Introduction to R, RStudio, and RStudio Server

The Data Lab

Powered by Alex's Lemonade Stand Foundation

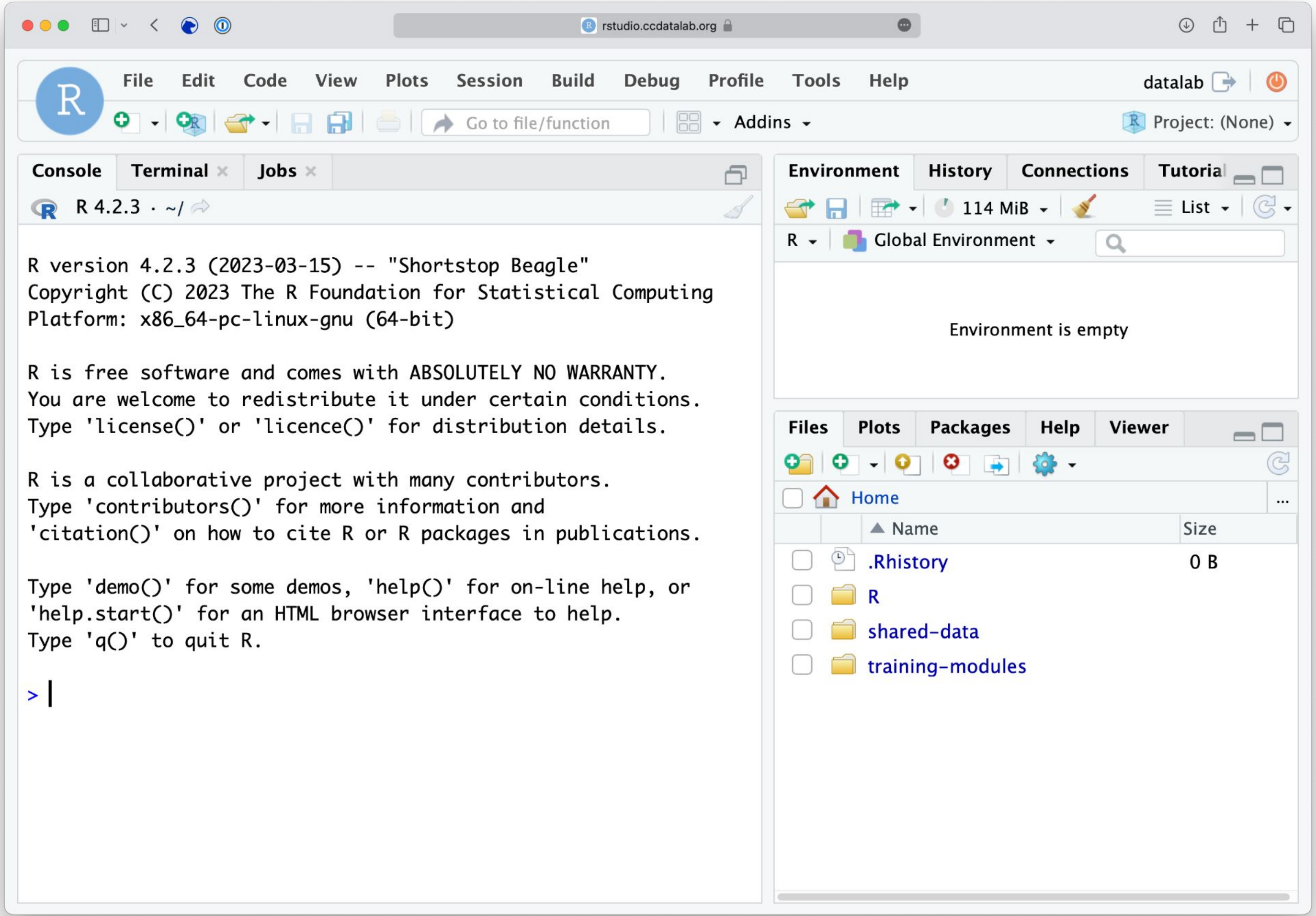
Command line vs GUI (graphics user interface)

- An interface is how you interact with a program

- GUI’s have buttons you can *click* to do things, but…
- Command-line interfaces (CLI) have you *type* out things to do them

RStudio Server: A basic guide

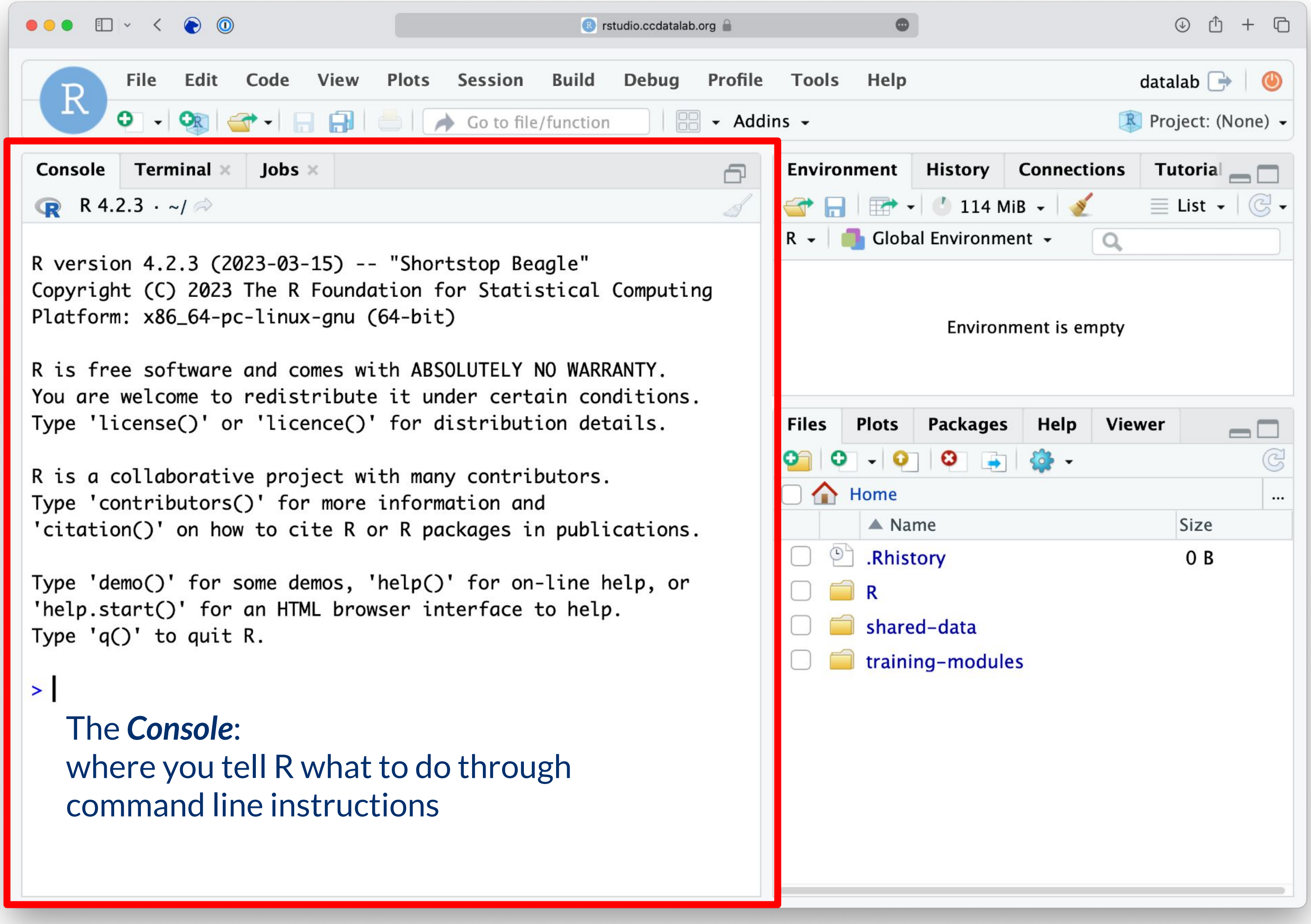
(FYI, we're now at `R 4.4.0`)



```
R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

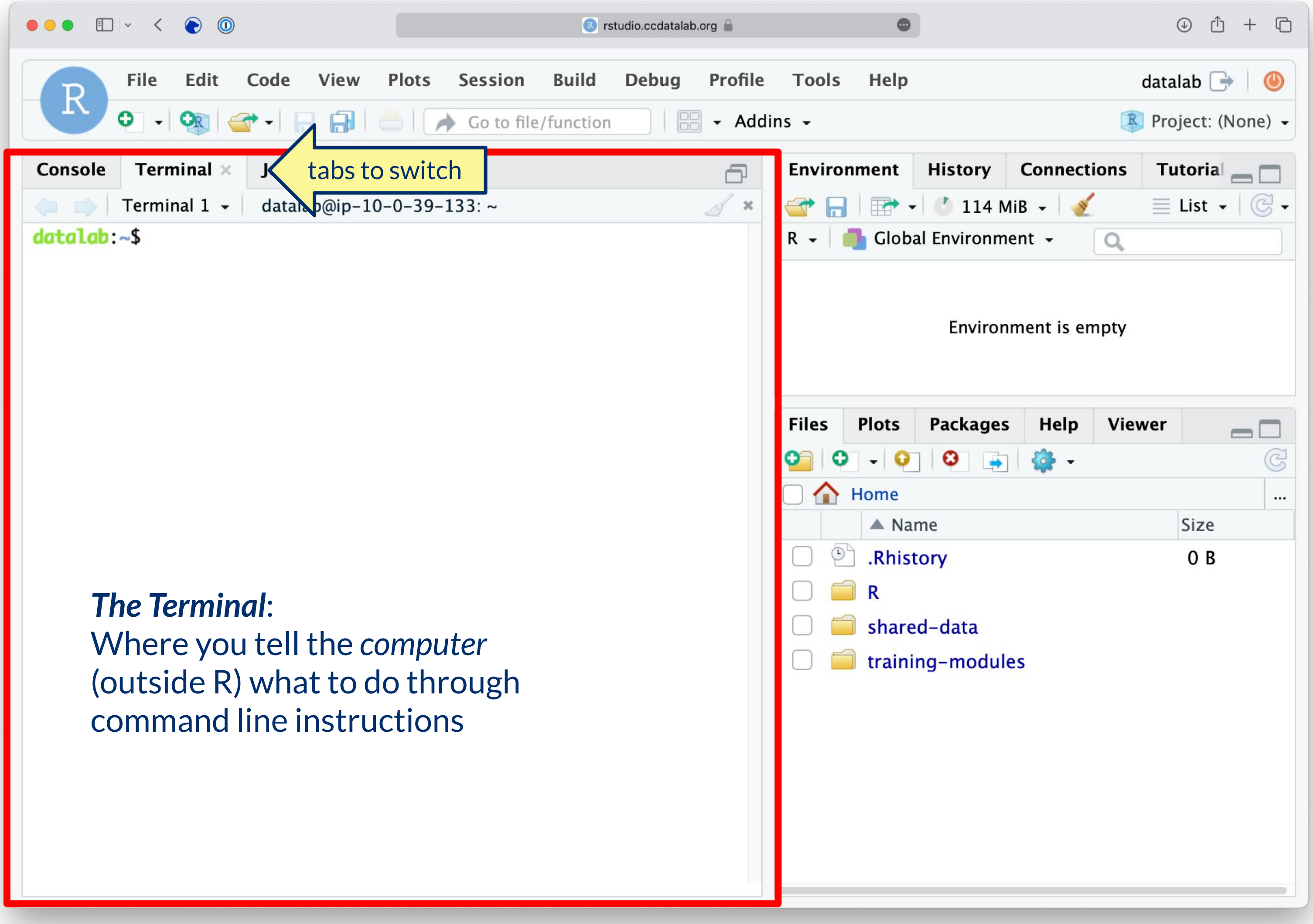
Environment is empty

▲ Name	Size
.Rhistory	0 B
R	
shared-data	
training-modules	



