

ANT 6973

DATA VISUALIZATION AND EXPLORATION

PLAN FOR TODAY

1. Introductions

2. Orientation

- Syllabus and course website
- R and RStudio

3. Complete your first in-class activities!

WHO AM I?

Dr. Fernando Campos

- Department of Anthropology
- Email:
fernando.campos@utsa.edu
- Office Hours:
 - Wed, 12:30 pm – 1:30 pm
 - By appointment



WHAT DO I DO?

Dr. Fernando Campos

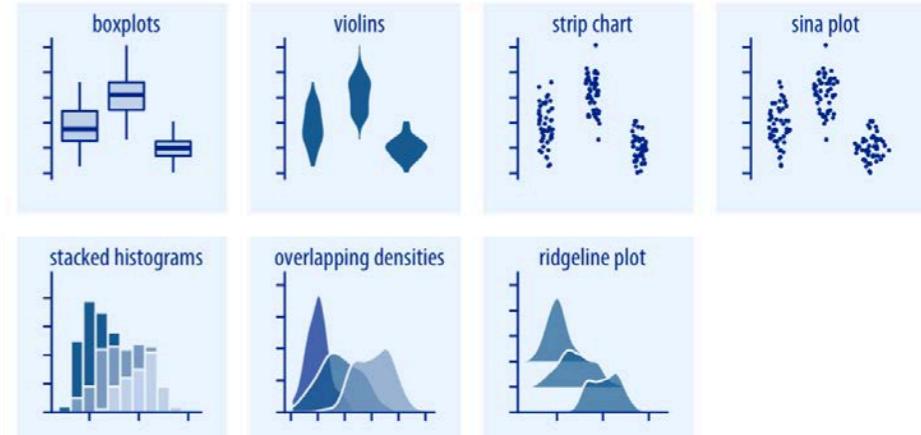
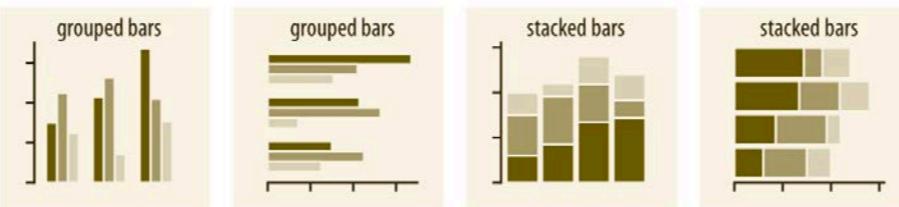
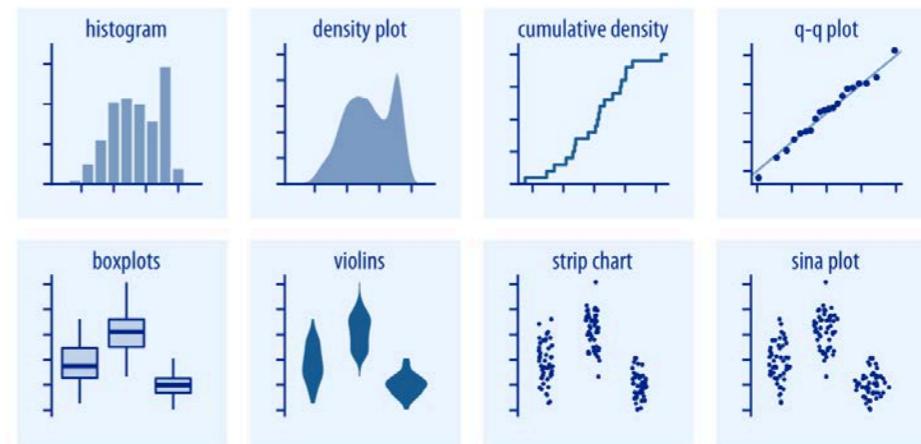
I'm a biological anthropologist, studying the behavioral ecology of wild primates.



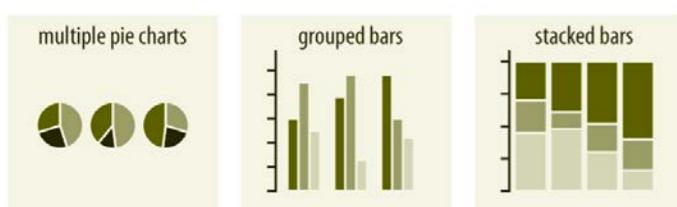
WHY AM I TEACHING
THIS COURSE?

