**Exam #1 Review**

Today’s lab is intended to help you think through the concepts we’ve covered thus far. What I want you to do is work in groups at your table to answer the question below corresponding to your table number (I’ll tell you your #). I give vague notions of what I want, and you need to perform a thorough analysis with this end in mind. You will each get about 20 minutes to work on these in your groups, at which point you will all need to present your results to the class. Try to keep your presentation to ~5 minutes, allowing for some questions. You can choose a designated speaker, or work together, but you should all be able to answer any questions I might have, and I will feel free to call on any member at any time to answer my questions. You should all also feel free to ask questions of the other presenting groups.

In addition to the data you pulled from GitHub, there is a document giving some description of what each variable measures. In any of the following questions, you might encounter little idiosyncrasies with the data – if you do, note how you handled them. If you alter the data in any way (replace an obvious typo of “nos” by “no”, for example), make a note and report this.

Q1: Look at variables #7, 10, 11, 17, 22, 35, and 49. Create appropriate graphical summaries of these variables, and report on any notable features. What have you learned? Give us some context. If the data in one of these variables is approximately bell-shaped, tell us the intervals that contain 68%, 95% and 99.7% of the data.

Q2: Look at variables 22 and 24. See if they are related; show us an appropriate graphical summary. You may also want to look at this relationship separately by gender (var. 1) to see what, if any, effect that has. Tell us what you have learned.

Q3: Look at variables 45 and 54. See if they are related; show us an appropriate graphical summary. You may also want to look at this relationship separately by Greek (var. 4) to see what, if any, effect that has. Tell us what you have learned.

Q4: Look at variables 4 and 6. See if they are related. Create an appropriate numerical summary of the data, and give some intuition as to whether or not they are related using this summary. Be careful which variable is placed where in this summary. Finally, do a hypothesis test about the relationship and report, in terms of the problem, what we can conclude and what we have learned.

Q5: Look at variables 46 and 47. See if they are related. Create an appropriate numerical summary of the data, and give some intuition as to whether or not they are related using this summary. Be careful which variable is placed where in this summary. Finally, do a hypothesis test about the relationship and report, in terms of the problem, what we can conclude and what we have learned.

Q6: Look at variables 46 and 47. Calculate conservative 95% CIs for the proportion who said “yes” to #47 separately for those who said “no” and “yes” to #46. Tell us what we learn from this analysis. Make some conclusions.

Do the same for variables 4 and 6. Calculate conservative 95% CIs for the proportion who responded in each category of #6 separately for those who said “no” and “yes” to #4. Tell us what we learn from this analysis. Make some conclusions.