

Install R and RStudio

Francesca Vitalini

10/12/2018

Introduction

These are the instructions for the workshop “Introduction to R with tidyverse”.

Workshop Objectives

This is what we will try to achieve in the workshop

- Become familiar with R and its community
- Learn the basics of R language and its concepts
- Import data in R
- Manipulate data using tidyverse (dplyr)
- Visualization of the data with tidyverse (ggplot2)
- Make a report out of your data with Rmarkdown
- Obtain good basis in R tidyverse for more advanced courses

Pre-requisites

In order to profit from the workshop, the following pre-requisites are important:

- No R knowledge is required, some is an advantage
- Suitable for complete beginners
- An understanding of programming of other programming languages can be beneficial
- Course participants are expected to bring their own laptop
- Make sure to have R and Rstudio installed

Instructions to install R

This workshop is in R. To install R 3.5.1 follow the instructions below for your platform.

R is a functional and object-oriented programming language and environment for statistical computing and graphics.

R is a free, open source, software and it is highly extensible through so-called *packages* which can be freely downloaded from CRAN

References:

- R: <https://www.r-project.org>
- CRAN: <https://cran.r-project.org/web/packages/>

UBUNTU 18.04 (and 16.04)

References

- R installation: <https://cran.r-project.org/bin/linux/ubuntu/README.html>
- Public key error: <https://stackoverflow.com/a/24732162>

Setup repository

From the terminal, open the file with the list of repositories

```
sudo gedit /etc/apt/sources.list
```

add the corresponding line at the end of the file

```
# 18.04 bionic beaver
deb https://cloud.r-project.org/bin/linux/ubuntu bionic-cran35/
# 16.04 xenial xerus
deb https://cloud.r-project.org/bin/linux/ubuntu xenial-cran35/
```

save and close.

Install R

In the terminal run

```
sudo apt-get update
```

Then install the R package by running

```
sudo apt-get install r-base
```

WINDOWS:

Download and execute the relevant installer from <https://stat.ethz.ch/CRAN/bin/windows/base/>

References:

- R installation <https://stat.ethz.ch/CRAN/bin/windows/>

MACOS

Download and execute R-3.5.1.pkg from <https://stat.ethz.ch/CRAN/bin/macosx/>

References:

- R installation <https://stat.ethz.ch/CRAN/bin/macosx/>

Instructions to install RStudio

In this workshop we will work in RStudio.

RStudio is a free and open-source integrated development environment for the R programming language.

To install RStudio IDE follow the instructions below.

References:

- RStudio: <https://www.rstudio.com>
- RStudio Cheat Sheets: <https://www.rstudio.com/resources/cheatsheets/>

