

# Introduction to R and RStudio

Meryl Brodsky, Communication & Information Librarian  
Hannah Chapman Tripp, Biosciences Librarian  
September 30, 2022

# What is R/R Studio?

## **R**

- Programming language
- Also refers to the software that interprets it's scripts
- For data analysis and graphics

## **RStudio**

- Wrapper around the R language
- User interface for working with R
- Extends what R can do and facilitates writing R code

# Why R and RStudio?

- Rely on written commands instead of point and click
- Great for reproducibility
- Interdisciplinary and extensible
- Works for data of all shapes and sizes
- Produces graphics
- R Studio is an Integrated Development Environment for working with R
- It allows you to write code, navigate files, inspect variables and visualize plots
- R Studio can be used for other things such as version control, developing packages or writing Shiny apps, but not today

# Setting up a project

The Working Directory

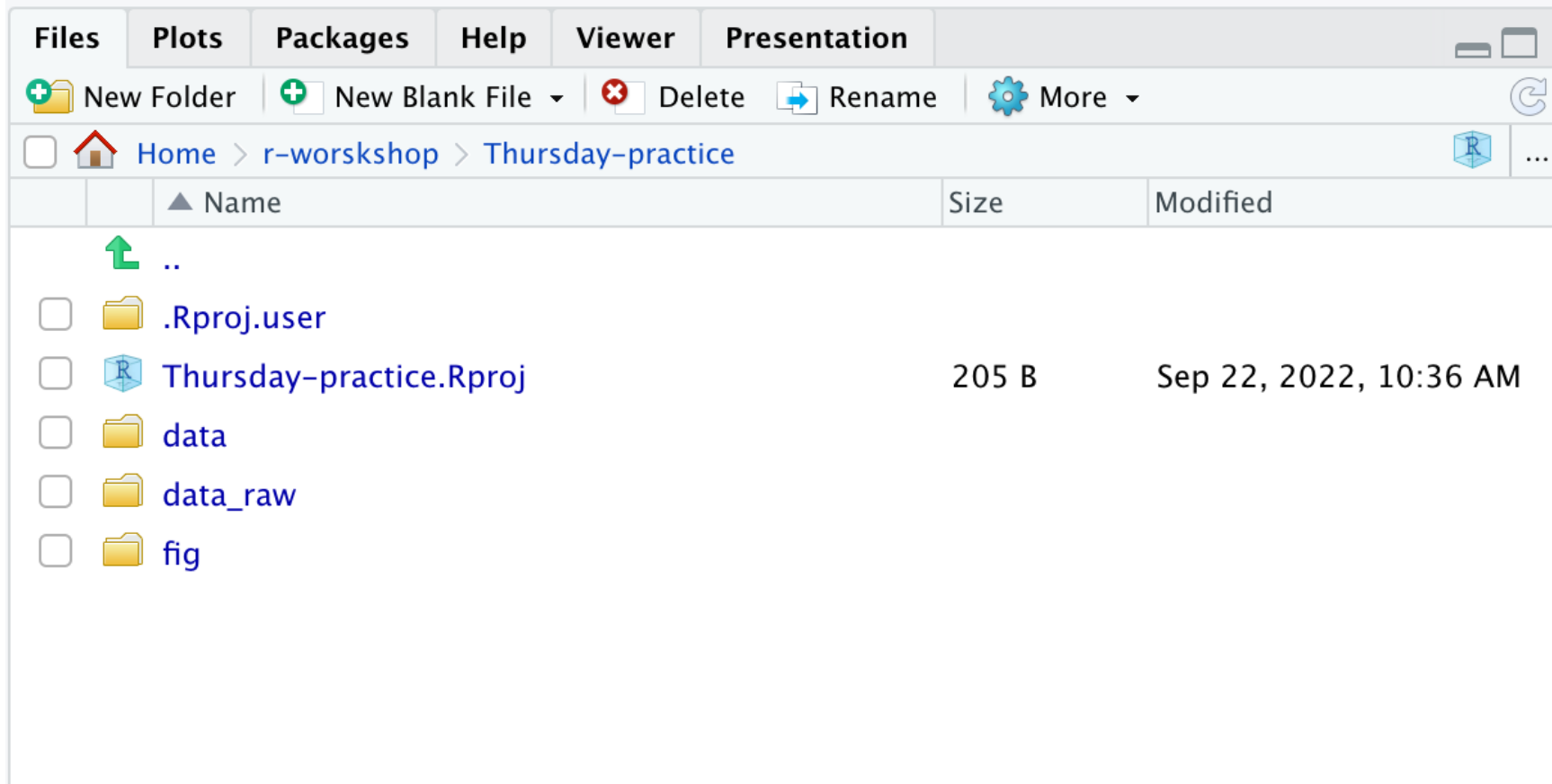
# Suggested subdirectories

- data\_raw                      for raw data
- data/                            for intermediate datasets
- fig/                              graphics generated by scripts







## Others for another time

- data\_output/                modified versions of raw data
- documents/                 outlines, drafts, other text
- scripts/                      R scripts for different analyses or plotting

# Your Working Directory



The screenshot shows the RStudio interface with the file browser pane open. The breadcrumb path is Home > r-worskshop > Thursday-practice. The file list includes a parent directory icon (green arrow), a .Rproj.user folder, a Thursday-practice.Rproj file (205 B, modified Sep 22, 2022, 10:36 AM), and three subfolders: data, data\_raw, and fig.

	▲ Name	Size	Modified
<input type="checkbox"/>	 ..		
<input type="checkbox"/>	 .Rproj.user		
<input type="checkbox"/>	 Thursday-practice.Rproj	205 B	Sep 22, 2022, 10:36 AM
<input type="checkbox"/>	 data		
<input type="checkbox"/>	 data_raw		
<input type="checkbox"/>	 fig		

# RStudio Interface

The screenshot displays the RStudio R-Workshop interface. The main editor shows a script with the following content:

```
1 |
2 | ### Creating objects in R
3 |
4 |
5 |
6 |
7 |
8 |
9 |
10 |
11 |
12 |
13 |
14 |
15 |
16 |
17 |
18 |
19 |
20 | ### Challenge
21 | ##
22 |
```

The Environment pane on the right shows the Global Environment, which is currently empty.