DIRECTORY OF VISUALIZATIONS

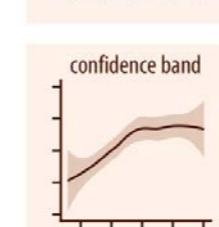
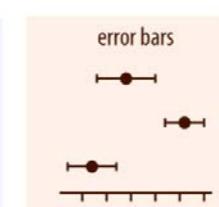
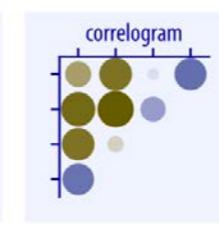
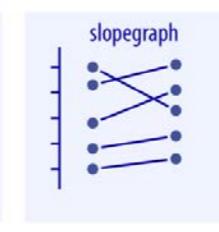
Amounts



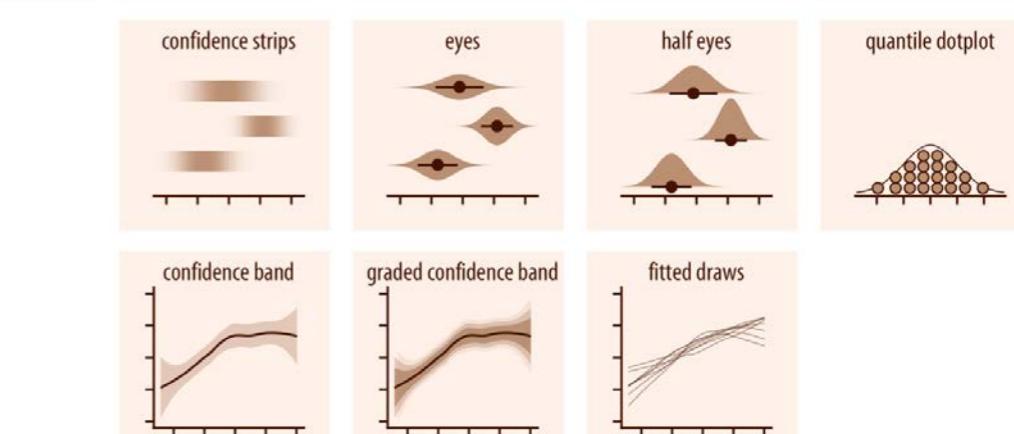
Proportions



Distributions



X-Y Relationships



DATA LITERACY



YOUR TURN

1. Who are you?
2. What are your research interests?
3. What is your experience with data?
4. Do you have any experience using R, and if so, how much?
5. Is there anything specific that you hope to learn in this course?



ORIENTATION

ORIENTATION

- Syllabus and course website
- R and RStudio

COURSE WEBSITE

- [https://www.campos-lab.net/
courses/data-viz/](https://www.campos-lab.net/courses/data-viz/)
- Bookmark it!
- Go here for up-to-date information, schedule, assignments, readings, etc.
- I'll use Canvas only for posting grades.
- Let's explore this a bit...

The screenshot shows a web browser window with the title bar 'ANT 6973: Data Visualization' and the URL 'localhost:6052/index.html'. The page content includes:

- A header section with a decorative graphic, a search bar, and navigation links for 'Home', 'Schedule', and 'Course Policies'.
- A main title 'ANT 6973: Data Visualization and Exploration'.
- A 'Course information' section listing:
 - Instructor: Dr. Fernando Campos
 - Location: Multidisciplinary Studies Bldg 2.03.18
 - Times: Wednesdays 9:00 am – 11:45 am
 - Course: ANT 6973.002
 - CRN: 24022
 - Office Hours: Wed, 12:30 pm – 1:30 pm (email to confirm meeting)
- An 'Overview' section describing the course as designed for graduate students in social or natural sciences to gain experience in data visualization and analysis using R, accessible to both novices and experienced users.
- An 'Approach' section stating the course uses a mix of traditional and flipped classroom approaches, with new material introduced in readings and class time focused on application.

On the right side of the page, there is a sidebar titled 'On this page' with links to 'Course information', 'Overview', 'Approach', and 'Textbooks'.

ORIENTATION

- Syllabus and course website
- R and RStudio

WHAT'S R VS. RSTUDIO?

- R is a programming language.
- RStudio is a nice user interface for R (an *integrated development environment* or IDE) that provides many convenient features and tools for using R.



