Stats
Steve Sawin

Guidelines For Group Projects

Ultimately each group will turn in a written project (typically 5-8 pages, including charts and graphs) on the last day of class and will do an oral presentation to the class (typically 8-10 minutes) in the last few days of the semester.

Groups and Topics

You will pick your group yourself (2-5 people unless you have permission otherwise) and will choose your topic in conjunction with me. My hope is that you will choose a subject that is of interest to you (perhaps related to your major, future career plans, hobby, passion, calling, or something somebody once said at a party), and ask questions to which you really want to know the answer. I will work with you to make the project reasonable, and my grading takes into account the varying difficulty of different projects, so don't worry about the project's being too hard or easy.

Here are the conditions. You need to look at how one variable (explanatory) affects another (response) in a population of some sort of individuals. You will have to gather some data and do descriptive and inferential statistics to explore the relationship between these variables. Emphasis is on not only doing the calculations, but analyzing the process of data gathering the design of the study to identify weaknesses that make the interpretation subtler. You should read the project rubric carefully when writing it up/presenting to the class, because that makes it extremely clear what I am looking for.

Difficulty and Grading

Different projects can vary considerably in the difficulty of the statistics involved and of the sampling and data gathering. My grading takes this into account. Your choice of variables determines the statistical procedure used, which affects the difficulty of the the statistics involved (also, some techniques we will learn so late in the semester that the short time frame for writing it up presents an additional challenge). The total possible points your project can earn depends on the procedure used. If you are using a simpler procedure and a typical sampling method, your total possible points may be less than 100. On the other had if you use a sophisticated procedure and a challenging sampling method (sampling people in downtown Bridgeport, setting up games of ping pong and recording the score, combing town records

of house sales) your total possible could be more than 100. The project rubric, which you should read carefully before you write up or present this project, offers a number of additional steps you can take that would earn additional points.

Steps

Early in the semester I have a deadline for "Group Members and Name." By this date you should decide who is in your group and make up a name for the group. But don't tell me in class or something because I will forget! Instead, download the planner at `faculty.fairfield.edu/ssawin/stats/planner.xls`. Open it in Excel and enter your names and a group name in the top (whimsy within the bounds of good taste is fine here. Please put your group name in all emails to me about the project, because I have a hard time keeping track of who is in which group otherwise!). Then email it to me. If you have an idea for the population you want to consider and the variables you want to relate, you can fill out the rest of the planner at the same time, or discuss it in the email with me. If you do not pick a group (or pick one that is too small) I will pick one for you.

Shortly afterwards is the deadline for "Initial Project Plans." By this date you should send the same planner with the two variables you want to consider, what type of variable they are (categorical with two values, categorical with more than two values, or numerical), the population and your plan for sampling all filled out. Notice this will determine the statistical procedure (which shows in green) you will use. You can come up with this in consultation with me or on your own. Since we will not have learned how to do the statistical analysis yet, your initial ideas may not be fully worked out or may not be practical, that is OK. I will reply with suggestions for how to amend it, and together we will work out your group's plan by the third date, "Final Project Plans." This final version of the plan contains the grading rubric I will fill out in the end, so you can see exactly how you will be graded.

After this you can begin gathering your data. You will take a sample, measuring the variables for each individual in the sample. That should be completed by at least two weeks before the end of the semester. Around that time through the end of the semester we will learn the statistical techniques to analyze the data.

The last three or four days of class are devoted do group presentations. Each group gives an 8-10 minute presentation to the class outlining their project, population, variables, and hypothesis, summarizing the data, detailing their inference calculations and conclusions, and suggesting potential problems with the study. Most people do powerpoint presentations with the group members each presenting a piece, and either bring a laptop or download the powerpoint presentation from their email (those who do both never have technical difficulty, just saying!). On the last day of class the written projects are due, and I will grade them by the final.

Of course you do not have the resources, time or knowledge to do a statistical study the way the pros do, and that is fine. Many of you will not find significant evidence for the hypothesis you were considering, and in almost all cases their will be serious flaw (biases, lurking variables, failures of the assumptions) that will make your conclusions less certain. That is OK, and in fact one of the most important components of the project is practicing recognizing these issues in real situations. You can see in the rubric that much of the points are based on identifying such issues. You will find that as you think about these questions deeply in your case, and then watch the other presentations discuss the issues that come up in their study design, you will master the skill of think critically about the underlying logic of statistical reasoning that you will see in your work and your life as an informed citizen. That may be the most important skill I can offer you, and this project is where that materializes.