

Introduction to R

Dates: 23-27 September 2024

Time: 13:00-15:00

Room: Sala del Capitolo, Badia Fiesolana

Instructor: [Joris Frese](#), PhD Researcher at EUI SPS (2022 Cohort)

Contact: Monika.Rzemieniecka@eui.eu & joris.frese@eui.eu

Description

This workshop is designed for individuals who are new to R or have limited experience with it.

It aims to provide participants with the fundamental skills to work effectively in the RStudio environment and conduct statistical programming tasks.

The course covers basic operations with R objects (vectors, data frames), data manipulation, data analysis and visualization, as well as best practices for coding with R.

For additional reading besides the slides used in the course, students can feel free to look at the e-book “R for Data Science”, co-authored by Hadley Wickham, the chief scientist at posit (the mother company behind RStudio). The book is freely available at: <https://r4ds.had.co.nz/index.html>. You can also check out this website listing various cheat sheets for popular R packages: <https://posit.co/resources/cheatsheets/>.

Format

Each session will consist of both lecture and workshop elements. Each researcher should have their own laptop with R and RStudio installed in order to participate interactively.

Before the workshop

For the workshop, you will need to have both R and RStudio working on your computers. If you don't have them already, you can download them for free from the following links:

- R: <https://cran.r-project.org/>
- RStudio: <https://posit.co/products/open-source/rstudio/>

Course structure

Day 1 (Monday, 23 Sep): Introduction, Basics of R & RStudio

In the first session, we will learn the very basics of R and RStudio, focusing on the layout of RStudio and some basic functions and features of R.

Day 2 (Tuesday, 24 Sep): Data Manipulation

In the second session, we will focus on data preparation, such as subsetting, selecting, summarizing, merging, and general data wrangling. We will get to know the tidyverse, which contains a series of useful packages for these purposes.

Day 3 (Wednesday, 25 Sep): Data Analysis & Visualization

The third session will focus on basic data analysis (data inspection, summary statistics, correlations, regression) and data visualization (bar charts, line graphs, scatter plots, ggplot).

Day 4 (Thursday, 26 Sep): Data Analysis & Visualization (cont.)

The fourth session will continue with slightly more advanced data analysis and visualization techniques.

Day 5 (Friday, 27 Sep): Reproducible Code, Coding Advice

In the fifth and final session, we will discuss ways to make your code more efficient, easier to read, and easier to reproduce. We will also talk about the best sources for good coding advice when you are stuck on a task.