Engine



Dashboard

WHY R?



- R is the premier data analysis and visualization platform.

```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

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.

> █
```

WHY LEARN R?

- **Pros (1 of 2):**

- Powerful, state-of-the-art, and used by many researchers, R is becoming a standard for data analysis
- Easy to extend, modify, and improve with add-on packages
 - **Programmable:** if R can't do a particular task right now, you (or someone) can program it
- Freely available for Windows, Mac, Unix



```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

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.
> |
```

WHY LEARN R?

- **Pros (2 of 2):**

- Reproducibility: R relies on a series of *written commands* rather than pointing and clicking (and this is a good thing!)
- Powerful plotting packages produce publication-quality graphics.
- R has a large and welcoming community.



```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

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.

> █
```

WHY LEARN R?

- **Cons:**

- Command-based rather than interactive in the point-and-click sense (like SPSS & Excel)
 - Can be significant learning hurdle
 - Specific commands easy to forget
- Documentation sometimes poor
- Still evolving: backward-compatibility has been an issue



```
● ● ● 1. R
Campos-Mac-Mini-2018:~ czx412$ r
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

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.
> |
```

- Let's make sure we all have a current installation of R.
- <https://cloud.r-project.org/>
- If you already have R installed, it's still a good idea to update.

The Comprehensive R Archive Network



CRAN
[Mirrors](#)
[What's new?](#)
[Search](#)
[CRAN Team](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Task Views](#)
[Other](#)

Documentation
[Manuals](#)
[FAQs](#)
[Contributed](#)

Download and Install R
 Precompiled binary distributions of the base system and contributed packages, **Windows** and **Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

 R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms
 Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2023-06-16, Beagle Scouts) [R-4.3.1.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)

Questions About R

- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

What are R and CRAN?
 R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc. Please consult the [R project homepage](#) for further information.

CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R. Please use the CRAN [mirror](#) nearest to

- And a current installation of RStudio.
- <https://posit.co/download/rstudio-desktop/>
- If you already have RStudio installed, check for a new version using menus:
 - Help > Check for Updates

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR MACOS 11+

This version of RStudio is only supported on macOS 11 and higher. For earlier macOS environments, please [download a previous version](#).

Size: 380.82 MB | [SHA-256: 184804EA](#) | Version: 2023.06.1+524 | Released: 2023-07-07

All Installers and Tarballs

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use [an older version of RStudio](#).

Linux users may need to import [Posit's public code-signing key](#) prior to installation, depending on the operating system's security policy.

OS	Download	Size	SHA-256
Windows 10/11	RSTUDIO-2023.06.1-524.EXE	212.77 MB	A8325AD5
Mac OS X	RSTUDIO-2023.06.1-524.DMG	212.77 MB	184804EA
Linux	RSTUDIO-2023.06.1-524-Linux.tar.gz	212.77 MB	184804EA

LET'S PAUSE UNTIL WE ALL
COMPLETE THESE STEPS



- If you already have RStudio installed, now would be a good time to **check for updates**.
- In RStudio, Help > Check for Updates...

The screenshot shows the RStudio interface with the Help menu open. The 'Check for Updates' option is highlighted with a blue background. The menu also includes links for R Help, Search R Help, About RStudio, Accessibility, RStudio Docs, RStudio Community Forum, Cheat Sheets, Keyboard Shortcuts Help, Markdown Quick Reference, Roxygen Quick Reference, and Diagnostics.

