

Chapter 1: Getting started

Dr Megan L. Wood

2022-11-30

Chapter Overview:

1. What is R?
2. Pros of R
3. Cons of R
4. Introduction to the Workspace
5. Packages
6. Creating a project
7. Using R as a calculator

What is R?

R is a programming language and environment for statistical computing and graphics. The software is “*open-source*” which means that it is not sold commercially and anyone can download and use for free. It also means that anyone can contribute towards the development of the software. Unlike SPSS or STATA, R is based on command-line coding. RStudio is a wrapper to accompany R which provides a (albeit limited) Graphical User Interface (GUI). Even when using RStudio, commands are executed via executable code rather than clicking and selecting commands.

Pros of using R

- Open-source (no fees or licences required)
- It is constantly being developed by contributors globally
- Versatile language (>10k packages!)
- Platform independent (Mac, Windows, Linux)
- Reproducible - useful for version control and collaboration

Cons of using R

- Slower than other alternatives
- Can take up more memory compared to other software (e.g., python)
- Can be difficult to learn

The Workspace

Open R Studio and take a look at the different windows. You should see:

- **The Environment:** This is where your saved objects, functions, and dataframes are stored.
- **The Console:** This can be used as a “workpad” to test your code before including in your script. It is also where the output of any code is displayed
- **The Editor:** Here, you can run your code, save it entire scripts, store particular functions/code chunks as objects to use later in your analysis. No code will get run in here until you **Execute** the code
- **The Viewing Pane:** There are multiple views to toggle through here such as your files within your project, any plots you may run, and the Help function which displays the documentation for a particular package or function.

