Using the Statistical Software R

Simon Heuberger Ryan T. Moore*

2025-08-16

1 Background

Political scientists use statistical software packages to conduct quantitative analyses. One such software package is R. R is free and open-source. Statistical analyses can be done using R alone, but it is more user-friendly to use an R platform, or IDE (*integrated development environment*). One such platform is RStudio. R and RStudio are separate pieces of software and need to be downloaded and installed separately. We will only open RStudio. R will run from within RStudio. We do our analysis with the statistical computing software R.

2 Download and Installation

Install R before RStudio, so that RStudio will automatically find your version of R.

If you cannot install R and RStudio, you can access them via American University's Virtual Apps here: https://aka.ms/wvdarmweb. Instructions are available here.

If you are using a Chromebook, check out the installation advice here.

R

Navigate to https://cran.r-project.org, then click `Download R for [your operating system]`.

*Department of Government, American University, Kerwin Hall 228, 4400 Massachusetts Avenue NW, Washington DC 20016-8130. tel +1 202 885 6470; `rtm` (at) `american` (dot) `edu`; http://ryantmoore.org.

- For Mac OS, depending on your version of Mac OS and whether your chip is from Apple or Intel, click on the appropriate installer to download it. For example, for Mac OS 11 (Big Sur) and higher, `R-4.5.1-arm64.pkg`.
- For Windows, click `base`, then `Download R-4.5.1 for Windows` (or follow the links there for other versions).

Once the download is complete, find the downloaded file, open it, and follow the installation instructions.

RStudio

Go to the Posit website. Select Download RStudio.

Select the installer appropriate for your operating system. This is likely to be `RStudio-2025.05.1-513.DMG` or `RStudio-2025.05.1-513.EXE`.

Once the download is complete, find the downloaded file, open it, and follow the installation instructions.

3 The Basics of R and RStudio

The Basic Workflow

Whenever you start a new analysis, you'll go through the following steps:

1. Open RStudio (by opening a `.Rproj` file)
2. Create a new file for your code (a `.qmd` or `.R` file)
3. Write your analysis code in the `.qmd` or `.R` file
4. Execute the code to perform your analysis
5. Save the file
6. (Repeat steps 3 to 5 many times)

At the end of the day, save your file and quit RStudio. Tomorrow, open RStudio by opening a `.Rproj` file (more on `.Rproj` below), open your `.R` or `.qmd` file, run your code, and you'll be back where you left off.

Preferences

In Tools - Global Options - General, we recommend you

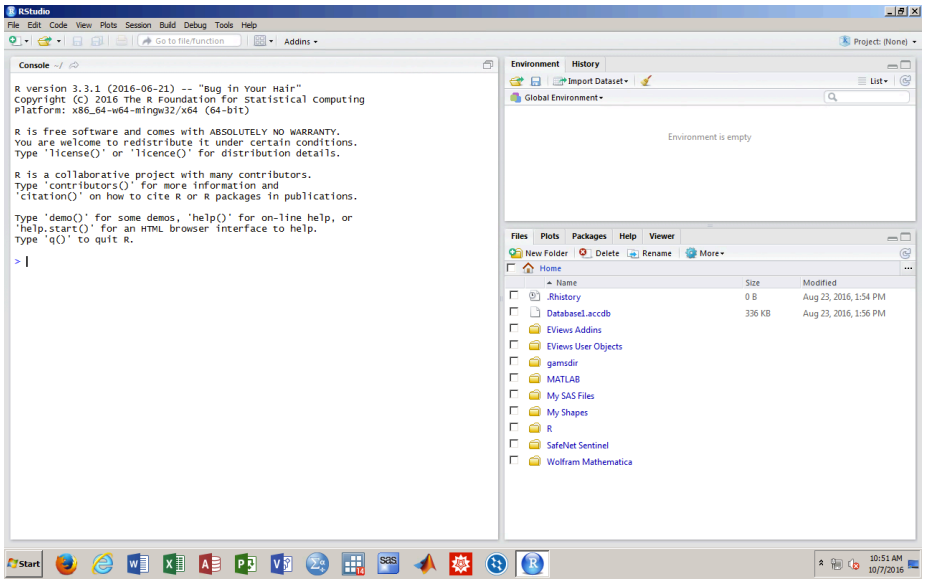
- **uncheck** "Restore .RData into workspace at startup"
- "Save workspace to .RData on exit:" **Never**

