Installing R

Now that we've got a handle on what a data scientist is, how to find answers, and then spent some time going over a data science example, it's time to get you set up to start exploring on your own. And the first step of that is installing R.

What is R? What is CRAN?

First, let's remind ourselves exactly what R is and why we might want to use it.

R is both a programming language and an environment, focused mainly on statistical analysis and graphics. It will be one of the main tools you use in this and following courses.

R is downloaded from the Comprehensive R Archive Network, or CRAN, and while this might be your first brush with it, we will be returning to CRAN time and time again, when we install packages - so keep an eye out!

Why should you use R?

Outside of this course, you may be asking yourself - why should I use R?

The reasons for using R are myriad, but some big ones are:

1) Its popularity

R is quickly becoming the standard language for statistical analysis. This makes R a great language to learn as the more popular a software is, the quicker new functionality is developed, the more powerful it becomes, and the better the support there is! Additionally, as you can see in the graph below, knowing R is one of the top five languages asked for in data scientist job postings!
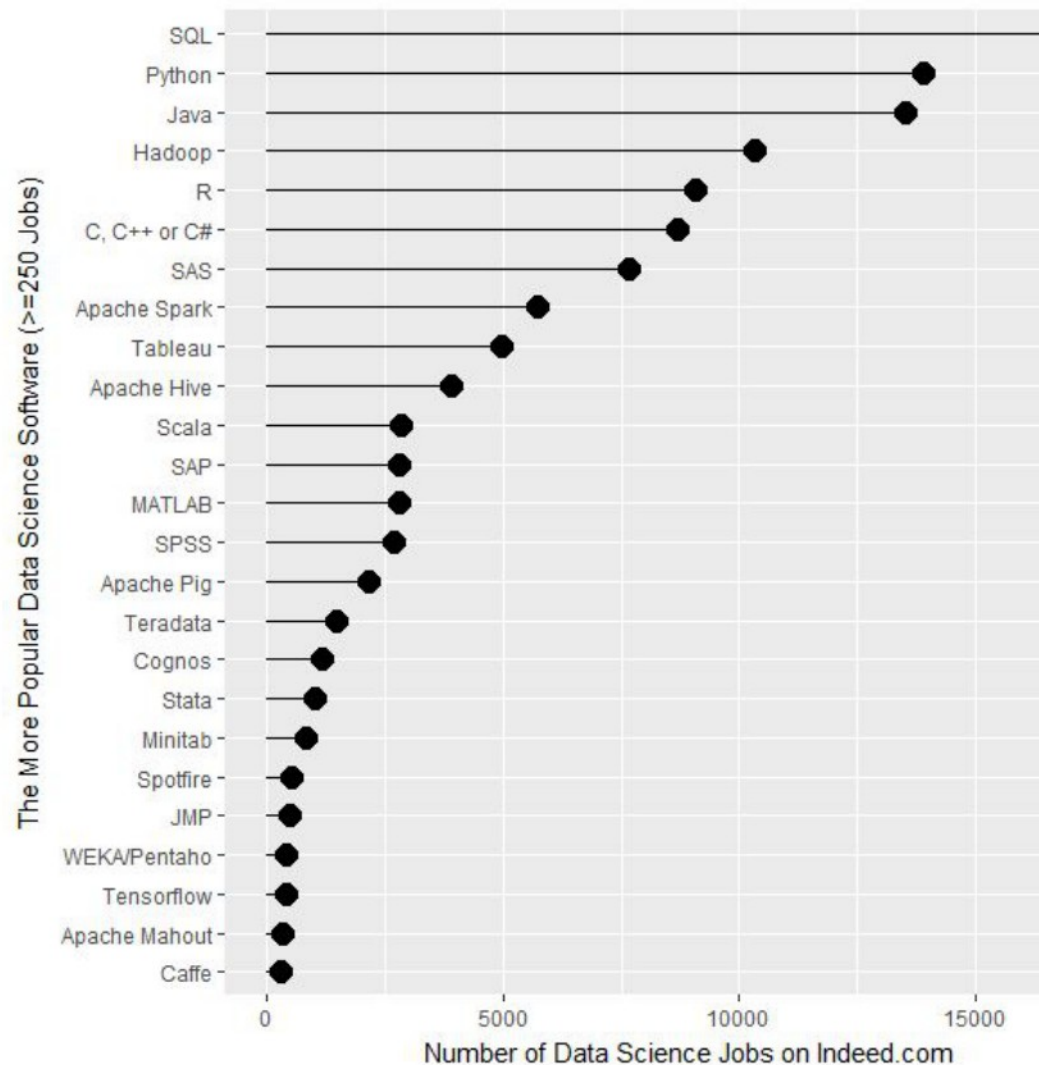
R's popularity among data scientists from r4stats.com

