Data Analysis Project

The project for this class will consist of a statistical analysis of one of the data sets provided or one of your choice. You should begin by reading the descriptions of the data sets, and looking at the sites listed. By May 24, 2002, you should turn in a project topic statement. This entails a one-page written statement about the data set you plan to use and the research questions you are going to analyze using the data. (If using a data set other than those provided, make sure you ”ok” it with me first.) You should make reference to any preliminary analysis you have done that shows that the data set is relevant to your research questions. I will review the project topic statements and will offer suggestions. However, I may ask you to revise and resubmit your project statement. It is important for you to have a clear understanding of what questions you plan analyze before beginning to write you report.

Once you have a clear idea of the of your research question, you should apply what you have learned about descriptive statistics, graphical methods, normal approximations, correlation and regression, and hypothesis testing to your data set. You must use JMP IN for this part of your project. You need to hand in pertinent computer output and a written description of the statistics you compute. This does not mean handing in formulas, but rather an interpretation of what you have found. The goal is not to do an exhaustive data analysis, i.e. do not calculate every statistic and procedure you have learned for every variable, but rather let me know that you are proficient at using JMP IN at a basic level, and that you are proficient at interpreting and presenting the results. Focus on methods that help you begin to answer your research questions. Also pay attention to your presentation. Neatness, coherency, and clarity will count.

A one to two page conclusion and discussion is required. This will require a summary of what you have learned about your research questions along with statistical arguments supporting your conclusions. Also critique your own methods and provide suggestions for improving your analysis. Issues pertaining to the appropriateness of the data set and statistical analysis should be discussed here. A paragraph on what you would do differently if you were able to start over with the project and of what you would do next if you were going to continue to work on the project should be included. The project should be 6-8 pages in length. Your final project will be due June 21, 2002, in class. (There will be no extensions of this deadline.)

Grading: Grading of the project will take into account the following areas:

Content: What is the quality of research and/or policy questions and relevancy of data to those questions?

Correctness: Are the statistical procedures carried out and explained correctly?

Writing and Presentation: What is the quality of the statistical presentation, writing and explanations?
Creativity: Is the project carefully thought out? Are the limitations carefully considered? Does it appear that time and effort went into the planning and implementation of the project?

Students may consult with one another, but each person must hand in an individual project which represents their own knowledge and skills. You should attend at least one office hour to ask questions and to discuss your project with me. In addition, you will need to sign up for a 5-10 minute slot during lab to give a brief presentation of your project to your classmates.

The entire data analysis project will account for 300 points or 15% of your final grade, with the following distribution of points:

Project Proposal/Description: 30 points

Office Hour: 20 points

Class Presentation: 50 points

Paper: 200 points