

A statistical project will count for 10% of your final grade. Here are the parameters of the project.

Working alone or in a group.— You may do the project *alone*, or in a group of *two or three students*. You may not have a larger group. If you plan to work in a group, you may choose your own groups, but all must agree to work together. Groups must be formed by the time that the *plan* for the report is due on April 18. If you work in a group, each student must do the *oral portion of the project individually*. In addition, each student will need to turn in a separate short document that answers some questions about how the group worked together which will help me decide if all of the group deserve equal credit for the grading of the project. My expectation is to give equal scores to all students, but I will not do so if there is evidence that one or two students did not contribute fairly.

When working in a group, it is essential that each person makes a significant contribution to the work and that all agree that the work is shared fairly: this does not mean that every task must be split equally, but it does mean that roughly equal time and effort should be contributed by each group member.

Project plan.— On Friday, April 18, a project plan is due. The project plan must include the following elements:

1. A brief statement of the question you wish to address, including, if appropriate, a definition of one or more populations of interest.
2. A brief description on how you will gather data including a detailed sampling plan or randomization plan, if appropriate, and a description of what the individual *cases* will be (see Chapter 1 for review).
3. A list of variables you will measure, how you will measure them, and the role these variables will play in your analysis (are these variables explanatory or response variables, discrete or continuous)?
4. A brief summary of the statistical method or methods you plan to use for inference.
5. If working in a group, a signed statement that all will contribute fairly to the project.

The project plan should be about one-page in length.

Project scope.— For your project, you must *gather your own data*. It is not acceptable to use a data set someone else has collected. Your project should be on a topic that *you find interesting*. Your project must be either: (1) an *observational study* where you define one or more populations and have a plan to take random samples from population(s) of interest, or, if not practical, a plan to take a representative sample; (2) a *randomized experiment*, where you compare at least two groups, and you use randomization, when possible, to assign

treatments or divide cases into groups; or (3) a *simulation study* where you compare two or more statistical methods from the course in multiple settings. In addition, your project must include at least one *inference question* where you will use a confidence interval, a hypothesis test, or both, to draw a conclusion.

It is not important that your question be of broad scientific interest. It is important that you take care to gather data that can address your question of interest, that you use appropriate methods from the course, that you interpret the results of your methods correctly, and that you can communicate what you have learned effectively.

There is no minimum sample size, but you must gather an adequate amount of data for it to be appropriate to use methods from the course. If using human subjects, you must obtain *informed consent* from each individual. Your project cannot require approval from a campus board to ensure safety of human or animal subjects, or ethical considerations. You must not break any laws in gathering data.

Written report.— Each group (or individual working alone) must turn in a single report by the beginning of the final examination on May 12. You should also have a draft of the report finished when you defend the report orally, but you may use feedback from the oral defense to edit and improve the final written report. The written report *cannot exceed 5 pages in total*. The format of the report is somewhat flexible, but must include the following items:

1. An *introduction* that describes the problem of interest, the question or questions you address, and other background information needed to understand what you intend to do;
2. A *methods section* in which you describe how you gathered data and what statistical methods you used to analyze the data you gathered;
3. A *results section* in which you state the results of your analysis and discuss and interpret in context what the results mean and what you have learned about your question of interest.
4. At least one *graph* which displays the collected data in a manner that is informative about the primary question(s) of interest.

Oral defense.— You will sign up individually for a ten-minute oral defense of your project. You should bring a draft of the written report with you (groups defend separately, but should use the same written report). You do not need to prepare a presentation, but each individual should be familiar with all aspects of the project. I will ask questions about how you gathered data, what methods of analysis you used, and how well you understand what you did and why you did it.

Project ideas.— The idea for your project is open-ended as long as it meets the criteria above. There are project ideas in each unit summary: pages 155–157, for Unit A, pages 317–319 for Unit B, pages 457–458 for Unit C, and 616–620 for Unit D (which we will begin to study after the next exam). Simulation examples could be comparing properties of

different methods for confidence intervals, or comparing the power of alternative methods for hypothesis testing, or considering the ramifications when model assumptions fail. Gathering data can involve finding information on the web, but you cannot simply find a complete data set. Many of the data sets from the textbook would be appropriate, if you picked your own topics and found corresponding data.

Grading.— The project will be graded on a ten-point scale. I will use the following rubric:

1. 20% Overall clarity and effectiveness of writing;
2. 20% Student understanding during the oral defense;
3. 15% Effectiveness of the introduction;
4. 15% Appropriateness of statistical methods and the correctness of their use;
5. 15% Appropriateness and correctness of discussion and interpretation of results;
6. 15% Effectiveness of the graph(s) and displays of data.