```
R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

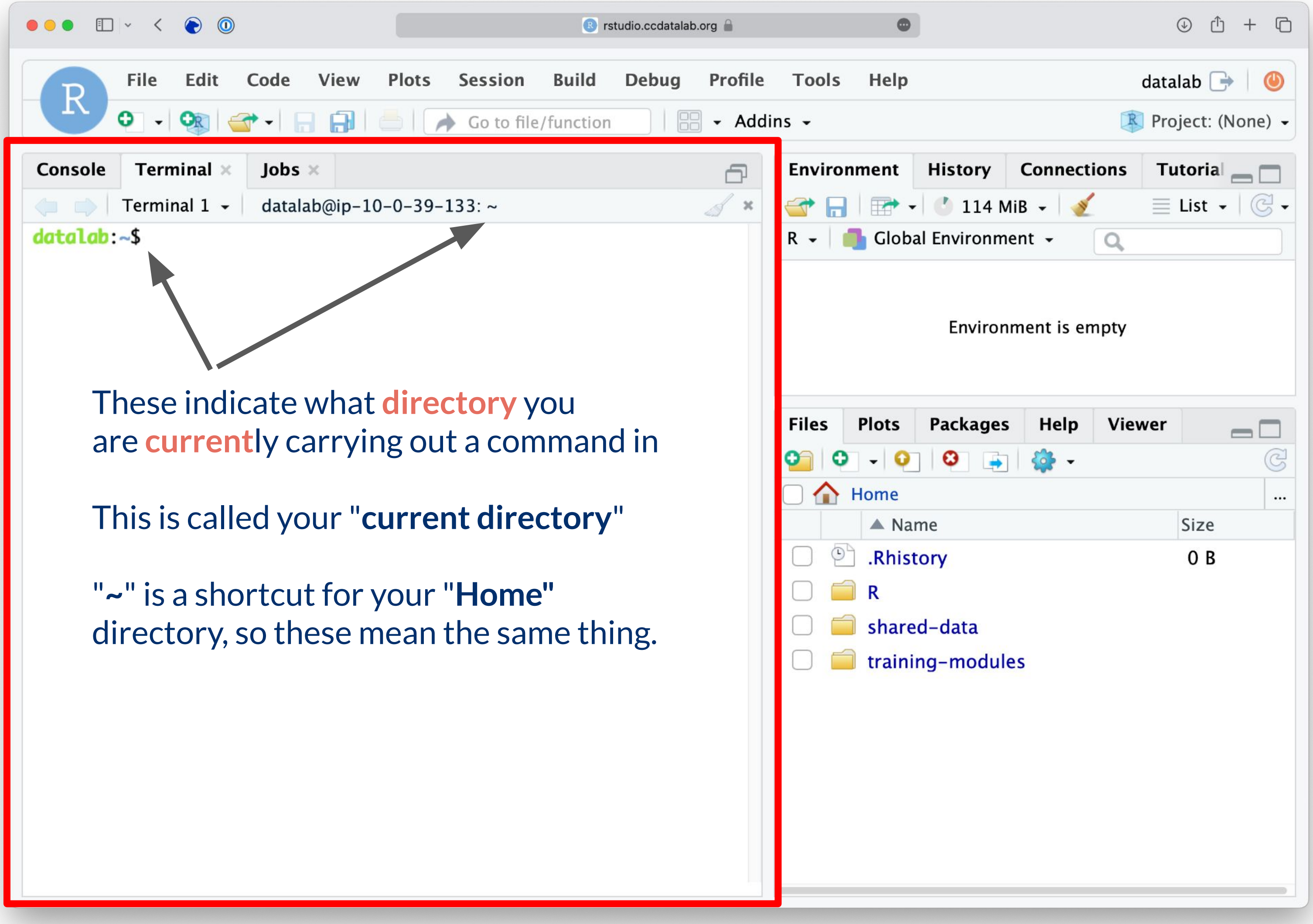
The ***Console***:
where you tell R what to do through command line instructions



```
datalab:~$
```

tabs to switch

The Terminal:
Where you tell the *computer* (outside R) what to do through command line instructions



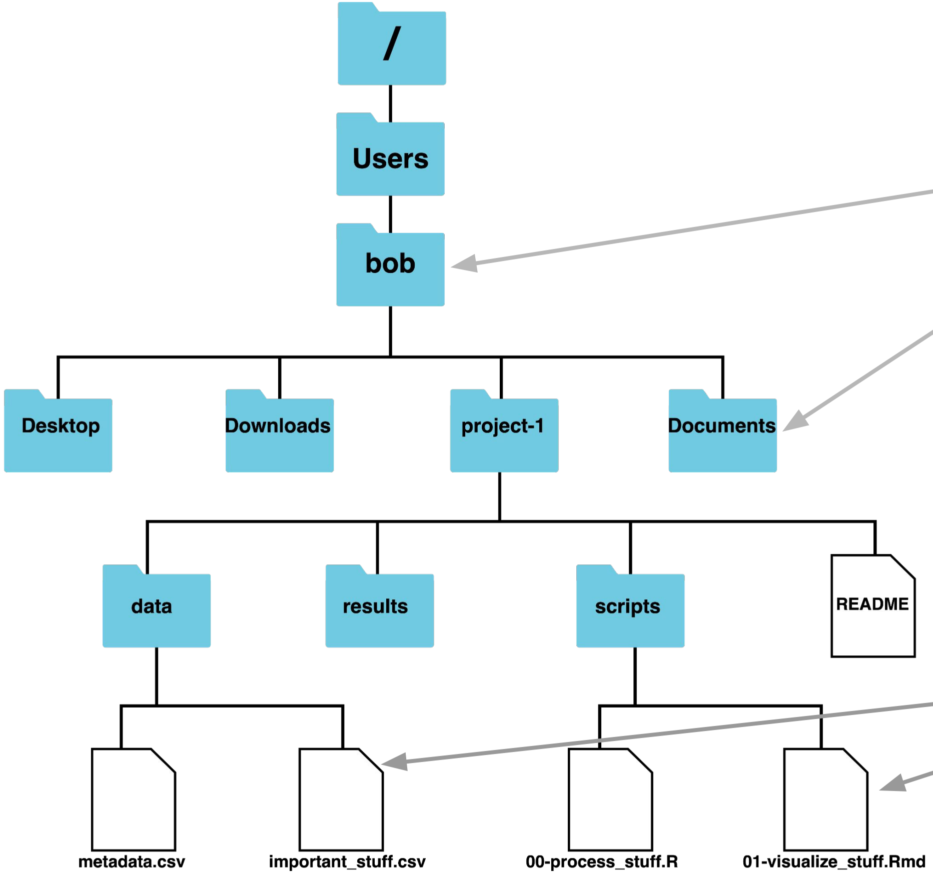
```
datalab:~$
```

These indicate what directory you are currently carrying out a command in

This is called your "**current directory**"

"**~**" is a shortcut for your "**Home**" directory, so these mean the same thing.

Example of a filesystem hierarchy



Folders aka directories

Files

We are always working somewhere!

