

Introduction to R

Open Online Course

University of Mannheim

Lion Behrens, M.Sc.
behrens@uni-mannheim.de

April 2022

Overview

This self-paced open online course is developed for B.A. and M.A. students at the University of Mannheim. The course provides an introduction to modern R for data analysis that underlies academic research in the social sciences. The material is constructed around seven thematic schemes which together cover all essential skills that need to be mastered for applied quantitative research. Every thematic scheme is constructed around (i) a set of slides, (ii) a set of short video lectures, (iii) illustrative code and (iv) short exercises for self-study.

The data that is going to be used throughout the course is the cumulative datafile of the European Social Survey (Round 1-9).

Time schedule

The construction of this course is going to start on Monday, June 27 and last until Sunday, August 28. This means that the material for the seven thematic building blocks is going to be developed over the course of nine weeks or 63 days. Hence, for each of the seven building blocks, there are nine days available to prepare the material. On Sunday, August 28, the material is ready to be placed online.

Course Content

Building Block	Topics	Prepared until
Getting started		
1 The R Environment	<p><i>Topics covered:</i></p> <ul style="list-style-type: none">○ Video: Welcome and what is R?○ Video: R, RStudio, Packages, basic functionality and how to get help○ Video: R Projects and other best practices○ Video: Object classes – vectors, matrices, factors, lists and data frames <p><i>Learning objectives:</i></p> <ul style="list-style-type: none">– Understanding the unique advantages of R for data analysis and how it relates to general purpose programming languages– Understanding the basic functionality of R/RStudio, variable assignment, selecting and indexing, appending and merging objects– Understanding how to set up an efficient workflow for data analysis (using R Projects)– Knowing how to structure one's script– Differentiating basic data types and when they are used	Tuesday, July 5

R Basics

2	Data Manipulation	<p><i>Topics covered:</i></p> <ul style="list-style-type: none">○ Video: Data management (load and store data)○ Video: Logical and relational operators○ Video: Data manipulation in base R and dplyr <p><i>Learning objectives:</i></p> <ul style="list-style-type: none">– Knowing how to read in structured datasets of various formats (dta, csv, txt) into R and how to store them in different formats– Understanding the basics of cleaning and preparing datasets (defining missing values, re-coding and transforming variables) using base R and dplyr– Knowing how to explore descriptive statistics– Knowing what tibbles are and how they relate to base R data types	Thursday, July 14
3	Control Flow	<p><i>Topics covered:</i></p> <ul style="list-style-type: none">○ Video: Conditional statements○ Video: How to write replicable code? <p><i>Learning objectives:</i></p> <ul style="list-style-type: none">– Knowing how use conditional statements (if, if-else) to structure the workflow (conditional code execution)	Saturday, July 23
4	Statistical Modeling	<p><i>Topics covered:</i></p> <ul style="list-style-type: none">○ Video: Linear regression (incl. interactions, polynomials, variable transformations)○ Video: Generalized linear models○ Video: Multilevel regression <p><i>Learning objectives:</i></p> <ul style="list-style-type: none">– Knowing how to apply the most commonly used linear additive models in the social sciences– Knowing how to inspect the output of these models– Knowing how to export and format regression results to publication-ready tables using <i>stargazer</i>	Monday, August
5	Data Visualization	<p><i>Topics covered:</i></p> <ul style="list-style-type: none">○ Video: base R vs. ggplot2 philosophies○ Video: Plotting in base R○ Video: Plotting in ggplot2 <p><i>Learning objectives:</i></p> <ul style="list-style-type: none">– Understanding how the approaches to visualizing data in base R and ggplot2 relate to each other– Knowing how to construct the most common plots in base R and ggplot2 (histograms, scatterplots) and how to visualize regression results	Wednesday, August 10

Efficient workflow

6 Loops & apply

Topics covered:

- Video: for loops, while loops
- Video: The apply family

Friday, August 19

Learning objectives:

- Understanding how to write efficient code
- Understanding how loops can shorten repetitive code
- Understanding how the apply family serves as a computationally efficient alternative to loops that avoids indexing

7 Functions

Topics covered:

- Video: Manually defining functions

Sunday,
August 28

Learning objectives:

- Understanding how own functions can customize one's workflow
-
-