

Lecture 02: Introduction to R and RStudio

Taylor Arnold

30 August 2017

Today

The goal of today is get you set up on your own computer with GitHub, R, and RStudio.

Hopefully you brought your own machine to work with today so that you'll have this set up for the remainder of the semester.

Reminder, all course materials (including the first set of notes) are available at:

<https://statsmaths.github.io/stat289>

GitHub

About GitHub

GitHub is a website that provides a public repository for code. In the open source software community it has become the de facto place to store code.

GitHub offers a free tier of its service; the only limitation is that all of your repositories will be publicly available.

While the site is a commercial enterprise, the underlying technology, **git**, is completely free and open source.

Profile

Click on the icon on the upper right hand side of the screen and select *Settings*.

Fill in the following information:

- ▶ Your name (first only is OK)
- ▶ A short bio (1-2 sentences)
- ▶ Location (either RVA or your hometown)
- ▶ A profile picture

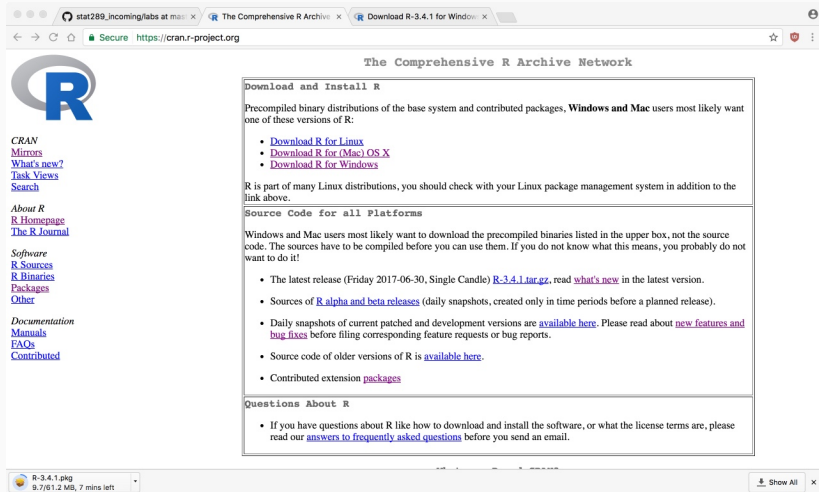
Make sure to save the information.

We'll be using GitHub classrooms for this course. I will send out a link right now that takes you to the first assignment.

R and RStudio

Download R

Next, we need to download the R programming language. To do this go to <https://cran.r-project.org/> and select your platform:



The screenshot shows a web browser window with the address bar displaying <https://cran.r-project.org/>. The page title is "The Comprehensive R Archive Network". On the left sidebar, there are links for "CRAN", "Mirrors", "What's new?", "Task Views", "Search", "About R", "R Homepage", "The R Journal", "Software", "R Sources", "R Binaries", "Packages", "Other", "Documentation", "Manuals", "FAQs", and "Contributed". The main content area is titled "Download and Install R" and contains the following text: "Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:". Below this, there is a list of links: "Download R for Linux", "Download R for (Mac) OS X", and "Download R for Windows". The text continues: "R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above." Below this, there is a section titled "Source Code for all Platforms" with the text: "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!". Below this, there is a list of links: "The latest release (Friday 2017-06-30, Single Candle) [R-3.4.1.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](#).", and "Contributed extension [packages](#)". Below this, there is a section titled "Questions About R" with the text: "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." At the bottom of the browser window, there is a download bar showing "R-3.4.1.pkg" with a size of "9.7/61.2 MB" and a time of "7 mins left". There is also a "Show All" button.

stat289_incoming/labs at masi x The Comprehensive R Archive x Download R-3.4.1 for Windows x

Secure <https://cran.r-project.org/>

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](#)
- [Download R for \(Mac\) OS X](#)
- [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 (Friday 2017-06-30, Single Candle) [R-3.4.1.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

