GRADUATE STATISTICS PSYC 640

FALL 2024

CLASS INFO – WHEN THE TEACHING HAPPENS

- We will meet T & R from 11:00am to 12:15pm in this room (LA - 3233)
- My drop-in office hours are Tuesdays and Wednesdays from 12:30 – 2pm
- My office is in EAS-3378
- Best way to contact? Email me. I will usually respond within 24 hours. If it has been longer, then you can send me a follow-up

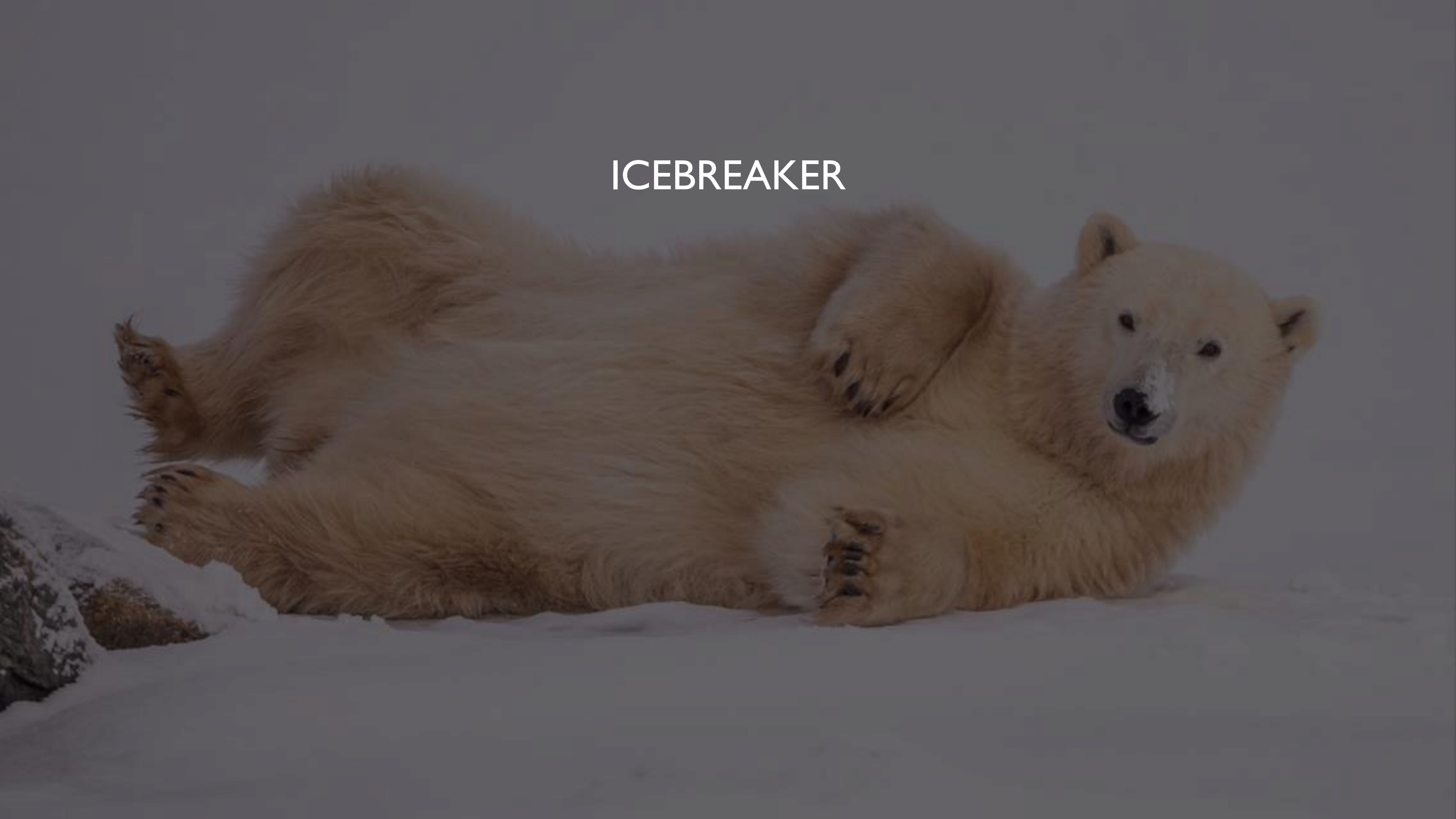
CLASS INFO – COMMUNICATION

- If you are having trouble with the course, please communicate them early
- Ask Questions!!
- Be honest with me. I will do what I can to support you

ABOUT ME

- Undergraduate degree at University of Denver
  - Psychology & Philosophy
  - Hated Statistics
- Graduate degree at University of Illinois at Urbana Champaign
  - Advisor Dr. Benjamin Hankin
- Internship & Postdoc at Rochester Institute of Technology

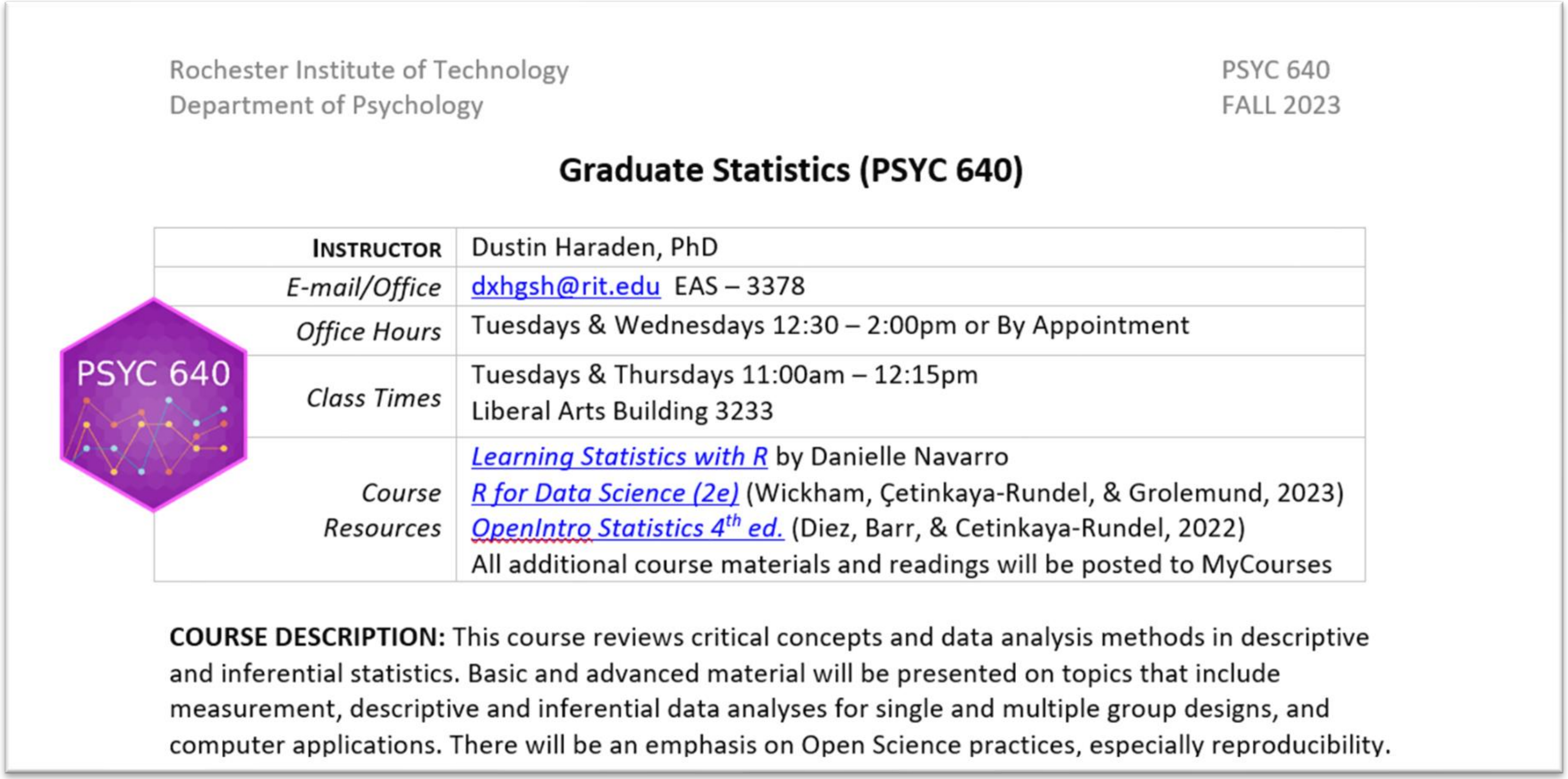
ICEBREAKER



TEACHING PHILOSOPHY

- I'm here for you and to support you however I can
- My job is to create a supportive, low-stakes environment where you can crash test ideas and develop new skills
- Your responsibility is to come to class prepared, ready to have fun, take some risks, and actively participate!
  - Ask questions, share relevant funny stories, seek clarification, question the status quo where applicable.
- Science is a team sport & failures/mistakes will happen, but it is okay! That's what science and discovery is!

SYLLABUS



Rochester Institute of Technology
Department of Psychology

PSYC 640
FALL 2023

Graduate Statistics (PSYC 640)

INSTRUCTOR	Dustin Haraden, PhD
E-mail/Office	dxhgsh@rit.edu EAS – 3378
Office Hours	Tuesdays & Wednesdays 12:30 – 2:00pm or By Appointment
Class Times	Tuesdays & Thursdays 11:00am – 12:15pm Liberal Arts Building 3233
Course Resources	*Learning Statistics with R* by Danielle Navarro *R for Data Science (2e)* (Wickham, Çetinkaya-Rundel, & Grolemund, 2023) *OpenIntro Statistics 4th ed.* (Diez, Barr, & Cetinkaya-Rundel, 2022) All additional course materials and readings will be posted to MyCourses

COURSE DESCRIPTION: This course reviews critical concepts and data analysis methods in descriptive and inferential statistics. Basic and advanced material will be presented on topics that include measurement, descriptive and inferential data analyses for single and multiple group designs, and computer applications. There will be an emphasis on Open Science practices, especially reproducibility.

COOL FANCY WEBSITE

TEXTBOOK & MATERIALS

- *Learning Statistics with R* by Danielle Navarro
- *R for Data Science (2e)* (Wickham, Çetinkaya-Rundel, & Grolemund, 2023)
- *OpenIntro Statistics 4th ed.* (Diez, Barr, & Cetinkaya-Rundel, 2022)
- R & R-Studio (will download later today)

EVALUATION/GRADING

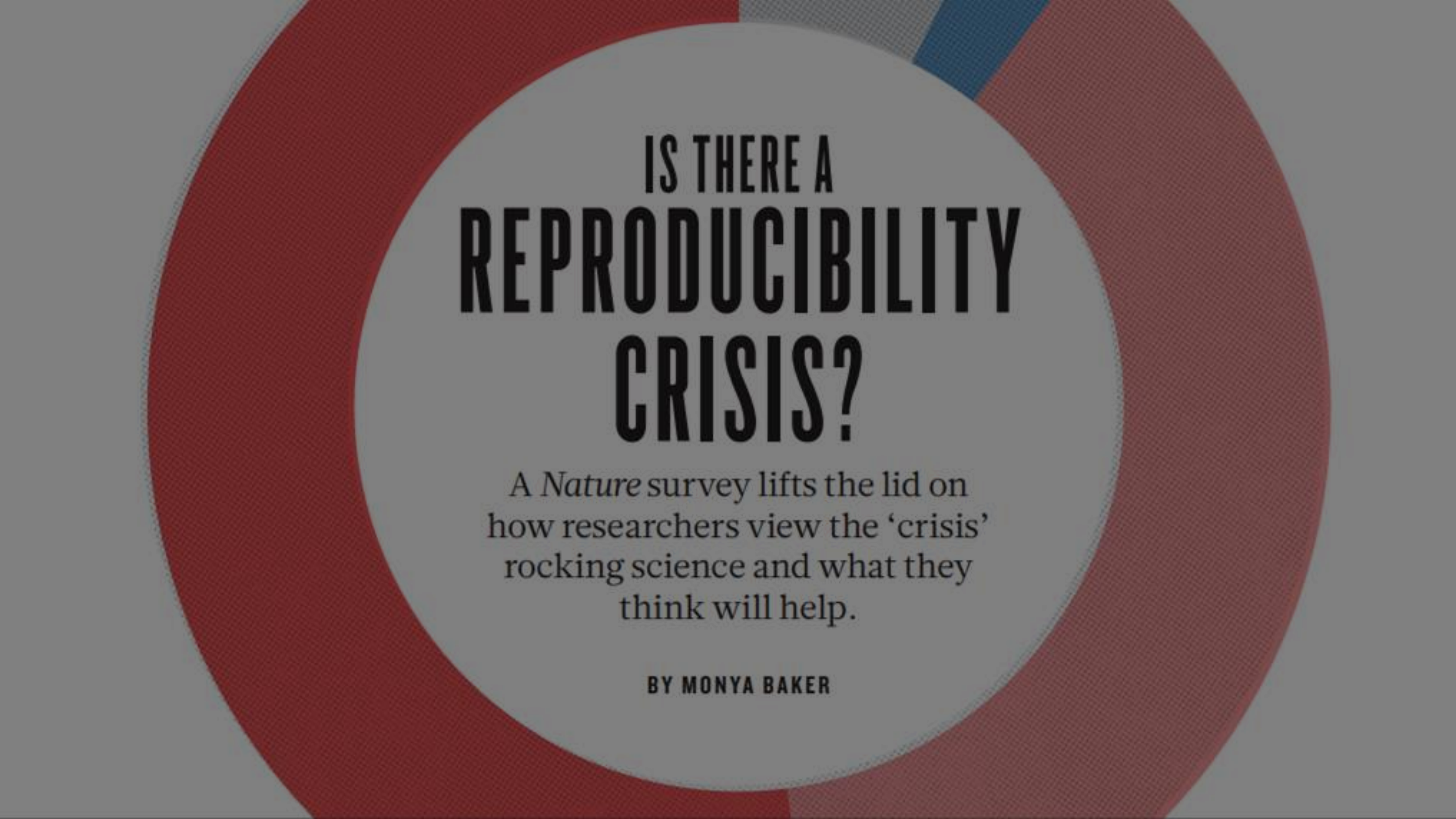
Component	Weight
Lab Assignments	35%
Journal Entries	15%
Presentations	25%
Oral Exam	25%

