Class Projects (15% of course grade)

**Goal:** To collect, describe, and analyze data to answer two questions of your choice. This will allow you to apply the skills you learn in this course to the world around you which will in turn enhance your appreciation and retention of the material.

**Teams:** For the class projects, you will work in groups of 3-5. It is up to the members of the group to make sure everyone contributes equally. Teams should be formed by Jan. 27.

**Topics:** You are free to choose your own questions. The questions may be related to your major or some other topic of interest. You should choose a topic so that it will be straightforward to gather the data. The easiest approach will be to design an experiment to compare two groups but the only rule is to make sure the topic is interesting to your group! Be creative! We will discuss some previous topics in class, and some previous project topics can be found on the course web pages (e.g., statweb.calpoly.edu/chance/stat217/projects.html). You will want to collect lots of data and then narrow in on two hypothesis pairs later. After the topics are selected, most of the work for the projects will take place outside of class.

**Project Reports:** The goal of the project reports is to keep you thinking about the projects as the term progresses. Keep in mind that your project may change and evolve as the course progresses. Still, with each project report I would like to hear about your progress and ideas. Turn in one project report for each team, including team members’ names and previous project reports, preferably typed. Below are some guidelines on what I would like to see in each report.

You should have your team formed by Jan. 27. Make sure you obtain phone numbers and email addresses for each other. If you think finding meeting times will be difficult, you may want to start dividing the workload into subgroups.

The **first project report** is due Feb. 3. For this report you should identify your topic/questions of interest, the variables you plan to measure, the population you plan to draw conclusions about, the sample and sampling frame you plan to use (if applicable), and the type of study (e.g., survey or experiment). Previous project topics can be found at: http://statweb.calpoly.edu/chance/stat217/projects.html

The **second project report** is due Feb. 10. Your data collection techniques should be more clearly defined. If an experiment, give your tentative design. If a survey, give the preliminary questionnaire. You should indicate why this study is appropriate to answer your question and what precautions you will take (e.g., nonresponse, nonsampling bias, wording). Each team is to bring in four copies of their typed report. These reports will be peer reviewed by other students in the course.

The **third project report** is due March 3. You should have finished collecting your data. Include a description of your observational units, your variables, their measurement units, possible ranges/responses of these variables, as well as preliminary descriptive statistics. You should also specify two research questions/sets of hypotheses that you plan to test using your data. You should review p. 164-165 and indicate which set of procedures you believe you will use. Include a justification for that choice of procedure. You should also outline how the remaining work will be completed (who, when). If you need to set an appointment with me to discuss inference for regression or anova, indicate who in your group will be responsible for speaking with me.

**Rough Draft** (optional). If you turn in a rough draft by March 5. I will review the paper, providing comments and suggestions for improving your final grade.
**Final Reports:** Final reports are due on or before March 12. See last page of syllabus for a checklist of what should be included in the report. Reports must be typed. Turn in one report per group and include previous project reports and evaluations. Incorporate computer output into the body of the paper. *Raw data should be turned in on a separate floppy disk.* You may assume your audience will understand all statistical terminology.

**Grading Criteria for Final Report:**
10%: Quality of written report
20%: Design of survey/experiment – was data collection adequately explained, were the appropriate data collected to answer the questions posed. Points here will include originality of topic chosen.
25%: Correctness of statistical analysis and checks of technical conditions
20%: Appropriateness of interpretations of the results of the statistical calculations and conclusions (is it a cause and effect relationship? what is a reasonable population?)

**Presentations:** (other 25%) Each group will give a 10 minute presentation (at the most) *in the evening* during the last week of classes. Any number of group members may present part of the report. The presentation should not include extensive details, but provide the audience with an overview of what was done, what conclusions can be drawn, any drawbacks of the techniques and further recommendations. Feel free to be creative! The presentation should be accompanied by PowerPoint, transparencies, or a poster board.

*Project Presentations will be given in the evening the week of 3/10. Your group must commit to being present for a 90 min session one evening. You will be asked to critique the other group presentations.*

**Evaluation:** Each person’s grade will be 75% group grade and 25% individual grade. Individual grades will be determined by the instructor and team member evaluations.

**CHECKLIST FOR FINAL REPORTS**
Title
Statement of purpose/question of interest (~ 2 lines)

I. Introduction
   Explain the purpose of the study in your own words, why you were interested in the question and what you expected to see from the data.

II. Summary of data collection (no such thing as “too much detail” in this section!)
   Identify sample, population, sampling frame (if applicable), response rate (if applicable)
   Describe sampling technique (if applicable) or experimental design
   Describe how were measurements taken. Operational Definitions?
   Discuss how you tried to reduce bias
   Identify possible remaining sources of bias and their possible effects on the results
   If a sample, why do you believe the sample was representative?
III. **Summarize and interpret the data** (concise, well-labeled, easy to read)

Categorical data:
- Graphs: pie chart or bar chart (with percents not counts)
- Numerical: Proportion and sample size in each category

Quantitative data:
- Graphs: e.g., histograms or dotplots
- Measures of Central Tendency: e.g., mean, median
- Measures of Spread: e.g., range, SD, IQR
- Checks and Identification of Outliers: boxplots

Interpret and discuss what you can learn from the pictures (e.g., compare shape…) and make sure you report sample sizes as well.

IV. **Statistical inference** (with justification for procedures chosen)

*For at least 2 interesting questions*, state and interpret both hypothesis pairs and/or define parameter to be estimated through confidence interval. I recommend starting with a test of significance and if it is significant, calculate a follow-up confidence interval.

**Quantitative data:**
- e.g., one sample $t$ procedures, two sample $t$ procedures, regression, ANOVA

**Categorical data:**
- e.g., one sample $z$ procedures, two sample $z$ procedures, chi-square test

**Include discussion of why your method was appropriate**

**Diagnostics**
- Verification of any technical conditions made during statistical inference
- Suggestions if technical conditions aren’t met

Summarize in your own words what these procedures reveal.

V. **Interpretation/Explanation of Results**

What does it all mean?

Use the above summaries to justify your interpretation
- Back up your conclusions with interpretations of the graphical summaries as well.
- Suggest reasons for what you’ve observed
  - e.g., why do you think these groups are different? or not different?

VI. **Overall Conclusions**

Review of what analyses revealed and any potential problems with your study.

Recommendations based on your analyses.

Future Questions for someone conducting a similar study in this area and/or discussion of what you could have done differently with your project to strengthen your conclusions