R & R Studio Tips and Best Practices

Jay Emerson

Tips and Best Practices

- Based on years and years of experience working with undergraduate and graduate students, high school students, government and industry professionals, . . .
- Will save you time
- Will help reduce confusion
- Will help avoid frustration
- Will make it more likely you're doing what you want to do
- Will save us time

Change some R Studio default options (critical for reproducibility)

- Open R Studio
- `Tools > Global Options...`
- Look for the first drop-down: `Save the workspace to .RData on exit`
- Change this first drop-down to `Never`
- All of the check-boxes ***above*** this dropdown should be ***unchecked***. That is, no, don't restore anything at startup.

Tip 1: Organize your data analysis world with folders

- Create a main folder someplace for your work. Maybe on your desktop?
- Use subfolders in a sensible way, and use a lot of them!
- Example: Each homework assignment (say for students) should have its own folder.
- Philosophy: one folder, one "data set" (even if a single data set might involve multiple files), and one or more scripts that are a complete record of your work.

Tip 2: Do all of your work in a script

- Avoid typing in the console unless it is really something interactive that you don't need a record of.
- Save your work regularly (that is, save the script)
- Only work on one thing at a time (more on this later)

Tip 3: Only work on one thing at a time!

- Constantly close up R Studio (don't just minimize)
- Whenever possible, start R Studio by double-clicking a script for your analysis – by doing this, R Studio will automatically be working in the right place (if you organize yourself as recommended above)
- Don't have multiple scripts open simultaneously relating to different data analysis problems

Demonstration

Getting started with a new script

Demonstration

Common error with read.csv() or read.table()

Demonstration

A file error with that CSV that wasn't a CSV

Tip 5: Learn a few shortcuts

- With your cursor on some line of a script, do `command-enter` on the Mac or `control-enter` on the PC. Some people say "return" instead of "enter", or "ctrl" instead of "control" or (on the Mac "squiggley" instead of "command").
  - This runs that line of the script, equivalent to a copy-paste from the script into the console
  - If it doesn't work or doesn't do what you expect, then perhaps you didn't run the preceeding code in your script? The script is a sequential record of your work. Actual work is done in the console.
- If you highlight a block of code, the same shortcut will run the block
- Tab-completion can sometimes be helpful
- The arrow keys can be used in the console to browse through past commands

Tip 6: Other customization (optional)

- Again in `Tools > Global Options...`
- Under the "Appearance" section, choose a font size large enough so you can read your code easily
- Under "Code" there are a bunch of options that are a matter of personal preference.
- Example: "Insert matching parens/quotes" is selected by default, but I turn it off (I don't like it). Some students love this. Your choice.

Tip 7: How much code should go on a line and is indentation important?

- Not more than 80 characters
- Under `Code > Display` there is an option to show a vertical line at column 80, and I strongly recommend sticking to this.
- A single command can be broken onto multiple lines.
- Indendatation is important and easy in R Studio. Don't be sloppy!
- `command-I` or `ctrl-I` can automatically do some nice indendatation of your code... another useful shortcut!

Demonstration or Exercise

Repetition!

Foreign language character set difficulties

This may only apply to a few of you. And it tends to be a Windows problem (I don't recall this on the Mac). If your operating system (Windows) is set up to use a foreign language, your installation of R might sometimes run into troubles with character encodings (kind of like fonts) – things like accents and special characters. My only advice is learn how to switch between languages. Also note that in R Studio's `Tools > Global Options... > Code > Saving` configuration options some common script types are listed. It may be that forcing a type as UTF-8 or ASCII will help resolve some issues when the occur.

Never use Microsoft Word to edit R scripts

Word tends to use special character encodings for things like quotes, which will definitely cause problems in R. Avoid it!