S371 Lab
Introduction to R

1. Class logistics
2. Attendance and participation
3. Install R and Rstudio
4. Rstudio interface

Something about me

- Ekaterina (Katya) Baldina
- I am your lab instructor
- I will
  - Help you go through the technical side of data analysis in this class
  - Help you with your homework assignments, exam preparation, and the final project
- My office hours:
  - By appointment
  - If you want to meet with me beyond the lab hours, please make an appointment with me via email
- My email: baldina@iu.edu

Rules in the lab

- Whenever you have a question, feel free to ask me questions about assignments and other things
- Or feel free to shoot me an email!

Rules in the lab

- I will post the slides on Canvas after lab

Rules in the lab

- You are not required to wear a mask, but please wear one in you feel sick, or you may skip the lab if you are unwell. Please, think about others around you! ☺

Announcement

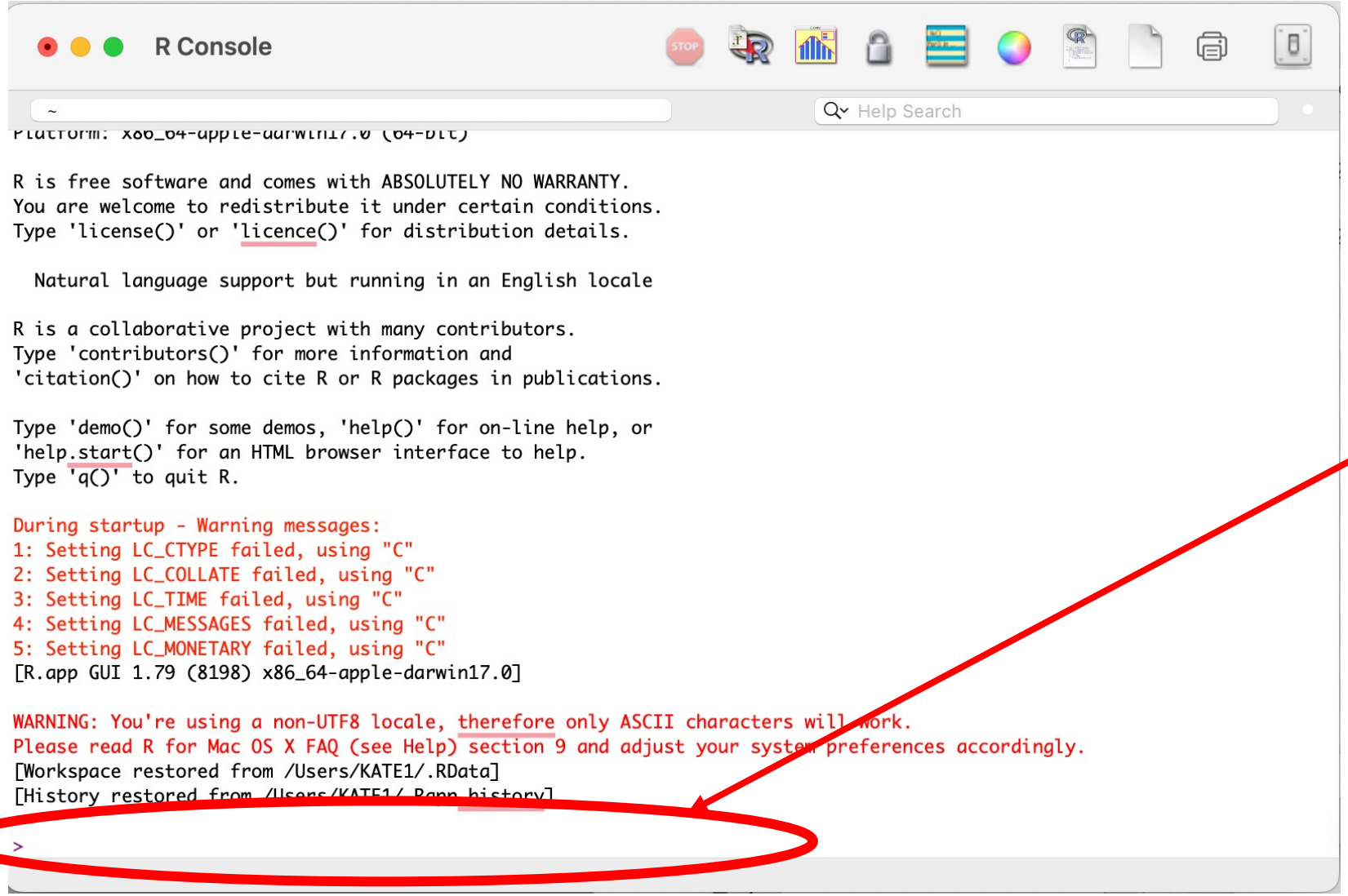
- Your Canvas grade may not be accurate
- If you have questions let me know!

Statistical Software: R

R

- A statistical software
  - Help you do complex calculations
- A programming language
  - You input some commands in R→R do the operations and give you the results
- Why do we use R?
  - Powerful functions
  - Free open source software

