Module 3: Installing Software

Week 0: Course Orientation

Module Preview

Hello! And welcome to Module 3: Installing Software.

In this module, we'll learn how to install all the software that we'll need for CSCI E-5a.

- In Section 1, we'll have an overview of the installation process.
- In Section 2, we'll install R.
- In Section 3, we'll install RStudio.
- In Section 4, we'll configure RStudio.
- In Section 5, we'll install the rmarkdown package, and knit an R notebook to an HTML document.
- In Section 6, we'll install tinytex, and knit an R notebook to a PDF document.

When you've completed this module, you'll be able to:

- Install R
- Install RStudio
- Configure RStudio
- Install the rmarkdown package
- Install tinytex
- Knit an R notebook to an HTML or PDF document

All right! Let's get started by having an overview of the installation process.

Section 1: Installation Overview

Main Idea: *We can provide an overview of the installation process*

In this section, we'll provide an overview of the installation process.

Before we get started, it's helpful to be clear about the difference between R and RStudio:

- R is a computer programming language that takes expressions, evaluates them, and returns an output value.
- RStudio is an integrated development environment (IDE) for R, which means that it provides a complete system for writing, testing, and executing R code.

RStudio features "R notebooks", which are documents that integrate text, code, output, and graphics, and which can be exported as a PDF. All of the lecture material for CSCI E-5a is distributed in this R notebook format, and your problem sets and examinations are also in this format. Our Gradescope grading system is configured to accept only submissions that result from an R notebook exported as a PDF. Thus, you must be able to work with R notebooks using RStudio, and in particular you must be able to render an R notebook to a PDF. Let's repeat that, just to be clear:

> In order to take CSCI E-5a: *Programming in R*, you **must** be able to render an R notebook to a PDF.

One of the primary goals of Week 0: Course Orientation is to make sure that everyone can work with R notebooks and render them to PDFs. That's why we have Problem Set 0, so that you can demonstrate that your system is set up properly and you can perform all the required technical operations to take the course.

By the way, although I've used the verb "render" to mean exporting an R notebook as a PDF, in the world of RStudio the preferred term is to "knit" the notebook. To be honest, I have no idea why this verb was chosen, but it's absolutely standard and is used throughout the R community. So alternatively we can say: in order to take CSCI E-5a: *Programming in R*, you **must** be able to knit an R notebook to a PDF.

For CSCI E-5a, I recommend that you install R and RStudio onto your local hard drive. The installation process is not particularly difficult, and most people find it to be straightforward. Once you've done this, then you'll have a complete R system installed on your system that you can use long after the course has ended. You'll also be able to extend and update your system when new versions are released. Finally, once you become familiar with the installation process you'll be able to install R and RStudio on a new computer.

However, installing R and RStudio onto your local hard drive is not your only option. RStudio also offers a cloud-based version of the IDE called RStudio Cloud that you can use for all your work in CSCI E-5a. If you are affiliated with an institution or company that has a license for RStudio Cloud then you are welcome to use that for the course. However, even if you have access to RStudio Cloud, then I still encourage you to install R and RStudio on your local hard drive. After all, you might not always have access to the service, and you might be in situations where you don't necessarily have Internet access but would still like to use R and RStudio. So, I encourage everyone to install the software on their hard drives, even if they have access to RStudio Cloud. However, it's not required, and if you want to use RStudio Cloud for the course, then that's fine too. Just remember: no matter what system you ultimately adopt, you **must** be able to render R notebooks into PDFs.

Section 2: Installing R

Main Idea: *We can install R on a local hard drive*

In this section, we'll see how to install R on a local hard drive.

The first step is to go to the R Project website, which is located at www.r-project.org. A simple way to do this is to simply search for the letter "r":

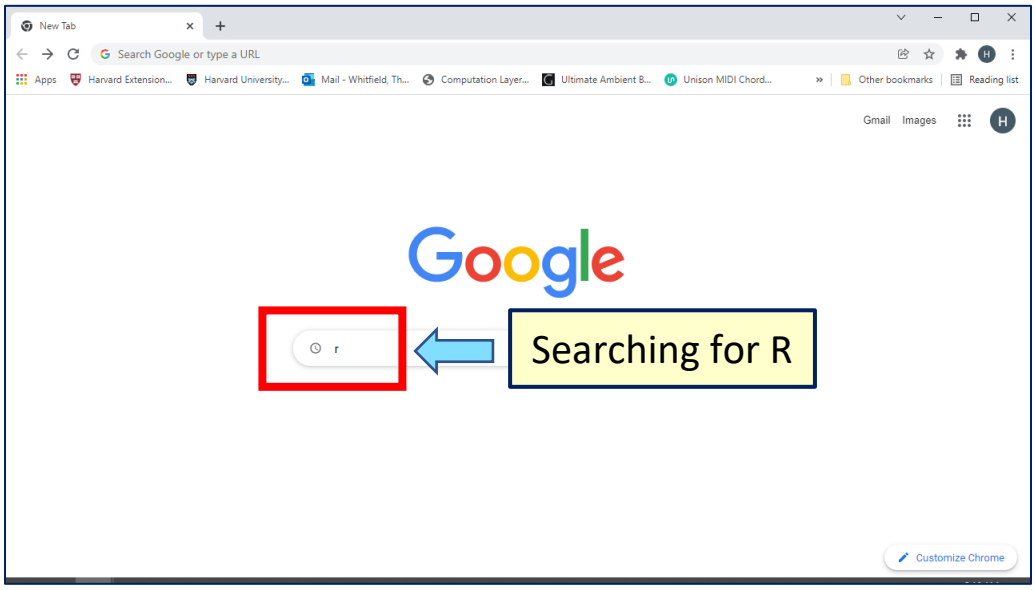

Figure 1: Searching for the R Project website

The R Project for Statistical Computing is the first item in my search results:

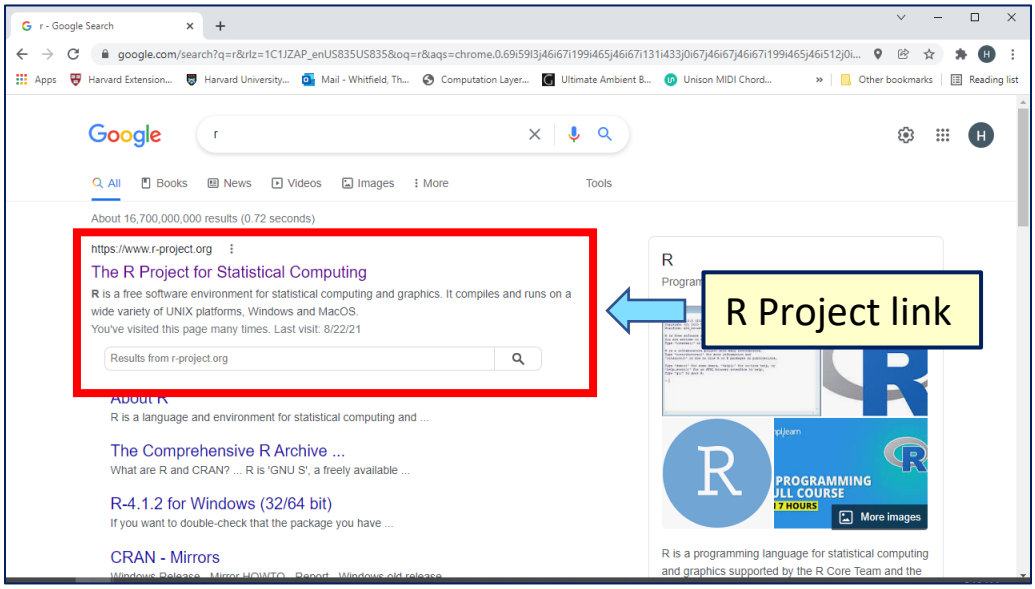

Figure 2: Search results for R Project website

When you click on this link, you arrive at the R Project website:

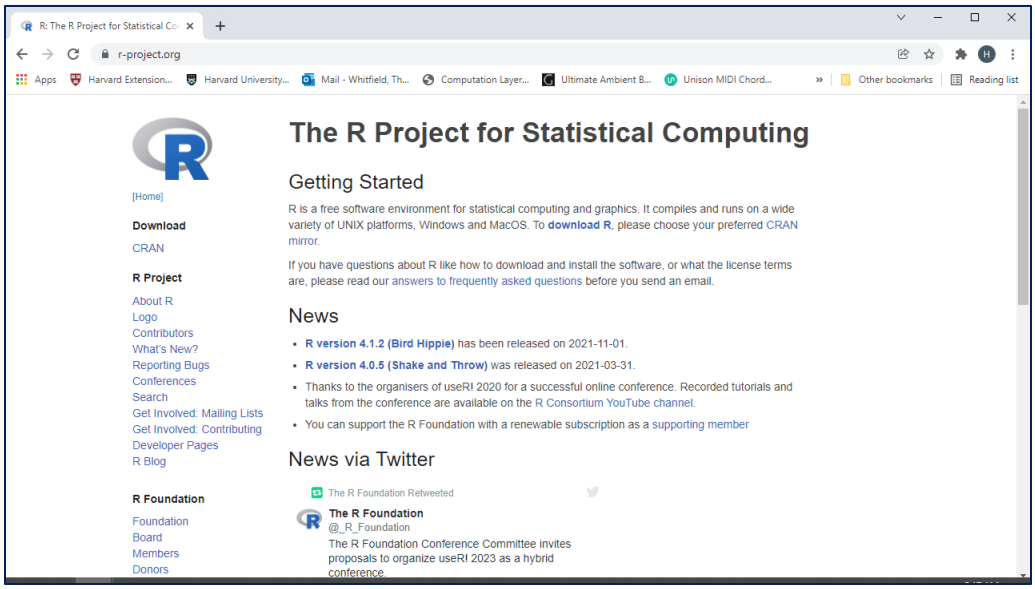

Figure 3: R Project website home page

The link for download is at the top of the left-hand vertical menu:

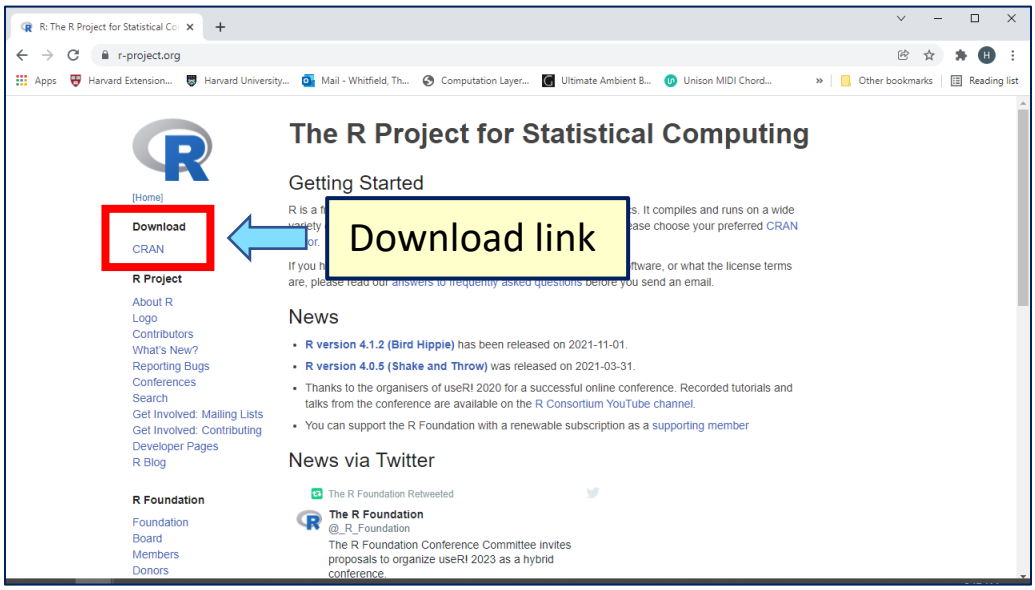

Figure 4: Download link

Let's take a closer look:

Figure 5: Download link (close-up)

When you click on this link, you'll go to a page to select a "mirror", which is basically just a server from which you can download R. The very top link on the page is supposed to automatically direct you to a local server, so I recommend clicking on this:

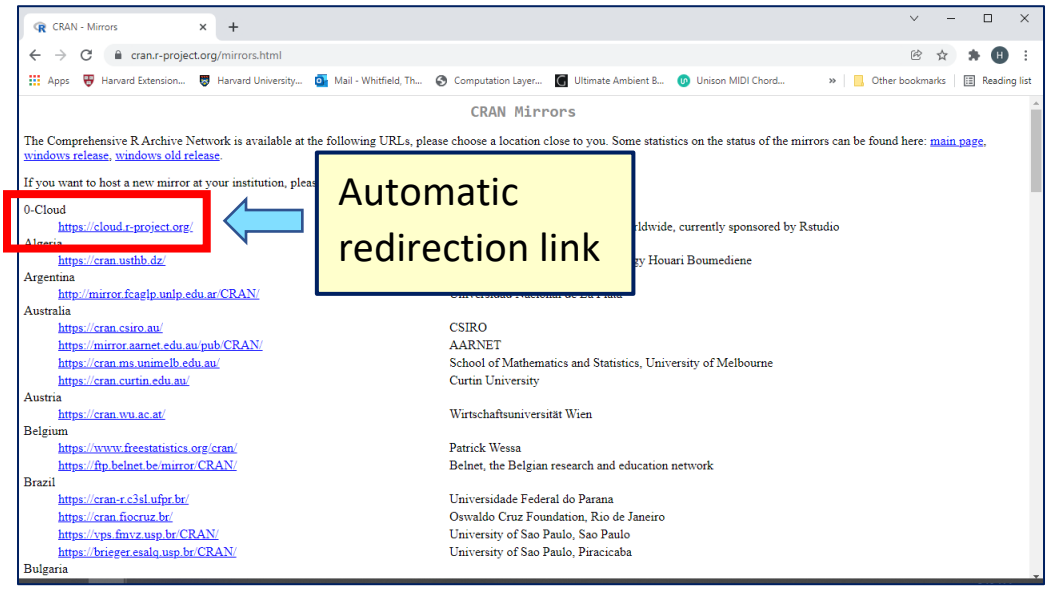

Figure 6: Selecting a mirror site

Finally, you'll arrive at the page for downloads:

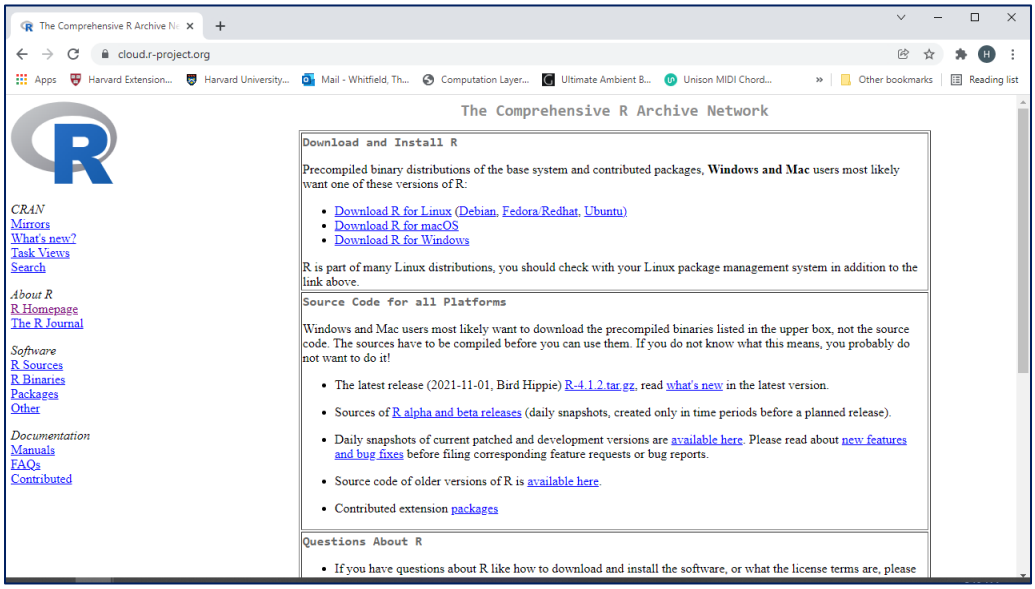

Figure 7: Main download page

For Windows users, click on the third link from the top:

