Online course learning objectives

This course teaches learners the required skills to use this flexible and multi-purpose platform for their own research.

By the end of this course, learners will:

- Have a good understanding of how R works.
- Be able to perform a wide range of data management tasks, with a focus on solving day-to-day conundrums that we all face as social scientists.
- Have the knowledge and skills to apply an extensive set of data exploratory and visualization techniques.
- Be able to use R to perform some of the most common statistical techniques used in the social sciences, namely a dimension reduction technique and OLS regression with interactions.

Language: English
Time to complete: 20 hours
Instructor: Andreea Moldovan

Online course full syllabus

MODULE ONE: WHAT IS R AND WHY USE IT?

- Why use R and what is it exactly?
- Downloading and installing R and RStudio.
- Introducing the user-friendly RStudio Integrated Development.
- Environment.
- Resources and repositories: a short introduction to CRAN, GitHub, and ‘the R Journal’.

MODULE TWO: THE R LANGUAGE SIMPLIFIED

- R scripts, environments, workspace.
- Setting a working directory.
- Basic R terms: functions, arguments to functions, objects, vectors, lists.
- Installing packages and loading a library of functions.
- Where can I go for help? Package documentation, the help button and the help function.

MODULE THREE: EVERYDAY DATA MANAGEMENT

- Matrices, arrays and data frames.
- Creating data.
- Importing and exporting data in various formats.
- How does R treat continuous and categorical data?
- Merging data sets.
- Data manipulation using package ‘car’ and an introduction to ‘tidyverse’.
MODULE FOUR: DESCRIPTIVE STATISTICS AND GRAPHS

- Measures of central tendency.
- Measures of dispersion.
- Descriptive graphics using basic functions: bar plots (simple, stacked, grouped); histograms (with and without a normal curve; density histograms); scatter plots (simple; enhanced scatter plot conditioning on a third variable using package 'car').

MODULE FIVE: SUMMATED SCALES IN R

- Case study: Measuring attitudes to immigration using the General Social Survey 2014.
- Summated scales.
- Reliability analysis.
- Data manipulation using 'tidyverse'.

MODULE SIX: ORDINARY LEAST SQUARES REGRESSION

- Case study: Exploring factors that are associated with attitudes to work in Wave 6 of the World Values Survey 2010-2014.
- Bivariate regressions.
- Moving to a multivariate regression model.
- Fitting interactions terms Visualizing interactions using 'ggplot2' and 'effects'.

SHORT TASTER SESSION (10 - 15 MINUTES) BIG DATA AND R IN BRIEF

- Accessing big data using R.
- Big data modelling options available in R.