Check for Updates

R Help
Search R Help ^ ⌘ F1

About RStudio

Accessibility >

RStudio Docs

RStudio Community Forum

Cheat Sheets >

Keyboard Shortcuts Help ⌘ K

Markdown Quick Reference

Roxygen Quick Reference

Diagnostics >

Console

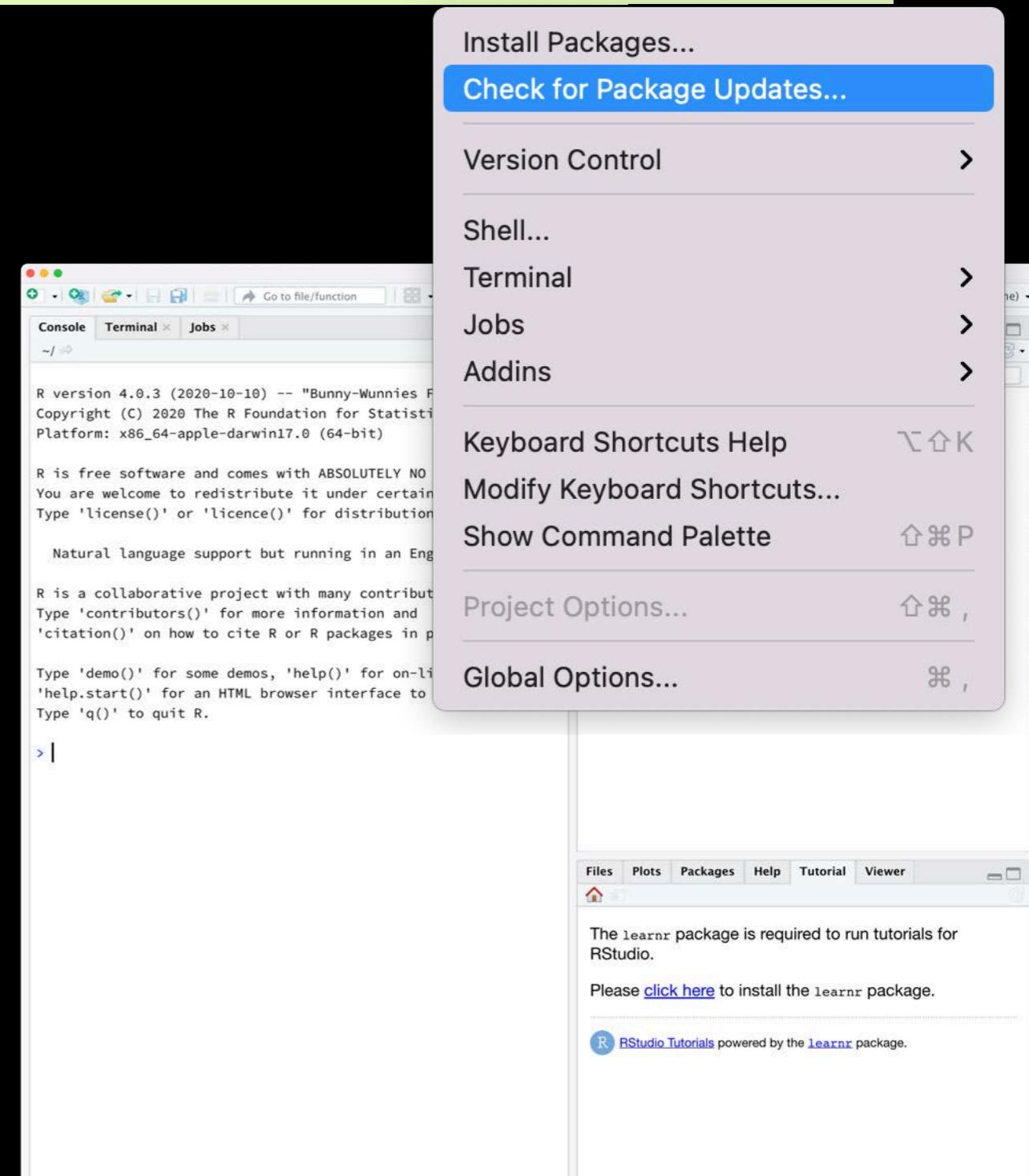
```
R version 4.0.3 (2020-10-10) -  
Copyright (C) 2020 The R Foundation  
Platform: x86_64-apple-darwin11.0.0  
  
R is free software and comes with  
no warranty. You are welcome to redistribute  
it under certain conditions.  
Type 'license()' or 'licence()' for more  
information.  
  
Natural language support but  
no machine translation.  
  
R is a collaborative project where  
improvements are welcome.  
Type 'contributors()' for more  
information.  
  
Type 'citation()' on how to cite R  
or 'demo()' for some demos,  
'help.start()' for an HTML  
browser-based help system,  
or 'q()' to quit R.  
  
> |
```