Install RStudio

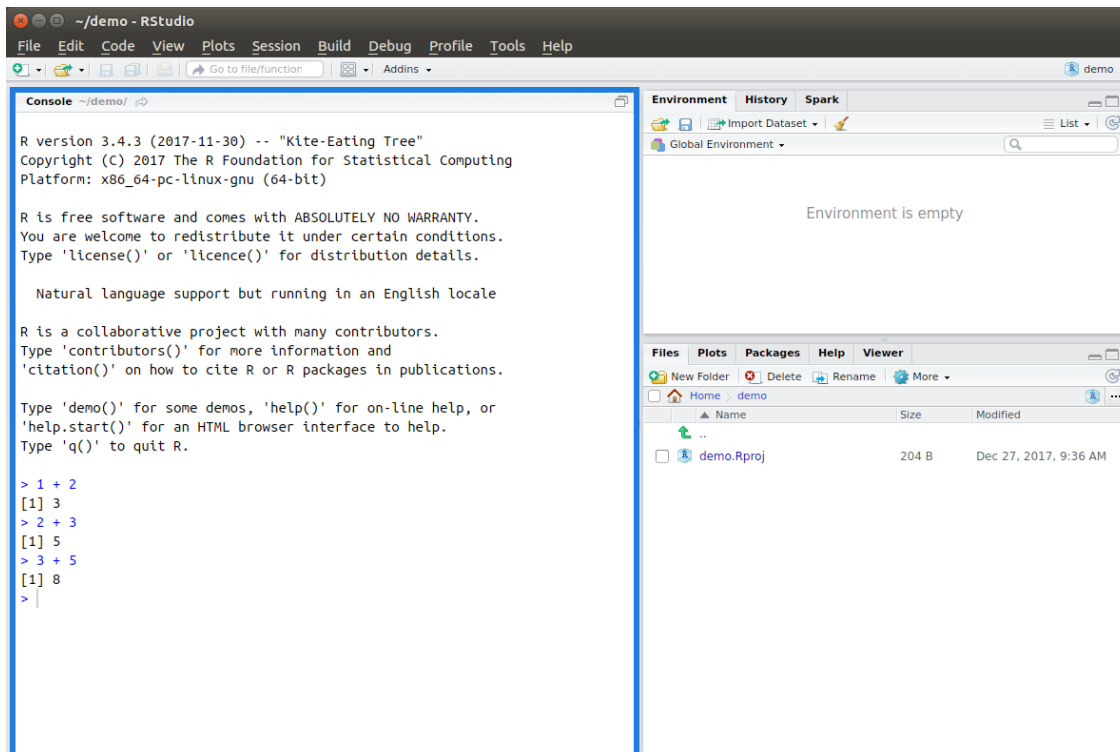
Download and execute the relevant installer for RStudio Desktop FREE from <https://www.rstudio.com/products/rstudio/download/#download>

Get Familiar wit R Studio