GENERATIVE AI

- Academic Integrity
- Generative AI and statistics/psychology
- Limitations & Bias

ChatGPT

QUESTIONS ABOUT THE COURSE?



IS THERE A REPRODUCIBILITY CRISIS?

A *Nature* survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

BY MONYA BAKER

Reproducibility vs. Replicability

Reproducibility:

- Taking all materials from a study and coming to the same conclusions

Replicability:

- The process of applying the same methodology with a different sample and research group

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

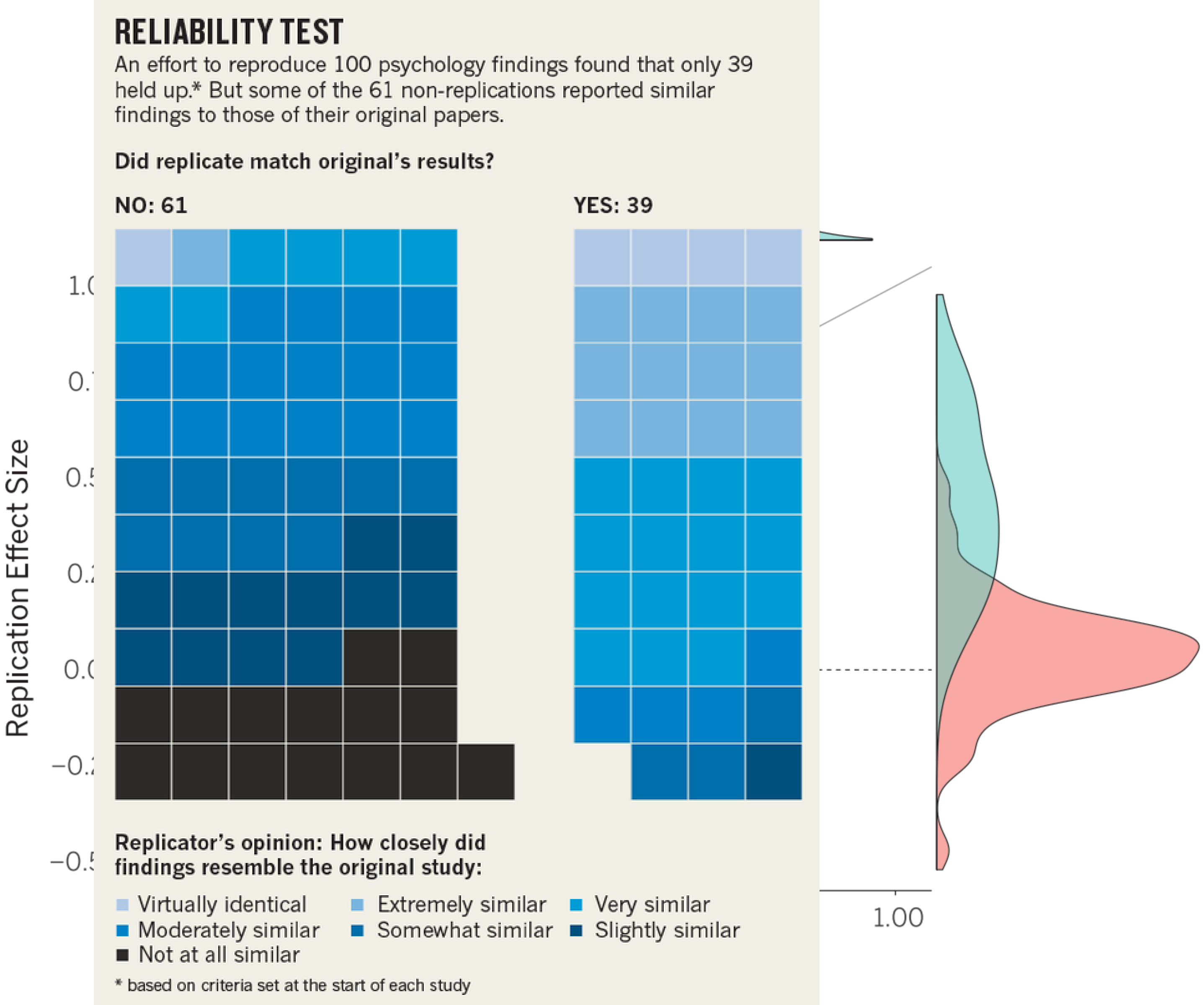
Open Science Collaboration*,†

+ See all authors and affiliations

100 Replication Studies
Adequately Powered

Original Studies:
Mean Effect: 0.403
% with p<.05: 97%

Replication Studies:
Mean Effect: 0.197
% with p<.05: 47%



RELIABILITY TEST

An effort to reproduce 100 psychology findings found that only 39 held up.* But some of the 61 non-replications reported similar findings to those of their original papers.

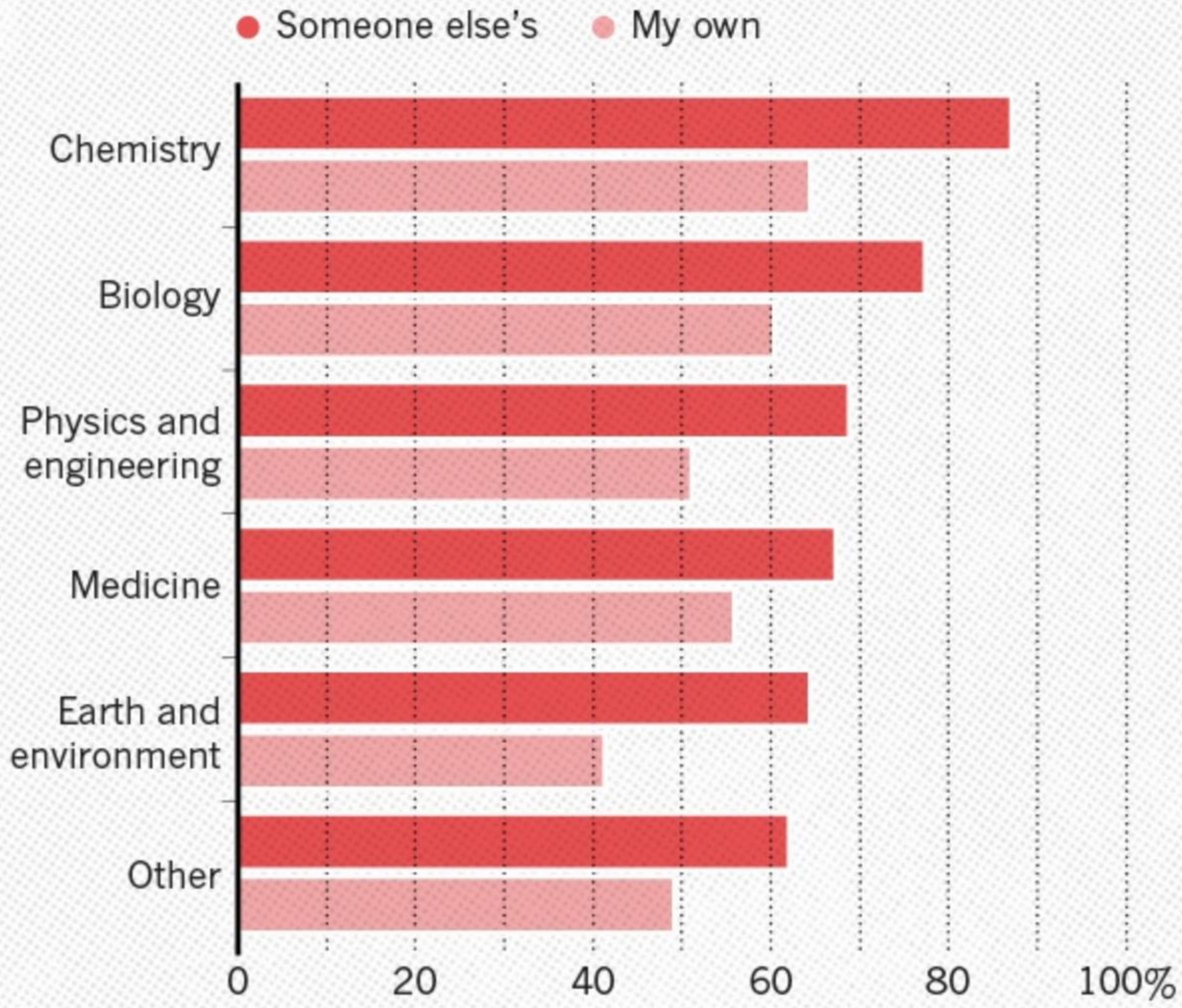
Did replicate match original's results?

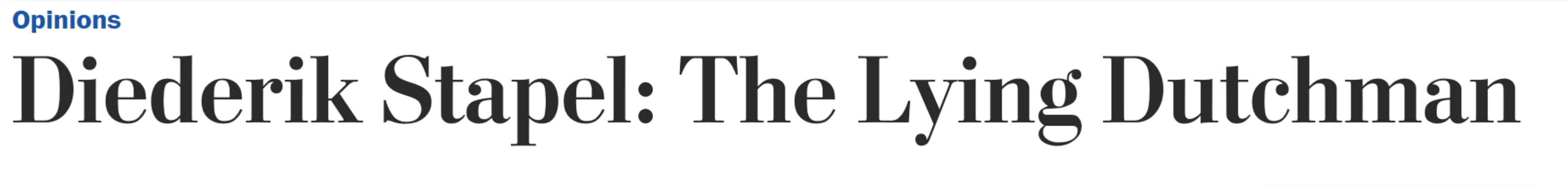
Replicator's opinion: How closely did findings resemble the original study:

* based on criteria set at the start of each study

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

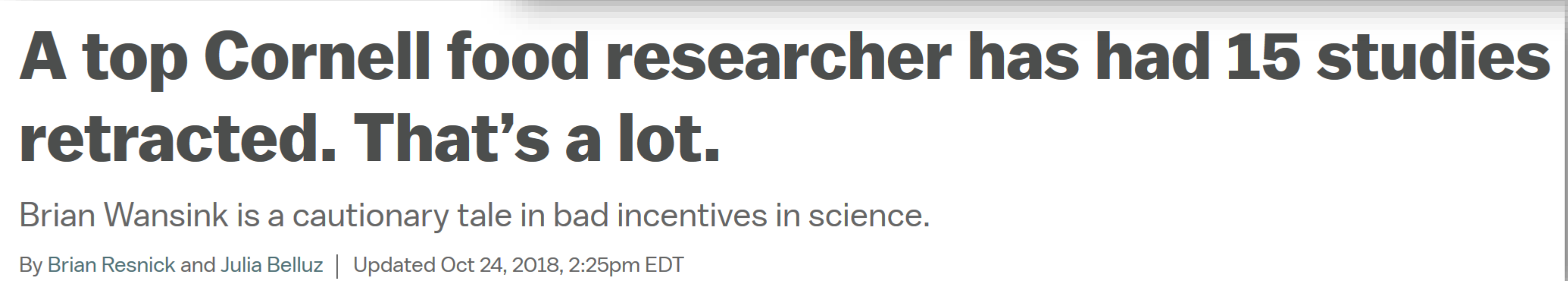
Most scientists have experienced failure to reproduce results.