2) Its cost

FREE!

This one is pretty self-explanatory - every aspect of R is free to use, unlike some other stats packages you may have heard of (eg: SAS, SPSS), so there is no cost barrier to using R!

3) Its extensive functionality

R is a very versatile language - we've talked about its use in stats and in graphing, but its use can be expanded to many different functions - from making websites, making maps using GIS data, analysing language... and even making these lectures and videos! For whatever task you have in mind, there is often a package available for download that does exactly that!

4) Its community

And the reason that the functionality of R is so extensive is the community that has been built around R. Individuals have come together to make "packages" that add to the functionality of R - and more are being developed every day!

Particularly for people just getting started out with R, its community is a huge benefit - due to its popularity, there are multiple forums that have pages and pages dedicated to solving R problems. We talked about this in the Getting Help lesson; these forums are great both for finding other people who have had the same problem as you, and posting your own new problems.

Installing R

Now that we've spent some time looking at the benefits of R, it is time to install it! We'll go over installation for both Windows and Mac below, but know that these are general guidelines and small details are likely to change subsequent to the making of this lecture - use this as a scaffold.

For both Windows and Mac machines, we start at the CRAN homepage: https://cran.r-project.org/

Installation - for Windows

If you are on a Windows computer, follow the link Download R for Windows, and follow the directions there - if this is your first time installing R, go to the base distribution and click on the link at the top of the page that should say something like "Download R [version number] for Windows." This will download an executable file for installation.

R for Windows

Subdirectories:

base	Binaries for base distribution. This is what you want to **install R for the first time**.
contrib	Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). Th information on third party software available for CRAN Windows services and correspon environment and make variables.
old contrib	Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; man Ligges).
Rtools	Tools to build R and R packages. This is what you want to build your own packages on V build R itself.

R-3.4.3 for Windows (32/64 bit)

