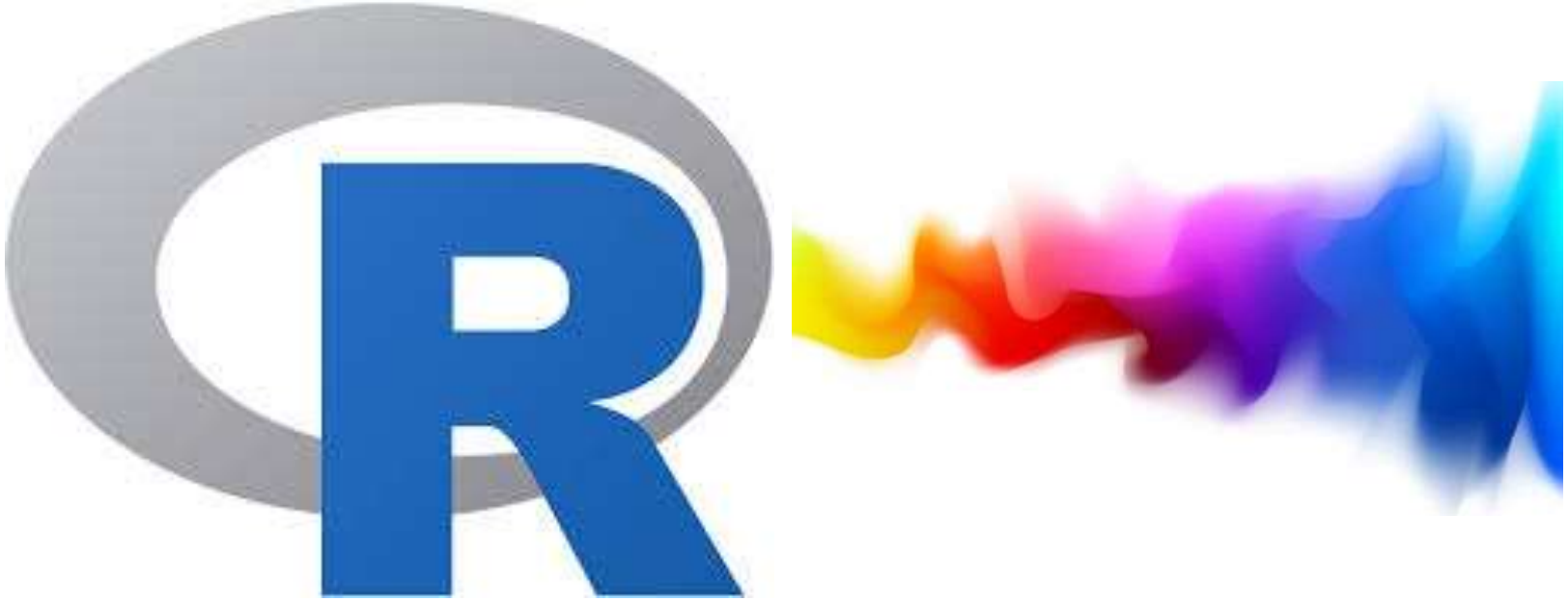


Basics of R Programming

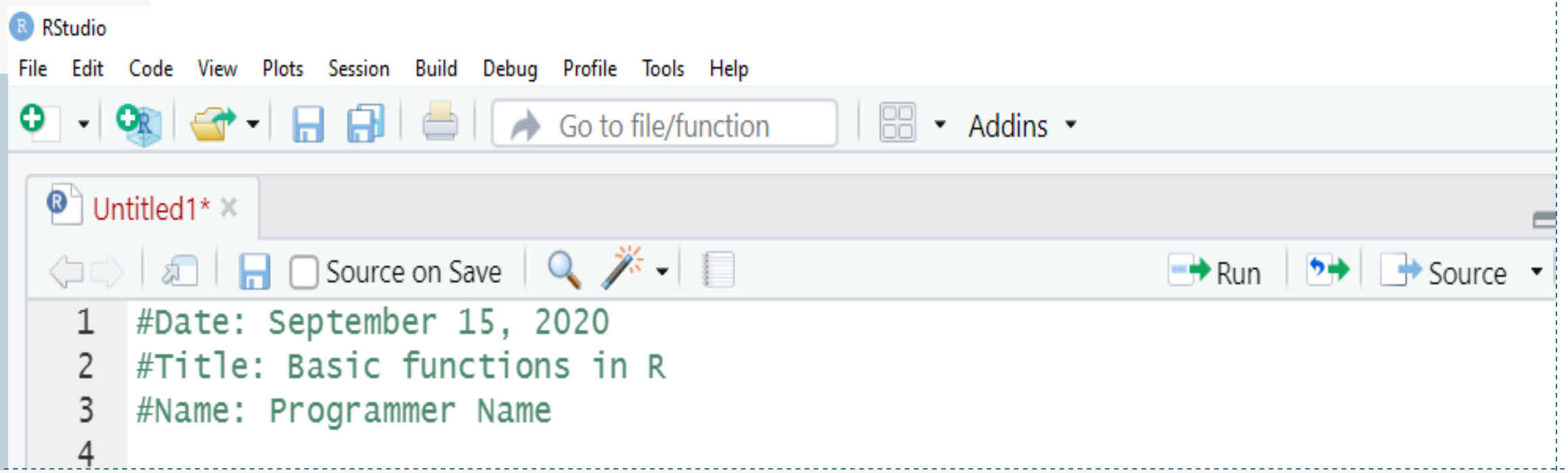
VN-Biostat Pre-Workshop

Basics of R Programming

- Basic arithmetic operations in R
- Saving an R script
- How to import/upload databases
- R packages, how to use them and where to find them

Commenting in R

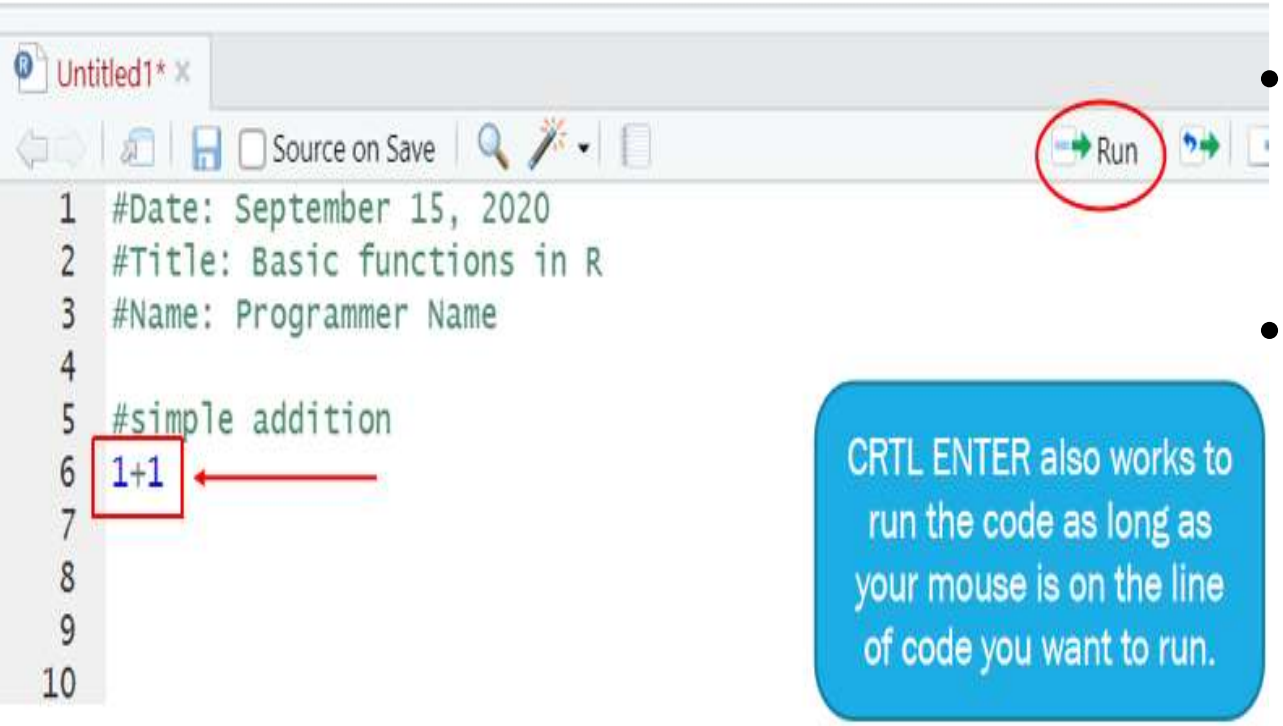
- You can add comments by including the hashtag symbol before your comments. This will turn the font of the comments green. Example shown below:



Commenting in R and case sensitivity

- **It is worth noting that R is very much case sensitive.**
- **The variable “Age” is different to the variable “age”.**
- **R also ignores anything on a line that follows a # symbol. This is very handy, as it is possible (and highly recommended from a reproducible research perspective!) to add comments to your scripts using the # symbol.**
- **In complex scripts it's a great idea to make many comments explaining what you are doing.**
- **Then you don't have to try to remember weeks (or years!) later, and it’s much easier for others to understand what you did! It’s also important to note that R ignores spaces, so its okay to use lots of spaces to make your scripts look nice.**

