Lecture 1

Daniel Ballesteros-Chávez

September 25, 2021

Contents

1 Introduction to R

- What is R?

R is an open source statistical software.

- How to get it?

You can get R directly on their website `https://www.r-project.org/`

- What is RStudio?

RStudio is an integrated development environment (IDE) for R.

It is a good tool to get started if you are not used to scripting programming. It has very helpful visual features.

I use emacs and it is difficult for me to get used to Rstudio. My advice: use whatever you find more comfortable.

- Basic commands
  - getwd()
  - setwd()
  - install.packages()
  - library()
  - df <- read.dbf()
  - df$NAME
  - dim(df)
  - class(df)
  - ls()
  - rm(list=ls())
  - sum(df$NAME)
  - as.numeric()
  - as.character()
  - df$NEWVAR <- df$NAME
  - table()
  - tapply, sapply, lapply, etc.

2 The Working Directory

Get my Working Directory (a.k.a. where am I?)

```
getwd()
```

Set my Working Directory (a.k.a. change directory)

```
setwd("/home/ennaniux/Documents/R_Modelling")
```

How does this work in Windows?

```
setwd("C:/home/ennaniux/Documents/R_Modelling")
```

or maybe

```
setwd("C:\\home\\ennaniux\\Documents\\R_Modelling")
```

3 Installing packages

The Packages are sets of tools that can be downloaded from different servers around the world. Different packages have different R functions for specific purposes.

For example, the foreign package allows R to read different data set files like .sav, .dbf, and other file extensions.

In order to install the package foreign we type in the console

```
install.packages("foreign")
```

this will provide a list of possible servers to choose from, and you can select one close to your location.

4 Simple manipulations; numbers and vectors

The simplest data structure R operates on is the **numeric vector**, which is a single entity consisting of an ordered collection of numbers.

```
x <- c(1,3,5,9)
x
```

The syntax is equivalent to

```
c(1,3,5,9) -> y
y
```

and

```
assign("z",c(1,3,5,9) )
z
```

5 Writing a data frame

Write a data frame by specifying the columns:

```
df <- data.frame(
"NAME" =  c("Aleksandra", "Hugo", "Piotr", "Ewa"),
"AGE"  =  c(29,35, 39, 33),
"HEIGHT"= c(1.68, 1.83, 2.03, 1.66) )
df
```

What is the dimension of the data frame?

```
dim(df)
```

What are the variable names of the data frame?

```
names(df)
```

6 Reading a data frame

From a .csv file

```
df <- read.csv('./path_to/file.csv')
```

From a .dbf file

```
library(foreign)
df <- read.csv('./path/to/file.dbf')
```

From a .sav file

```
library(foreign)
     df <- read.spss(file='./path/to/file.sav', to.data.frame=TRUE)
     str(df)   # show the structure of the data frame
```

From the clipboard

```
df2 <- read.table(file = "clipboard", sep = "\t", header=TRUE)
```

7 Creating a new variable

- Graphics
- Reading data
- Markdown