Assume we are working “from” the **bob** directory.
This means **bob** is the ***current (working) directory***

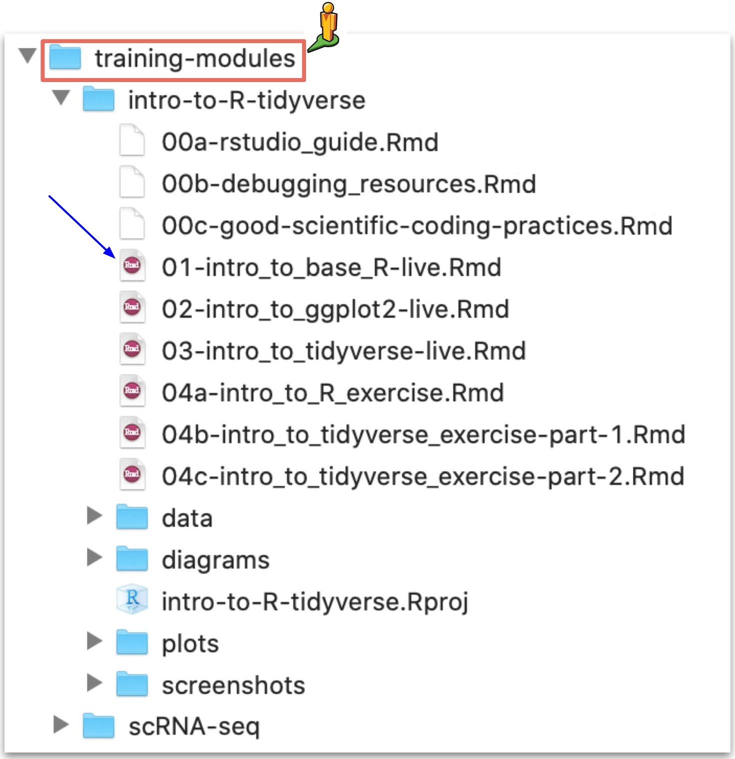
We therefore need to know the **paths** to files we are working with, relative to our working directory, to be able to use those files in our code.

The file we are working on

Relative path: `project-1/scripts/00-process_stuff.R`

Let's look at our workshop files

Let’s say we want access to `01-intro_to_base_R-live.Rmd`



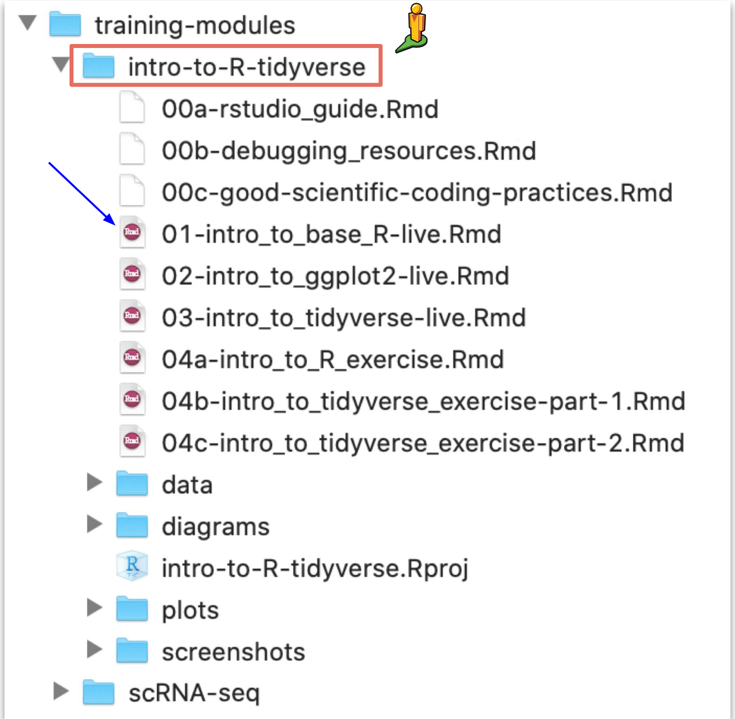
- training-modules
  - intro-to-R-tidyverse
    - 00a-rstudio_guide.Rmd
    - 00b-debugging_resources.Rmd
    - 00c-good-scientific-coding-practices.Rmd
    - 01-intro_to_base_R-live.Rmd
    - 02-intro_to_ggplot2-live.Rmd
    - 03-intro_to_tidyverse-live.Rmd
    - 04a-intro_to_R_exercise.Rmd
    - 04b-intro_to_tidyverse_exercise-part-1.Rmd
    - 04c-intro_to_tidyverse_exercise-part-2.Rmd
    - data
    - diagrams
    - intro-to-R-tidyverse.Rproj
    - plots
    - screenshots
  - scRNA-seq

Current/working directory = `training-modules`

File path = `intro-to-R-tidyverse/01-intro_to_base_R-live.Rmd`

Relative paths depend on your working directory

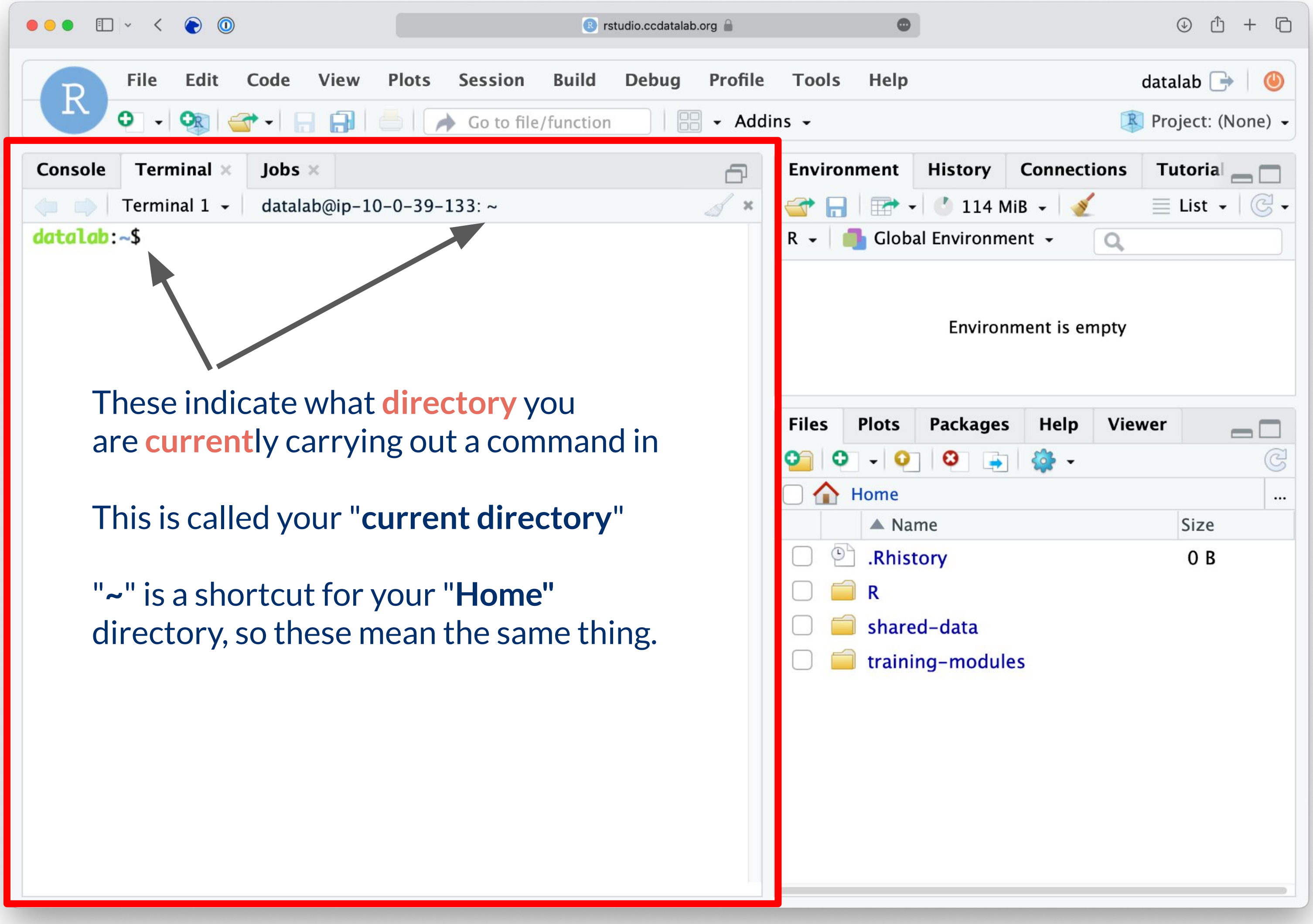
Let's say we want access to `01-intro_to_base_R-live.Rmd`



```
▼ training-modules
  ▼ intro-to-R-tidyverse
      00a-rstudio_guide.Rmd
      00b-debugging_resources.Rmd
      00c-good-scientific-coding-practices.Rmd
      01-intro_to_base_R-live.Rmd
      02-intro_to_ggplot2-live.Rmd
      03-intro_to_tidyverse-live.Rmd
      04a-intro_to_R_exercise.Rmd
      04b-intro_to_tidyverse_exercise-part-1.Rmd
      04c-intro_to_tidyverse_exercise-part-2.Rmd
    ▶ data
    ▶ diagrams
      intro-to-R-tidyverse.Rproj
    ▶ plots
    ▶ screenshots
  ▶ scRNA-seq
```