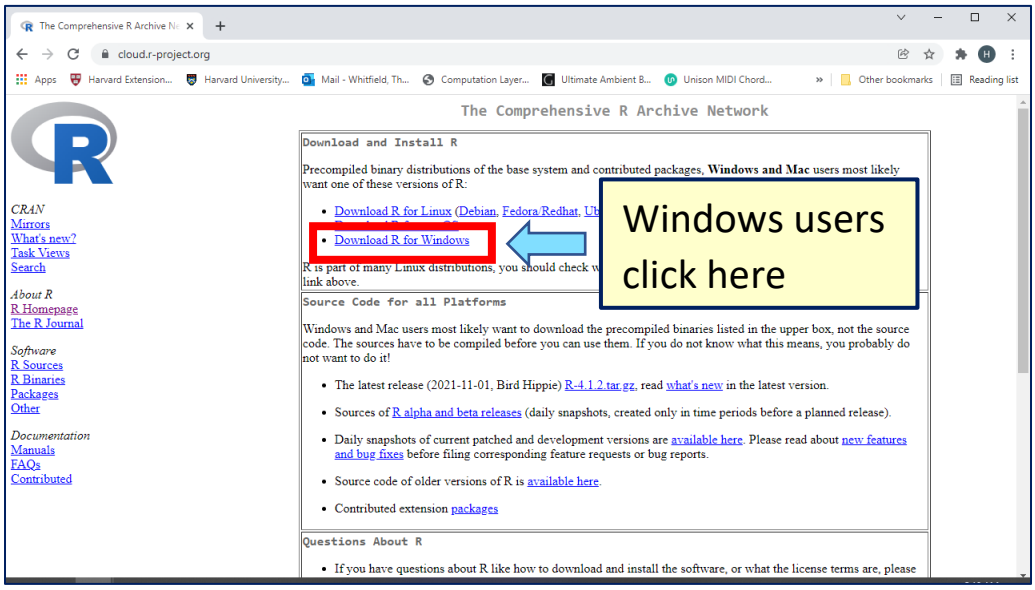

Figure 8: Windows download page link

Next, you'll be presented with a set of choices; if you're installing for the first time, you should select "base":

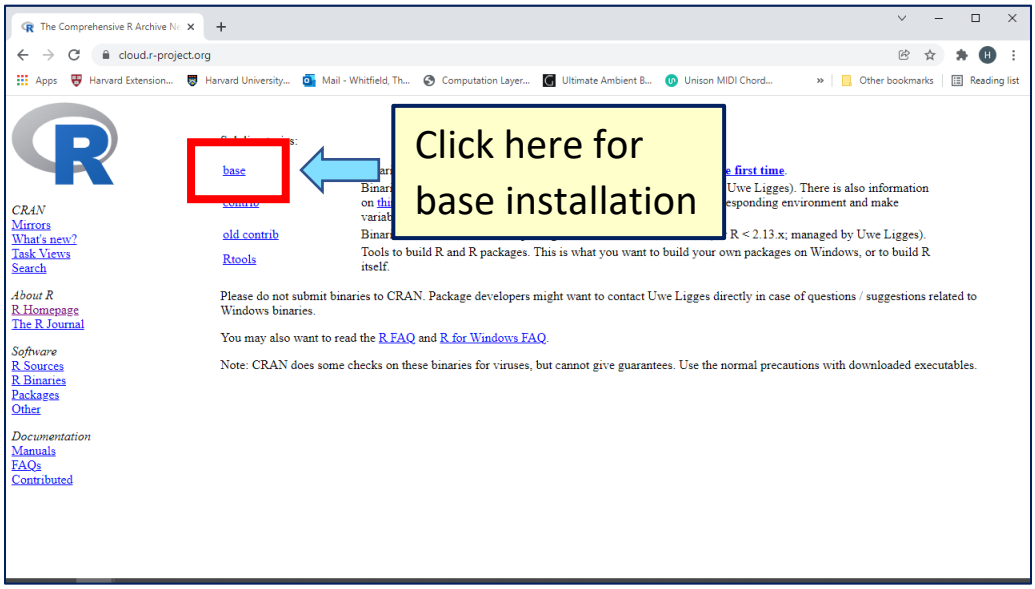

Figure 9: Windows download options

Finally, you'll get to the download page, and you can click on the link at the top to download the installer:

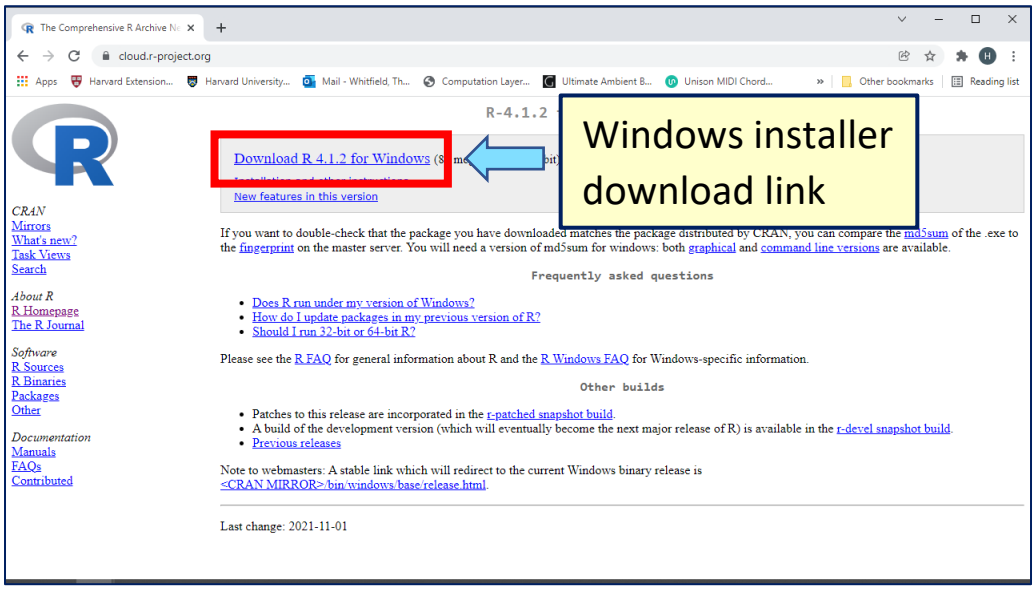

Figure 10: Windows download file link

For Mac users, let's start again from the download page:

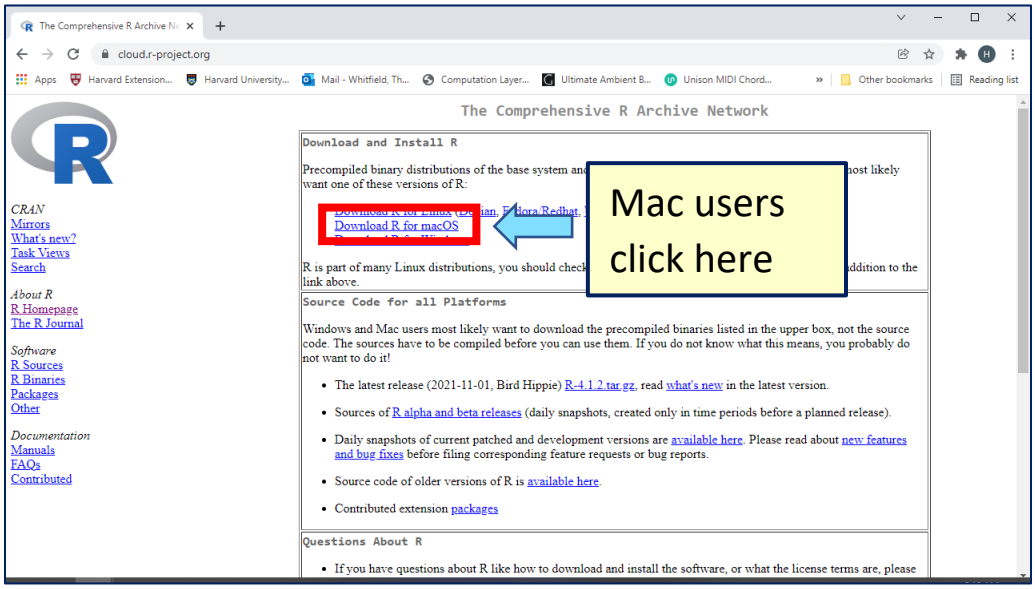

Figure 11: Mac download page link

The download page for Macs is:

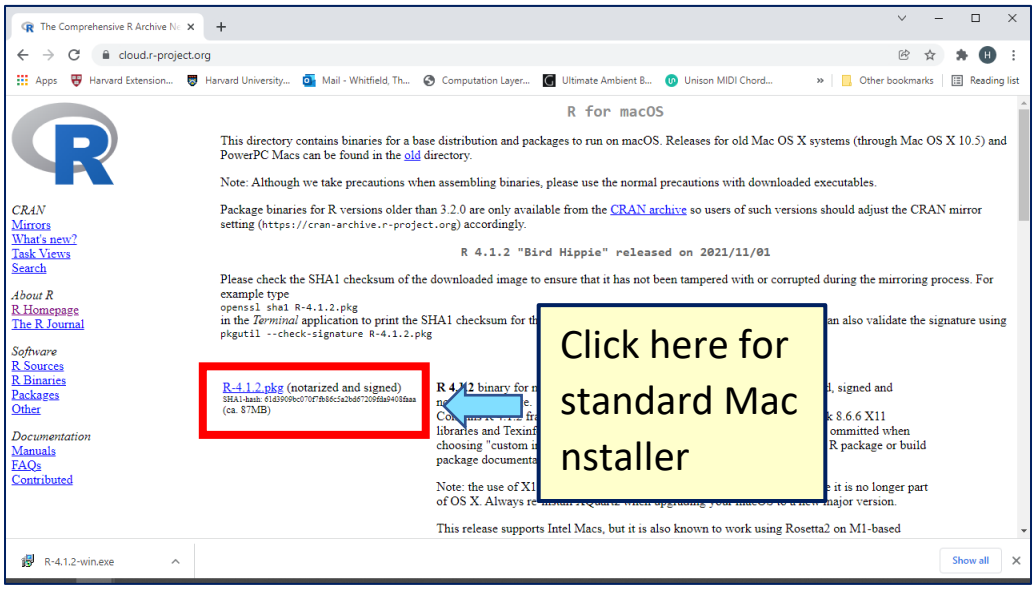

Figure 12: Mac download page

The installation process for Macs is usually fine, but occasionally there can be issues. Note that the current release is for macOS 10.13 (High Sierra), which was released in September 2017, so if you have an older system this might be an issue. I've also heard of some issues with the new M1 Macs, but I don't have any further information. If you're a Mac user and you're having problems with this installation, contact the teaching staff and we'll try to help you.

Once you've downloaded the installer, run it to install R:

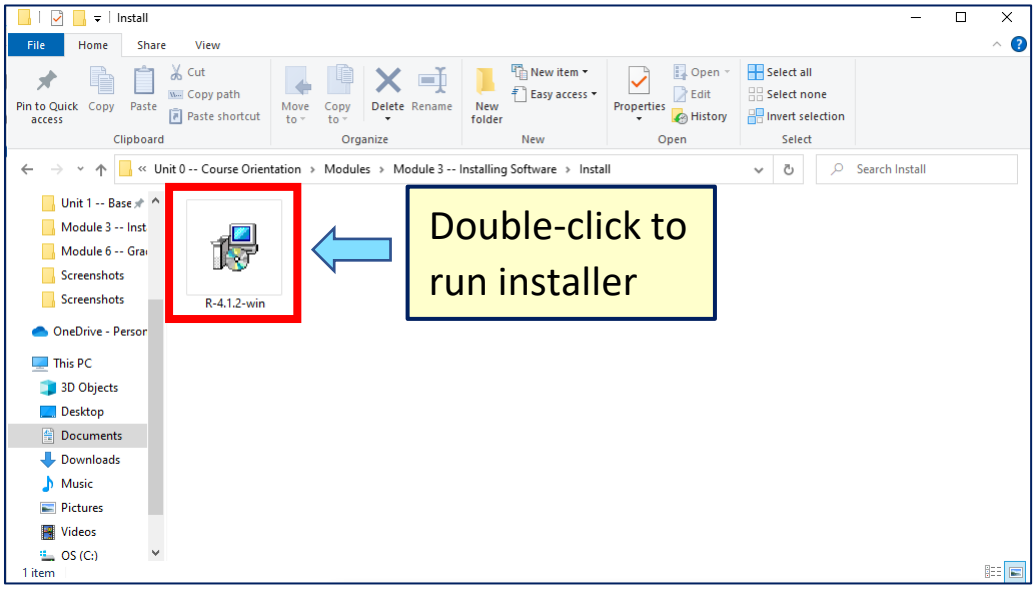

Figure 13: Running the installer

I recommend accepting all the defaults unless you have a very very very good reason for not doing so.

When the installation process is done, the R computation engine will be installed, as well as a simple user interface called "RGui". In Windows, you can see this in the Start menu:

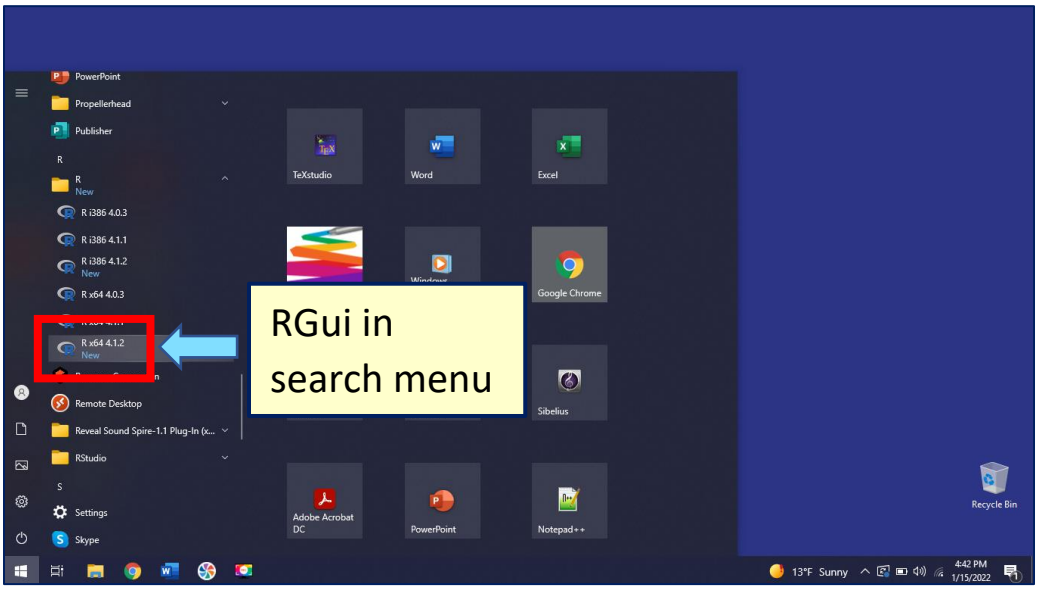

Figure 14: RGui in Start menu

When you run the RGui program, you'll see a simple interface:

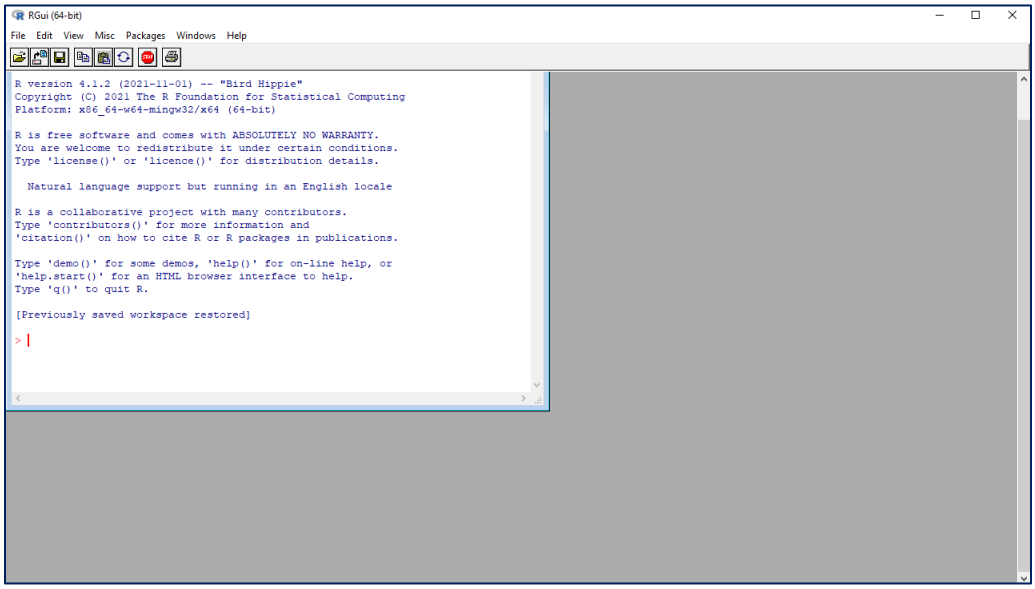

Figure 15: RGui display

While the RGui interface has limited functionality, it can nonetheless be quite useful in some situations. We won't be using this program because we will be focused on RStudio, which provides many more features, but it's good to know about the RGui program, and there are situations where it's the preferred approach.

