

Quantitative Methods II

Assignment 5

September 25, 2011

This is the fifth homework assignment for the course. Its purpose is to continue increasing your proficiency in using a statistical program, producing presentation-worthy graphics, and writing to get your point across.

Remember that all statistics are an attempt to gather information about a process (or population) from a sample of data generated by the process. As such, we will never be able to answer a statistical question with absolute certainty, only with statements of confidence ranges and of expected Type I Error rates.

This assignment covers Linear Modeling — including testing all assumptions, selecting the best variable combination, and determining if the research model should be additive or interactive. Ask yourself how you will accomplish these goals. Sketch an outline of what you will need to do before you do it. Include graphs as appropriate. The final graph with the prediction curve(s) needs to be presentation quality. The other graphics only need to be utility quality.

When you hand in this assignment, you will email to me two separate files, your **typed solutions** to the questions asked in the homework and a separate **script file**. The script file allows me to check that you did the correct analysis. The solution file allows me to see that you can answer the questions in complete and coherent sentences, weaving in graphics and statistics appropriately. As always, the email needs to include the subject line:

POLS6123: Assignment 5

Note. *Make sure you include neither code nor raw results in the write-up. The code needs to be attached to the email in the separate script file.*

PROBLEM: WHAT IS THE PROBABILITY THAT IT WILL PASS?

[[25]]

The data file `ssm2` contains several variables:

- `state` (the name of the US state in which the vote took place),
- `year` (the year in which the vote took place),
- `myear` (the year after 2000 in which the vote took place),
- `pctFavor` (the percent of the vote in favor of the ballot measure),
- `propFavor` (the proportion of the vote in favor of the ballot measure),
- `civilunion` (that the ballot measure also banned single-sex civil unions),
- `south` (that the state is in the South),
- `religiosity` (an index, 1–5, representing the level of religious strength in the state),
- `gspcap` (the GSP per capita in the state),
- `popdensity` (population density in the state),
- `povertyRank` (the rank of the poverty rate in the state),
- `povertyRate` (the actual poverty rate in the state), and
- `hdi` (the human development index, a summary measure of human development in the state).

Your job is to predict the outcome of the Washington ballot measure using one dichotomous independent variable and one continuous independent variable — it is your choice as to which variables you use.

Remember, you have to decide which variable combination is best (and how you know) and whether there should be an interaction or not. You must explain in your solutions as to why you made your decisions.

Make sure you test the assumptions of Ordinary Least Squares. Make adjustments as necessary. Predict the actual vote in favor of the ballot measure and estimate the probability that it will pass.

The values for Washington are as follows:

Variable	Value
state	Washington
year	2009
myear	9
civilunion	1
south	0
religiosity	1
gspcap	41,751
popdensity	97.2
povertyRank	36
povertyRate	10.2
hdi	5.2