STOR 664: Project Assignment: FALL 2022

The project is to complete an analysis of an applied dataset of your own choosing. The dataset must come from some applied scientific or social science field, and must involve multiple regression analysis in some way. Beyond that, it's open-ended. I would like to see a proposal from you, giving an outline of your project and the dataset(s) you propose to use, so that I can give you feedback whether I think it's suitable for this course.

You may work in teams of up to three students per team (but a one-student or two-student team are also fully acceptable). I will expect you to certify that the students have made an equal contribution to the project and the grade for the project will also be the same for each student (e.g. if I give the project a grade of 19/25, that means all members of the team will receive a score of 19). Students on the same team are expected to work closely with each other and it must be an independent project, i.e. no overlap with any other team's work or with any other project you may be doing for another course (if any doubt about the latter point, please ask me). As with any student assignment or exam, you are allowed to ask me for help. While I am not going to forbid you from discussing your project with other students or professors, in general I discourage that – the project is expected to be essentially your own work, and any substantial help you receive from anyone else should be acknowledged.

The project will be graded on the basis of: 50% for technical content, 50% for the application.

Technical content means: selection of appropriate statistical analysis, choosing the model, intelligent use of diagnostic and graphical methods, correctness of any subsidiary calculations you make as part of the analysis (e.g. hypothesis tests, predictions).

Application means: explaining the scientific or social science context in which the data arose, describing the problem that you are trying to solve, how well your analysis goes towards solving that problem, and the quality of your final write-up. The final conclusions should be written in non-statistical language so that a researcher in the applied field could understand your contribution.

The final report is expected to be 10-15 pages in length: you may submit additional material as technical appendices but the bulk of the grading will be based on the main report.

Contribution to final grade for the course: this project is worth 25% of your total grade, as announced in the syllabus.

Deadlines:

Proposal Due: Tuesday, October 25. Each team should submit ONE proposal indicating clearly who the team members are.

Final Due Date: Tuesday, November 29 (last day of class).