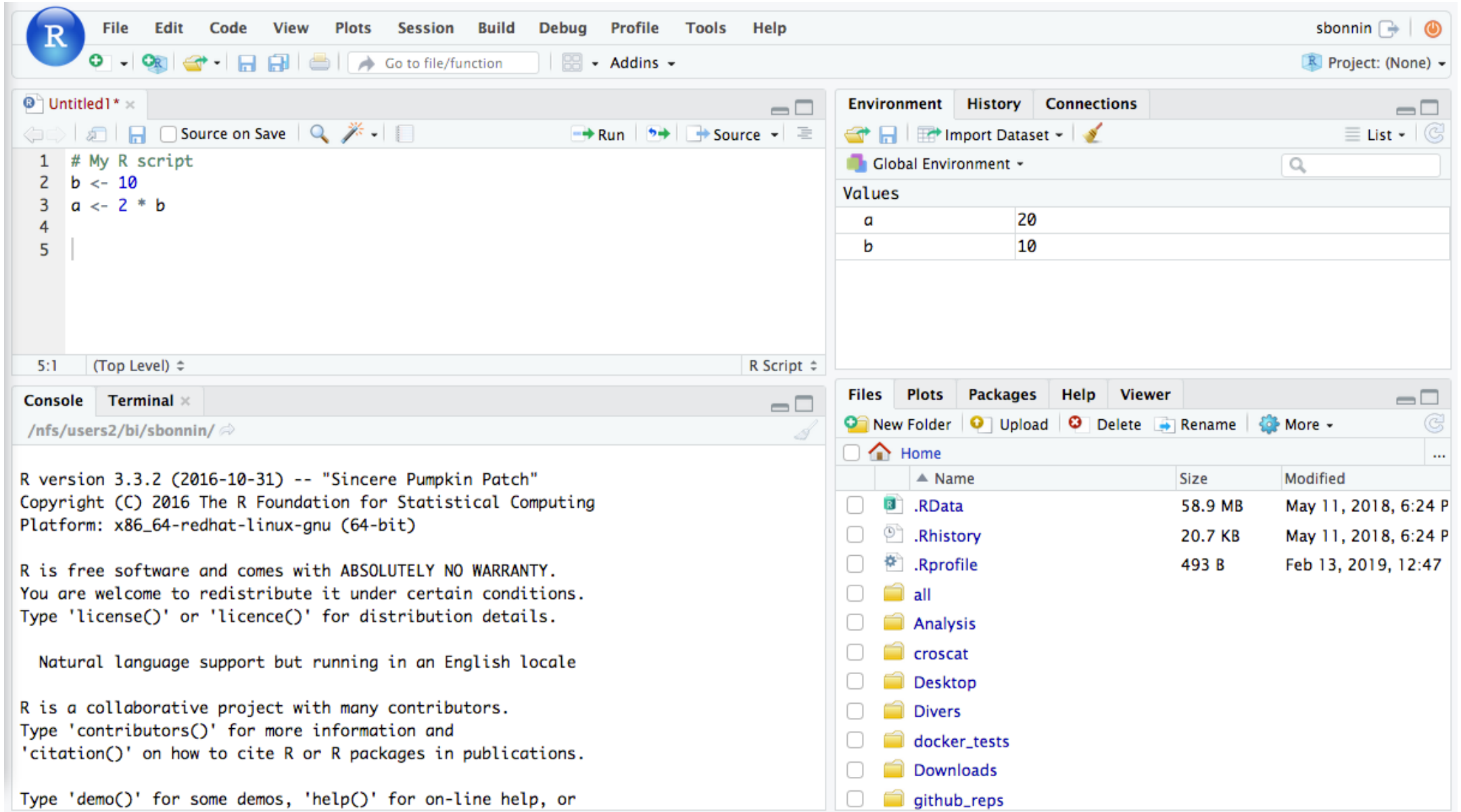
R Console

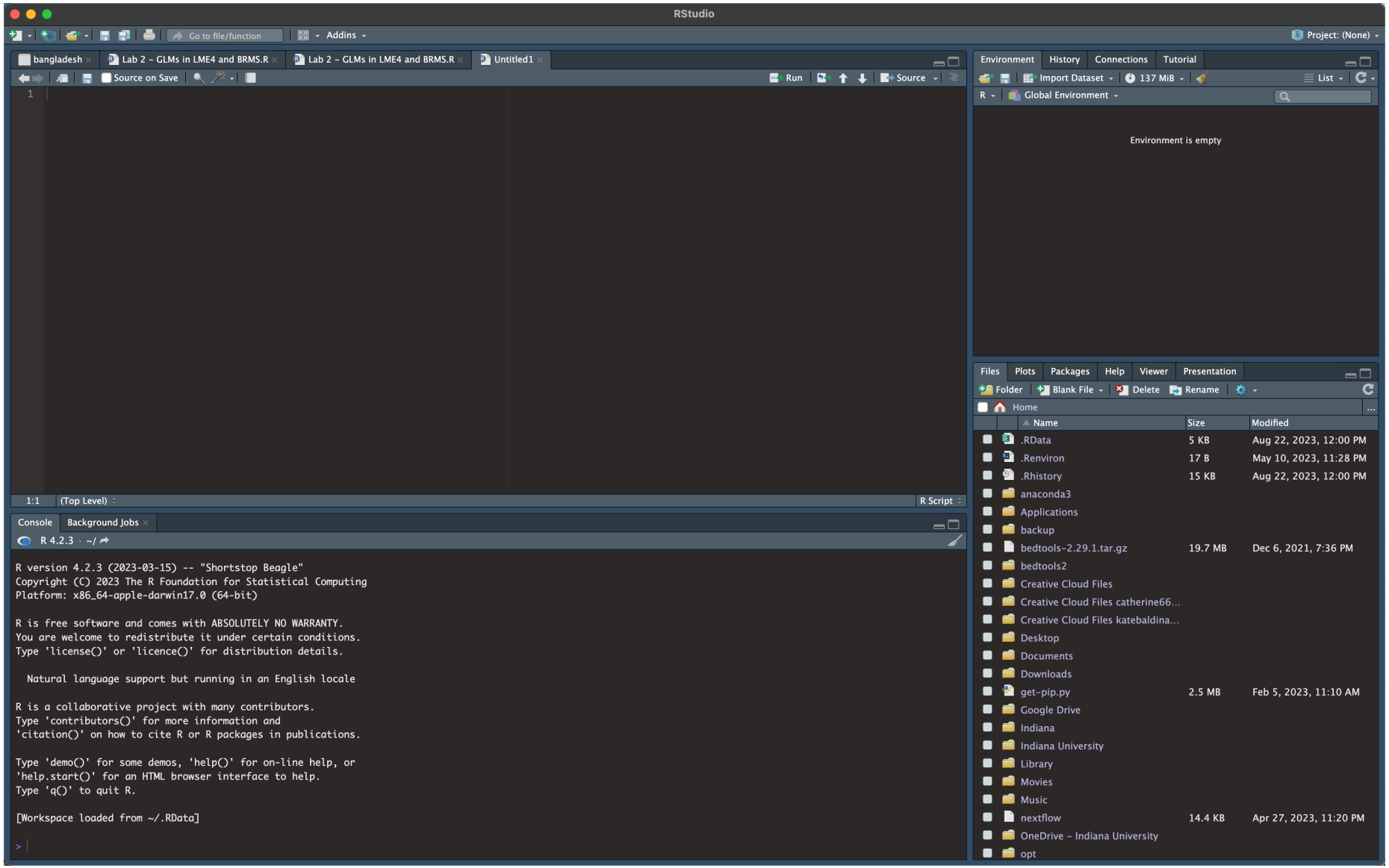


You can input commands here and then hit Enter

RStudio

- An integrated development environment (IDE)
  - Better looking interface
  - Easier to access different functions
  - Give you hint in programming
  - Make programming in R much easier
- It is also free
- Rstudio is an add-on to R
- **In order to use Rstudio, you need to install R first**


File Edit Code View Plots Session Build Debug Profile Tools Help
sbonnin
Go to file/function
Addins
Project: (None)
Untitled1*
Source on Save
Run
Source
1 # My R script
2 b <- 10
3 a <- 2 * b
4
5
5:1
(Top Level)
R Script
Environment History Connections
Import Dataset
List
Global Environment
Values
a 20
b 10
Console Terminal
/nfs/users2/bi/sbonnin/
R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-redhat-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.
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
Files Plots Packages Help Viewer
New Folder
Upload
Delete
Rename
More
Home
Name
Size
Modified
.RData 58.9 MB May 11, 2018, 6:24 P
.Rhistory 20.7 KB May 11, 2018, 6:24 P
.Rprofile 493 B Feb 13, 2019, 12:47
all
Analysis
croscat
Desktop
Divers
docker_tests
Downloads
github_reps


RStudio
Go to file/function
Addins
Project: (None)
bangladesh
Lab 2 - GLMs in LME4 and BRMS.R
Lab 2 - GLMs in LME4 and BRMS.R
Untitled1
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Console
Background Jobs
R 4.2.3 · ~/
R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 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.
[Workspace loaded from ~/.RData]
>
Environment
History
Connections
Tutorial
Import Dataset
137 MiB
List
R
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
Presentation
Folder
Blank File
Delete
Rename
Home
Name
Size
Modified
.RData 5 KB Aug 22, 2023, 12:00 PM
.Renviron 17 B May 10, 2023, 11:28 PM
.Rhistory 15 KB Aug 22, 2023, 12:00 PM
anaconda3
Applications
backup
bedtools-2.29.1.tar.gz 19.7 MB Dec 6, 2021, 7:36 PM
bedtools2
Creative Cloud Files
Creative Cloud Files catherine66...
Creative Cloud Files katebaldina...
Desktop
Documents
Downloads
get-pip.py 2.5 MB Feb 5, 2023, 11:10 AM
Google Drive
Indiana
Indiana University
Library
Movies
Music
nextflow 14.4 KB Apr 27, 2023, 11:20 PM
OneDrive - Indiana University
opt

Installing R and RStudio

- R and RStudio are installed in all computers on campus
- You can also download R and RStudio to your own computer and install them yourself (Windows, Mac, and Linux)
- Download R:
- https://cloud.r-project.org/
- Download RStudio (download the free version):
- https://rstudio.com/products/rstudio/download/

Installing R and RStudio

- Step 1: Download R
  - https://cloud.r-project.org/
- Step 2: Install R
- Step 3: Download Rstudio
  - https://rstudio.com/products/rstudio/download/
- Step 4: Install Rstudio
- Step 5: Open Rstudio and start doing some data analysis!

Installing R

- Download R that suits your operating system

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

The Comprehensive R Archive Network

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

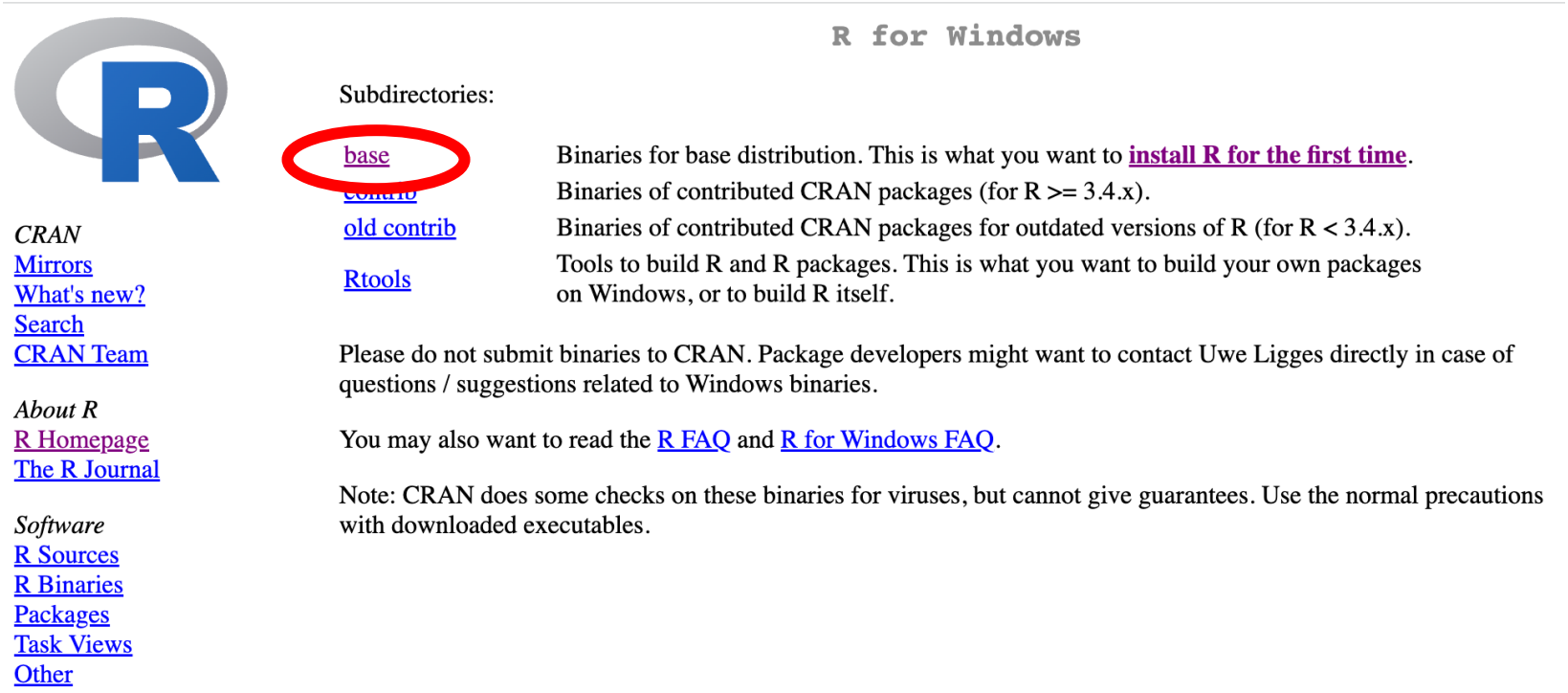
Questions About R

- 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?