For our purposes, the appeal of the RGui program is that if you can run it then that indicates that the R installation process was successful.

So that's how to install R on your local hard drive.

Now let's see how to install RStudio.

Section 3: Installing RStudio

Main Idea: *We can install RStudio on a local hard drive*

In this section, we'll see how to install RStudio on a local hard drive.

The first step is to go to the RStudio website at www.rstudio.com. You can search for "RStudio" in a search engine:

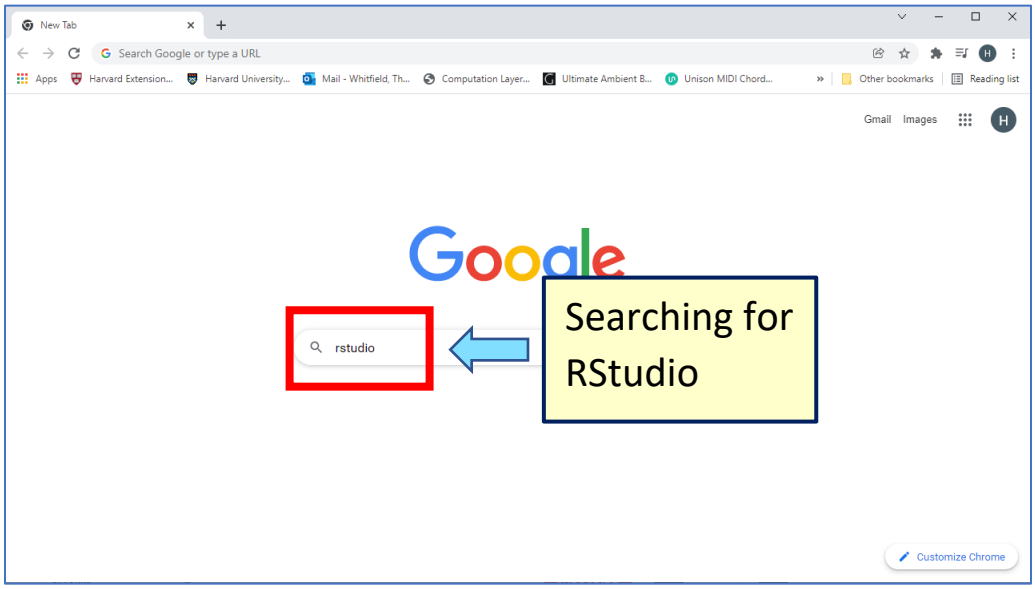

Figure 16: Searching for RStudio

RStudio is the first result of this search:

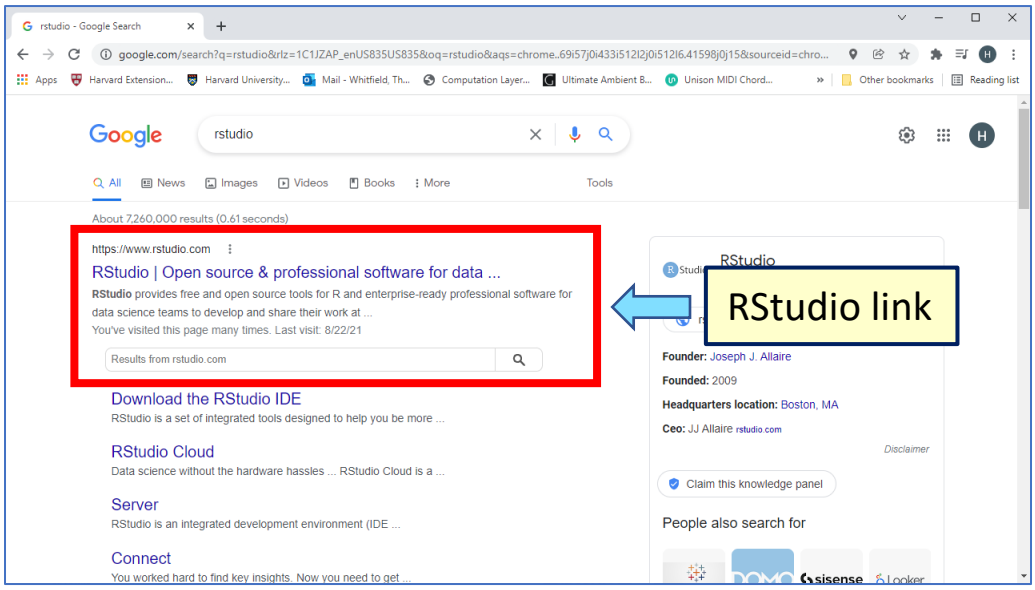

Figure 17: RStudio search results

Clicking on the link takes you to the website, and the link for the Download page is the first item in the very top menu:

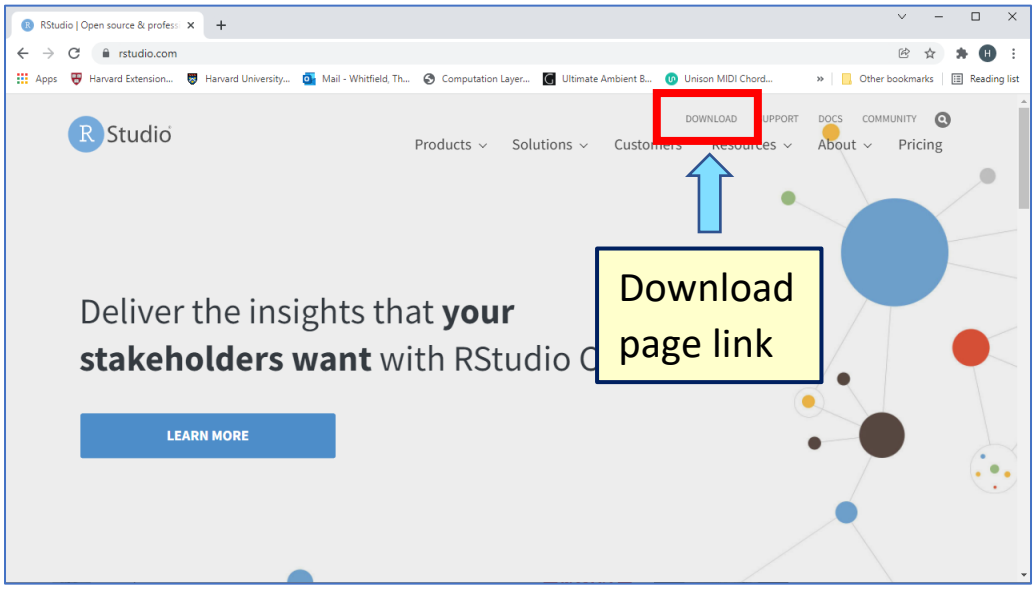

Figure 18: RStudio home page and Download page link

Let's take a close-up look at that:

Figure 19: RStudio home page and download link (close-up)

Clicking on this link takes you to the Downloads page:

Figure 20: RStudio downloads page

Scrolling down, you have a choice of options, and you should download the RStudio Desktop version, which is the left-most option:

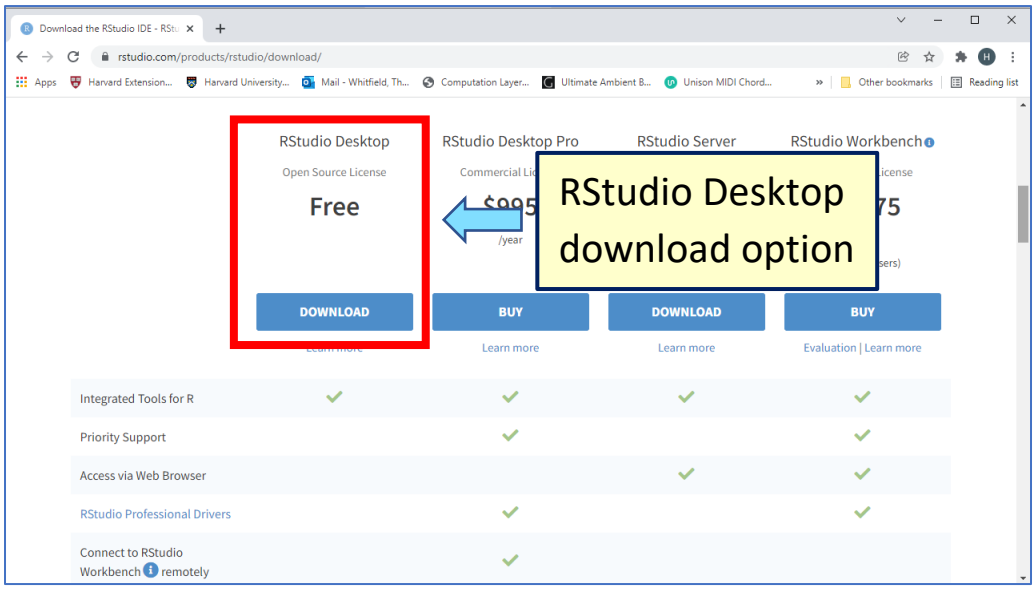

Figure 21: RStudio Desktop download option

When you click on the download link, the website detects your local system and makes an automatic recommendation:

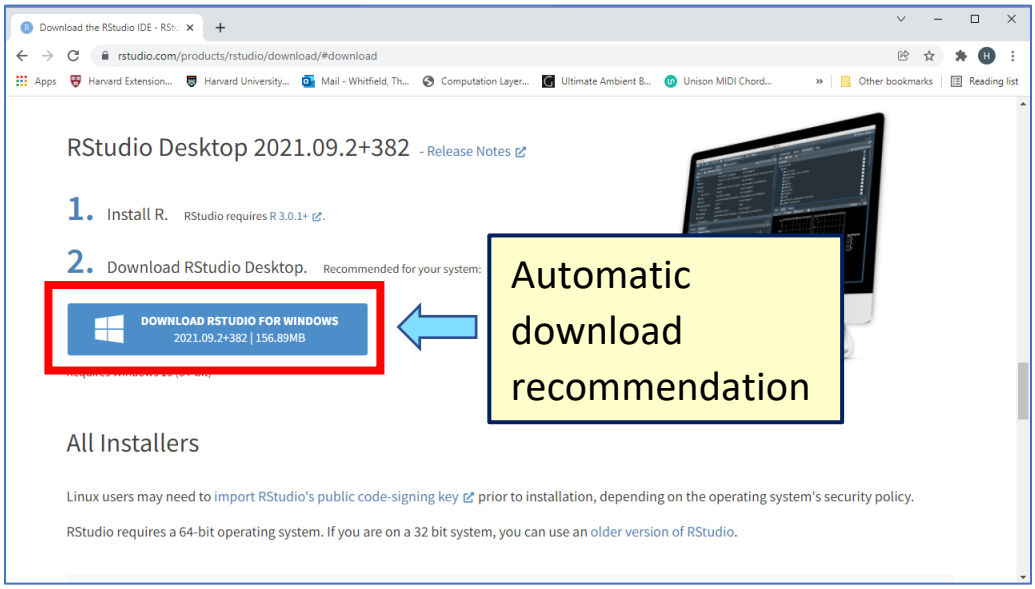

Figure 22: Automatic download recommendation

If for any reason you don't want to use the automatic suggestion, you can scroll down to find a complete list of available installers:

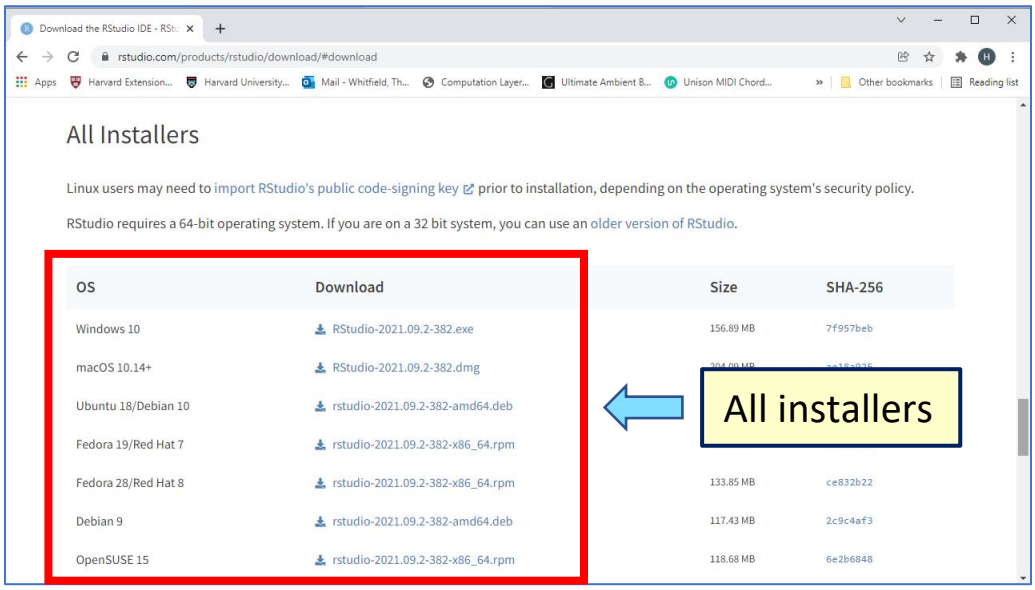

Figure 23: List of all available installers

Once again, click on the link for the download, and save the file to your hard disk. Then run the installation process; I recommend accepting all the defaults unless you have a very very very good reason for not doing so.

So that's how to install RStudio on your local hard drive.

Now let's learn how to configure RStudio.

Section 4: Configuring RStudio

Main Idea: *We can configure a number of options in RStudio*

In this section, we'll see how to configure RStudio.

