

CSEP 2021: Data Science for the Modern Exercise Physiologist

Pre-workshop instructions

Congratulations on starting your journey into R and RStudio!

This instruction page will provide you specific details on how to install the free programs before the workshop, as well as provide some activities to familiarize yourself with the interface if you have some additional time.

By the end of these instructions, you will have:

1. Downloaded and installed both R and RStudio
2. Installed and opened the tidyverse library package
3. Written one line of code in a script – congrats, you’re now a hacker-in-training!

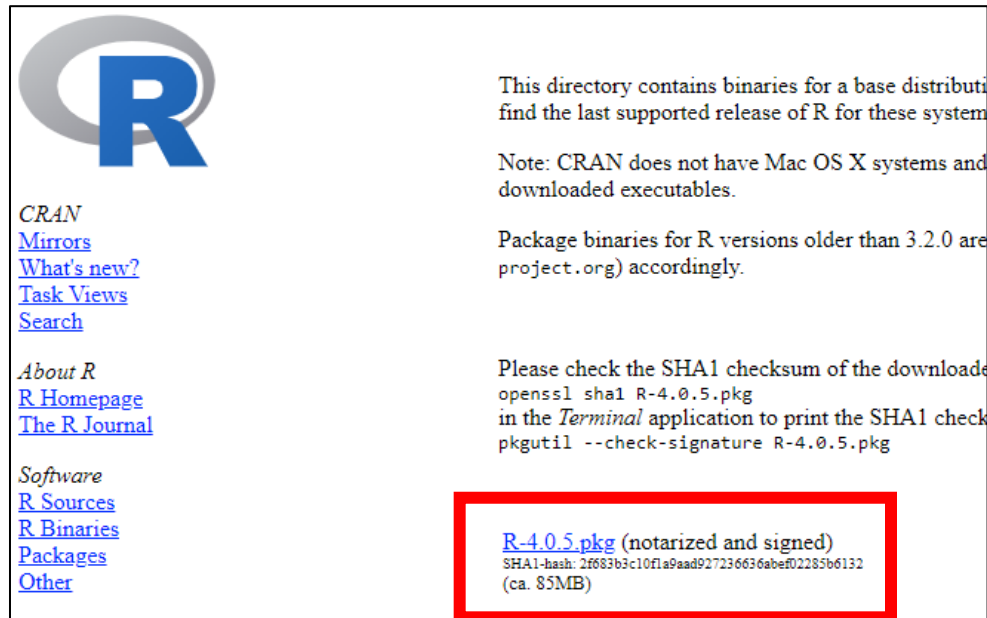
Step 1: Terminology-at-a-glance

R	The actual coding ‘language’ in which you will be learning. R is Free Software, meaning that the user (you) has much greater freedom in how you choose to use the program, which has led to amazing community-built statistical and plotting packages that are free to use and continually maintained.
RStudio	The interface that a separate company (RStudio) has created to make it easier to navigate around the R language. While there are paid versions available for large companies, the basic version of RStudio is freely available for individual download and use.
Function	A specialized coding command that takes some inputs (e.g., x-axis, y-axis, grouping), and creates an output (e.g., a plot).
Package	A group of functions that have been created by one group/person and has been made freely available to share and download online.
Library	When you want to open a package in your RStudio window, it will be opened into your active library. This library is cleared each time you open RStudio.
Script	A group of code that can be saved, worked on later, and shared like any other document file (e.g., like a .doc or .xlsx, but it’s called a .R).

Step 2: Download R and RStudio

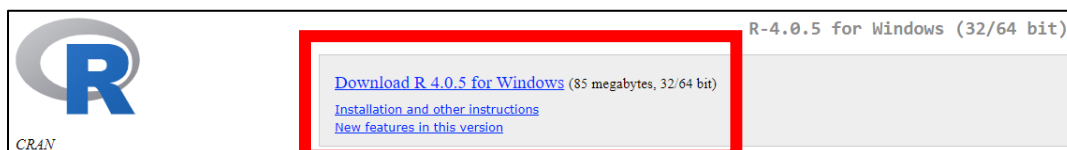
1. **First, let’s download R.** Head to the following page: <https://cran.rstudio.com/>
 - a. If you are downloading for Mac OS X, click the Download R for (Mac) OS X link. At the time of writing these instructions, the latest release version is R-4.0.5.pkg.
 - i. If you are a visual person, here is a quick video walking you through the install:

https://www.youtube.com/watch?v=lcawuhf0Yqo&ab_channel=RogerPeng



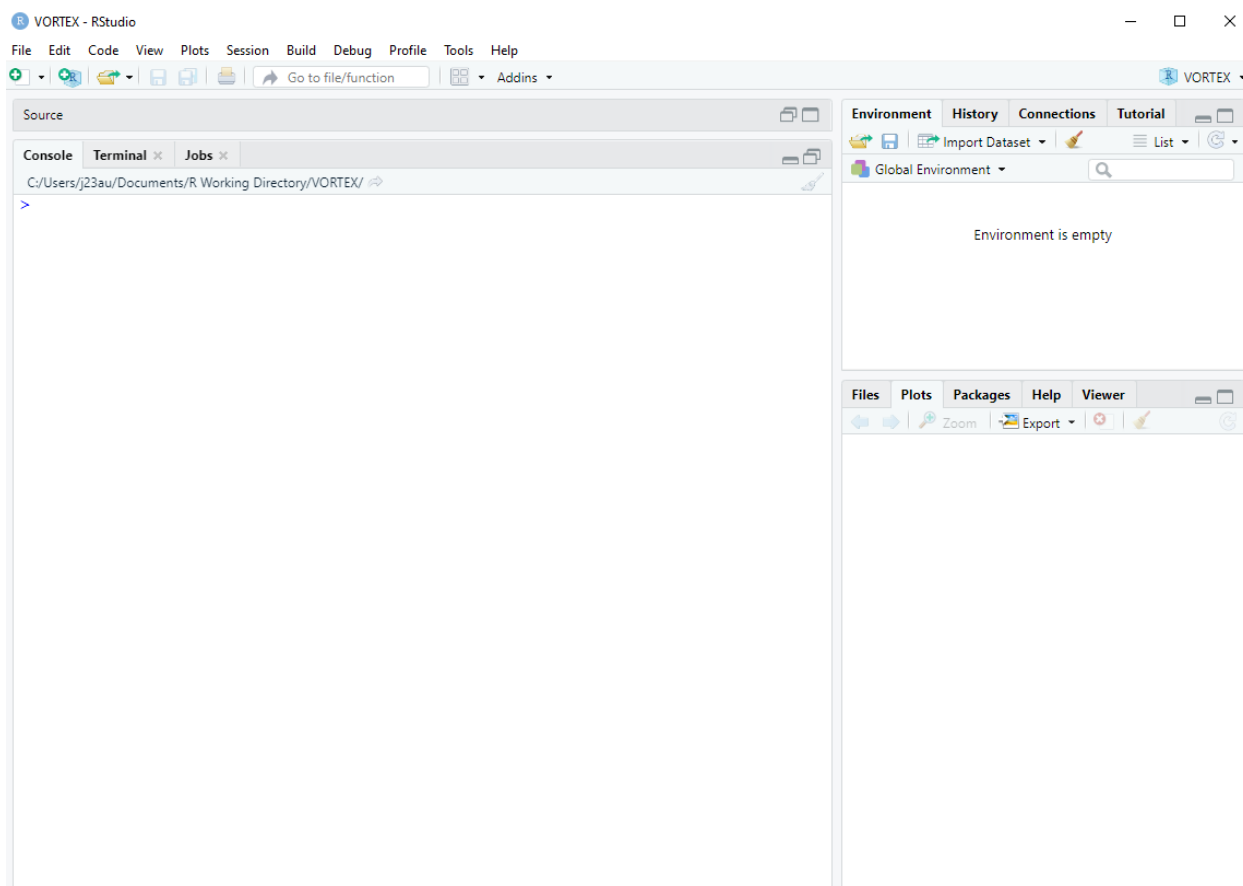
The screenshot shows the CRAN R Binaries page. On the left, there is a sidebar with the CRAN logo and links: [CRAN](#), [Mirrors](#), [What's new?](#), [Task Views](#), [Search](#), [About R](#), [R Homepage](#), [The R Journal](#), [Software](#), [R Sources](#), [R Binaries](#), [Packages](#), and [Other](#). The main content area on the right contains the following text: "This directory contains binaries for a base distribution. To find the last supported release of R for these systems, see the [CRAN website](#)." Below this, a note states: "Note: CRAN does not have Mac OS X systems and does not provide downloaded executables." Further down, it says: "Package binaries for R versions older than 3.2.0 are not available (see [project.org](#)) accordingly." A section titled "Please check the SHA1 checksum of the downloaded package" provides instructions: "openssl sha1 R-4.0.5.pkg" and "in the *Terminal* application to print the SHA1 checksum" and "pkgutil --check-signature R-4.0.5.pkg". At the bottom, a red box highlights the download link: [R-4.0.5.pkg](#) (notarized and signed), with the SHA1 hash: 2f683b3c10f1a9aad927236636abef02285b6132 (ca. 85MB).

