BIOS 663: Intermediate Linear Models
Data Analysis Project

Important Dates:

- Monday April 8: Project proposals due in class

- Monday April 22 and Wednesday April 24: Short presentations in class

- Monday April 29: Project report (which should be no more than 4 pages) due by 5pm.

Data Sets: You may choose to work on the following data or on your own data:

1. Plasma Retinol data: `http://lib.stat.cmu.edu/datasets/Plasma_Retinol`

2. Wage data: `http://lib.stat.cmu.edu/datasets/CPS_85_Wages`.

3. Body Fat data: `http://lib.stat.cmu.edu/datasets/bodyfat`.

4. Boston House Price data: `http://lib.stat.cmu.edu/datasets/boston`.

5. Air Pollution data: `http://lib.stat.cmu.edu/datasets/NO2.dat`.

6. Football data: `http://lib.stat.cmu.edu/datasets/profb`

Project Proposal: The project proposal should include the names of the group members (up to 5 members per group) and two data sets that your group would like to work on. Please rank the two data sets according to your preference, based on which I'll assign one data set to each group to ensure that at least two groups will work on the same data set provided. Groups that work on the same data set are not allowed to communicate with each other and should perform their analysis independently.

Short Presentations: Each group will give a $\sim$ 10 minute presentation which covers

- description of the problem and data of interest

- description of the analysis and justification of the analysis

- results (including relevant descriptive statistics, tables, and graphs)

- conclusions and discussions

Each group will be graded by other groups after their presentation, based on the appropriateness of statistical methods used for analysis, effectiveness in communicating results, and clarity of the presentation. Final grade will be a weighted score of my grade (50%) and the other groups' grades (50%).