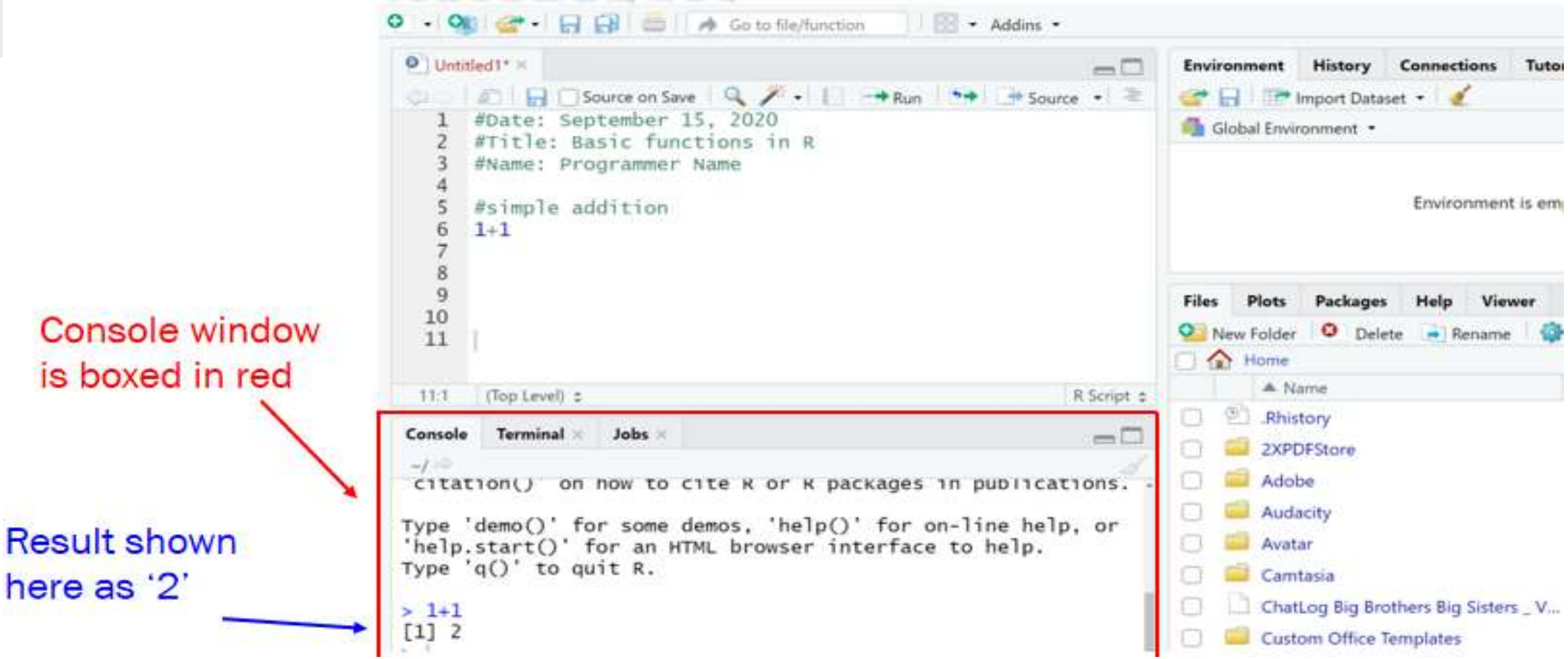
Basic Arithmetic Operations in R: Addition



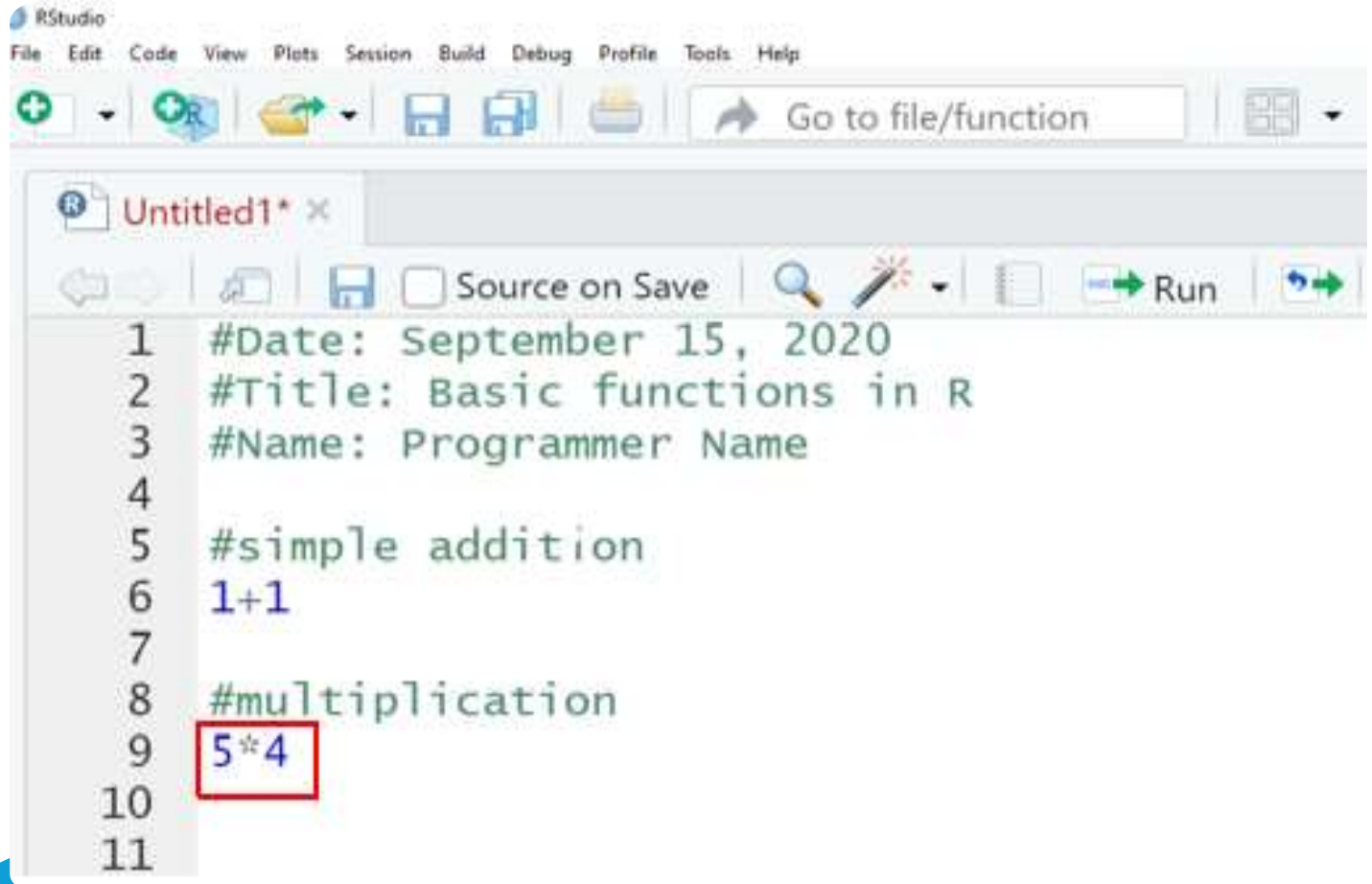
- Simply add 1+1 in the RScript. Type in the following code, highlight the code and select ‘Run’
- If you don’t select a certain section of your code before clicking ‘run’, all codes present in the RScript file will be executed. You can either highlight the section of code or place your mouse on the same line as the code (i.e., place your mouse on line 6) then click run.

Addition Results

Your results will be displayed in the Console Window



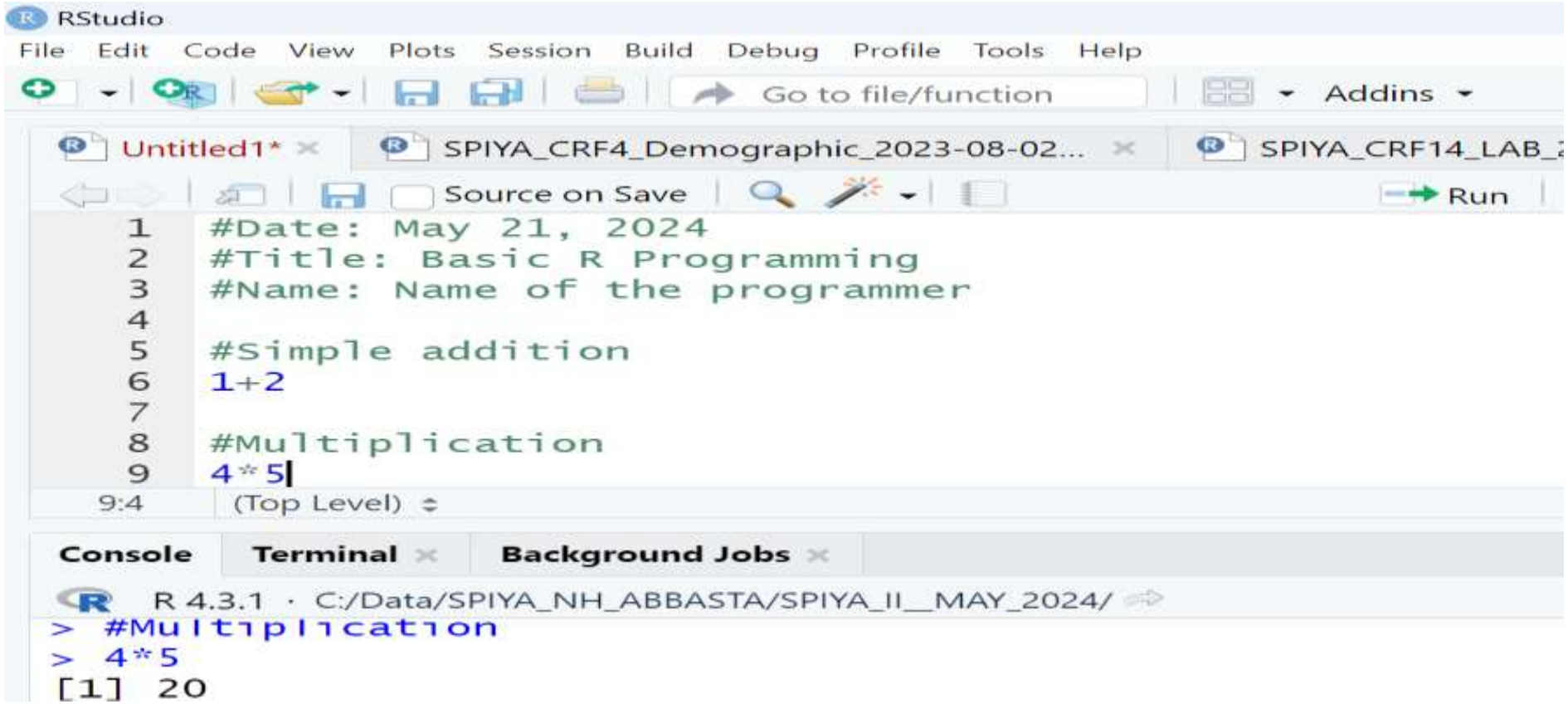
Multiplication



- Simply multiply two numbers by adding the asterisk (*) sign in between them, highlight the code, and click ‘run’