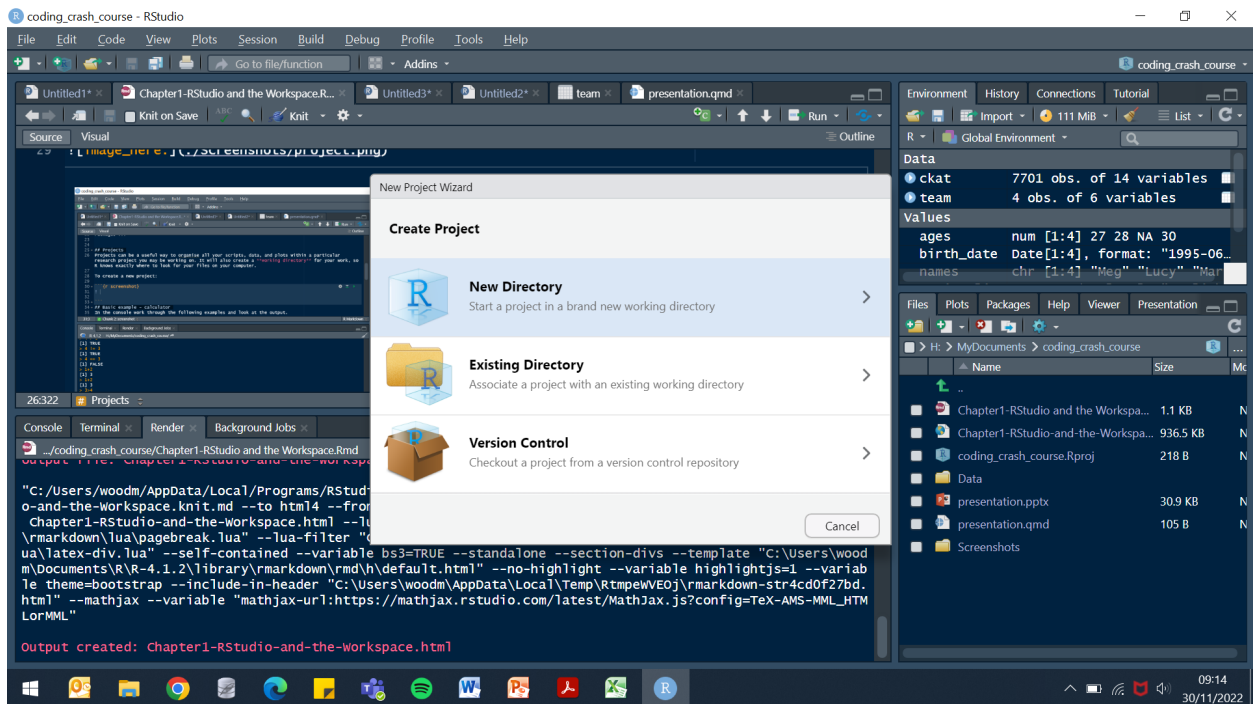
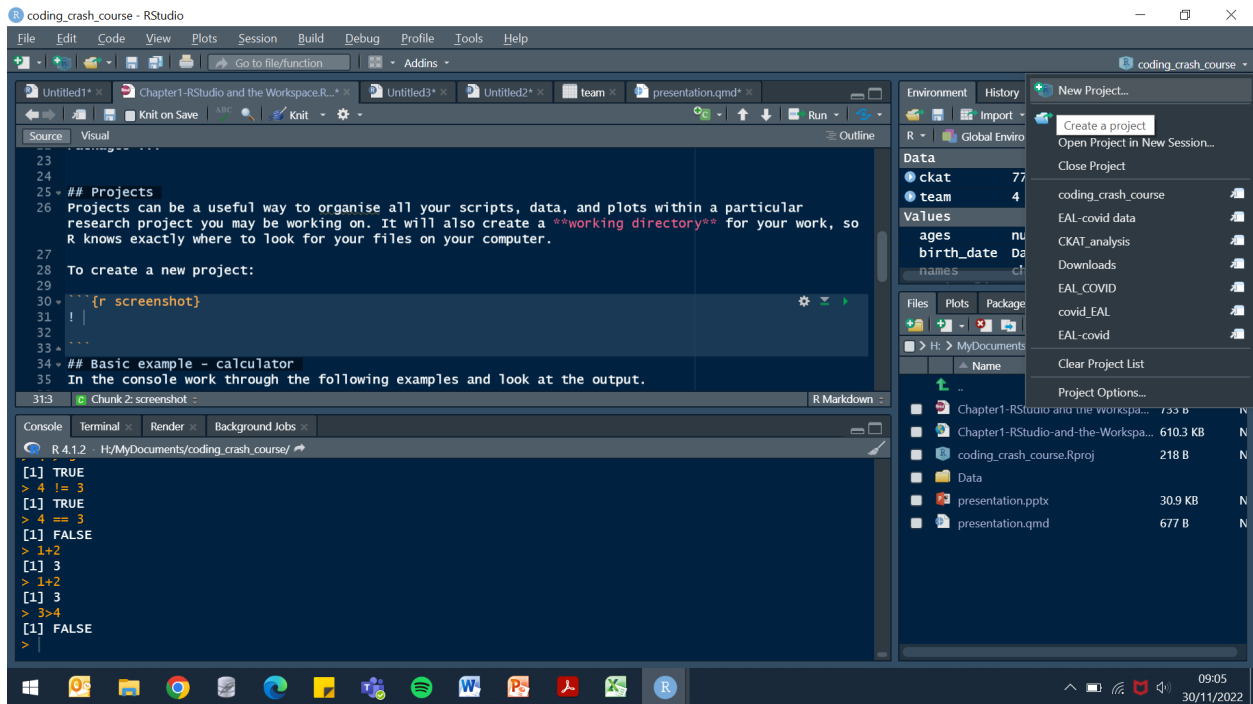
Packages