Installing R for Windows users

- Download **base**



Installing R for Windows users

- Click the download link

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

R-4.3.1 for Windows

Download R-4.3.1 for Windows (79 megabytes, 64 bit)
README on the Windows binary distribution
New features in this version

This build requires UCRT, which is part of Windows since Windows 10 and Windows Server 2016. On older systems, UCRT has to be installed manually from here.

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the md5sum of the .exe to the fingerprint on the master server.

Frequently asked questions

- Does R run under my version of Windows?
- How do I update packages in my previous version of R?

Please see the R FAQ for general information about R and the R Windows FAQ for Windows-specific information.

Other builds

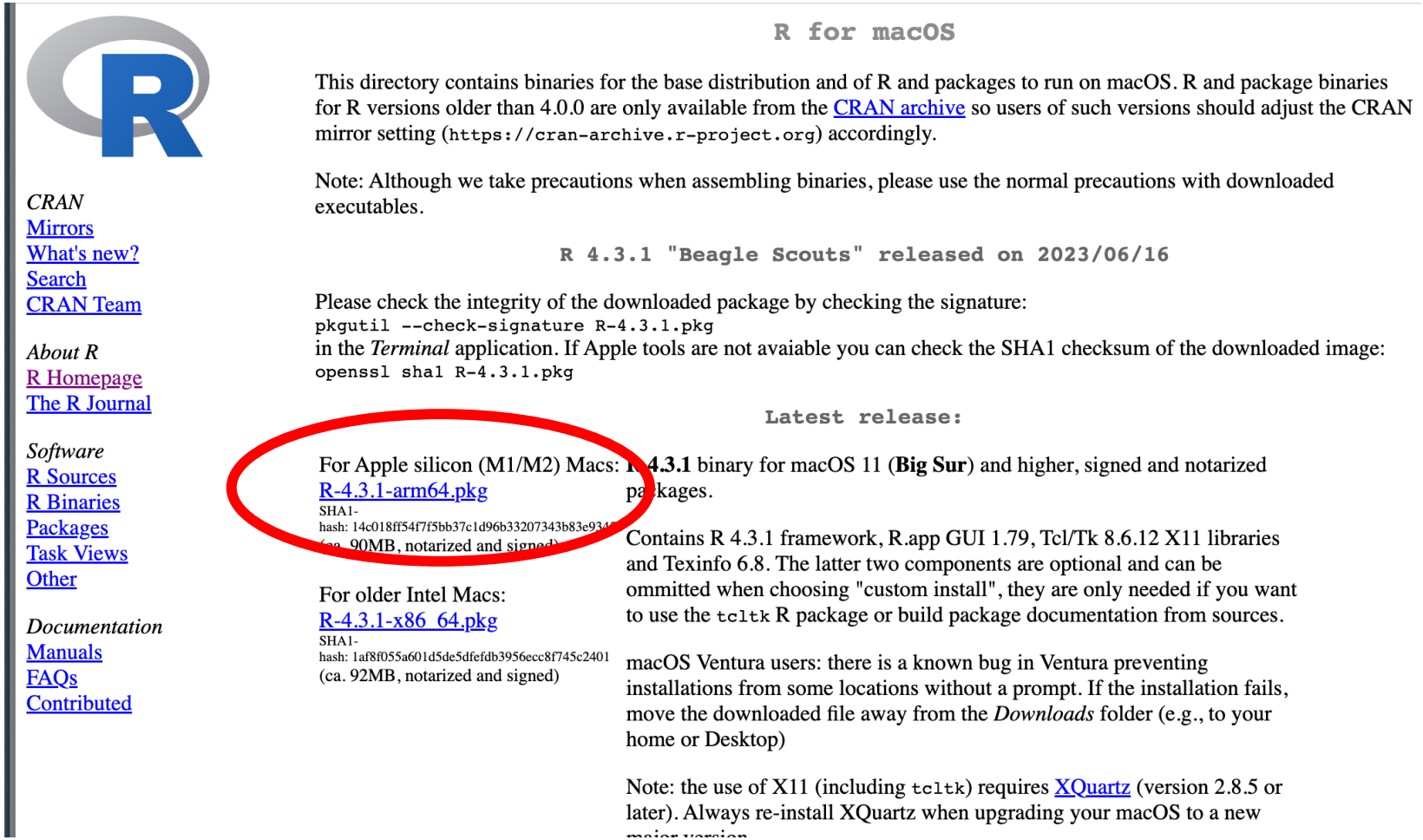
- Patches to this release are incorporated in the r-patched snapshot build.
- A build of the development version (which will eventually become the next major release of R) is available in the r-devel snapshot build.
- Previous releases

Note to webmasters: A stable link which will redirect to the current Windows binary release is <CRAN MIRROR>/bin/windows/base/release.html.

Last change: 2023-06-16

Installing R for Mac Users

- Click the download link



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

R for macOS

This directory contains binaries for the base distribution and of R and packages to run on macOS. R and package binaries for R versions older than 4.0.0 are only available from the CRAN archive so users of such versions should adjust the CRAN mirror setting (`https://cran-archive.r-project.org`) accordingly.

Note: Although we take precautions when assembling binaries, please use the normal precautions with downloaded executables.

R 4.3.1 "Beagle Scouts" released on 2023/06/16

Please check the integrity of the downloaded package by checking the signature:
`pkgutil --check-signature R-4.3.1.pkg`
in the *Terminal* application. If Apple tools are not avaiable you can check the SHA1 checksum of the downloaded image:
`openssl sha1 R-4.3.1.pkg`

Latest release:

For Apple silicon (M1/M2) Macs:
R-4.3.1-arm64.pkg
SHA1-
hash: 14c018ff54f7f5bb37c1d96b33207343b83e934

90MB, notarized and signed)

For older Intel Macs:
R-4.3.1-x86_64.pkg
SHA1-
hash: 1af8f055a601d5de5dfefdb3956ecc8f745c2401
(ca. 92MB, notarized and signed)

4.3.1 binary for macOS 11 (**Big Sur**) and higher, signed and notarized packages.

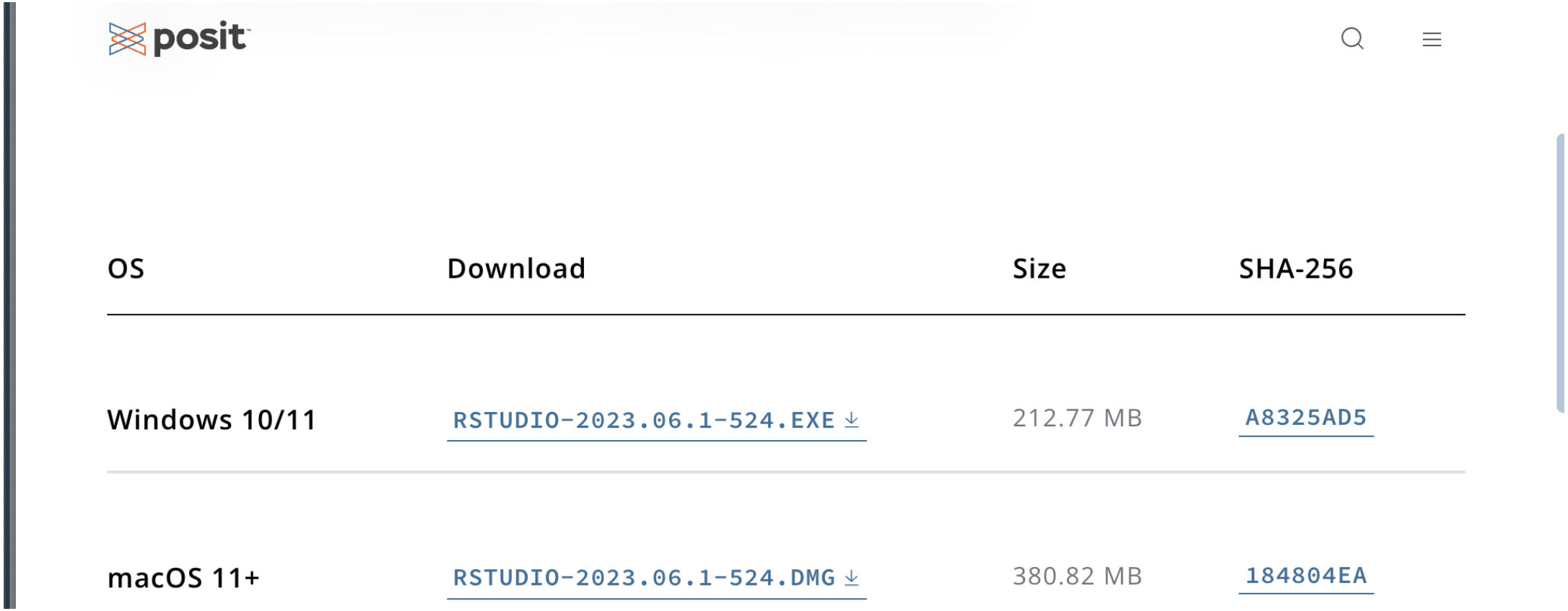
Contains R 4.3.1 framework, R.app GUI 1.79, Tcl/Tk 8.6.12 X11 libraries and Texinfo 6.8. The latter two components are optional and can be ommitted when choosing "custom install", they are only needed if you want to use the `tcltk` R package or build package documentation from sources.

macOS Ventura users: there is a known bug in Ventura preventing installations from some locations without a prompt. If the installation fails, move the downloaded file away from the *Downloads* folder (e.g., to your home or Desktop)

Note: the use of X11 (including `tcltk`) requires XQuartz (version 2.8.5 or later). Always re-install XQuartz when upgrading your macOS to a new

Installing RStudio

- https://rstudio.com/products/rstudio/download/



Why do we need to use statistical software?

