Pre-workshop Preparation

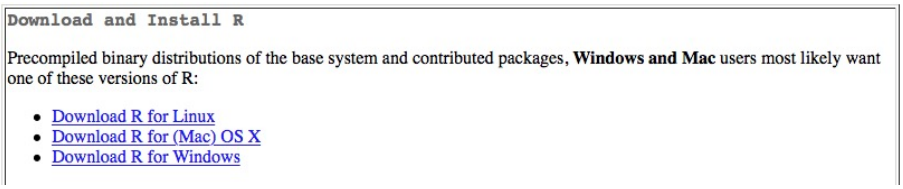
Before taking part in the workshop it would be very helpful if you could download **R** and **R Studio**, and install the packages that will be used in the workshop as this can take a little time. The programs are free and are available for both Mac and Windows.

I have outlined the steps needed to download the programs and packages below.

1 Loading R

To download and install **R**:

1. Go to **https://cran.r-project.org**
2. Select the Mac or Windows download option on the R Download Page
3. Follow the instructions in the sections below for your operating system



Download and Install R

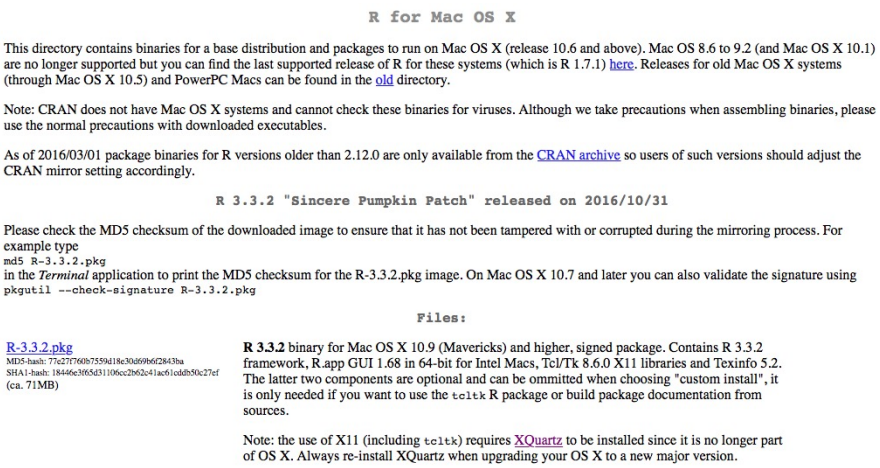
Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

Figure 1: R Download Page

For Mac Users

The following page should now be shown:



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 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 3.3.2 "Sincere Pumpkin Patch" released on 2016/10/31

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-3.3.2.pkg`
in the *Terminal* application to print the MD5 checksum for the R-3.3.2.pkg image. On Mac OS X 10.7 and later you can also validate the signature using
`pkgutil --check-signature R-3.3.2.pkg`

Files:

R-3.3.2.pkg
MD5-hash: 77e27f760b7559d18e30d69b6f2843ba
SHA1-hash: 18446e3f65d31106cc2b62c41ac61cddb50c27ef
(ca. 71MB)

R 3.3.2 binary for Mac OS X 10.9 (Mavericks) and higher, signed package. Contains R 3.3.2 framework, R.app GUI 1.68 in 64-bit for Intel Macs, Tcl/Tk 8.6.0 X11 libraries and Texinfo 5.2. The latter two components are optional and can be ommitted when choosing "custom install", it is 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.

Figure 2: R Download Page for Mac

Download and install R

The download you choose depends on the operating system running on your Mac. To find out your operating system click on the *Apple icon* on your menu bar and select *About This Mac*.

If you are running Mavericks (OS X 10.9) or higher click on the first link to download. Double click on the downloaded .pkg file and follow the installer instructions to download. This may take a few minutes.

If you are running Snow Leopard or Mountain Lion click on the second link - if this is the case, you will not need to download XQuartz.