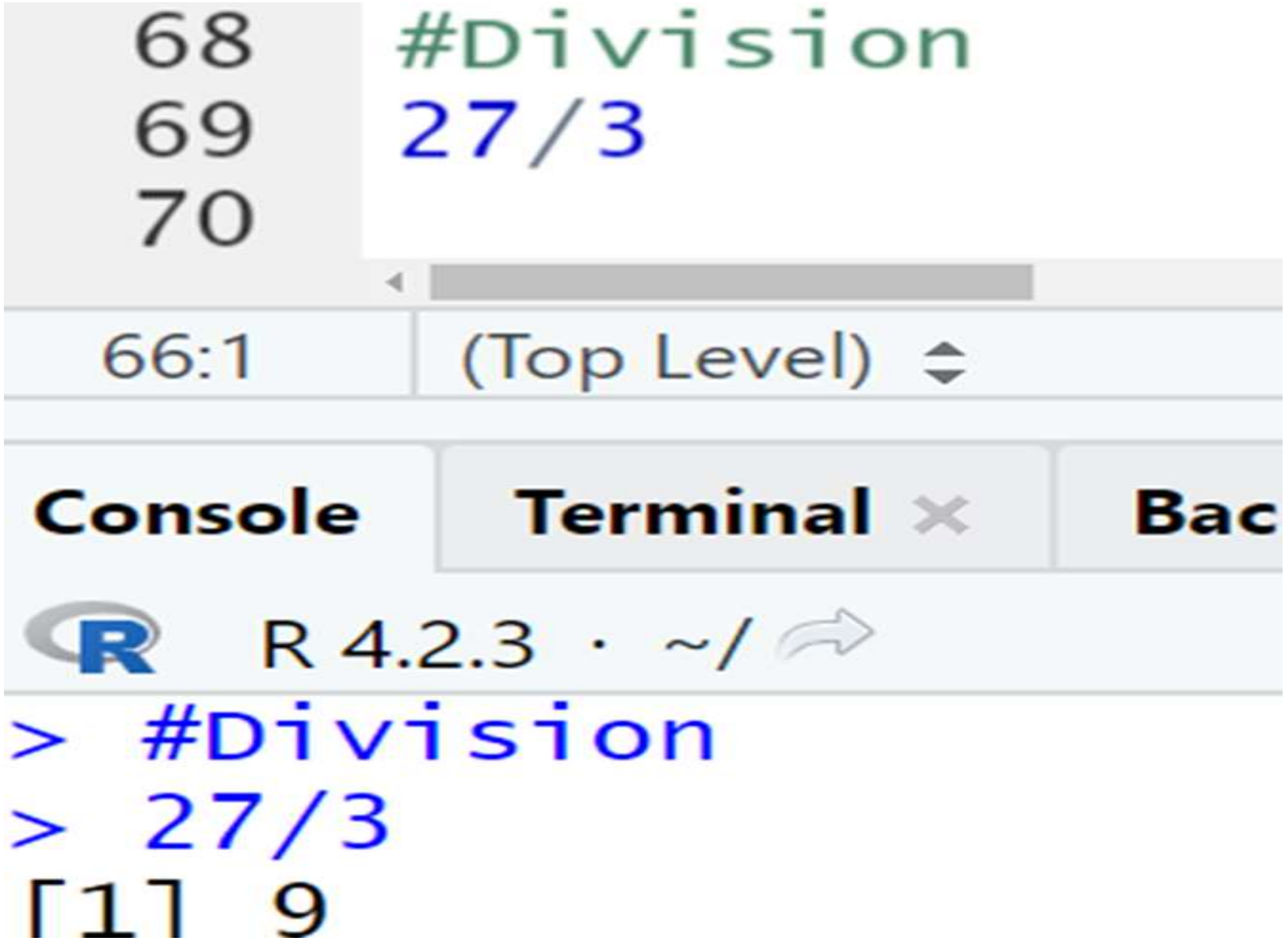
Multiplication results

Your results will be displayed in the Console Window



Division

You can divide two numbers by including a forward slash (/) in between them



Logical and Arithmetic Operators

Operator	Description
<	less than
<=	less than or equal to
>	greater than
>=	greater than or equal to
==	exactly equal to
!=	not equal to
!x	Not x
[redacted]	[redacted]
x & y	x AND y

Operator	Description
+	addition
-	subtraction
*	multiplication
/	division
^ or **	exponentiation
x %% y	modulus (x mod y) 5%%2 is 1
x %/% y	integer division 5%/%2 is 2

Saving an R script

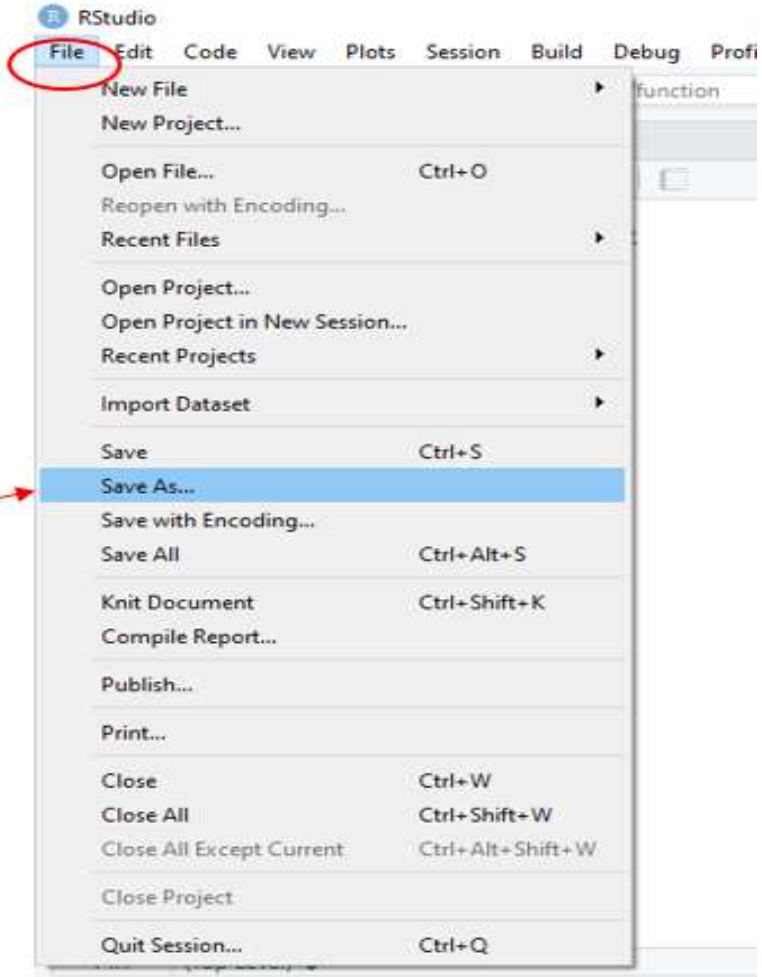
Saving Your Work

- ❑ Remember to save your R scripts regularly. To save your script:
- Click on **File** in the menu bar.
- Select **Save** or **Save As** to choose a location and filename for your script.