Let's launch RStudio:

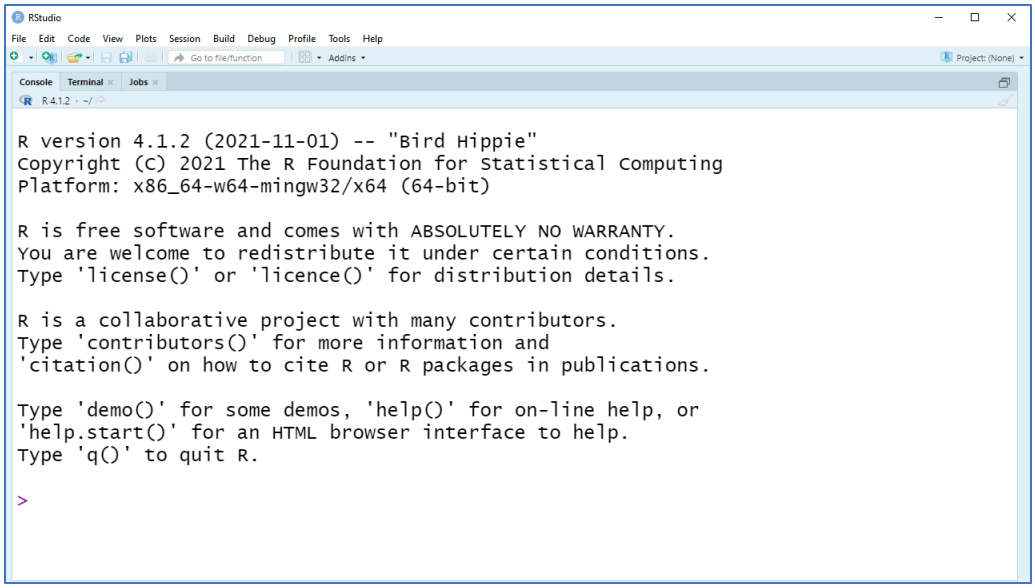

Figure 24: RStudio initial display

First, let's go to the Tools item in the top menu bar:

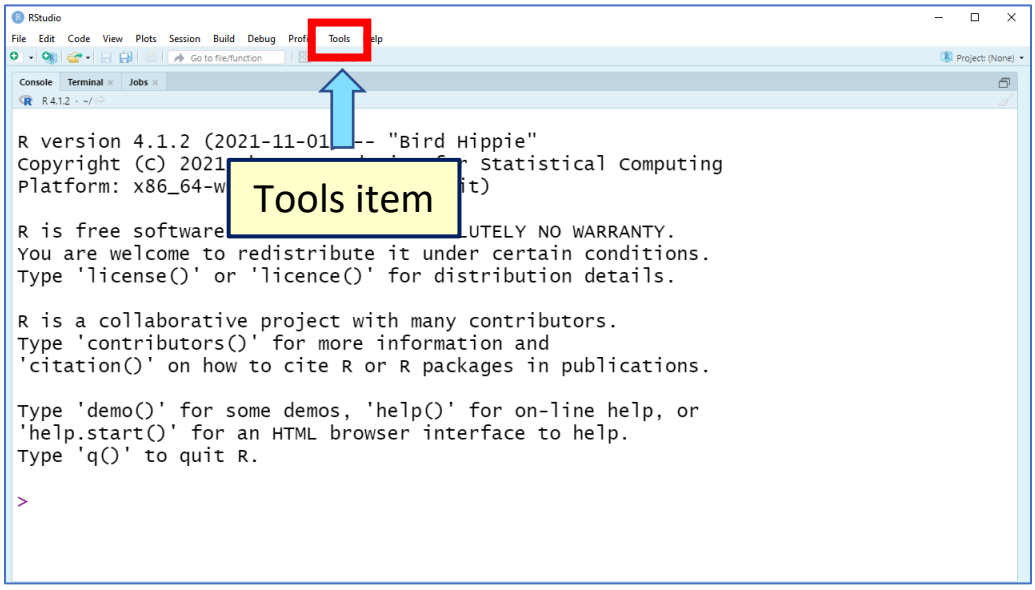

Figure 25: RStudio Tools menu item

When you click on this link, you'll get a drop-down menu, and you should click on the entry at the very bottom named "Global Options":

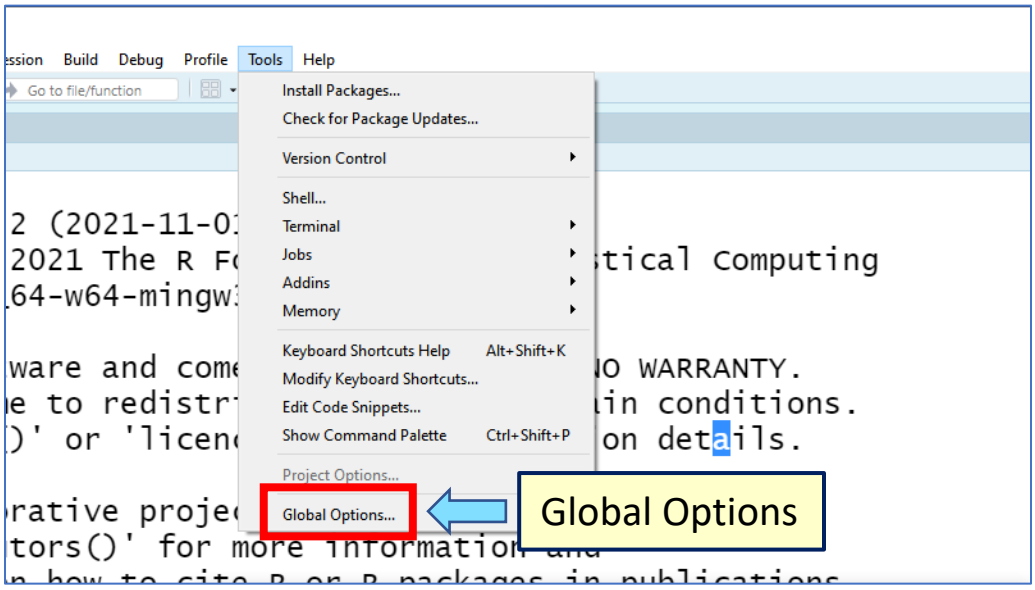

Figure 26: Global Options selection

The Global Options dialog box offers a large number of options for you to customize RStudio:

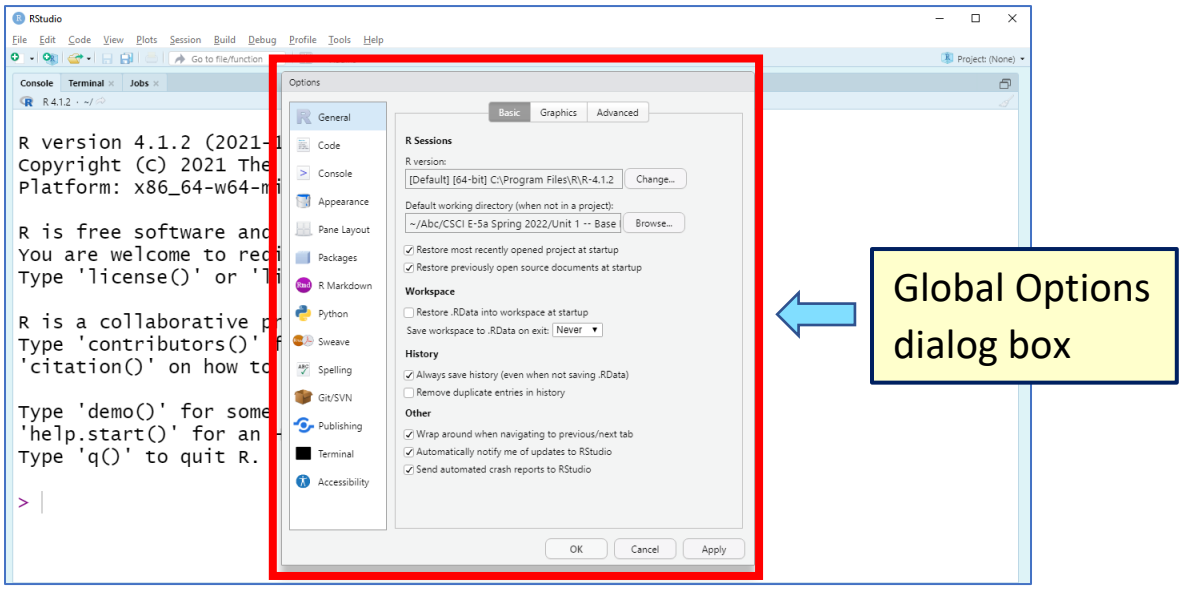

Figure 27: RStudio Global Options menu

The top entry for the Global Options is "General", and here I suggest changing the "Current working directory" to the CSCI E-5a folder:

To do this, first click on the "Browse" button for this option:

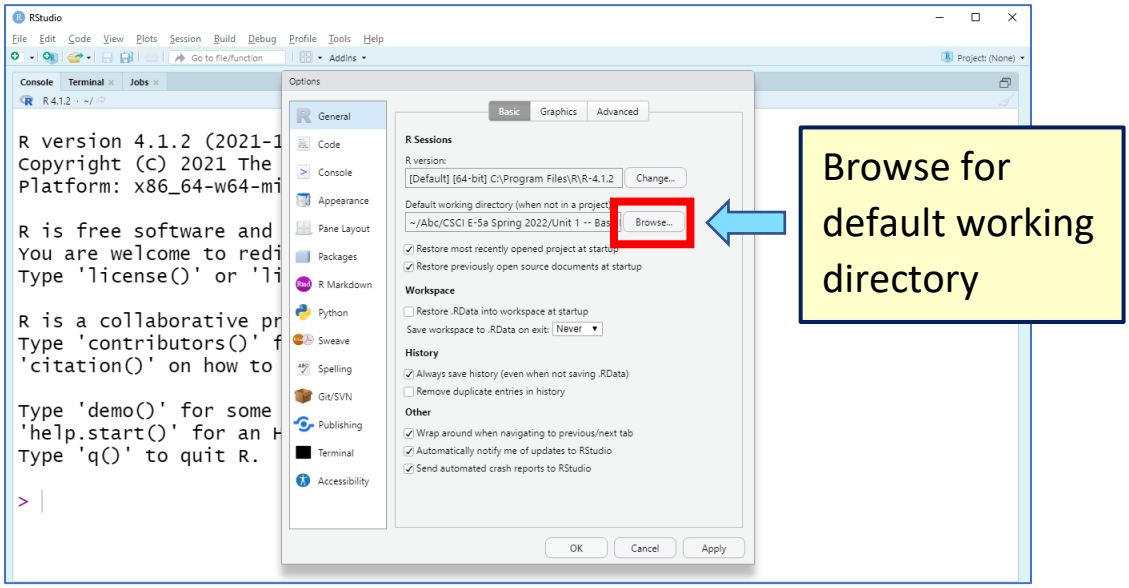

Figure 28: Browsing for the default working directory

Then you'll get a dialog box that you can use to navigate to the CSCI E-5a folder:

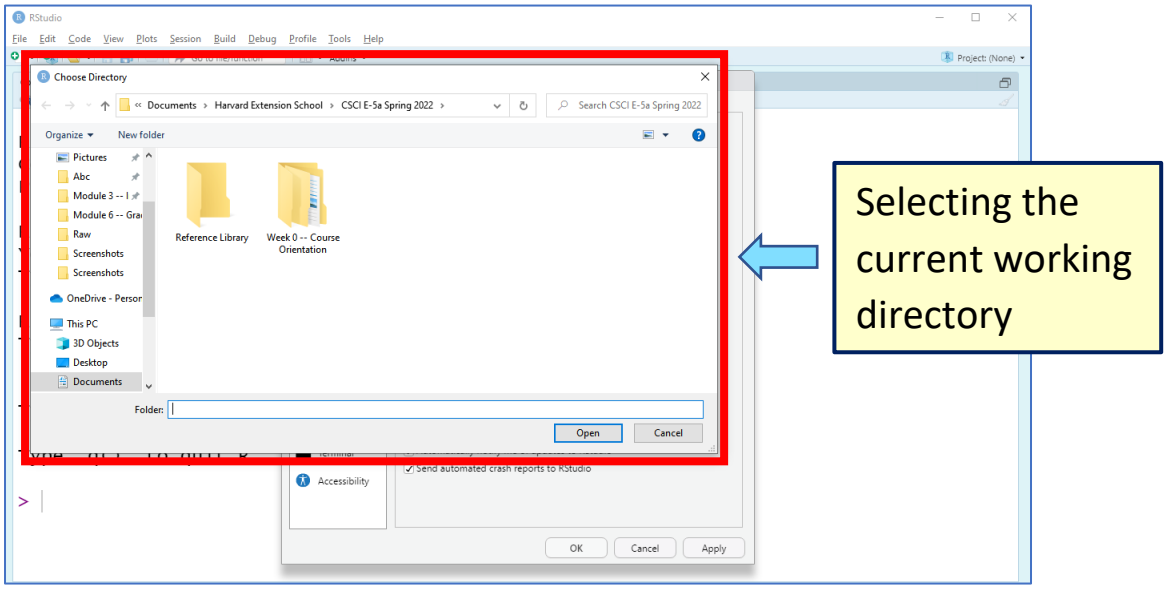

Figure 29: Selecting the current working directory

Next, let's select the "Code" tab in the Global Options box:

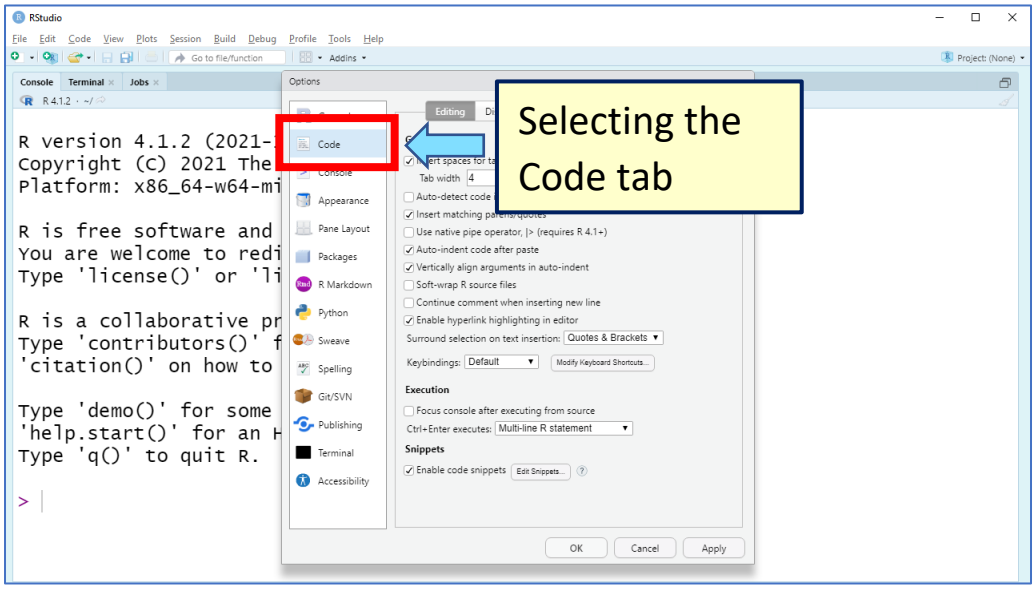

Figure 30: Selecting the Code tab

Set the Tab Width to 4:

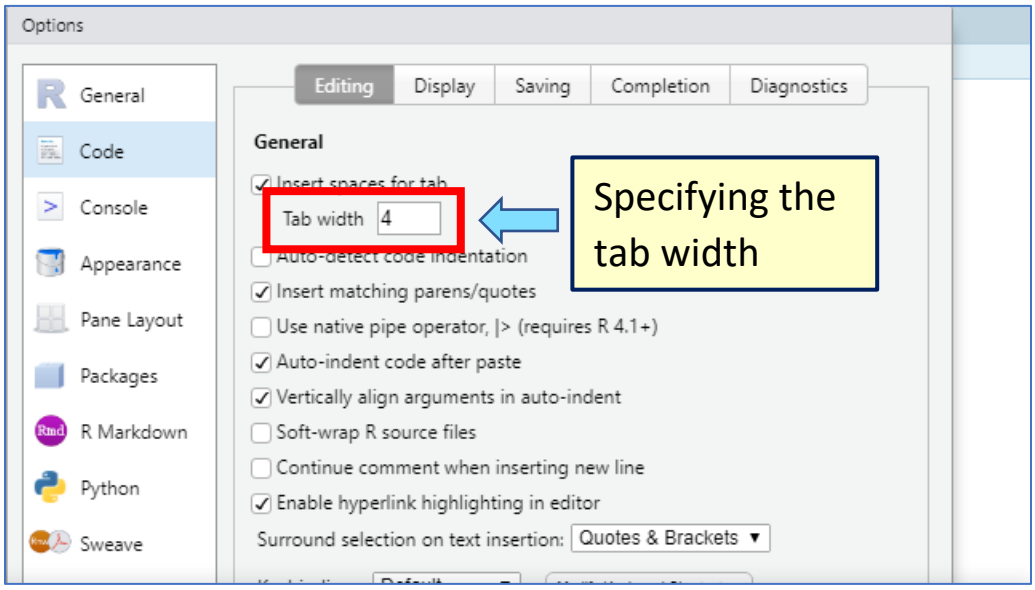

Figure 31: Specifying the tab width

Now let's select the Appearances tab:

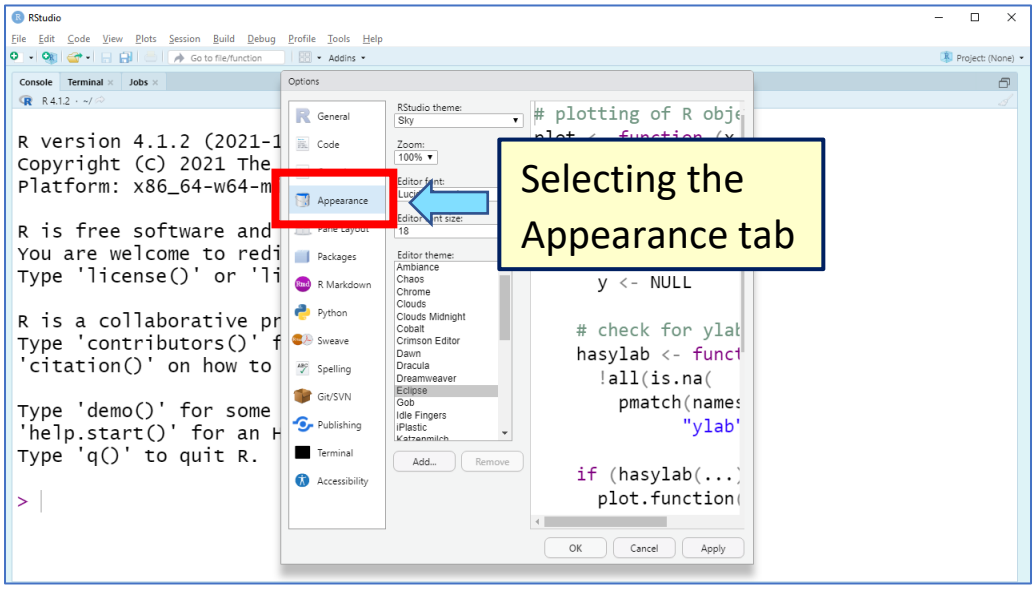

Figure 32: Selecting the Code tab

Here you can specify the RStudio theme, font, font size, and editor theme:

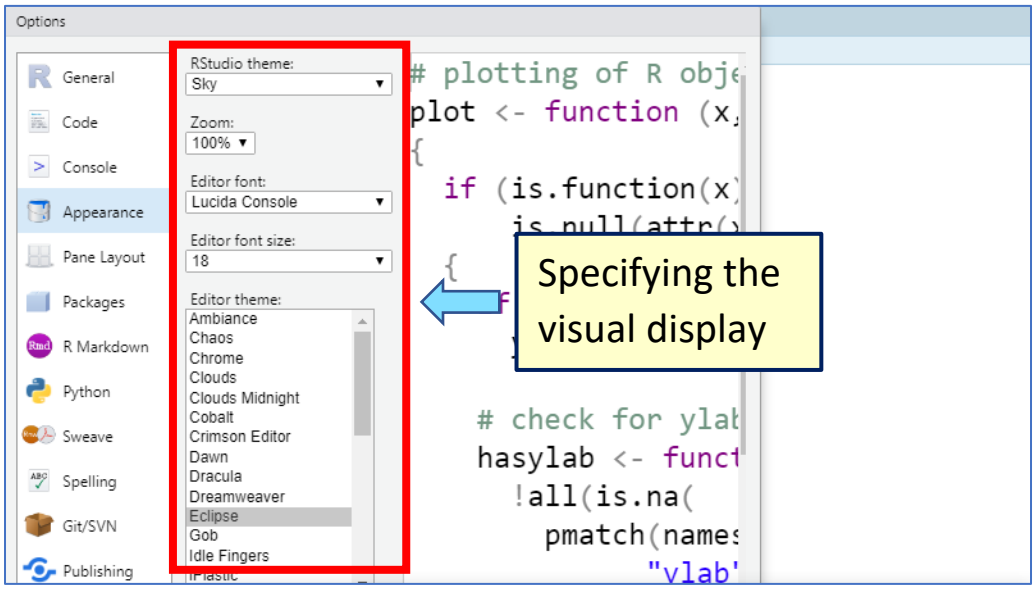

Figure 33: Customizing the visual display

For all demonstrations in CSCI E-5a, I'll use the Sky theme for RStudio, with Lucida Console as the editor font and an editor font size of 18. You are welcome to choose your own settings, and it's really up to you and what you like. However, I will encourage you to use a large font size – when you're debugging your code it helps to be able to see everything very clearly, and I've notice that students who use small font sizes sometimes have trouble detecting errors.

Finally, let's select the Pane Layout tab:

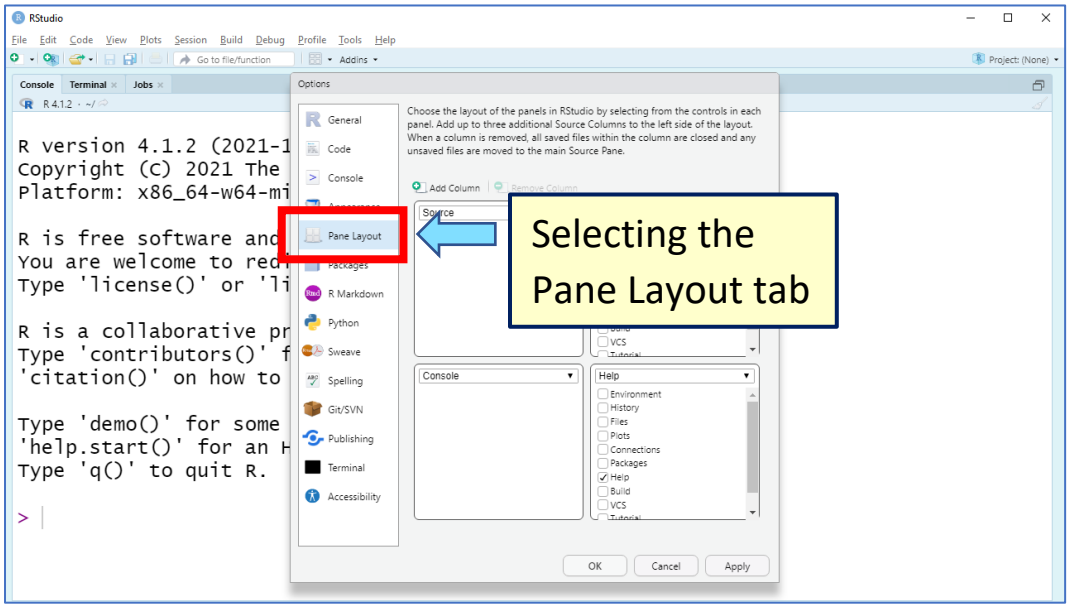

Figure 34: Selecting the Pane Layout tab

This determines the content of the panes of the display. You can choose these however you like, but for all demonstrations in CSCI E-5a I will use a standard layout. First, the upper left-hand pane should be set to "Source":

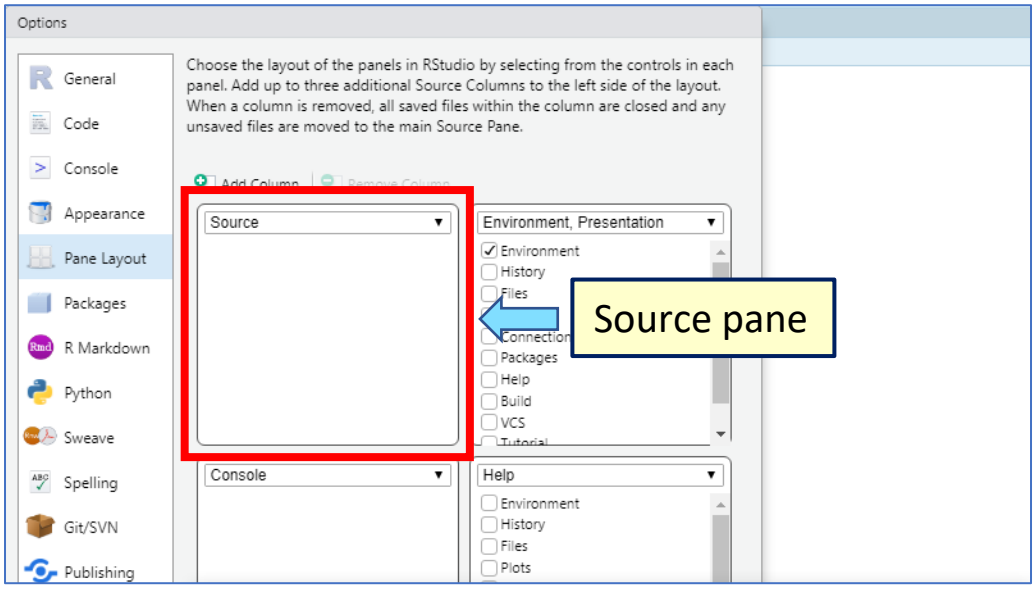

Figure 35: Specifying the Source pane

The lower left-hand pane should be set to "Console":

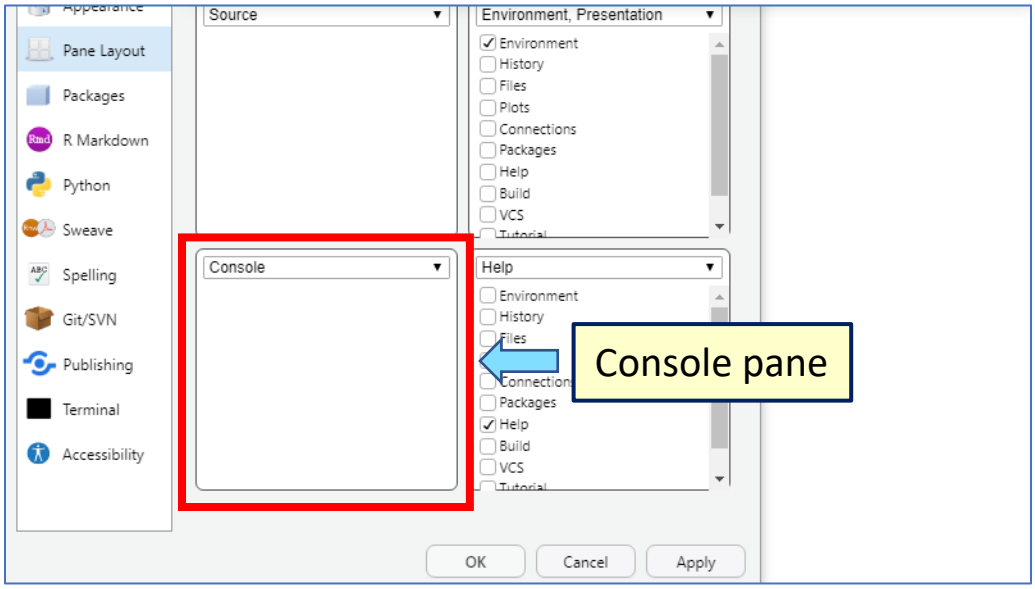

Figure 36: Specifying the Console pane

For the upper right-hand pane, make sure that the box for "Environment" is checked:

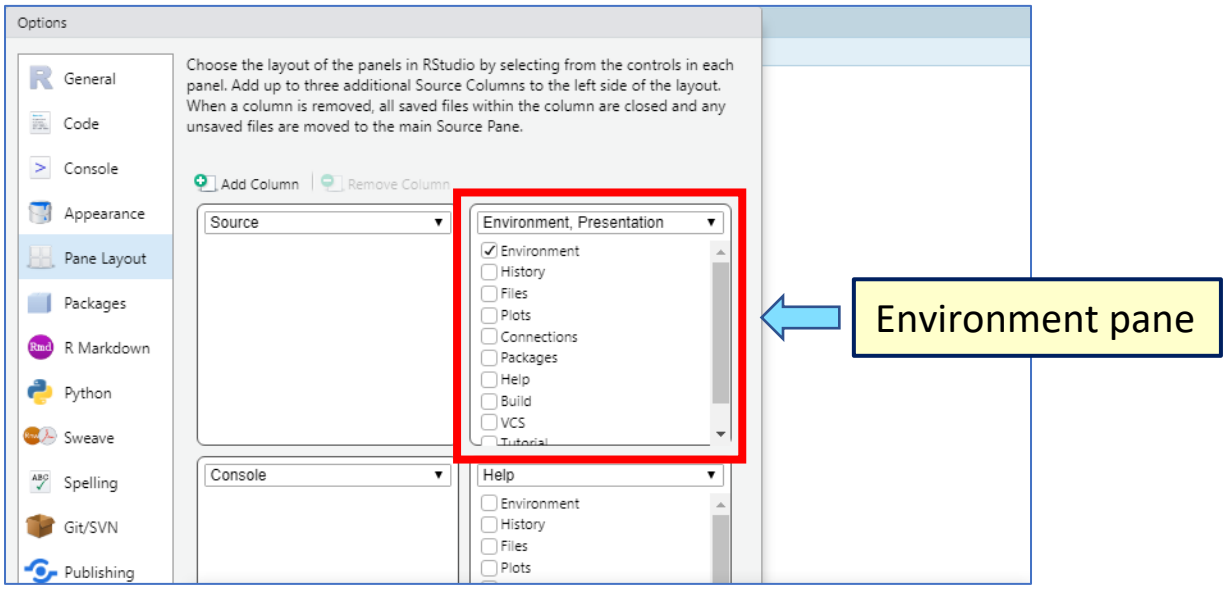

Figure 37: Specifying the Environment pane

For the lower right-hand pane, make sure the box for "Help" is checked:

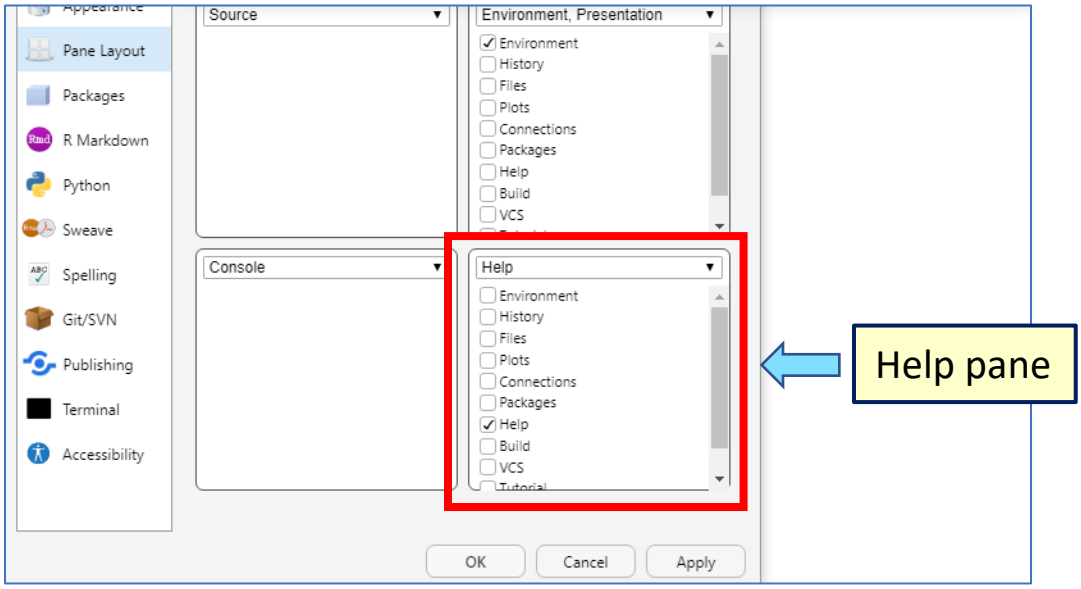

Figure 38: Specifying the Help pane

You can check other boxes for the right-hand panes if you want to, but I recommend this setup as the minimum.