Due to time constraints during the workshop, it would be beneficial if you could start getting familiar with R Studio.

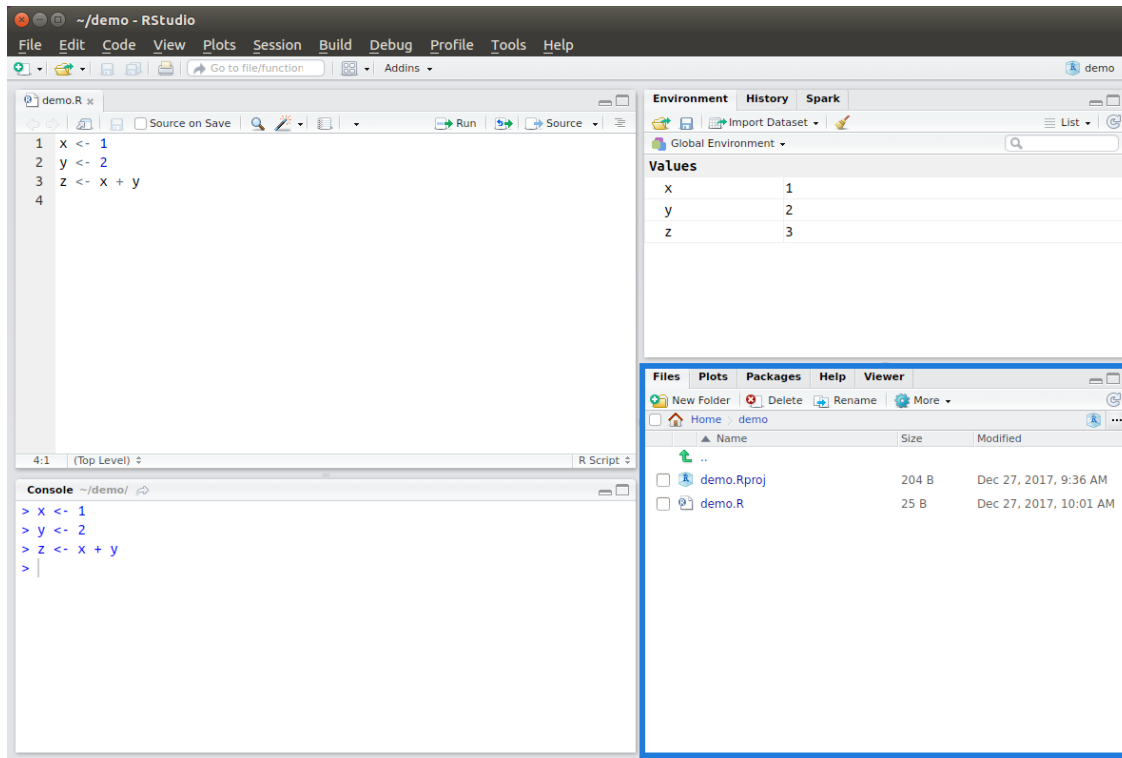
R Console

Here is where you execute R code.



