R Lab 1: Introduction to R

Ikuma Ogura

August 20, 2019

Why Do You Need to Learn R?

- Why do you need to learn programming?
  - Because it makes your life easier!
  - It is helpful in all the steps in data analysis.
- Why R?
  - Free
  - Relatively easy to learn compared to other languages
  - Many people are using it $\rightarrow$ there are lots of helpful resources out there!
  - It's the standard software for quantitative political science (and other social science) research.

Basic Steps of Data Analysis

1. Specify research question
2. Collect data
3. Preprocess data
   - The data you have can contain errors, be incomplete etc.
   - Clean and tidy the dataset so that it can be used for analysis
4. Summarizing/Visualizing data
5. Statistical analysis
6. Summarizing/Visualizing the results

R is Helpful...

- For collecting data
  - Web scraping
  - Text extraction
  - ...
  - By automating the process, we can reduce the costs and errors in data collection.
- For data preprocessing
- For data visualization
- For statistical analysis

Example: Effort and Team Performance in Japanese Professional Soccer

- Do total distance run matter for soccer team performance?
  - Russia in 2018 World Cup?
- Data collection, preprocessing, & visualization with R

What You Need to Have

- R
  - Download from CRAN website
  - https://cran.r-project.org/
- RStudio
  - An interface to use R
  - Optional but very helpful
  - Download from https://www.rstudio.com/

