

Basic Introduction to



Simon Knüsel

Forest Dynamics, WSL

simon.knuesel@wsl.ch

Goals

- Start to code
- Learn how to use R for the analysis of your data
- Know where and how to get help if you are stuck

Why use R?

- Reproducibility
- Keep original data untouched
- It is open source... with a large community
- Advanced statistics
- State-of-the-art graphics
- Powerful data manipulation
- Supports large datasets
- Fast computation
- Easier automation
- Anyone can contribute
- ...

RStudio

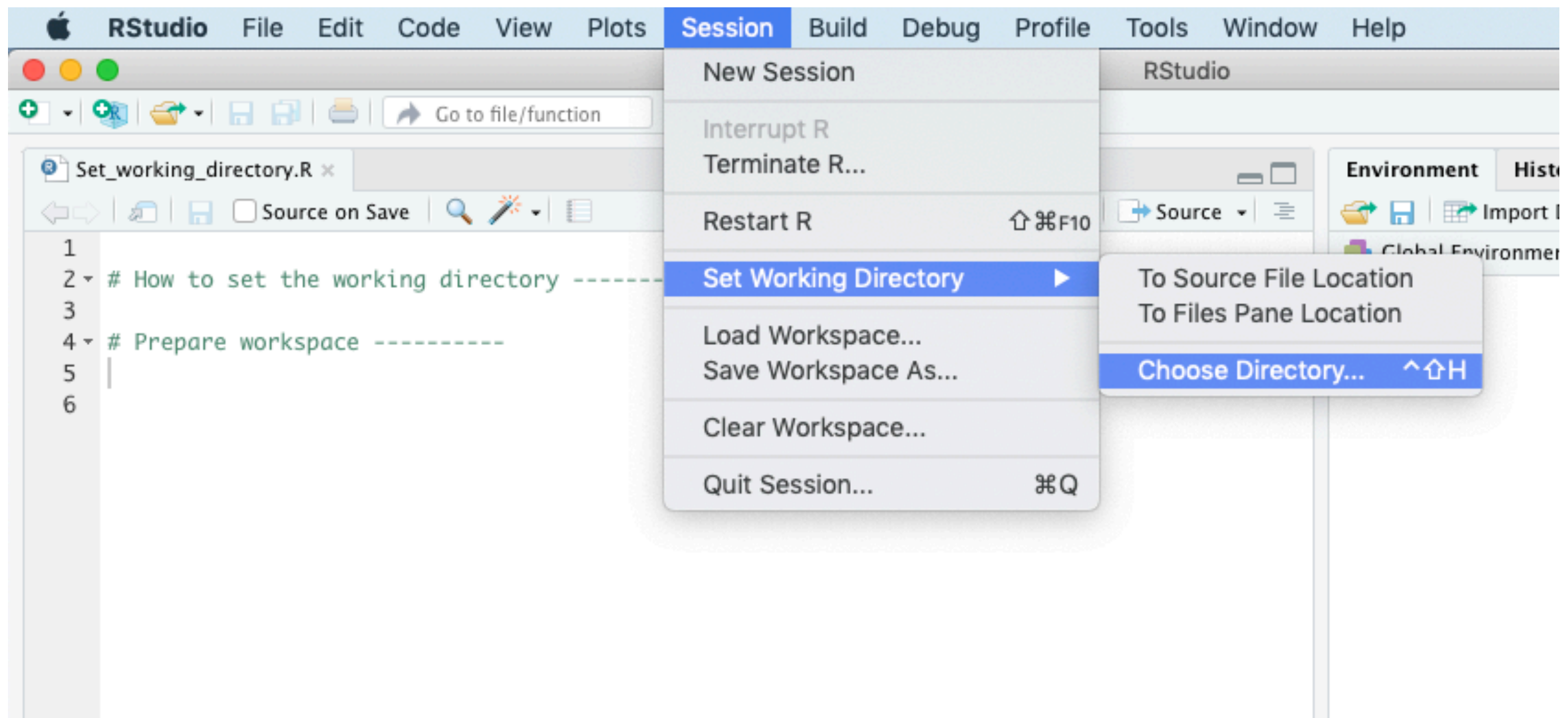


The image shows the RStudio desktop application interface. The top toolbar includes icons for file operations and a search bar. The main window is divided into four panels:

- Script editor (write code):** The top-left panel shows an R script file named 'Why_use_R.R'. The code includes comments and commands for reading a CSV file, replacing values, deleting data, and plotting.
- Environment / History (loaded objects):** The top-right panel shows the 'Global Environment' with a search bar. Below it, the 'Data' pane shows 'data_L1' with 998 observations of 4 variables.
- Console (executed code):** The bottom-left panel shows the output of the executed code, including the same commands as the script editor.
- Files / Plots / Packages / Help:** The bottom-right panel shows a file explorer view of the project directory, listing files like '.gitignore', '.Rhistory', '01_Data', '02_Slides', '03_Script', 'Introduction_to_R.Rproj', 'LICENSE', and 'README.md'.

