Introduction to R Workshop

Spring 2021

Instructor: Burcu Kolcak
Email: bk457@polisci.rutgers.edu
Dates: Tuesday, January 12 − Friday, January 15
Morning Session: 10:00 am − 12:00 pm
Afternoon Session: 1:30 pm − 3:30 pm
Location: Zoom
Office Hours: 7:00 − 9:00 pm

Description

This workshop is designed to introduce students to the statistical programming language R. The main focus will be on covering topics, concepts, and R commands that will help you get familiar with R for your quantitative methods courses and start using R for your own research. We will use real datasets to explore how data analysis and statistical analysis using R looks like in action. We will also briefly discuss various types of analyses using R: geospatial, text, and network analysis.

No prior knowledge of R or any other programming language is required for the workshop. We will start learning R from scratch. More advanced students are more than welcome to join whenever they wish.

Structure

Each day of the workshop will be divided into two sessions. In the morning session and the first half of the afternoon session, we will cover new topics, concepts, and R commands. We will work on short exercises after we cover some topics. The second half of the afternoon session will be devoted to putting the skills covered into practice.

Daily Problem Sets

We will have daily problem sets to maximize the amount of practice we do throughout the workshop. Although problem sets are highly recommended for facilitating your learning process, they will be optional.

Slack

We will be using Slack to communicate. Announcements, meeting links, handouts, solutions to problem sets, and R resources will be posted on Slack. Our Slack channel is called *Polisci R @Rutgers*. You can join the channel by following this link. You can find a short Slack tutorial here. Please feel free to start using Slack to post any questions you have.

Getting Started with R

We will be using R via R Studio − a powerful user interface − to help make using R easier. The only thing you need to do before the workshop is to install R and R Studio. A *Getting Started Handout* on R and RStudio will be sent to the participants of the workshop.

Materials

The workshop will not follow a specific textbook. The following books are recommended. Relevant chapters from textbooks for each day can be found in the outline.

Imai, Kosuke. 2018. *Quantitative social science: An introduction.* Princeton University Press. *(QSS)*

Healy, Kieran. 2018. *Data Visualization: A Practical Introduction.* Princeton University Press. *(DViz)*

Wickham, Hadley and Garrett Grolemund. 2016. *R for data science.* O'Reilly Media, Inc. *(R4DS)*

Chris Bail at Duke University has a series of short videos introducing R. These videos have been created as a resource for participants in the Summer Institute in Computational Social Science (SICSS) without much prior programming experience. If you want to practice after the sessions, these videos can be useful. Relevant videos for each day can be found in the outline.

You can find Bail's R Videos here.

Outline

Before the Workshop

- A *Getting Started Handout* on R and RStudio will be sent by email before the workshop.

Day 1: Tuesday, January 12

Recommended Reading

Vance, Ashley. 2009. "Data Analysts Captivated by R's Power." *The New York Times.*

Topics

- Getting Familiar with R

  Using R as a calculator; working with vectors; reading data into R; packages; summarizing data; the structure of R; programming tips

Relevant Material

QSS, Chapter 1
Bail, R Basics

Problem Set I

Posted on Tuesday, January 12, 12:30 pm; due on Tuesday, January 12, 11:59 pm.

Day 2: Wednesday, January 13

Topics

- Subsetting Data: Logical Operators, Relational Operators & Conditional Statements
- Descriptive Statistics
- Visualizing Univariate Data

Relevant Material

QSS, Chapter 2 (2.2, 2.6); Chapter 3 (3.3)
Bail, Data 'Wrangling'

Problem Set II

Posted on Wednesday, January 13, 12:30 pm; due on Wednesday, January 13, 11:59 pm.

Day 3: Thursday, January 14

Topics

- Correlation and Linear Regression
- Visualizing Bivariate Relationships
- More on Visualization: Visualization with ggplot2

Relevant Material

QSS, Chapter 3 (3.6); Chapter 4 (4.2)
DViz, Chapter 3
Bail, Data Visualization; Basic Programming; Modeling

Problem Set III

Posted on Thursday, January 14, 12:30 pm; due on Thursday, January 14, 11:59 pm.

Day 4: Friday, January 15

Topics

- Loops
- Discovery: Geospatial Analysis
- Discussion: Textual Analysis and Network Analysis
- A Very Brief Introduction to R Markdown

Relevant Material

QSS, Chapter 5
DViz, Chapter 7
Bail, Communicating & Collaborating

Problem Set IV

Posted on Friday, January 15, 12:30 pm; due date is flexible.