

Instructions: Group Project Progress Reports
STAT 340: Data Modeling II
Due November 12, 2021 at 11:59pm

The purpose of this assignment is to serve as a check-in on your group's project. Think of this as a sort of first draft of your group's final report (though strictly speaking your "final report" will take the form of a poster). Of course, we have no expectation that you have definitively answered your statistical question of interest or anything to that effect. Rather, the goal of this assignment is to ensure that you are making progress toward that goal, and to give us an opportunity to see potential issues well ahead of the final project deadline at the end of the semester.

By Friday, November 12 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 summary of your statistical question(s) of interest (you may copy-paste this from your proposal, if you wish).
- A brief description of why your group found this data set in particular to be of interest and why your reader should care.
- 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 (you may copy-paste this from your proposal, if you wish).
- A block of code showing how to load the data into R (if your data set is too large to permit this, please speak to us promptly).
- At least one plot summarizing some aspect of the data set that is interesting to you and your group. Your plot should be well-labeled, have an appropriate title, and your document should include a reasonably detailed description of what the plot shows as well as any discussion/interpretation of the contents of the plot (e.g., if the plot shows a clear trend in your data, discuss that trend, why it is or isn't surprising, etc.).
- A brief discussion of the progress and/or challenges faced so far in answering your statistical question(s) of interest. This may include a discussion of the methods and models used; issues that arose when downloading and cleaning the data; shortcomings of the methods/models used so far, etc.
- A summary of your next steps (e.g., your goals for the remainder of November, other methods/models you want to try, etc.).

There is no particular length requirement for this document, but five pages total is a good target (including figures). 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 been working consistently on your project.