Check that R has downloaded correctly

Go to *Applications* and double click on the R icon and the page below should open:

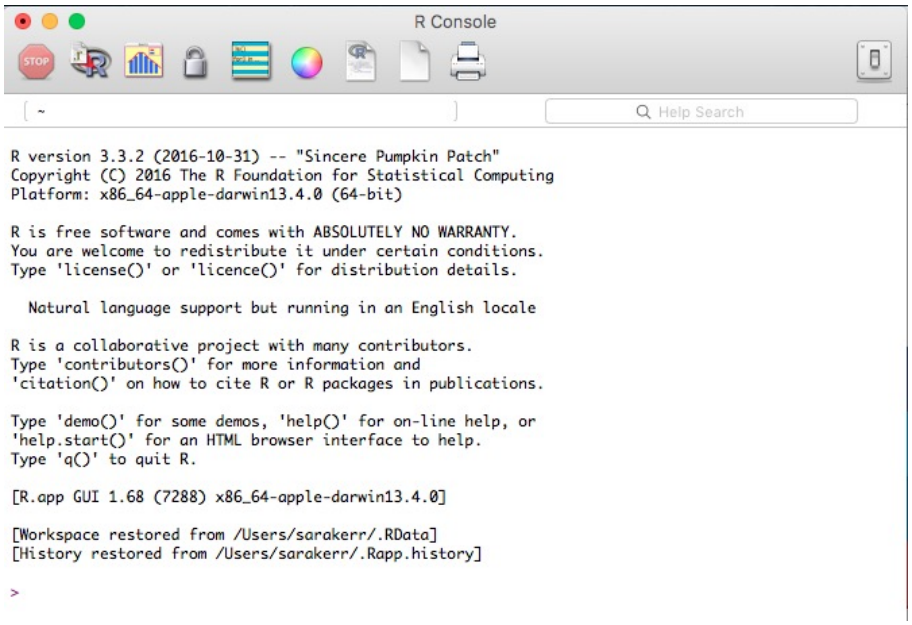


Figure 3: R Console

You can now close the R console.

If this does not work, try installing R again.

Download and install XQuartz

Once you have downloaded R you will also need to download the latest version of **XQuartz**. Click on the link on the R download page or go to **https://www.xquartz.org**. Follow the install instructions.

Figure 4: XQuartz Download Page

Download Xcode and accept licence

Finally, go to the App Store and download Xcode.

Once it has been downloaded, go to *Applications* and open Xcode by double clicking. You may be prompted to install additional elements, if so click *install*. Open Xcode and accept licence. You can now close Xcode.

For Windows Users

The following page should now be shown:

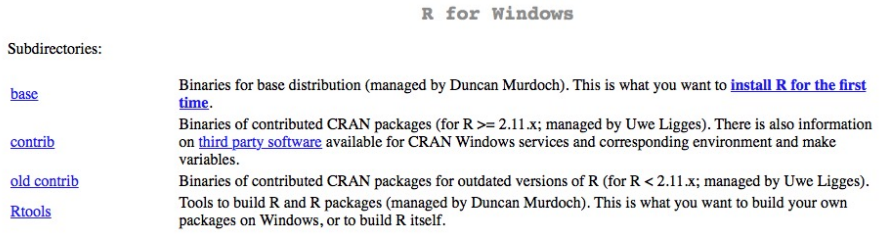


Figure 5: R for Windows

Download and install R Click on the link *install R for the first time*. The following page will be shown:

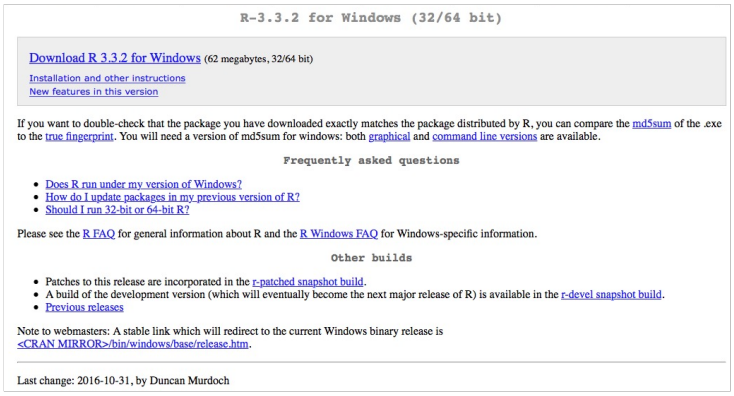


Figure 6: R Download Page for Windows

Click on the link *Download R 3.3.2. for Windows*. Once the .exe file has downloaded, double click on it and follow the installation instructions.

For a video tutorial walk through of the download process see: https://youtu.be/maQJMgQkqoE.

