**Applied Data Analysis Project** (50 points)

In this project, you will use data from the 2018 *General Social Survey* (GSS) to answer a question about criminal justice in the real world. The 2018 GSS data is a sample of 2,348 adults living in the United States who answered survey questions about their attitudes and experiences.

The GSS is a nationally representative survey of U.S. residents. This means that findings from the GSS can be generalized to the U.S. population. This means that when you do a hypothesis test using GSS data, you can test hypotheses about the U.S. as a whole.

In this project, you will use the 2018 GSS to answer a research question. Please choose ONE of the options below, then follow the instructions to write a short report that will answer the research question using the data.

**Research Question Options**

**Option #1: Do men and women have the same amount of trust in American courts?**

**Option #2: Do people identifying as White, Black, and “other” races have the same amount of trust in American courts?**

**Option #3: Is there a relationship between years of education and trust in American courts?**

**Option #4: Is there a relationship between supporting police use of force and supporting the death penalty?**

For each option, Stata output showing descriptive statistics for each variable and the results of the appropriate hypothesis test (t test, ANOVA, chi square test, or correlation) will be provided.

**Report**

Your report should consist of the following sections. Each one should be separated by a heading. Each section can be answered using 1-2 short paragraphs.

**Research question (5 points).** Briefly describe the research question you selected. Describe how you think the variables may be related. (*NOTE: You do NOT need to cite any sources—just describe what you think might be the case.*)

**Variables (10 points).** Describe the variables in your research question, including what the variable attributes are and how the variables are coded. State which you think is the independent variable (cause) and which you think is the dependent variable (effect). State what the level of measurement is for both variables.

**Descriptive statistics (15 points).** Provide *descriptive* *statistics* for the variables in your sample. For nominal and ordinal level variables, include frequency distributions and describe the relative frequencies. For interval or ratio level variables, describe the means and standard deviations. Explain what the descriptive statistics mean in words.

**Inferential statistics (15 points).** Answer your research question using the appropriate *hypothesis test* (alpha level =.05). State which hypothesis test is being presented (t test, ANOVA, chi square test, or correlation) and write out your null and alternative hypothesis. Report the test statistic (t, F, χ2, or r) and the p-value. Explain whether you reject the null hypothesis, as well as what the results of your hypothesis test tell you about how the variables are related in the U.S. population.

**Limitations (5 points).** Briefly describe one of the limitations of hypothesis testing (e.g., false positives, spuriousness) and how it may apply to your results.

In all, the research project should be about 2-3 pages in length, double spaced. You may copy and paste the Stata results into the paper to help you describe the results. Note that the report will be graded on content, NOT length.