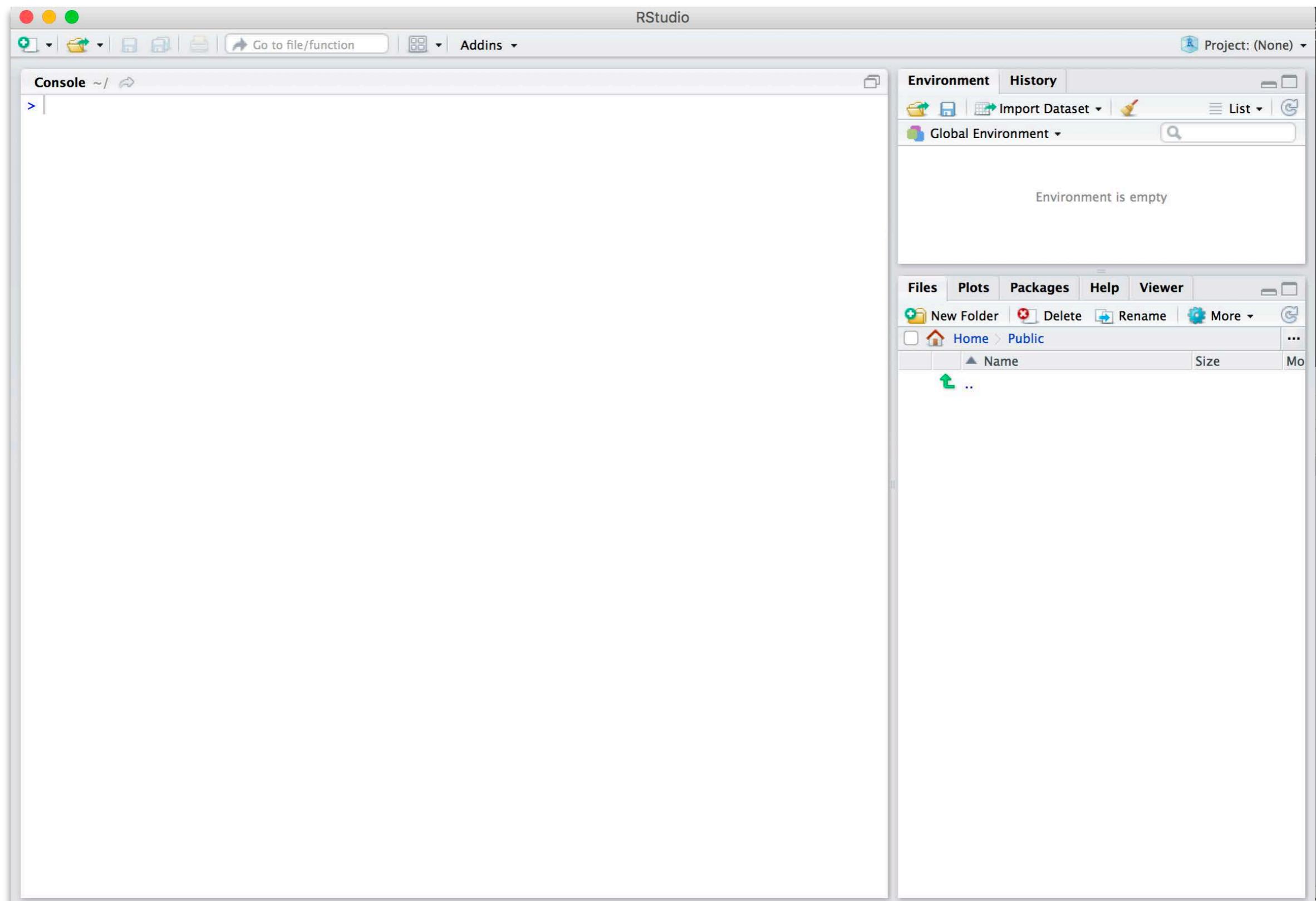
Tutorial

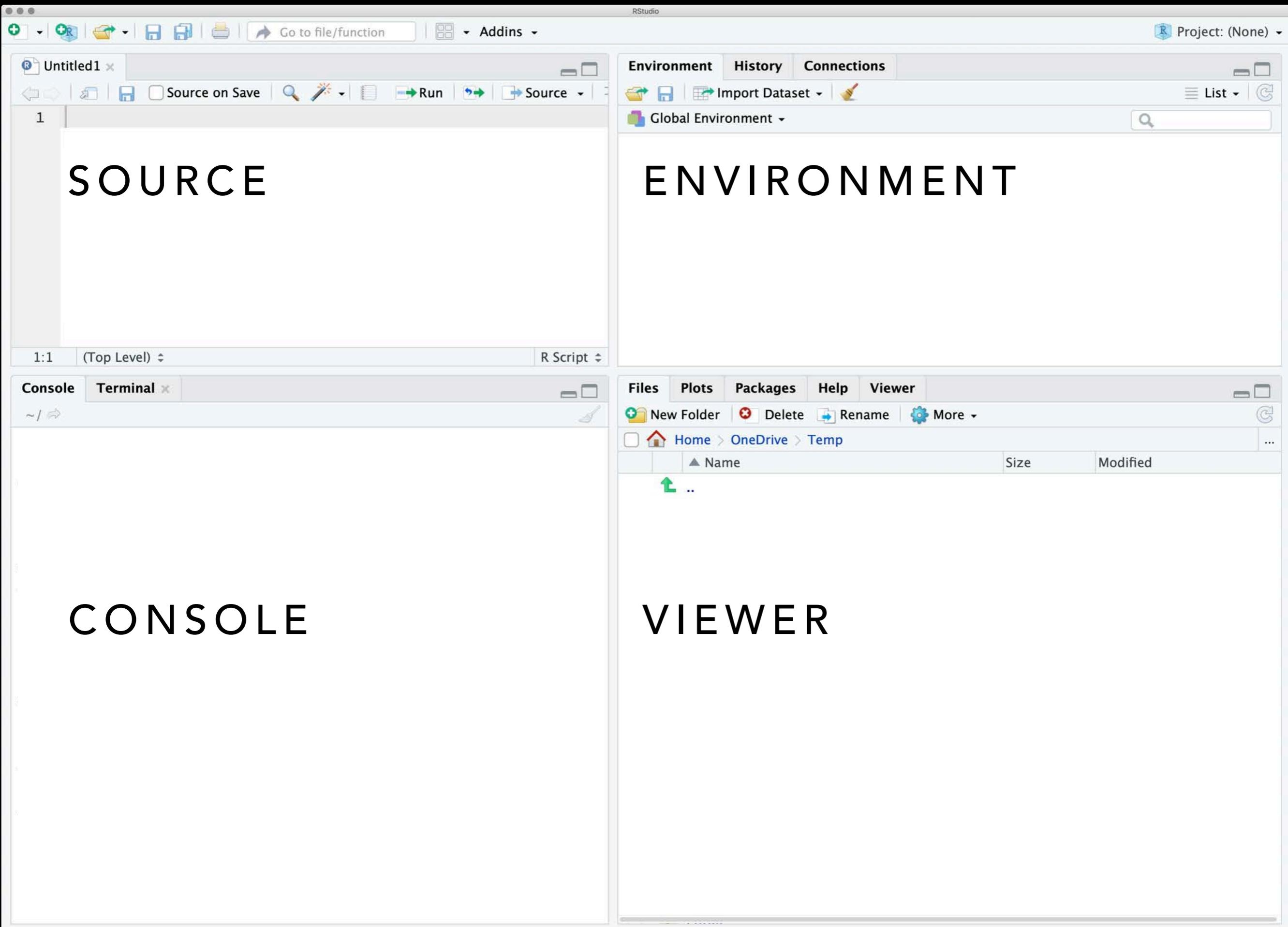
The `learnr` package is required to run tutorials for RStudio.
Please [click here](#) to install the `learnr` package.