The Files pane on the bottom right shows the project structure:

Name	Size	Modified
..		
.Rproj.user		
R-Workshop.Rproj	205 B	Sep 1, 2022, 12:24 PM
Script.R	0 B	Sep 1, 2022, 12:25 PM
data_raw		
data		
fig		

The Console pane at the bottom shows the following output:

```
R 4.2.1 · ~/r-worskshop/R-Workshop/
Natural language support but running in an English locale
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.
> get wd()
Error: unexpected symbol in "get wd"
> getwd()
[1] "/Users/mb59937/r-worskshop/R-Workshop"
>
```

# Interacting with R

Console vs Script



# Accepting commands

- Execute commands directly from the script editor by using the Ctrl + Enter shortcut (on Macs, Cmd + Return will work, too).
- The command on the current line in the script (indicated by the cursor) or all of the commands in the currently selected text will be sent to the console and executed when you press Ctrl + Enter.
- RStudio provides the Ctrl + 1 and Ctrl + 2 shortcuts which allow you to jump between the script and the console panes.
- When ready: >
- If waiting: +
- Cancel commands by pressing Esc.
- Stop sign when working

# Creating Objects

- To create an object, give it a name, followed by the assignment operator, followed by the value
- Assignment operator <-
- `X <- 3`
- Value of X is 3
- Can also use `x =3`, but not recommended
- `weight_kg <- 55`
- Shortcut "alt +" on PC, "option +" on Mac

# Object names

- Can be given any name
- Should be explicit but not too long
- Cannot start with a number
- R is case sensitive
- Try to avoid function names as object names
- Avoid dots “.” in object names
- Be consistent in code styling—style guides available online

# Objects vs Variables

and Saving your script

# Functions

- R is a functional programming language
- Functions are canned scripts
- Predefined
- Call a function by entering it into the console
- Ex: Sum of 3 and 4

```
> sum(3,4) #sum() is the function and 3,4 are the arguments
```

```
> weight_kg <- sqrt(10)
```

## Passing the argument to the function

- Arguments can be specified or take on default value
- You can specify a value instead of the default
- This is passing the argument to the function
  - > round(3.14159)
  - > ?round or args(round)
  - > round(3.14159, digits =2) or round(3.13159, 2)

# Help

Files Plots Packages Help Viewer

R: Arithmetic Mean rank

mean {base}

## Arithmetic Mean

### Description

Generic function for the (trimmed) arithmetic mean.