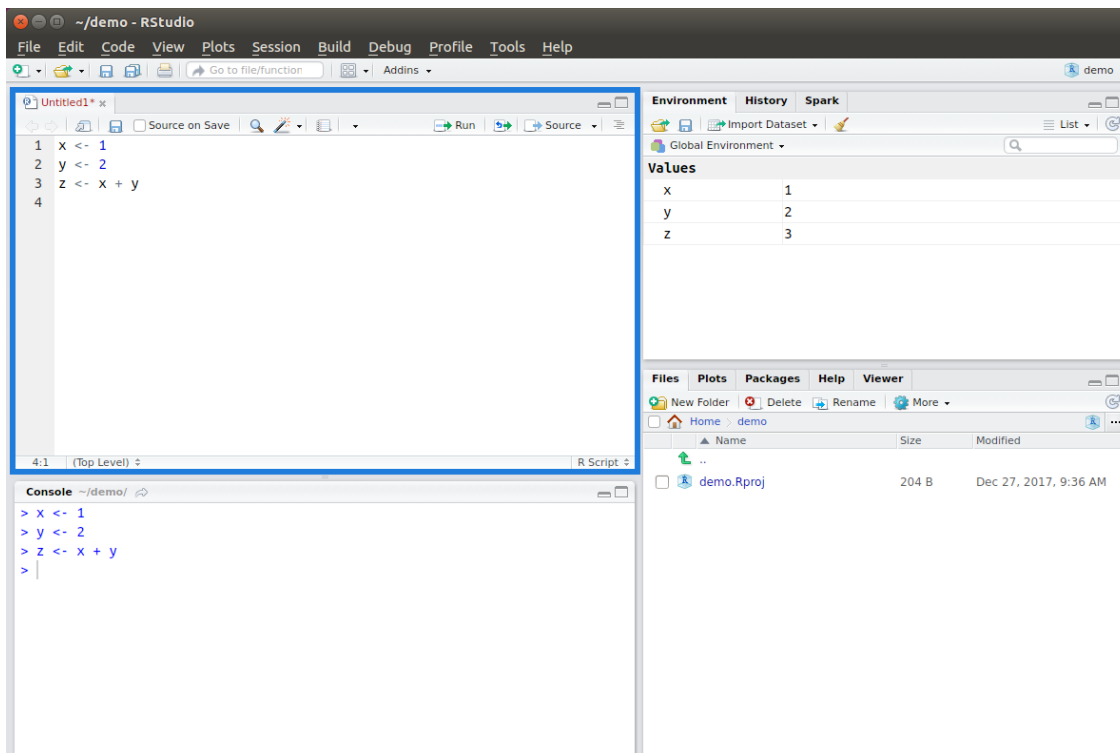
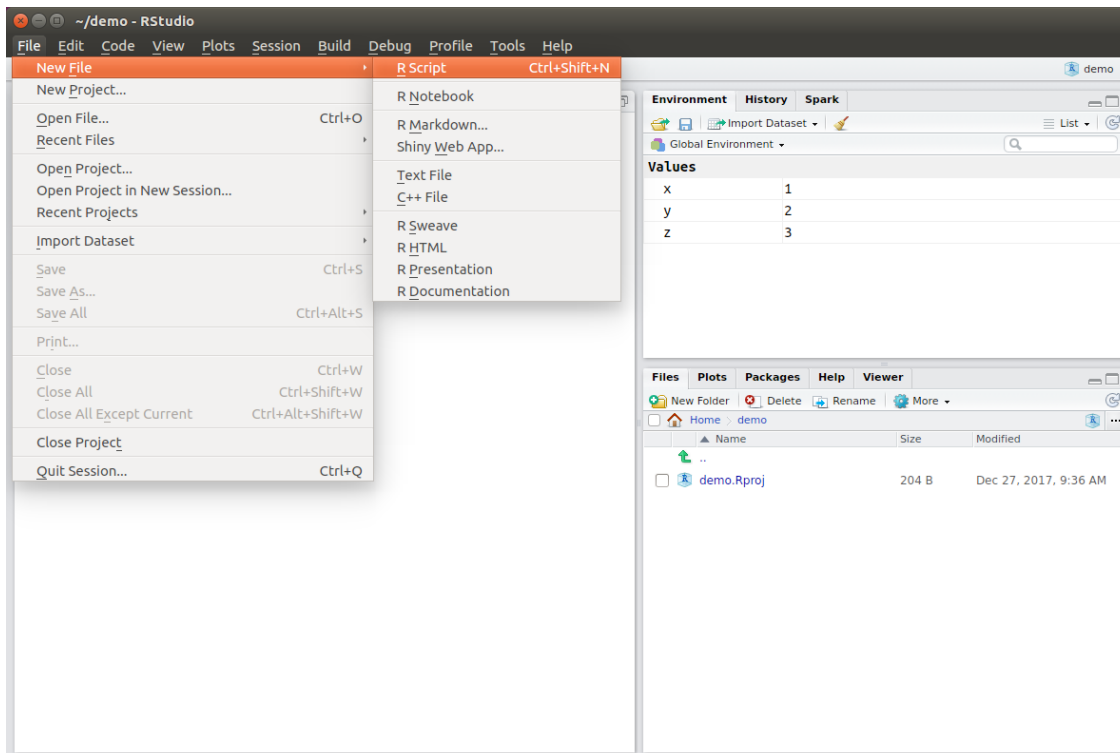
File Browser