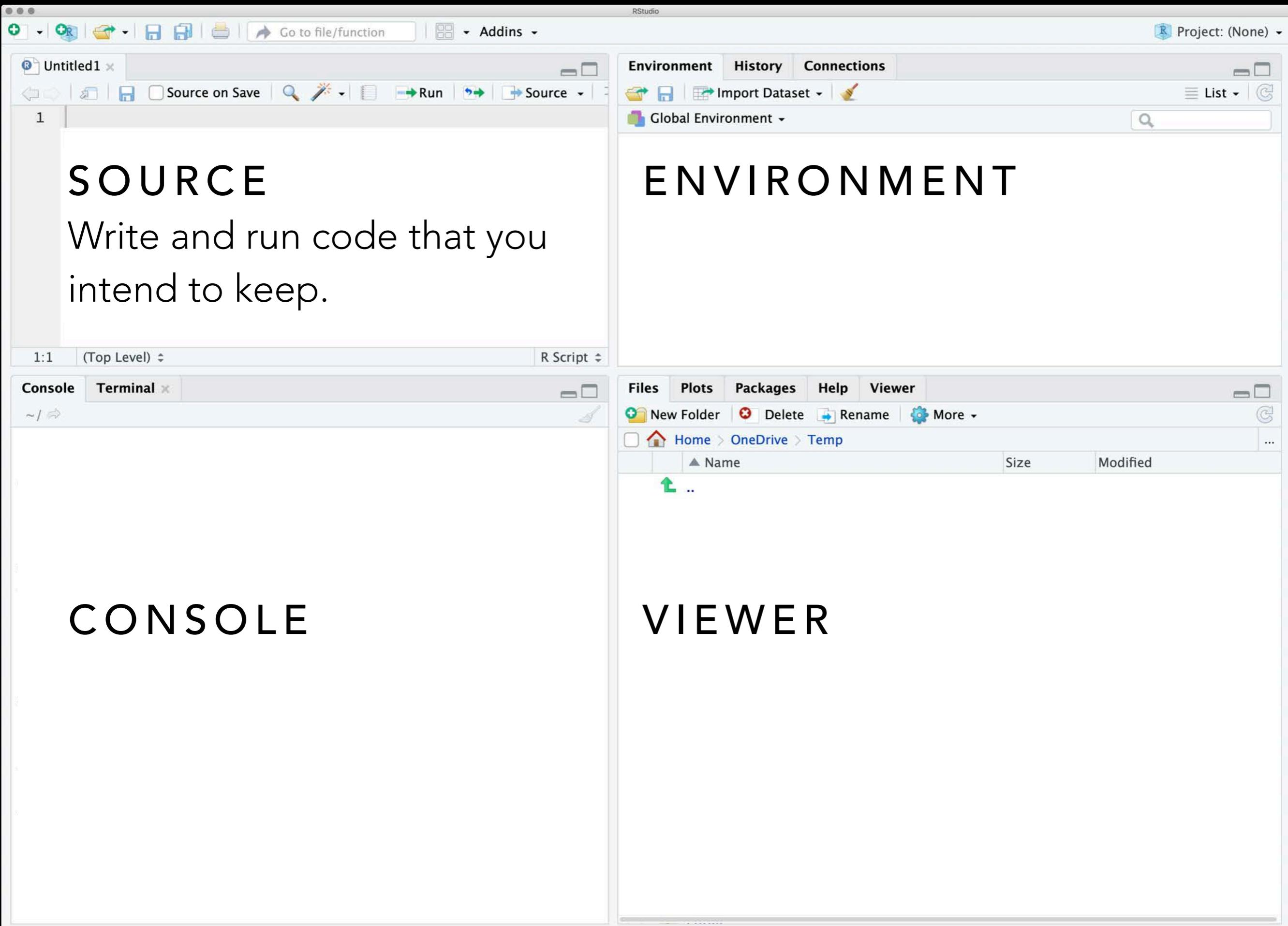
RStudio Tutorials powered by the `learnr` package.

- ... and while you're at it, update all your packages.
- In RStudio, Tools > Check for Package Updates...
- It is good practice to periodically update R and start with a clean slate
 - Outdated packages are a common source of errors









The image shows the RStudio IDE interface with four main panels:

- SOURCE**: A panel for writing and running code. It includes a code editor with a tab labeled "Untitled1 x", a toolbar with icons for file operations, and a status bar indicating "1:1 (Top Level) R Script".
- ENVIRONMENT**: A panel for managing global variables. It shows tabs for "Environment", "History", and "Connections", and a dropdown menu for "Global Environment".
- CONSOLE**: A panel for viewing executed code. It includes tabs for "Console" and "Terminal", and a status bar indicating "~ /".
- VIEWER**: A panel for file management. It shows tabs for "Files", "Plots", "Packages", "Help", and "Viewer". The "Files" tab is active, displaying a file tree under "Home > OneDrive > Temp" and a list view with columns for "Name", "Size", and "Modified".

The image shows the RStudio IDE interface with four main sections labeled: SOURCE, ENVIRONMENT, CONSOLE, and VIEWER.

SOURCE
Write and run code that you intend to keep.

ENVIRONMENT
View objects, functions, etc.
View command history.

CONSOLE
View executed code. Use sparingly for experimentation.

VIEWER

The RStudio interface includes a top menu bar with File, Edit, View, Insert, Cell, Run, Tools, Help, and Addins. The left pane features a file browser, a source editor with tabs for Untitled1 (R Script), and a console tab. The right pane includes an environment viewer, a file browser, and a viewer pane.

