R Basics for Social Sciences

Universidad Carlos III de Madrid

Semester:	Spring 2025
Room:	5.1.01
Time:	Wednesdays, 10:30–12:30

Contact Information

Instructor:	Patrick Kraft, PhD
Office:	18.2.A19
Email:	patrickwilli.kraft@uc3m.es

I Course Description

An introduction to R course typically aims to provide students with a solid foundation in the R programming language and its applications in data analysis and statistics. This course is intended to help faculty in the Social Sciences Department to learn the R language to be able to teach their courses using this software. The main objectives of such a course include:

- Familiarity with R Environment: Help students become comfortable with the R programming environment, including the R console and RStudio.
- Basic Syntax: Teach the fundamental syntax and data structures in R, such as vectors, data frames, and lists.
- Data Import and Export: Instruct students on how to import data from various file formats (e.g., CSV, Excel, and text) and export results.
- Data Manipulation: Cover data manipulation techniques, including subsetting, filtering, merging, and reshaping data.
- Statistical Analysis: Introduce basic statistical concepts and demonstrate how to perform common statistical tests and analyses using R.
- Data Visualization: Teach students how to create informative and visually appealing data visualizations using R packages like ggplot2.
- R Packages: Familiarize students with the concept of packages and how to install and use them to extend R's capabilities for specific tasks.
- Reproducible Research: Emphasize the importance of writing well-documented and reproducible R scripts and reports using tools like R Markdown and Quarto.
- Basic Programming Concepts: Introduce students to programming concepts in R, such as loops, conditional statements, and functions.
- In-Class Applications: Provide practical examples and possible exercises to do in class with their students and how to solve them.

These objectives are designed to give the faculty a solid foundation in R that allows them to teach basic statistics, either descriptive or inferential, and incentivize them to apply this language in their research activities.

II Prerequisites

This course has no prerequisites, although prior experience with other statistical programming environments such as SPSS or Stata may be helpful. Please bring a laptop to every meeting since most of the course will consist of interactive lab sessions. You can get started by installing R (`https://cran.rstudio.com/`) and RStudio (`https://posit.co/download/rstudio-desktop/`) on your computer.

III Schedule

Date	Topic	Room
22/01	Introduction to base R & RStudio	5.1.01
31/01	Data Visualization using ggplot2 (1)	5.1.01
14/02	Data Visualization using ggplot2 (2)	5.1.01
28/02	Data Manipulation using dplyr (1)	5.1.01
13/03	Data Manipulation using dplyr (2)	5.1.01
10/04	Cleaning messy data with tidyr	5.1.01
24/04	Working with (generalized) linear models in R	5.1.01
08/05	Creating regression tables for latex, markdown/quarto, and word	5.1.01
22/05	Simulating & visualizing quantities of interest (1)	5.1.01
05/06	Simulating & visualizing quantities of interest (2)	5.1.01

Note: Schedule may be subject to change depending on our progress during the semester.

IV Resources

Here are is a selection of free online resources that will help you get started using R. Make sure to bookmark these links, as they can serve as useful references beyond this course when you continue working with R.

Getting Started

- Posit recipes: https://posit.cloud/learn/recipes
- Posit cheat sheets: https://posit.co/resources/cheatsheets/

Introduction to R & Data Wrangling

- R for Data Science (2e): https://r4ds.hadley.nz/
- R Cookbook (2e): https://rc2e.com/
- Tidyverse Cookbook: https://rstudio-education.github.io/tidyverse-cookbook/

Basic Plotting & Data Visualization

- Data Visualization: https://socviz.co
- Fundamentals of Data Visualization: https://clauswilke.com/dataviz/
- R Graphics Cookbook (2e): https://r-graphics.org/

Statistical Modeling & Reporting

- `estimatr`: Fast Estimators for Design-Based Inference: https://declaredesign.org/r/estimatr/
- `modelsummary`: Summarize statistical models in R: https://modelsummary.com/
- `marginaleffects`: Simulate quantities of interest in R: https://marginaleffects.com/
- R Markdown: The Definitive Guide: https://bookdown.org/yihui/rmarkdown/
- Quarto: An open-source scientific and technical publishing system: https://quarto.org/

Books to consult if you don't like the tidyverse (not available online)

- Matloff, Norman. 2011. *The art of R programming: a tour of statistical software design*. No Starch Press
- Fox, John, and Sanford Weisberg. 2018. *An R companion to applied regression*. 3 ed. Sage Publications