**DATA 557**

**Homework Assignment 5**

**Instructions**

Submit your written solutions in pdf format to the dropbox on the canvas page by **5:00, Wednesday February 13.**

For Question 1, work in a group (either the same group that you worked with during class or a different group). One member of the group should include the group’s solution to Question 1 in their HW 5 submission and list all the members of the group.

For question 2, you should submit your own work; as usual, you may work together to solve problems, but you should submit your own work for Question 2.

You do not need to submit anything for Question 3. This question is designed to get you started thinking about your final project.

It is not necessary to include R code with your solutions.

**Question 1**

(This is a continuation of the work on the “process.csv” data from class)

Compare the mean number of defects for the 4 methods. Use the ANOVA F-test to test the null hypothesis that the mean number of defects per ball bearing is the same for all 4 methods. What is the conclusion from the test? Assess the assumptions of the ANOVA F-test. Is the F-test valid for this study? Verify your answer (i.e, whether or not the F-test is valid) using an appropriate simulation study.

**Question 2**

A researcher is designing an experiment to compare the mean life length of three different types of cell-phone batteries (A, B, and C). The null hypothesis to be tested is that the mean life length is the same for all three battery types. Based on previous experiments of the life length of similar batteries it is expected that life length has an exponential distribution with mean approximately 3 days. There are two alternative hypotheses of particular interest. For the first alternative hypothesis, the mean life-length is 3 days for A and B, but only 2 days for C. For the second alternative hypothesis of interest, the mean life-lengths are 2, 3, and 4 days for A, B, and C, respectively. Your assignment is to design the experiment, i.e., determine appropriate sample sizes for each battery type, in such a way that the following two requirements are met: 1) the ANOVA F-test is a valid test of the null hypothesis with type I error rate of approximately 0.05, 2) power of 90% to detect the first alternative hypothesis of interest, and 3) power of 90% to detect the second alternative hypothesis of interest.

**Question 3**

Develop a possible idea for a final project to discuss with your group in class on Feb 13. Propose a possible topic, research question or hypothesis to investigate, and explore possible sources of data for this project. Be prepared to discuss with your group in class on Feb 13.