**Guide to R Studio**

**(ECON 339)**

Table of Contents

[Importance of Learning R 3](#_Toc93049936)

[Introduction to R (Basics) 3](#_Toc93049937)

[Simple Operations in R 4](#_Toc93049938)

[Advanced Operations 5](#_Toc93049939)

[Reading in Data on Mac 6](#_Toc93049940)

[How to find Observations & Variables 8](#_Toc93049941)

[Scatter Plots 8](#_Toc93049942)

[Histogram 11](#_Toc93049943)

[Regressions 13](#_Toc93049944)

[Additional Resources/Trouble Shooting 14](#_Toc93049945)

# Importance of Learning R

There is a growing demand for data scientists and economic research assistants. A decent number of entry-level well-paying jobs require knowledge of a programming language and a solid foundation in economic theory. A few job title examples include economic consultant analyst, research assistant, data analytics analyst, and etc.

Why R instead of another language? R studio is a widely used statistical analysis and data visualization program. R is a growing language and becoming an industry standard. In addition, it is an open-source language which differs from another widely accepted statistical coding language called STATA.

The most important reason that R is incorporated into the classroom is to help students better understand economic research. When reading peer-reviewed papers it is important to have a basic understanding of the econometric models used in each paper. Using R allows for students to have hands on experience that will further their comprehension of papers read in class.

# Introduction to R (Basics)

The first step is of course to download R onto your computer/laptop. While learning a coding language can be intimidating it is an incredibly useful tool and can be used almost as a powerful graphing calculator. The best place to begin in coding is in a familiar environment ­­­­­­like simplistic math equations. Please keep in mind that when learning anything new it seems hard in the beginning but gets easier overtime.

\*This guide focused and directed to be a resource for ECON 339. It includes all instructions and direction needed for the course.

# Simple Operations in R

Since coding is unfamiliar the best place to begin is with familiar situations which is why this guide will start with simple operations. Most operations are easily predictable, but there are exceptions. Another benefit and tool of R is the ability to be a calculator and R follows the operation rules (PEMDAS). The R script window is where you type the codes (upper left window), and the console window produces the output of the run code (lower left window). Highlight the code that you want to run before clicking “Run”.

Addition: \_ + \_

Example: 2+2

Subtraction: \_ - \_

Example: 4-2

Multiplication: \_\*\_

Example: 2\*2

Division: \_/\_

Example: 3/2

Exponents and roots: \_\*\*\_ or \_^\_

Example: 2^3 and 4^(1/2)

All these examples can be seen on the photo below:

Graphical user interface, application

Description automatically generated

# Advanced Operations

Now that we understand simple operations the next step is learning the power of rational operations. These are similar to the Excel commands (i.e., IF, AND, OR). If you want to figure out is something is true or false, then R can tell you. The example below is very basic but will provide insight to rational operations.

\*Note: Remember to highlight code before you click “Run” (upper right corner of R script file)

Graphical user interface, application

Description automatically generated

# Reading in Data on Mac

The next step after familiarizing yourself with the operation function is to learn how to read in the data. The term “reading in the data” simply uploads the data, usually an excel/csv file, into R studio.

1) Download the data

2) Save in an easily accessible folder

3) Click “Finder”

4) Click “View” in the upper bar

5) Click “Show Path Bar”

6) Locate the data file and “control + click”

7) Click the ‘Copy “data file name” as Pathname’

8) Open R studio

9) data name = read.csv(“data file pathway name.csv”) -> type into the first line of R script

Graphical user interface, text, application

Description automatically generated

10) Click “Run” in R studio

11) [Optional Step] type View(“data name”) to see the data in grid form (like excel)

Graphical user interface, application

Description automatically generated

The data is read in as long as there are no errors stated in the R Console (one of the 4 windows). Now the data is in R, then we can start to learn simple commands in R.

\*Please note in step 9 that “data name” can be any name the coder chooses meaning it could be called “data” or “data1” or etc. Typically, in R studio if there is a \_\_\_\_ = \_\_\_\_ then the coder is simply naming something (data, variable, etc.).

# How to find Observations & Variables

1) Read in Data

2) Click “Environment” tab (upper right window)

3) Make sure it is set to “Global Enviroment”

4) Look for data set name

5) Look at right column and find the number of variables and number of observations

There is an example located below where you can clearly see for the given data set the number of observations (GovernorData) is 3150 and the number of variables is 9.

Graphical user interface, text, application

Description automatically generated

# Scatter Plots

The ability to visually see the relationships between variables is very imperative for predicting/checking the regression results (look for linear relationship). Remember that capitalization matters with every single command and to pay attention to detail. It is important to remember when running any line of code that involves a specific variable to reference the data set first. It is the data set’s name, then $, and finally the variable name.

Basic Scatter Plot Steps:

1) Read in data

2) View the data (optional but encouraged)

3) Plot command: plot (x = dataname$variablename, y = dataname$variablename)

4) Look at the “plot tabs” (bottom right) to see the scatter plot

Example Code: plot (x = GovernorData$year, y = GovernorData$GDPgrowth

Graphical user interface, text, application

Description automatically generated

More Advanced (Labeling) Scatter Plot Steps:

1) Do steps 1 and 2 of the basic scatter plot steps

2) Plot command with labels: plot (x = dataname$variablename, y = dataname$variablename, main = ‘title name’, xlab = ‘x axis name’, ylab = ‘y axis name’)

3) Look at “plot tabs”

Example Code: plot(x = GovernorData$year, y = GovernorData$GDPgrowth, main = 'GDP of States from 2005-2020', xlab = 'Year', ylab = "State GDP")

Graphical user interface, text, application

Description automatically generated

\*Compare the labeled graph to the basic graph

# Histogram

The ability to plot histograms is a great asset to interpreting data points and to see if the data is evenly distributed. The commands and syntax are relatively similar to creating scatter plots.

Steps to Create a Simple Histogram:

1) Read in data

2) Histogram Command: hist(dataname$variablename)

3) Look under “plots tab”

Example Code: hist(GovernorData$GDPgrowth)

Graphical user interface, text, application

Description automatically generated

More Advanced (Labeling) Histogram Steps:

1) Read in data

2) Histogram Command: hist(datename$variablename, main = ‘title name’, xlab = ‘x axis name’)

3) Look at “plots tab”

Example Code: hist(GovernorData$GDPgrowth, main = 'Histogram of State Quarterly GDP Growth', xlab = 'State Quarterly GDP Growth')

Graphical user interface, application

Description automatically generated

# Regressions

**How to Run a Simple Regression:**

1) Read in data

2) Determine the two variables need for the regression (Dependent & Independent)

3) Regression Command: regression name = lm(dependentvariable ~ independentvariable, data = data set name)

4) Summary Regression Command: summary(regression name)

5) Interpret results of summary (in consule)

Example Code: simplereg = lm(GDPgrowth ~ year, data = GovernorData)

summary(simplereg)

Graphical user interface, application

Description automatically generated

# Additional Resources/Trouble Shooting

Here’s are other guides that may be of use:

<https://bookdown.org/ndphillips/YaRrr/>

[https://moderndive.com](https://moderndive.com/)

If you ever run into any errors, then try to look them up in forums because chances are the answer lies within. Youtube videos are also extremely helpful.