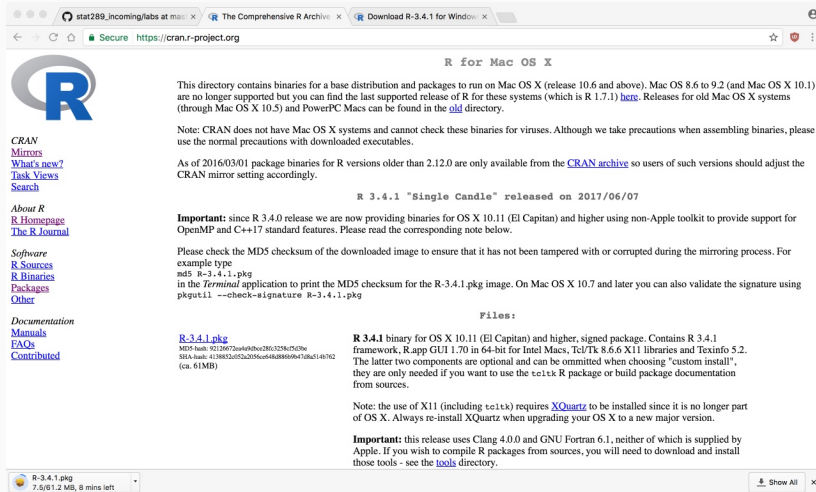
- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

R-3.4.1.pkg
9.7/61.2 MB, 7 mins left

Show All x

Download R

For macOS, just download R-3.4.1 (or whatever is the most recent):



The screenshot shows a web browser window with the address bar displaying "https://cran.r-project.org". The page title is "R for Mac OS X". The main content area contains the following text:

This directory contains binaries for a base distribution and packages to run on Mac OS X (release 10.6 and above). Mac OS 8.6 to 9.2 (and Mac OS X 10.1) are no longer supported but you can find the last supported release of R for these systems (which is R 1.7.1) [here](#). Releases for old Mac OS X systems (through Mac OS X 10.5) and PowerPC Macs can be found in the [old](#) directory.

Note: CRAN does not have Mac OS X systems and cannot check these binaries for viruses. Although we take precautions when assembling binaries, please use the normal precautions with downloaded executables.

As of 2016/03/01 package binaries for R versions older than 2.12.0 are only available from the [CRAN archive](#) so users of such versions should adjust the CRAN mirror setting accordingly.

R 3.4.1 "Single Candle" released on 2017/06/07

Important: since R 3.4.0 release we are now providing binaries for OS X 10.11 (El Capitan) and higher using non-Apple toolkit to provide support for OpenMP and C++17 standard features. Please read the corresponding note below.

Please check the MD5 checksum of the downloaded image to ensure that it has not been tampered with or corrupted during the mirroring process. For example type

```
md5 R-3.4.1.pkg
```

in the *Terminal* application to print the MD5 checksum for the R-3.4.1.pkg image. On Mac OS X 10.7 and later you can also validate the signature using

```
pkgutil --check-signature R-3.4.1.pkg
```

Files:

R-3.4.1.pkg
MD5-hash: 92126672ea49d8ec286c32258c55c9e
SHA1-hash: 4138853a052a205cc64d886969a71d85148762
(ca. 61MB)

R 3.4.1 binary for OS X 10.11 (El Capitan) and higher, signed package. Contains R 3.4.1 framework, R.app GUI 1.70 in 64-bit for Intel Macs, Tcl/Tk 8.6.6 X11 libraries and Texinfo 5.2. The latter two components are optional and can be omitted when choosing "custom install", they are only needed if you want to use the `tcltk` R package or build package documentation from sources.

Note: the use of X11 (including `tcltk`) requires [XQuartz](#) to be installed since it is no longer part of OS X. Always re-install XQuartz when upgrading your OS X to a new major version.