Here you can browse your files.



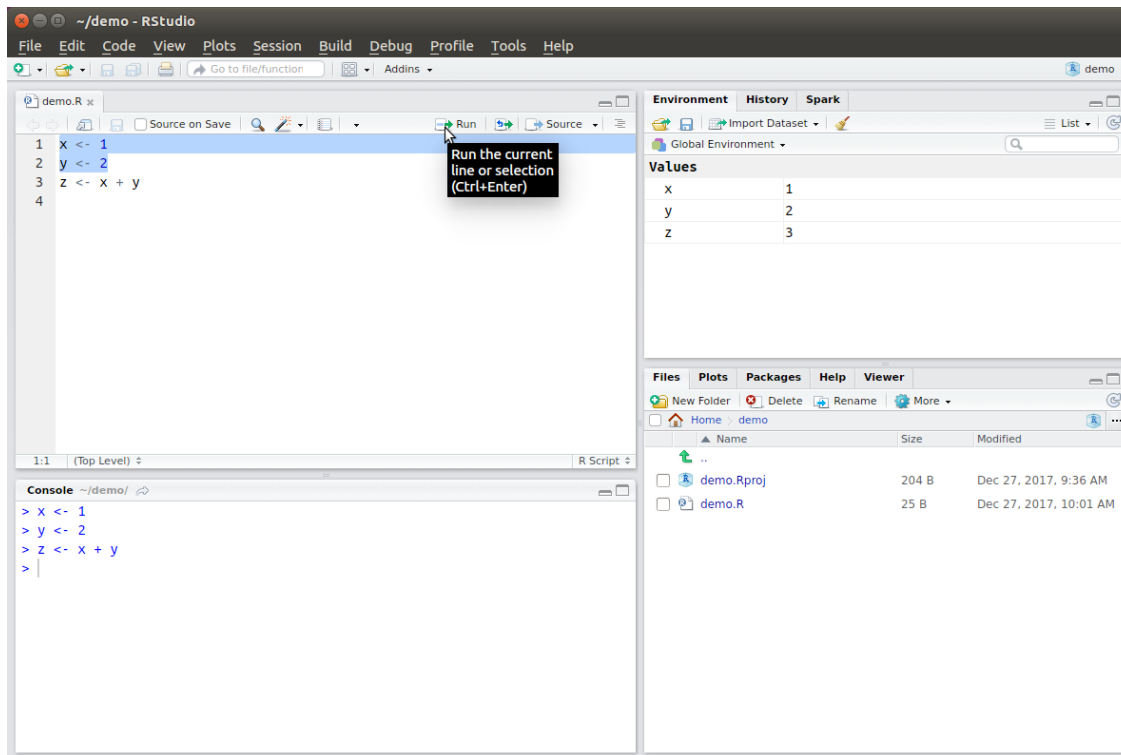
Script Files

Here is how to make a script where to write your R code.



To execute R code from a script.

Hit **Ctrl+Enter** to run the current line or selection:



The tidyverse package

In this workshop we will use a series of very known packages, which share a common way to represent data, so that they work in perfect harmony. They define the tidyverse framework. It includes

- ggplot2 - to make plots
- dplyr - for data manipulation
- tidyr - to clean the data
- readr - to read data stored in rectangular formats (e.g. csv)
- purrr - for enhanced functional programming in R
- tibble - introduces the tibble structure, an “efficient” dataframe
- stringr - for string manipulation
- forcats - tools to work with factors, i.e. categorical variables

In this course we will focus on **dplyr** and **ggplot2**.

References:

- tidyverse: <https://www.tidyverse.org>

Install tidyverse package

Go to the Console.

Type `install.packages("tidyverse")` and click **Enter** to install tidyverse.

Note that the installation might take some time.