Set working directory

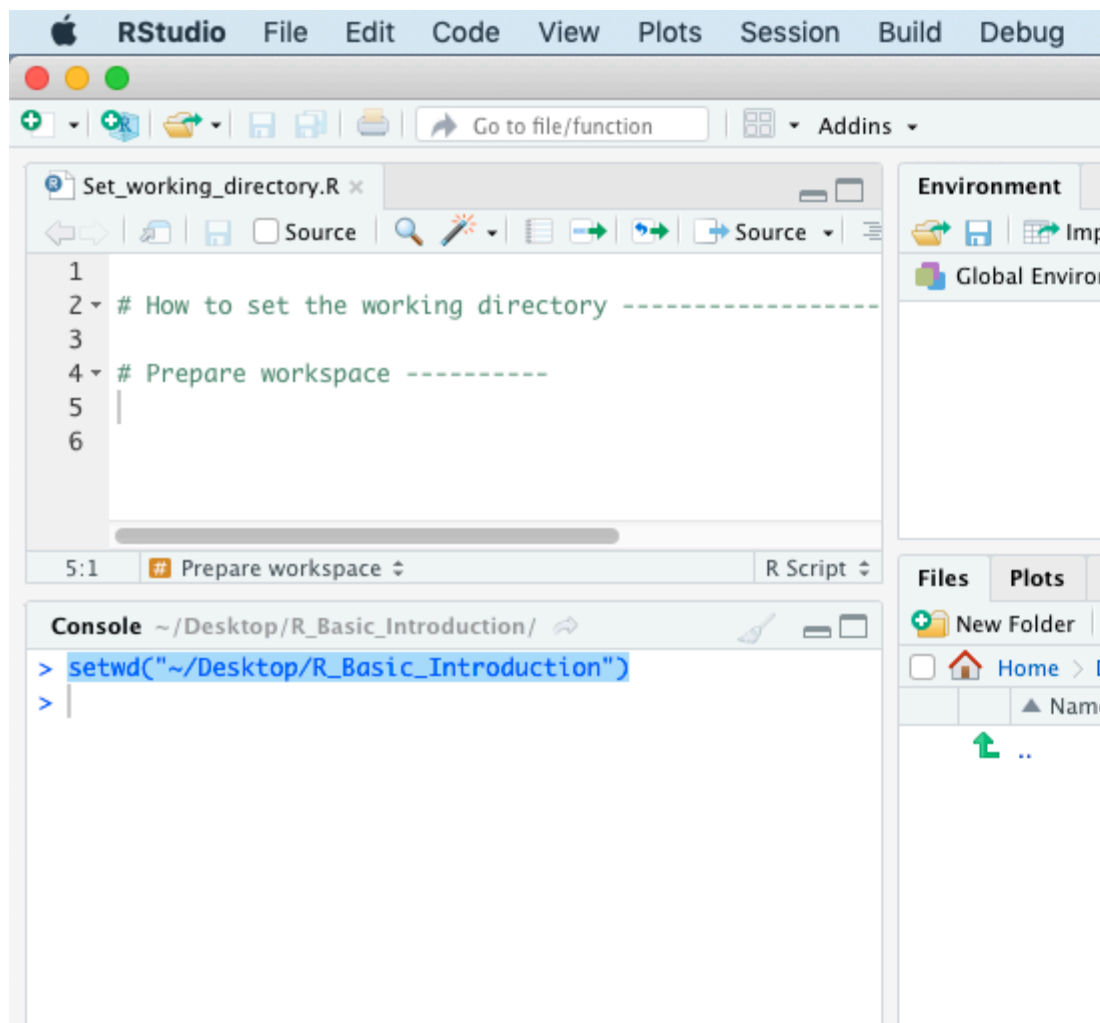
- 1) Choose the working directory from file
Session > Set Working Directory > Choose Directory...