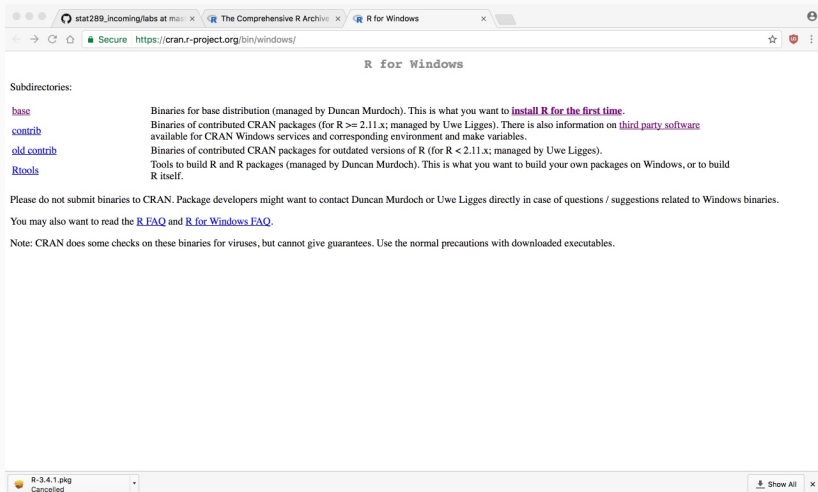
Important: this release uses Clang 4.0.0 and GNU Fortran 6.1, neither of which is supplied by Apple. If you wish to compile R packages from sources, you will need to download and install those tools - see the [tools](#) directory.

The left sidebar contains links: CRAN, Mirrors, What's new?, Task Views, Search, About R, R Homepage, The R Journal, Software, R Sources, R Binaries, Packages, Other, Documentation, Manuals, FAQs, Contributed.

The bottom of the page shows a download progress bar for "R-3.4.1.pkg" (7.5/61.2 MB, 8 mins left) and a "Show All" button.

Download R

For Windows, first select **base**



The screenshot shows a web browser window with the address bar displaying `https://cran.r-project.org/bin/windows/`. The page title is "R for Windows". Under the heading "Subdirectories:", there are four links: [base](#), [contrib](#), [old contrib](#), and [Rtools](#). Each link is followed by a description of the subdirectory's contents. The [base](#) link is highlighted in purple. Below the subdirectories, there is a paragraph advising not to submit binaries to CRAN and a link to the [R FAQ](#) and [R for Windows FAQ](#). A note at the bottom states that CRAN does some checks for viruses but cannot give guarantees. At the bottom of the browser window, a download bar shows a file named "R-3.4.1.pkg" with a status of "Cancelled" and a "Show All" button.

R for Windows

Subdirectories:

- [base](#) Binaries for base distribution (managed by Duncan Murdoch). This is what you want to **install R for the first time**.
- [contrib](#) Binaries of contributed CRAN packages (for R \geq 2.11.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.
- [old contrib](#) Binaries of contributed CRAN packages for outdated versions of R (for R $<$ 2.11.x; managed by Uwe Ligges).
- [Rtools](#) Tools to build R and R packages (managed by Duncan Murdoch). This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Duncan Murdoch or Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.

R-3.4.1.pkg
Cancelled

Show All

Download R

And then *Download R 3.4.1*

The screenshot shows a web browser window with the address bar displaying `https://cran.r-project.org/bin/windows/base/`. The page title is "R-3.4.1 for Windows (32/64 bit)". The main content area has a grey background and contains the following text:

[Download R 3.4.1 for Windows](#) (62 megabytes, 32/64 bit)
[Installation and other instructions](#)
[New features in this version](#)

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. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit 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.htm](#).