Check that R has downloaded correctly

Open R - a page similar to the page below should open:

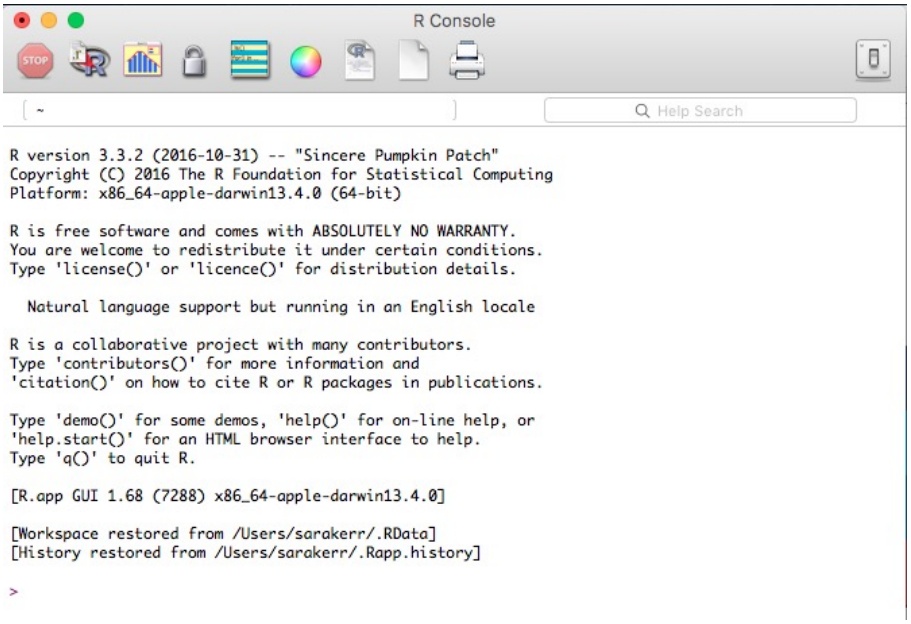


Figure 7: R Console

You can now close the R console.

If this does not work, try installing R again.

2 Loading 'R Studio'

Although R can be used directly, it is easier to use it via an integrated development environment (IDE). One of the most popular IDEs for R is **R Studio**. R Studio is made up of four main panes which allow simultaneous access to the console, document editor, environment and help.

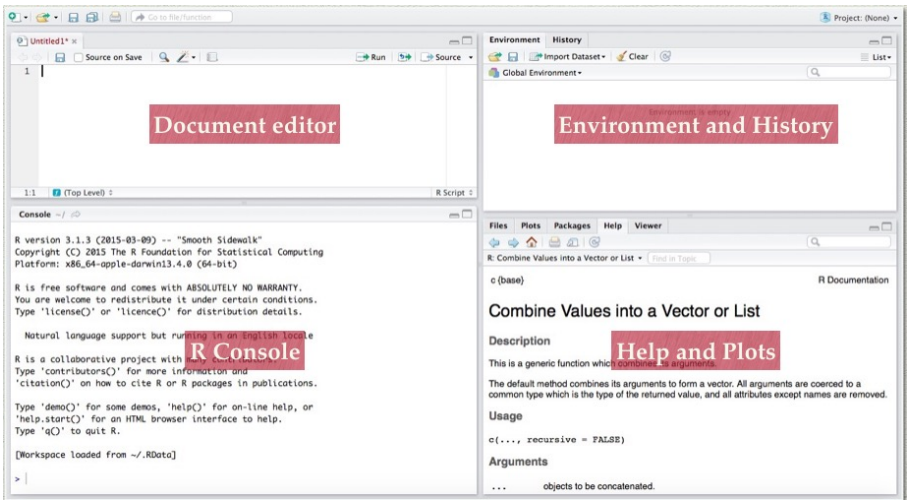


Figure 8: The R Studio Panes

R Studio can be downloaded by going to **https://www.rstudio.com/products/rstudio/download/**. Scroll down the page until you reach the heading *Installers for Supported Platforms*, then select the option for your operating system.

The downloaded file will be .dmg for Mac or .exe for Windows. Double click and follow the installer instructions.

Check that R Studio has downloaded correctly by opening the program.

Updates: From time to time both R and R Studio are updated. This can have an impact on how packages work, especially if you are running an outdated version of R. Make sure you check for updates every 3–6 months.

3 Loading the packages for the workshop

R is set up with **'base' R** which includes a variety of functions and built in data sets. To improve the funtionality of R there is an increasingly broad range of packages available for specific uses. These can be easily added to R as needed. The majority of packages (almost 10000) are available via **https://cran.r-project.org/web/packages/**.