Opinions

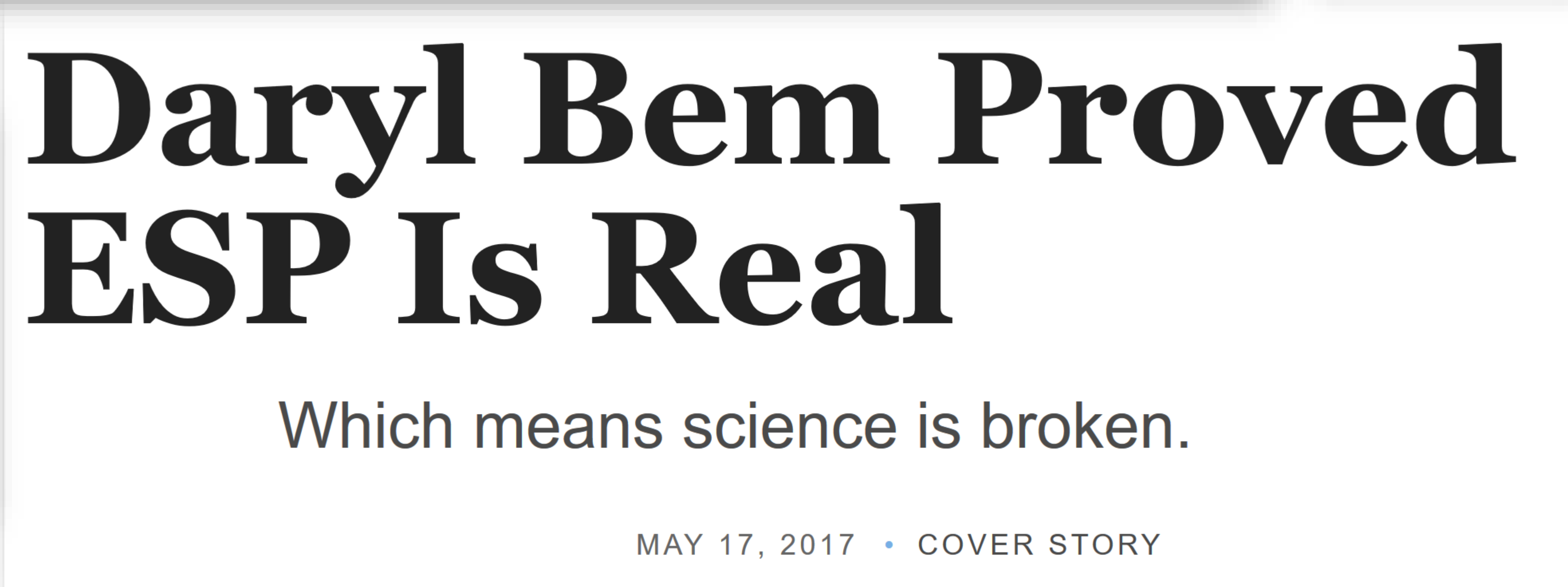
Diederik Stapel: The Lying Dutchman



A top Cornell food researcher has had 15 studies retracted. That’s a lot.

Brian Wansink is a cautionary tale in bad incentives in science.

By Brian Resnick and Julia Belluz | Updated Oct 24, 2018, 2:25pm EDT



Daryl Bem Proved ESP Is Real

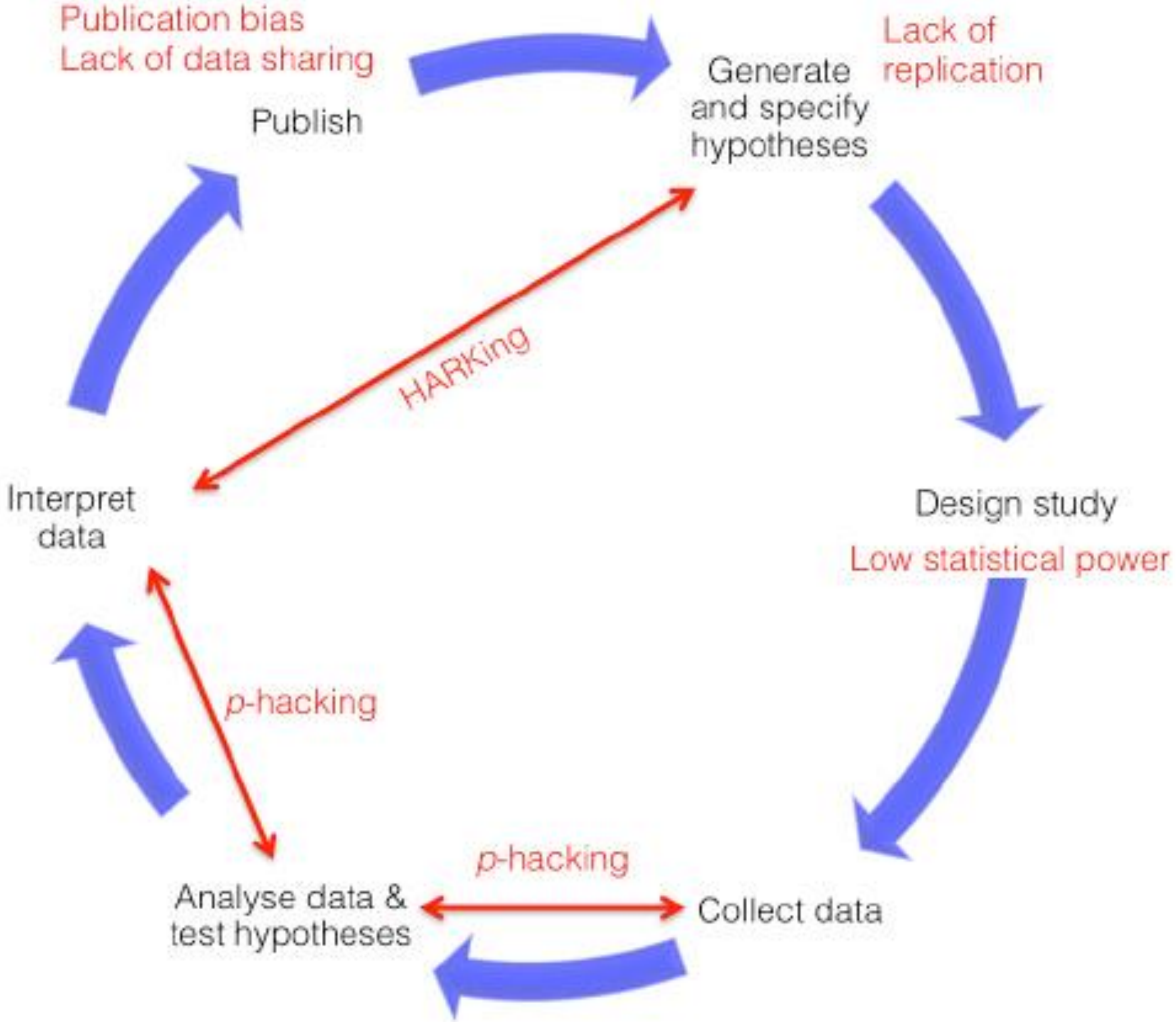
Which means science is broken.

MAY 17, 2017 • COVER STORY

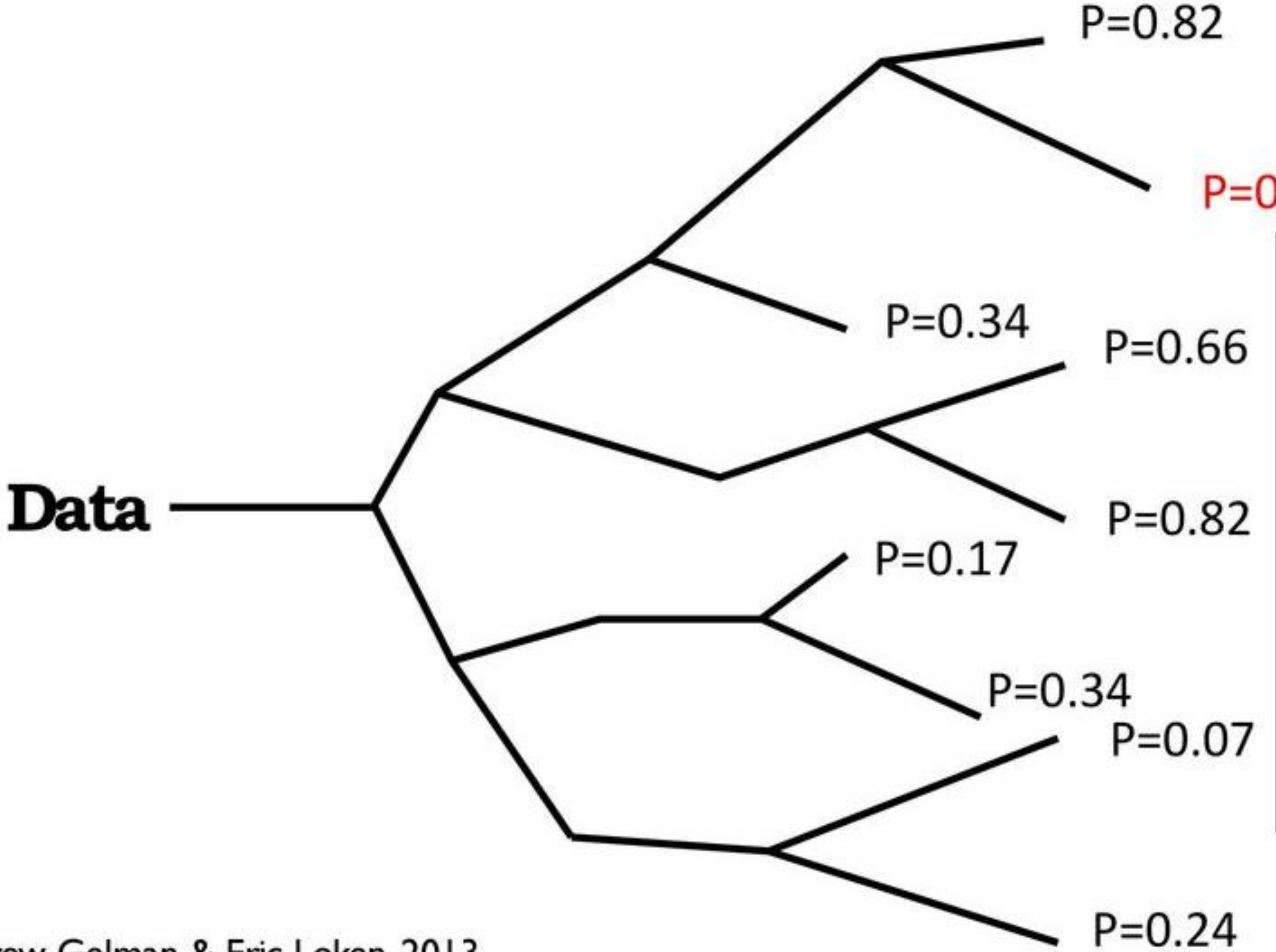
Questionable Research Practices

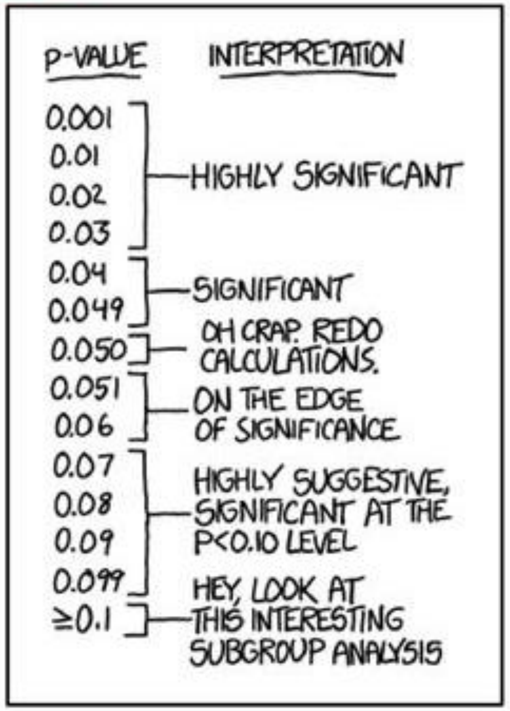
- Selective reporting of dependent variables
- Deciding whether to collect more data after looking to see whether the results will be significant
- Failing to disclose experimental conditions
- In a paper reporting selectively studies that worked
- Reporting an unexpected finding as having been predicted from the start
- Falsifying Data