Last change: 2017-06-30, by Duncan Murdoch

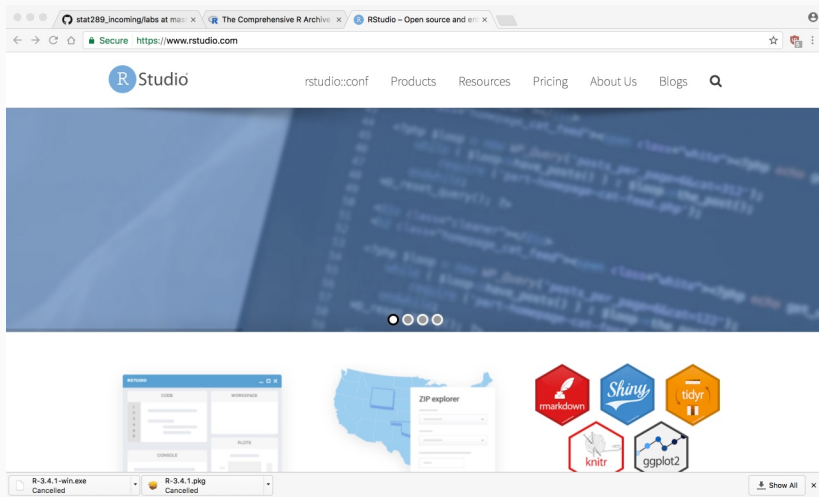
At the bottom of the browser window, there is a download bar showing two files: "R-3.4.1-win.exe" and "R-3.4.1.pkg", both with a status of "Cancelled". A "Show All" button is visible on the right side of the download bar.

Once you have the .pkg (macOS) or .exe (Windows) file, install this on your computer according to the default settings.

The files we just downloaded are the core R language files doing all the hard work of processing data. Next, we'll install a helpful GUI frontend that make calling R easier.

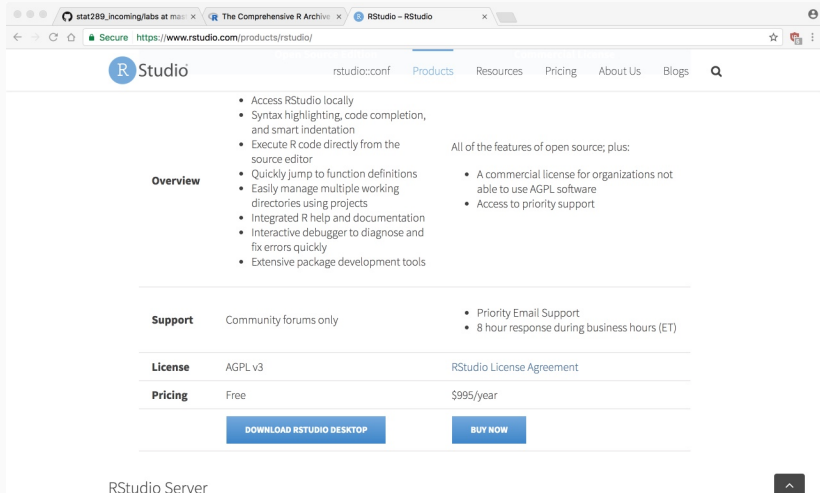
Download RStudio

Go to <https://www.rstudio.com/>. Click on Products => RStudio.



Download RStudio

Scroll down to the **DOWNLOAD RSTUDIO DESKTOP** button and click on it.



The screenshot shows the RStudio website in a web browser. The browser's address bar displays the URL <https://www.rstudio.com/products/rstudio/>. The website's navigation bar includes the RStudio logo, a link to `rstudio::conf`, and links for **Products**, Resources, Pricing, About Us, and Blogs. A search icon is also present.