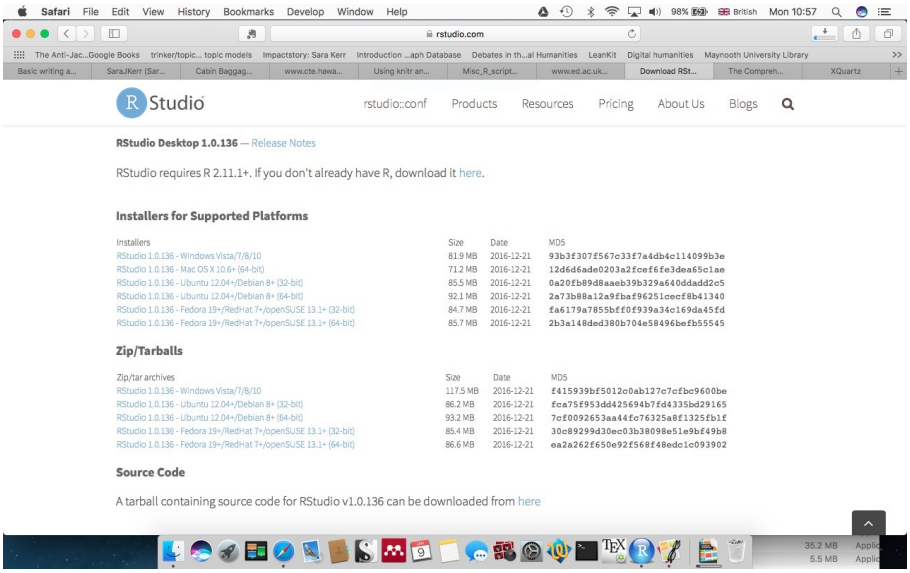


Figure 9: R Studio Download Page

There are two ways to add packages to R in R Studio, the first is by using the *Tools - Install Packages* option in the top menu in R Studio.

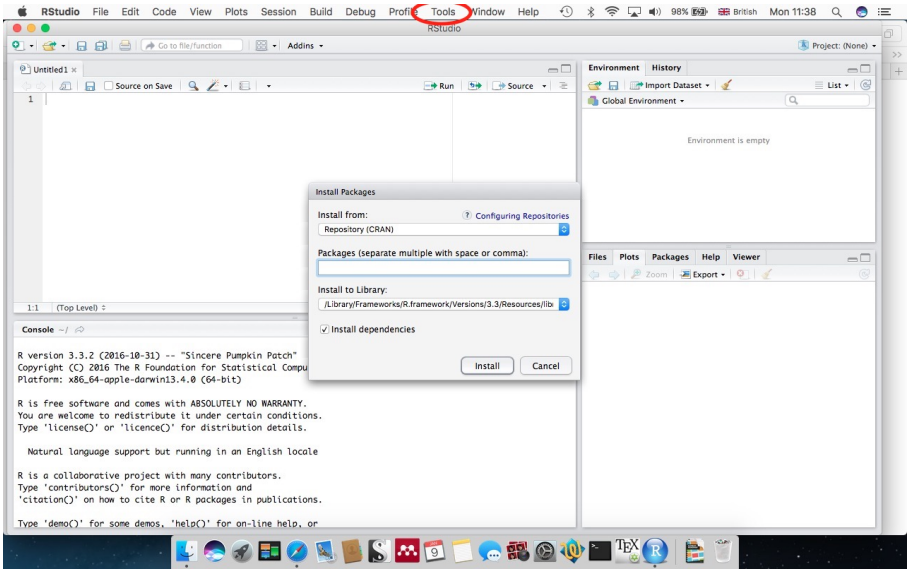


Figure 10: R Studio Install Packages

You will need to write the name of each package you wish to download separated by a comma or space. However, this will only work for packages which are available from the CRAN website.

You will need to install the following R packages for the workshop: "tm", "RColorBrewer", "ggplot2", "wordcloud", "tsne", "Rtsne", "devtools", "stringi", "ggrepel" and "wordVectors".

Run the following piece of code in the R Console to install the packages required for the workshop:

```
# Create a vector of the required packages available from CRAN
needed <- c("tm", "RColorBrewer", "ggplot2", "wordcloud", "tsne",
                    "Rtsne", "devtools", "stringi", "ggrepel")

# install the packages to your laptop
install.packages(needed)

# install the wordVectors package from GitHub
library(devtools)
install_github("bmschmidt/wordVectors")
```

It may take a little time for all the packages to install. Once this has been done, and the command prompt > is showing, you can close R.

4 Learning 'R'

During the workshop you will be given all the code you need to carry out the analysis.

If you wish to get an overview of how R works and familiarize yourself with some of the commands **https://www.datacamp.com/courses/free-introduction-to-r** or **http://tryr.codeschool.com** give a brief practical introduction.

For a more detailed and complex introduction see **https://www.r-bloggers.com/how-to-learn-r-2/**.