Current/working directory = `training-modules/intro-to-R-tidyverse`

File path = `01-intro_to_base_R-live.Rmd`



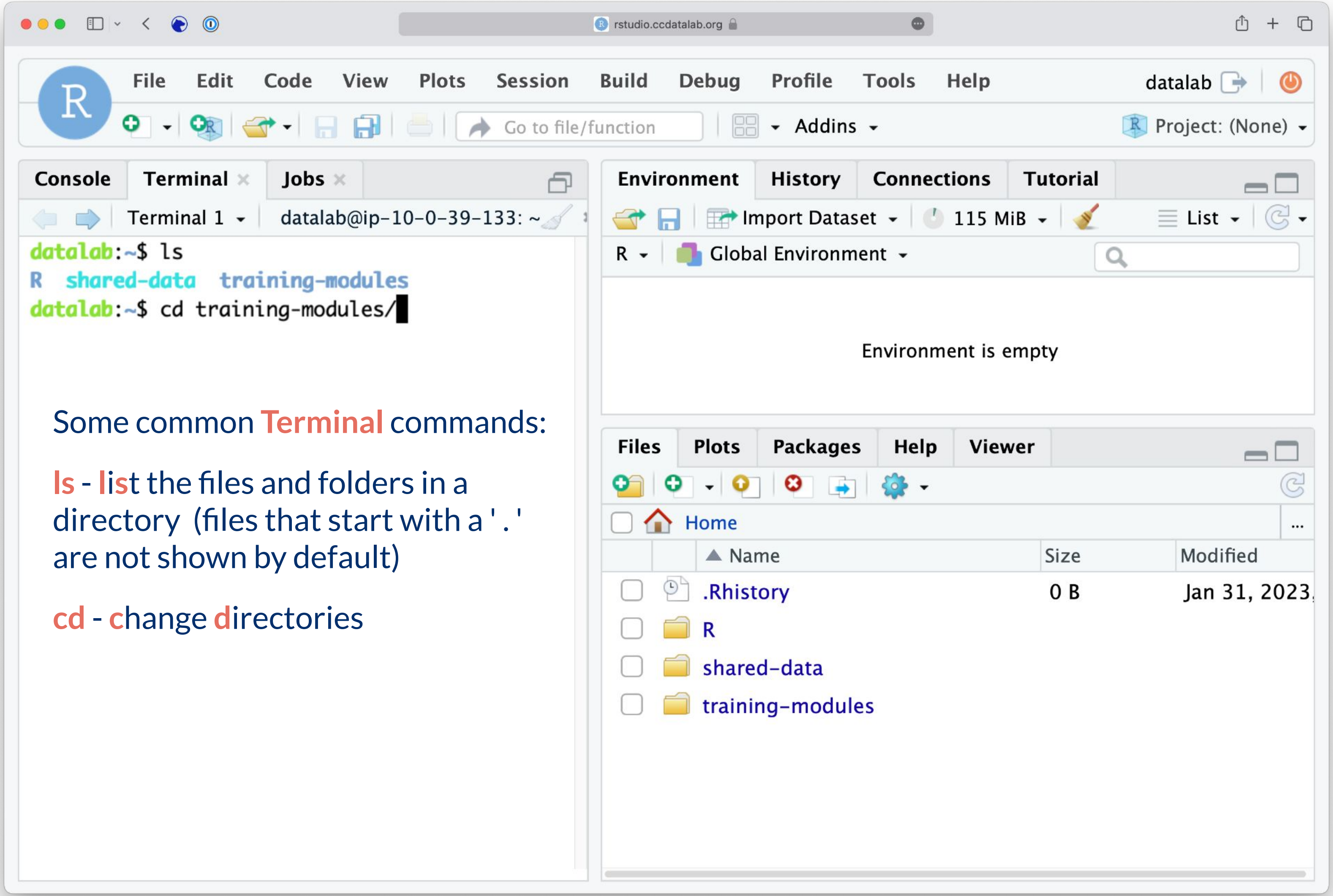
```
datalab:~$
```

Name	Size
.Rhistory	0 B
R	
shared-data	
training-modules	

These indicate what directory you are currently carrying out a command in

This is called your "**current directory**"

"**~**" is a shortcut for your "**Home**" directory, so these mean the same thing.



```
datalab:~$ ls
R  shared-data  training-modules
datalab:~$ cd training-modules/
```

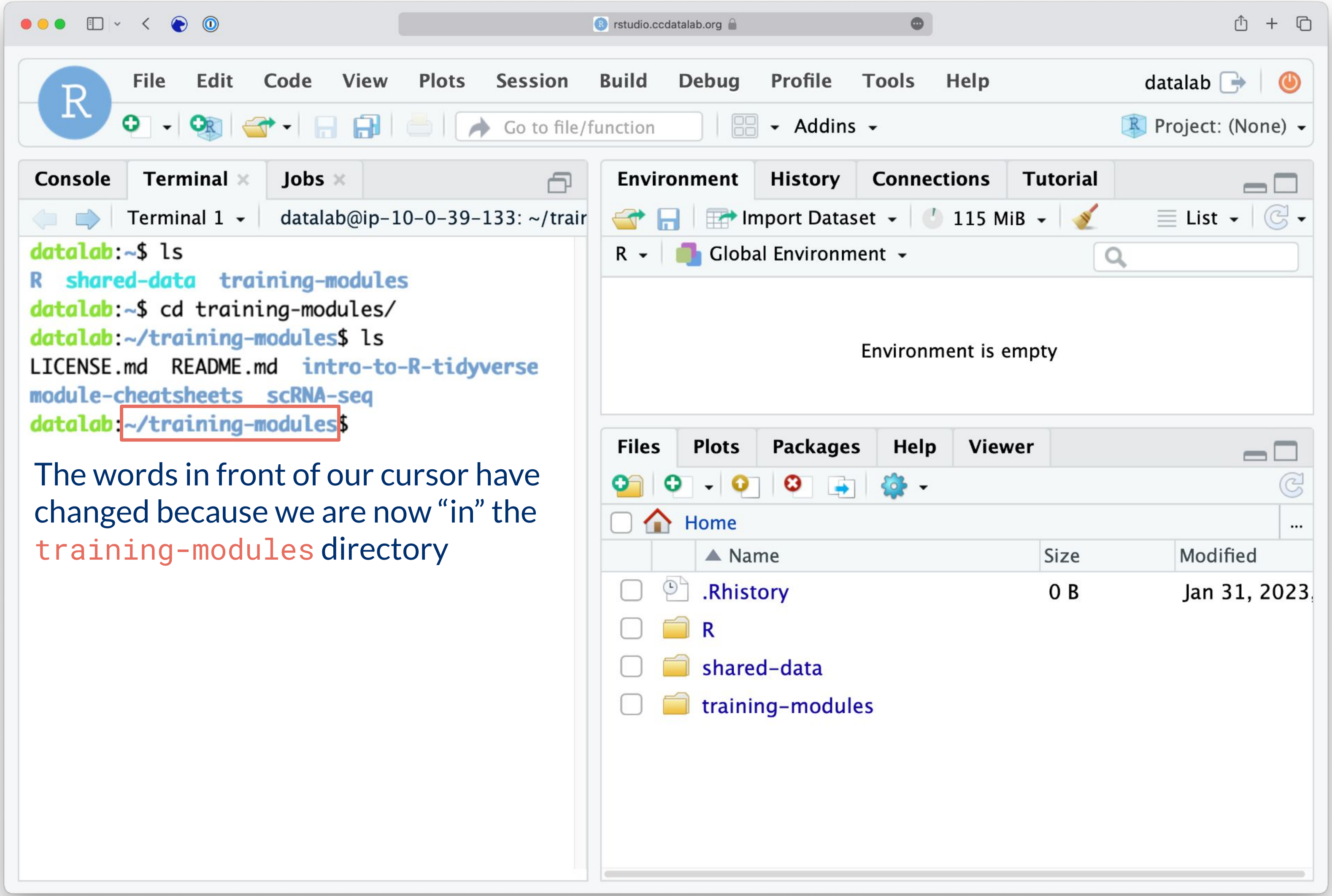
Environment is empty

Name	Size	Modified
.Rhistory	0 B	Jan 31, 2023,
R		
shared-data		
training-modules		

Some common Terminal commands:

ls - list the files and folders in a directory (files that start with a ' . ' are not shown by default)

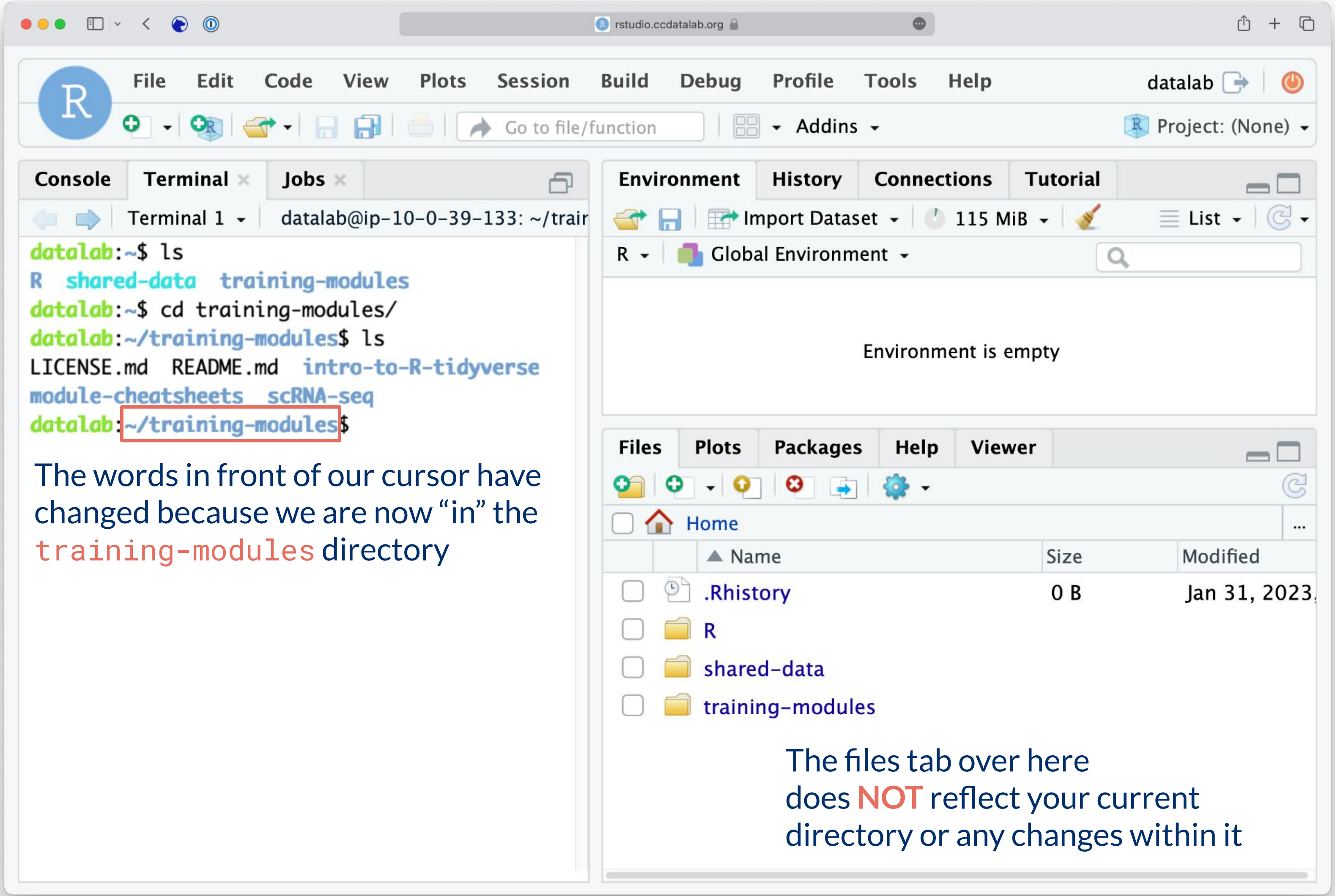
cd - change directories



```
datalab:~$ ls
R  shared-data  training-modules
datalab:~$ cd training-modules/
datalab:~/training-modules$ ls
LICENSE.md  README.md  intro-to-R-tidyverse
module-cheatsheets  scRNA-seq
datalab:~/training-modules$
```

