
Homework for Lecture 1: What is Data?

The following homework assignment should be completed as a Word document submitted to Canvas. Please use complete sentences. You may discuss the assignment with your classmates, but your write-up and work should be done independently of your classmates. The use of AI is considered a violation of academic policy for this assignment. When in doubt, ask me what is and is not allowed.

Answer all of the following questions:

1. Choose one real question that could be answered using data. This question can be about campus life, a hobby, sports, health, entertainment, work, social media, or something else you care about. Do not use the dining hall example from class or examples directly from the notes. In one paragraph, explain:
 - what question you would want to answer,
 - why that question is interesting or useful,
 - who might care about the answer,
 - and what decision, behavior, or change the data could help inform.
2. Imagine you were creating/collecting a small data set to help answer your question from Question 1. Describe what one row of your data set would represent. Then create a table with at least 5 variables you could collect. For each variable, include:
 - the variable name,
 - what the variable measures or records,
 - whether it is quantitative continuous, quantitative discrete, qualitative nominal, or qualitative ordinal,
 - and one sentence explaining why it fits that category.

Your table should include at least 2 quantitative variables and at least 2 qualitative variables.

3. Data does not automatically tell the full story. For the data set you described in Question 2, write one paragraph explaining at least two ways the data could be incomplete, misleading, biased, or easy to misinterpret. Then explain one piece of context that someone would need before making a conclusion from the data.
4. Write a short reflection answering the following question: What is one thing you think people should be careful about when collecting, interpreting, or presenting data? Explain your answer using an example connected to your proposed data set.