### Usage

```
mean(x, ...)
```

## Default S3 method:  
mean(x, trim = 0, na.rm = FALSE, ...)

### Arguments

**x** An R object. Currently there are methods for numeric/logical vectors and [date](#), [date-time](#) and [time interval](#) objects. Complex vectors are allowed for `trim = 0`, only.

**trim** the fraction (0 to 0.5) of observations to be trimmed from each end of `x` before the mean is computed. Values of `trim` outside that range are taken as the nearest endpoint.

mean  
mean.Date  
mean.default  
mean.difftime  
mean.POSIXct  
mean.POSIXt

# Vectors

Data Types



# Types of vectors

- “character”
- “numeric” (or “double”)
- “logical”
- “integer”
- “complex”
- “raw”

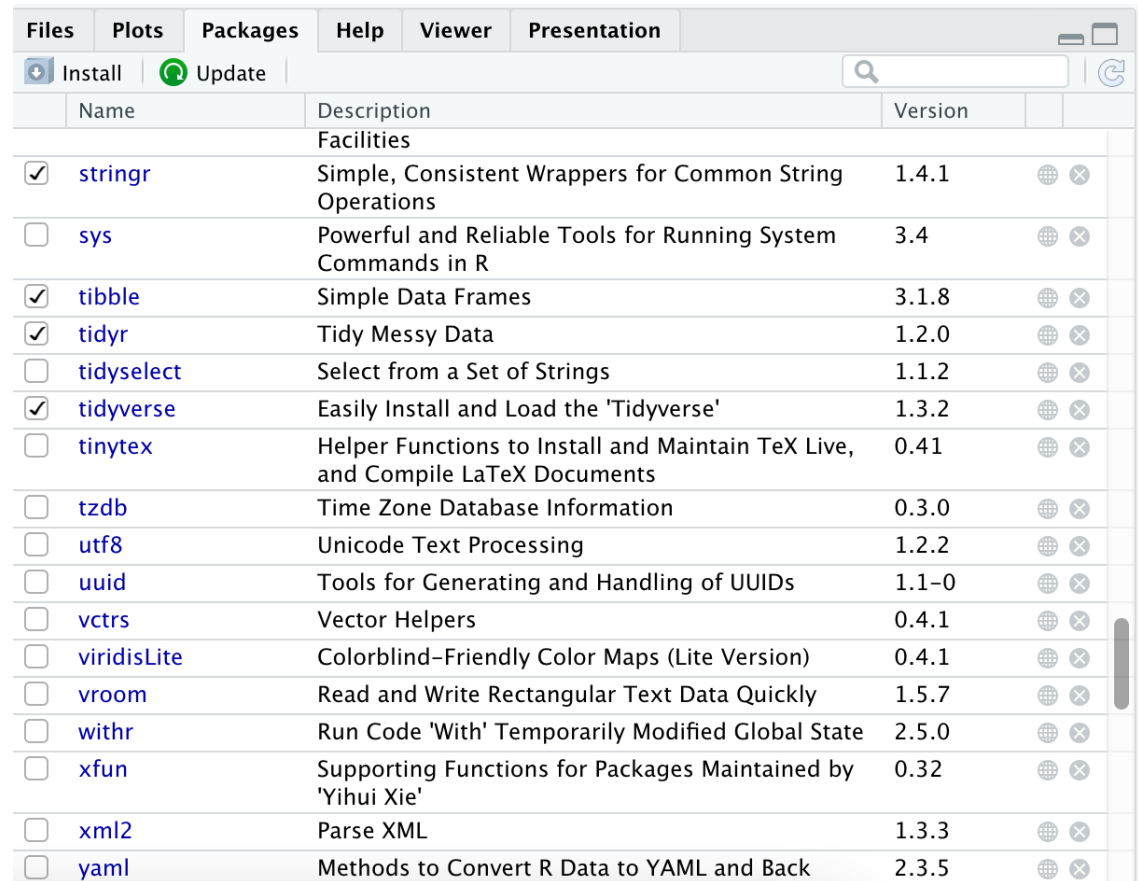
# Packages

Adding functionality to R

# Install the tidyverse package

```
install.packages("tidyverse")  
library(tidyverse)
```

# Installing Tidyverse visual



Files	Plots	Packages	Help	Viewer	Presentation
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/> Install	<input checked="" type="checkbox"/> Update		
Name	Description	Version			
Facilities					
<input checked="" type="checkbox"/> stringr	Simple, Consistent Wrappers for Common String Operations	1.4.1	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> sys	Powerful and Reliable Tools for Running System Commands in R	3.4	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/> tibble	Simple Data Frames	3.1.8	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/> tidyr	Tidy Messy Data	1.2.0	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> tidyselect	Select from a Set of Strings	1.1.2	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/> tidyverse	Easily Install and Load the 'Tidyverse'	1.3.2	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> tinytex	Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents	0.41	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> tzdb	Time Zone Database Information	0.3.0	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> utf8	Unicode Text Processing	1.2.2	