The image shows the RStudio IDE interface with four main sections labeled: SOURCE, ENVIRONMENT, CONSOLE, and VIEWER.

SOURCE
Write and run code that you intend to keep.

ENVIRONMENT
View objects, functions, etc.
View command history.

CONSOLE
View executed code. Use sparingly for experimentation.

VIEWER
View files in working directory, packages, image outputs, etc.





1. ORGANIZE YOUR WORK INTO PROJECTS

(more on this next week)



2. ALWAYS START A SESSION
WITH A CLEAN SLATE

RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help

Install Packages...
Check for Package Updates...

Version Control >

Shell...
Terminal
Jobs
Addins

Keyboard Shortcuts Help ⌘K
Modify Keyboard Shortcuts...
Show Command Palette ⌘P

Project Options... ⌘,

Global Options... ⌘,

RStudio

Console Terminal Jobs

~/Documents/academic-website/ ↵

R version 4.0.3 (2020-10-10) -- "Bunny-Wunnies Freak Out"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (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.

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.

> |

Environment History Connections

R Global Environment

Env

Files Plots Packages Help Tutorial Viewer

New Folder Delete Rename More

Home > Documents > academic-website

	Name	Size	Modified
..			
	.gitignore	53 B	Apr 12, 2020, 9:53 AM
	.Rhistory	2 KB	Jan 14, 2021, 10:44 AM
	academic.Rproj	258 B	Jan 14, 2021, 10:44 AM
	assets		
	config		
	content		
	data		
	LICENSE.md	1.1 KB	Apr 12, 2020, 9:51 AM
	public		
	README.md	3.3 KB	Apr 12, 2020, 9:51 AM
	resources		

R version 4.0.3 (2020-10-10) -- "Bunny"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (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 details.

Natural language support but running

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

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

>

Project: (None)

Console Terminal Jobs

~/Documents/academic-website/

Options General

Code Console Appearance Pane Layout Packages R Markdown Sweave Spelling Git/SVN Publishing Terminal Accessibility Python

Basic Graphics Advanced

R Sessions

Default working directory (when not in a project):
~/Documents/academic-website

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:

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

OK Cancel Apply

resources

Size	Modified
53 B	Apr 12, 2020, 9:53 AM
2 KB	Jan 14, 2021, 10:44 AM
258 B	Jan 14, 2021, 10:44 AM
1.1 KB	Apr 12, 2020, 9:51 AM
3.3 KB	Apr 12, 2020, 9:51 AM



3. TWO WAYS OF INTERACTING WITH R: CONSOLE AND SCRIPTS

(most of the time, use scripts)

RStudio

Project: (None)

Untitled1*

```
1 2 > 3
2
3 2 + 2 == 4
4
5 isTRUE(2 + 2 == 5)
6
```

Source on Save | Run | Source | Environment | History | Connections | Import Dataset | Global Environment | List | Search

Environment is empty

6:1 (Top Level) R Script

Console Terminal Jobs

```
~/Documents/academic-website/
> 2 > 3
[1] FALSE
> 2 + 2 == 4
[1] TRUE
> isTRUE(2 + 2 == 5)
[1] FALSE
>
```

Files Plots Packages Help Tutorial Viewer

New Folder Delete Rename More

Home > Documents > academic-website

	Name	Size	Modified
..			
	.gitignore	53 B	Apr 12, 2020, 9:53 AM
	.Rhistory	2.2 KB	Jan 19, 2021, 12:51 PM
	academic.Rproj	258 B	Jan 14, 2021, 10:44 AM
	assets		
	config		
	content		
	data		
	LICENSE.md	1.1 KB	Apr 12, 2020, 9:51 AM
	public		
	README.md	3.3 KB	Apr 12, 2020, 9:51 AM
	resources		



4. BECOME COMFORTABLE WITH ERROR MESSAGES



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.

```
> my_data <- read.csv("my_data.csv")
Error in file(file, "rt") : cannot open the connection
In addition: Warning message:
In file(file, "rt") :
  cannot open file 'my_data.csv': No such file or directory
> Can you say that in plain English please?
Error: unexpected symbol in "Can you"
> :(
Error: unexpected ':' in ":"
```

> |



5. USE R STUDIO'S TOOLS
TO MAKE LIFE EASIER

YOUR TURN

Please complete the activity
for this week listed on the
course website under:

Schedule > Assignments

It's short!



20:00