

Pol Sci 630: Problem Set 11 - Diagnostic Techniques and Imputation of Missing Data

Prepared by: Jan Vogler (jan.vogler@duke.edu)

Due Date: Tuesday, November 10th, 2015, 10 AM (Beginning of Class)

It is absolutely essential that you show all your work, including intermediary steps, and comment on your R code to earn full credit. Showing all steps and commenting on code them will also be required in future problem sets.

Please use a *single* PDF file created through knitr to submit your answers. knitr allows you to combine R code and \LaTeX code in one document, meaning that you can include both the answers to R programming and math problems. Also submit the source code that generates the PDF file (i.e. either .Rnw or .Rmd files)

Make sure that the PDF files you submit do not include any references to your identity. The grading will happen anonymously. You can submit your answer at the following website: <http://ps630-f15.herokuapp.com/>

R Programming

Problem 1 (4 points)

Do the following in R:

Load the *LDC_IO_replication* dataset that was introduced in the tutorial. Imputation of missing data.

Problem 2 (6 points)

Do the following in R:

Diagnostic techniques.

Problem 3 (6 points)

Do the following in R:

Load the *National Accounts Data* (*na_data*) by the *Penn World Tables* that was used in the last homework. Diagnostic techniques.

Note: If you have the dataset in the same folder as your .Rnw file, you do not need to set a working directory when you compile your PDF. Note that you are not supposed to set a working directory as this might reveal your identity to the grader.