Scientific Process & QRP’s



Garden of Forking Paths





P-VALUE	INTERPRETATION
0.001	HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04	SIGNIFICANT
0.049	
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	
0.07	HIGHLY SUGGESTIVE, SIGNIFICANT AT THE P<0.10 LEVEL
0.08	
0.09	
0.099	
≥0.1	HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS

Andrew Gelman & Eric Loken, 2013

TOOLS TO PROMOTE TRANSPARENCY

INSTALLING R & R-STUDIO

LINK TO INSTALL GUIDE

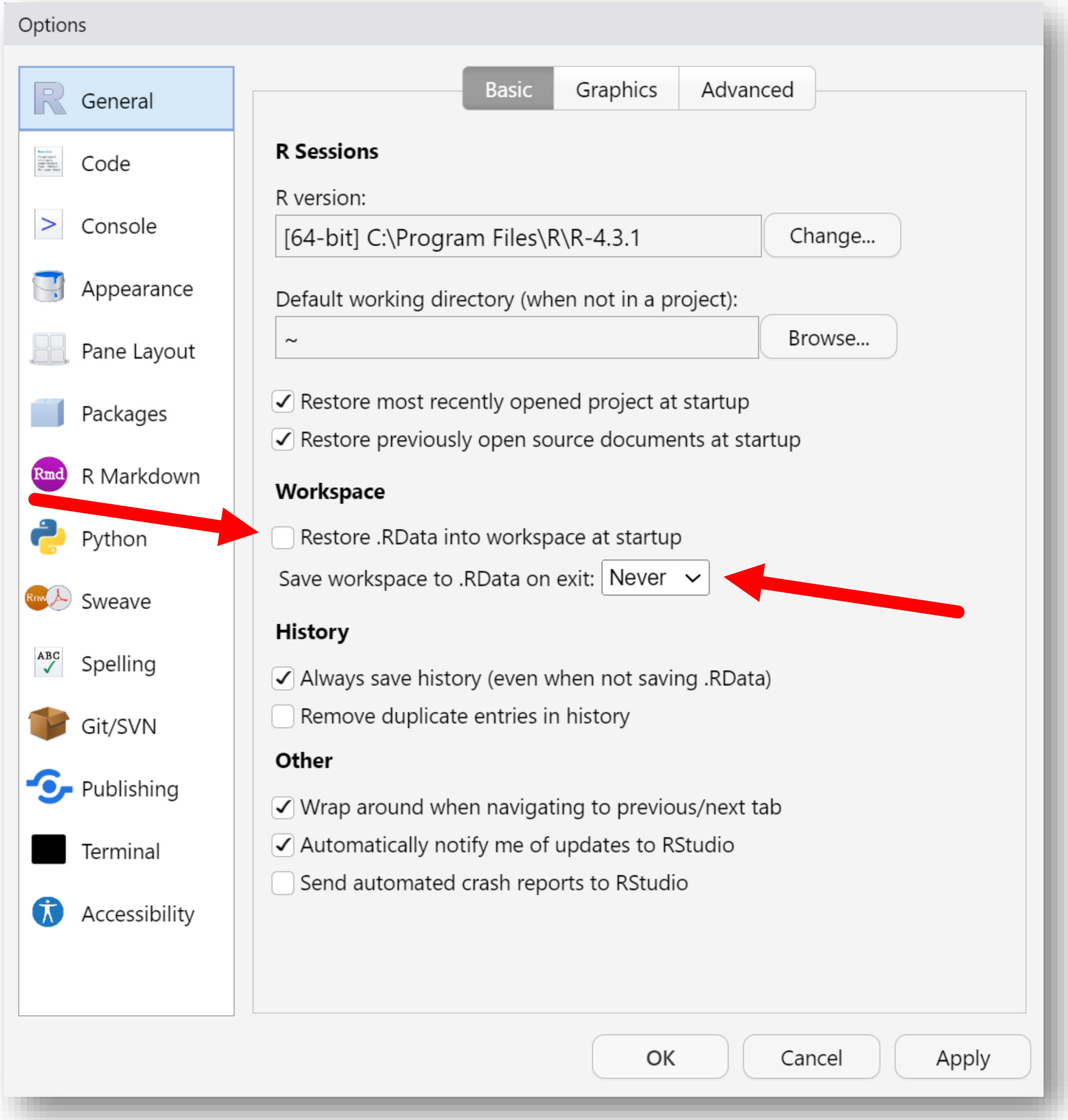
GETTING STARTED

- Start with a Blank Slate
- Make the code work for you
- This is your space – Redecorate!

START WITH A BLANK SLATE

- Navigate to ***Tools > Global Options***

- By default, R Studio saves all of the objects in your environment. In general, this is not ideal, because it means that you may have taken steps interactively that are not documented in your code.



R Sessions

R version:

[64-bit] C:\Program Files\R\R-4.3.1

Default working directory (when not in a project):

~

☑ Restore most recently opened project at startup

☑ Restore previously open source documents at startup

Workspace

□ Restore .RData into workspace at startup

Save workspace to .RData on exit: Never

History

☑ Always save history (even when not saving .RData)

□ Remove duplicate entries in history

Other

☑ Wrap around when navigating to previous/next tab

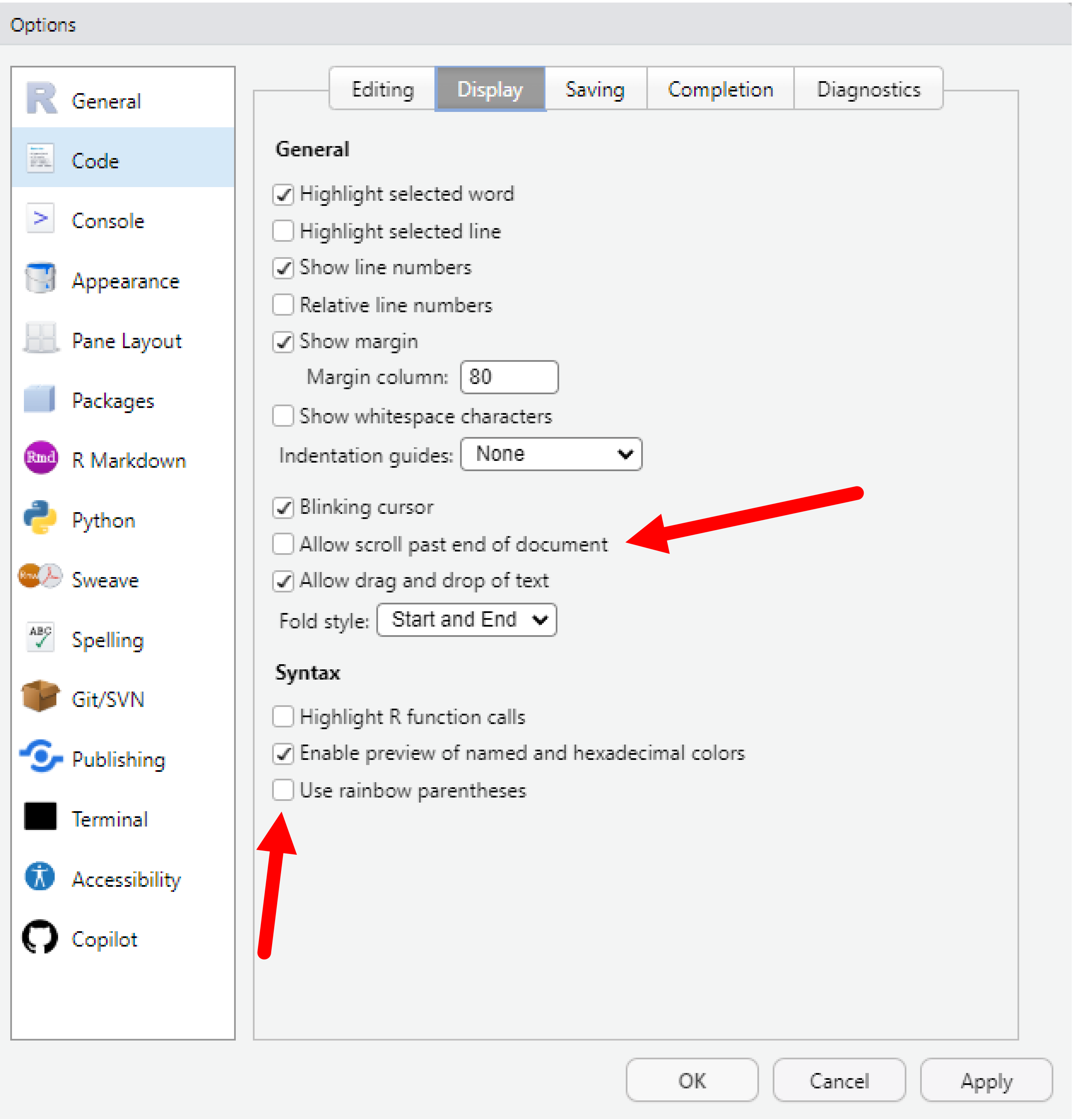
☑ Automatically notify me of updates to RStudio

□ Send automated crash reports to RStudio

CODE UPDATES

- Navigate to ***Tools > Global Options > Code > Display***

- Make sure these options are checked:
  - “Allow scroll past end of document”
  - “Use rainbow parentheses”



General

☑ Highlight selected word

☐ Highlight selected line

☑ Show line numbers

☐ Relative line numbers

☑ Show margin

Margin column: 80

☐ Show whitespace characters

Indentation guides: None

☑ Blinking cursor

☐ Allow scroll past end of document

☑ Allow drag and drop of text

Fold style: Start and End

Syntax

☐ Highlight R function calls

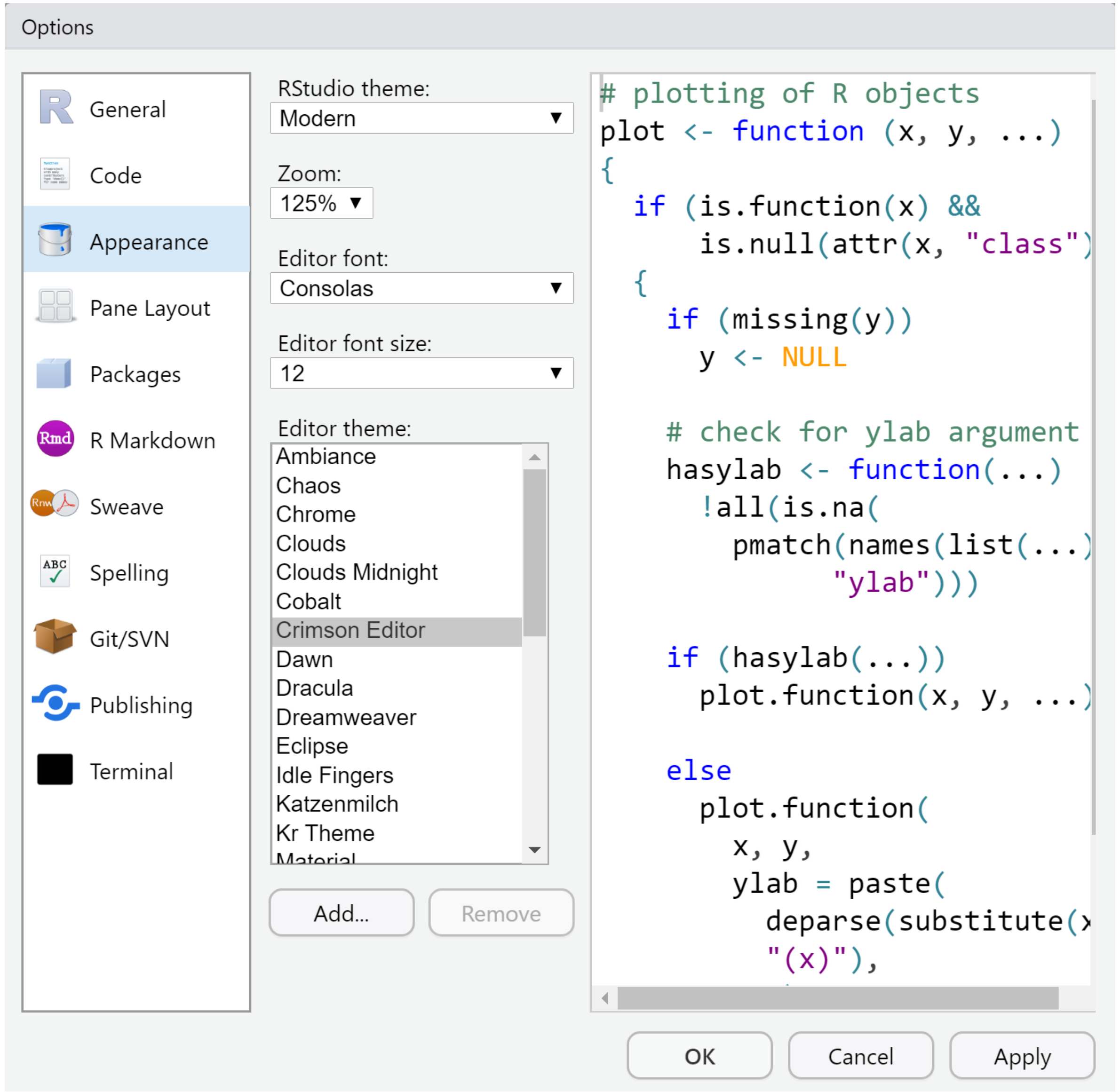
☑ Enable preview of named and hexadecimal colors

