

**What is the project?** The final project will be a poster display on any reasonable topic involving data analysis or statistics. Posters will be displayed in the style of a professional conference, scheduled for Wednesday, December 10 (during “Reading Period”). Projects are to be done in groups of three or four students. A proposal (worth 10% of your project grade) is due on Wednesday, November 12. A 15-minute meeting with me (worth 10% of your project grade) is required during November 12–19. Late projects are not accepted.

**What are possible types of projects?** A project may be one of the following:

(1) **A research report** on a topic that interests you—perhaps a topic we have touched on in class, or something you learned about in another class or through your own reading or experience. Such a project might involve reading articles in journals, newspapers, magazines, and books; summarizing and discussing the statistical issues involved; and drawing conclusions. Examples of possible topics include the use of DNA fingerprinting in court cases, the use of statistics in standardized testing, the probabilities involved in gambling, the use of election polls in the media, the evidence for linking second-hand smoke and cancer, advanced methods of statistical graphics, the use of regression and forecasting for the stock market, etc.

(2) **An original survey or experiment.** You might conduct a survey of Swarthmore students on some topic that interests you. Such a project might involve developing a survey or set of questions, selecting a simple random sample as carefully as possible, and gathering the information. For example, one group from a previous year surveyed students on their favorite breakfast cereals and then checked if their claimed preferences matched quantities actually consumed in dining halls. Or you might design and carry out some experiment or taste test: can people distinguish between day-old and fresh-baked goods, will people respond differently to a news story in which the subject is implied to be of one race as opposed to another race, is one style of Frisbee throwing more accurate than another, etc.

(3) **An analysis of an existing dataset.** You might apply some technique we have learned in the course to a dataset that has already been collected. For example, the WWW is a vast source of datasets on almost any subject, such as demographics, disease, economics, geography, entertainment, science, etc. Some interesting sites include [www.baseball-reference.com](http://www.baseball-reference.com) for baseball data, [www.census.gov](http://www.census.gov) and [www.city-data.com](http://www.city-data.com) for demographic data, [www.leesmovieinfo.net](http://www.leesmovieinfo.net) for movie earnings, etc. You could also critique or expand upon a dataset you have seen in the media.

If you have another project idea that does not fall into one of these categories, feel free to talk to me. I encourage interesting and innovative ideas!

**What kind of data should we collect?** At an absolute minimum, you should collect at least one variable that is numerical and one that is categorical (e.g., gender, class year, major). Typically people doing surveys will collect 5–10 variables on each subject. Most projects should have at least one confidence interval and one hypothesis test and one graph.

**What should the poster include?** The poster display should include the following components:

(1) **A statement of the question or purpose.** What problems or questions did you set out to investigate? What were the key issues raised?

(2) **The background and preparation** for conducting the project. How did you prepare for the project? What sources or background readings did you consult? What information did you use in developing your ideas from the conceptual stage to the finished project?

(3) **Methodology.** What did you do, and how did you do it? How did you gather information using a survey, experiment, or other method? If you did a survey, how did you try to take a

random sample? What statistical techniques did you use – for instance, scatterplots, correlation, hypothesis tests, confidence intervals, regression, ANOVA?

(4) **Results and conclusions**, the summary and presentation of your data analyses. What did you find out? This might include tables, graphs, or verbal summaries. What did you learn about the problem or question you set out to investigate?

(5) **Discussion and critique**. What did you learn about the process of carrying out your project? What went wrong, and how could you improve it next time? For instance, did any sources of bias creep into your survey or experiment? What advice would you give future students?

Not all points will be relevant to all projects. For example, research reports may not have any “results” as such.

The content of your project should be equivalent to roughly that of a 10 page paper. (You do *not* need to hand in a paper.) You should use large print on your poster and aim for an attractive and well-organized display. Try to think of a catchy title that captures the spirit of your project.

**How will the project be graded?** All group members receive the same grade; the project counts for 18% of your course grade. The following three areas will receive roughly equal weight:

(1) **Presentation:** Is the poster attractively presented and well organized? How well does it address the five points listed above (if relevant to the project)?

(2) **Statistical content:** Does the project incorporate concepts and methods from this course? Are they used appropriately and correctly?

(3) **Subject matter content:** Did you learn something about the subject of the project? Is the project interesting, fun, provocative, compelling, educational, fascinating, etc.?

**How do I find a project partner?** Groups should have three or four people. If you’re looking for partners, send me an email with a short description of your interests and contact info, and I’ll post it on the project web site.

### **What are the deadlines?**

*5:00 pm, Wednesday, November 12:* Email me your project proposal at [scwang@swarthmore.edu](mailto:scwang@swarthmore.edu).

Who is in your group? What is your topic? What issues or questions are you addressing?

How will you collect the data? For instance, what is your sampling method, how many people are you surveying, what questions are you asking? Or, how are you designing your experiment, how are you randomizing, how are you avoiding bias? What statistical techniques are you using? Be as specific as possible; your email should be about a paragraph or two.

**IMPORTANT:** In your email, put **Stat 1 Project** in your **Subject:** line. Please do not email other questions in project email; send those separately. Please send your proposal as a plain email, **not as an attachment**. Completing the project proposal email on time is worth 10% of your project grade, with a 5% deduction for each day or partial day late. *Please follow these instructions to earn full credit for your project proposal.*

*Wednesday, November 12 through Wednesday, November 19:* Sign up for a 15-minute time slot to talk to me to get feedback on your proposal. The meeting is worth 10% of your project grade.

You may also meet with me earlier, if you’d like.

*Wednesday, December 10:* Project day: posters due. *No late posters accepted!*

**A final word:** I encourage innovation, creativity, and humor in your projects. I want you to have fun with this assignment, and to create a piece of work that you will be proud of.