

## STAT512 Project Proposal (10 pts)

**Name(s):**

**Department(s):**

**Due:** Tues 4/02 by midnight

See detailed notes/discussion on the next page, but only this first page needs to be submitted.

1. Considering the topic descriptions on the next page and your answers to the questions below, which topic area do you think is most appropriate for your project data and research goals?
  - A. Multiple Regression
  - B. ANCOVA
  - C. Logistic Regression
  - D. Factorial ANOVA
  - E. Blocking
  - F. Repeated Measures
  - G. Not sure
2. **Response variable**
3. **Is the response variable numerical?**                      Yes                      No  
  
If you circled no, please provide further detail.
4. **Predictor variables.** Identify each variable as continuous or categorical.
5. **Sample size**
6. **Design.** Is there any blocking or repeated measures?
7. **Other comments?**

**Notes:**

1. Students may work individually or in pairs.
2. This proposal will be used to group students and for my own record. It is fine if your data and/or analysis changes for the final project. I will not use the proposal information when I am grading the final projects.
3. Very large data sets ( $n > 1000$  or  $k > 50?$ ) may not be appropriate for the project. Similarly very small data sets ( $n < 10?$ ) may not be appropriate. For large datasets, you can consider using a subset of data for the project.
4. It is fine to use one part or a simplified version of a larger project.
5. If you have questions about whether a data set/analysis will be appropriate, come see me during office hours.

**Description of Topics:**

- A. Multiple Regression: Analysis will include a small number of continuous predictors up to a large number of predictors (can be mixed categorical and continuous).
- B. ANCOVA: Analysis will include a small number (up to 3?) of predictors but at least one categorical and one continuous.
- C. Logistic Regression: Binary response variable with continuous and/or categorical predictors.
- D. Factorial ANOVA: Analysis will include a small number (up to 4?) of categorical predictors. Typically from a designed experiment. More details in ExpDesign2.
- E. Blocking: Blocking refers to when experimental units are grouped in some way. For example: A number of plots (blocks) are selected. Then treatments are randomly assigned within blocks so that each treatment is observed within each block. More details in ExpDesign1 and Random2.
- F. Repeated Measures: Repeated measures refers to studies where repeat observations are taken on the same subject or experimental unit. Most often the repeated measures are across time, but other situations are possible. For example: subjects are randomly assigned to either a treatment or control group. Then measurements are taken weekly for each subject. More details in Random3.