

Lab 1: Introduction to R

Business Statistics

Miguel Salema

Católica-Lisbon SBE

October 31, 2023



Section 1

Introduction

- Presentation
- Why R
- Why Quarto
- The Basics of R
- The Basics of Quarto

- Who am I?
- Contact information:
 - email: `misalema@ucp.pt`;
 - Office: 503;
 - Office hours: 1 p.m. on Thursdays;
- In case you email me a question, use this guide about how to ask a good code related question.
- Who are you?
- Moodle materials.

This course will provide you with the tools to answer business- and economics-related questions based on **real world data**. You will be able to answer questions such as:

- How does education affect wages?
- What is the value of implementing a new CRM system?
- Which prices maximizes the profits?
- Which customers should be targeted for your next campaign?

Most importantly, you will be aware of the potential pitfalls you may incur if the analysis is not performed carefully.

This will prepare you for your thesis.

- Basics of R
- Basics of Quarto

R is a free software, easy to install and runs in multiple OS.

- 1 R is **free**;
- 2 R is **popular**: in both companies and academia;
- 3 R is **powerful**: can handle large data and long computations, and the graphical visualizations are unmatched in other programs;
- 4 R is **flexible**: R-related environments allow to develop a large variety of projects such as these notes, interactive websites, research papers, etc. . .
- 5 R is **well supported**: a lot of documentation and forums. Excellent documentation on packages.

- Replicable documents;
- No formatting problems;
- Organization: everything in the same place;
- Dynamism saves time;
- ...

When you first open RStudio, you will be greeted by three panels:

1. The interactive R console (entire left)

- Here you can type commands and see their output.

2. Environment / History (tabbed in upper right)

- Environment: This tab stores the things you create during your R session.
- History: This is a record of your activity, you can save it for later use.

3. Files / Plots / Packages / Help / Viewer (tabbed in lower right)

4. The R script(s) and data view

- The R script is where you keep a record of your work.

Everybody has R installed? Let's pause for 5 minutes, to finish the installation.

Exercise 1

- Create a folder for today's class
- Create an R project there
- Open this class RScript
- Create a Quarto Document

Section 2

Getting Started

There is also an R R cheat sheet for reference.

- Use # to comment. Everything in the right of a 'hashtag' will be omitted by the compiler;
- Comment all the things you (try to) do!
- There are data classes: numeric, character, dates, logical, ...
- We can create different ordered structures: a vector, a matrix, a list, ...
- We make use of functions.

You can use the Console just like a calculator. Try performing some calculations:

```
1+1  
## [1] 2
```

```
2*(10-2)*4  
## [1] 64
```

```
8^(1/3)  
## [1] 2
```

Function	Description
<code>abs(x)</code>	Takes the absolute value of x
<code>log(x, base=y)</code>	Takes the natural logarithm of x
<code>exp(x)</code>	Returns the exponential of x
<code>sqrt(x)</code>	Returns the square root of x
<code>factorial(x)</code>	Returns the factorial of x ($x!$)

R has built-in mathematical functions. For example:

```
sqrt(25) # square root  
## [1] 5
```

```
exp(2) # e^2  
## [1] 7.389056
```

```
log(1) # natural log  
## [1] 0
```

R does not have a command for declaring a variable. A variable is created the moment you first assign a value to it. To assign a value to a variable, use the `<-` sign. To output (or print) the variable value, just type the variable name:

```
name <- "John"
age <- 40

name      # output "John"
## [1] "John"
age       # output 40
## [1] 40
```

Once you store an object in your environment you can interact with it directly by using its name

```
age + 10 # interact with object
## [1] 50
```

Note that the previous command does not change the value of age because we did not store the result in a variable. The following command would change the value of age.

```
age <- age + 10
age
## [1] 50
```

Objects or variables can also contain text (aka strings). Everything in R is an object, and an object can contain anything.

```
my_class <- 'BRM is awesome'
```

R is case sensitive and spelling sensitive.

```
# try running  
# My_class
```

Coding is just applying functions

```
vec <- 1:10  
mean(vec)  
## [1] 5.5  
seq(1, 10, 2)  
## [1] 1 3 5 7 9
```

Exercise 2

Build the following vector:

- 1 starting in -10;
- 2 Ending in 89;
- 3 Going 3 by 3.

```
# this is a vector
c(3, 5)
## [1] 3 5

#this is the mean of that vector
mean(c(3, 5))
## [1] 4
```

Exercise 3

Search help and identify the arguments of the `mean()` function. Then, calculate the mean and standard deviation of the vector `c(3, NA, 16, 20)`

```
# I can apply functions to every element of a vector  
c(3, 5) + 1  
## [1] 4 6
```

Exercise 4

Country A has a GDP of 100 in $t = 0$ and grows 3% a year. Build the time series of that GDP

```
#For Windows/Linux (make sure you have forward slashes)
setwd("C:/path/to/directory/containing/the/files/") # change the file's path to you

#For Macs
setwd("/path/to/directory/containing/the/files/")

getwd()
```

You can check what files you have in your working directory by running the command

```
list.files()
```

```
load("use another data")
```

Section 3

Quarto

First thing: Render

Now let's go over the details

Exercise 5

Add a new citation to the file.

Exercise 6

Place the average wage as inline R code.

- Complete DataCamp's first chapter of the Introduction to the Tidyverse course: Data wrangling