Once you've finished configuring everything in the Global Options dialog box, you can save the settings by clicking on the "OK" button:

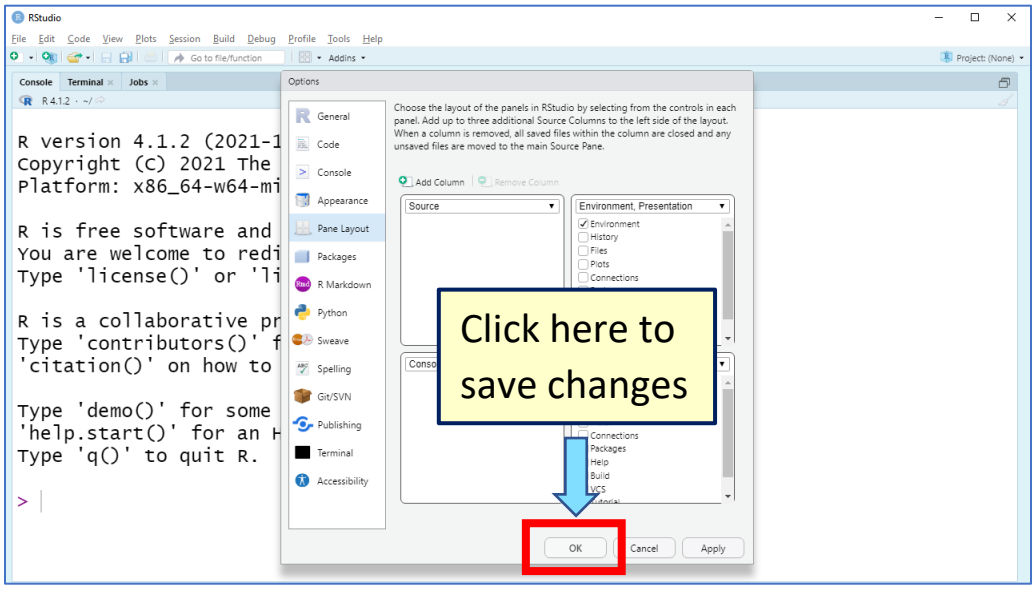

Figure 39: Saving your changes

OK! That's how to configure RStudio. Now let's see how to install the rmarkdown package.

Section 5: Installing the rmarkdown package

Main Idea: *We can install the rmarkdown package and knit an R notebook to an HTML document.*

In this section, we'll install the rmarkdown package and knit an R notebook to an HTML document.

Now that we've configured RStudio, we need to install the rmarkdown package. Packages enable us to extend the functionality of R and to customize our system. In particular, the rmarkdown package is required in order to work with R notebooks, so you must install this package!

In order to install the rmarkdown package, first make sure that you are connected to the Internet. Then, at the prompt at the bottom of the Console window, type the command:

```
install.packages( "rmarkdown" )
```

You screen should look like this:

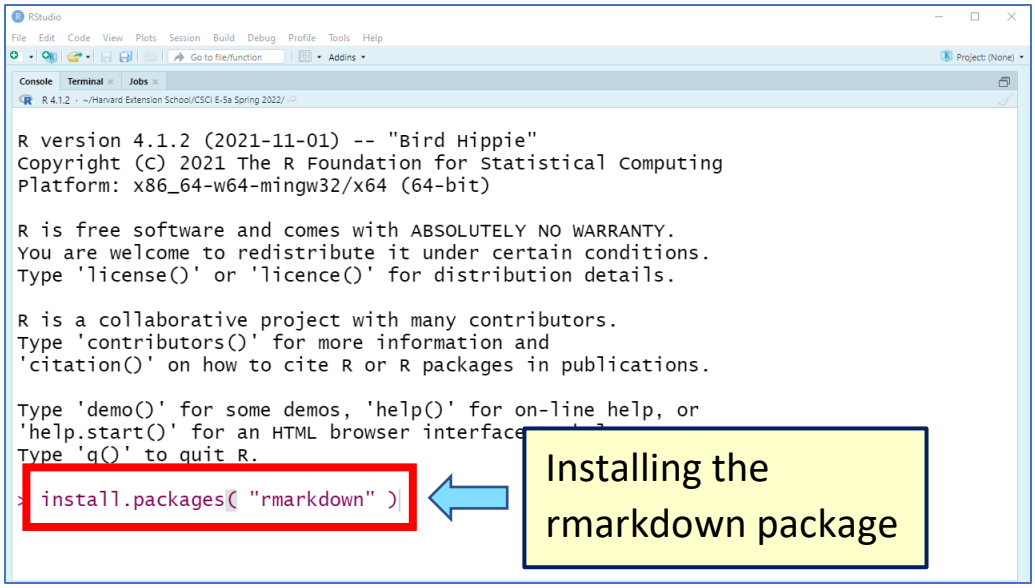

Figure 40: Installing the rmarkdown package

When you're all done, you should get the message "package 'rmarkdown' successfully unpacked and MD5 sums checked":

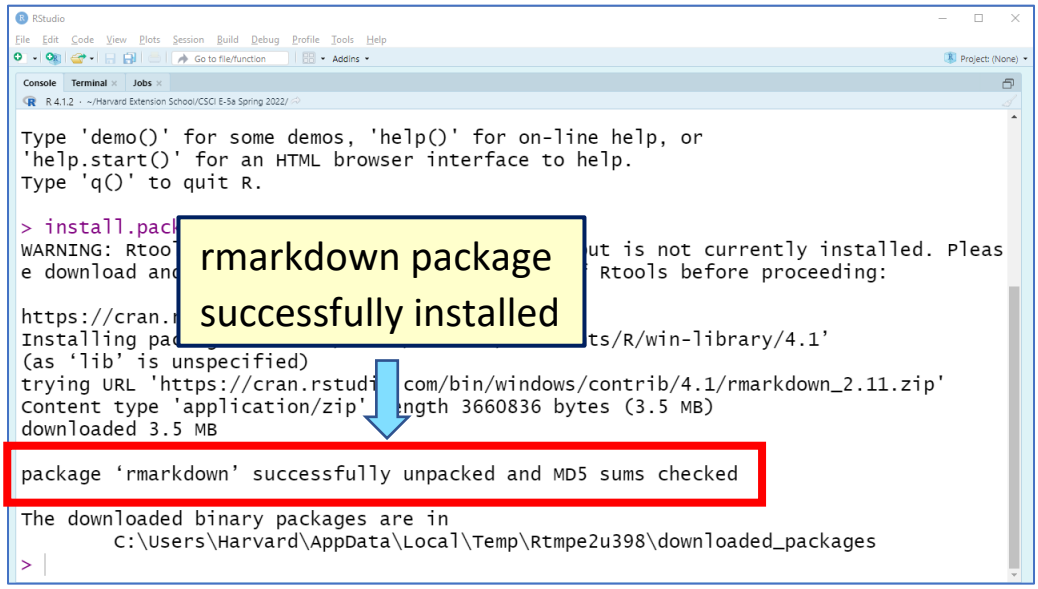

Figure 41: Finished installing the rmarkdown package

Now that you've installed the rmarkdown package, you should be able to knit an R notebook to an HTML document. To create a new R notebook document, click on the New File icon in the upper left-hand corner of RStudio:

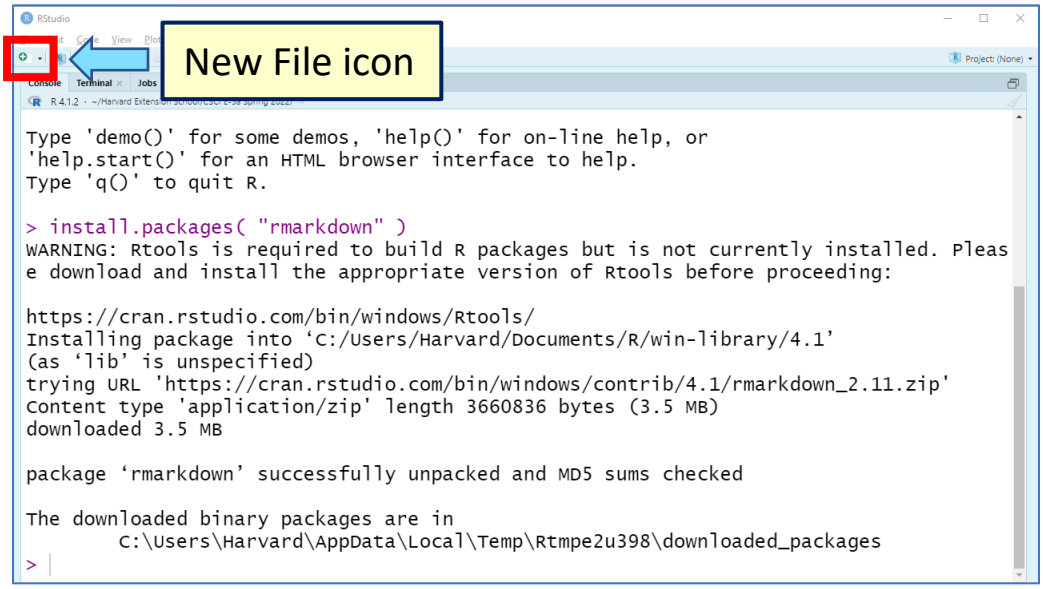

Figure 42: New File icon

I admit – that New File icon is very small! Let's take a closer look:

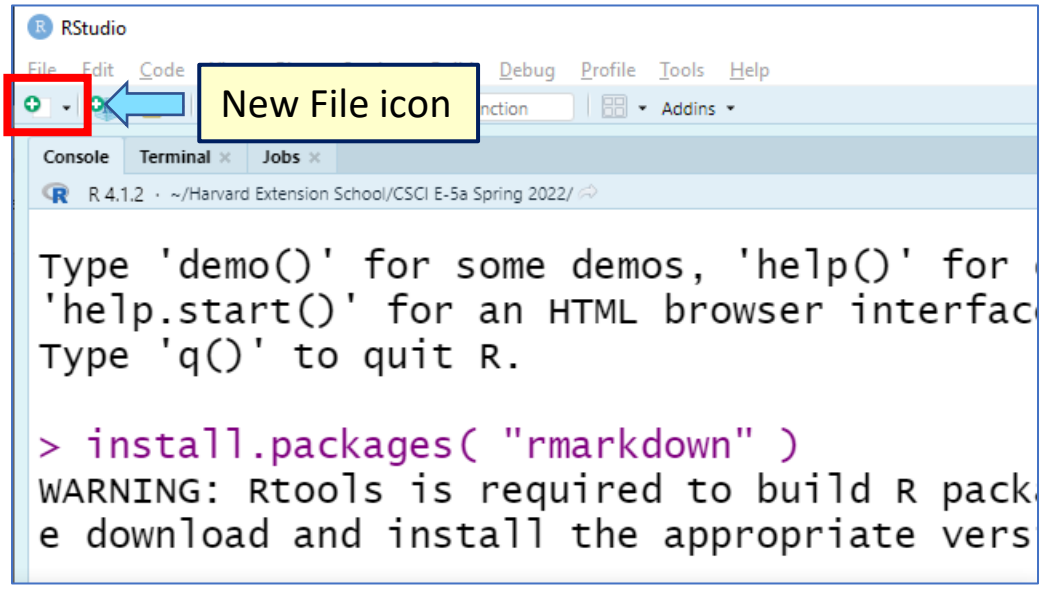

Figure 43: New File icon (close-up)

When you click on the link, you'll see a drop-down menu:

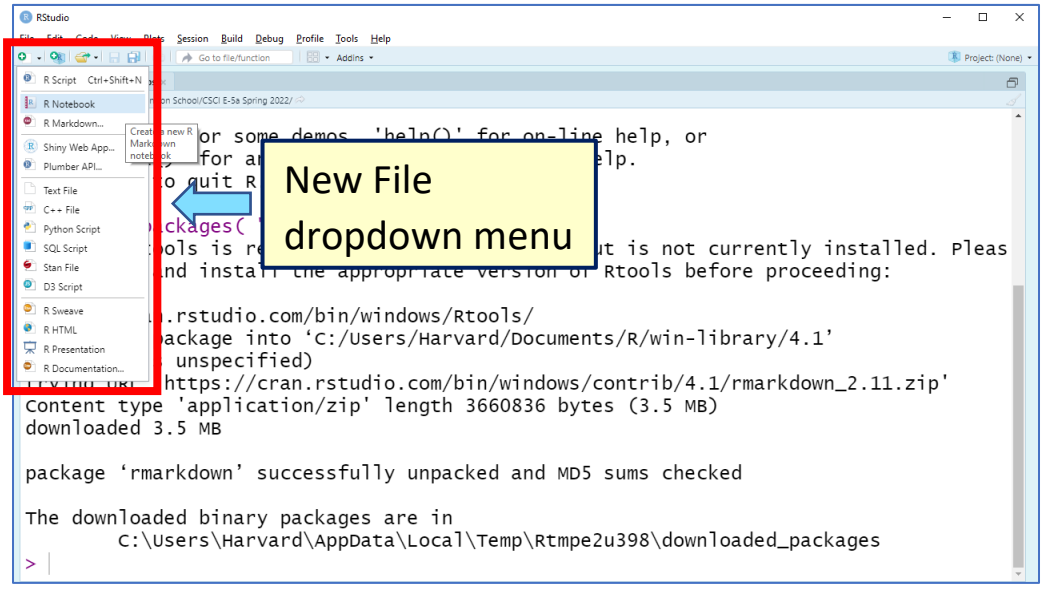

Figure 44: New File drop-down menu

Let's take a closer look:

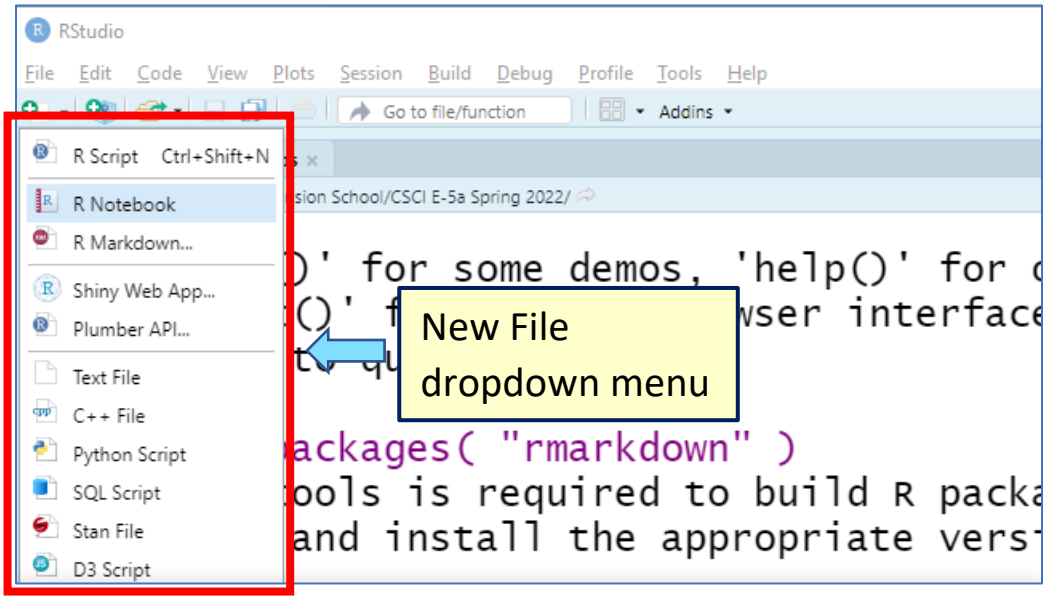

Figure 45: New File drop-down menu (close-up)

When you create a new R notebook, you'll see this:

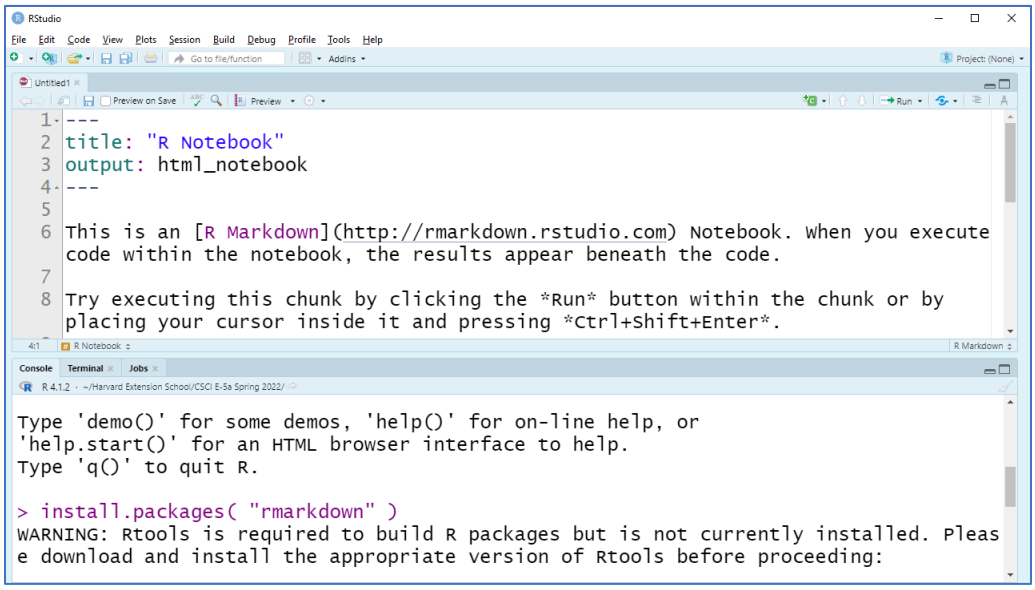

Figure 46: New R notebook

Notice that the Console window has now been reduced in size, and there is a new window on top. This new window is called the "Source" window, because it contains the source code for the R notebook.