The main content area features an **Overview** section with a list of features:

- Access RStudio locally
- Syntax highlighting, code completion, and smart indentation
- Execute R code directly from the source editor
- Quickly jump to function definitions
- Easily manage multiple working directories using projects
- Integrated R help and documentation
- Interactive debugger to diagnose and fix errors quickly
- Extensive package development tools

To the right of the features list, it states: "All of the features of open source; plus:"

- A commercial license for organizations not able to use AGPL software
- Access to priority support

Below the overview, there is a table with the following information:

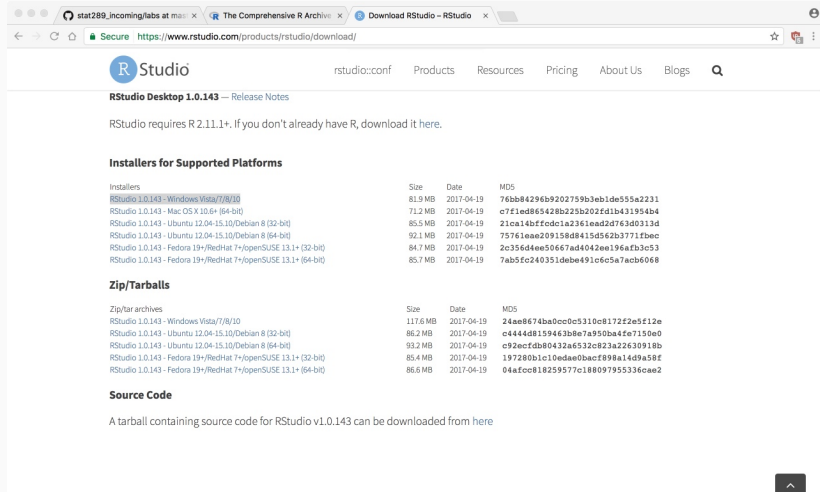
Support	Community forums only	<ul style="list-style-type: none">• Priority Email Support• 8 hour response during business hours (ET)
License	AGPL v3	RStudio License Agreement
Pricing	Free	\$995/year

At the bottom of the page, there are two prominent blue buttons: **DOWNLOAD RSTUDIO DESKTOP** and **BUY NOW**.

The footer of the page includes the text "RStudio Server" on the left and a small upward-pointing arrow icon on the right.

Download RStudio

Scroll down again to the **Installers for Supported Platforms**.
The Windows link gives you an exe:



The screenshot shows a web browser window with the URL <https://www.rstudio.com/products/rstudio/download/>. The page features the RStudio logo and navigation links: `rstudio::conf`, Products, Resources, Pricing, About Us, and Blogs. A search icon is also present. The main heading is **RStudio Desktop 1.0.143** with a sub-link for Release Notes. A note states: "RStudio requires R 2.11.1+. If you don't already have R, download it here." Below this is the section **Installers for Supported Platforms**, which contains a table of download links, sizes, dates, and MD5 hashes. The first row is highlighted. Another section, **Zip/Tarballs**, contains a similar table. At the bottom, the **Source Code** section mentions a tarball for RStudio v1.0.143.

RStudio Desktop 1.0.143 — Release Notes

RStudio requires R 2.11.1+. If you don't already have R, download it here.

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.0.143 - Windows Vista/7/8/10	81.9 MB	2017-04-19	76bb84296b9202759b3eb1de555a2231
RStudio 1.0.143 - Mac OS X 10.6+ (64-bit)	71.2 MB	2017-04-19	c7f1ed865428b225b202fd1b431954b4
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	85.5 MB	2017-04-19	21ca14bffcddc1a2361ead2d763d0313d
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	92.1 MB	2017-04-19	75761eae209158d8415d562b3771fbec
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	84.7 MB	2017-04-19	2c356d4ee50667ad4042ee196afb3c53
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	85.7 MB	2017-04-19	7ab5fc240351debe491c6c5a7acb6068