☐ Use rainbow parentheses

MAKE IT PRETTY

Tools > Global Options > Appearance

Update the appearance! This is your program and it should look how you want it to



RStudio theme: Modern

Zoom: 125%

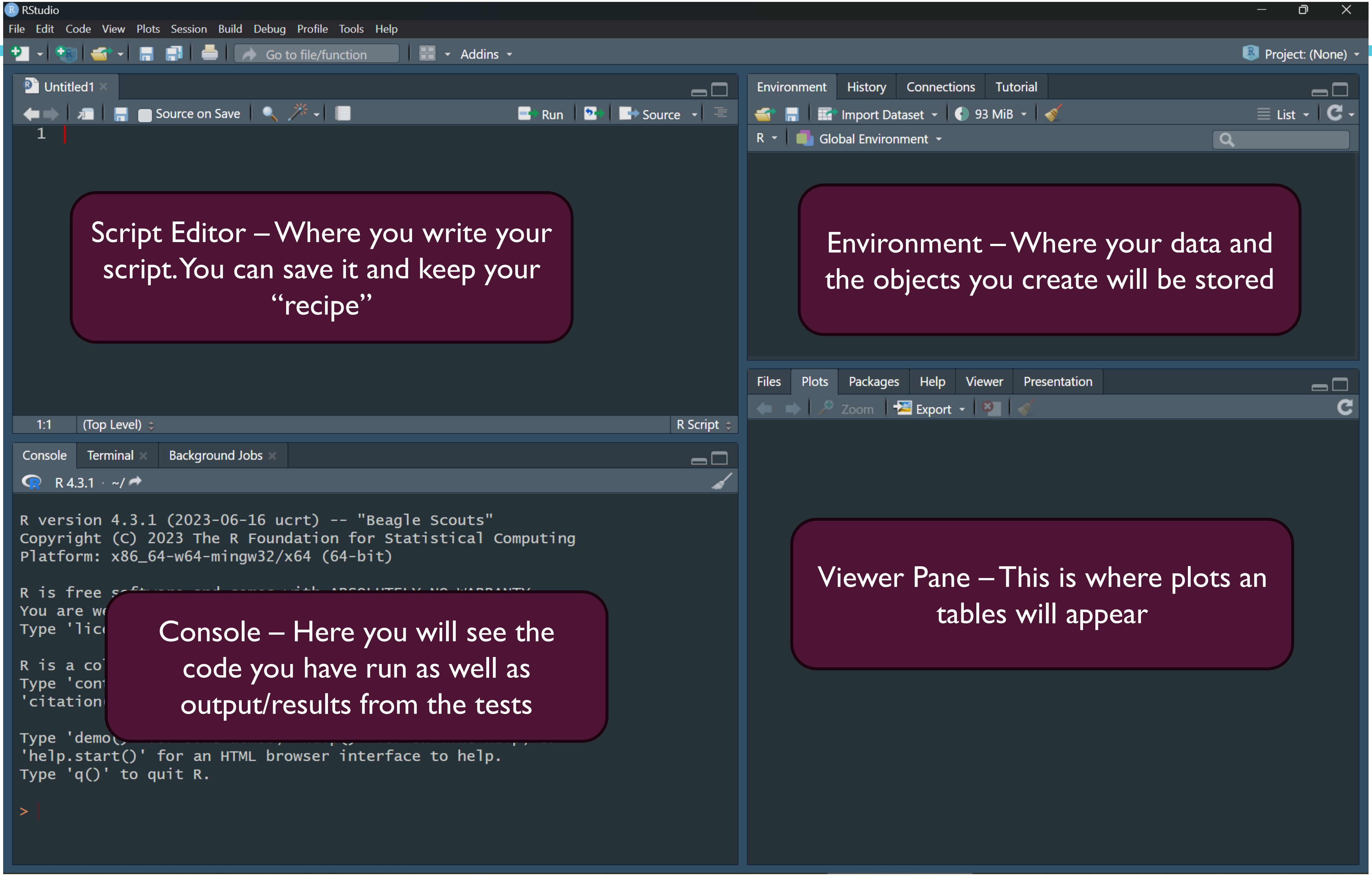
Editor font: Consolas

Editor font size: 12

Editor theme:

Ambiance
Chaos
Chrome
Clouds
Clouds Midnight
Cobalt
Crimson Editor
Dawn
Dracula
Dreamweaver
Eclipse
Idle Fingers
Katzenmilch
Kr Theme
Material

```
# plotting of R objects
plot <- function (x, y, ...)
{
  if (is.function(x) &&
      is.null(attr(x, "class")
  {
    if (missing(y))
      y <- NULL

    # check for ylab argument
    hasylab <- function(...)
      !all(is.na(
        pmatch(names(list(...)
               "ylab")))

    if (hasylab(...))
      plot.function(x, y, ...)

    else
      plot.function(
        x, y,
        ylab = paste(
          deparse(substitute(x
          "(x)"),
```



RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
1:1
(Top Level)
R Script
Environment History Connections Tutorial
Import Dataset
93 MiB
List
R
Global Environment
Files Plots Packages Help Viewer Presentation
Zoom
Export
Console Terminal Background Jobs
R 4.3.1 · ~/

```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are we
Type 'lic

R is a co
Type 'con
'citation

Type 'demo
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

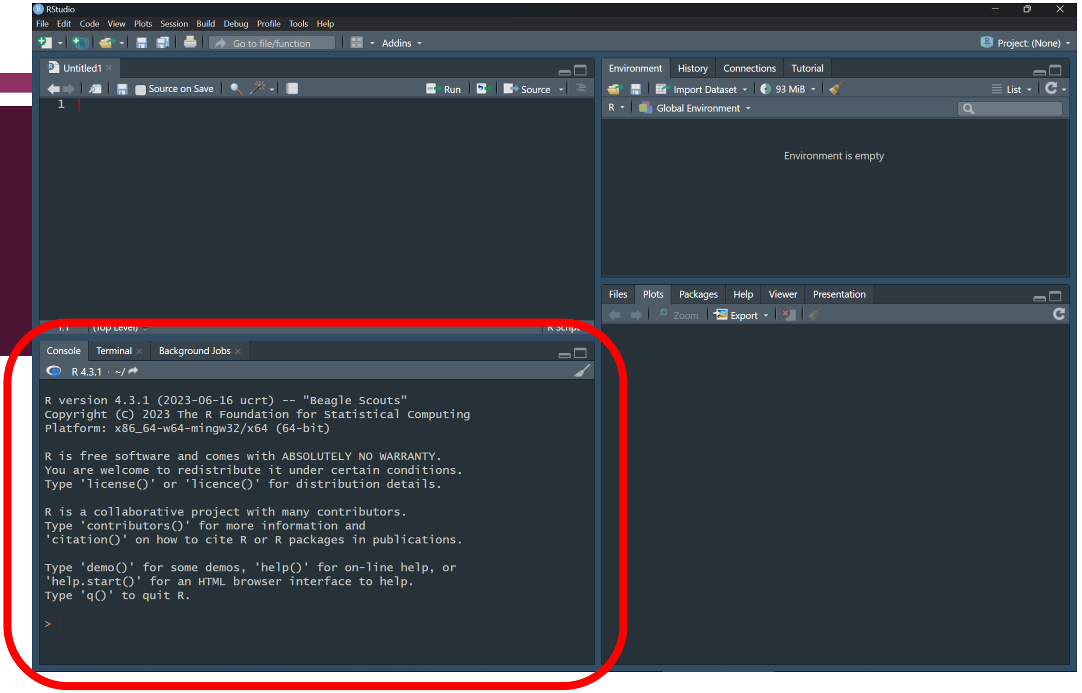
Script Editor – Where you write your script. You can save it and keep your "recipe"

Environment – Where your data and the objects you create will be stored

Console – Here you will see the code you have run as well as output/results from the tests

Viewer Pane – This is where plots an tables will appear

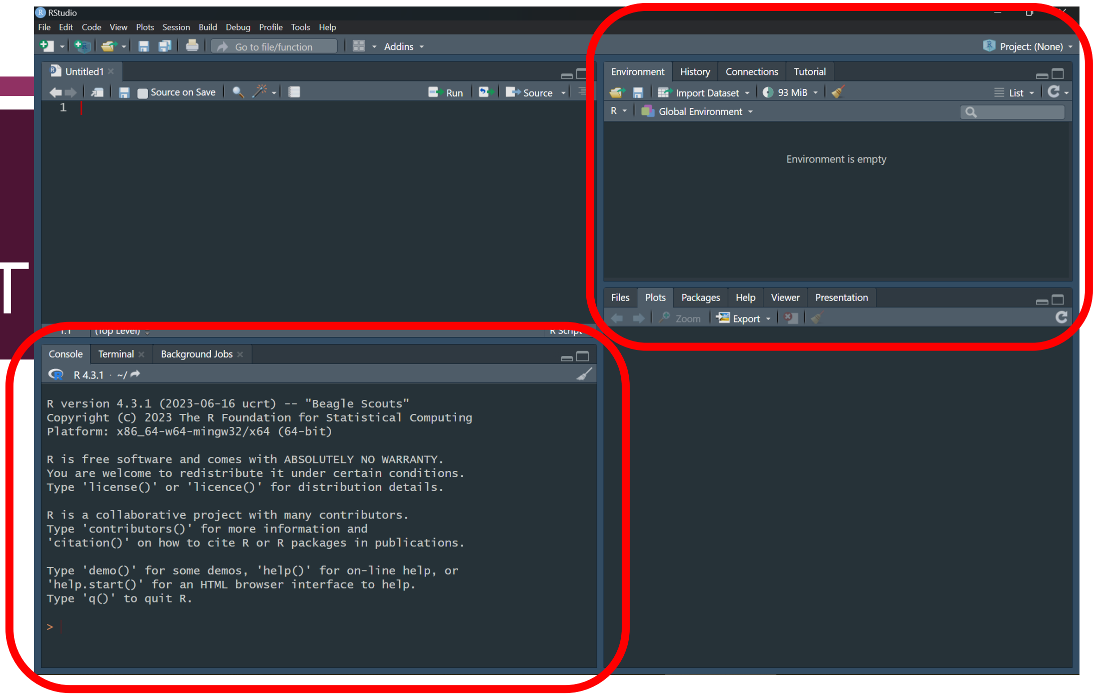
RSTUDIO - CONSOLE



```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (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.