RStudio

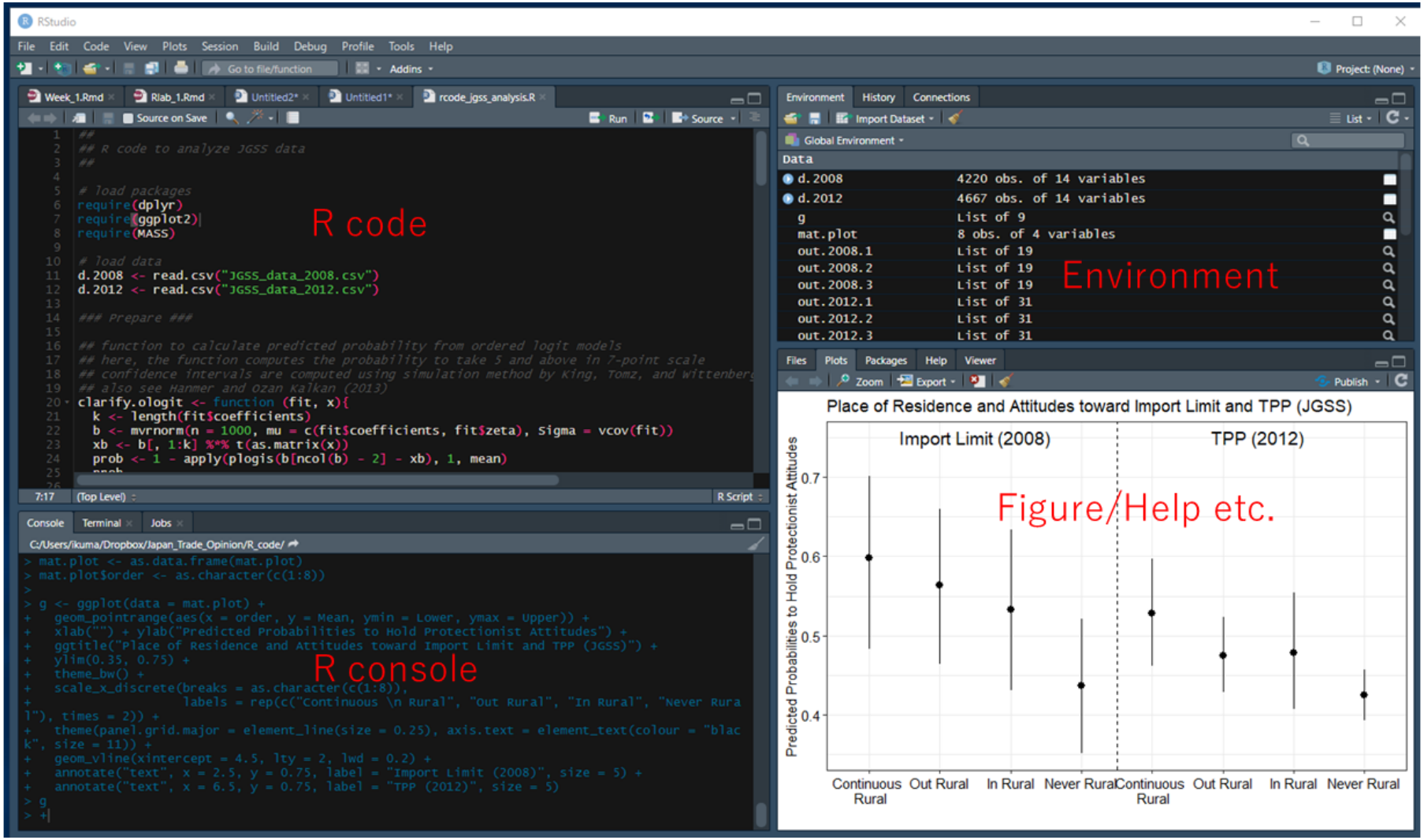

Figure 1: RStudio

Tips for Learning/Programming R

1. **Keep Record**
   - Don't directly type your R codes into the console.
   - Store your R codes as .R files.
2. **Make the Code Easy to Understand**
   - Add comments to your code
     - ★ R ignores everything written after # in implementing the codes
   - Follow coding guidelines
     - ★ e.g., Google's R Style Guide
     - ★ https://google.github.io/styleguide/Rguide.xml
3. **Practice Makes Perfect!**

Glossary

- **Object**
  - Named "box"/"container" that we store values/data etc. in R
- **Assignment**
  - The process of creating/modifying objects
- **Command/Function**
  - We use a command/function to perform some tasks on an object/objects
- **Argument**
  - The definitions, directions, or objects that are passed to a command/function
- **Package**
  - A collection of functions, data, and documentations which is publicly shared to enhance the functionality of R.

Example

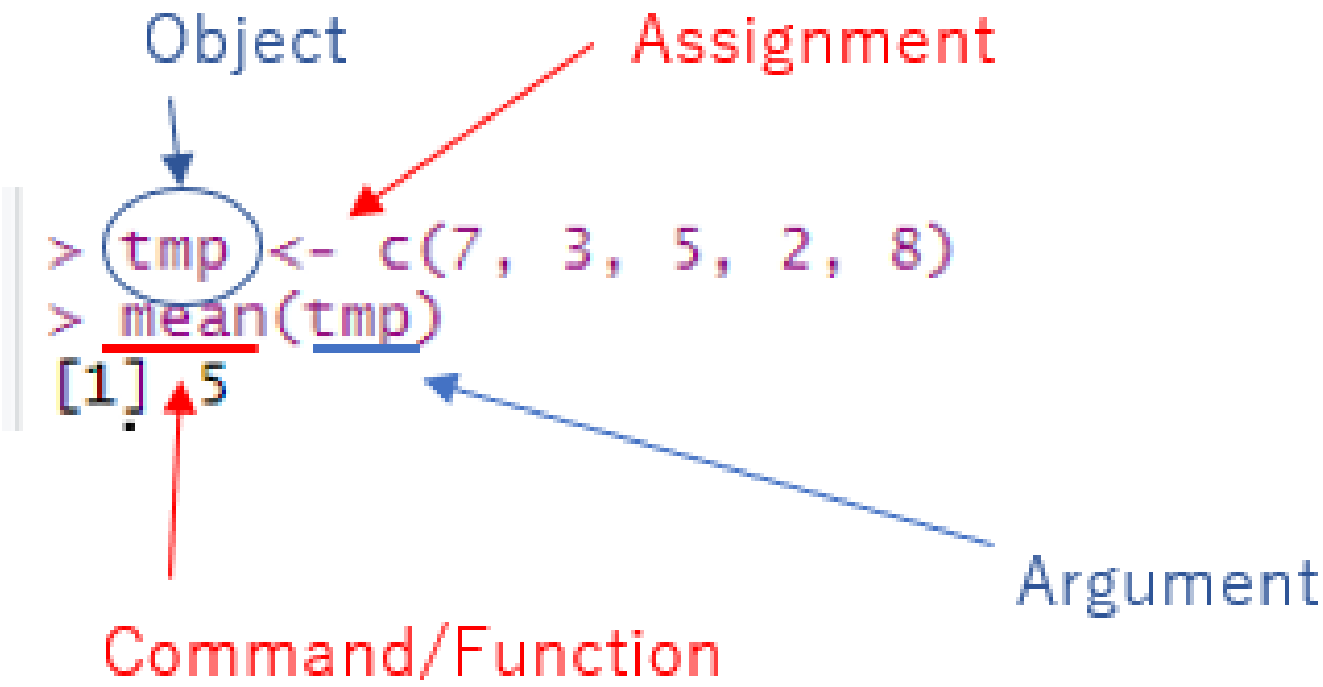

Figure 2: Example

Homework & Tommorow

- Homework
  - Download R and RStudio to your laptop if you haven't done so.
- Tommorow
  - Working with objects and functions
  - Please bring your laptop to the class!