Social Science Computational Resources Series

June 30, 2020

Brooke Luetgert, PhD

THE UNIVERSITY OF CHICAGO

Office of Research and
National Laboratories
Research Computing Center

About this Series

This R workshop and tomorrow's session in Python are designed to expose participants to working with data in various software environments. While many tools will "get the job done", knowing the tips and tricks of multiple alternatives helps the user to work more efficiently and confidently through data analysis and presentation. Our focus is on good data practice, documentation and reproducibility across software platforms.

This is an introductory series for absolute beginners!

THE UNIVERSITY OF CHICAGO | Office of Research and National Laboratories Research Computing Center

Contact Details

Brooke Luetgert, PhD. Computational Scientist at RCC

Email: luetgert@uchicago.edu
Office: TAAC 2, 5607 South Drexel
Telephone: (773)-834-5313

RCC Help Desk: Reg. 216, Mon-Fri 9AM-5PM

Materials on GitHub- use search bar, enter user:luetgert
Our folder is **luetgert/introR**

Plan for Today

Our focus will be two-fold:

1. We want to download RStudio and set up our local workspace

2. We want to work our way through several descriptive data analysis and plotting tasks with a focus on exporting our results for reporting.

Download RStudio

R and RStudio are separate downloads and installations. R is the underlying language and compute environment, but RStudio is the IDE (integrated development environment) that makes R easier and interactive.

<u>First, we install R, then RStudio:</u>

1. Go to the CRAN website, download and install R
2. Go to the RStudio download page and download

*Always confirm your OS: **Windows vs. Mac**

THE UNIVERSITY OF **CHICAGO** | **Office of Research and National Laboratories Research Computing Center**

Tidyverse

<u>AFTER</u> RStudio is running on your laptop, we will install our first package:

At the console type:
install.packages(c("tidyverse"))

THE UNIVERSITY OF CHICAGO | Office of Research and National Laboratories Research Computing Center

Why R?

Learning R may be more challenging than STATA or SPSS, but typing your commands rather than depending on point and click is a very good thing!

R scripts will make the steps in your analysis clear. The code you write will be transparent. You will gain a deeper understanding of what you are doing and the assumptions that you are making.

Code is great for reproducibility, fixing mistakes, updating assumptions and data and meeting publication standards!

More R Benefits

R produces beautiful plots that are easy to label

R has a tremendous support community with a great deal of code snippets and insights

R is open-source and cross-platform

R offers a large selection of add-on packages which enhance the R experience for discipline specific research.