



RonR! First Seminar Series

R is a free programming language designed to apply statistical techniques and produce data visualization. Among its endless capabilities, R is famous among statistical software for its flexibility in creating functions, packages, data visualization, for being open source and for offering a wide range of statistical tools. In the first series of seminars we aim at introducing you to the R programming language, making the transition as seamlessly as possible.

The outline of the seminar series is as follows:

In the first seminar we will show you the power of R. We will go through the strengths of R and how you can use it to obtain meaningful results. In the next session we will go through the building blocks of the R language and understand some traditional R features. We will guide you through simple examples on how the R language works.

After you've mastered these building blocks, in the third seminar we will start to get our hands dirty by cleaning a dataset. We will go through each step of manipulating a dataset and making it as workable as possible. Next, we'll produce everything from contingency tables to regression output.

The fourth seminar will be specialized on R graphical capabilities. Lastly, we will learn some commands that will allow you to make your data analysis much easier such as loops, conditional executions, among others.

We intend to make some, if not all the seminars, as interactive as possible. For that, we plan to lean R by using R. We will develop the exercises inside the program so we can play and practice anytime we want. Each session might have different or even paired instructors that will present and practice the examples with you.

Don't worry if you have:

1. Little to no experience with R.
2. No previous programming experience or no previous programming with other statistical commercial software.
3. Limited statistical knowledge

However, we do have a requirement:

4. You are motivated enough to put up with some initial (and continuing) frustration :)

Seminars Program:

Please, bring your laptop with R already installed (and possibly with the battery fully charged).

All seminars are free but If you're interested in participating, please email Jorge Cimentada at jorge.cimentada@upf.edu to be included in our mailing list.

Seminar 1. An overview of the strengths of R – April 13, 2016 from 15.00 to 16.30; room: 24.120.

Instructor: Jorge Cimentada

- Welcome to the group and seminar series
- Brief intro to R
 - First steps using R
 - The webpage of the R project
 - Installation of R
 - Contributed packages and their installation

- How to work with R: Workspaces and scripts
- Graphs (ggvis, ggplot2)
- Interactive graphs(ggvis, plotly, shiny, Rcharts)
- Maps
- Summary statistics output(stargazer package)
- Regression output(stargazer package)
- Making reports, PDF's and powerpoints through R (RMarkdown package)
- Programming exercises and exams in R (swirl and exam package)

Seminar 2. Getting familiar with R objects – April 20, 2016 from 15.00 to 16.30; room: 24.S01

Instructors: Basilio Moreno and Jorge Cimentada

- Basic objects: Vectors
 - Numerical, logical and character vectors
 - Generation of regular sequences
 - Vector arithmetic
 - Matrices and lists
- Basic objects: Matrices and lists
 - Creation and modification of matrices
 - Creation and modification of lists
 - Subsetting matrices and lists
- Basic objects: Data frames (I)
 - Creation of data frames
 - Adding new variables
 - Subsetting data frames
- Basic objects: Data frames (II)
 - Working with data frames: Functions attach and with
 - Merging data frames
 - Labelling variables

Seminar 3. Manipulating and processing data – April 27, 2016 from 15.00 to 16.30; room: 24.120

Instructor: Daniel Ciganda

- Importing data from SPSS, EXCEL and STATA
- Basic data manipulation:
 - Creating new variables
 - Recoding variables
 - Renaming variables
 - Selecting and removing variables
 - Filtering and subsetting variables
 - Grouping variables
 - Merging datasets
- Basic descriptive analyses:
 - Computation of numerical indicators (mean, median, variance, etc.)
 - Frequency tables for categorical variables
 - Correlations of two numerical variables
 - Contingency tables for two categorical variables
 - Stargazer package for **outputs**
- Simple and multiple regression models **output**
- Other regression outputs: Logistic regression model.

Seminar 4 An introduction to R programming – May 11, 2016 from 15.00 to 16.30; room: 24.120

Instructor: Bruno Arpino

- Basic programming:
 - Conditional execution.
 - Repetitive execution.
 - Apply(), sapply() and lapply() functions.
 - Writing your own functions.
 - Monte Carlo simulations.

Seminar 5. An introduction to sequence analysis – May 17 2016 from 15.00 to 16.30; room: 24.120.

Instructor: Albert Julià

PROGRAM TO BE DEFINED

The organizers of the RonR! working group

Jorge Cimentada (jorge.cimentada@upf.edu) and Bruno Arpino (bruno.arpino@upf.edu)

Research and Expertise Centre on Survey Methodology (RECSM)

Universitat Pompeu Fabra