Set working directory

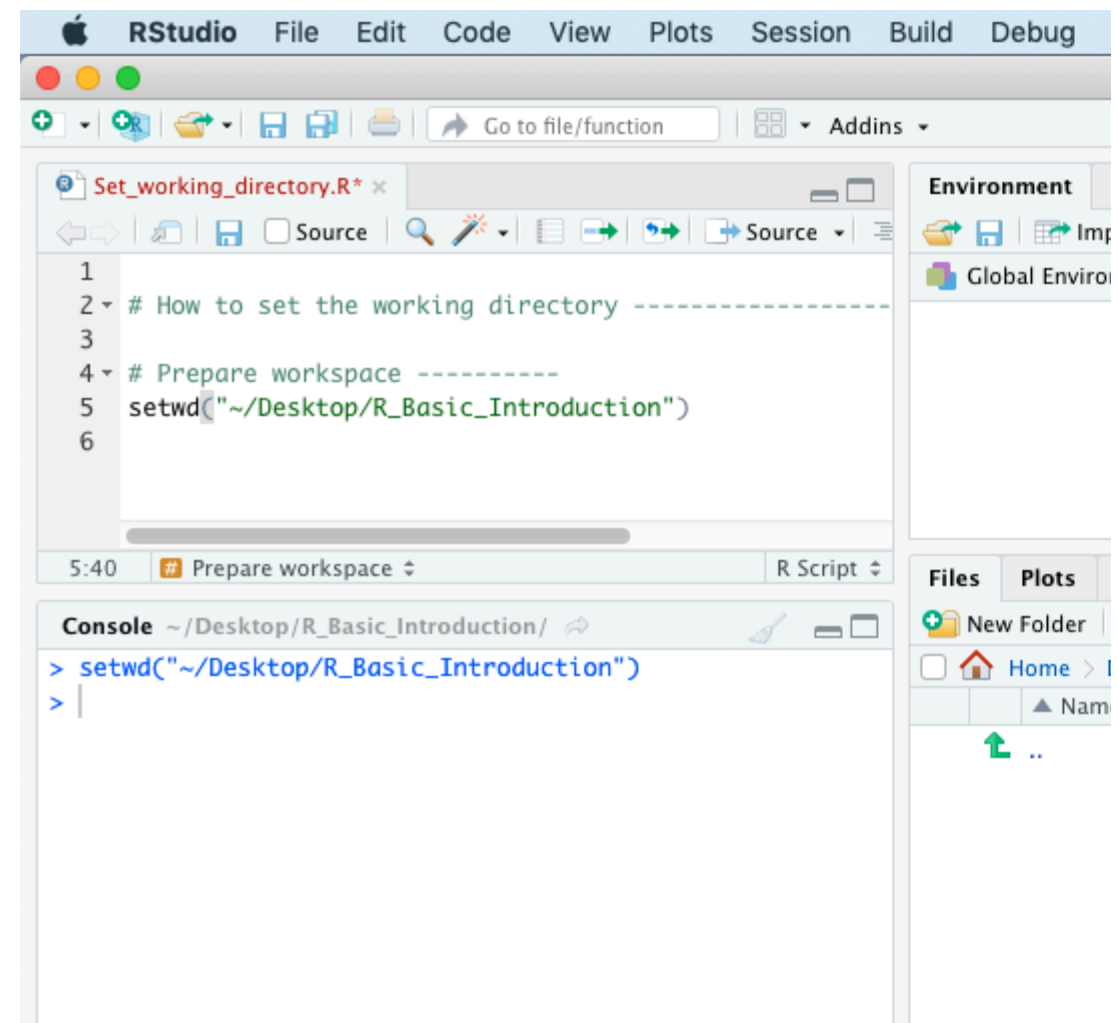
2) Copy the code output in the console

3) Paste it into the R-script



This screenshot shows the RStudio interface. The source editor on the left contains a script with comments: "# How to set the working directory" and "# Prepare workspace". The console at the bottom shows the command `setwd("~/Desktop/R_Basic_Introduction")` being executed. The status bar at the bottom indicates the current working directory is `~/Desktop/R_Basic_Introduction/`.

```
1  
2 # How to set the working directory -----  
3  
4 # Prepare workspace -----  
5  
6  
5:1 # Prepare workspace R Script  
Console ~/Desktop/R_Basic_Introduction/  
> setwd("~/Desktop/R_Basic_Introduction")  
>
```