Zip/Tarballs

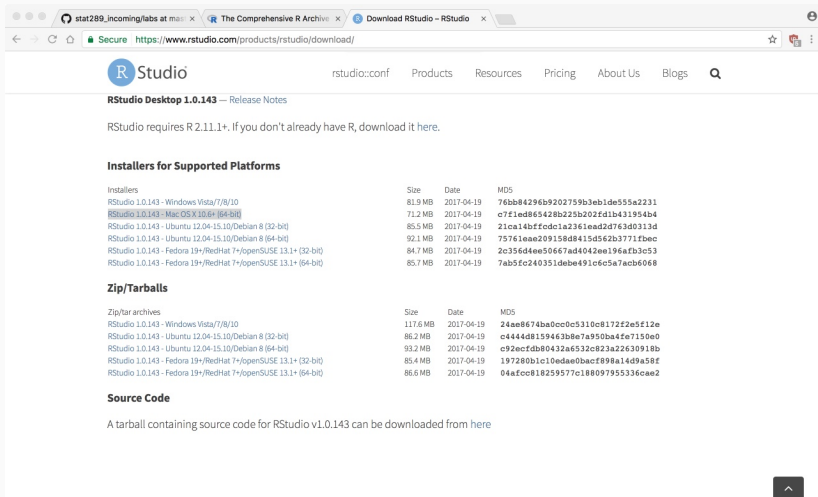
Zip/tar archives	Size	Date	MD5
RStudio 1.0.143 - Windows Vista/7/8/10	117.6 MB	2017-04-19	24ae8674ba0cc0c5310c8172f2e5f12e
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	86.2 MB	2017-04-19	c4444d8159463b8e7a950ba4fe7150e0
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	93.2 MB	2017-04-19	c92ecfdb80432a6532c823a2263091b8
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	85.4 MB	2017-04-19	197280b1c10edae0bacf898a14d9a58f
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	86.6 MB	2017-04-19	04afcc818259577c188097955336cae2

Source Code

A tarball containing source code for RStudio v1.0.143 can be downloaded from here

Download RStudio

And the macOS link gives a dmg:



The screenshot shows a web browser window with the URL <https://www.rstudio.com/products/rstudio/download/>. The page features the RStudio logo and navigation links: `rstudio::conf`, Products, Resources, Pricing, About Us, and Blogs. A search icon is also present.

RStudio Desktop 1.0.143 — Release Notes

RStudio requires R 2.11.1+. If you don't already have R, download it here.

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.0.143 - Windows Vista/7/8/10	81.9 MB	2017-04-19	76bb04296b9202759b3eb1de555a2231
<u>RStudio 1.0.143 - Mac OS X 10.6+ (64-bit)</u>	71.2 MB	2017-04-19	c7f1ed865428b225b202f3d1b431954b4
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	85.5 MB	2017-04-19	21ca14bffc0c1a2361ead2d763d0313d
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	92.1 MB	2017-04-19	75761eae209158d8415d562b3771fbec
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	84.7 MB	2017-04-19	2c356d4ee50667ad4042ee196afb3c53
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	85.7 MB	2017-04-19	7ab5fc240351debe491c6c5a7acb6068

Zip/Tarballs

Zip/tar archives	Size	Date	MD5
RStudio 1.0.143 - Windows Vista/7/8/10	117.6 MB	2017-04-19	24ae8674ba0cc0c5310c8172f2e5f12e
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	86.2 MB	2017-04-19	c4444d8159463b8e7a950ba4fe7150e0
RStudio 1.0.143 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	93.2 MB	2017-04-19	c92ecfdb80432a532c823a22630918b
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	85.4 MB	2017-04-19	197280b1c10edae0baccf898a14d9a58f
RStudio 1.0.143 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	86.6 MB	2017-04-19	04afcc818259577c188097955336cae2