Check that everything went fine by running in the Console: `library(tidyverse)`

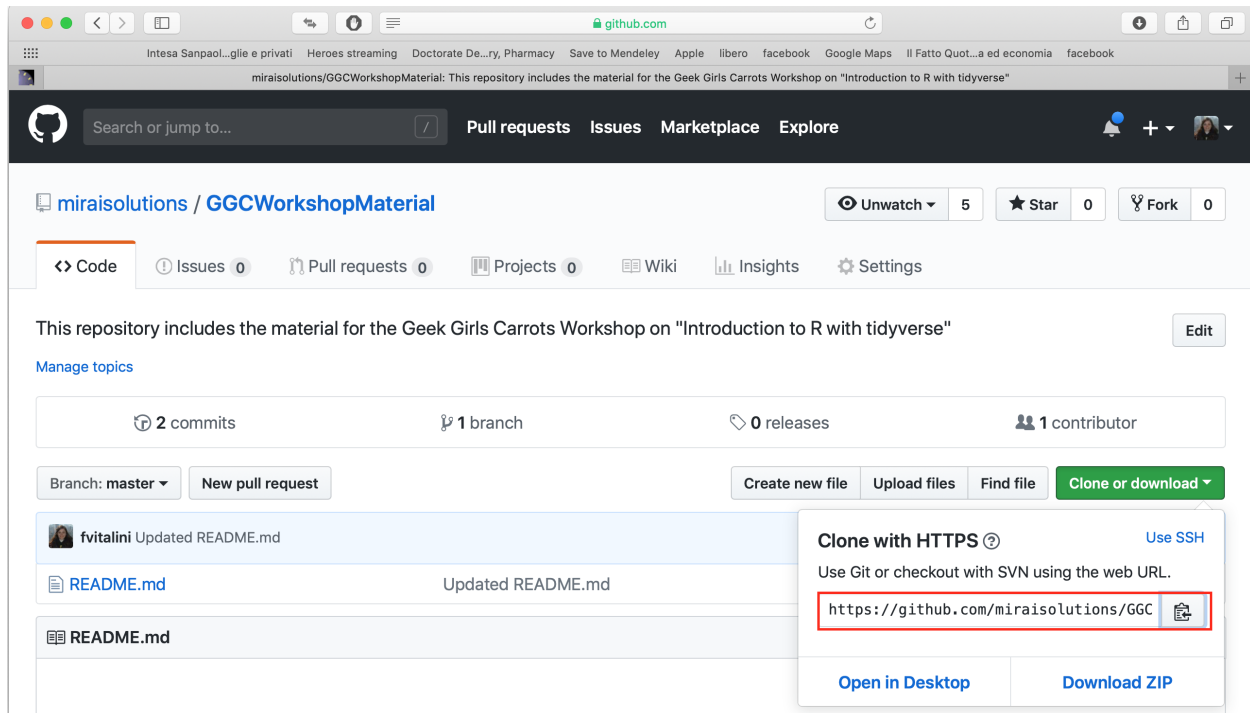
Get the material for the course

The material is stored on GitHub.

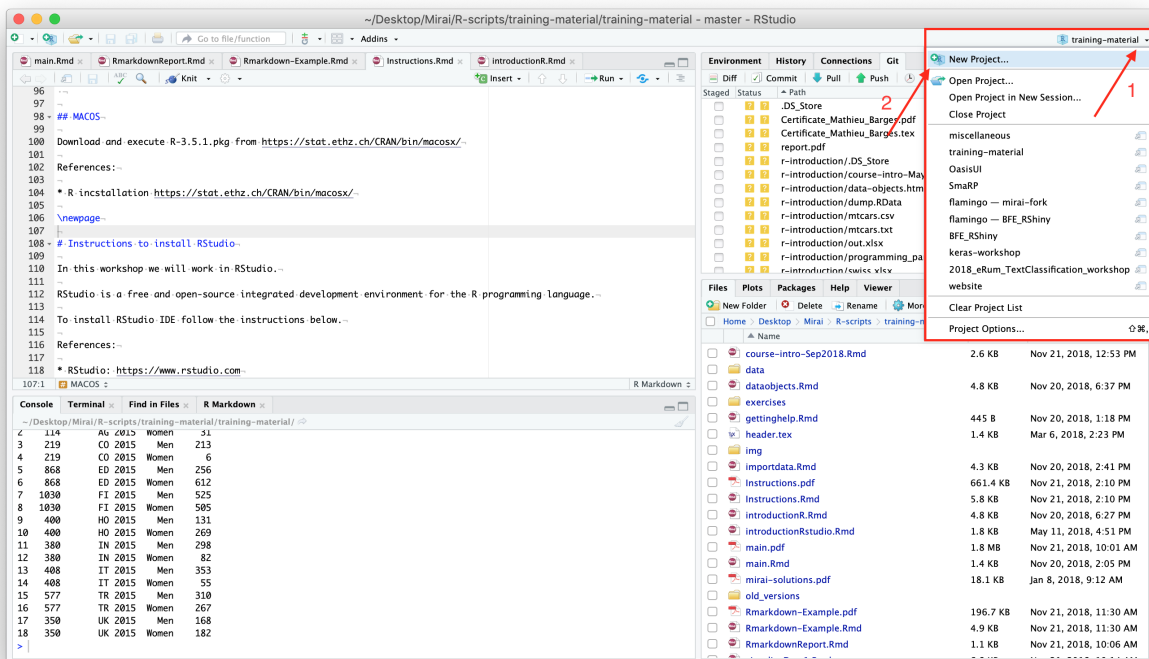
To access it you need a GitHub login. If you do not have one you can create it at <https://github.com> and here is the GitHub documentation.

Once you have logged in, navigate to <https://github.com/miraisolutions/GGCWorkshopMaterial>.