This screenshot shows the RStudio interface after the command from the console has been pasted into the script editor. The source editor now shows the `setwd` command on line 5. The console remains the same, showing the command was executed successfully.

```
1  
2 # How to set the working directory -----  
3  
4 # Prepare workspace -----  
5 setwd("~/Desktop/R_Basic_Introduction")  
6  
5:40 # Prepare workspace R Script  
Console ~/Desktop/R_Basic_Introduction/  
> setwd("~/Desktop/R_Basic_Introduction")  
>
```

Functions

```
function_name(first argument, second argument, ...)
```

Example:

```
vec <- c(3, 6, 7, 4.2, NA, 6, 8)
```

```
mean(x = vec, na.rm = TRUE)
```

Help on functions:

```
?mean (?function_name)
```

Data structures

Vector

Elements separated by a comma

```
vec <- c(3, 6, 7, 4.2, 6, 8)
```

Data frame

Table with rows and columns

Species	Length (cm)	Height (cm)	Weight (kg)
Lynx lynx	112	65	22.5
Canis lupus	145	78	34

Indexing of vectors

```
vec <- c(3, 6, 7, 4.2, 6, 8)
```

```
# select the first element of the vector
```

```
> vec[1]
```

```
3
```

Indexing of vectors

```
vec <- c(3, 6, 7, 4.2, 6, 8)
```

```
# select the first element of the vector
```

```
> vec[1]
```

```
3
```

```
# select elements 1 to 3
```

```
vec[1:3]
```

```
3 6 7
```

Indexing of data frames

```
df <- airquality  
view(df)
```

```
# select first row
```

```
df[1, ]
```

```
# select third column
```

```
df[, 3]
```

```
df[row, column]
```



	Ozone	Solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
5	NA	NA	14.3	56	5	5
6	28	NA	14.9	66	5	6
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
10	NA	194	8.6	69	5	10
11	7	NA	6.9	74	5	11
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18

Indexing of data frames

select custom range


```
df[c(3:5, 7), 3:4]
```

	Ozone	Solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
5	NA	NA	14.3	56	5	5
6	28	NA	14.9	66	5	6
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
10	NA	194	8.6	69	5	10
11	7	NA	6.9	74	5	11
12	16	256	9.7	69	5	12
13	11	290	9.2	66	5	13
14	14	274	10.9	68	5	14
15	18	65	13.2	58	5	15
16	14	334	11.5	64	5	16
17	34	307	12.0	66	5	17
18	6	78	18.4	57	5	18

Shortcuts

	Mac	Windows / Linux
Run line	Cmd + Enter	Ctrl + Enter
# (Comment line)	Alt + 3	#
%>% (Pipe sign, used in dplyr package)	Cmd + Shift + M	Ctrl + Shift + M
Show keyboard shortcuts	Alt + Shift + K	Option + Shift + K

VS. **Studio**[®]

- R is a programming language that runs in the program 
- RStudio adds a more structured user interface to R and uses the R language (*RStudio is an Integrated Development Environment (IDE) in which one can use the R language*)
- RStudio can not run independently of R!

Update

If you see a variation of the following warning message after you load a package:

```
library(dplyr)
```