You can minimize the Console window by clicking on the top bar and dragging, or by clicking on the minimize icon in the upper right-hand corner of the Console window:

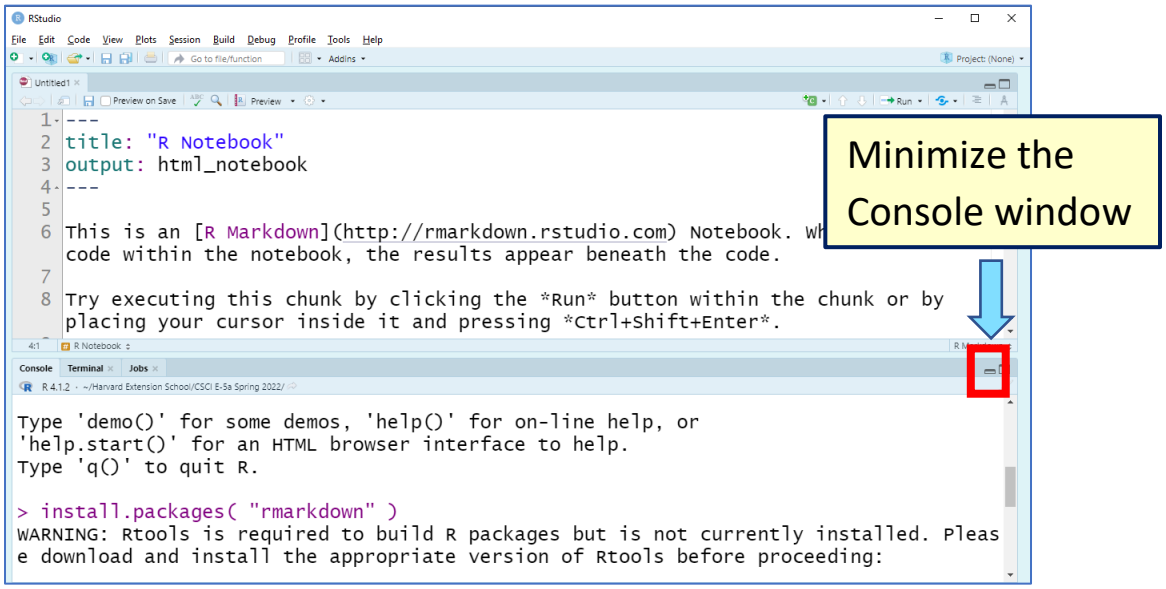

Figure 47: Minimizing the Console window

Once you've minimized the Console window, your display will look like this:

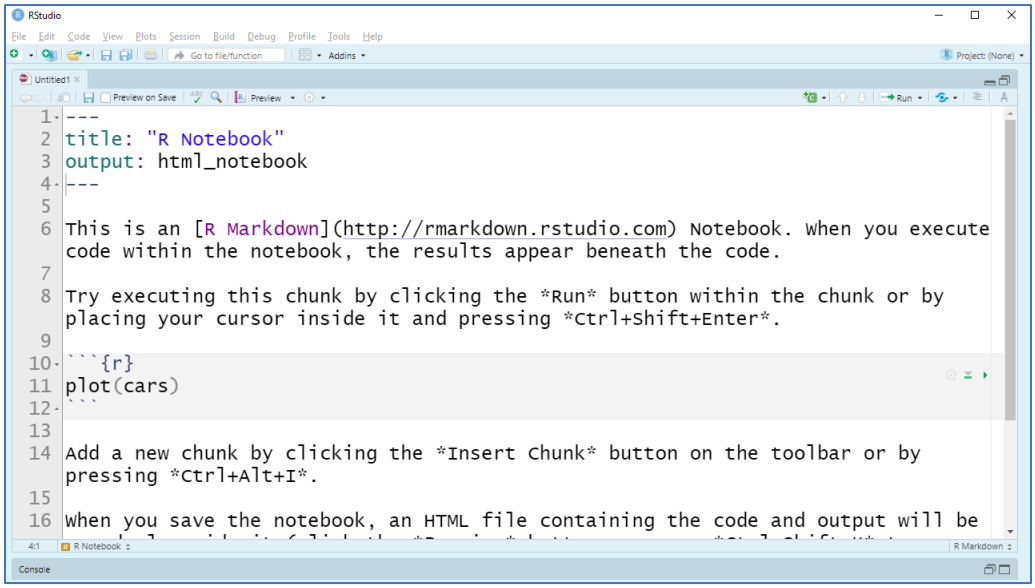

Figure 48: Full screen R notebook

Before we can knit this document to an HTML file, we first have to save it to disk. There are a number of different ways to save this document, but one way is to click on the Save Current Document icon, which is to the right of the New File icon:

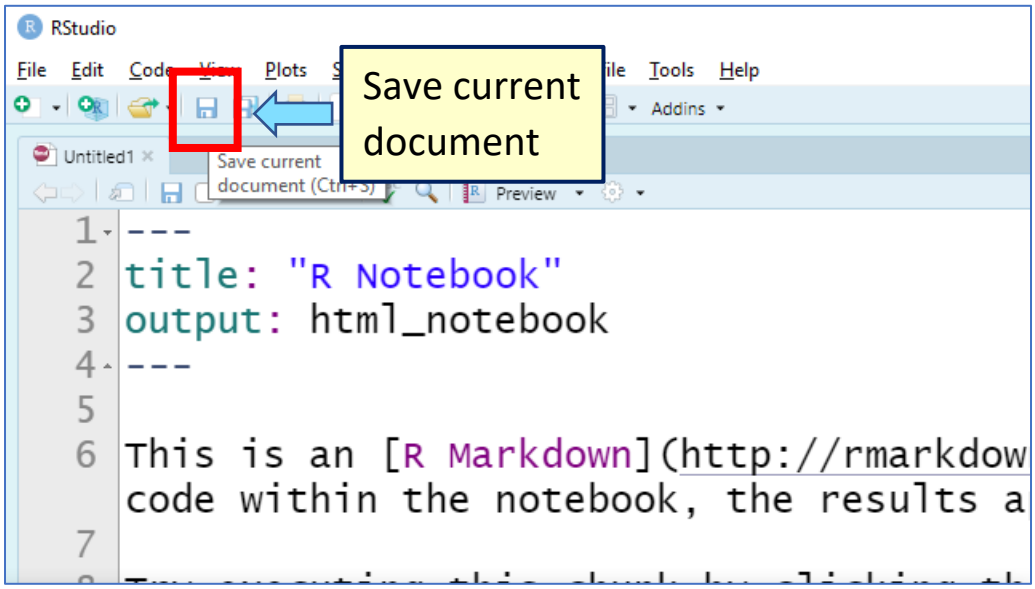

Figure 49: Save Current Document icon

When you click on this icon, a dialog box will appear. I recommend creating a folder called "Test" in the CSCI E-5a folder and saving the R notebook there; you can also call the R notebook something like "Test":

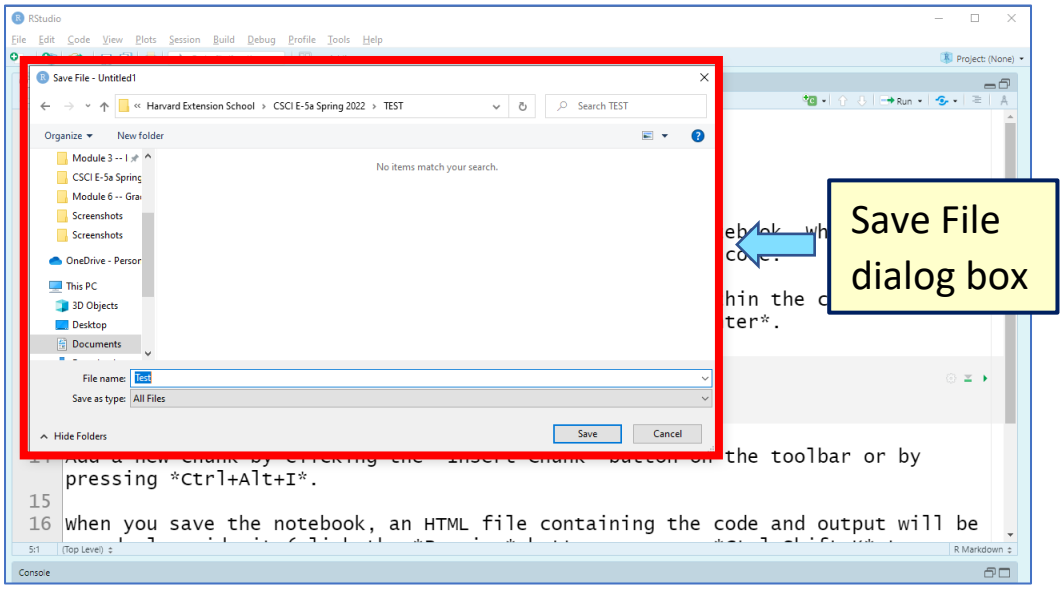

Figure 50: Saving the R notebook

Now we can knit the R notebook! The icon for knitting is at the top of the source window, and it currently says "Preview":

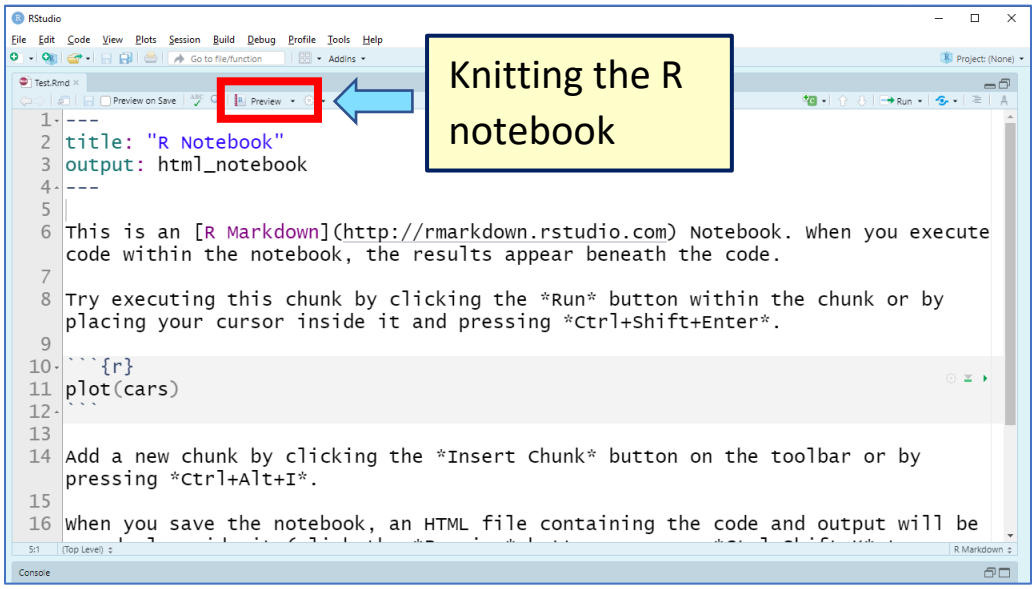

Figure 51: Knitting preview icon

When you click on this icon, you'll get a drop-down menu:

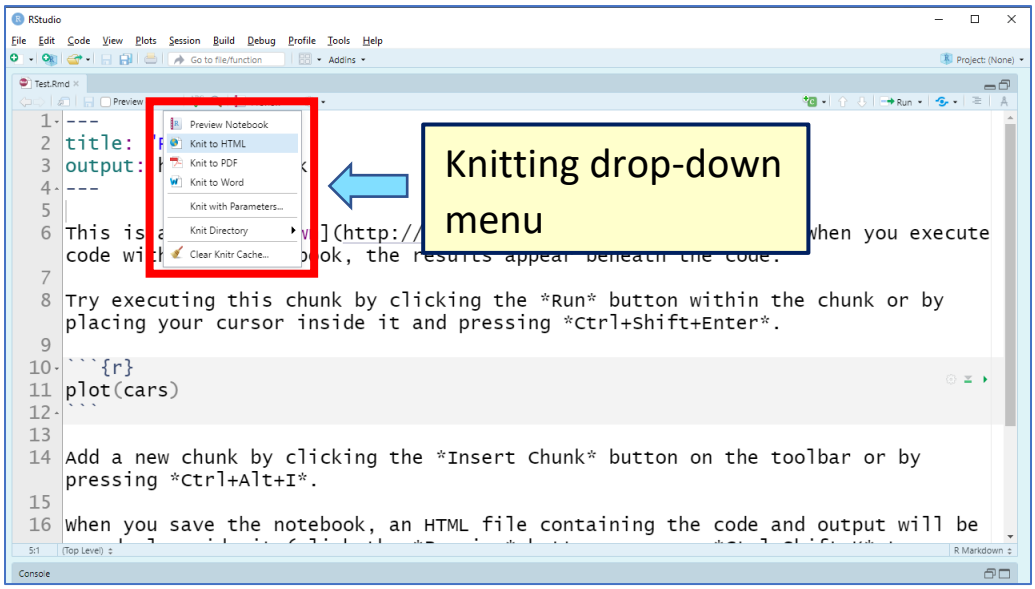

Figure 52: Knitting drop-down menu

Let's zoom in for a closer look:

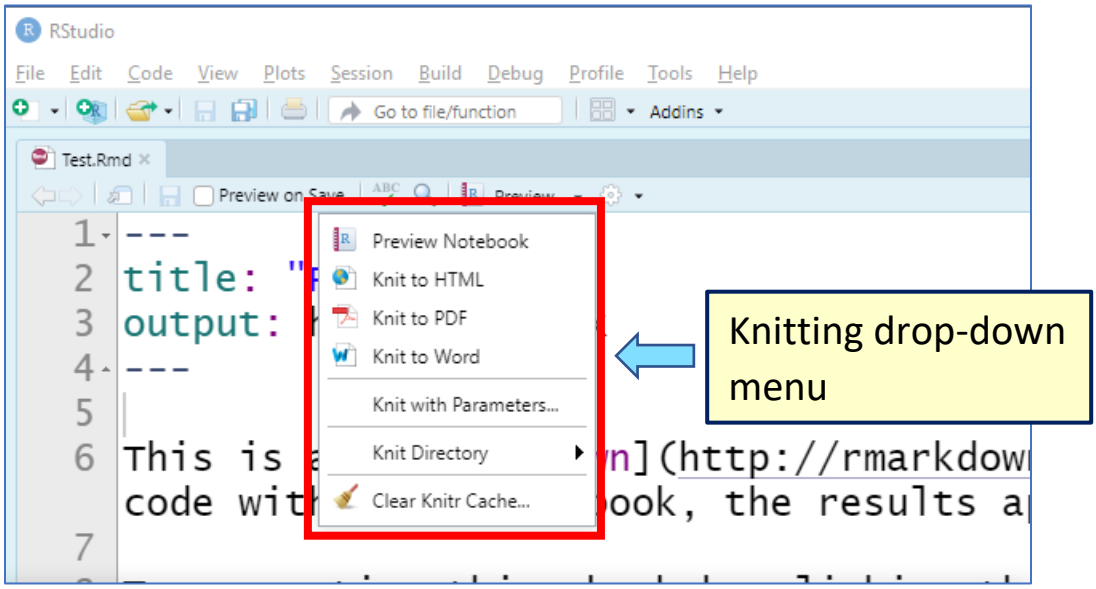

Figure 53: Knitting drop-down menu (close-up)

