ask the students to rate their understanding on a scale of 1 to 5 and mark the ratings on slips of paper, which they would deposit in a box. The teacher can then use these ballots to decide whether to review the challenging material at the next class.

5. Loaded Question:

Steer an individual toward certain responses. In other words, when a question contains wording or information which influences the respondent's answer that is called a loaded question.

General example:

The question "do you support your MP's thoughtless decision to introduce anti-abortion legislation" suggests that the MP is acting without doing much research.

In your opinion, is MDM4U the best math course you have ever taken or is there another less valuable math course that you would consider to be the best.

To avoid loaded questions, do not use words or phrases that are based on individual opinions.

6. Leading Question:

Contain wording or information intended to sway a respondent's opinion. In other words, a leading question is a question that suggests an answer which otherwise may not have been chosen by the respondent without prompting.

General example:

The question "do you prefer Coke, Pepsi or another brand of cola best?" is a leading question because a respondent might choose one of the listed colas simply because they were mentioned.

One way to avoid leading questions is to avoid listing specific options.

Summary:

- Sampling, measurement, response, and non-response bias can all invalidate the results of a survey.
- Intentional bias can be used to manipulate statistics in favour of a certain point of view.
- Unintentional bias can be introduced if the sampling and data-collection methods are not chosen carefully.
- Leading and loaded questions contain language that can influence the respondents' answers.