Warning message: package 'dplyr' was built under
R version 3.5.1

- It is time to update R

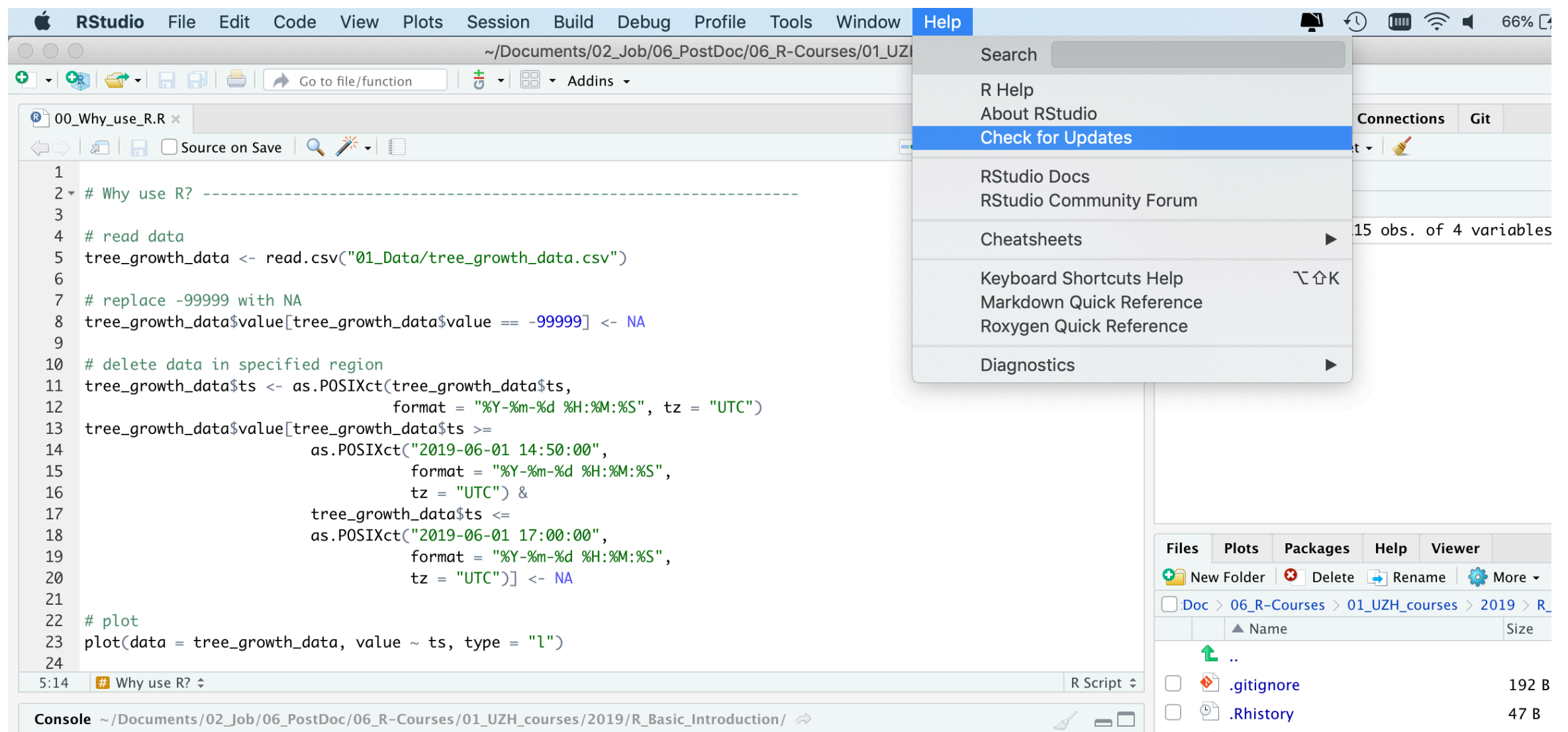
Update

- You need to download and re-install R from:
<https://www.r-project.org/>
- After the installation RStudio will automatically recognise the new R version.
- You have to re-install the packages after the installation of a new version of R.

(Attention: Few old packages might not run on the new R version. However, you can have multiple versions of R on your computer or downgrade to a previous version)

Update Studio[®]

Check from time to time whether there is a new version of RStudio available (Help > Check for Updates)



Capabilities of R

- Nice graphs
https://www.r-graph-gallery.com/violin_and_boxplot_ggplot2.html
- Maps with R
<https://www.r-graph-gallery.com/choropleth-map.html>
- Interactive web applications (R Shiny)
<https://shiny.rstudio.com/gallery/movie-explorer.html>
- R and databases
<https://db.rstudio.com/getting-started/connect-to-database>

Where to get help

- Websites
 - RStudio Cheat Sheets
<https://www.rstudio.com/resources/cheatsheets/>
 - Google
e.g. „r how to merge two data frames“
 - <https://stackoverflow.com/>
 - <https://www.r-bloggers.com/>
- Online book
 - R for Data Science
<https://r4ds.had.co.nz/>

...more help

- Other R courses
 - R: tidyverse for data science (UZH)
<https://app.connect.uzh.ch/apps/id/kurse.nsf/veranstaltungen.xsp>
 - Specialised R-Workshops (2 days, 1ECTS for PhD students, Plant Science Center)
<https://www.plantsciences.uzh.ch/en/teaching/phdplantscience/coursecatalogue.html>
 - Zurich R Courses
<https://www.zhrcourses.uzh.ch/en.html>
 - Specialised R-Workshops (1 day intensive course, small groups)
<https://ethz.ch/services/de/it-services/katalog/support-weiterbildung/it-training/kurse.html>

Sources

For the development of this course I was mainly inspired by the course material of Jan Wunders *R* course

Introduction to R (Wunder, 2016)