- **uncheck** "Always save history (even when not saving .RData)"

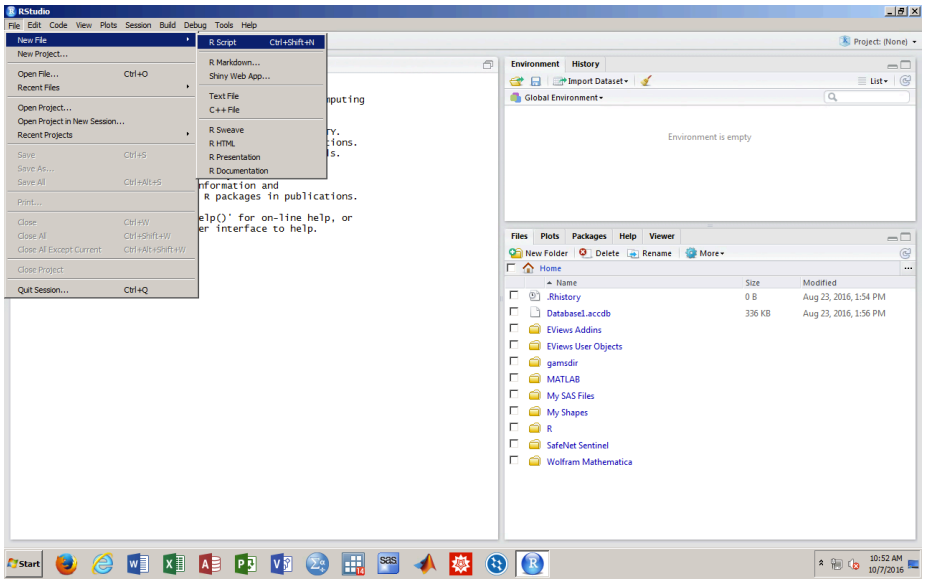
Also in the General tab, you can control the layout of RStudio with Pane Layout".

Creating a `.R` or `.qmd` File

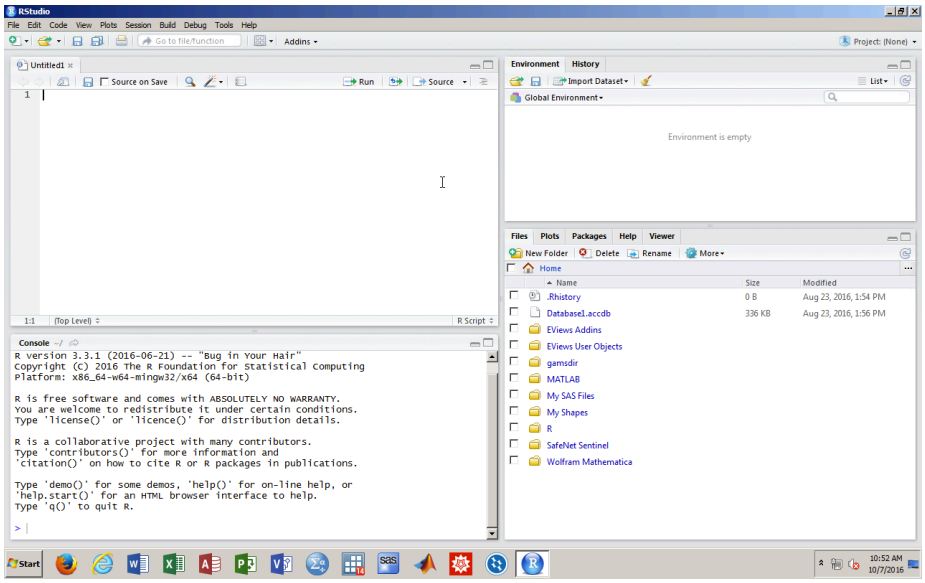
Open the RStudio application. Below is the default RStudio interface.



