

Introduction to R

Introduction

We recommend that you use RStudio and create script files in RMarkdown, as this makes it easy to read data files and gives you a good base to develop your R programming.

Installation of R

Before using R for the first time you need to install both programs. Find installers here:

For Mac:

- base R: <https://cran.r-project.org/bin/macosx/>
- RStudio: <https://rstudio.com/products/rstudio/download/#download> (Installers)

For Windows:

- base R: <https://cran.r-project.org/bin/windows/base/>
- RStudio <https://rstudio.com/products/rstudio/download/#download> (Installers)

The RStudio structure

After installing both R and RStudio, open RStudio. You will see three windows, to the left the console, to the upper right one with the environment (imported datasets and variables) and to the lower right one with files in your working directory. If you open a script file this will extend to four windows as below in Figure 1.

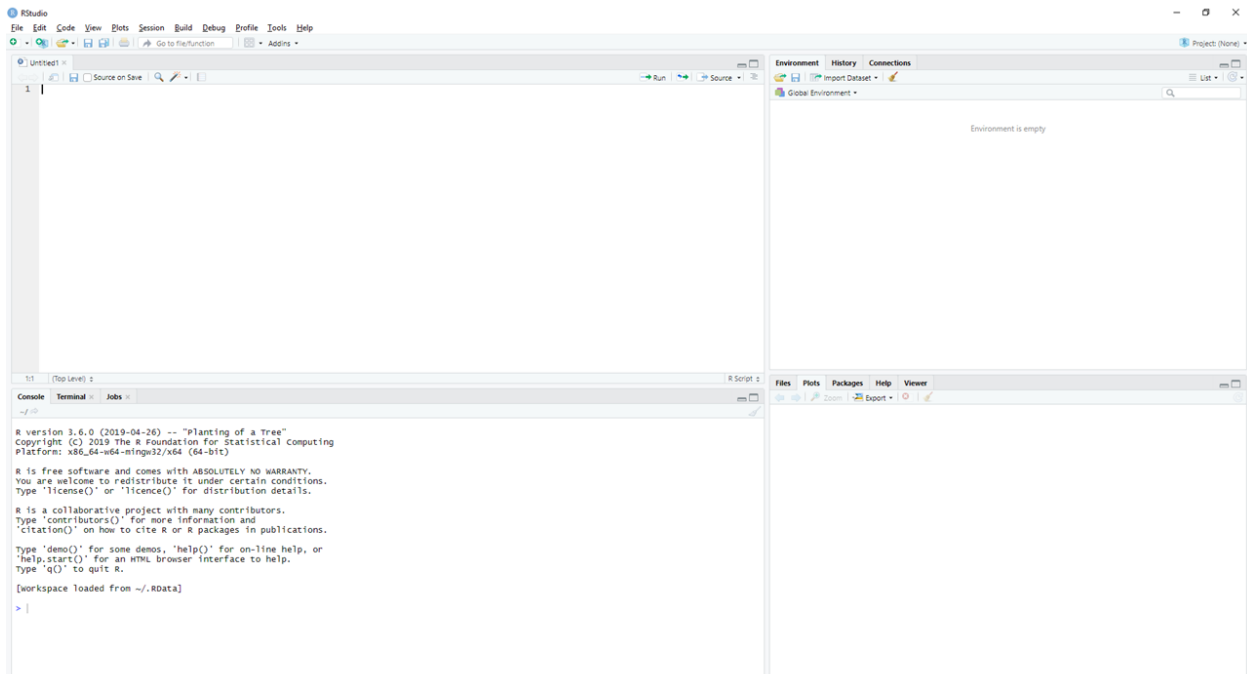


Figure 1: The opening window in RStudio.

The windows are:

Upper left: R code is written in this window (see later)

Upper right: Loaded data material, a history over run scripts, ...

Lower left: The console. Code can be run directly there, but it is not possible to save its contents when the session is closed. Therefore code is usually written in a script (e.g. Markdown) file in the upper left window.

Lower right: Shows plots and help files and allows the installation of installing and loading of packages.

Installation of packages

In R some functions are present in the base installation, while other, more specialised functions, are available in packages. Package installations are only done once before using the package for the first time. To install a package go to the *lower right* window in RStudio and choose the tab called *Packages* and click *Install*. In the line *Packages* type the package name and install the package. A list of installed packages are given in the *User_library*.

The tidyverse package

Most of our instructions use the tidyverse syntax and it is therefore necessary to install the tidyverse package. Observe that this is a large package and can take long time (up to an hour) to install.

Follow the instruction to install a package given above to install this package (write `tidyverse` in the line that specifies the package).

The readxl package

The `readxl` package simplifies reading and writing data files and is often used in the R instructions on our page. Install this package as well.