- Because we need to do data analysis
- What is data analysis?

Why do we need to use statistical software?

- An analogy: cooking
- What is cooking?
- We turn raw food into a dish

Why do we need to use statistical software?

- What are the things you need for cooking?
  - You need a kitchen, kitchenware and ingredients

- What is data analysis?
  - We turn the raw data (ingredients) into meaningful outputs (dishes)
  - R is the kitchen with all kinds of kitchenware
  - R gives us all the tools we need to turn the raw data into meaningful outputs
- There are other statistical software that can do the same things as R
    - But they are not free

Why do we need to use statistical software?

- Professor Schultz will teach you different ways to cook dishes (scrambled eggs, pasta, pizza, fried rice…)
  - i.e. different kinds of statistics and the underlying theoretical/mathematical principles
- I will teach you how to use different kinds of kitchenware to make cooking possible

Why do you need to attend the lab?

- Even though your lab attendance doesn't affect your final grade directly, lab attendance affects your final grade indirectly.
- If you don't come to the lab
  - You don't know how to use R to do data analysis → you can't do assignments and project correctly
  - You lose the opportunity to ask me
- I will give you a hands-on experience on how to use R for data analysis
- I will answer your questions on upcoming homeworks, exams, project

What is the relationship between R and Rstudio?

- R is the programming language that comes with a very basic terminal (interface to do programming)
- Rstudio is an add-on IDE to the R programming language
  - Fancier interface
  - Make programming in R much easier

RStudio

- Rstudio is just fancier and easier to use than the default R terminal
- So we will stick with Rstudio throughout this semester
- When you need to analyze data, just open Rstudio. You don’t need to bother with R (just leave it installed on your computer. That’s it)

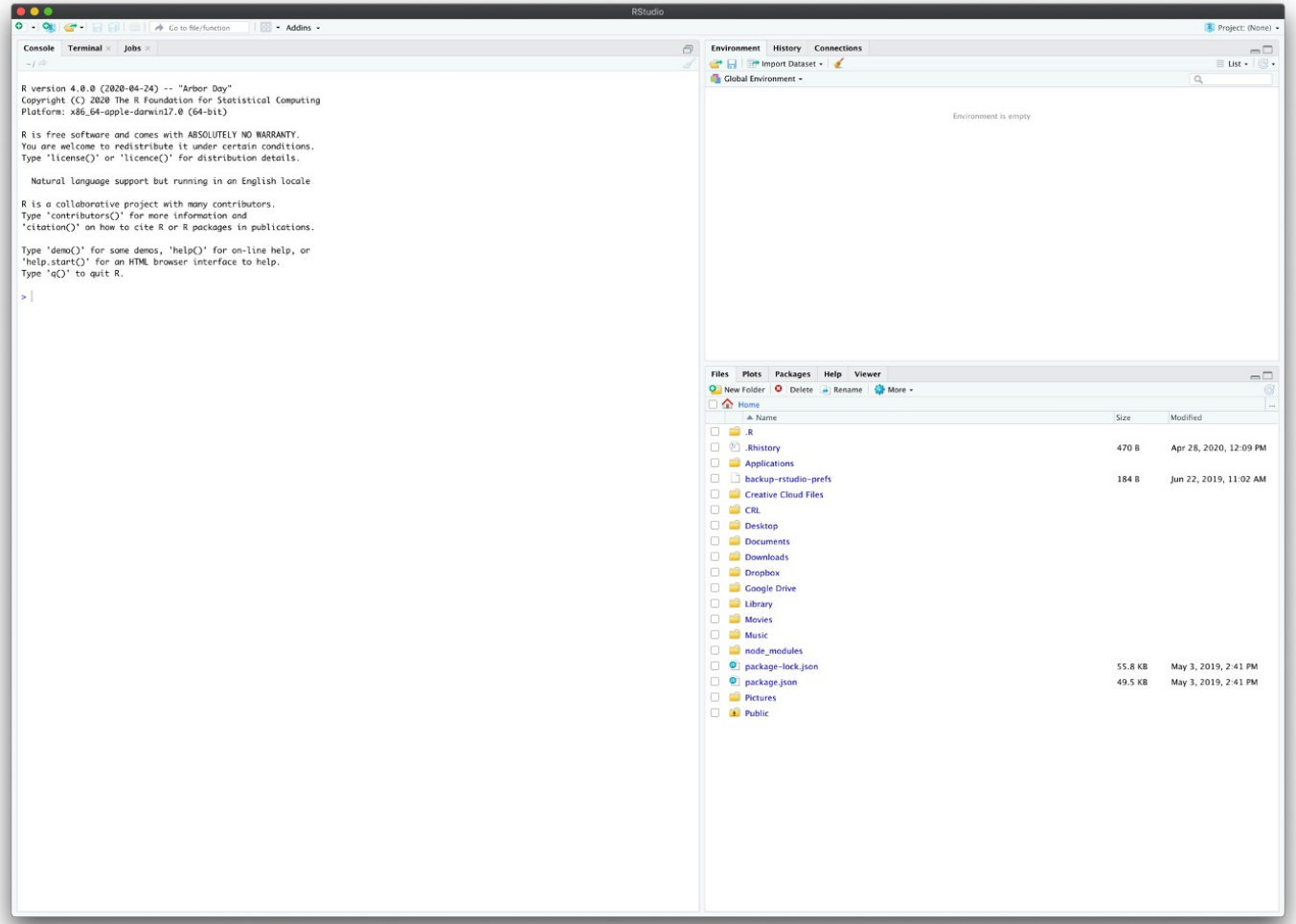
A Note

- When I use the term R, I am reffering to the R programming language
  - Example: “The R code/package is...”
- When I use the term Rstudio, I am referring to the interface that we use
  - Example: “In the command window of Rstudio...”

Opening RStudio

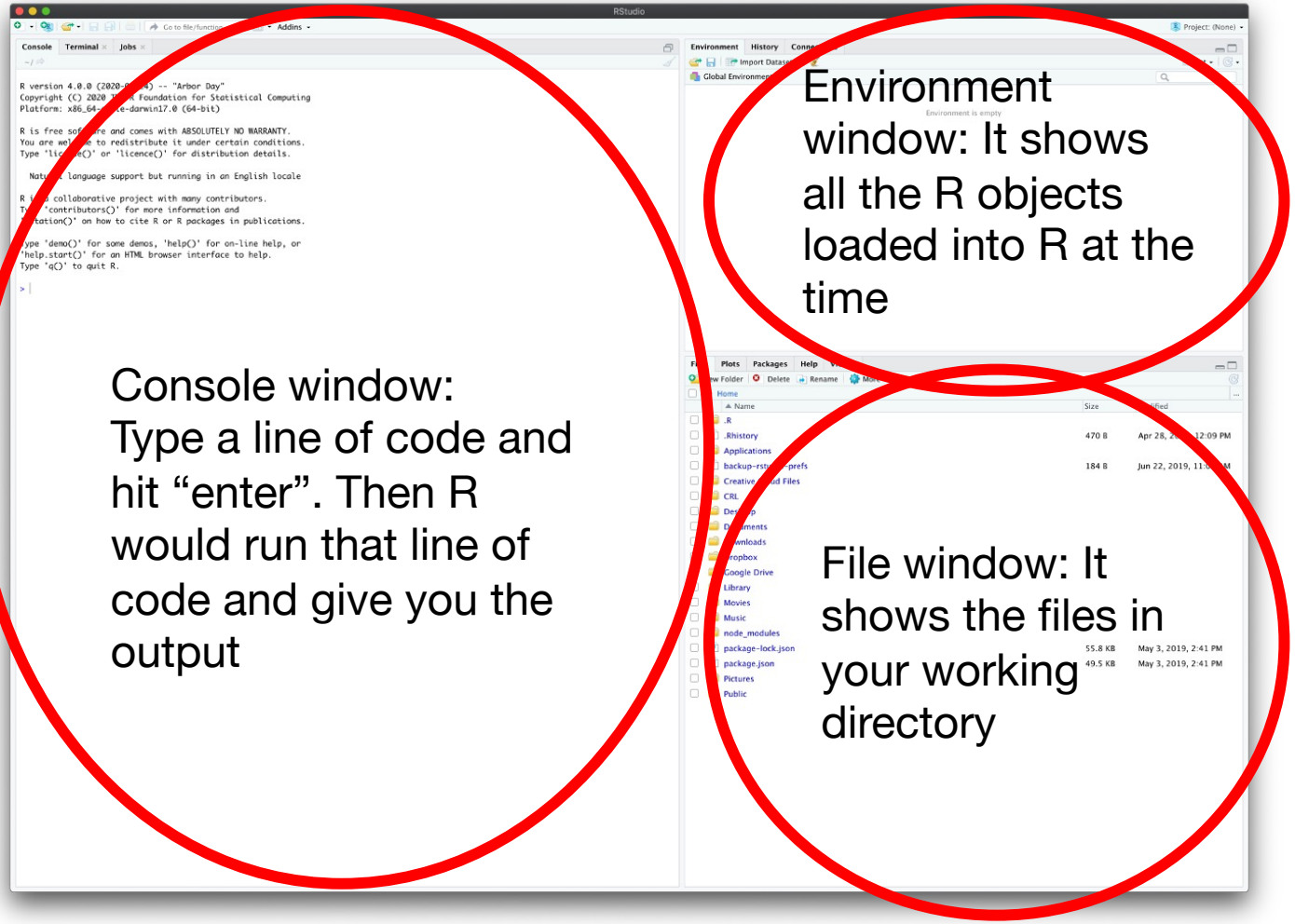
- Search for “Rstudio” in Windows status bar
- For Mac, you can try to search your launch pad or click Command+Space and enter “Rstudio” in the search window