Open a new R Script (a `.R` file, as shown below), or a new "Quarto" file (a `.qmd` file). We will use `.R` files when we only want to record our code. We will use `.qmd` files to integrate analysis with English prose, figures, etc., such as for problem sets and papers.



Then the RStudio interface will look like this:



The top left window is the new "R Script" or "Quarto" file. This is where you type and save the code you want to preserve. The saved file is called a `.R` or `.qmd` file.

The bottom left window is the **Console**. This is where all of your code is executed, and some results displayed.

The top right window shows you the **Environment** and the **History** tabs.

Environment lists all the data and functions you have loaded in your current session. History lists all of the code commands you have used in your current session.

The bottom right window shows you the **Files**, **Plots**, **Packages**, **Help**, and **Viewer** tabs. Files lists all the files and folders that are in your default workspace (how to set the workspace is shown below). Plots shows the plots you have created in your current session. Packages lists all the R packages you have loaded (how to load packages is shown below). Help shows on information on any R package or function if you request it.

How Does RStudio Work?

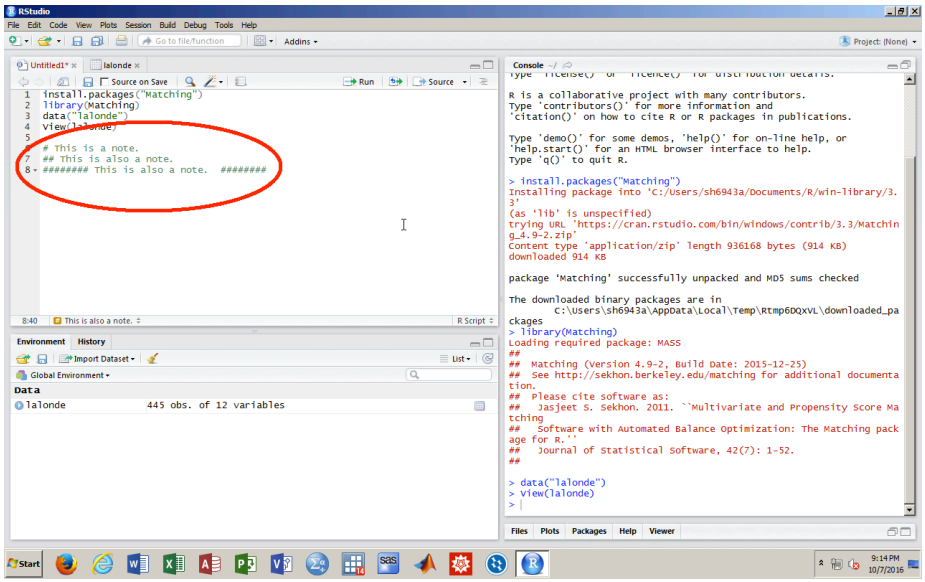
R is code-driven statistical software. The tabs in the task bar are kept to a minimum, which means that you will enter all code and conduct all analyses by typing on the keyboard. For example, to take the mean of a list of numbers `x`, you will type `mean(x)`.

Write the code you want to preserve in the `.R` or `.qmd` file and save this file frequently. This allows you to easily recreate and resume your analyses at a later time.

The code you write is executed in the Console. Some output of your code will also appear in the Console. Each new line in the Console is indicated by a blue `>` symbol.