Name	Size	Modified
.Rhistory	0 B	Jan 31, 2023,
R		
shared-data		
training-modules		

The words in front of our cursor have changed because we are now "in" the `training-modules` directory



```
datalab:~$ ls
R  shared-data  training-modules
datalab:~$ cd training-modules/
datalab:~/training-modules$ ls
LICENSE.md  README.md  intro-to-R-tidyverse
module-cheatsheets  scRNA-seq
datalab:~/training-modules$
```

Environment is empty

Name	Size	Modified
.Rhistory	0 B	Jan 31, 2023
R		
shared-data		
training-modules		

The words in front of our cursor have changed because we are now “in” the `training-modules` directory

The files tab over here does **NOT** reflect your current directory or any changes within it

Introduction to R

The Data Lab

R programming

Programming: making executable scripts for accomplishing a task
(in this case, data analysis is our task)

Scripts allow others to see, step-by-step, what you did.

Why we use R:

- It's free and open-source
- People make cool packages that do stuff for us
- Many researchers in genomics use it (as well as Python)

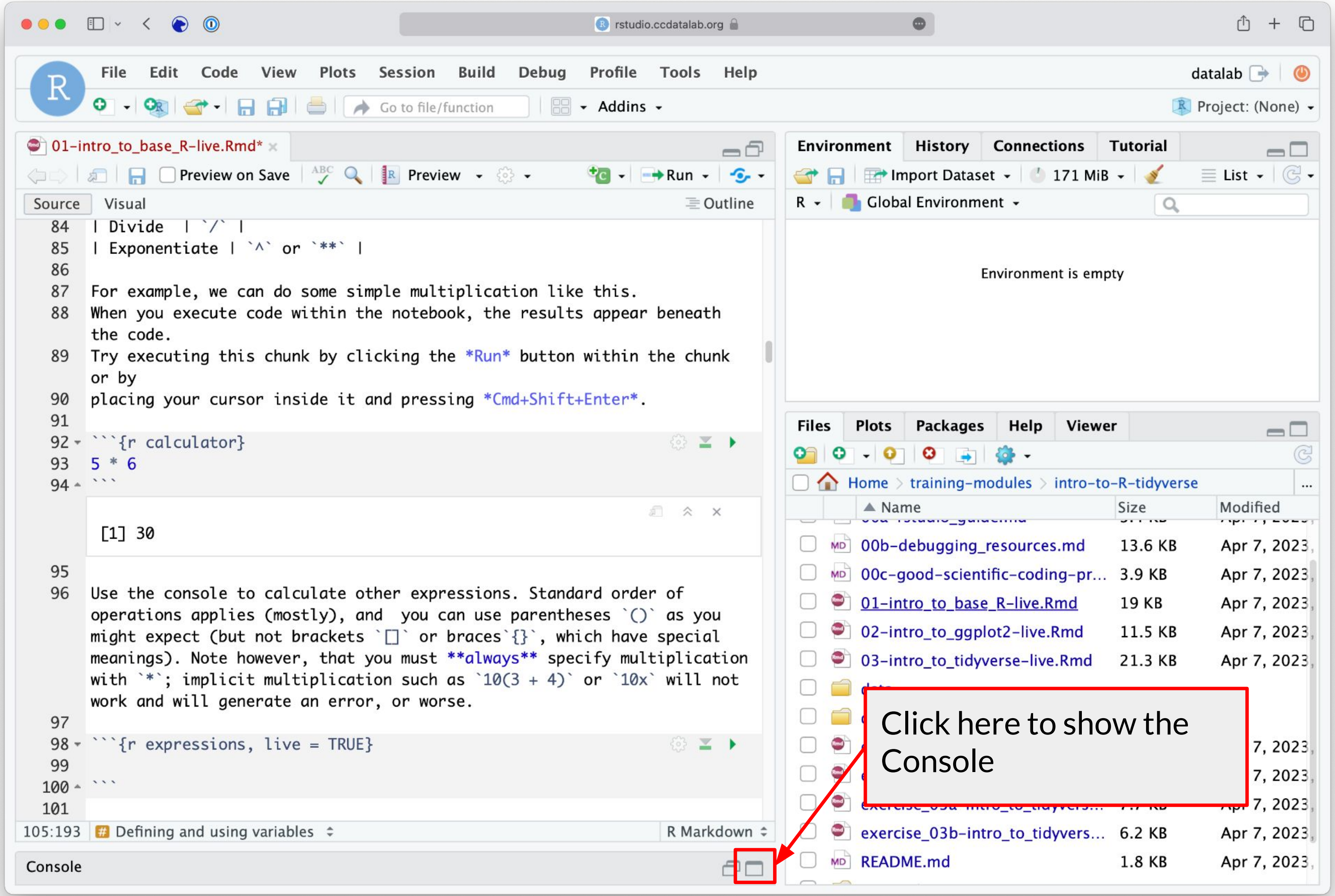
R, RStudio, and RStudio Server

R is a statistical programming language.

RStudio is an IDE for working in R

- IDE: Integrated Development Environment
- We write R code using the (free!) RStudio IDE

RStudio Server allows us to run the **RStudio IDE** from a browser



~~~~
| Divide   | `/` |
| Exponentiate | `^` or `**` |

For example, we can do some simple multiplication like this.
When you execute code within the notebook, the results appear beneath the code.
Try executing this chunk by clicking the *Run* button within the chunk or by
placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```{r calculator}
5 * 6
```
~~~~

```
[1] 30
```

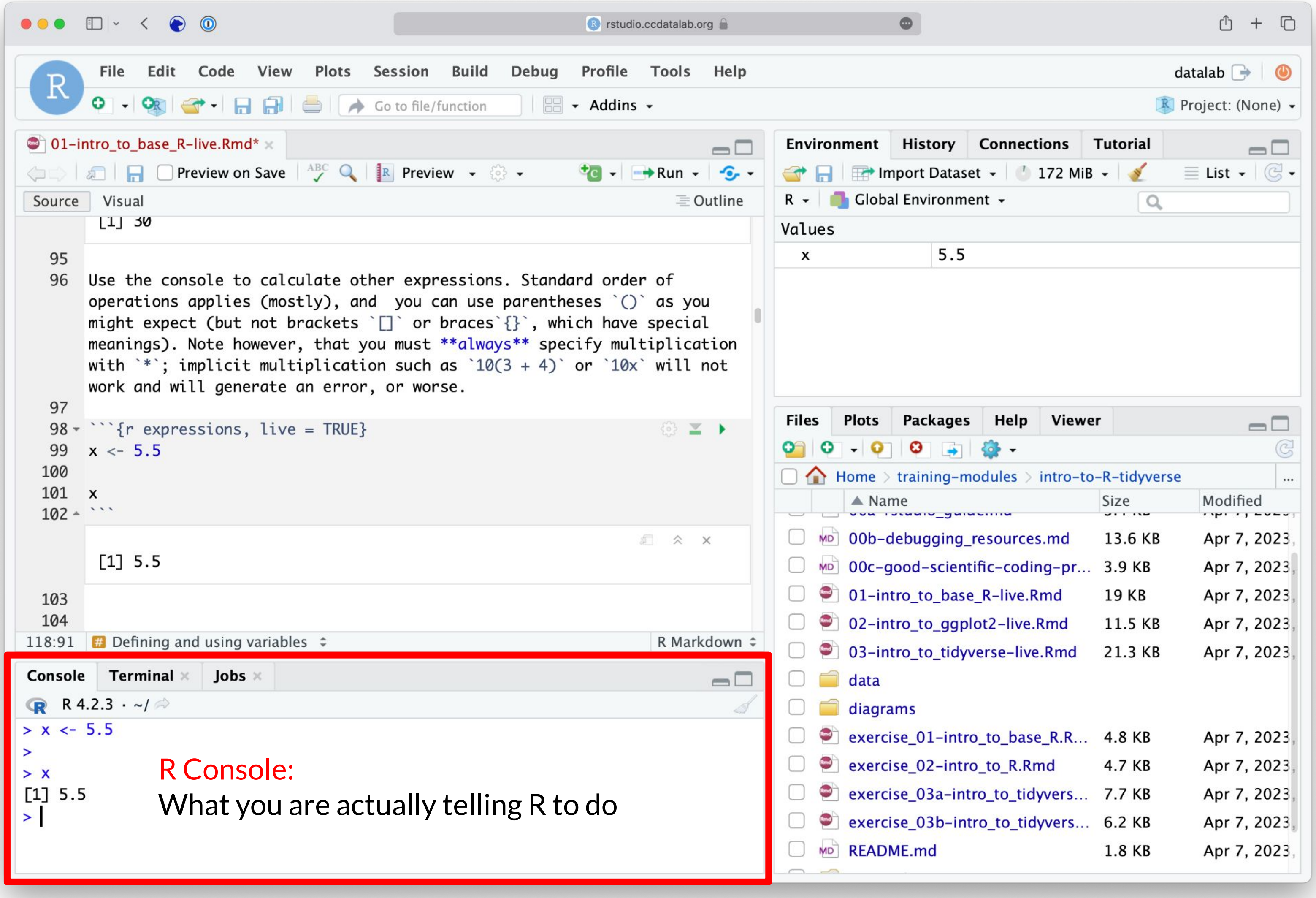
~~~~
Use the console to calculate other expressions. Standard order of operations applies (mostly), and  you can use parentheses `()` as you might expect (but not brackets `[]` or braces`{}`, which have special meanings). Note however, that you must **always** specify multiplication with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not work and will generate an error, or worse.

```{r expressions, live = TRUE}

```
~~~~

Environment is empty

Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023,
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023,
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023,
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023,
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023,
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023,
README.md	1.8 KB	Apr 7, 2023,

Click here to show the Console



```
[1] 30
```

Use the console to calculate other expressions. Standard order of operations applies (mostly), and you can use parentheses `()` as you might expect (but not brackets `[]` or braces`{}`, which have special meanings). Note however, that you must **always** specify multiplication with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not work and will generate an error, or worse.

