

Quantitative Methods II

Assignment 8

October 16, 2011

This is yet another homework assignment for the course (I lost count). Its purpose is to continue increasing your proficiency in producing presentation-worthy graphics and to give you practice in fitting both classical linear models (CLMs) and generalized linear models (GLMs). Recall that the `glm()` command works amazingly like the `lm()` command when your error terms are Normally distributed as we assumed before this week.

This assignment has you perform four analyses on the data. You select the most appropriate of the four, determine (and explain using statistics) why it is most appropriate, and produce prediction graphs that illustrate the point the models make.

When you turn in this assignment, you will email to me *two* files: your write-up and your R script. The write-up will provide a graph and a results table for the model you selected. It will also provide explanations of the four possible models, and why you selected the model you did.

As always, the email needs to include the subject line:

POLS6123: Assignment 8

Also as always, if you have problems, let me know.

PROBLEM: GDP PER CAPITA, AGAIN

[[25]]

In Assignment 3, you modeled the GDP per capita for several States around the world using the State's region and its level of honesty in government. In Assignment 7, you modeled the GDP per capita using the level of honesty in government and the category of government in the State. Finally, in the Mid-Semester Review Laboratory Activity you modeled GDP per capita with the level of honesty in government and OPEC membership status. In Assignment 3, you did not transform the dependent variable, but you did (or should have) in the other two.

For this assignment, you will again model the GDP per capita in a State. You will have four possible modeling paradigms to use: linear modeling without transforming the dependent variable, linear modeling *with* transforming the dependent variable appropriately, generalized linear modeling with the identity link (equivalent to linear modeling without transforming the dependent variable), and generalized linear modeling with the logarithm link.

The research model you will use will be

$$\text{gdpcap} \sim \text{democracy} * \text{region}$$

To be clear: You will fit the data using each of the four modeling techniques. You will select the best of the four. In your paper, you will first discuss the four options. Then, how you will select the best option and why that is an appropriate method (you may need multiple methods). Then, specify which you select (include relevant model-selection statistic), and why the others were not selected (again, include the relevant model-selection statistics). Then, you will provide a prediction plot and a coefficients table. Finally, you will write a couple pages explaining what the table and graph indicate.

Note. *There are a couple questions you should think about and write in your notes (not in the assignment you turn in). The first is why linear modeling and generalized linear modeling are equivalent when we use the identity link and the Gaussian distribution for the*

GLM. The second is why transforming the dependent variable with a transformation and fitting that with linear modeling is not equivalent to fitting the original data with generalized linear modeling, a Gaussian distribution, and that same transformation as the link.