Opening RStudio



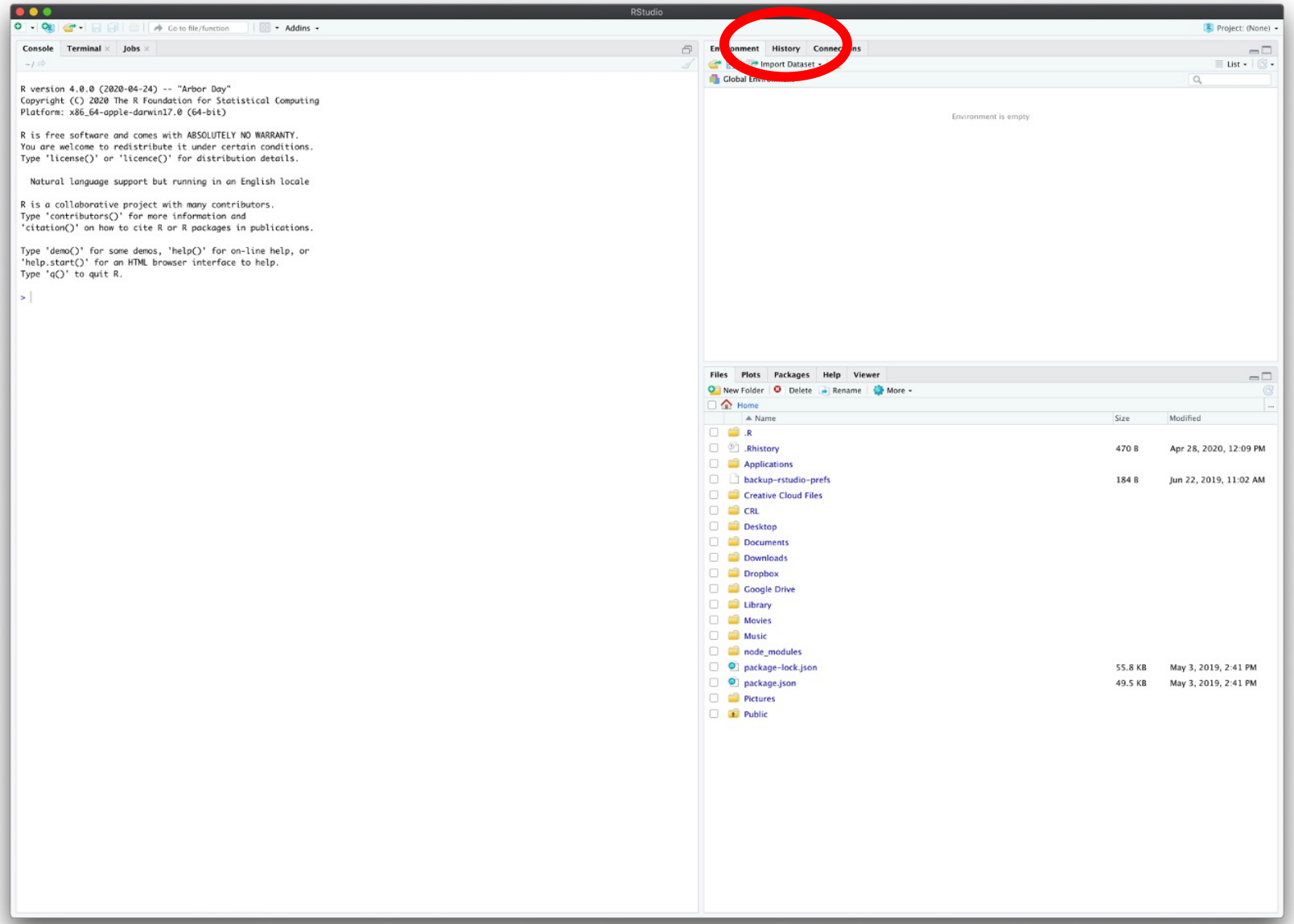
The default RStudio user interface.

RStudio


Console window:
Type a line of code and hit “enter”. Then R would run that line of code and give you the output
Environment window: It shows all the R objects loaded into R at the time
File window: It shows the files in your working directory

RStudio

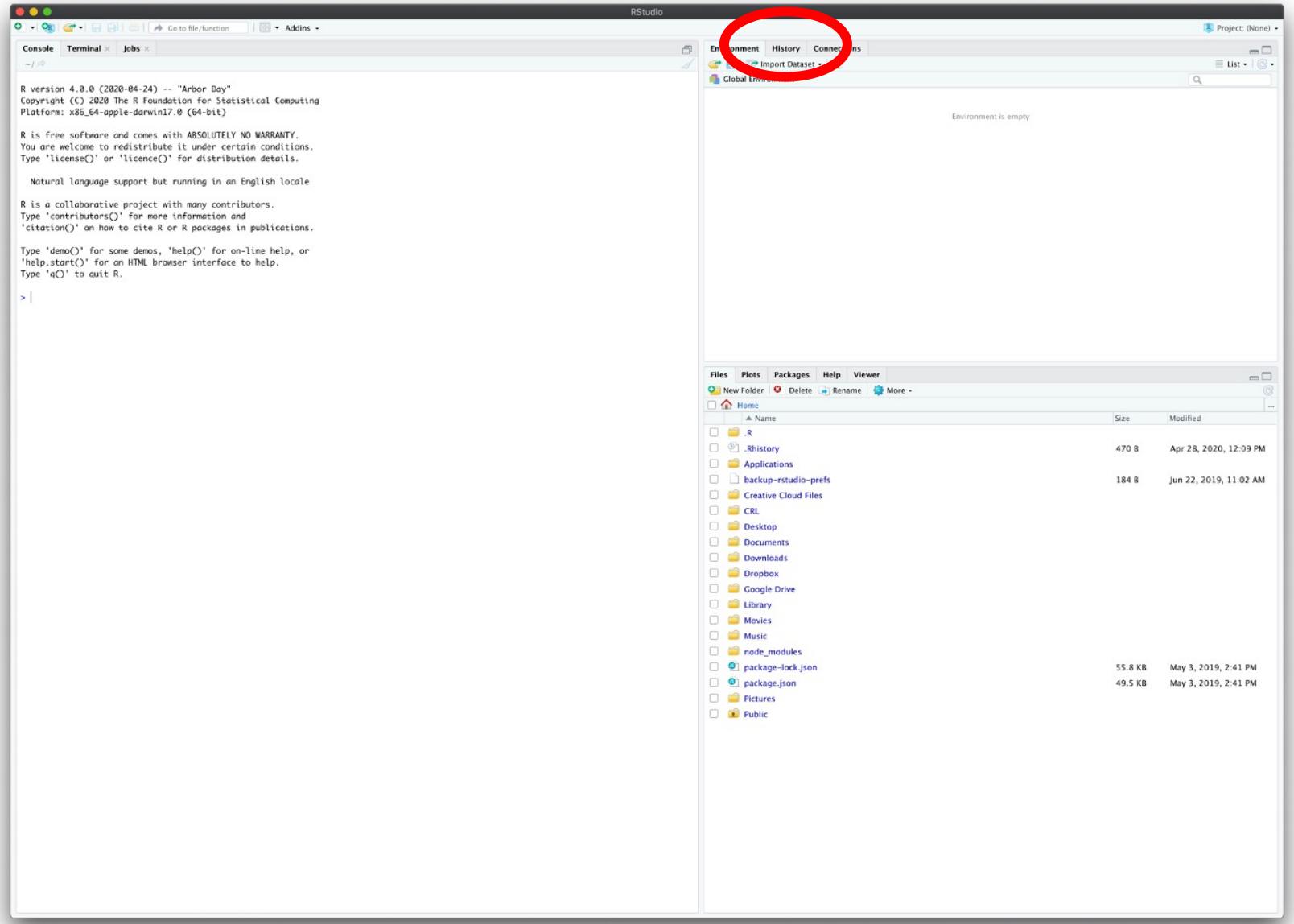
Click the History Tab




```
R version 4.0.0 (2020-04-24) -- "Arbor Day"
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.

>
```

