Outline of R Workshop

Yue Hu
July 29, 2015

1 Purpose

- Displaying the features, advantage, and disadvantages of R (comparing to STATA).
- Introducing the basic knowledge of statistical analysis and programming in R.
- Showing the means and resources to self learn R.

2 Schedule (Brief)

Day 1 (1 hour) Overview of R:

1. What is R?
2. Why do we learn R (as a beginner of programming/ as a STATA user)?
3. What R can do but STATA can't?
4. Install R and Rstudio
5. How to incorporate R with Github.

Day 2 (1.5 hours) Basic R

1. How to create vectors, lists, matices, and datasets?
2. Use R as an advanced calculator.
3. How to input, output data and set working directions?
4. Introduction of common R packages.

Day 3 (1.5 hours) Data Manipulation

1. How to use filter rows?
2. How to select columns (variables)?
3. How to create new variables based on the existing ones and how to arrange them?
4. How to bind data (long to wide, wide to long, merge, append)?
5. An introduction of piping programing

Day 4 (1.5 hours) Data Analysis and Programming

1. How to do descriptive statistics in R?

2. How to do t-test and ANOVA in R?
3. How to do OLS in R?
4. Introduction of basic MLE models in R.
5. Multiple Imputation and Factor Analysis (optional)

Day 5 (1.5 hours) Visualization and Output

1. Introduction of ggplot2.
2. How to do descriptive graphs (histgram, density, frequency)?
3. How to graph the regression results?
4. How to present interaction?
5. A very brief introduction of Latex.
6. How to incorporate R output (tables and figures) with Latex.

3 Examples

Table

Table 1

	Dependent variable:
	mpg
cyl	-0.942^{*}
	(0.551)
hp	-0.018
	(0.012)
wt	-3.167^{***}
	(0.741)
Constant	38.752^{***}
	(1.787)
Observations	32
R^2	0.843
Adjusted R^2	0.826
Residual Std. Error	2.512 (df = 28)
F Statistic	50.171^{***} (df = 3; 28)
Note:	*p<0.1; **p<0.05; ***p<0.01