

Economics 381, Fall 2016

Problem Set 3 – Introduction to Stata

Due Date: September 27

Instructions: Complete the problem set answers in the spaces given. Keep track of your Stata commands, print them out, and **staple** your Stata commands (and output, e.g. histogram in Q5) to this sheet.

Name (printed):

This problem set gets you used to looking through a dataset. This is good practice for when you pick your own dataset to analyze for the applied project. You will need to use the `sum`, the `corr`, the `tab`, the `graph twoway`, and the `reg` commands.

- Q1. Download and use the `gss381.dta` file available on Canvas.¹ Using the command `desc` look through the variables and their descriptions. Which variable do you find most interesting?
- Q2. How many observations are there from 2008?
- Q3. What is the age of the oldest person in the data?
- Q4. What are the different categories for the ‘race’ variable?
- Q5. Generate a `hist` of how many siblings people have.
- Q6. What percent of respondents think their federal income tax burden is too low?
- Q7. What is the average age of people who think radical Muslim clergymen should be allowed to speak freely?
- Q8. What is the average age of people who do NOT think radical Muslim clergymen should be allowed to speak freely?

¹This dataset is an edited version of the General Social Survey, which focuses on people in the south-east since 2000.

- Q9. What is the standard deviation of the age of people who voted for Obama as President in 2008?
- Q10. Is there a positive or negative correlation between ideology (as measured by ‘polviews’) and happiness (as measured ‘happy’)? Interpret this result: who consider themselves more happy, liberals or conservatives?
- Q11. What is the coefficient of correlation between ideology and income? Who is richer, liberals or conservatives?
- Q12. What is the coefficient of correlation between ideology and education? Who is more educated, liberals or conservatives?
- Q13. What is the coefficient of correlation between income and education?
- Q14. Draw a scatterplot of the relationship between income and education. On the same graph, include a line of best fit between these variables.
- Q15. Run a regression with people’s views about their federal tax burden being the dependent variable, and your independent variables being income, education, and ideology. Include the results in your answers.