>
```

- You can type directly into this
- It will not be saved
- Simplest thing is to use it as a calculator! Go ahead and try it out
  - Type 42 + 13 and hit ENTER ← This is a ***command***
  - You are commanding the computer to perform the task that you want
  - Note: Take a look at 3.2 in Learning Statistics with R
    - Another resource is R for Data Science Chapter 3

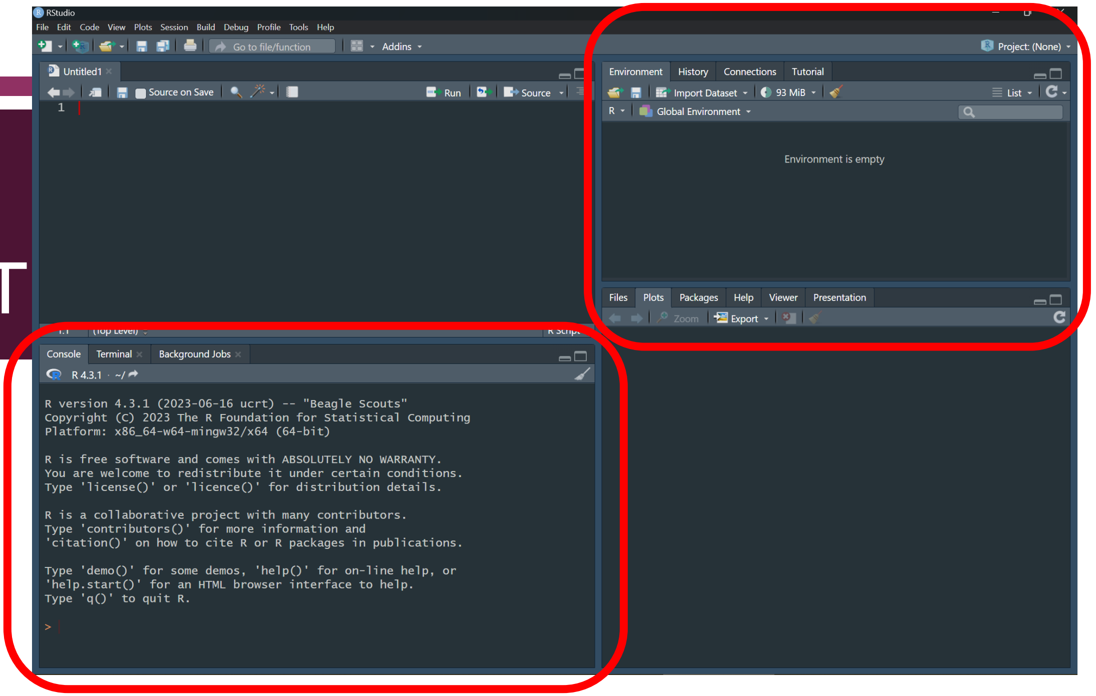
RSTUDIO – CONSOLE & ENVIRONMENT



```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (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.

>
```

- Creating Objects
  - "Objects" are containers for information
  - They are created by using the assignment operator ←
    - Shortcut for inserting the operator: Alt + - (minus sign)
  - Try it out! Create an object named "class" and assign it the number of students (16)
  - Once it is run, you will not see it in the console, but it will appear in the Environment
- Manipulating Objects
  - You can use object names in calculations
  - What happens if 4 students leave the class, how many are left?
  - Call the object again…did it change?
    - Updating objects

RSTUDIO – CONSOLE & ENVIRONMENT

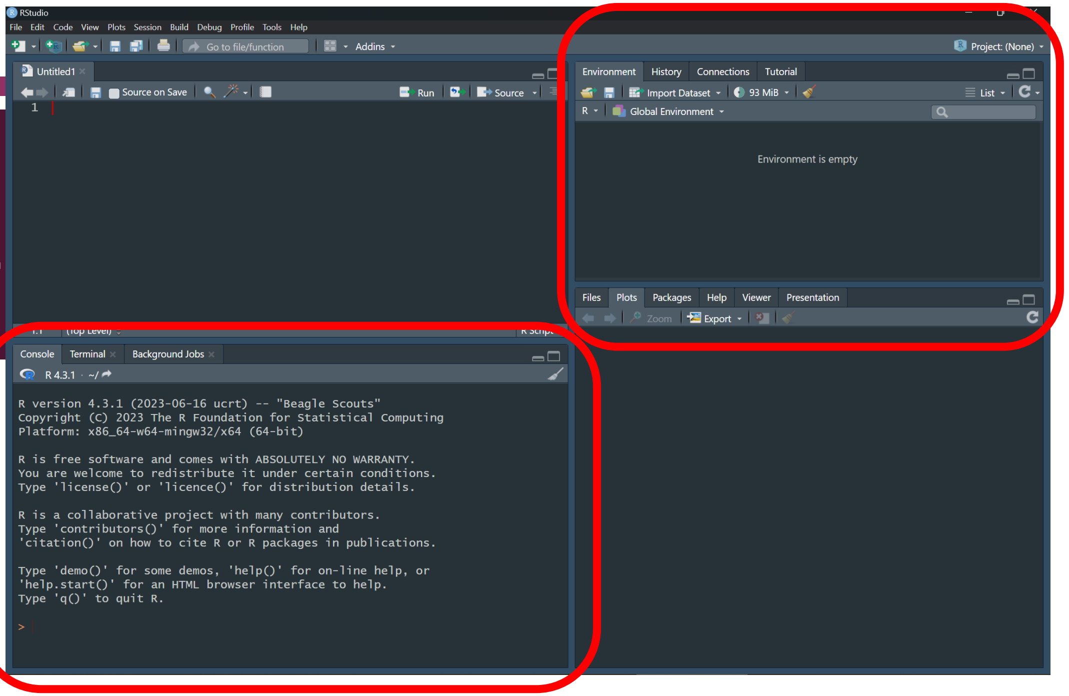


Environment is empty

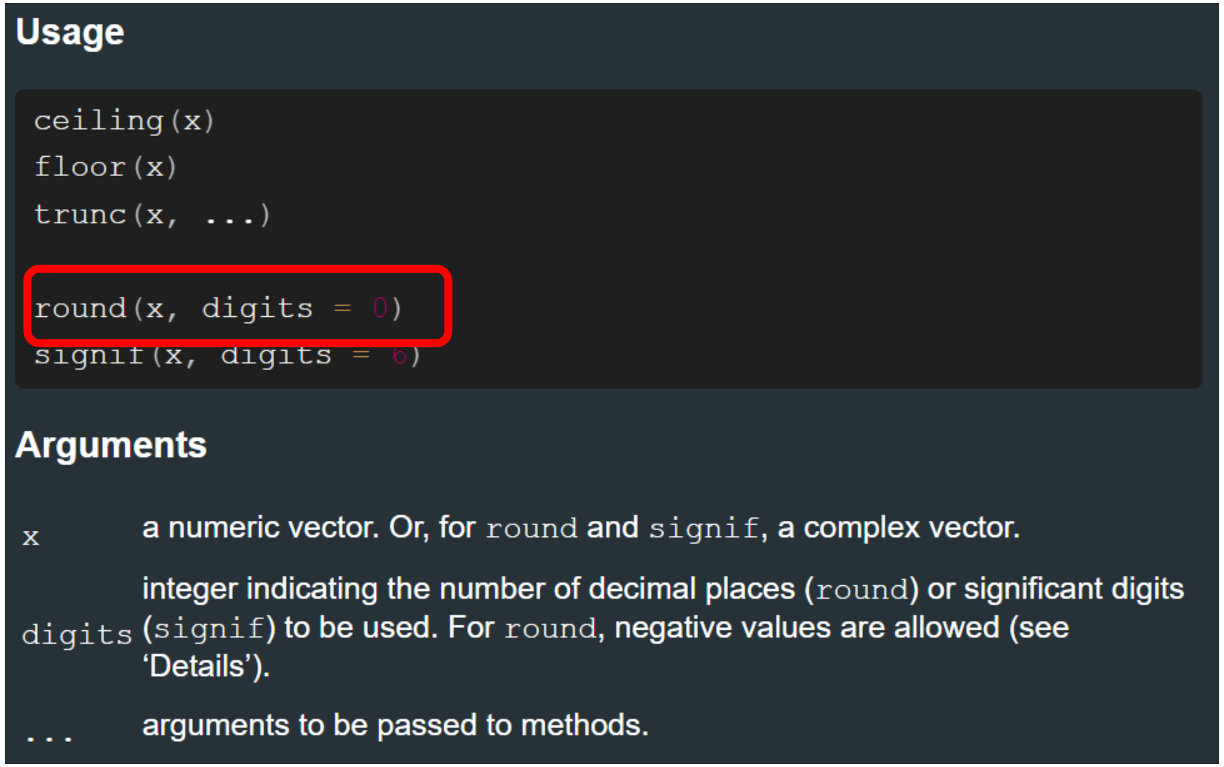
```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

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.

>
```

- Using Functions
  - Create object `class_grade` and assign it a list of numbers (separated by commas) using the combine function `c()`
  - Use a function to find the average of the grades using a function (Use Google)
  - Calculate the average manually (with R) to see if they match!
  - Create an object `class_average` that has the class average that you calculated
    - Maybe you don’t want that many decimal places…

RSTUDIO – CONSOLE & ENVIRONMENT



```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (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.

>
```



Usage

```
ceiling(x)
floor(x)
trunc(x, ...)

