

What is Missing Data?

In real data, even though all the data is supposed to be collected for future analysis and record, not all the data will be completely collected. Many situations can cause the absence of certain data, such as the respondents not responding to the survey, or the responses are not in the correct form of what the survey needs. Other situations that may end up with missing data are situations such as the design of the survey has some flaws, such as the survey is too long, and ends up with respondents not answering all the questions in the survey. Also, sampling bias can also cause respondents not answering the survey. Sampling Bias exist when the sample is not a representation of the overall target population. When such bias happens, the responses may end up with missing data, or the answers are inappropriate to use for future analysis.

There are different types of missing data. The first type of missing data is the data that are completely missing at random, which means that the reason for such data to miss is not about any variables that can affect the survey. For example, if the data is deleted accidentally, then the data can be considered missing at completely random. The second type of data is missing at random, which means that the missing data are missing due to the variables that are related to the survey, but the actual reasons are not due to the variables. For example, if the respondent chooses not to answer a question because they do not want to give any information about this variable due to their age, then the missing data can be considered as missing at random, because the actual reason for missing is not about the variable itself, but about age. Finally, the last type of missing data is the type of data that is not missing at random, which means that the missing data is directly related to the variables related to the survey. A good example for this situation is the people not giving information about their personal information such as income because they want to keep their privacy.

[Shiji Zhang will provide feedback for my essay. And you can find in the Github link: <https://github.com/JayCao2048/Mini-Essay8.git>]