

Introduction to Programming with R for Political Scientists

Markus Freitag
Geschwister Scholl Institute of Political Science
LMU Munich

July 12-13, 2021

E-mail: markus.freitag@campus.lmu.de

Office Hours: Flexibly available upon request

Course Website: TBD

Web: markusfreitag.netlify.com

Venue: Zoom or at the GSI

Course Description

R is a programming language and free software environment for statistical computing: for data analysis, data manipulation and data visualization. Always a cornerstone in statistics, R has also become the “lingua franca” of quantitative political science in recent years. Moreover, it has become increasingly popular in industry. This course will introduce political scientists to R, with an emphasis on programming (as compared to statistical analysis) and reproducible workflows including version control with GitHub. Starting with “Base R”, the course will also cover various recently popular packages for data manipulation and visualization (e.g. from the tidyverse). The objective of this course is to lay the foundations for conducting rigorous and reproducible quantitative research in R.

Learning Objectives

1. Using the R-Studio IDE and version control with GitHub
2. Data wrangling and visualization
3. Writing functions (and clean code)
4. Conducting reproducible research with Rmarkdown

Prerequisite Knowledge

This course is intended for advanced B.A. students, M.A. students and junior faculty who want to work with R or switch to R from another programming language (e.g., Stata). No prior experience with R is required. However, basic knowledge of statistics is necessary to follow the (practical) examples of this course (BA students must have completed “Methoden der Politikwissenschaft II”). Proficiency in another programming language (e.g. Stata or Python) may decrease the learning curve.

Class Schedule

The course will take the format of an intense 2 day workshop with “chalk and talk” blocks as well as breakout sessions in which groups of participants work together on problem sets.

Topics	Time-Schedule
Day 1 (12.07.2021)	
Kick-Off and First Steps	
R Studio IDE and version control with github	09:00 - 10:30
Break	10:30 - 10:45
Base R and into the Tidyverse	
Base R & Tidyverse Basics	10:45 - 11:45
breakout session 1	11:45 - 12:30
Break	12:30 - 13:00
Data Wrangling I	13:00 - 14:00
breakout session 2	14:00 - 14:30
Data Wrangling II: Advanced data manipulations	14:30 - 15:30
breakout session 3	15:30 - 16:15
Day 2 (13.07.2021)	
Data Visualization	09:00 - 10:30
breakout session 1	10:30 - 12:15
Break	12:15 - 12:45
Writing functions (and clean code)	12:45 - 13:45
breakout session 2	13:45 - 15:00
A complete scientific workflow with R	15:00 - 16:15
Wrap-up	

Preparation

In order to avoid a sluggish and organization-heavy start of the course, installation of all software components used in this course is to be completed in advance. Preparatory readings are not mandatory but may decrease the learning curve.

Setting up R, R-Studio and (Git)Hub

If you've never used R, R-Studio or (Git)Hub before, start by installing each software component. I suggest you to do so in the following order:

1. Install [R](#).
2. Install [R-Studio](#).
3. Create a [GitHub](#) account. Take some care with the user-name if you want to keep this account throughout your career. You can't change it afterwards.
4. Install [Git](#).

4.1 Open a shell. On windows, Git comes with its own shell. Open it by searching for "Git Bash". On Mac, just search for your "Terminal" application. Enter the following commands:

```
git config --global user.name 'Your name'
git config --global user.email 'Your E-mail linked to GitHub'
```

4.2 (optional) Sometimes, using a GUI can be convenient for interacting with Git. [GitHub Desktop](#) has gotten pretty decent recently. The GitHub Desktop installation wizard also helps you with configuring Git such that you can skip step 4.1.

If you already want to familiarize yourself with using Git(Hub) together with R-Studio, you can read into it in more depth [here](#) (however, we will also cover this in class).

A couple of days before the course you will receive an invitation to the course repo which you have to accept.

Readings (optional)

Preparatory

Monogan III, James E. 2015. *Political Analysis Using R*. 1st ed. 2015. Cham: Springer International Publishing: Imprint: Springer.

Wickham, Hadley, and Garrett Grolmund. 2016. *R for data science: import, tidy, transform, visualize, and model data*. First edition. Sebastopol, CA: O'Reilly. Free online version available: <https://r4ds.had.co.nz/index.html>.

Further Reading

Urdinez, Francisco, and Andres Cruz. 2020. *R for Political Data Science: A Practical Guide*. 1st ed. Chapman and Hall/CRC.