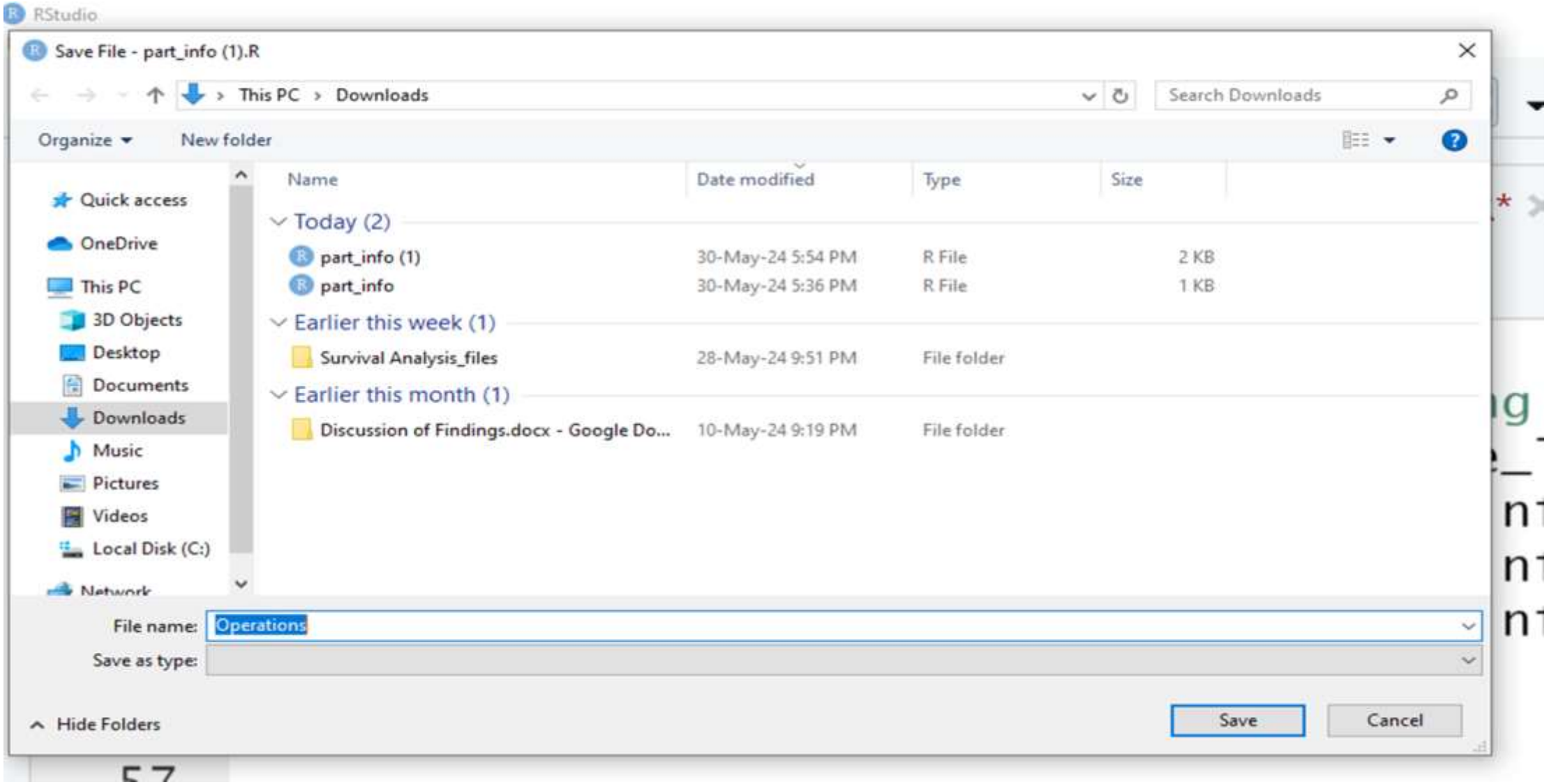
Please save your RScript

- It is good practice to save your R Script so that you can continue where you left off just in case you need to leave the session or if you want to come back to the code at a later point.

- Click ‘file’ → ‘save as’