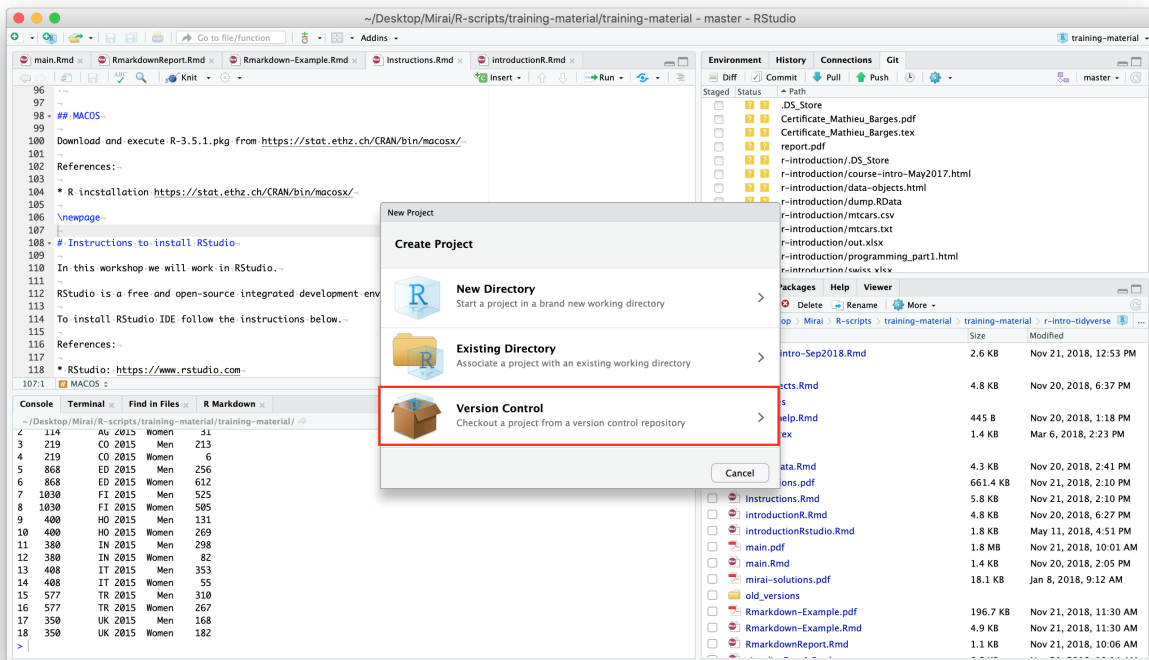
Click on Clone or download (green button to the right) and copy the url.



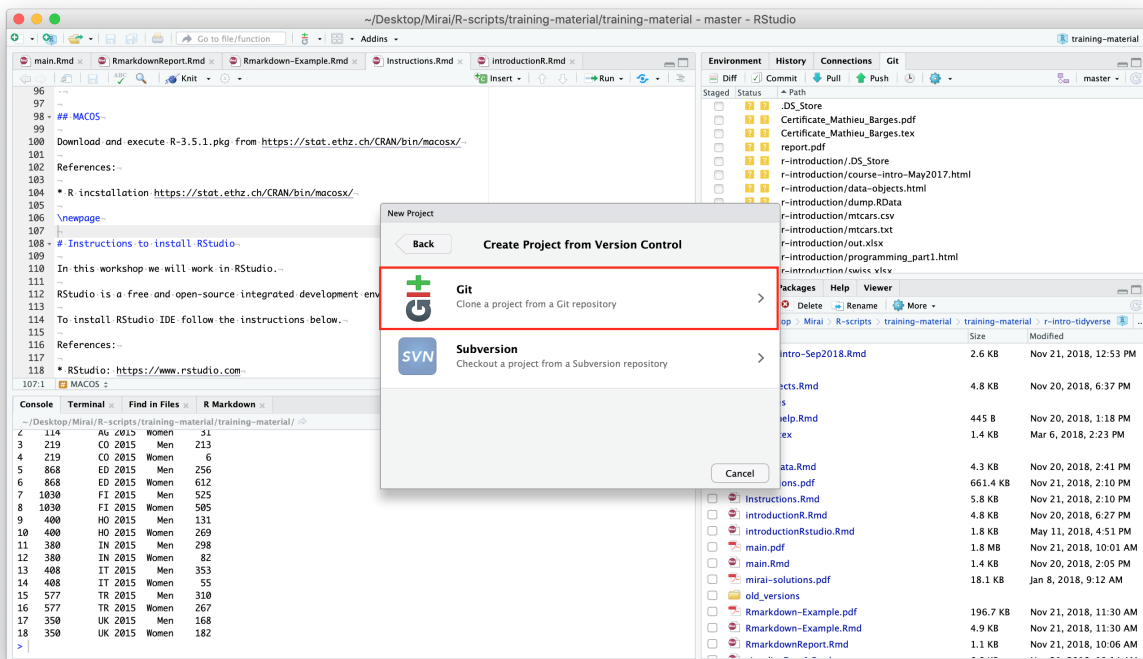
Then open RStudio. Click on the small arrow on the top right, next to the RStudio icon and the Project Name (if you do not have any project open will state Project (none)), and select New Project.