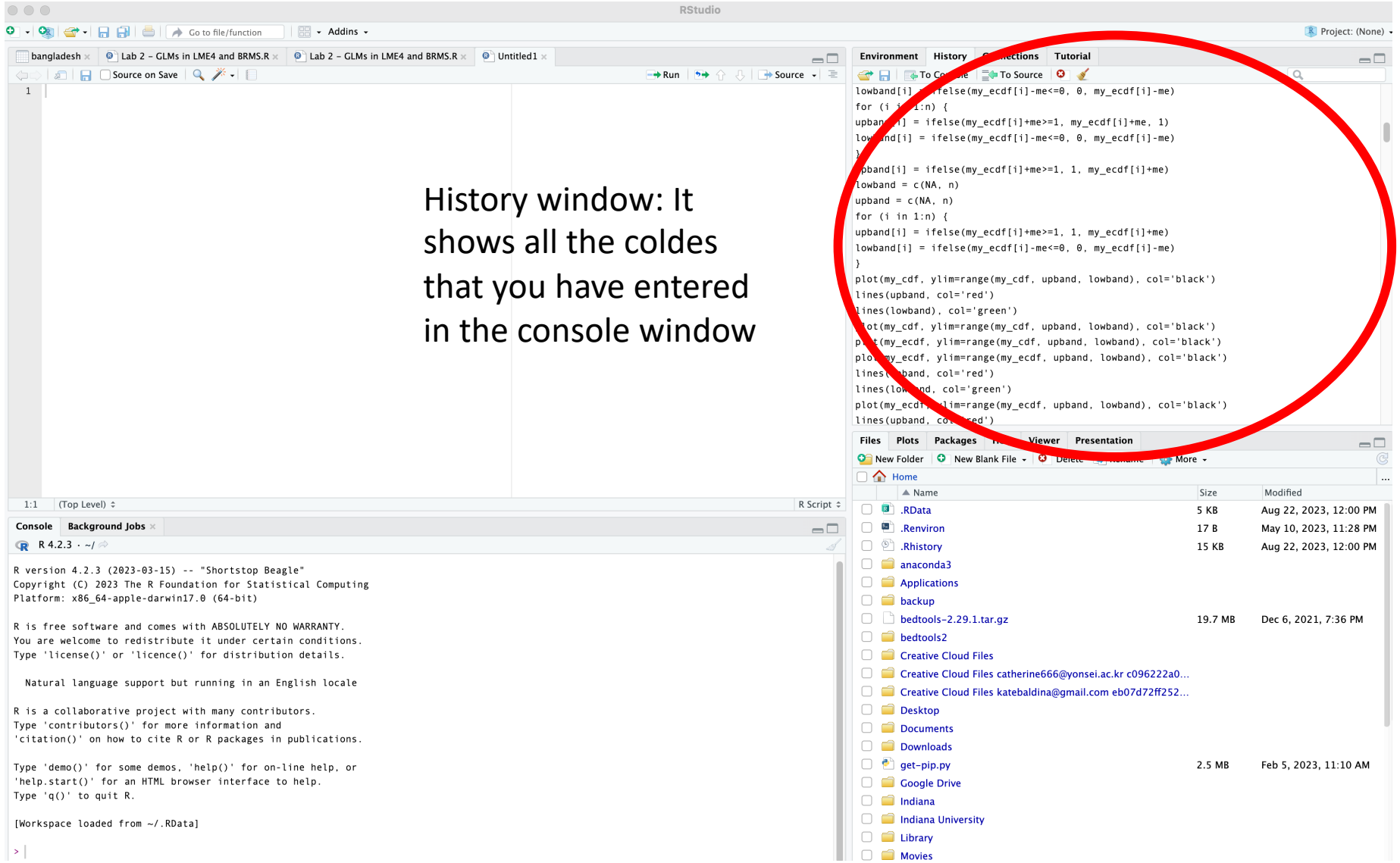
RStudio

Click the History Tab

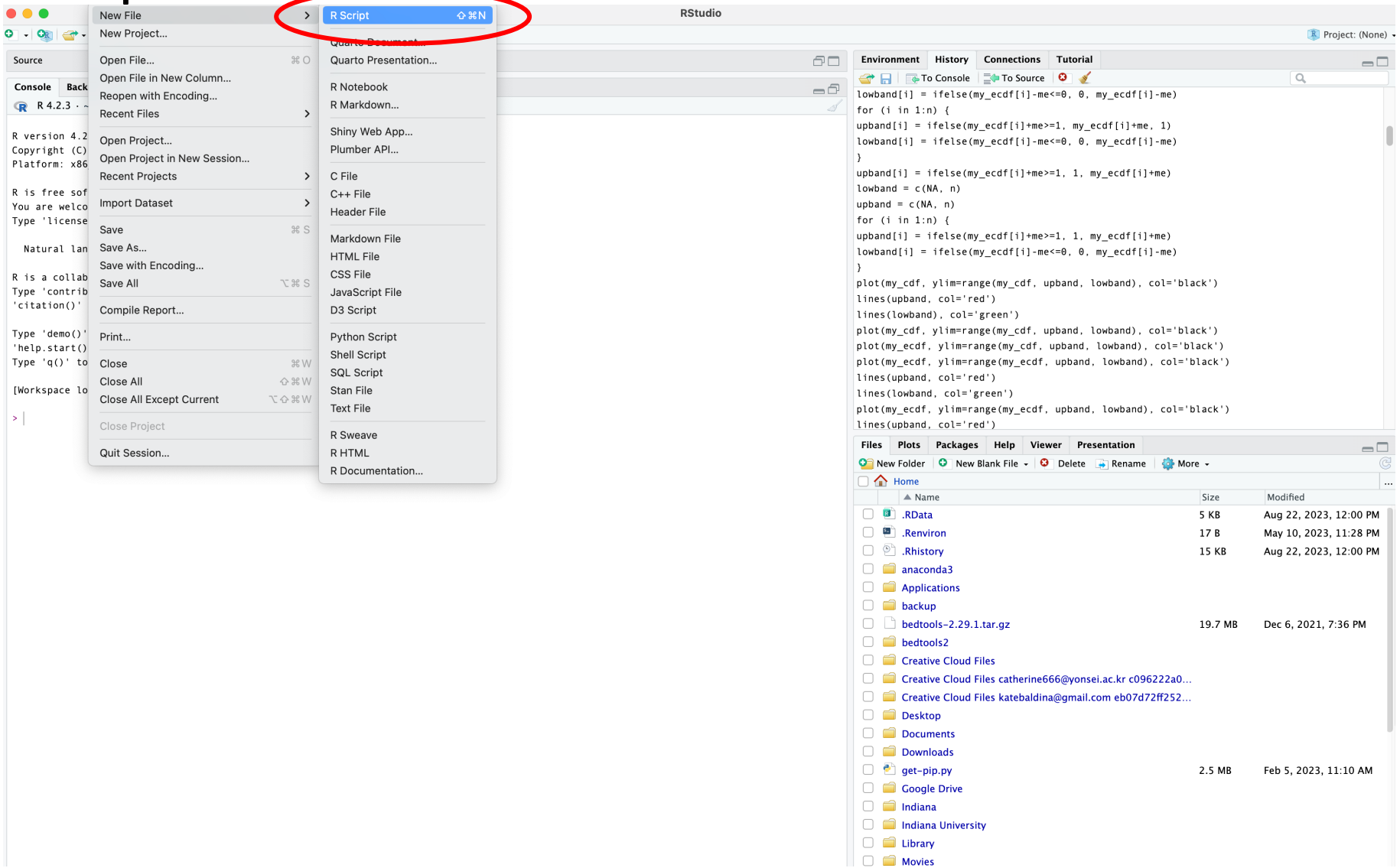



```
R version 4.0.0 (2020-04-24) -- "Arbor Day"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (64-bit)

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.

>
```