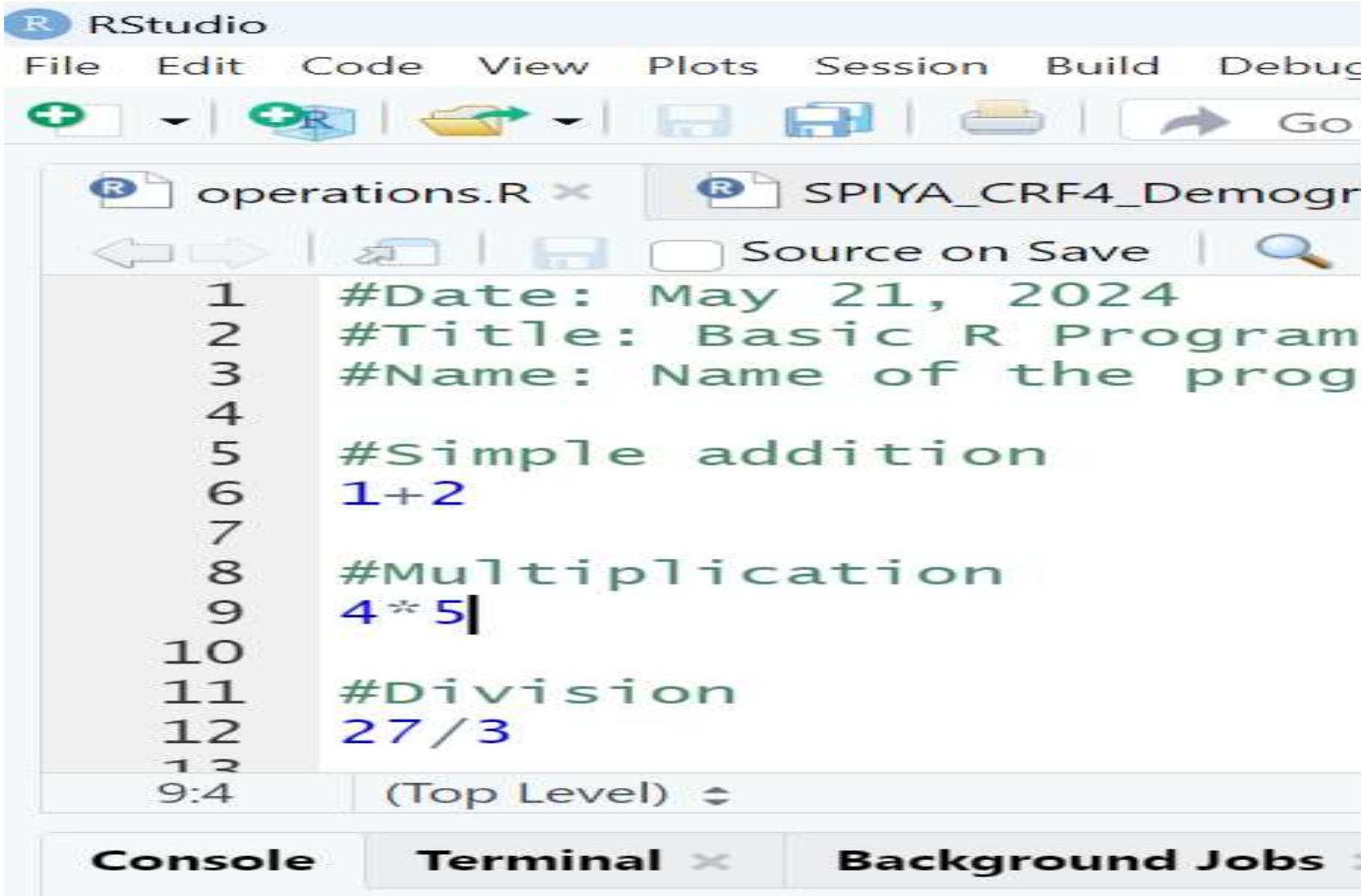


Name your script

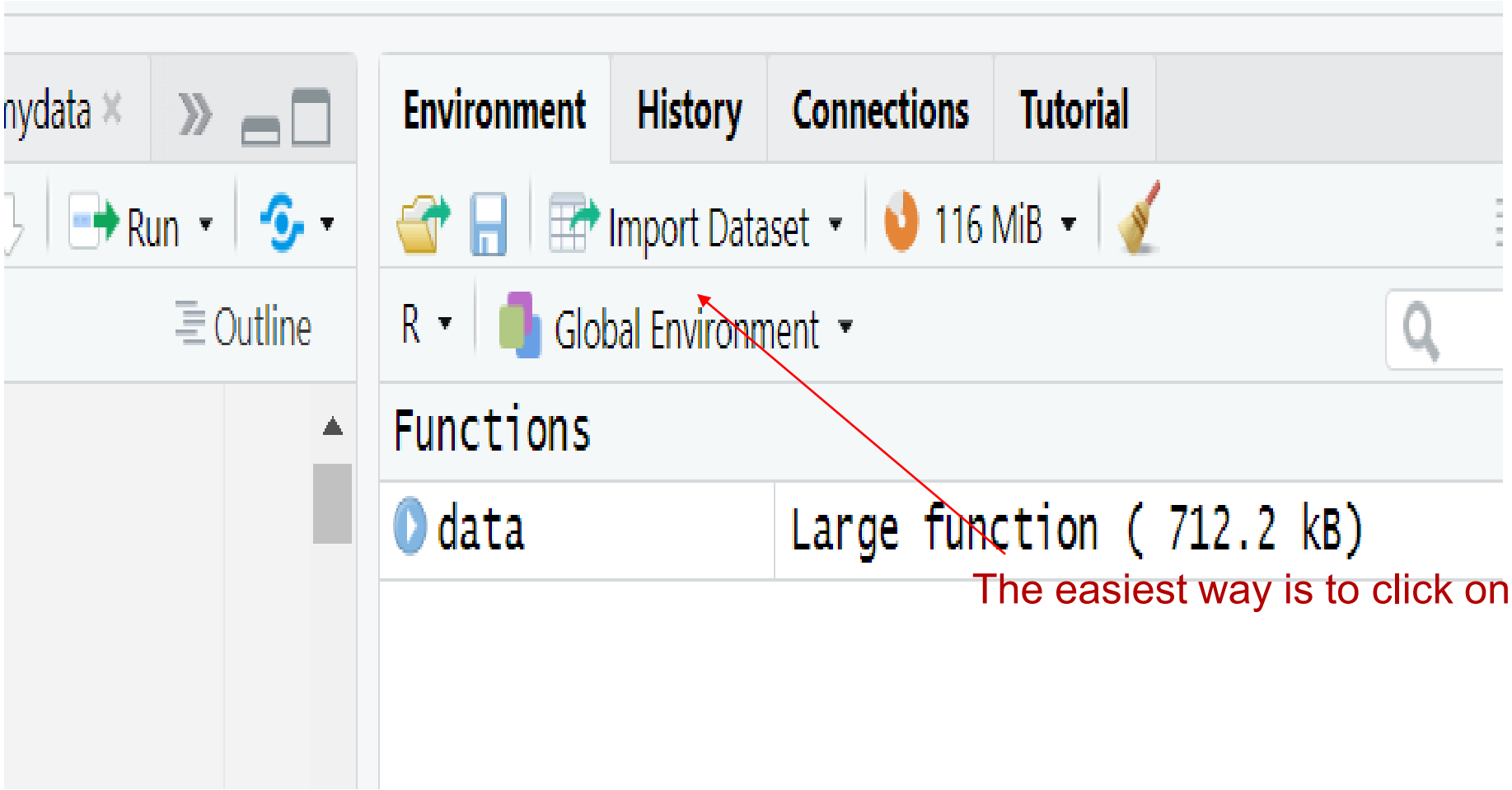


Please Save your R Script

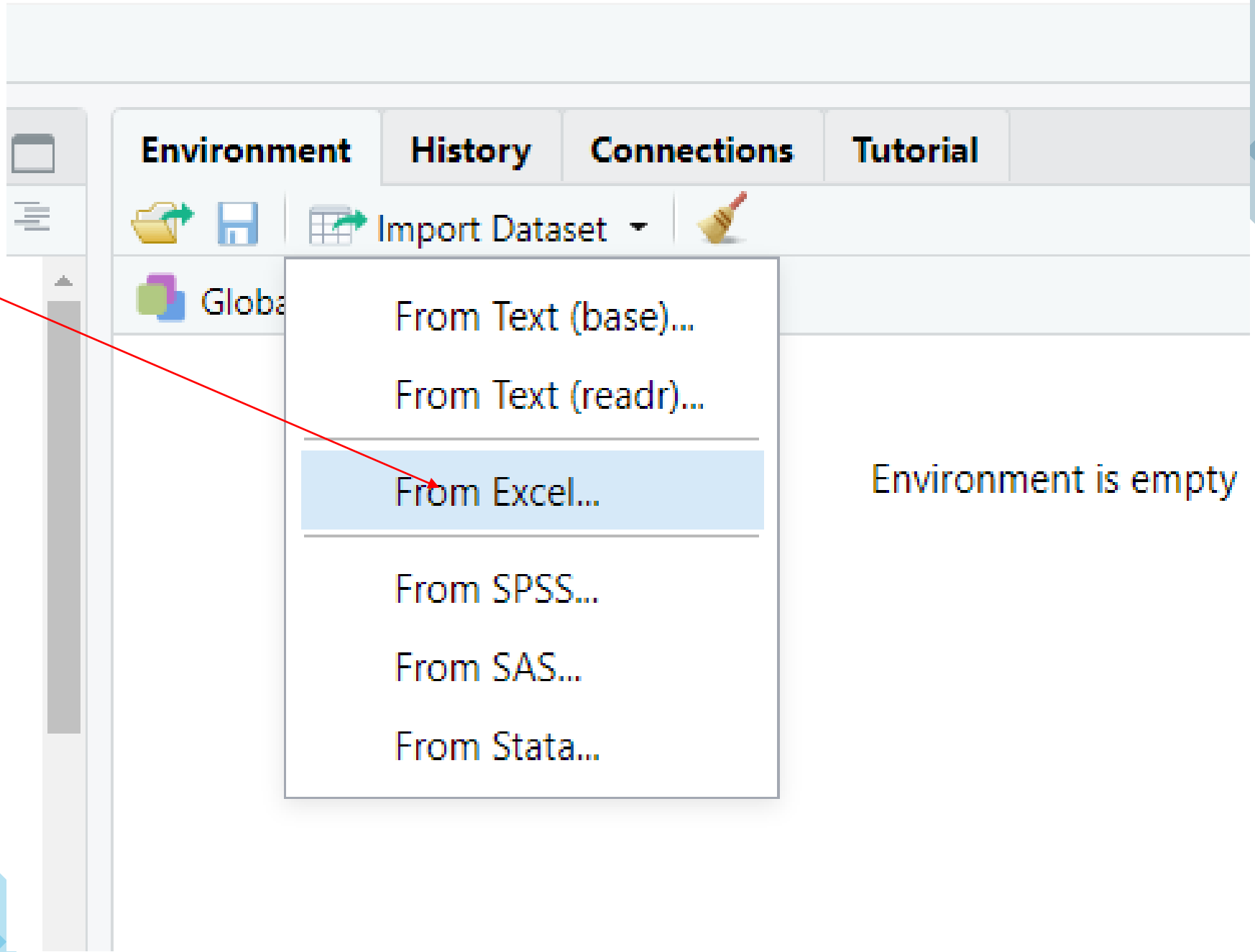
- Notice how the heading is now called 'operations.R'
- RScripts have a '.R' extension
  - Similar to how word documents end in '.doc' or '.docx'



