**STATA Exercise**

*SOC 4881 – Population Studies Research Practicum*

**DUE**

*March 13*

**INSTRUCTIONS**

Complete this assignment in your group. Be sure to agree---early---on a division of labor within the group; make sure everyone has a meaningful role and has a way to learn from the assignment. It might be a good idea to meet as a group (in person or virtually) to make a plan.

Your task is to do very basic analyses of the full class dataset. The purpose of this exercise is (1) to get you more familiar with working in Stata and (2) to get you more familiar with your data and measures.

For all the questions below, use the full class data file and include all of the boys in the analysis.

1. ***Describe Two Outcome Variables***
2. Report the mean, standard deviation, and median of wages. Also report the number of cases with a valid value for wages.
3. Include a histogram that shows the distribution of wages.
4. Report the mean, standard deviation, and median of longevity (in days). Also report the number of cases with a valid value for longevity.
5. Include a histogram that shows the distribution of longevity (in days).
6. ***Describe a Name-Based Variables***
7. Report the mean, standard deviation, and median of “pct\_BLACK\_FIRST” — the percentage of en with each first name who are African American. Also report the number of cases with a valid value for pct\_BLACK\_FIRST.
8. Include a histogram that shows the distribution of pct\_BLACK\_FIRST
9. Now report the mean value of pct\_BLACK\_FIRST separately for whites and blacks.
10. How do you interpret the difference between the two means you reported in “c” above? What do those numbers tell you?
11. ***Do Some Basic Multivariate Analyses***
12. Report the correlations between three variables: wages, pct\_BLACK\_FIRST, and longevity (in days).
13. Ignoring the sibling structure of the data for now, regress longevity on wages. How do you interpret the coefficient representing the effect of wages on longevity?
14. Estimate a within-brother-pair regression of longevity on wages. Now how do you interpret the coefficient representing the effect of wages on longevity?
15. How do the estimates of the effects of wages on longevity differ between “b” and “c” above? What do you make of that difference … that is, how do you interpret it?

Turn in 2 files for this assignment:

1. The Stata code (aka, “do” file) used to produce all your results. (We will run it to make sure it does what you say it does!)
2. A Word file with short written answers to the questions above.

**GRADING**

This assignment is worth 6 points. Everyone in the group will receive the same score unless I receive (and verify) information that some in the group deserve more/less than others in the group. In those instances, I will communicate with everyone in the group to come to a fair resolution.

You will earn 2 points for each of questions 1, 2 and 3. For each, you will get full credit if your answer is complete and accurate; you will get 1 point if your answer is semi-complete and/or semi- accurate; and you will get 0 points for answers that are incomplete and inaccurate.

Assignments turned in late will receive half credit. Assignments should be emailed to me at warre046@umn.edu.