Download R 3.4.3 for Windows (62 megabytes, 32/64 bit)
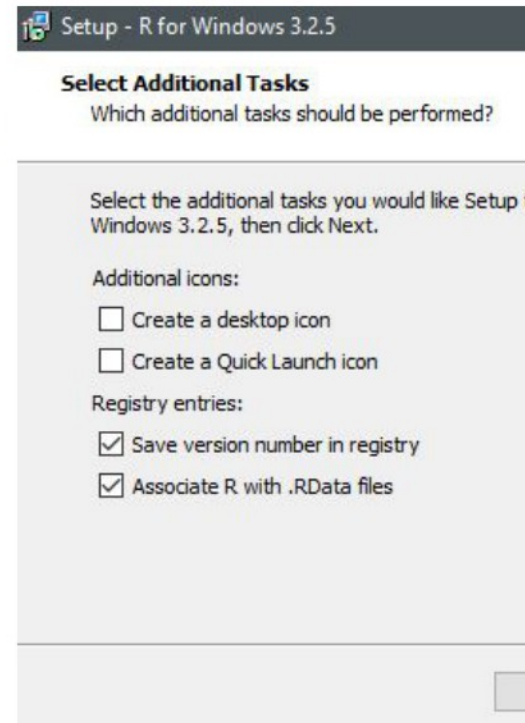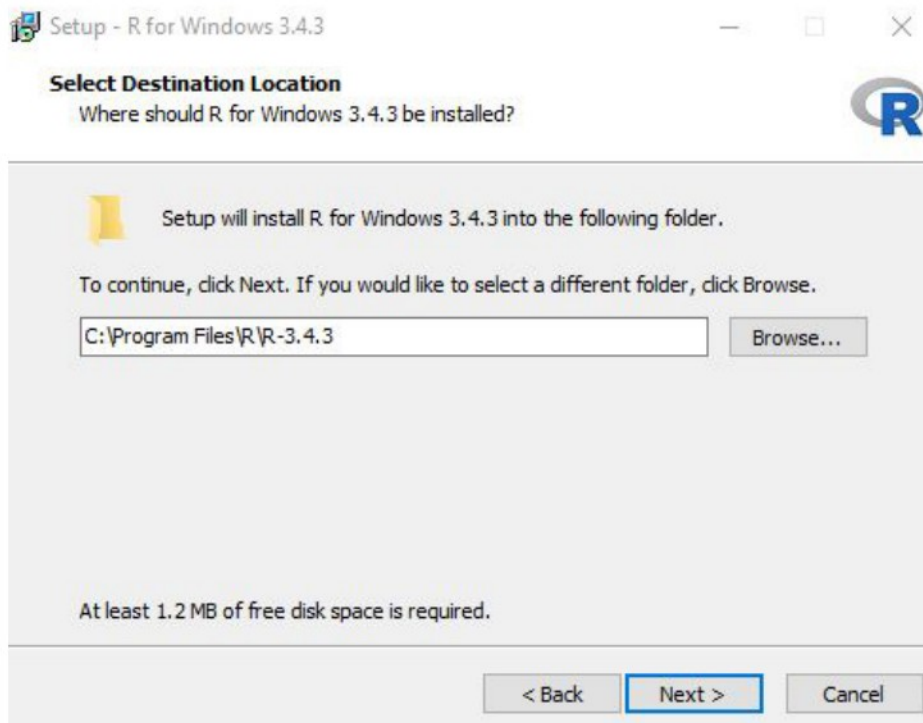Installation and other instructions
New features in this version

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, y of the .exe to the fingerprint on the master server. You will need a version of md5sum for windows: both graphic are available.

https://cran.r-project.org/bin/windows/; https://cran.r-project.org/bin/windows/base/

The download page for R on Windows

Open the executable, and if prompted by a security warning, allow it to run. Select the language you prefer during installation and agree to the licensing information. You will next be prompted for a destination location - this will likely be defaulted to Program Files, in a subfolder called R, followed by another directory of the version number. Unless you have any issues with this, the default location is perfect.
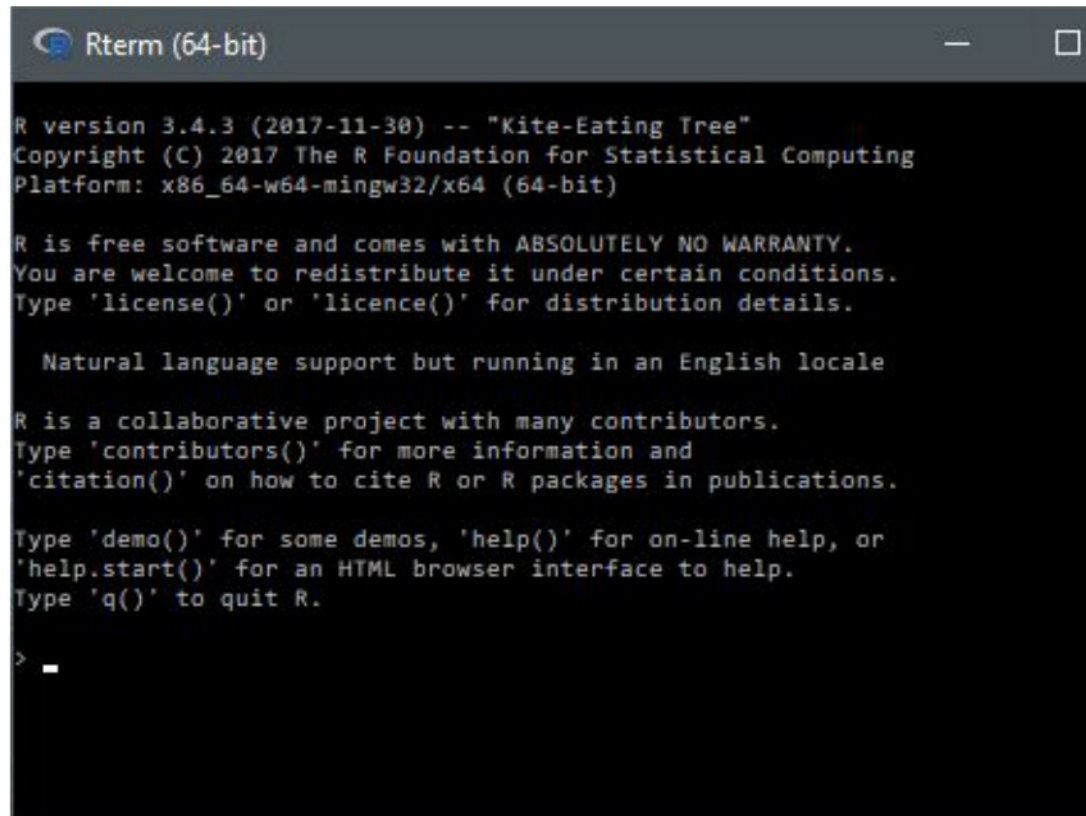
Select Destination Location
Where should R for Windows 3.4.3 be installed?

