Portfolio 3
PA 606: Seminar in Quantitative Techniques (Spring 2020)

25 points

Due: April 25, 2020

Topics:

Analysis of Variance (F)

Data Sources:

`GSS2014_final.csv` (**data set**: 2014 General Social Survey – GSS)
`GSS_Codebook.pdf` (**codebook**: description of 2014 GSS and measurement of variables)
`GSS_Variables_and_Descriptions.pdf` (**descriptions**: names and titles for 2014 GSS)

Overview:

Answer each of the questions below, **in full sentences/paragraphs**, and show your tables/plots (if necessary) inline (interspersed within the text). For full credit, you must append (copy and paste) your R Script at the end of this portfolio as the final page(s).

Problems

Data Set Information

- Describe the data set, including the name and who administered the data/survey (e.g. which survey research firm), year administered, the unit of analysis, and the number of observations.

ANOVA

In this test, you'll be examining mean differences in **occupational prestige** of a respondent's job (`sei10`) by their **type of occupation** (`workfor1`).

- Describe the variables. Using the codebook (`GSS_Codebook.pdf`) and the list of variable descriptions (`GSS_Variables_and_Descriptions.pdf`) for the `GSS2014_final.csv` data set, for each variable, describe its text (e.g. the question asked for each variable), the level of measurement, and the values/categories within the variable.
- Define a research question for the variables (e.g. "*Is there a mean difference in Y by categories of X*").
- Define the null hypothesis (H_0) and the alternative hypothesis (H_1) for this test.

- Describe the various assumptions of the ANOVA (F) and how you would assess them. Next, show all necessary tables/plots that demonstrate your assessment of whether or not you've met the assumptions of the test. (*Note: if showing tables, create them, do not simply copy and paste from your output.*)
- Run the test. Fully and correctly report the test. Whether or not you find significance, follow your results with a means plot to demonstrate the comparison of the means, and describe what you find in terms of mean differences by group (e.g. compare the group means).
- Describe when and why you would run a post-hoc test?

Extra Credit

- Run a post-hoc test (specifically, Tukey's HSD) and describe your findings.