~~~
```{r expressions, live = TRUE}
x <- 5.5

x
```
~~~

```
[1] 5.5
```

Values

x	5.5

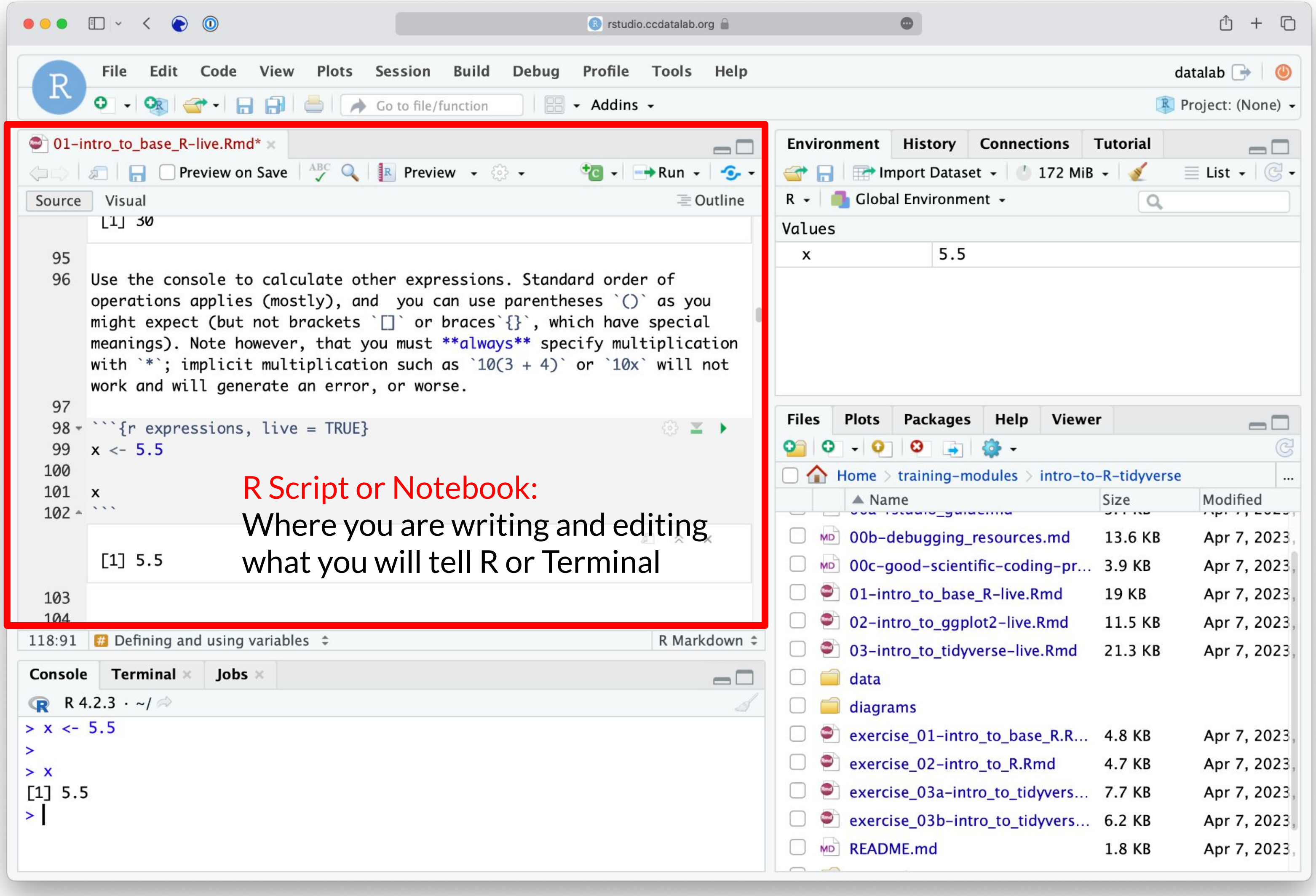
Home > training-modules > intro-to-R-tidyverse

Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023,
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023,
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023,
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023,
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023,
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023,
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023,
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023,
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023,
README.md	1.8 KB	Apr 7, 2023,

```
> x <- 5.5
>
> x
[1] 5.5
>
```

R Console:

What you are actually telling R to do



~~~
[1] 30

Use the console to calculate other expressions. Standard order of operations applies (mostly), and  you can use parentheses `()` as you might expect (but not brackets `[]` or braces`{}`, which have special meanings). Note however, that you must **always** specify multiplication with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not work and will generate an error, or worse.

```{r expressions, live = TRUE}
x <- 5.5

x
```

 [1] 5.5
~~~

R Script or Notebook:

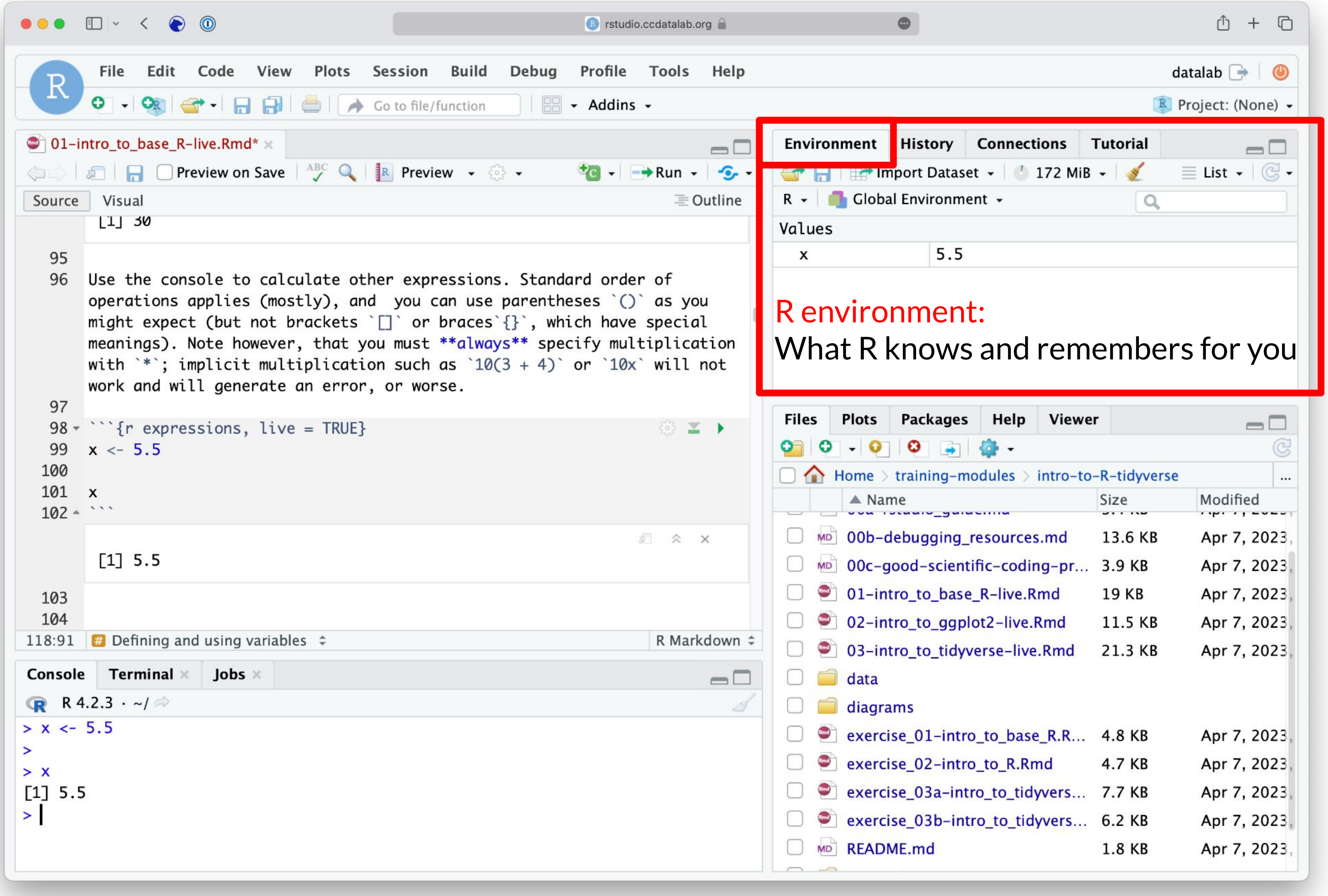
Where you are writing and editing what you will tell R or Terminal

```
> x <- 5.5
>
> x
[1] 5.5
>
```

Values

x	5.5

Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023,
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023,
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023,
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023,
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023,
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023,
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023,
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023,
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023,
README.md	1.8 KB	Apr 7, 2023,



~~~
[1] 30

Use the console to calculate other expressions. Standard order of
operations applies (mostly), and  you can use parentheses `()` as you
might expect (but not brackets `[]` or braces`{}`, which have special
meanings). Note however, that you must **always** specify multiplication
with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not
work and will generate an error, or worse.

```{r expressions, live = TRUE}
x <- 5.5

x
```