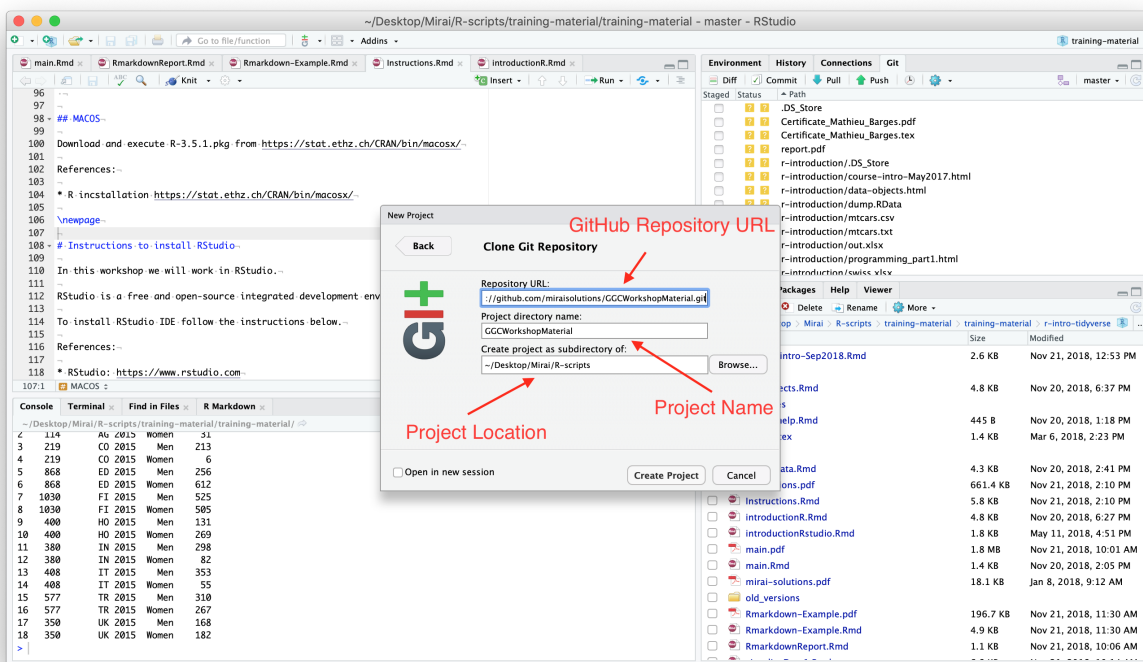
Select Version Control.



Select Git.



As Repository URL paste the GitHub URL. With Project Directory Name you can provide the name of the project. It is custom to use the same name as the GitHub repository (in this case GCWorkshopMaterial). The third field is where your project will be created. Choose a convenient path.



Click Create Project. RStudio will automatically open a window with the new project.

Outline of the Course:

Introduction 1. What is R? What is Rstudio? 2. How to get usefull packages? What is tidyverse? 3. Getting help. The CRAN community 4. Importing data in R 5. Data objects in R

Hands-on with Tidyverse: a Case Study 1. Filter and manipulate data with tydiverse (dplyr) 2. Add some statistics to the data 3. Plot data with tydiverse (ggplot2) 4. Make a report with Rmarkdown

The data

The dataset we will consider in this course is provided by the Swiss Federal Office.

It describes the occupation by gender in Switzerland between 1970 and 2015.

We have pre-processed the data for you so that it would be easier to work with.

The data is provided in two formats: `.xlsx` and `.csv`. Please have a look at it prior to the workshop.