round(x, digits = 0)
signif(x, digits = 6)
```

Arguments

`x`	a numeric vector. Or, for `round` and `signif`, a complex vector.
`digits`	integer indicating the number of decimal places (`round`) or significant digits (`signif`) to be used. For `round`, negative values are allowed (see 'Details').
`...`	arguments to be passed to methods.

- Using Functions
  - The `round()` function (`?round`)
  - Function name & corresponding arguments

CLOSING RSTUDIO

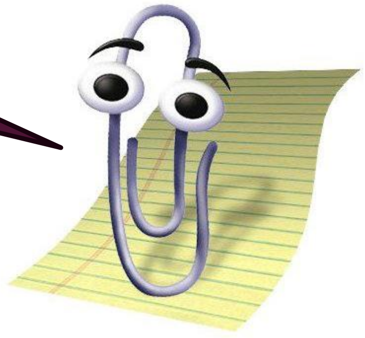
- What will happen with the objects we created once we close and re-open Rstudio?
- What about all the other work that you did?

CLOSING RSTUDIO

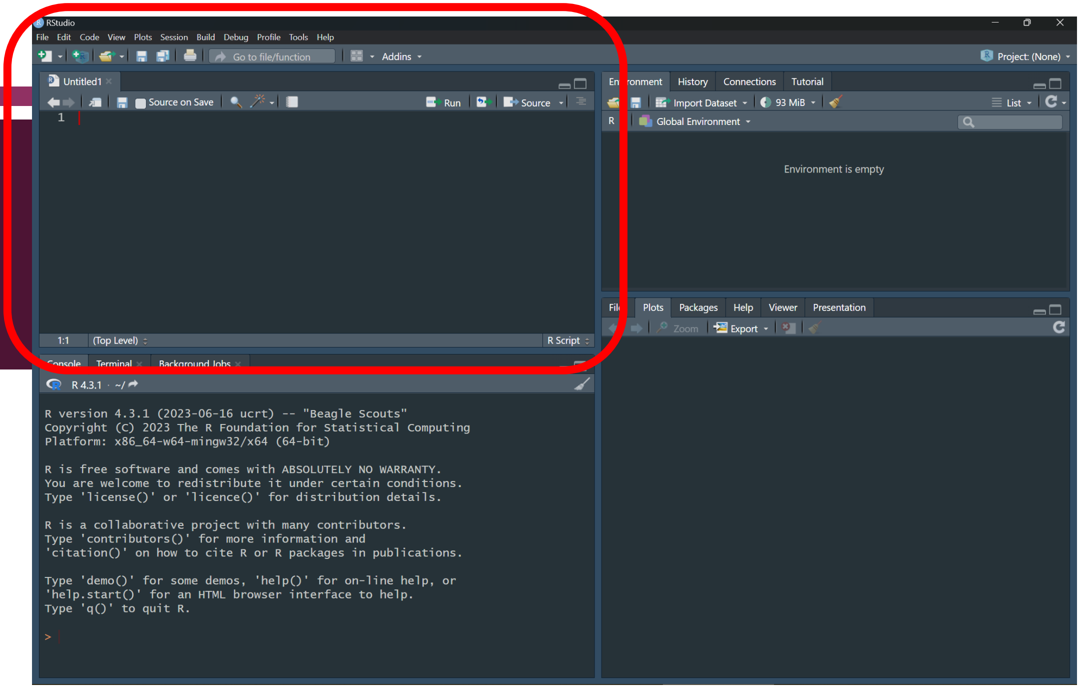
- What will happen with the obje nd re-open Rstudio?
- What about all the other work

It looks like you are trying to use Rstudio but all of your information was deleted. Maybe you shouldn't just use the console. Would you like some help?

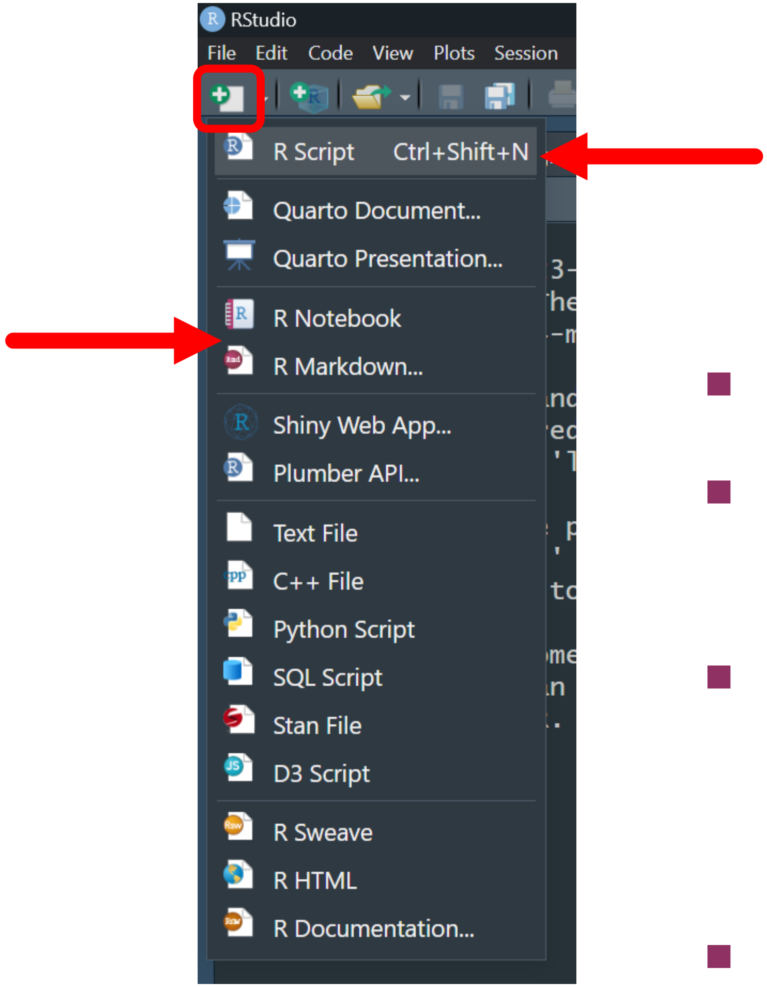
Yes No



RSTUDIO – SCRIPTS & MARKDOWN / NOTEBOOK

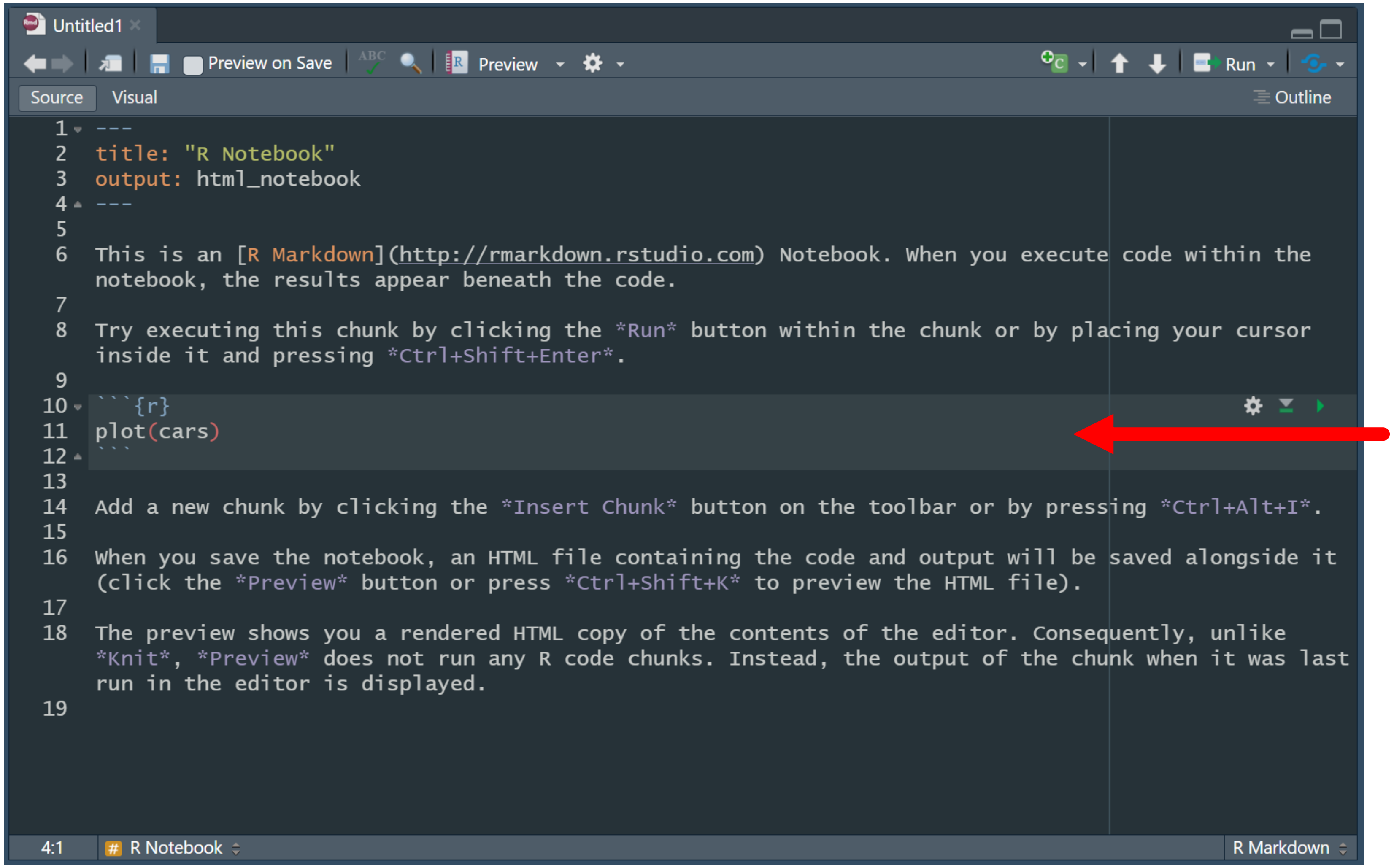


```
R version 4.3.1 (2023-06-16 ucrt) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

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.

>
```



- Scripts & Markdown/Notebooks are like the recipes
- These are the documents that you will be saving in order to continue your projects
- Create one script and one notebook
  - When creating a notebook, you may be prompted to update "a version of the markdown package" – Select Yes
- What are some differences that you notice between the two?

RSTUDIO - NOTEBOOK



~~~
---
title: "R Notebook"
output: html_notebook
---

This is an [R Markdown](http://rmarkdown.rstudio.com) Notebook. 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 *Ctrl+Shift+Enter*.

```{r}
plot(cars)
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

~~~

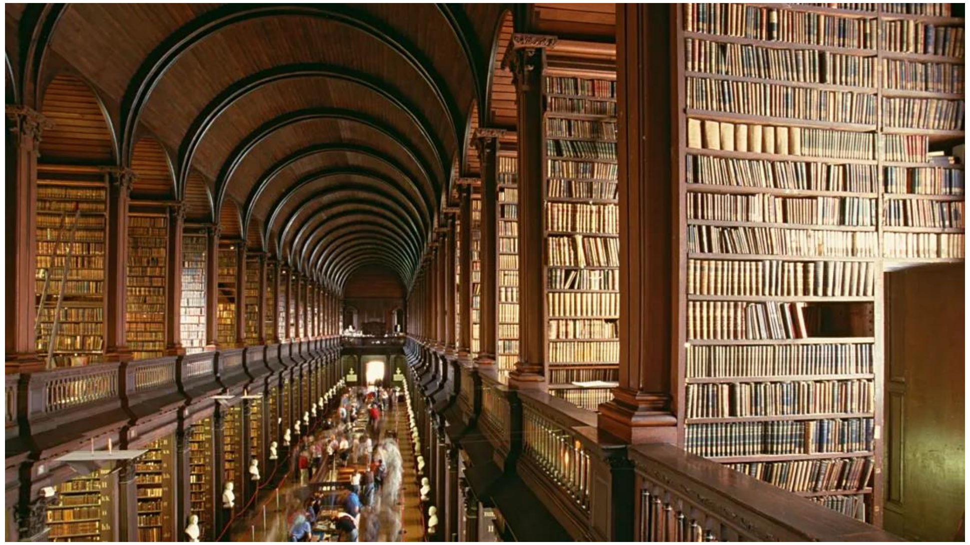
Normal Text

Code chunk

THE LIBRARY OF R

"A package is a like a book, a library is like a library; you use library() to check a package out of the library"

Hadley Wickham, Chief Scientist, R Studio



THE LIBRARY OF R - INSTALL.PACKAGES()

"Packages" are shareable collections of R code that provide functions (i.e., a command to perform a specific task), data, and documentation. Packages increase the functionality of R by improving and expanding on base R (basic R functions).

Installing and Loading Packages

To download a package, you must call install.packages():

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

Put this in the console

You can also navigate to the Packages pane, and then click "Install", which will work the same as the line of code above. This is a way to install a package using code or part of the R Studio interface.

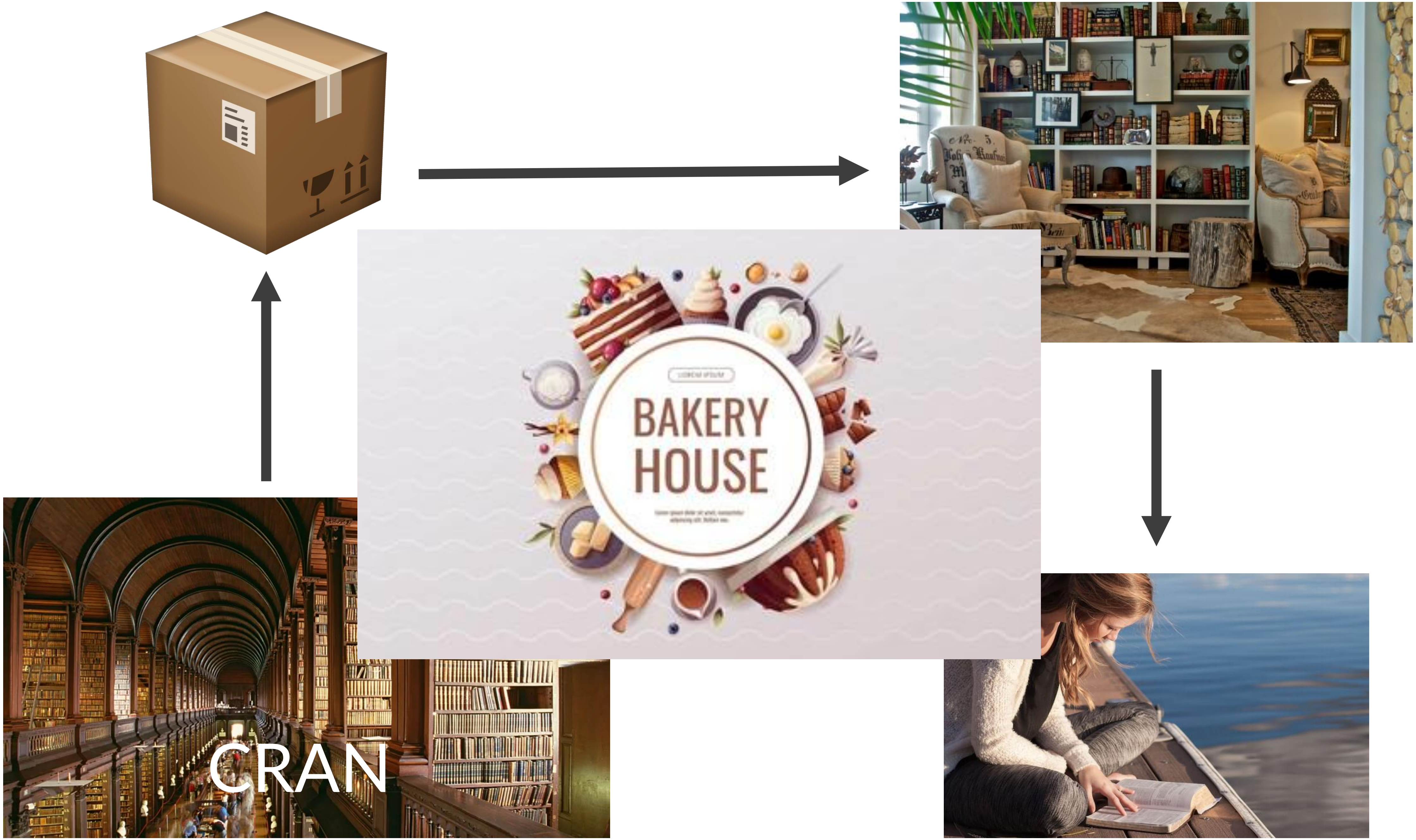
Usually, writing code is a bit quicker, but using the interface can be very useful and complementary to use of code.