[1] 5.5
~~~

Console:

```
> x <- 5.5
>
> x
[1] 5.5
>
```

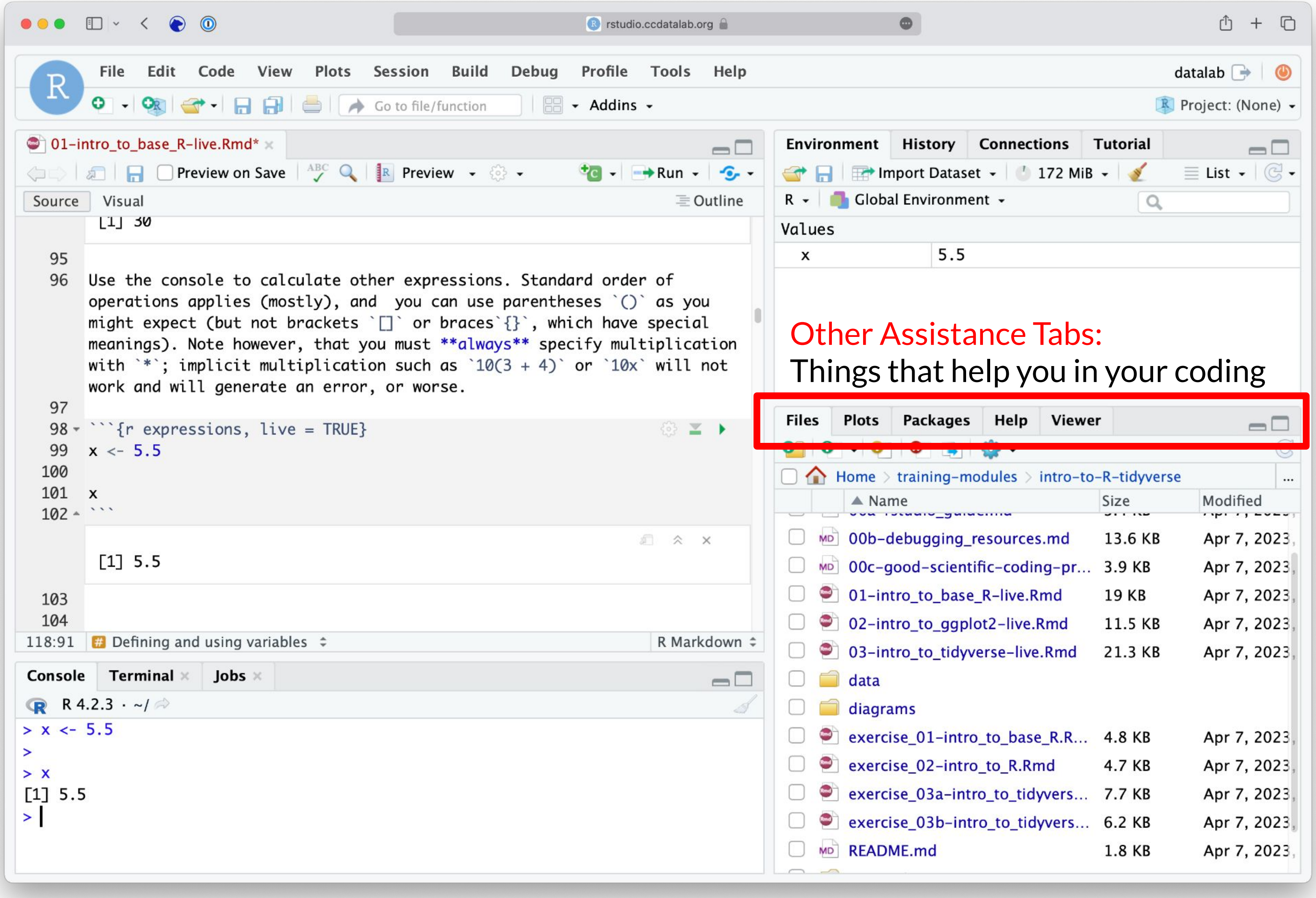
Values

x	5.5

Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023
README.md	1.8 KB	Apr 7, 2023

R environment:

What R knows and remembers for you



Other Assistance Tabs:

Things that help you in your coding

```
[1] 30
```

Use the console to calculate other expressions. Standard order of operations applies (mostly), and you can use parentheses `()` as you might expect (but not brackets `[]` or braces`{}`, which have special meanings). Note however, that you must **always** specify multiplication with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not work and will generate an error, or worse.

~~~
```{r expressions, live = TRUE}
x <- 5.5

x
```
~~~

```
[1] 5.5
```

Values

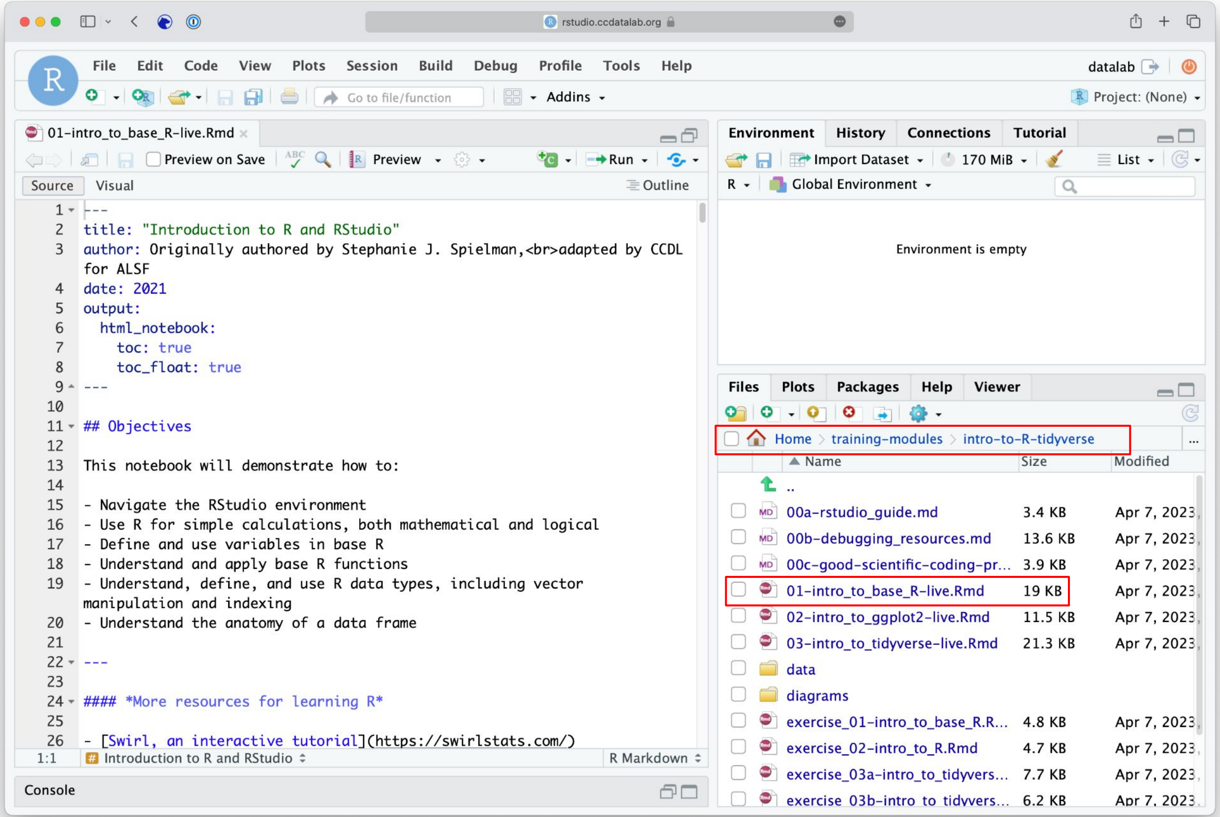
x	5.5

```
> x <- 5.5
>
> x
[1] 5.5
>
```

Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023,
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023,
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023,
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023,
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023,
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023,
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023,
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023,
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023,
README.md	1.8 KB	Apr 7, 2023,

R Notebooks

Use the "**Files**" tab to open: `training-modules/intro-to-R-tidyverse/01-intro_to_base_R-live.Rmd`



```
---
title: "Introduction to R and RStudio"
author: Originally authored by Stephanie J. Spielman,<br>adapted by CCDL for ALSF
date: 2021
output:
  html_notebook:
    toc: true
    toc_float: true
---

## Objectives

This notebook will demonstrate how to:

- Navigate the RStudio environment
- Use R for simple calculations, both mathematical and logical
- Define and use variables in base R
- Understand and apply base R functions
- Understand, define, and use R data types, including vector manipulation and indexing
- Understand the anatomy of a data frame