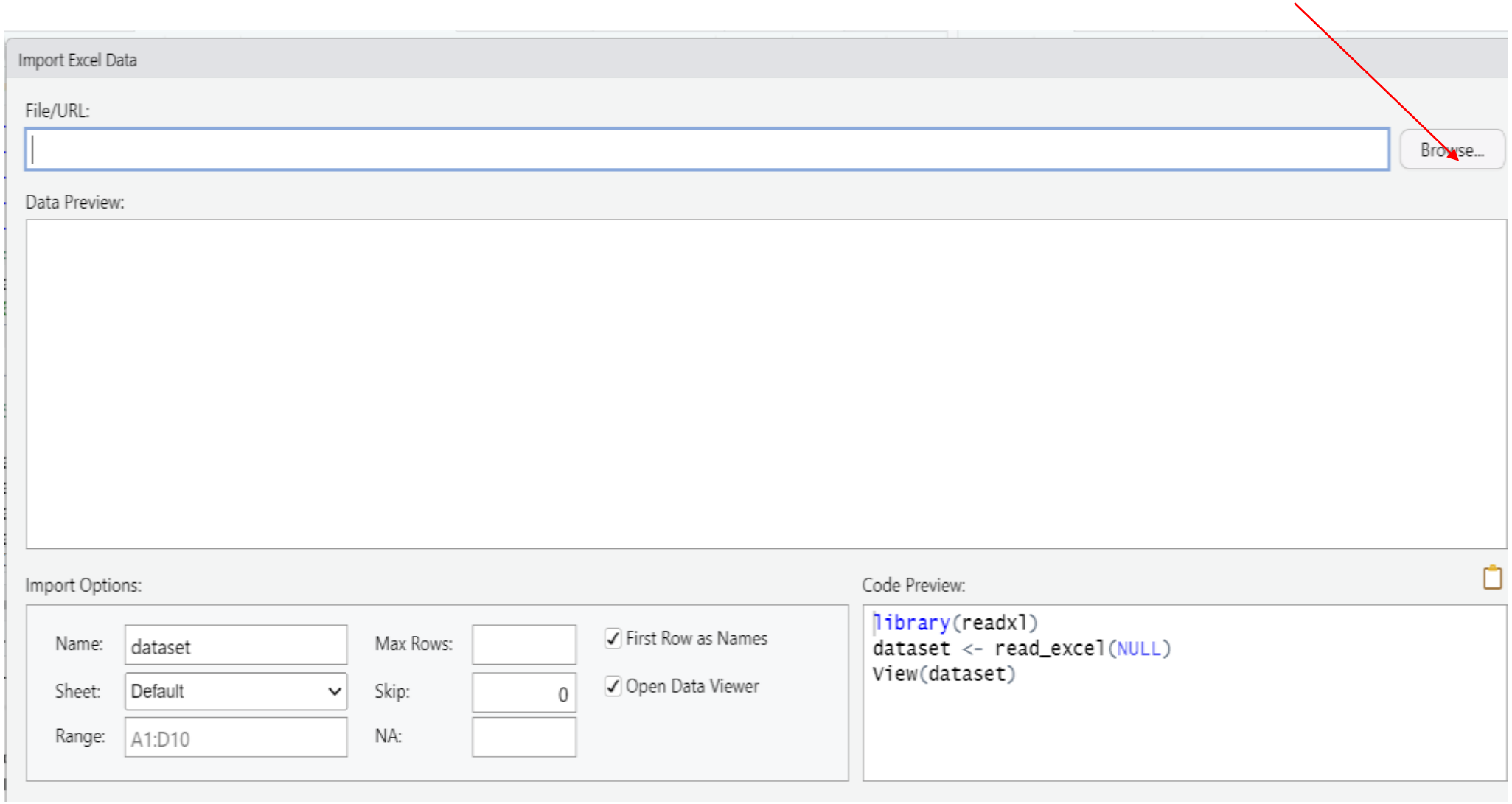
How to upload an excel file in R

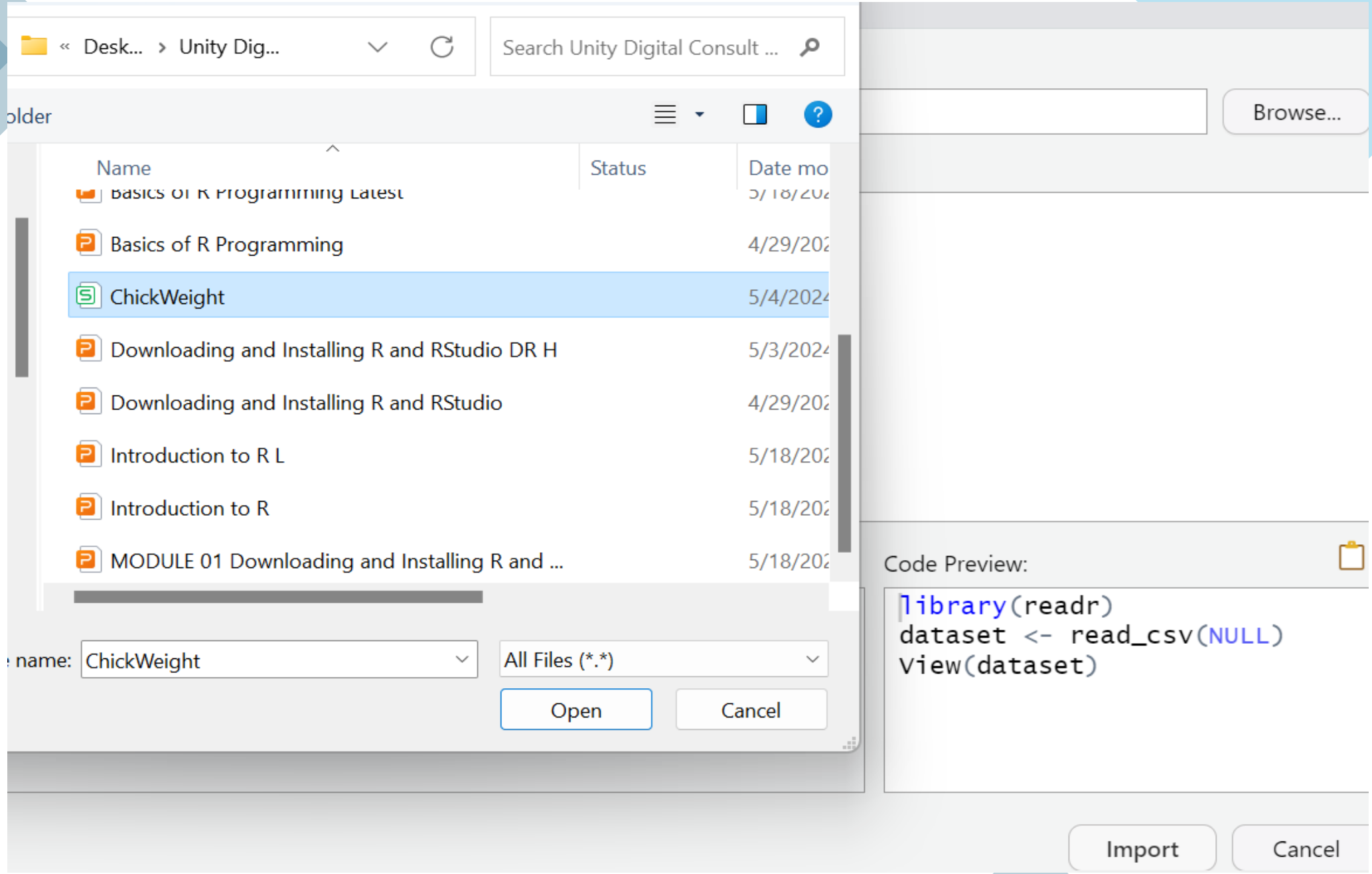


Select “from Excel”

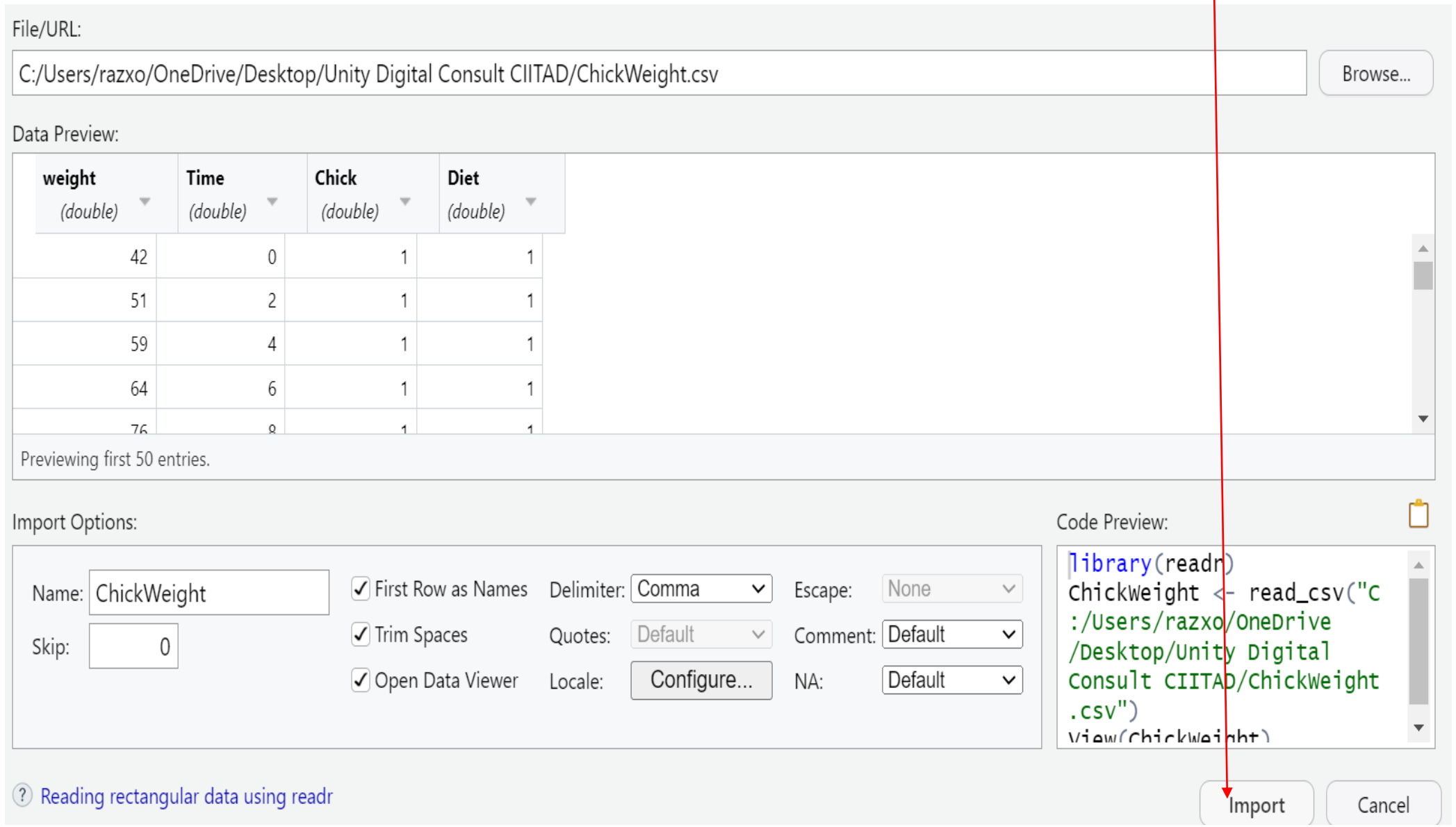


- Select “browse” and search for your excel file. Make sure it’s saved on your computer!



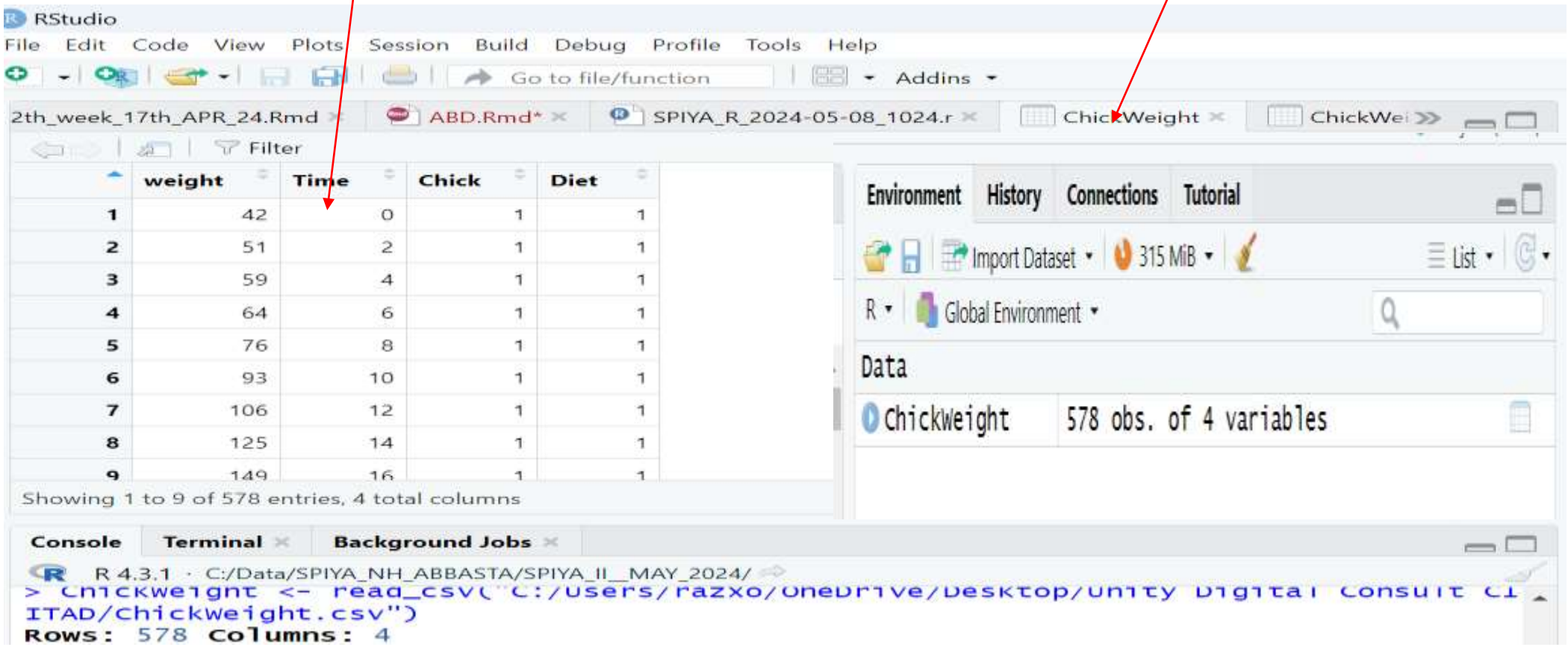

« Desk... › Unity Dig...
Search Unity Digital Consult ...
Name
Status
Date mo
Basics of R Programming
4/29/202
ChickWeight
5/4/2024
Downloading and Installing R and RStudio DR H
5/3/2024
Downloading and Installing R and RStudio
4/29/202
Introduction to R L
5/18/202
Introduction to R
5/18/202
MODULE 01 Downloading and Installing R and ...
5/18/202
name:
ChickWeight
All Files (*.*)
Open
Cancel
Browse...
Code Preview:
library(readr)
dataset <- read_csv(NULL)
View(dataset)
Import
Cancel

Once you select your excel file, you will see the preview of the data. Next, select "import". Press import to import your dataset



You see a new tab that has opened your dataset in R.

