Political Science 15 Homework #4
Due Friday, March 7th, by 5pm

This homework consists of three multiple-part questions, and is due Friday, March 7th, by 5pm (you can turn it in either in class, to my mailbox, to the TAs mailboxes, or in my office). Late homeworks will be penalized one letter grade per day (with homeworks received after 5pm that day counted as late). Due to the large enrollment in our class (approximately 90 people), we will not accept homeworks by email. Homeworks should be roughly 2 typed, double-spaced pages in length.

To complete this homework you’ll need to use SPSS. You will probably need to do this in the computer lab (the same place where we hold section each week). The data for this homework is on our class website. It consists of 9 variables observed in 561 voting precincts in San Francisco in the 2008 general election. The variables are (1) the precinct number, (2) the percentage of voters in that precinct that voted for Barack Obama, (3) the percentage of voters in that precinct that voted yes on Proposition 8, (4) the percentage of registered voters in that precinct that are male, (5) the percentage of registered voters in that precinct that are under the age of 35, (6) the percentage of registered voters in that precinct that are Hispanic, (7) the percentage of registered voters in that precinct that are Black, (8) the percentage of households in that precinct that are same-sex couples, and (9) the percentage of registered voters in that precinct that have incomes over $75,000.

Use a 5% significance level and two-tailed tests when testing the following hypotheses.

**Question 1:** Proposition 8 banned gay marriage in the state of California (a yes vote was a vote to ban gay marriage, a no vote did not want to institute such a ban). Support for this proposition was fairly evenly divided across the state, and Proposition 8 ultimately passes with 52% of the vote.

- Test the hypothesis that the mean vote percentage for Proposition 8 across these San Francisco precincts was 52%. Report the mean vote percentage for Proposition 8, the t score of the test, and the significance level of this test.
- Would you reject or accept (fail to reject) the null hypothesis that the mean vote percentage for Proposition 8 across these precincts was 52%? Why?

**Question 2:** Suppose you were interested in testing the hypothesis “Precincts with higher percentages of same-sex couple households voted for Barack Obama at higher rates.”

- Run a regression to test your hypothesis. What is the intercept of the regression line, and what does it tell you in this case? What is the slope of the regression line, and what does it tell you in this case?
• What is the R squared of this regression line? What does the R-squared of this regression line tell you?

• Based on these regression results, would you accept or reject this hypothesis? Explain why you made this decision.

• Suppose someone criticizes your hypothesis test, and argues that it is actually income that is driving these results. This person thinks that in San Francisco higher income households supported Barack Obama, and coincidentally same-sex couple households tend to have higher incomes than other types of households. Run another regression to test the hypothesis above, but this time include income as a control variable. Report the intercept, slope coefficients, and R-squared of this regression line.

• Did your original decision to accept or reject this hypothesis change once you included the control variable? Why or why not?

**Question 3:** Now generate your own original hypothesis about a relationship between two variables in this data.

• Run a regression to test your hypothesis. What is the intercept of the regression line, and what does it tell you in this case? What is the slope of the regression line, and what does it tell you in this case? What is the R-squared of this regression line, and what does this tell you in this case?

• Based on these regression results, would you accept or reject this hypothesis? Explain why you made this decision.

• Think of a plausible control variable that you might want to include in your regression. Run another regression to test the hypothesis above, but this time including this control variable. Report the intercept, slope coefficients, and R-squared of this regression line.

• Did your conclusion about your hypothesis change once you included the control variable? Why or why not?

• Print out your SPSS output for this question and attach it to your homework.