About the Project

You are expected to either (1) conduct a thorough analysis of a data set using statistical modeling techniques, (2) use statistical methodology to gain insights into some theory, or (3) explain and apply some (nontrivial) statistical method that we have not discussed in class (e.g., Bayesian linear regression, kriging). You may work in small groups if you wish. I hope that you choose a data set or problem based on your interests (e.g., for graduate students, it would be a great idea to choose a data set relevant to your own research interests), rather than just “going through the motions”. You will be expected to describe your results in a formal paper, and briefly present your results to the class (date TBD). The project grade will be determined on the basis of the accuracy of the statistical analysis and the quality of the paper and presentation. I expect graduate students to have a more sophisticated understanding of statistical techniques, procedures, etc., and thus, I will hold graduate students to a (slightly) higher standard. You will be required to attend all presentations and will be graded accordingly. If you miss the final project presentations will receive a zero for their project grade.

Grading Procedure

The project grade will be determined on the basis of the quality of the statistical analysis, paper, and presentation. The project will be graded out of one-hundred points. A specific rubric follows:

Fifteen points will be related to oral communication.
Sections and Topics for Data Analysis Paper

I suggest that you include the following sections and topics in your paper and presentation. If you are not analyzing a specific data set (e.g., if you are explaining a modeling method, you should be able to adjust these sections accordingly):

1. **Introduction/Background**
   - Why are you interested in this problem?
   - What do people need to know to understand? Assume that your audience is not an expert in the application field.
   - Is there any prior research on your topic that might be helpful for the audience?
   - From where did the data come? Is this an experiment or observational study? Who collected the data? Why was the data collected (if you weren’t the one doing the collecting)?
   - What are the questions of interest that you hope to answer?

2. **Methods/Results** (experimental design and data collection)
   - How did you obtain the data?
   - Summarize and explore the data. Are any transformations needed? Graphics would be good here.
   - What analyses are most appropriate to answer the question of interest?
   - Describe the analyses used. Check assumptions!
   - Present relevant graphics and interpret results.

3. **Conclusions**
   - What are your conclusions? What did you learn?
   - What would you do to improve the design of your study/experiment?
   - How would you extend this research? What future research ideas come to mind based on your results and experience with this analysis?

**Fifteen points** will be related to written communication. You will receive a grade of 0-3 in each of the following categories: context and purpose of writing, content development, genre and disciplinary conventions, sources and evidence, control of syntax and mechanics.

**Seventy points** will be related to problem solving. You will receive a grade between 0-10 on how well you define the problem/question of interest. You will receive a grade between 0-10 on how well you propose solutions to answer the questions. You will receive a grade between 0-30 on implementing solutions to answer the question of interest. This is the proper application of statistical methods to the problem at hand. You will receive a grade between 0-20 on how well you evaluate outcomes, explain conclusions, interpret results, etc.