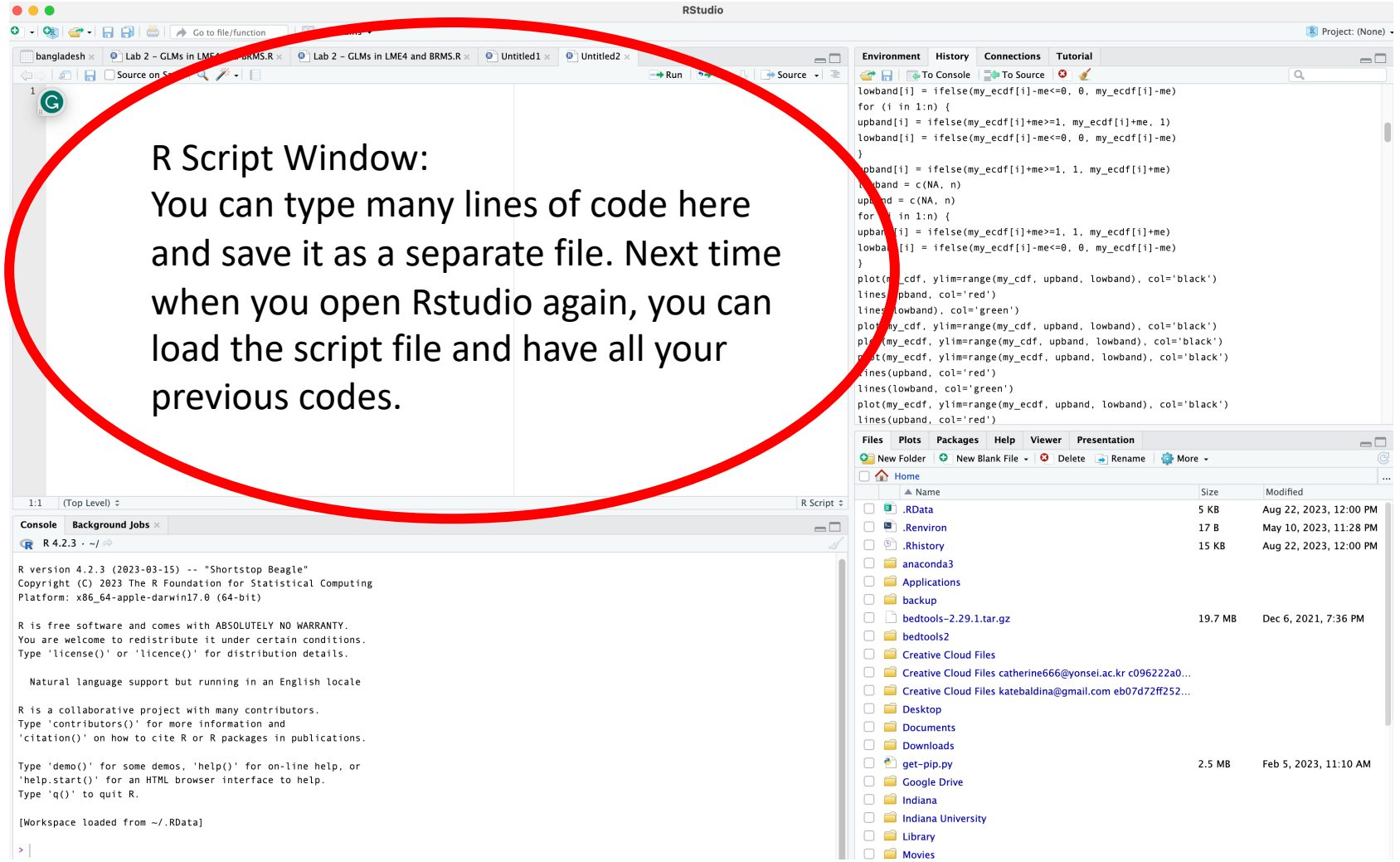
RStudio


RStudio
Go to file/function
Addins
Project: (None)
bangladesh
Lab 2 – GLMs in LME4 and BRMS.R
Lab 2 – GLMs in LME4 and BRMS.R
Untitled1
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
History window: It shows all the coldes that you have entered in the console window
Environment
History
Connections
Tutorial
To Console
To Source
lowband[i] = ifelse(my_ecdf[i]-me<=0, 0, my_ecdf[i]-me)
for (i in 1:n) {
upband[i] = ifelse(my_ecdf[i]+me>=1, my_ecdf[i]+me, 1)
lowband[i] = ifelse(my_ecdf[i]-me<=0, 0, my_ecdf[i]-me)
}
upband[i] = ifelse(my_ecdf[i]+me>=1, 1, my_ecdf[i]+me)
lowband = c(NA, n)
upband = c(NA, n)
for (i in 1:n) {
upband[i] = ifelse(my_ecdf[i]+me>=1, 1, my_ecdf[i]+me)
lowband[i] = ifelse(my_ecdf[i]-me<=0, 0, my_ecdf[i]-me)
}
plot(my_cdf, ylim=range(my_cdf, upband, lowband), col='black')
lines(upband, col='red')
lines(lowband), col='green')
plot(my_cdf, ylim=range(my_cdf, upband, lowband), col='black')
plot(my_ecdf, ylim=range(my_cdf, upband, lowband), col='black')
plot(my_ecdf, ylim=range(my_ecdf, upband, lowband), col='black')
lines(upband, col='red')
lines(lowband, col='green')
plot(my_ecdf, ylim=range(my_ecdf, upband, lowband), col='black')
lines(upband, col='red')
Files
Plots
Packages
Viewer
Presentation
New Folder
New Blank File
Delete
More
Home
Name
Size
Modified
.RData 5 KB Aug 22, 2023, 12:00 PM
.Renviron 17 B May 10, 2023, 11:28 PM
.Rhistory 15 KB Aug 22, 2023, 12:00 PM
anaconda3
Applications
backup
bedtools-2.29.1.tar.gz 19.7 MB Dec 6, 2021, 7:36 PM
bedtools2
Creative Cloud Files
Creative Cloud Files catherine666@yonsei.ac.kr c096222a0...
Creative Cloud Files katebaldina@gmail.com eb07d72ff252...
Desktop
Documents
Downloads
get-pip.py 2.5 MB Feb 5, 2023, 11:10 AM
Google Drive
Indiana
Indiana University
Library
Movies
Console
Background Jobs
R 4.2.3 · ~/
R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 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.
[Workspace loaded from ~/.RData]
>