- b. If you are downloading for PC, click the Download R for PC link. On the next page, click the first link called 'base' OR 'install R for the first time', which will both bring you to the same page. At the time of writing these instructions, the latest release version is R 4.0.5.
 - i. Choose the 32-bit version if indicated – this should be fine for most systems.
 - ii. If you are a visual person, here is a quick video walking you through the install: https://www.youtube.com/watch?v=mfGFv-iB724&ab_channel=RogerPeng
- c. NOTE: You will never actually open this base version of R, but it needs to be installed to be able to use RStudio.



The screenshot shows the R-4.0.5 for Windows (32/64 bit) download page. On the left, there is a sidebar with the CRAN logo and links: [CRAN](#), [R-4.0.5 for Windows](#), [Installation and other instructions](#), and [New features in this version](#). The main content area on the right contains the following text: "Download R 4.0.5 for Windows (85 megabytes, 32/64 bit)". A red box highlights the download link: [Download R 4.0.5 for Windows](#) (85 megabytes, 32/64 bit).

2. **Next, we need to download RStudio.** Head to the following page: <https://www.rstudio.com/products/rstudio/download/>
 - a. Select the Download option for the free RStudio Desktop. **YOU DO NOT NEED THE PAID VERSION**; this is for large organizations only.
 - b. The download page will automatically detect if you have a PC or Mac and will offer the appropriate download and install.
3. If everything has gone well, when you open RStudio, it should look like this:



Step 3: Install the required packages

1. In the bottom left quadrant of the RStudio screen, you will find the Console area.
Type the following line of code next to the > symbol:

```
> install.packages("tidyverse", dependencies = T)
```

- a. This will take some time to install; if RStudio is working on something, you will see this 'Stop sign' in the top right of the Console window.
 - b. If you are given a 'y/n' (i.e., yes/no) prompt to add an additional or missing package, type y into the Console area to continue.
2. This is the step where many different errors might occur depending on your computer. In most cases, someone else has also encountered the same problem and you **can likely Google possible solutions** (as with most problems with R). The following is a common list of errors you might encounter

Error	What it means	Solution
-------	---------------	----------

There is no package called 'xx'	You are missing a package to be able to install something	Usually installing that other package first will fix the problem. Typical command is <code>install.packages("[packagename]")</code>
Installation of package 'xx' had non-zero exit status	The package install ran into an unspecified problem	This usually isn't your fault, but something wrong with the webpage you are trying to download it from. Try the <code>install.packages</code> command again, or try again the next day.
Most other problems	Tidyverse is a larger combination of packages, so there are a lot of other things that could go wrong	Individually install the packages required for this tutorial: <code>install.packages("dplyr")</code> <code>install.packages("ggplot2")</code> <code>install.packages("tidyr")</code> <code>install.packages("readr")</code> <code>install.packages("purrr")</code> <code>install.packages("tibble")</code>

Step 4: Test out the library and create a script

1. **Create a new document**, and select 'Script'. This will open up a new window in the top left of RStudio
2. Quick code writing tips to get started
 - a. Use the `#` symbol to write comments. This will become coloured text and will allow you to write notes to yourself on that line of code, but will not be identified as 'code'
 - b. **Start your script by writing the following title:** `## Tidyverse Test Script`
 - c. **Give you script a datestamp:** `# Date: April 14, 2021`
 - d. **Indicate the authors of your script:** `Author: Jason Au`
3. For each package you want to use, you always need to load the associated libraries first to bring it into your RStudio environment. For this trial, try to open your tidyverse package: **`library(tidyverse)`**
 - a. A series of messages will run through the Console to indicate if you were able to load tidyverse properly. It will look like this:
4. The coding rite of passage: Enter these commands into your script to **write your first lines of code:**

```
x <- "Hello world!"
print(x)
```

Congratulations!

**You have successfully installed R
and RStudio and have started your
data science journey!**