Getting started with RMarkdown files

The advantage with RMarkdown is that it is easy to load data files and run R code directly in the file. It is also possible to compile your Markdown files to pdf, doc or html, which makes it possible to write your reports or home assignments directly in R¹.

Open a Markdown file by choosing File → New File → R Markdown. The upper left window should look like in Figure 2:

Most of the text in the file is not needed. Delete everything below the first grey area.

Save Markdown files

A Markdown file is saved as any other file using File → Save As. Save all your files connected to a certain project in the same folder. Especially saving data files and Markdown files in the same folder is necessary to simplify importing data to R.

Known problems

If the chosen folder lies on a Server or in the Cloud there can arise problems to save the Markdown files. Typically, after trying to save the file you will get an error message. If this is the case try the following solutions:

- a. Map your documents path to a unit letter: Go to the file explorer on your computer and right-click on “This computer”. Choose “Map network drive”(“Anslut nätverksenhet”), where you can specify the drive, e.g. use “H:”. As Folder (Mapp) give “\\storage.slu.se\home\$\[username]”, where [username] is your SLU username.
- b. Work locally on your computer: Save all files you need locally on you computer (typically on C:).

¹All our R material is written in R Markdown

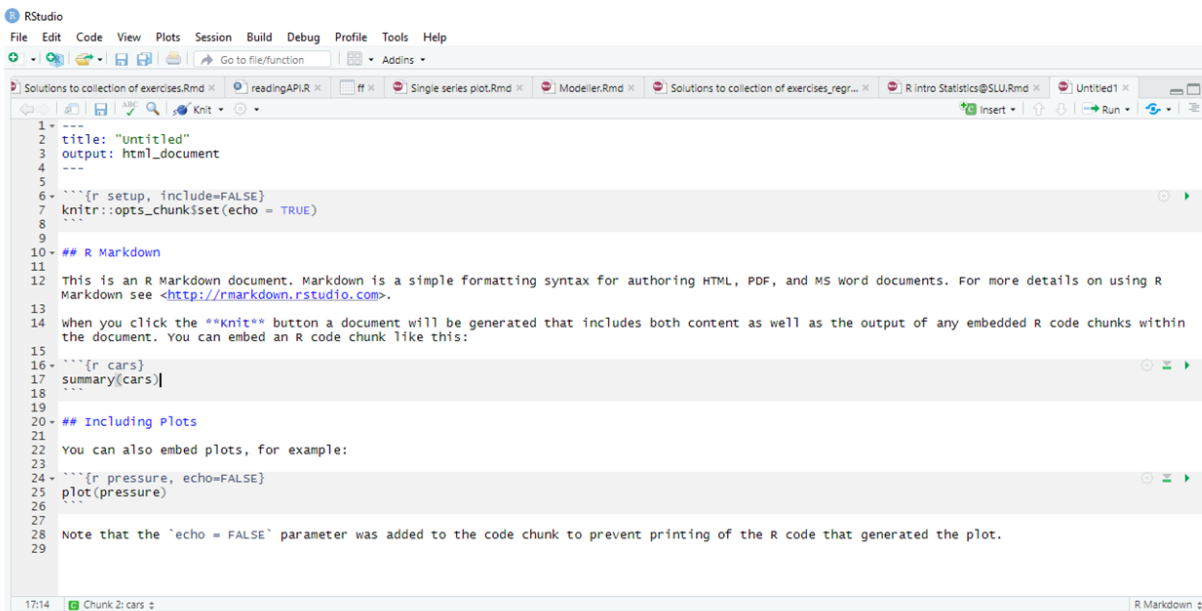


Figure 2: The opening window when you choose a RMarkdown file.

Code chunks

White areas are used to write text, such as comments or formulas. In grey areas the R code is written. They always have to start with ````{r}` and end with `````, as shown below. Only the middle part (`plot(lynx)` can be exchanged). This is called a code chunk and can be arbitrarily long.

```
```{r}
plot(lynx)
```
```

A new code chunk is inserted by `Ctrl+Alt+I` (Windows) or `Command+Option+I` (Mac).

Loading packages

When working with the R script we need to load the packages that we have installed earlier. This needs to be done for each session. A package is loaded with the function `library`. This needs to be placed in a code chunk, i.e. in a grey area.

```
library(tidyverse)
library(readxl)
```

Executing code

Put your cursor in the line you want to execute and press `Ctrl+Enter` (Windows) or `Command+Enter` (Mac). The selected line should transfer to the lower right window.

You can also click the green triangle in the code chunk to run the entire chunk.