The dataset is in the R environment as shown on your screen. Name of dataset is “Chickweight”



R packages

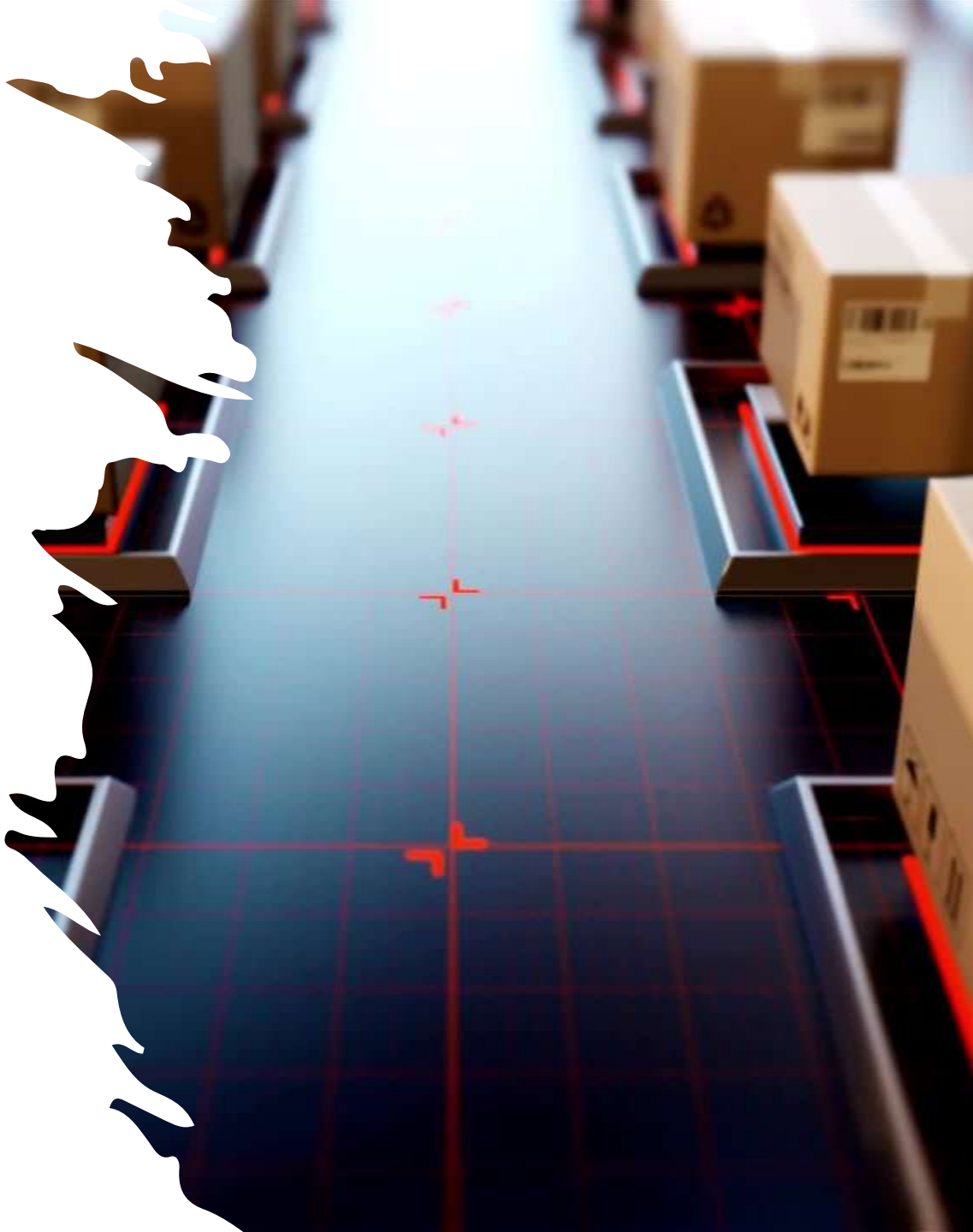
- R packages are collections of R functions, datasets, and documentation that extend the functionality of R. The Comprehensive R Archive Network (CRAN) and GitHub are primary repositories for R packages, where you can find thousands of packages for various purposes.

R packages

- Extending R's capabilities as you dive into projects, many thousands of R users have developed useful code and shared this code as installable packages.
- These packages can be downloaded from a variety of sources but the most popular are CRAN, Bioconductor and GitHub.
- Currently, CRAN hosts over 15000 packages and is the official repository for user contributed R packages.
- Bioconductor provides open source software oriented towards bioinformatics and hosts over 1800 R packages.
- GitHub is a website that hosts git repositories for all sorts of software and projects (not just R).

Installing Packages

- In order to install a package from an online repository like CRAN we have to first download the package files,
- All of this can be done at the Console using a single function:
  - Install.packages
- For example, if we want to install a package called **tidyverse**, we use:
  - install.packages(“tidyverse”)
  - Installing a package is a “do once” operation.



Installing Packages

- You can install more than one package at a time. Example,

```
pckg.names <- c(“tidyverse", "ggplot2")
install.packages(pckg.names)
```

Loading and attaching packages

```
14  package.names <- c("tidyverse", "ggplot")
15  install.packages(package.names)
16
```

16:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.3.1 · ~/

```
see the ideas at
https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.3/tidyverse_2.0.0.zip'
Content type 'application/zip' length 430817 bytes (420 KB)
downloaded 420 KB

package 'tidyverse' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\razxo\AppData\Local\Temp\RtmpWSkVEv\downloaded_packages
>
```

Loading and attaching packages

- Once we've installed a package or two we'll probably want to actually use them.
- Fortunately, it can be done in a single step with a function called "library"
- The library function works exactly as we might expect it to. If we want to start using the fortunes package—which was just installed above—all we need is:
  - library("fortunes")
- loading and attaching a package via library is a "do every time" operation.

Loading and attaching packages

```
14 package.names <- c("tidyverse", "ggplot")
15 install.packages(package.names)
16
17 library(tidyverse)
18
```

18:1 (Top Level) R Script

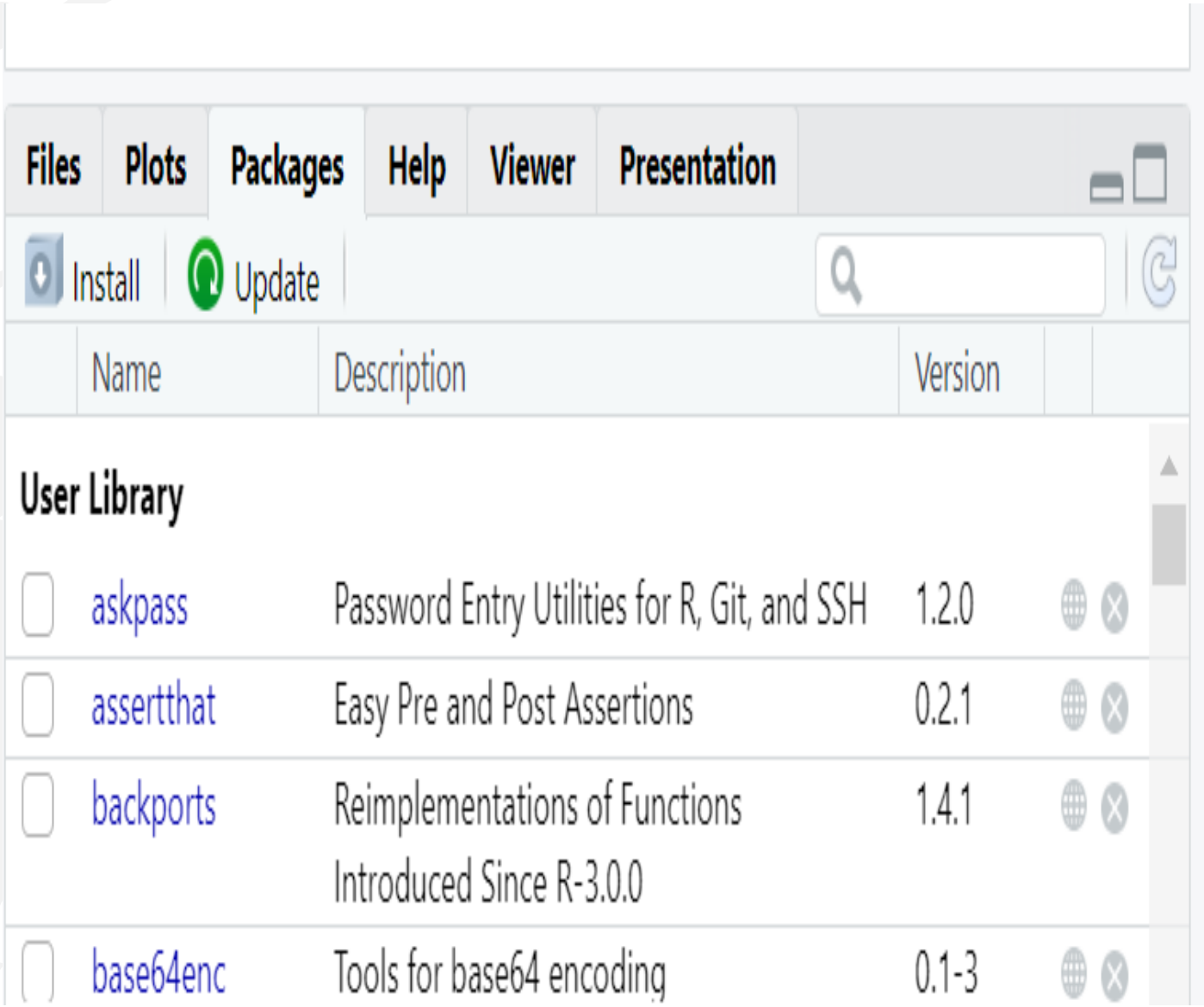
Console Terminal Background Jobs

R 4.3.1 · ~/

```
x dplyr::lag()    masks stats::lag()
i Use the conflicted package to force all conflicts to become errors
Warning messages:
1: package 'tidyverse' was built under R version 4.3.3
2: package 'ggplot2' was built under R version 4.3.3
3: package 'tidyr' was built under R version 4.3.2
4: package 'readr' was built under R version 4.3.2
5: package 'purrr' was built under R version 4.3.2
6: package 'dplyr' was built under R version 4.3.2
7: package 'stringr' was built under R version 4.3.2
```