Setup will install R for Windows 3.4.3 into the following folder.

To continue, click Next. If you would like to select a different folder, click Browse.

C:\Program Files\R\R-3.4.3   Browse...

At least 1.2 MB of free disk space is required.

< Back   Next >   Cancel

Select Additional Tasks
Which additional tasks should be performed?

Select the additional tasks you would like Setup
Windows 3.2.5, then click Next.

Additional icons:

☐ Create a desktop icon
☐ Create a Quick Launch icon

Registry entries:

☑ Save version number in registry
☑ Associate R with .RData files

The install wizard for installing R

You will then be prompted to select which components should be installed. Unless you are running short on memory, installing all of the components is desireable. Next you'll be asked about startup options, and again, the defaults are fine for this. You will then be asked where Setup should place shortcuts - this is completely up to you, you can allow it to add the program to the start menu, or you can click the box at the bottom that says to not create a start menu link. Finally, you will be asked whether you want a desktop or Quick Launch icon - up to you! I do not recommend changing the defaults for the registry entries though.

After this window, the installation should begin. Test that the installation worked by opening R for the first time!

The R terminal running!

Installation - for Macs

If you are on a Mac computer, follow the link Download R for (Mac) OS X. There you can find the various R versions for download. Note: If your Mac is older than OS X 10.6 ("Snow Leopard"), you will need to follow the directions on this page for downloading older versions of R that are compatible with those operating systems.

Click on the link to the most recent version of R, which will download a .pkg file.

```
                              R for Mac OS X
```

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

Note: CRAN does not have Mac OS X systems and cannot check these binaries for viruses. Although we ta
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 arch
versions should adjust the CRAN mirror setting accordingly.

```
             R 3.4.3 "Kite-Eating Tree" released on 2017/11/30
```

Important: since R 3.4.0 release we are now providing binaries for OS X 10.11 (El Capitan) and higher us
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 c
mirroring process. For example type
```
md5 R-3.4.3.pkg
```
in the *Terminal* application to print the MD5 checksum for the R-3.4.3.pkg image. On Mac OS X 10.7 and
validate the signature using
```
pkgutil --check-signature R-3.4.3.pkg
```

```
                                        Files:
```

R-3.4.3.pkg
MD5-hash: d51d9869f3cbe0d782eede113897393a
SHA-
hash: d2694cd4b8d5539deab0e68a73bd79eb715fe62f
(ca. 74MB)

R 3.4.3 binary for OS X 10.11 (El Capitan) and higher, signed package.
R 3.4.3 framework, R.app GUI 1.70 in 64-bit for Intel Macs, Tcl/Tk 8.6
libraries and Texinfo 5.2. The latter two components are optional and ca
ommitted when choosing "custom install", they are only needed if you
use the `tcltk` R package or build package documentation from sources

