The exam is scheduled for Friday, 6:30 PM in the Mac Lab. The exam will be approximately twice as long as a test but you can have all three hours to complete it. As usual it will have a closed book portion and an open book portion. On the open book portion, you may use your book, your notes, and R. However you may not use the internet or other resources. The test will consist of the following type of questions.

1. There will be questions that are repeated from the first three tests. These might be repeated verbatim or be somewhat changed (e.g., same question, different numbers).

2. There will be questions (on the closed book portion) about Chapter 8, the material covered since the last test.

3. There will be some data to analyze (much like the last test). There will again be three such questions, the data and stories for which are below.

4. There may be a couple of questions that don’t fit any of the above categories. For example, there probably will be a relatively open-ended question that will address your understanding of a fundamental concept of the course and that you will need to use R to answer completely.

Data Stories

The data are available by using `getdata()` from the course package. Otherwise get the data from http://www.calvin.edu/~stob/data/exam1.csv

1. The dataframe exam1 contains data on the air discharged from a water treatment facility. The air is passed through a membrane that causes the contaminants to dissolve in water. The variable removal contains the percentage of contaminants removed from each of several samples of air. The question is whether the temperature of the air at the inlet to the membrane has some effect on the efficiency of the removal process. The temperature (in degrees centigrade) of each of the samples is also in the dataframe.

2. The dataframe exam2 contain results of a survey of a random sample of commuters in a large metropolitan area. These drivers were asked (among other things) about the size of their car and the length of their commute. The counts of drivers in each category are in the dataframe.

3. The dataframe exam3 contains data on an experiment performed on a number of office workers. Researchers noted that office workers whose arms are elevated a large percentage of the time that they are working have a greater chance of injury. The researchers measured the percentage of time that the office workers had their arms elevated above 30°. The job conditions were then changed over the course of several months and then the measurement was taken again. Of course the researchers were hoping that their changes reduced the percentage of time that the workers needed to elevate their arms.