

PSC 200
Final Data Analysis

Due: Dec 17, 2003 by 5pm

On the course website, you will find the dataset **blackvotereg1960.sav**. The dataset contains variables on black and white voter registration in Southern counties in 1960, along with a number of demographic variables. Unless otherwise noted, assume the variables represent measurements taken in 1960.

Use this data and all of the statistical tools in your power to explain levels of black voter registration (**blackreg**) in the South during this turbulent period. In conducting your research, be sure to address how the following affected black voter registration in 1960:

- income
- education
- religion
- employment demographics
- political organizations
- racist obstacles to black voting (e.g. poll tax or literacy requirements)

Formulate at least five specific hypotheses. Begin by conducting exploratory analysis, using graphs, comparison of means, correlations, and/or bivariate regressions. To control for various factors, use multiple regression to test your hypotheses. Interpret the results, including both the statistical significance and the substantive effect of the variables in your model. Which hypotheses are confirmed? Which variables appear to have the largest affect on the level of black voter registration?

In your analysis, also control for the following three areas of the South: *Deep South* (Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina), *Border States* (Tennessee, Virginia), and *Texas*. You may need to use other dummy variables or interaction terms to test your hypotheses.

Finally, although there is no need to conduct a literature review, write your analysis as a research paper (see the course readings for examples). The paper should be 7–10 pages of text (plus any tables and graphs). It should be double-spaced in 11 or 12-pt font, with margins of exactly 1 inch on each side. All pages should be stapled.