───────────────────────────○───────────────────────────

Downloading the .pkg file for Macs

Open the pkg file and follow the prompts as provided by the installer. First, click continue on
the welcome page and again on the important information window page. Next you will be
presented with the software license agreement, again, continue. Next you may be asked to
select a destination for R, either available to all users or to a specific disk - select whichever
you feel is best suited to your setup. Finally, you will be at the "Standard Install" page; R
selects a default directory and if you are happy with that location, go ahead and click install.
At this point, you may be prompted to type in the admin password, do so, and the install will
begin!
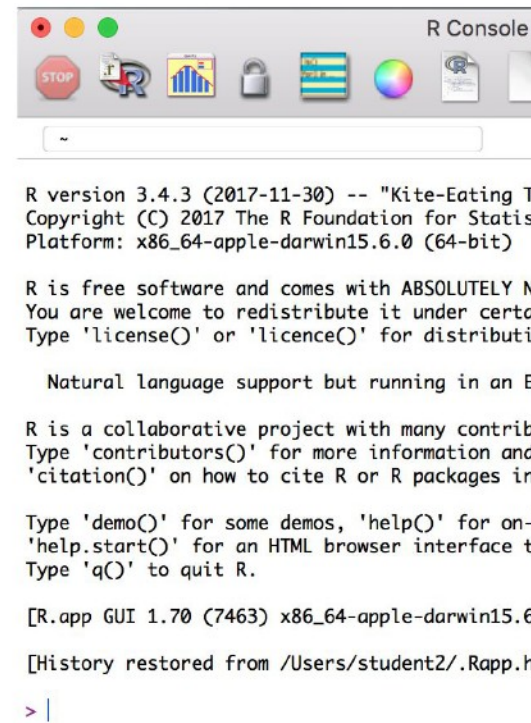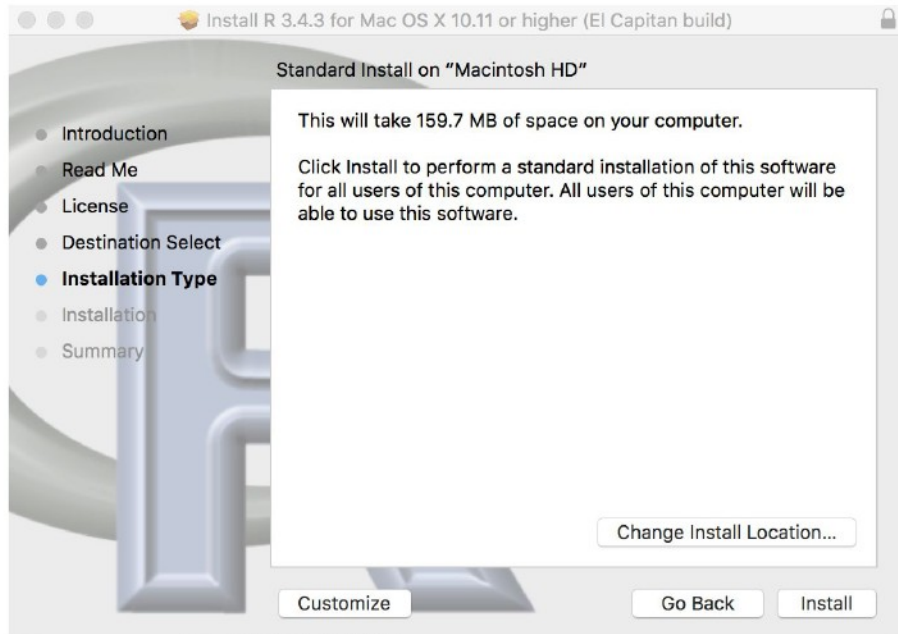
Installation of R on a Mac

Once the installation is finished, go to your Applications and find R. Test that the installation worked by opening R for the first time!

Summary

In this lesson we first looked at what R is and why we might want to use it. We then focused on the installation process for R on both Windows and Mac computers. Before moving on to the next lecture, be sure that you have R installed properly.