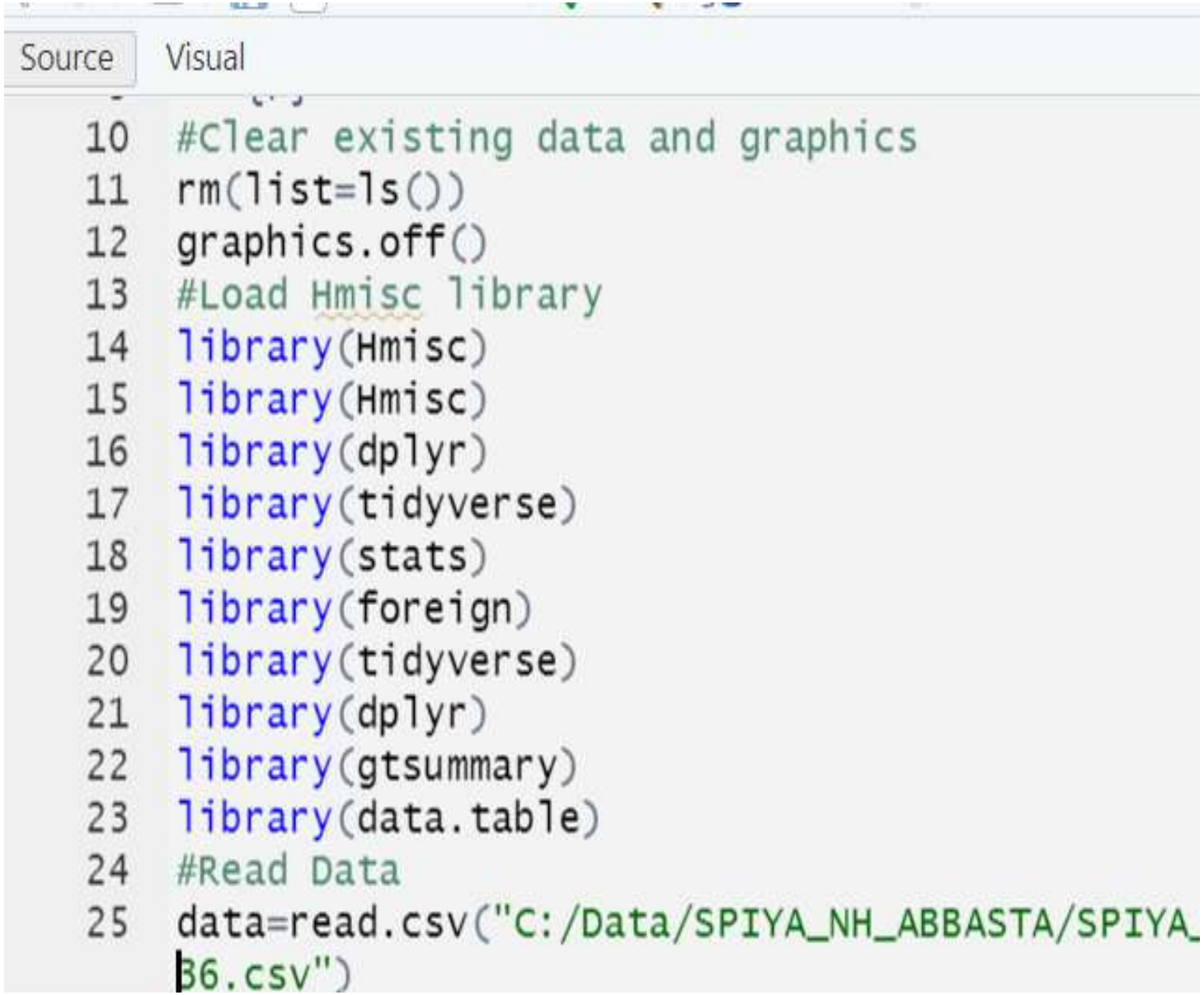
Updating packages from last window of Rstudio

- Once again, if we really don't like working in the Console RStudio can help us out.
- There is a small button next to each package listed in the Packages tab
- Packages that have been loaded and attached have a blue check box next to them, whereas this is absent from those that have not.
- Clicking on an empty check box will load up the package



One last tip

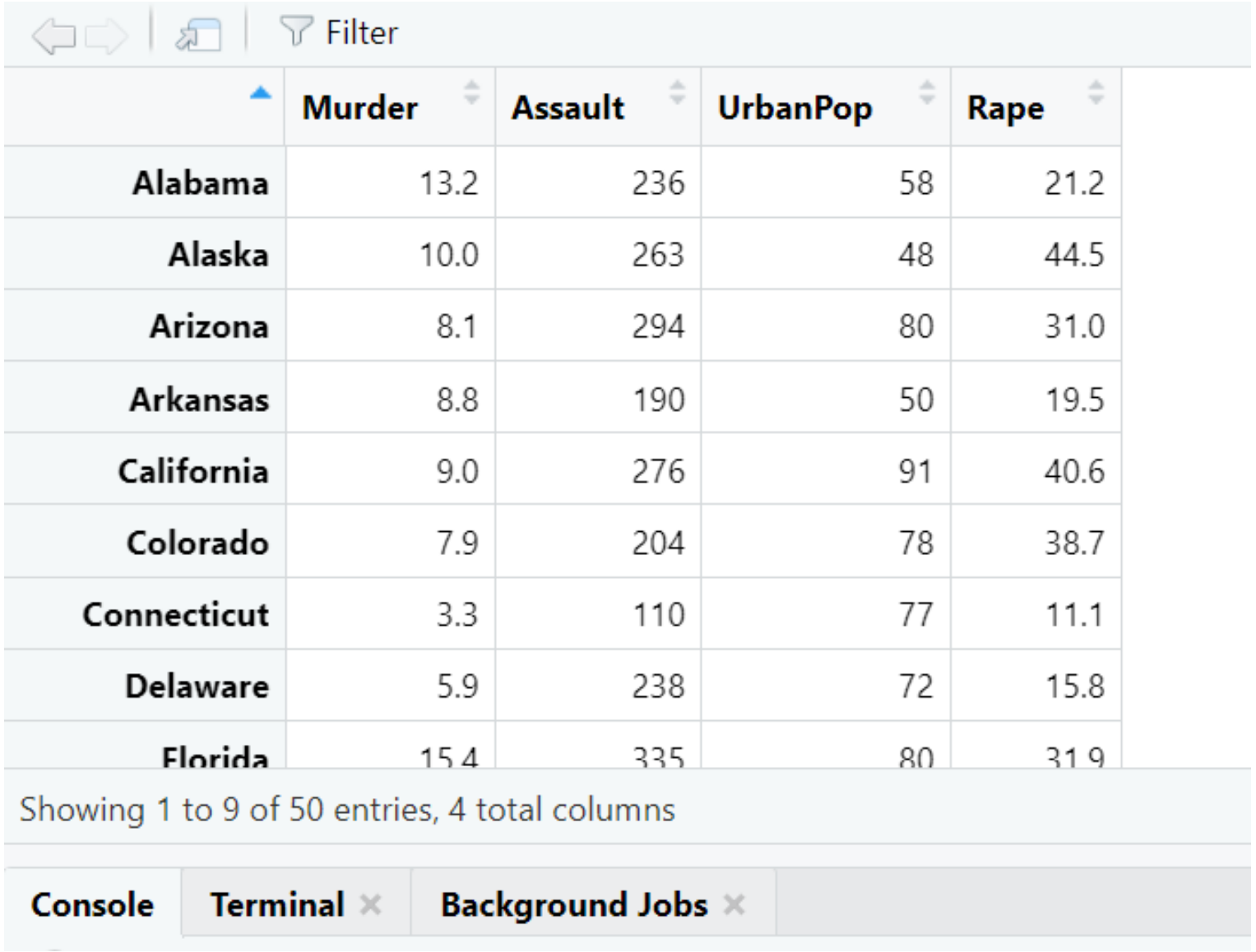
- We can use library anywhere, but typically the library expressions live at the very beginning of a script so that everything is ready to use later on.



Exercise

1. Import USArrests Data using from “Text(readr)”
2. Save your R Script and save it as myname.R.

USArrests Data View



Filter

	Murder	Assault	UrbanPop	Rape
Alabama	13.2	236	58	21.2
Alaska	10.0	263	48	44.5
Arizona	8.1	294	80	31.0
Arkansas	8.8	190	50	19.5
California	9.0	276	91	40.6
Colorado	7.9	204	78	38.7
Connecticut	3.3	110	77	11.1
Delaware	5.9	238	72	15.8
Florida	15.4	335	80	31.9

Showing 1 to 9 of 50 entries, 4 total columns

Console | Terminal | Background Jobs

References

- https://dzchilds.github.io/eda-for-bio/packages.html
- https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf
- https://cran.r-project.org/doc/manuals/R-intro.pdf

Thank you