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> uuid	Tools for Generating and Handling of UUIDs	1.1-0	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> vctrs	Vector Helpers	0.4.1	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> viridisLite	Colorblind-Friendly Color Maps (Lite Version)	0.4.1	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> vroom	Read and Write Rectangular Text Data Quickly	1.5.7	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> withr	Run Code 'With' Temporarily Modified Global State	2.5.0	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> xfun	Supporting Functions for Packages Maintained by 'Yihui Xie'	0.32	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> xml2	Parse XML	1.3.3	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/> yaml	Methods to Convert R Data to YAML and Back	2.3.5	<input type="checkbox"/>	<input type="checkbox"/>	

## Bringing a dataset into RStudio

```
download.file(url = "https://ndownloader.figshare.com/files/2292169", destfile = "data_raw/portal_data_joined.csv")
```

# Other ways to get data into RStudio

- **Manually:** You can manually create it using the `data.frame()` function in Base R, or the `tibble()` function in the tidyverse.
- **Import it from a file:** Below is a very incomplete list
  - Text: TXT (`readLines()` function)
  - Tabular data: CSV, TSV (`read.table()` function or `readr` package which contains `read_csv()`)
  - Excel: XLSX (`xlsx` package)
  - Google sheets: (`googlesheets` package)
  - Statistics program: SPSS, SAS (`haven` package)
  - Databases: MySQL (`RMySQL` package)

# Read the data into R

```
> read_csv("data_raw/portal_data_joined.csv")
```

```
Rows: 34786 Columns: 13
```

— Column specification —

```
Delimiter: ","
```

```
chr (6): species_id, sex, genus, species, taxa, plot_type
```

```
dbl (7): record_id, month, day, year, plot_id, hindfoot_length, weight
```

```
i Use `spec()` to retrieve the full column specification for this data.
```

```
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# A tibble: 34,786 × 13
```

