POL S 411/511
Assignment 3
Total Marks: 16

Below, you will formulate a realistic, theoretically informed research question based on a relationship between two variables (an independent and dependent variable) in the dataset. You will use the **Comparative Study of Electoral Systems (CSES) dataset (module 5)** for this assignment. You will find the codebook and other documentation particularly useful for this.

Ensure your variables are either continuous or ordinal with at least 5 meaningful categories (NOT including "don't know", "other", etc), but no more than 15 (i.e. do not use the "age" variable). (1pt)

Write a do-file that (in the following order):

1. Contains the frontmatter for the do-file
    a. Contains a comment at the top of the file with your name, the course, date and instructor (each on separate lines)
    b. Sets the working directory and opens the dataset

2. Contains in comments (2pt):
    a. A strong, falsifiable research question
    b. The null hypothesis
    c. An alternative hypothesis that **explains why you think it is theoretically-justifiable**

3. Summarizes your two variables
    a. In comments, identify the mean, median, minimum, maximum, standard deviation and skewness (1pt)

For questions 4-6, make sure all your graphs are appropriately titled, labelled and use a helpful scheme. (1pt)

4. Create either a **bar graph** or **stacked bar graph** of all the different values of (i.e. not the means of) your dependent variable (1pt)
    a. Export this graph as a PDF file.

5. Create either a **bar graph** or **stacked bar graph** of all the different values of (i.e. not the means of) your independent variable (1pt)
    a. Export this graph as a PDF file.

6. Create either a **grouped bar graph** or **stacked bar graph** that shows all the values of your dependent variable by all the values of your independent variable (1pt)
    a. Export this graph as a PDF file.

7. Run a correlation matrix with significance test for your two variables (1pt).
    a. Explain in comments what you find. How strong is the relationship? What direction is it in? Is it statistically significant? (1pt)

8. Run a regression for your two variables (1pt).
    a. Explain in comments what you find. Is the relationship statistically significant? How do you interpret the findings of the regression? (2pts)

9. In comments, reflect on whether these findings favour rejecting or failing to reject the null hypothesis you identified in question 2 (1pt).

10. Draw a function/line in Stata (*help twoway function*) for the function corresponding to your regression line (1pt)
    a. Title the graph with the value of the function itself (i.e. “y = 0.5x + 10” – for example) (1pt)
    b. Export this graph as a PDF file.

Make sure to submit a log file containing the comments, code and output from above through eClass. **Make sure to submit a .log file, not a .smcl file, as well as the PDF graphs from your output above.**

To create a log file:

1. Start your log through the Stata menu (File → Log → Begin), making sure to select .log and not .smcl
2. Run your do file.
3. End your log through the Stata menu (File → Log → Close).
4. Open your .log file to verify the file has saved correctly with all of your output.