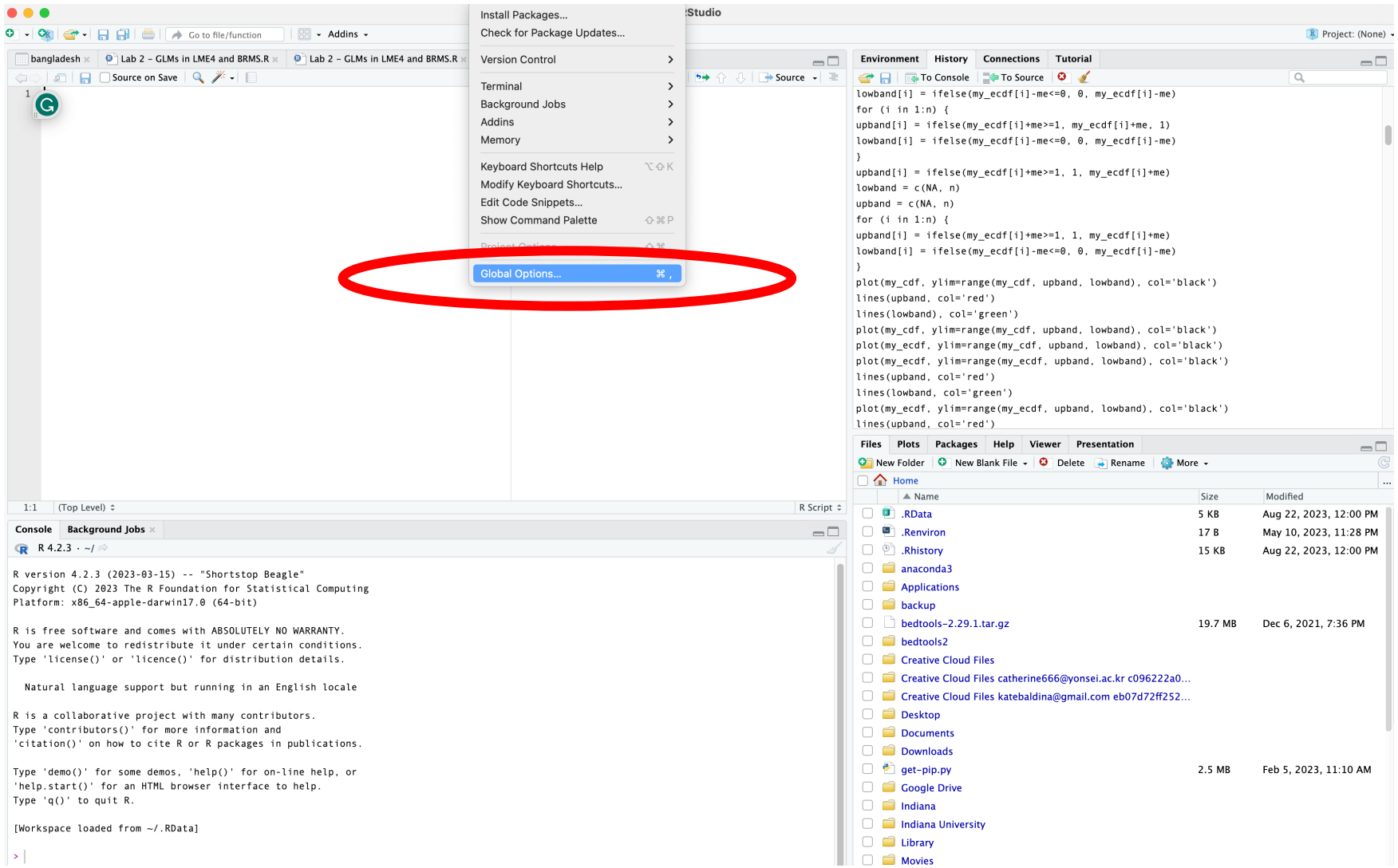
R Script Window



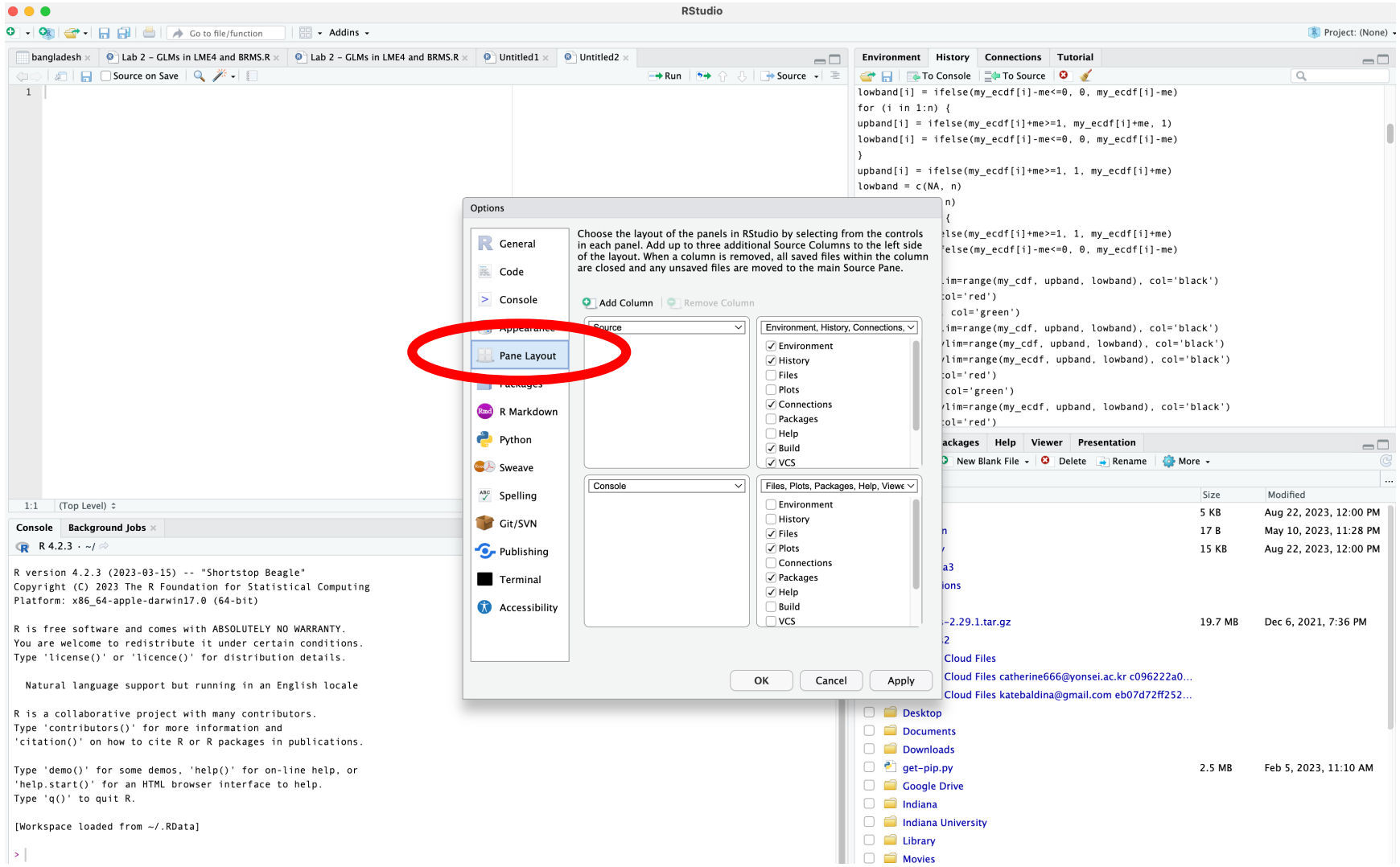
R Script Window



Rearrange windows in RStudio



Rearrange windows in RStudio



Command/Code in R

- A complete line of code tells R what to do:
  - Which function to use
  - What data/object you are referring to
- A line of code in R stats with the name of the function with open and close parentheses
- It looks like this:

XXX()

The name of the function

- Within the parenthses, you need to type the data/object you are referring to

Command/Code in R

- An example:

```
> mean(1:8)
[1] 4.5
>
```

  - mean() is the function
  - It calculates the mean of the numbers in the parentheses
  - 1:8 means all integers from 1 to 8 (inclusive)
  - So the entire line of code tells R to calculate the mean (average) of integers from 1 to 8 → this is the data that R need to work on
  - The output is 4.5

R packages

- R consists of packages
  - Similar to apps on the phone. One package do certain things. Another package can do other things.
  - Each package has one or more commands to be used
- After installing R, you only have the base package (just like pre-installed apps on your phone)
  - The base package gives you some basic functions (e.g., calculate the mean, load a data file in .txt or .cvs format)

R packages

- Whenever you need to use a function beyond the base package, you need to install the package and load it into R (just like opening an app before using its functions)
- To install packages in R, run the following command
  - install.packages('XXX')
- To load an installed package into R, run the following command
  - library(XXX)

Getting help in RStudio

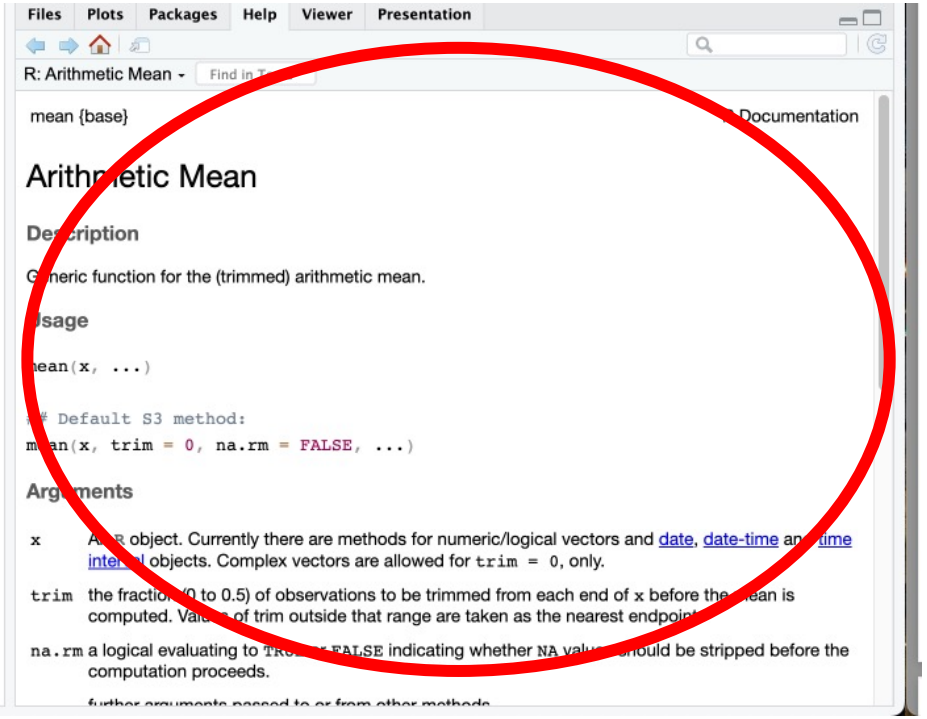
- You can always type “?” at the beginning of a command that you want to get help
- For example,

```
> ?mean
>
```

- This line of code asks R to pull out the documentation of the function “mean”

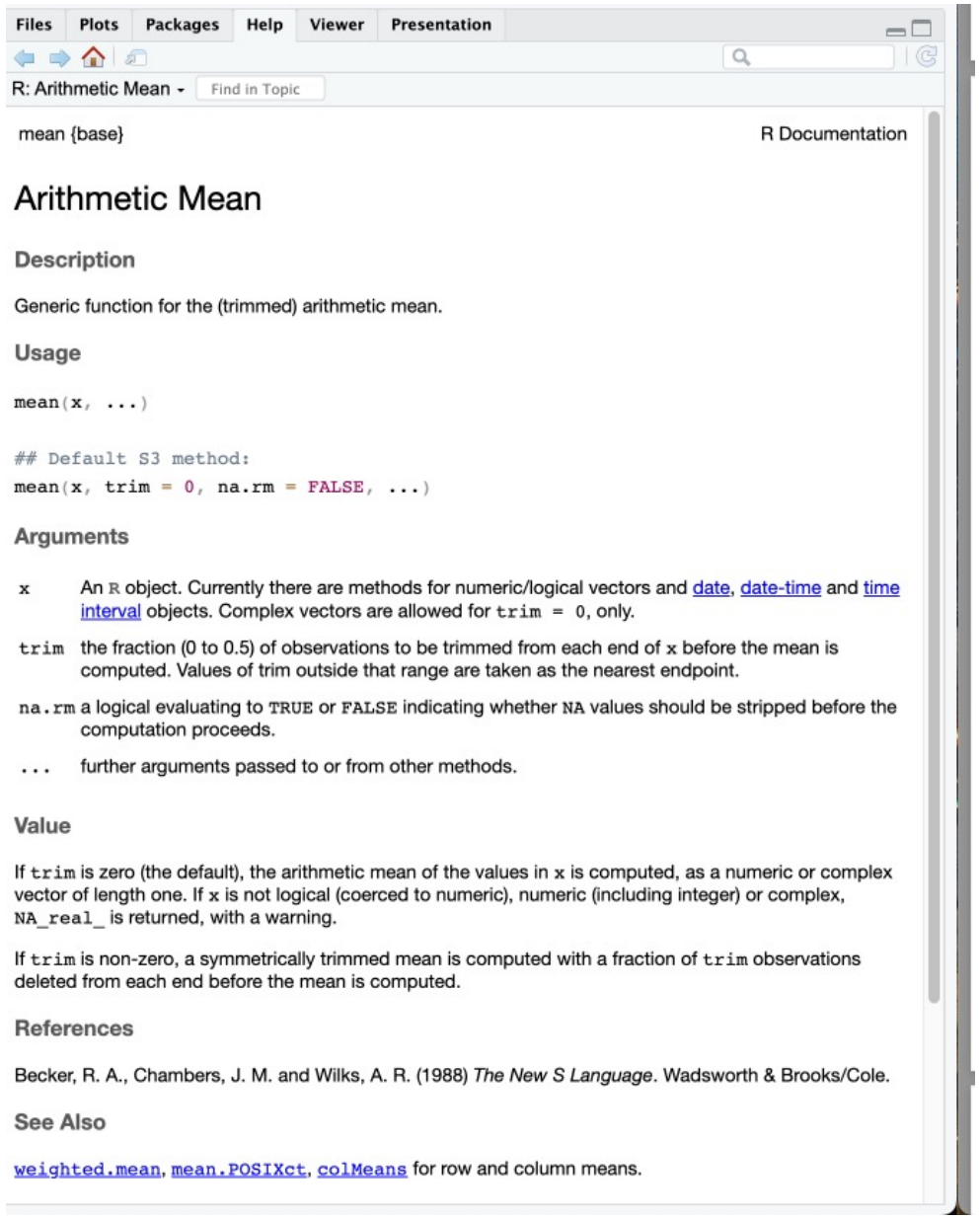
Getting help in RStudio

```
> ?mean
>
```



Here is the documentation explaining the mean function in R

Getting Help in RStudio



Getting Help

- You can always ask me or Professor Schultz about R programming
- Also, try googling the command/function in R that you want to know

- Some resources:
  - https://stackoverflow.com/questions/tagged/r
  - R IDE Cheatsheet (will be posted in Canvas)
  - https://education.rstudio.com/learn/beginner/