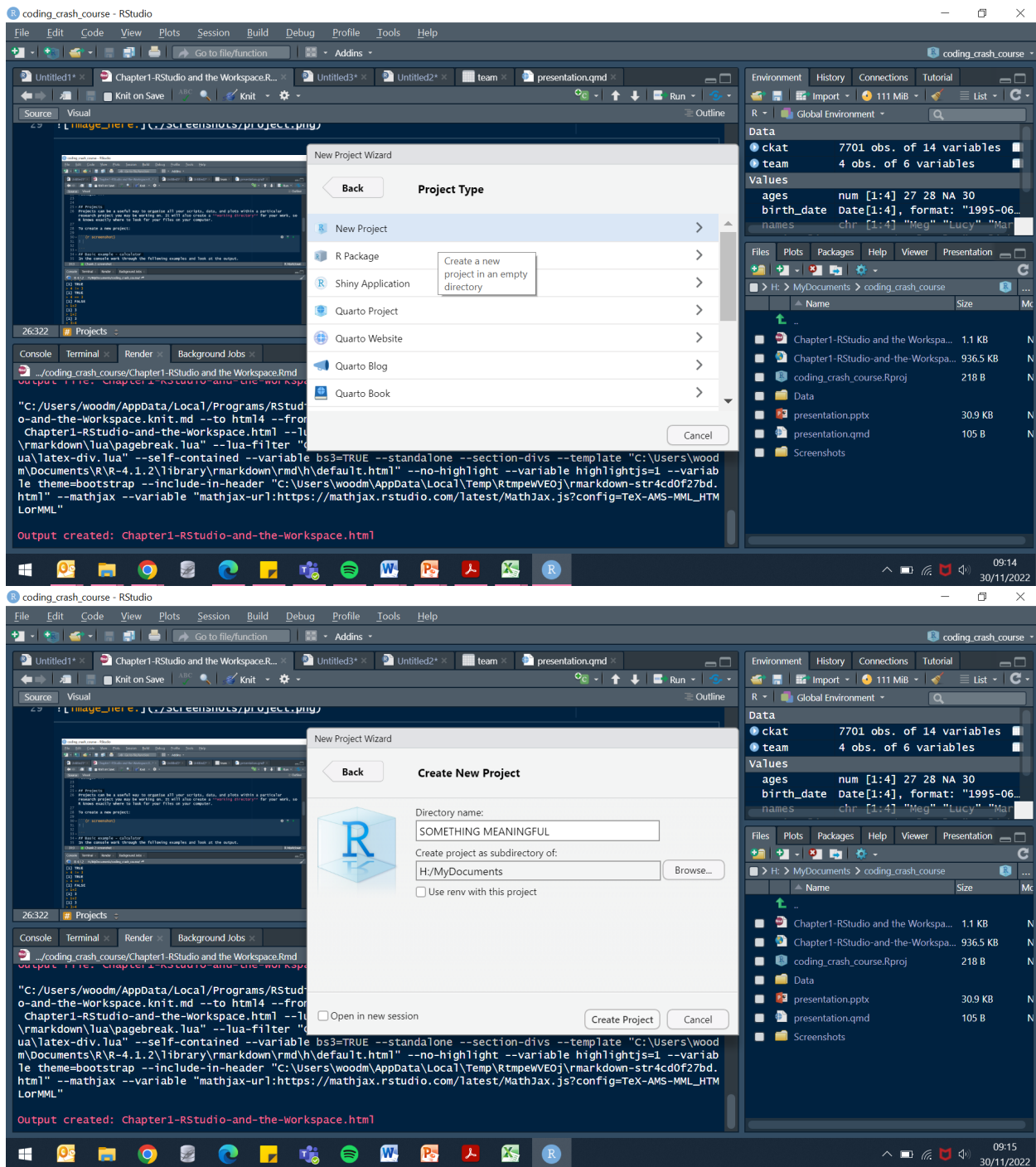
Packages are extensions to the R statistical programming language. R packages contain code/functions, data, and documentation. They are often created with specific purposes to allow us to solve more complex problems that cannot be achieved using R’s default functions. Anyone can build a package and upload to the Comprehensive R Archive Network (CRAN) for others to install and use. One of the most common and useful packages is called the **tidyverse** developed by Hadley Wickham (see <https://www.tidyverse.org/>). This is a special type of package, as it contains several packages within it. It is widely used in data science and will be the package we primarily use today. All the packages within tidyverse share a universal design in terms of the syntax and data structure which is intuitive to understand and therefore great for beginners.

Projects

Projects can be a useful way to organise all your scripts, data, and plots within a particular research project you may be working on. It will also create a **working directory** for your work, so R knows exactly where to look for your files on your computer. This will save you a lot of time and effort in the long run. It also links R directly into your Documents on your computer.

To create a new project:





Once you have created your project, open it up in your file explorer and create the following folders:

- data
- scripts
- plots

Basic example - calculator

In the console work through the following examples and look at the output.

Addition

```
1 + 2
```

Multiplication

```
3 * 2
```

Minus

```
3 - 2
```

Division

```
10 / 2
```

Booleans

```
4 > 3  
4 != 3  
4 == 3
```