Final project

For the final project, you can either come up with your own project or do the last homework. The following information is only if you are doing a project instead of doing the last homework.

The format of the final project can be one of

1. class presentation (plan roughly 10 minutes, or 10 minutes plus 5 minutes per person collaborating (up to 3 people from the class). Presentations should be in slide format (not just chalkboard) and converted to pdf and emailed to me (jamdeg@unm.edu) to avoid switching in the middle of class and relying on potentially incompatible software.

2. poster (the poster doesn't have to printed–you can send the poster as a pdf via email)

3. analyze a journal paper and do a 2-3 page summary, not including quoted figures or tables. The summary should include tables and figures from the paper as an appendix to be discussed in your write up. In the write up, describe the goals, methods, and conclusions of the study as well as statistical methods used, how the data was gathered/sampled, sample sizes uses, and major conclusions in terms of estimates, confidence intervals, or model selection. Indicate strengths and weaknesses of the paper and any way you think that the paper could have been improved.

4. analyze your own data, for example data you generate by counting word frequencies from text, job advertisements, twitter, amazon reviews, etc. If you generate your own data, explain how the data is generated, how samples were collected, whether randomization was used, etc. No experimenting on humans, please. Examples of text analysis include

   - collecting abstracts from one particular journal and seeing if there are patterns for male versus female authors (sex might have to be assumed based on names)
   - writing samples from the same author to see if their style changed over time, such as early versus later Harry Potter novels, Isaac Asimov novels, etc.
   - writing samples from the same topic from different newspapers, such as the New York Times and Wall Street Journal.
   - YouTube comments for anonymous looking versus nonanonymous looking user names

Whatever you choose for your final project, you should include methods used from ADA 2, such as ANOVA with multiple predictors, multiple regression, model selection, testing for interactions, logistic regression, principal components, cluster analysis, and classification analysis (which is starting this week, although logistic regression can also be used for classification).

With an open-ended project, grading is difficult, but consider the following rubric for a 25 point scale:

1. (5 points) presentation clarity (i.e. easy to follow ideas, nice slides, use of figures, etc.)

2. (5 points) good grammar

3. (5 points) use of statistical methods relevant for the class

4. (5 points) correct description of statistical methods/conclusions

5. (5 points) appropriate length