Source Code

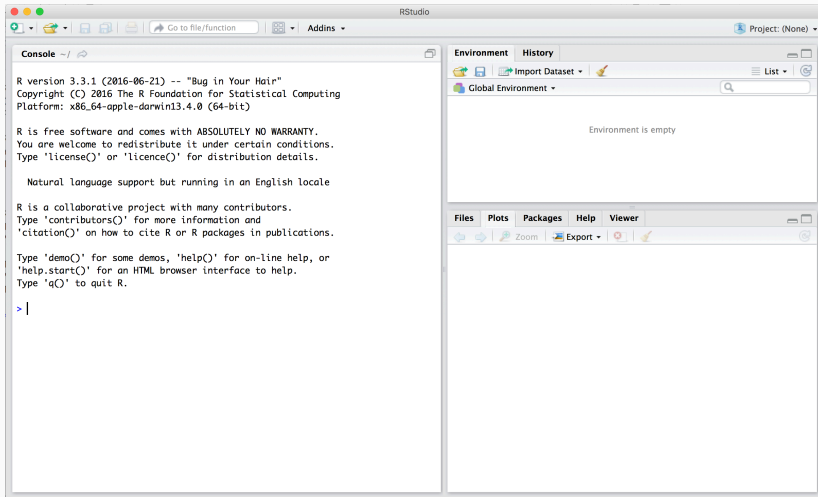
A tarball containing source code for RStudio v1.0.143 can be downloaded from [here](#)

Now, install R or RStudio as you would any other program. It should link automatically to the version of R you just installed.

Running R Scripts

Launch RStudio

Go ahead and launch RStudio. You should see a window that looks like this, as we saw on the class computers last time:

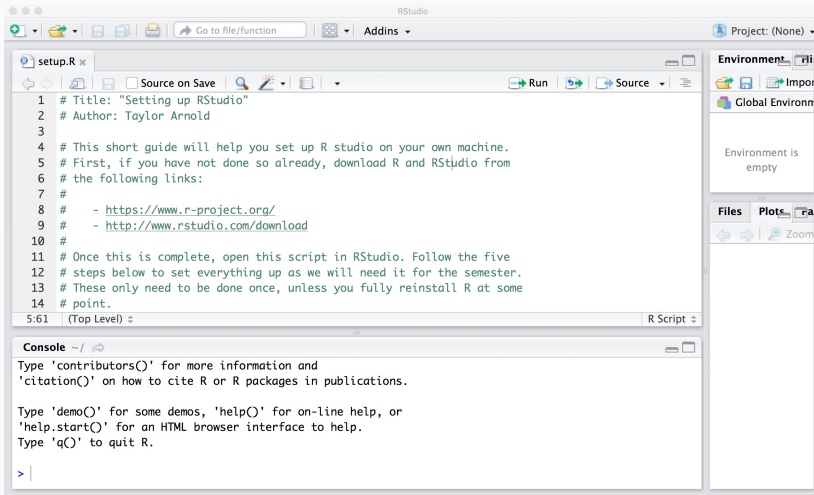


Open setup.R

Download the file *setup.R* from the course website and save it somewhere on your computer (perhaps the desktop?)

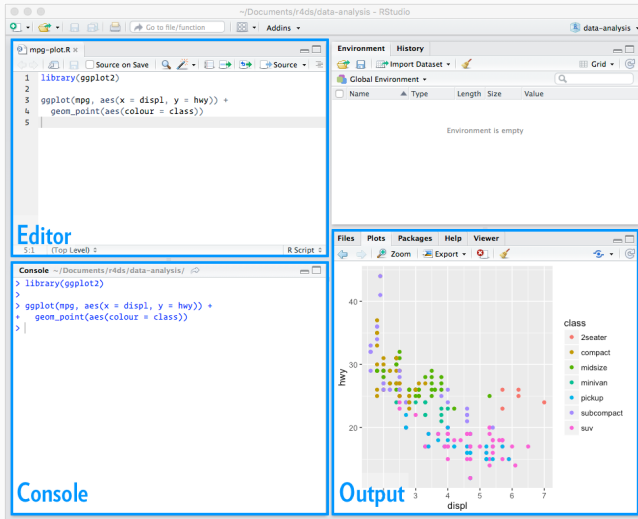
Select File => Open File and open the *setup.R* that you just downloaded.