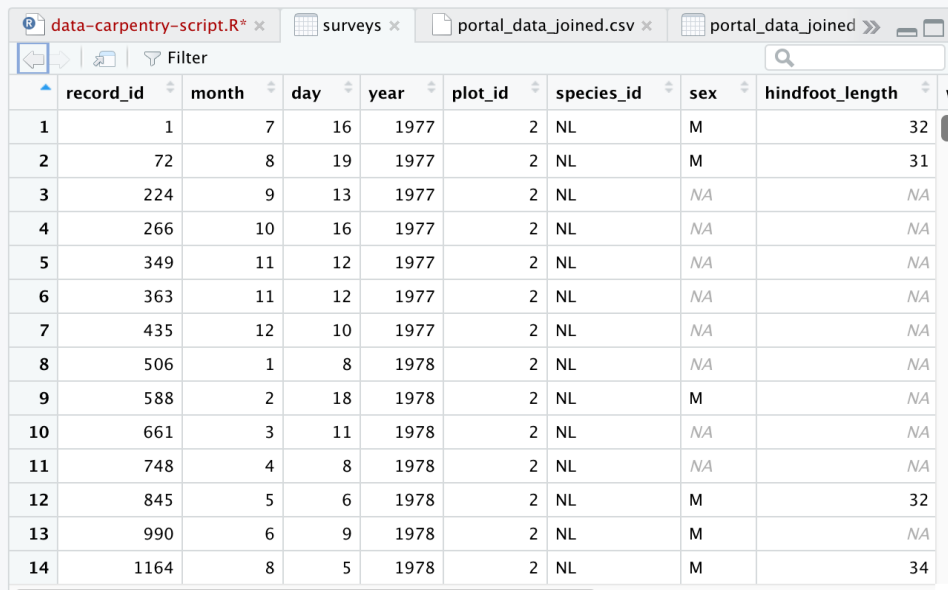
```
  record_id month   day  year plot_id specie...1 sex  hindf...2 weight genus species  
    <dbl> <dbl> <dbl> <dbl> <dbl> <chr>    <chr>    <dbl> <dbl> <chr> <chr>
```

```
read_csv("data_raw/portal_data_joined.csv")
```

# Look at the contents of the data file

```
head(surveys)
```

```
#> # A tibble: 6 × 13
#>   record_id month   day year plot_id species_id sex  hindfoot_length weight
#>   <dbl> <dbl> <dbl> <dbl> <dbl> <chr>      <chr>      <dbl> <dbl>
#> 1     1     7    16 1977     2 NL         M         32     NA
#> 2    72     8    19 1977     2 NL         M         31     NA
#> 3   224     9    13 1977     2 NL        <NA>     NA     NA
#> 4   266    10    16 1977     2 NL        <NA>     NA     NA
#> 5   349    11    12 1977     2 NL        <NA>     NA     NA
#> 6   363    11    12 1977     2 NL        <NA>     NA     NA
#> # ... with 4 more variables: genus <chr>, species <chr>, taxa <chr>,
#> #   plot_type <chr>
```



record_id	month	day	year	plot_id	species_id	sex	hindfoot_length	weight
1	7	16	1977	2	NL	M	32	NA
2	8	19	1977	2	NL	M	31	NA
3	9	13	1977	2	NL	NA	NA	NA
4	10	16	1977	2	NL	NA	NA	NA
5	11	12	1977	2	NL	NA	NA	NA
6	11	12	1977	2	NL	NA	NA	NA
7	12	10	1977	2	NL	NA	NA	NA
8	1	8	1978	2	NL	NA	NA	NA
9	2	18	1978	2	NL	M	NA	NA
10	3	11	1978	2	NL	NA	NA	NA
11	4	8	1978	2	NL	NA	NA	NA
12	5	6	1978	2	NL	M	32	NA
13	6	9	1978	2	NL	M	NA	NA
14	8	5	1978	2	NL	M	34	NA



# Links

Data Carpentries

<https://datacarpentry.org/R-ecology-lesson/00-before-we-start.html>

<https://datacarpentry.org/R-ecology-lesson/01-intro-to-r.html>

Texas Tech University R For Beginners

<https://guides.library.ttu.edu/R4Beginners>

University of Wyoming Introduction to R

[https://wyoscholar.uwyo.edu/articles/presentation/Introduction to R and RStudio/14589519?backTo=/collections/Digital\\_Scholarship\\_Center\\_Programming/5420781](https://wyoscholar.uwyo.edu/articles/presentation/Introduction%20to%20R%20and%20RStudio/14589519?backTo=/collections/Digital_Scholarship_Center_Programming/5420781)

[Meryl.Brodsky@Austin.utexas.edu](mailto:Meryl.Brodsky@Austin.utexas.edu)

[HChapmanTripp@Austin.utexas.edu](mailto:HChapmanTripp@Austin.utexas.edu)