For this step, we want to knit to an HTML document, so select "Knit to HTML":

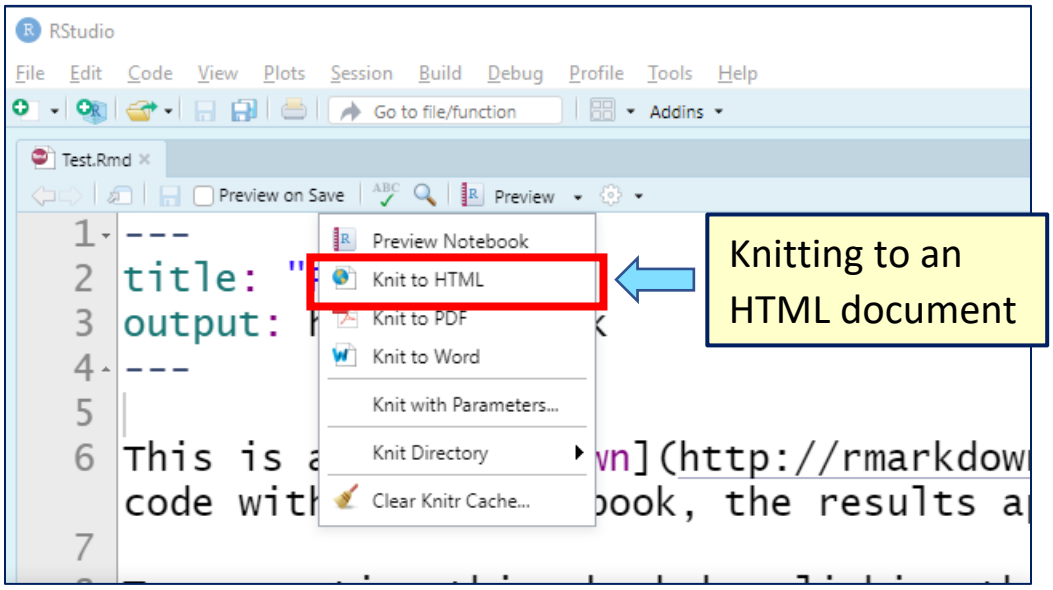

Figure 54: Knitting to an HTML document

R then creates a HTML document:

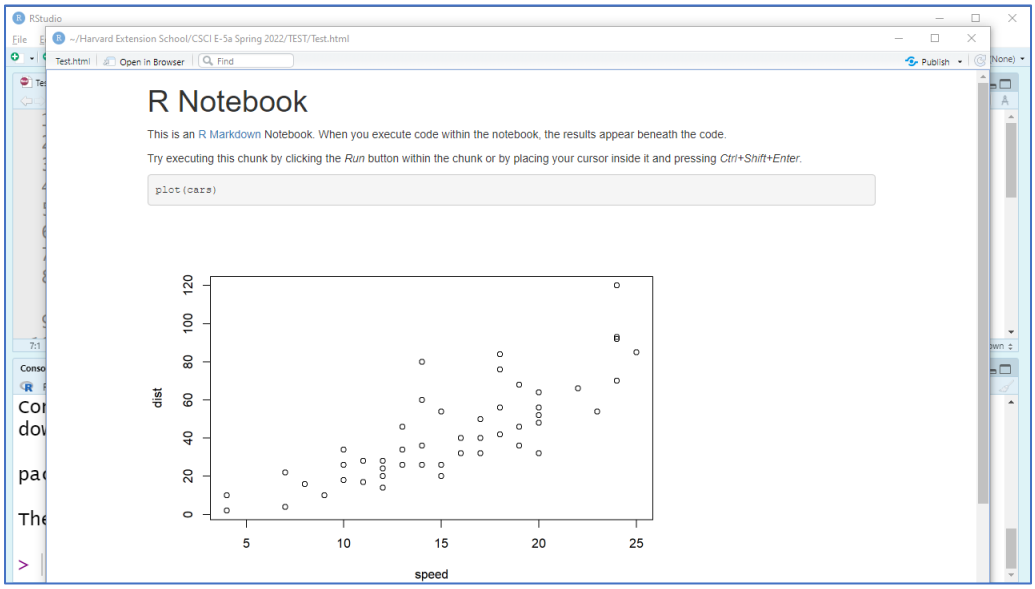

Figure 55: Knitting to an HTML document

I recommend taking a moment to read through the text of the R notebook, and then comparing that with the visual display of the HTML document.

So that's how to install the rmarkdown package and knit an R notebook to an HTML document. Now let's see how to install LaTeX and knit an R notebook to a PDF document.

Section 6: Installing LaTeX

Main Idea: *We can install LaTeX and knit an R notebook to a PDF document*

In this section, we'll see how to install LaTeX and knit an R notebook to a PDF document.

So far, we've been able to knit an R notebook to an HTML document. That's great! If you can successfully knit to an HTML document, it's a strong indication that the installation process has gone well so far. However, for CSCI E-5a that's not enough: we need to be able to knit to a PDF, and to do this we have to install LaTeX, which is our goal in this section.

What is this curious LaTeX thing? It's a powerful system for typesetting complex mathematics texts and formulas, and in the math/data science/machine learning world it is universally used. Once you become familiar with the formatting of LaTeX, you'll notice that many math and science books use this system. For CSCI E-5a you don't have to be able to use LaTeX to format text, and we'll only use it because it enables us to knit R notebooks to PDF

documents. But I encourage you to learn more about this fascinating tool, and if you're serious about further work in statistics, data science, or machine learning, it's a valuable skill to have. The best way to get started with LaTeX is to use the cloud-based system Overleaf, which you can find at www.overleaf.com. However, for our purposes this isn't necessary, and our interest in LaTeX is simply that it enables us to knit to a PDF.

To install LaTeX, we first have to download the tinytex package. Just as before, to do this you go to the Console window and at the prompt type in:

```
install.packages( "tinytex" )
```

Once you've installed the tinytex package, you should see this:

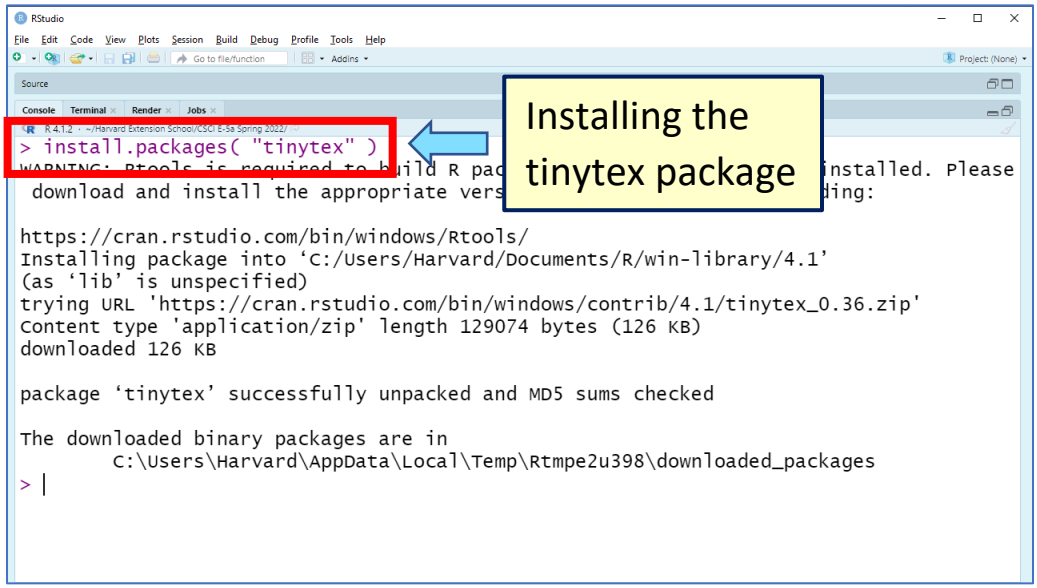

Figure 56: Installing the tinytex package

The tinytex package itself isn't a LaTeX installation. Instead, tinytex manages the LaTeX installation process, which can be a little complex. To install LaTeX, type this command into the prompt in the Console window:

```
tinytex::install_tinytex()
```

Make sure all the letters are lowercase, and be sure to use double colons and an underscore. Also, don't forget the paretheses at the end! You must type this command in precisely, and any deviations will result in an error. Once you've typed in the command exactly, hit return. At first, not much will happen, and you might wonder if there's some problem. Don't worry, it's working! However, the download and installation process will take a while to run (on my system it took about 7 to 8 minutes), so be patient! When it's all over, you should see a screen that looks like this:

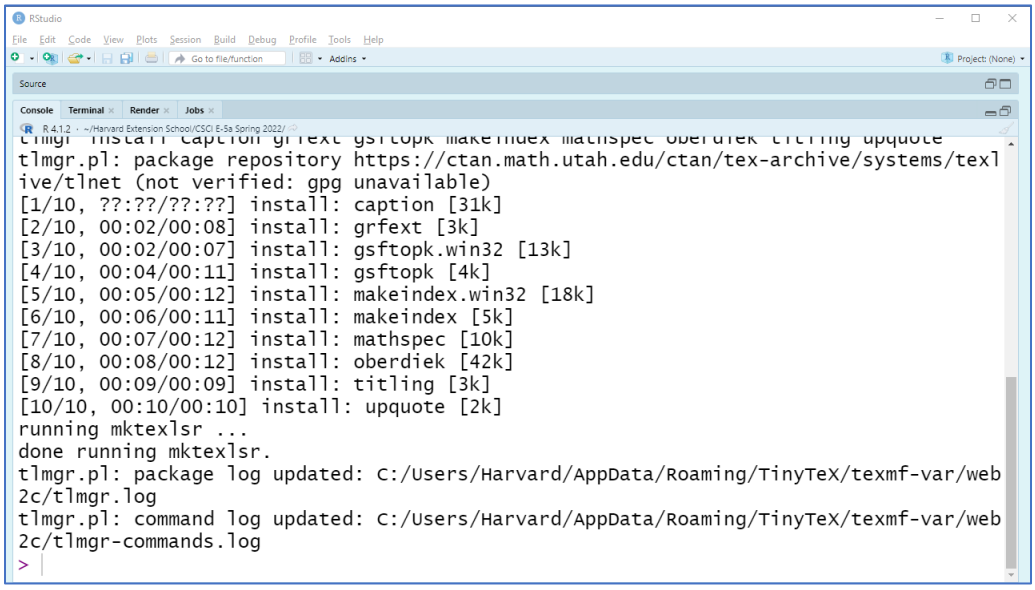

Figure 57: Running the tinytex installer

Almost done! Now let's return to the Source window with the Test R notebook. Once again, we'll go to the icon for knitting the R notebook, but now we'll select "Knit to PDF":

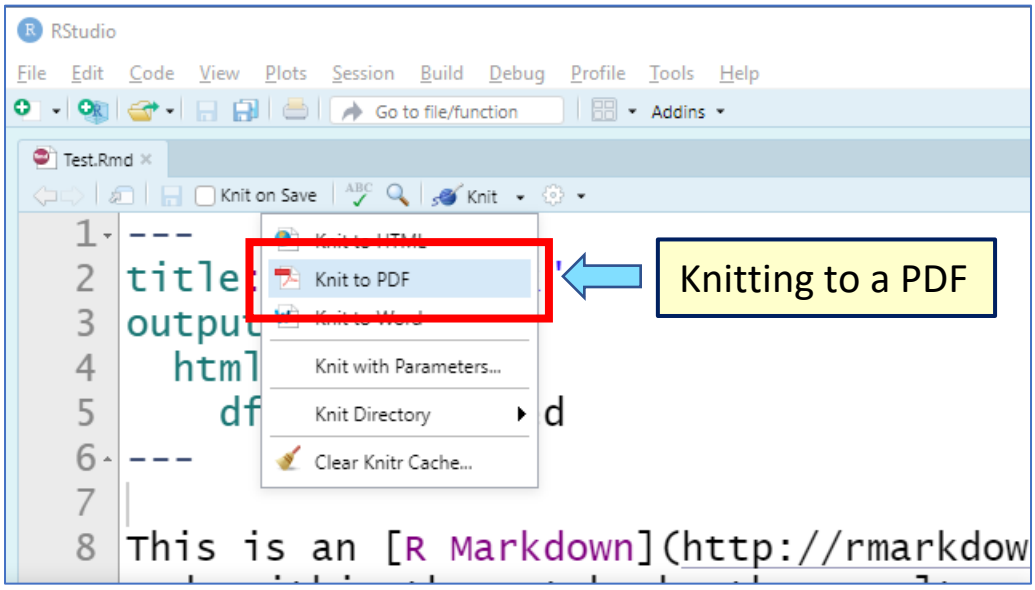

Figure 58: Knitting to an PDF document

And this is the result:

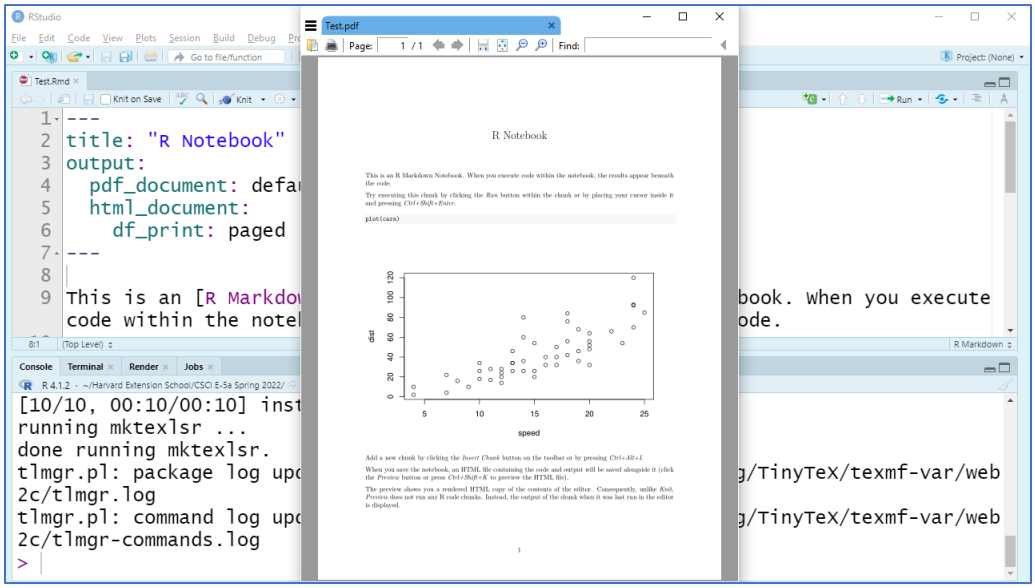

Figure 59: Finished knitting to PDF document

And you're done!!!

So that's how to install LaTeX and knit an R notebook to a PDF.

Now, let's review what we've learned in this module.

Module Review

In this module, we learned how to install all the software that we'll need for CSCI E-5a.

- In Section 1, we had an overview of the installation process.
- In Section 2, we installed R.
- In Section 3, we installed RStudio.
- In Section 4, we configured RStudio.
- In Section 5, we installed the rmarkdown package, and knitted an R notebook to an HTML document.
- In Section 6, we installed tinytex, and knitted an R notebook to a PDF document.

Now that you've completed this module, you should be able to:

- Install R
- Install RStudio
- Configure RStudio
- Install the rmarkdown package
- Install tinytex
- Knit an R notebook to an HTML or PDF document

All right! That's it for Module 3: Installing Software.

Now let's move on to Module 4: Working with R Notebooks.