Open setup.R



Running setup.R

We now have a fourth panel in RStudio, which contains an editor for the code:



We'll cover more about what types of code happen in the code editor. For now, notice that there are two types of code lines:

- ▶ comment lines starting with #, which contain plain text, and are highlighted in green
- ▶ code lines, which contain R commands in a specific format

To run a line of code, highlight the line or put your cursor in it and then hit the **Run** button. The code will run in the Console.

Lab 02

Finally, let's download the lab02.R file from the website (note: there is no lab01.R)

Open the file in RStudio, add your name, and save it.

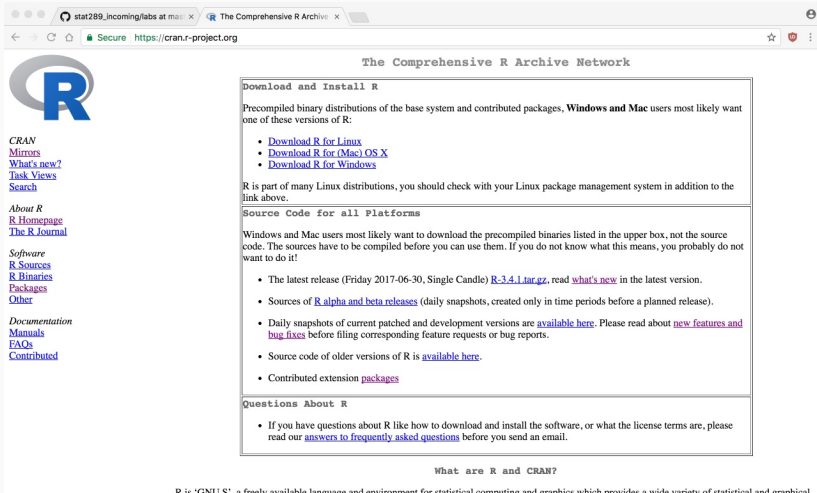
Upload the File

Now, go back to the new GitHub repository you created when you accepted the GitHub classroom invitation.

We want to add **lab02.R** to this repository. On Chrome you can drag and drop the file from your computer; otherwise, click on Upload files and navigate to the file on your computer.

Commit Changes

Once the file loads, scroll down to the bottom of the page and select **Commit changes** (Note: changes will be lost otherwise)



The screenshot shows the CRAN (Comprehensive R Archive Network) website in a web browser. The browser's address bar shows the URL <https://cran.r-project.org>. The page has a navigation menu on the left with links for CRAN, Mirrors, What's new?, Task Views, Search, About R, R Homepage, The R Journal, Software, R Sources, R Binaries, Packages, Other, Documentation, Manuals, FAQs, and Contributed. The main content area is titled "The Comprehensive R Archive Network" and is divided into two main sections: "Download and Install R" and "Source Code for all Platforms".

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](#)
- [Download R for \(Mac\) OS X](#)
- [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 (Friday 2017-06-30, Single Candle) [R-3.4.1.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

Commit Changes

If you navigate to the labs directory, you'll see the **lab02.R** file added to the repository. You can see its contents by clicking on the file.

The screenshot shows a web browser window displaying the GitHub repository page for `statsmaths/stat289_incoming`. The browser's address bar shows the URL `https://github.com/statsmaths/stat289_incoming/tree/master/labs`. The repository page header includes the repository name, a search bar, and navigation links for Pull requests, Issues, Marketplace, and Gist. Below the header, the repository statistics are shown: 1 Unwatch, 0 Stars, and 0 Forks. The main content area shows the file tree for the `labs` directory, with files `lab01.R` and `lab02.R` listed. `lab02.R` is highlighted, indicating it was added 3 minutes ago. The footer of the page shows the GitHub logo and links for Terms, Privacy, Security, Status, Help, Contact GitHub, API, Training, Shop, Blog, and About.

For Next Time

Complete the questions in **lab02.R** and upload the completed file to GitHub (just follow the same steps and make sure you commit the results).