THE LIBRARY OF R - LIBRARY()

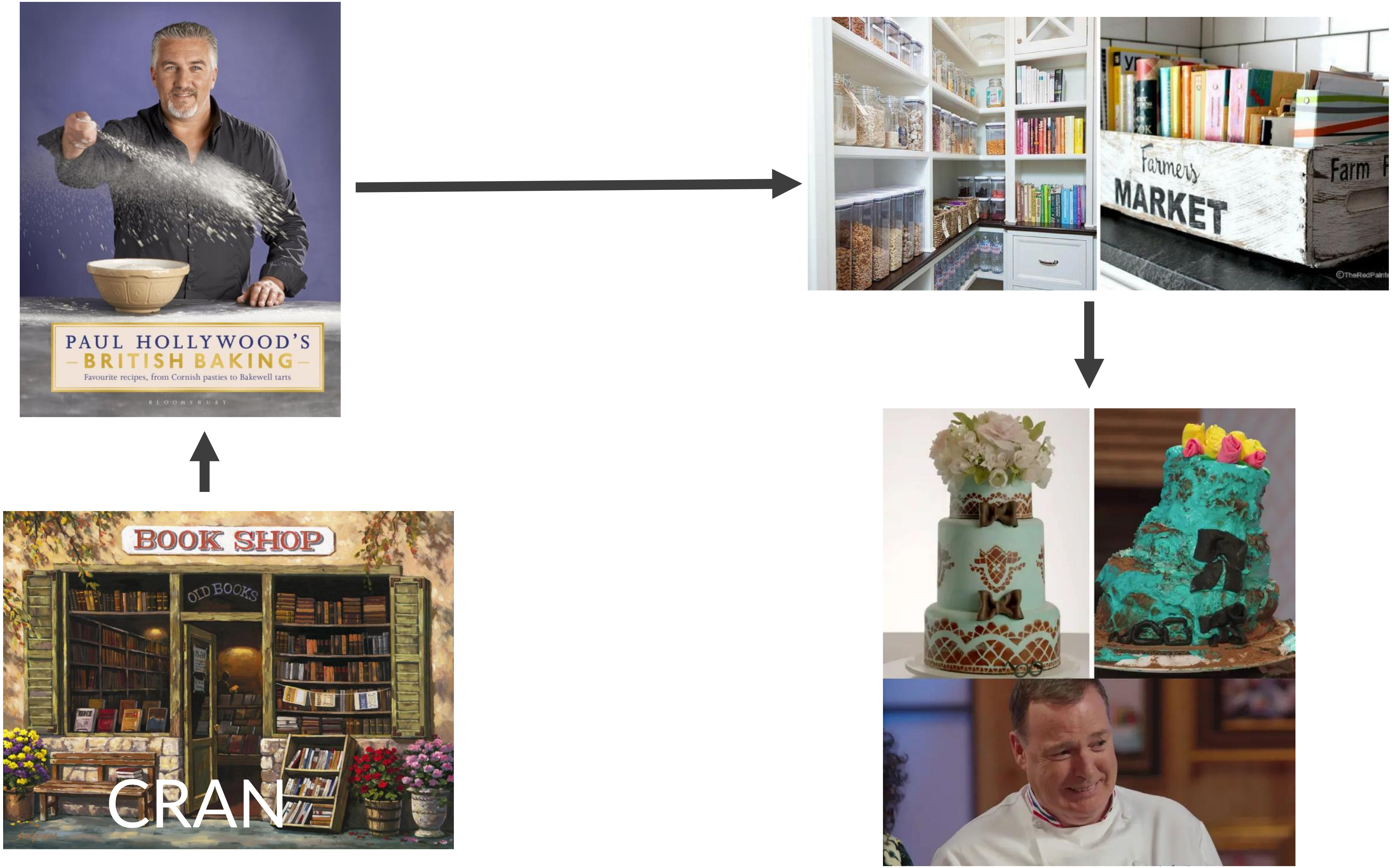
After the package is installed, it must be loaded into your R Studio session using library():

```
library(tidyverse)
```

We only have to *install* a package once, but to use it, we have to load it each time we start a new R session.

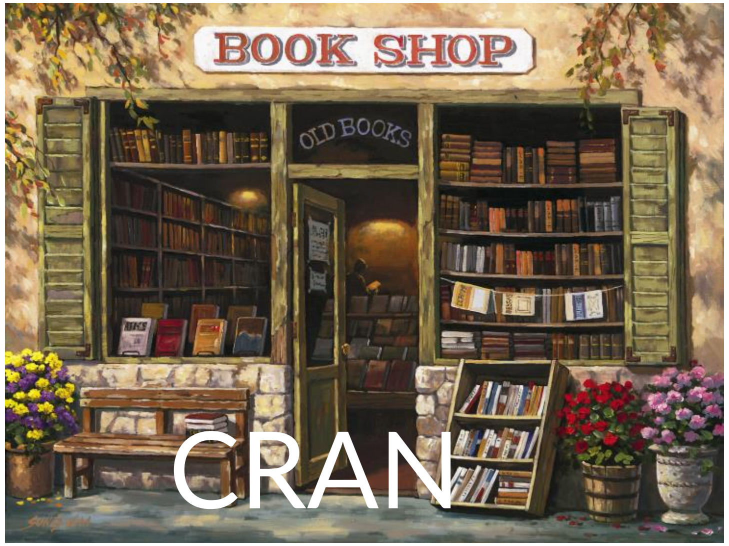

BAKERY
HOUSE
CRAN

THE GREAT BRITISH
BAKING SHOW

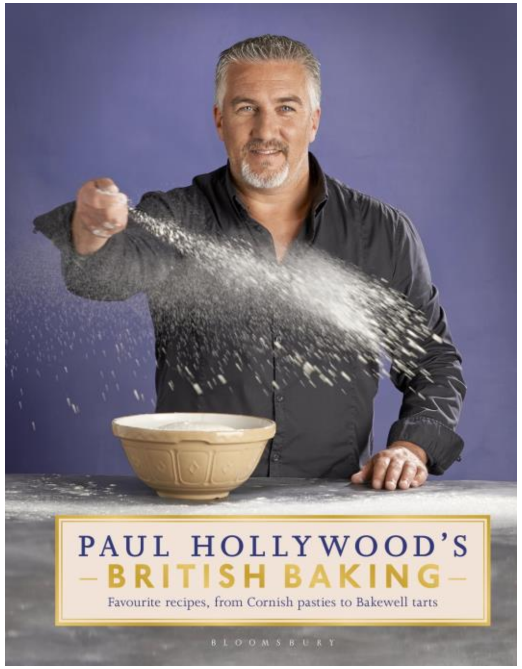


PAUL HOLLYWOOD'S
BRITISH BAKING
Favourite recipes, from Cornish pasties to Bakewell tarts
BLOOMSBURY
Farmers
MARKET
BOOK SHOP
OLD BOOKS

CRAN



```
install.packages(“tidyverse”)
```



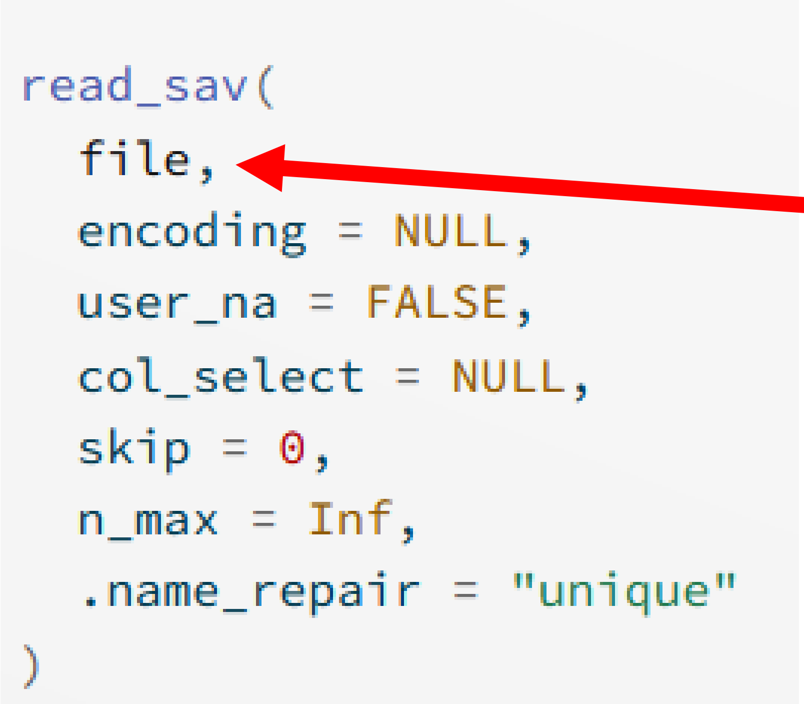
```
library(here)
```

PUTTING IT ALL TOGETHER

PUTTING IT TOGETHER - OVERVIEW

- Step 1: Create File Structure
  - One folder for the current project
  - Download the data to put in there
- Step 2: Create Notebook (helpful resource for further reading)
  - Install packages – `here()`
  - Put libraries in the code chunk
- Step 3: Load in Data
  - Using haven or rio package (https://haven.tidyverse.org/)

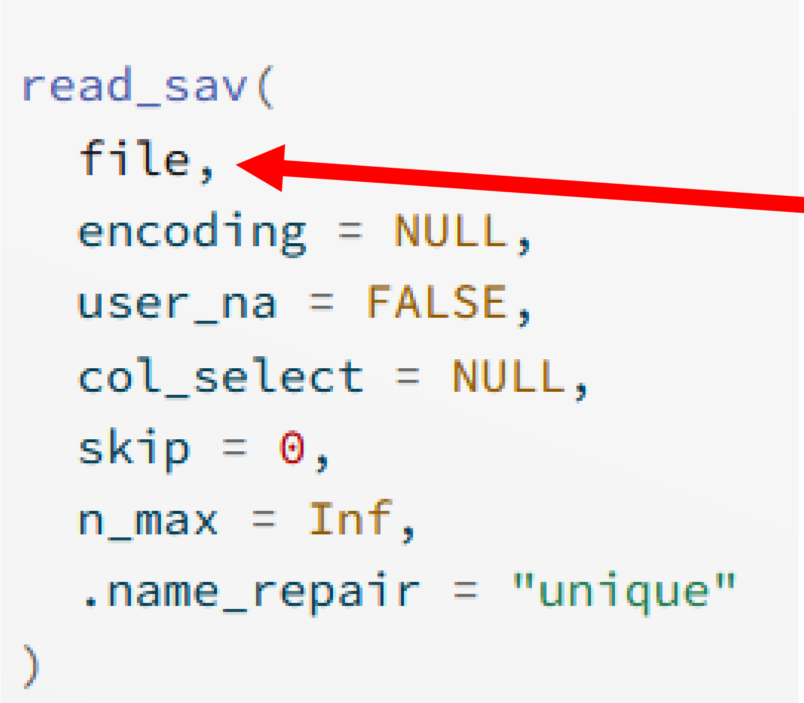
RSTUDIO – IMPORTING/READING DATA



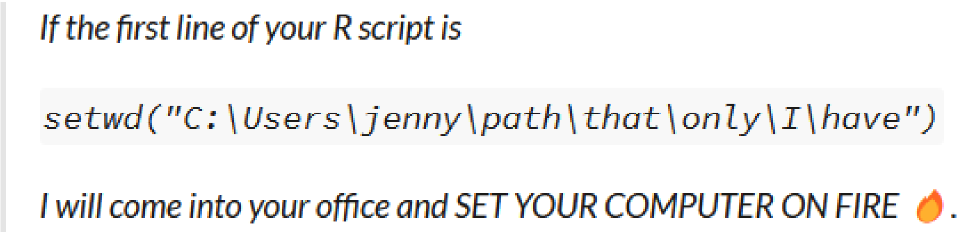
```
read_sav(
  file,
  encoding = NULL,
  user_na = FALSE,
  col_select = NULL,
  skip = 0,
  n_max = Inf,
  .name_repair = "unique"
)
```

file – refers to where on your computer this file lives. How can you give directions to get to this specific data file?

RSTUDIO – IMPORTING/READING DATA



```
read_sav(
  file,
  encoding = NULL,
  user_na = FALSE,
  col_select = NULL,
  skip = 0,
  n_max = Inf,
  .name_repair = "unique"
)
```



> *If the first line of your R script is*
>
> `setwd("C:\Users\jenny\path\that\only\I\have")`
>
> *I will come into your office and SET YOUR COMPUTER ON FIRE* 🔥.