

THE FINAL PROJECT

WHAT IS THE PROJECT? The final project will be a paper on any topic in statistics, data analysis, or graphics that is related in any way to the course. Projects are to be done in groups of two to four students. A proposal (worth 10% of your project grade) is due at 5:00 PM on Thursday, November 18; the finished project is due at 5:00 PM on Friday, December 17. Late projects will not be accepted except in case of a serious and unforeseeable emergency.

WHAT TYPES OF PROJECTS ARE POSSIBLE? 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, something mentioned in the book but not discussed in class, or something you learned about through your own reading or experience. Such a project might involve reading articles in journals or books, learning about the subject in greater detail than was covered in the course; summarizing and discussing the statistical issues involved; and applying the methodology to a real dataset. For example, you might explore nonlinear regression models and their applications, strengths and weaknesses of new high-dimensional visualization methods, etc.

(2) **An analysis of an existing dataset.** For example, the WWW is a vast source of datasets on almost any subject, such as disease, economics, geography, entertainment, science, etc. Sports is a good source for datasets; try, for instance, www.baseball-reference.com. You can also find impressive sources for demographic data; try, for instance, www.city-data.com or www.census.gov. You could also criticize or expand upon an article or dataset you have seen in the media. Or you might have a dataset you have come across in your own research.

(3) 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 SHOULD THE PROJECT INCLUDE? In general, your project should include the following sections. Every project will be different, of course, and some of these sections may not apply to some topics. These sections are listed as a general guideline.

(1) **A explanation of the problem or purpose.** What problems or questions did you set out to investigate? What are the key issues raised? How were the data collected?

(2) **Background or literature review.** What sources or background readings did you consult? What literature exists on your topic?

(3) **Methodology.** What did you do, and how did you do it? What statistical and graphical techniques did you use?

(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.

(5) **Discussion and critique.** What did you learn about the problem or question you set out to investigate? What were weaknesses and strengths of your analysis? If you had more time or resources, how could your project be improved?

Your project should be roughly 10 pages long. This is meant to give you an idea of the scope of the assignment, rather than to prescribe minimum or maximum limits on your work. What's important is the completeness and appropriateness of your analysis, rather than its length.

HOW WILL THE PROJECT BE GRADED? All people in a group will receive the same grade. The project counts for 14% of your course grade. Your grade will be based on the following three areas:

(A) Presentation (roughly 25%): Are the analyses clearly presented and well organized? How well does it address the five points listed above (if relevant to the project)? Is the text well-written? Are the graphs appropriate and informative?

(B) Statistical content (roughly 50%): Does the project incorporate relevant concepts and methods? Are they used appropriately and correctly?

(C) Subject matter content (roughly 25%): Does the reader learn something about the subject of the project? Is the project interesting, fun, provocative, compelling, educational, fascinating, etc.?

WHAT ARE THE DEADLINES?

5:00 PM, Thursday, November 18: Email me your project proposal addressing the following questions:

- Who is in your project group?
- What is your topic? What issues or questions are you addressing?
- How will you collect or obtain the data?
- What statistical techniques are you using? What kind of analyses do you plan to do?

Be as specific as possible; your email should be about one or two paragraphs. You should then sign up to meet with me briefly (15 minutes should suffice) so I can give you feedback on your proposal.

Important: Please read and follow these directions. In your project proposal email, put Stat 31 project in your Subject: line. Please do not email other questions in project email; send those in separate messages. Also, please do not send attachments; just send your proposal as plain text in the email. Completion of the project proposal email on time is worth 10% of your project grade, with a 5% deduction for each day or partial day late.

5:00 PM, Friday, December 17: The final project is due. Late projects will not be accepted except in case of a serious and unforeseeable emergency. Feel free to hand in the project earlier than the deadline.

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.