PROGRAMMING WITH R

WHAT IS R?

- «R is a [programming] language and environment for statistical computing and graphics» (https://www.r-project.org/about.html)

- It is a free software (GNU GPL license)

- Created in 1995 by
  Ross Ihaka (New Zeland)
  and Robert Gentleman (Canada)

- The name? Cfr. the initials of the creators
  also, it is a was based on another programming language, called «S»

R AND RSTUDIO
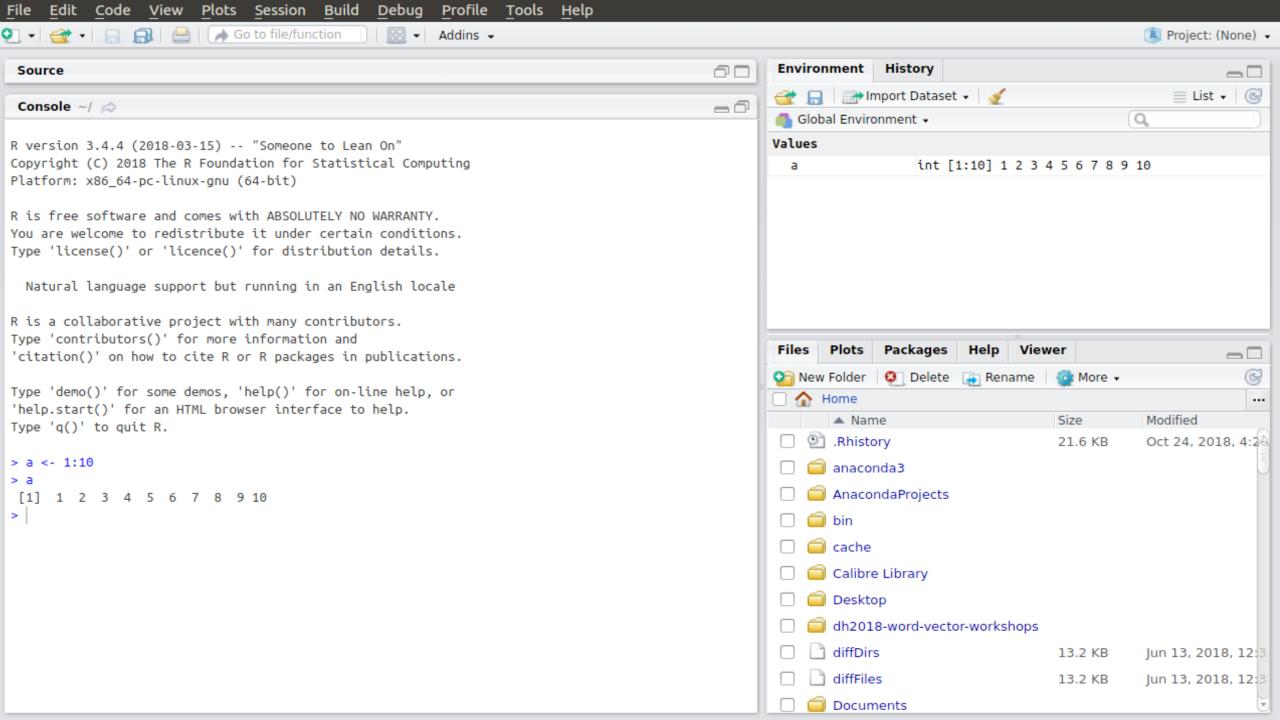
- R is the programming language

R

- R is the programming language

```
rsimone@rsimone-Inspiron-13-5378:~$ R

R version 3.4.4 (2018-03-15) -- "Someone to Lean On"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> a <- 1:10
> a
 [1]  1  2  3  4  5  6  7  8  9 10
> for(i in 1:10)
+ print(i:10)
 [1]  1  2  3  4  5  6  7  8  9 10
[1]  2  3  4  5  6  7  8  9 10
[1]  3  4  5  6  7  8  9 10
[1]  4  5  6  7  8  9 10
[1]  5  6  7  8  9 10
[1]  6  7  8  9 10
[1]  7  8  9 10
[1]  8  9 10
[1]  9 10
[1] 10
>
```

R AND RSTUDIO

- R is the programming language
- Rstudio is a «working environment», that makes it easier to work with R

Go to file/function　　Addins

Project: (None)

Source

Console　~/

```
R version 3.4.4 (2018-03-15) -- "Someone to Lean On"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> a <- 1:10
> a
 [1]  1  2  3  4  5  6  7  8  9 10
> 
```

Environment　　History

Import Dataset　　　　　　List

Global Environment

Values

a　　　　　　　int [1:10] 1 2 3 4 5 6 7 8 9 10

Files　Plots　Packages　Help　Viewer

New Folder　　Delete　　Rename　　More

Home

	Name	Size	Modified
	.Rhistory	21.6 KB	Oct 24, 2018, 4:2
	anaconda3		
	AnacondaProjects		
	bin		
	cache		
	Calibre Library		
	Desktop		
	dh2018-word-vector-workshops		
	diffDirs	13.2 KB	Jun 13, 2018, 12:3
	diffFiles	13.2 KB	Jun 13, 2018, 12:3
	Documents		

GETTING STARTED

- Open Rstudio and test some basic functionalities

- Mathematical operations

- Definition of variables
  with the arrow function:
  <-

- Operations with variables

- Basic commands: «paste» and «c» (concatenate)

SETTING A WORKING DIRECTORY

- In three different ways:
  - in the «Files» panel on the bottom-right of the screen
    (navigate, click on «MORE» -> «Set as Working Directory»
  - in the «Session» menu
    («Set Working Directory» -> «Choose Directory»)
  - (for advanced users) directly in the Console
    setwd(«my/working/directory")

Try to set as working directory the «2_Stylometrie» directory in your computer

CODE REUSE (AND R PACKAGES)

- In many cases, you don't need to write all the code by scratch:
  you can simply reuse (and adapt) code written by others
    today, we'll re-use code that I wrote

- The non-plus-ultra of code reuse, are the R Packages:
  R packages are collections of scripts, that make it possible to run complex operations
  through simple commands
    for example, see the «Stylo» package:
    https://github.com/computationalstylistics/stylo/tree/master/R

CODE REUSE (AND R PACKAGES)

- To call the 2,000 lines of code in Stylo, you can use a «simple» command:

stylo(gui = FALSE, corpus.lang="French",
analysis.type="CA", mfw.min=2000,
mfw.max=2000, mfw.incr=100,
distance.measure="dist.wurzburg")

CODE REUSE (AND R PACKAGES)

- To call the 2,000 lines of code in Stylo, you can use a «simple» command:

- Some packages (like Stylo) come with a graphical user interface

FIRST EXERCISE

- Let's use the «XML» package in R!

  …open the XML_Exercise.R file in RStudio

SECOND EXERCISE

- Stylomerty with the «stylo» package in R

    …open the Stylo_Exercise.R file in RStudio

THIRD EXERCISE

- Keyness analysis with the «quanteda» package in R

  …open the Keyness_Exercise.R file in RStudio