Loading data into R

Download the `cats.txt` data set as test data. Save the data file in the same folder as the RMarkdown file. This is important as R otherwise cannot find the file without specifying the entire path. Read the `cats` data with the code below. You should get some text that indicates which variables are read from the data set and which type they are (e.g. `col_character` indicates a text variable and `col_double` a numeric variable).

```
cats<-read_tsv("cats.txt")

## Parsed with column specification:
## cols(
##   Name = col_character(),
##   Breed = col_character(),
##   Sex = col_character(),
##   Neutered = col_character(),
##   Age = col_double(),
##   Weight = col_double()
## )
```

Other files can be read with the appropriate functions, e.g. `read_csv` for comma-separated files or `read_excel` for standard excel-files. Also the more general `read_delim` function can be used to read any delimited file.

Typical error messages

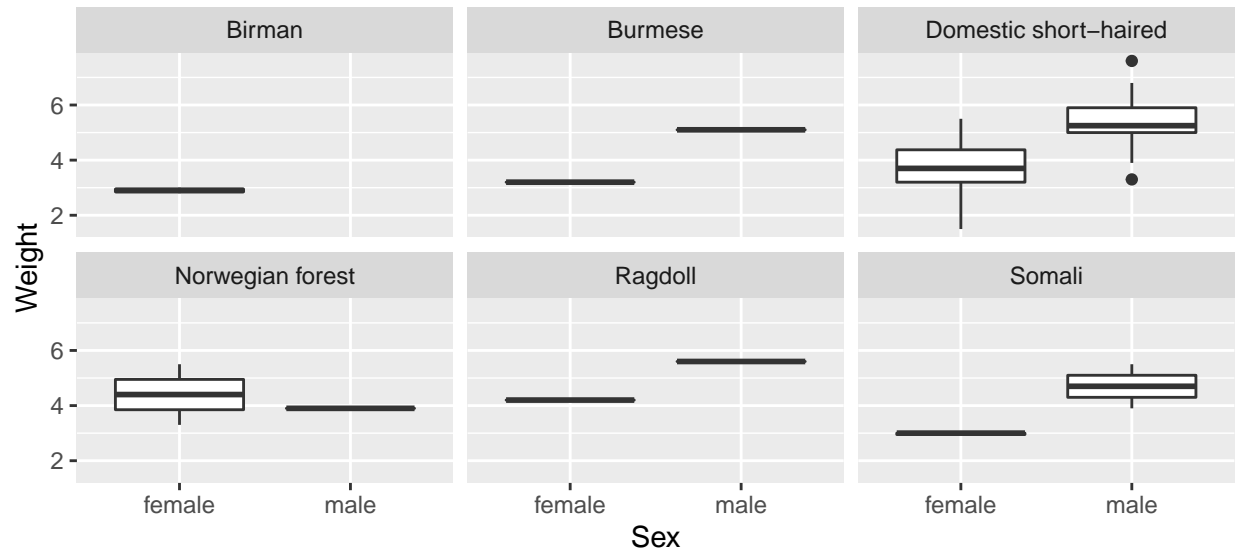
'cats.txt' does not exist in current working directory: You have not saved the `cats.txt` dataset in the correct location, i.e. in the same folder where the Markdown file is saved.

object 'cats2.txt' not found: You have written `cats.txt` without citations around. It should be "`cats.txt`".

Piping code

In RMarkdown you can write code with the general base R syntax or with R tidyverse syntax. We will mainly use the second. Typically for tidyverse is that code is "piped". For example, it starts by specifying the data frame which is sent further (`\texttt{>}`) to a plot function. R code is case-sensitive, i.e. it is important to distinguish between capital letters and lower case letters.

```
cats%>%
  ggplot(aes(x=Sex, y=Weight))+
  geom_boxplot()+
  facet_wrap(~Breed)
```



For explanation of the code and further examples see the material on Statistics@SLU webpage or computer exercises in our courses.

Knitting RMarkdown files

By default, it is possible to create html files from RMarkdown files. Save the code you have produced so far and click the button called “Knit” in one of the upper rows of the code window (the upper right window). This should produce a html file with your code and output.

Shortcuts in R Markdown

A full list of shortcuts for R Markdown in Windows and MAC: <https://support.rstudio.com/hc/en-us/articles/200711853-Keybaord-Shortcuts>. Here are some of the most important:

Run current line/selection Ctrl+Enter Command+Enter

Run current document Ctrl+Alt+R Command+Option+R

Insert chunk (Sweave and Knitr): Ctrl+Alt+I Command+Option+I

Knit Document (knitr) Ctrl+Shift+K Command+Shift+K

Finding errors

RStudio will give you hints if some of your code is wrong (Figure 3).

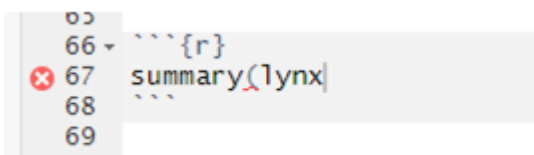


Figure 3: A bracket is missing in the code.