---

#### *More resources for learning R*

- [Swirl, an interactive tutorial](https://swirlstats.com/)
```

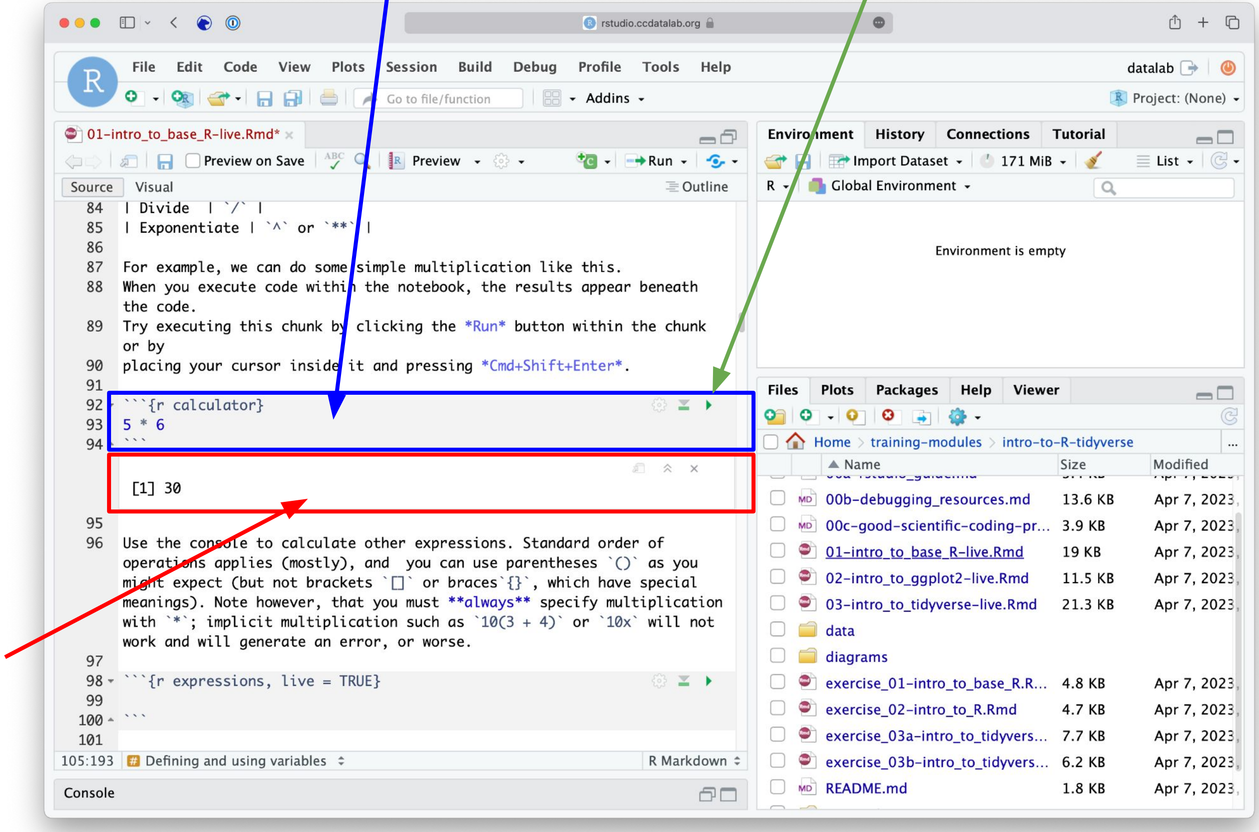
Name	Size	Modified
..		
00a-rstudio_guide.md	3.4 KB	Apr 7, 2023
00b-debugging_resources.md	13.6 KB	Apr 7, 2023
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023

R Notebooks

R Notebooks allow you to have files that show both your code and results

Executable code chunk

Can click here to run a code chunk



~~~~
| Divide  | `/` |
| Exponentiate | `^` or `**` |

For example, we can do some simple multiplication like this.
When you execute code within the notebook, the results appear beneath the code.
Try executing this chunk by clicking the *Run* button within the chunk or by
placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```{r calculator}
5 * 6
```

[1] 30


Use the console to calculate other expressions. Standard order of operations applies (mostly), and  you can use parentheses `()` as you might expect (but not brackets `[]` or braces`{}`, which have special meanings). Note however, that you must **always** specify multiplication with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not work and will generate an error, or worse.

```{r expressions, live = TRUE}

```
~~~~

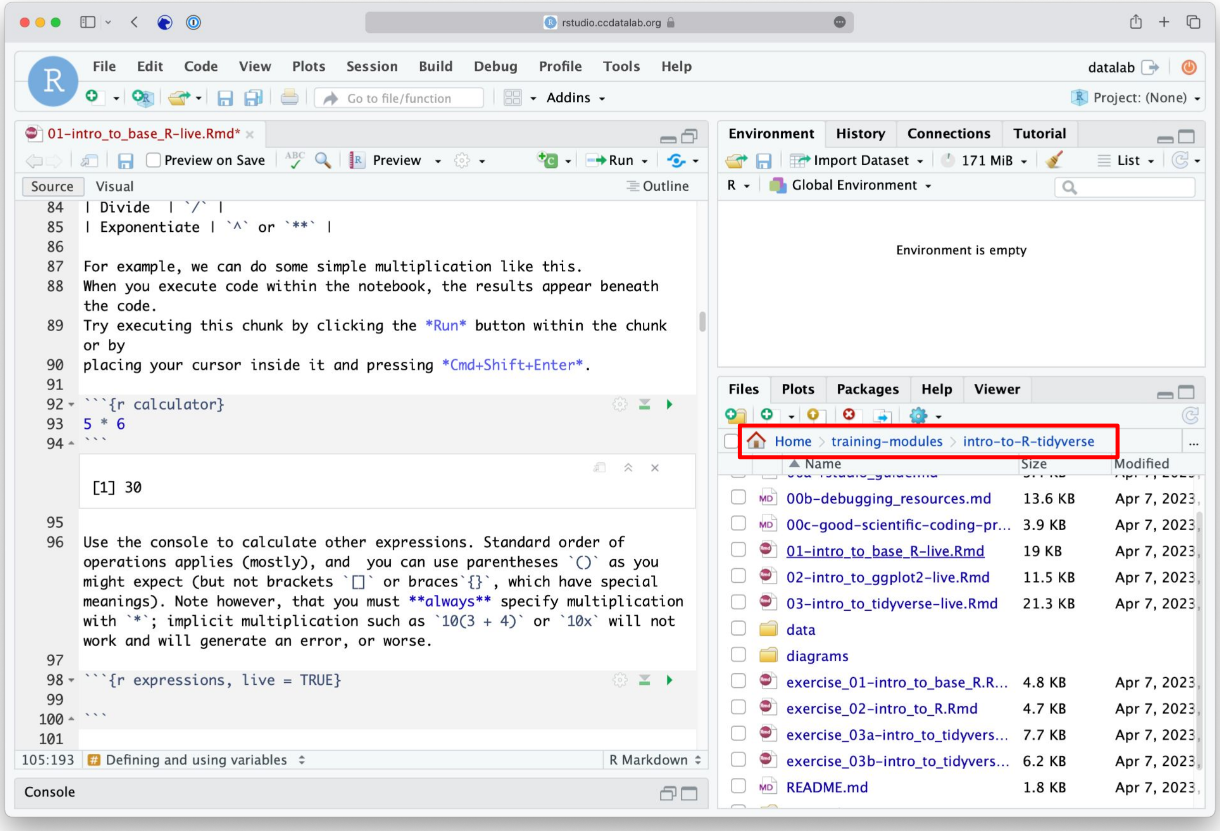
Name	Size	Modified
00b-debugging_resources.md	13.6 KB	Apr 7, 2023
00c-good-scientific-coding-pr...	3.9 KB	Apr 7, 2023
01-intro_to_base_R-live.Rmd	19 KB	Apr 7, 2023
02-intro_to_ggplot2-live.Rmd	11.5 KB	Apr 7, 2023
03-intro_to_tidyverse-live.Rmd	21.3 KB	Apr 7, 2023
data		
diagrams		
exercise_01-intro_to_base_R.R...	4.8 KB	Apr 7, 2023
exercise_02-intro_to_R.Rmd	4.7 KB	Apr 7, 2023
exercise_03a-intro_to_tidyvers...	7.7 KB	Apr 7, 2023
exercise_03b-intro_to_tidyvers...	6.2 KB	Apr 7, 2023
README.md	1.8 KB	Apr 7, 2023

Output from above code chunk

R Notebooks

- Code that runs in R Notebooks uses wherever the file is saved as its *current directory*

- Warning! That may not be the directory shown in the files pane *or* the console!



~~~
| Divide   | `/` |
| Exponentiate | `^` or `**` |

For example, we can do some simple multiplication like this.
When you execute code within the notebook, the results appear beneath
the code.
Try executing this chunk by clicking the *Run* button within the chunk
or by
placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```{r calculator}
5 * 6
~~~

[1] 30

Use the console to calculate other expressions. Standard order of
operations applies (mostly), and you can use parentheses `()` as you
might expect (but not brackets `[]` or braces`{}`, which have special
meanings). Note however, that you must **always** specify multiplication
with `*`; implicit multiplication such as `10(3 + 4)` or `10x` will not
work and will generate an error, or worse.

```{r expressions, live = TRUE}

```

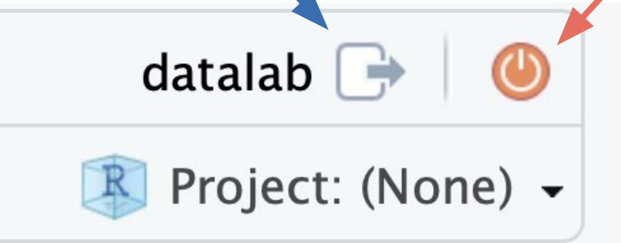
```

Environment is empty

Home > training-modules > intro-to-R-tidyverse

| Name | Size | Modified |
|---|---|---|
| 00b-debugging_resources.md | 13.6 KB | Apr 7, 2023 |
| 00c-good-scientific-coding-pr... | 3.9 KB | Apr 7, 2023 |
| 01-intro_to_base_R-live.Rmd | 19 KB | Apr 7, 2023 |
| 02-intro_to_ggplot2-live.Rmd | 11.5 KB | Apr 7, 2023 |
| 03-intro_to_tidyverse-live.Rmd | 21.3 KB | Apr 7, 2023 |
| data | | |
| diagrams | | |
| exercise_01-intro_to_base_R.R... | 4.8 KB | Apr 7, 2023 |
| exercise_02-intro_to_R.Rmd | 4.7 KB | Apr 7, 2023 |
| exercise_03a-intro_to_tidyvers... | 7.7 KB | Apr 7, 2023 |
| exercise_03b-intro_to_tidyvers... | 6.2 KB | Apr 7, 2023 |
| README.md | 1.8 KB | Apr 7, 2023 |
```

RStudio Sessions

- On the server, R is running many times at once
    - Each user has their own “**Session**” running, with its own memory and processes
- We will usually want to start new sessions between notebooks to keep the environment clean

Log out of website

End the current session and start new session


datalab
Project: (None)