R-intro: course contents

João Gonçalves

7 de Julho de 2018

Contents

Overview **1**

Pre-requisites **1**

Course content by session **1**

Session 1 2
Session 2 2
Session 3 2
Session 4 2

Overview

In this introductory course to R you will learn how to master the basics of this beautiful and immensely useful open-source language. R currently has a growing number of users and its applications in both academia and in the private sector are also increasing in a yearly basis.

The course will provide an introduction to the fundamental elements of R, starting from basic data structures and functions up to more complex stuff such as `if` conditionals and `for` loops. Some basic notions regarding the *tidyverse* approach to data manipulation (`dplyr` package) and visualization (`ggplot2`) will also be explored.

With the knowledge acquired from this course, you will be ready to undertake your own data analysis and (hopefully) be capable of pursuing more advanced topics.

Our exploration will start by focusing on some key aspects of R as programming language, its syntax and slowly introduce some of the functions used for basic statistics such as those used for summarizing data (e.g., mean, standard-deviation), pairwise correlation and linear models.

Pre-requisites

Don't worry we will start slowly! :-) As such, you don't really need any prior knowledge of computer programming (although some previous experience will make things easier) however, some familiarity with basic mathematical and statistical concepts and notions is assumed (e.g., distributions, centrality and dispersion measures, quantiles).

Course content by session

From vectors to sophisticated graphical data summaries with the `ggplot2` package these are the contents of this course:

Session 1

- Exploring RStudio environment and workflow
- Understanding the basics of *Rmarkdown* format
- Starting a work session
- Using the help system
- R as a (sophisticated) calculator
- Vectors and matrices - understanding the differences
- Vector and matrix indexation

Session 2

- Time for recap session #1
- Missing data: `NA` values
- Data frames (as an extension of matrices)
- Exploratory Data Analysis: scatter-plots and correlation analyses
- Factor variables
- Plotting and comparing distributions with boxplots
- Hypothesis testing (t-Test)
- Analysis of variance (one-way ANOVA)

Session 3

- Lists
- User-defined/custom functions
- If conditionals
- For loops
- Combining it all together

Session 4

- A brief tour to the `tidyverse`
- Data manipulation with the `dplyr` package
  - Main `dplyr` verbs
  - The pipe `%>%` operator
- Making pretty graphics with `ggplot2` - design principles
- Introduction to main types of plots/graphics:
  - Scatter-plots
  - Regression, trend and smooth lines
  - Histograms
  - Boxplots

- Modifying plots
- Making plots by groups