**POLS/GEOG** **418: Quantitative Methods**

Assistant Professor Joshua Zingher

Due June 27th, 2020

*Problem Set #7: Multivariate Regression*

In this problem set, we will use the 2016 ANES dataset (the same dataset we used on problem set four and five, available on Blackboard) to test two hypotheses about an individual’s orientations toward political protests. The data comes from the National Election Survey (NES) of 2016, which asked respondents about their opinions on a range of topics. For the following hypotheses, we will use the dataset’s “V162113” variable to measure a respondent’s view of the Black Lives Matter movement. This is a thermometer variable ranging from 0 to 100, with higher values indicating a more favorable view of BLM.

General note: make sure ***all*** tables, charts and graphs are labeled appropriately, so that an unfamiliar reader can interpret you variables and quantities or percentages you report in tabular and graphical form

*Ha1*: In comparing individuals, members of ethnic and racial minority groups will have a more favorable view of the Black Lives Matter (BLM) movement than whites even when controlling other competing explanations (use the “V162113” and “race” variables.)

1. Make sure the necessary variables contain only appropriate values. If necessary, replace Inap/Missing etc. equal to missing.
2. Generate a new dummy variable for each ethnic and racial minority group and add these dummy variables as independent variables in a regression with BLM attitudes as the DV. Are some minority groups more favorable towards BLM than others?
3. Add controls for the respondent’s level of education, income, gender, and region (note: there are multiple variables in the dataset that measure some these concepts in different ways. Be sure to explain which variables you chose and the rationale behind these choices). How does the addition of these control variables affect the original findings? Report the findings in the table and discuss the direction and magnitude of the effect that race has on attitudes towards BLM when controlling for these other factors.
4. Add an additional variable you feel might be important to the regression model from part ‘c’ (note: add this variable in addition to all of the variables you already included). Report the findings in the table and discuss the direction and magnitude of the effect that race and ethnicity has on attitudes towards BLM when controlling for these other factors.
5. Overall, what can you conclude about what effect race and ethnicity has on attitudes towards the BLM movement?
6. *Review*: For each hypothesis, please provide each of the following:
7. Set up the data to construct the necessary test (t-test or chi2). This might require you to generate a new variable or replace some of the values in an existing variable.
8. Use the correct tabular analytic technique: crosstabulation or means comparisons. Include your table with appropriate titles, labels and other relevant details.
9. Produce a bar chart or line graph, as appropriate to the hypothesis and data. Assess whether the distribution of your variables is skewed. Tell readers if there are any notable outliers.
10. Label all tables, charts and graphs appropriately, so that an unfamiliar reader can interpret you variables and quantities or percentages you report in tabular and graphical form.
11. Make a conclusion about the hypothesis. Does your analysis support the hypothesis? Does it refute it? Or is it inconclusive?

Chi2— “In comparing individuals, those who are under 50 years of age (age) are less likely to support legal abortion (V161232) than are those who are over 50.”

T-Test— “In comparing individuals, those who are Democrats will have less favorable assessments of Donald Trump (V161087) than Republicans”