Getting started guide for the SER workshop on open science

Sam Harper

2020-10-30

Getting started

I'm excited to meet you during our workshop on Oct 30. Since the workshop is virtual there will be perhaps less hands-on activity than if we were in person, but I still hope you'll come away with a strong grasp of why and how to engage in reproducible research. Given limited time and our virtual environment, it's going to be difficult to troubleshoot any technical issues, and that is not the core focus of the workshop. Nevertheless, you may find it useful to get a little familiar with software for conducting reproducible research so it does not look completely foreign.

R and RStudio

i) Download and install R from http://cran.r-project.org. Simply click the link to download R for Mac or Windows. If you already have R, consider updating to the latest version.

ii) Download and install the free version of RStudio from https://www.rstudio.com/products/rstudio/download/.

iii) Open up RStudio and type the following lines of code into the console (bottom left window):

```
install.packages(c("rmarkdown", "tidyverse", "here",
  "estimatr", "ggeffects", "kableExtra", "xtable", "stargazer"))
```

This code installs a few R packages that we will be using to generate reproducible research documents.

If you want a gentle introduction to using Rmarkdown, check out this tutorial: https://ourcodingclub.github.io/tutorials/rmarkdown/

Help!

Each function comes with a help page telling you what it does and what arguments you need to provide the function. There are two ways to ask to see the help page of a function:

```
help("library")
?library
```

When we execute either of the above lines of code, the Help tab will open and navigate to the help page for the function `library()`. Calling a help page in this way is most useful if you know the name of the function you'd like to use but forget exactly how to use the function. Don't worry if you can't understand what is written there right now.

If you would like help, but can't remember then name of the function you need help with, you can type a double question mark before a word that is related to the function you are looking for:

```
??histogram
```

This shows you all the packages and their functions which discuss histograms. For example ggplot2::geom_freqpoly(), means that there is a function `geom_freqpoly()` from the `ggplot2` package that is used to make histograms. You can click through the functions until you find the one you're looking for.

Finally, get used to using Google to debug. It's just part of the R universe.

Stata

If you are new to Stata or have not used it in some time, please see a couple of great overviews/introductions:

- German Rodriguez's Stata tutorial.
- Geocenter introduction.

For Stata enthusiasts, please see the detailed guidance on how to install and use `markstat` at the website of German Rodriguez. It also involves installing some dependencies (including LaTeX if you want to produce pdf documents.)

To reproduce the example in the tutorial, you'll also need the following packages, which can be installed by typing the following into your Stata command line:

```
ssc install tabout, replace
ssc install estout, replace
```

Github and OSF

Sign up for a free account at Github.
Sign up for a free account at the Open Science Foundation.

You can always delete these accounts later, but they may be useful for following along.