

Instructions: Group Project Proposals
STAT 340: Data Modeling II
Due October 22, 2021 at 11:59pm

The first step in your group projects is to select a data set and a statistical question of interest for you to explore. We have posted an assortment of possible data sources on Canvas to give you inspiration, though you are not required to use one of these resources for your project.

Having selected a data set to work on and a question that you seek to answer, your first job is to write a brief proposal explaining your data set and the question(s) that you hope to answer.

By Friday, October 22 at 11:59pm, you should submit via Canvas a document in PDF format that includes

- The names and NetIDs of all group members.
- A description of your data set and a URL at which to access it.
- One to three statistical questions that you intend to answer (fewer questions is, in a sense, better– it will help you focus your efforts in the future)
- A description of the variables available in your data set, with more details for the variables that are likely to be relevant to your question.
- A description of the statistical and/or data science methods that you expect to use. Of course, this is just a plan. You are going to learn a lot more stuff this semester, so if you later find that there are better methods for your problem, that is okay.

There is no particular length requirement for this document, but two pages total is a good target. If you can address all of the above in less space, that's fine. If you need a bit more space, that's fine too. Give enough detail that your instructors and TAs will see that you have thought carefully about your statistical question(s) and that you have actually looked at the data set's documentation.