Statistics 411/511
Homework 5
Due in Lab Monday and Tuesday, October 31 and November 1

- **Homework Guidelines** Homework must be typed and stapled. Pease label your paper with your name and lab time. See the syllabus for more guidelines.
- **Academic Integrity** You are encouraged to *discuss* the homework with other students, but what you turn in must be your own work in your own words. The syllabus contains more details and links to OSU's Student Conduct Code.

1. Read through problem 24, pages 146-7, then do the following.
   (a) Check boxplots of the data, and compare to boxplots of the log-transformed data. Turn in both sets of boxplots.
   (b) Perform a one-way ANOVA of the data. As the data analyst, you must decide whether or not to use a log transformation. Either will be acceptable, but you'll need to provide a short (one- or two-sentence) justification of your decision below. Turn in an ANOVA table (copy and paste from R output OK).
   (c) Answer the research question posed in the last sentence in problem 24. If you analyze the log-transformed data, you'll estimate "by what percent does the distribution of 2005 incomes..." whereas if you analyze the data on the original scale, you'll be estimating "by how many dollars..." Produce a "statistical conclusion" and "scope of inference" for this problem. In the scope of inference, include the brief justification of the log transformation (or lack thereof).
   (d) Mimic the `plot()` command on page 9 of Outline 5 to obtain a residual plot after your analysis of the data in part (b). Comment on the symmetry of each group's residuals about 0. Is the spread of the residuals approximately equal for all four groups?