POL S 411/511
Assignment 2
Total Marks: 19

Using the Quality of Government Basic Dataset (Jan 2020 version - see Week 5 Canvas page for a link to the data and codebook):

Part 1 – 9 points

Formulate a realistic, theoretically informed research question based on a relationship between two variables (an independent and dependent variable) in the dataset. You will find the codebook particularly useful for this.

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

1. Contains a comment at the top of the file with your name, the course, date and instructor (each on separate lines)
2. Opens the dataset.
3. Contains in comments:
    a) A strong, falsifiable research question (1pt)
    b) The null hypothesis (1pt)
    c) An alternative hypothesis that explains **why you think it is theoretically-justifiable** (1pt)
4. *Describe* and run a *codebook* for your variables (1pt).
5. Summarizes your two variables
    a) In comments, identify the mean, median, minimum, maximum, standard deviation and skewness (1pt)
6. Creates a two-way table between the two variables that describes, in comments, what the individual cells mean (1pts)
    a. Create a new labelling system for your variables (1pts)
    b. Give me some examples – i.e. "there are X countries with A value of my IV and B value of my DV"
7. Comments on whether the data appear to support the null or alternative hypothesis (1pt)

Part 2 – 5 points

In the same do-file as for Part 1:

1. Change the scheme for graphs in Stata to something other than the default (1pt)
2. Create two histograms – one each for Freedom House's Freedom of the Net score and Freedom of the Press Score (hint: both variables should be out of 100) (1pt).
    a. Make sure the y axes display frequency and the x axes have an appropriate number of bins (1pt)
3. Create a scatterplot between these two variables. In comments, describe the relationship you find. What can you conclude about the causal relationship between these variables? (2pts)

Part 3 – 6 points

In the same do-file as for Parts 1 and 2:

1. In the QOG codebook, find, **summarize and describe** variable *pei_peii* and variable *qs_proff*. In comments, provide a brief explanation for what these variables represent and what their descriptive statistics mean (2pt).

2. Create a scatterplot between these two variables. Using the codebook and your answer to the previous question to guide your interpretation, explain in comments what you see. (2pts)
3. In comments, briefly (2-3 sentences) comment on the **causal** relationship between these two variables (2pts)

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.**

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.