Organization

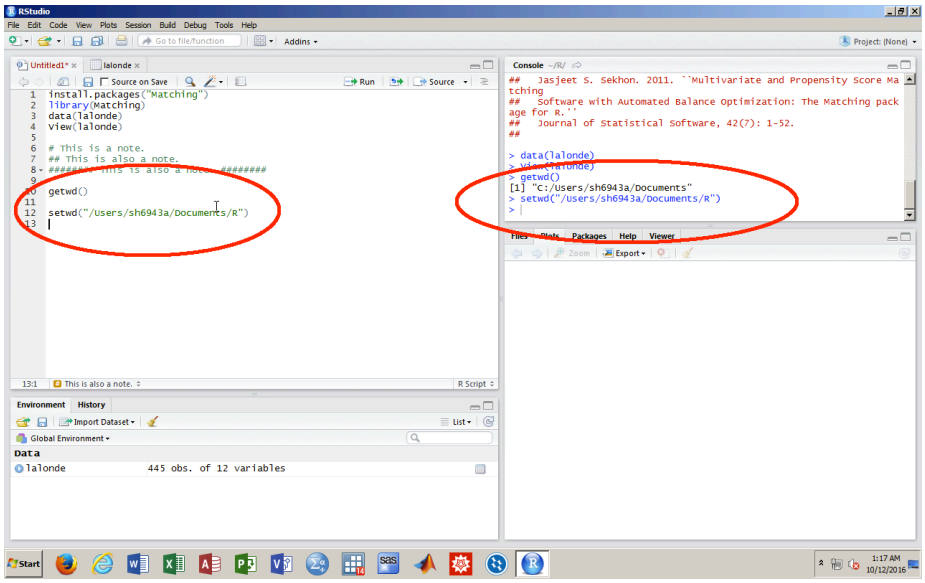
Keep your `.R` or `.qmd` file tidy. Delete lines that you no longer need. Write notes in your code for us, for collaborators, and most importantly, for your future self. Anyone looking at your code should be able to easily figure out what you are doing. Notes can be inserted by using the `#` symbol. Notes are shown in green font and are not executed by RStudio. You can use as many octothorpes as you wish.



Setting the Working Directory

R is always pointed at a specific folder on the computer you are working on. This folder is called the "working directory". To find out which directory RStudio is pointed at, run the `getwd()` (get working directory) function. To change your working directory, you can use the `setwd()` function and specify the path to the desired folder.

We recommend using R "projects" and the `{here}` package, rather than `setwd()`, to manage the working directory. See https://j.mp/32he2nf for details. You can only set the working directory to be a folder that already exists on your machine. It is good practice to start your R session by checking the working directory. That way, you know where any files or figures you save are located. Below we check and set the working directory:



Saving a `.R` or `.qmd` File

You need to save your `.R` or `.qmd` file in order to revisit it tomorrow, or, in the case of `.qmd`, to render the file. Simply click on the "Save" button and choose a folder. Alternatively, you can use the saving keyboard shortcut.

Where is my file?

If you do not specify a folder when you save, then your `.R` or `.qmd` will (usually) be in your working directory. Type at the R prompt `getwd()` to find the current working directory. Also, see the course FAQ and use R projects and the `{here}` package.

But where is my `.pdf` output from after I render?

If you did not specify a folder when you saved your `.R` or `.qmd`, then your `.pdf` will (usually) be in the same folder as your `.qmd`. Type `getwd()` at the R prompt to find the current working directory. Check your top-level user directory as well. Better, use R projects and the `{here}` package.

Installing and Loading Packages

Packages are sets of additional functions or data not installed with base R. You can install and load R packages with commands like:

```
install.packages("someNewPackage")
library(someNewPackage)
```

Package names like `{someNewPackage}` are case sensitive and quotation marks are required in the first command above. Installing a package is only necessary once (per version of R and the package), but loading it with `library()` needs to be done each time you restart R. It is good practice to begin your `.R` or `.qmd` file with a set of `library()` commands loading the packages needed for the analysis. For example,

```
library(here)
library(tidyverse)
```

Maintenance

RStudio will save all changes you make to the settings, such as changing the pane layout. Periodically, you may want to update both R and RStudio. This can be done via Help - Check for Updates in RStudio and R/Check For R Updates in your installation of R.