Data Visualization Summer Institute
East Stroudsburg University
East Stroudsburg, PA

Installation of R and RStudio

Part I: Installation of R

R is a free software environment for statistical computing and graphics (https://www.r-project.org). It is a GNU project (an operating system that is free software) which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. The name S refers to statistics. A commercial version of S is marketed as S-PLUS by the company Insightful (http://www.insightful.com/). The name R was chosen because R is the letter in the alphabet next to S, and because it was developed by two people whose first names started with the letter R (Robert Gentleman and Ross Ihaka, at the University of Auckland, New Zealand). The interfaces of R and S-PLUS are very different. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS. For more information, see https://www.r-project.org/about.html.

The major web resource for R is http://www.r-project.org/. The latest version is R-3.6.0 (2019-4-26). The download sites are available with mirrors in many countries (https://cran.r-project.org/mirrors.html). For example, some sites in USA

https://cran.cnr.berkeley.edu/	University of California, Berkeley, CA
http://ftp.ussg.iu.edu/CRAN/	Indiana University
http://cran.case.edu/	Case Western Reserve University, Cleveland, OH
http://lib.stat.cmu.edu/R/CRAN/	Statlib, Carnegie Mellon University, Pittsburgh, PA
https://cran.fhcrc.org/	Fred Hutchinson Cancer Research Center, Seattle, WA

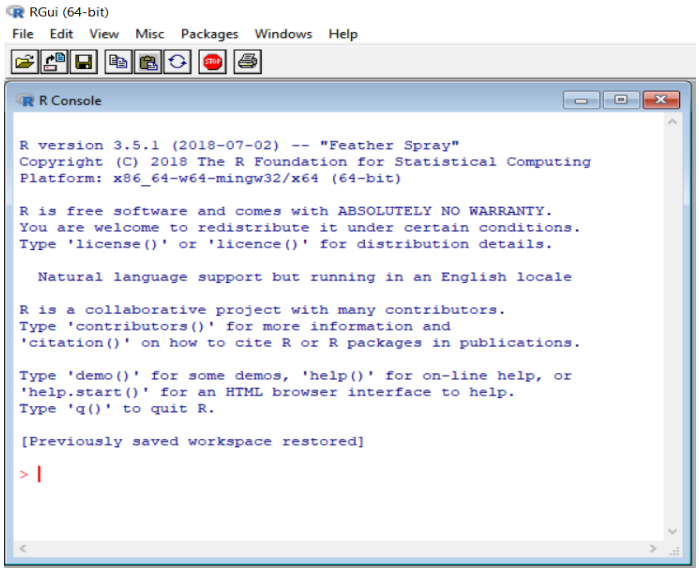
One of the great benefits of R is its community of users and the many user-contributed packages they're willing to share with the rest of us. Currently, the CRAN package repository features about 15,000 available packages (https://cran.r-project.org/web/packages/).

The following shows you how to install R and R packages.

Step 1: Download R by going to one of the mirrors such as https://cloud.r-project.org/. Choose the Linux, Mac or Windows version depending on your operating system. You

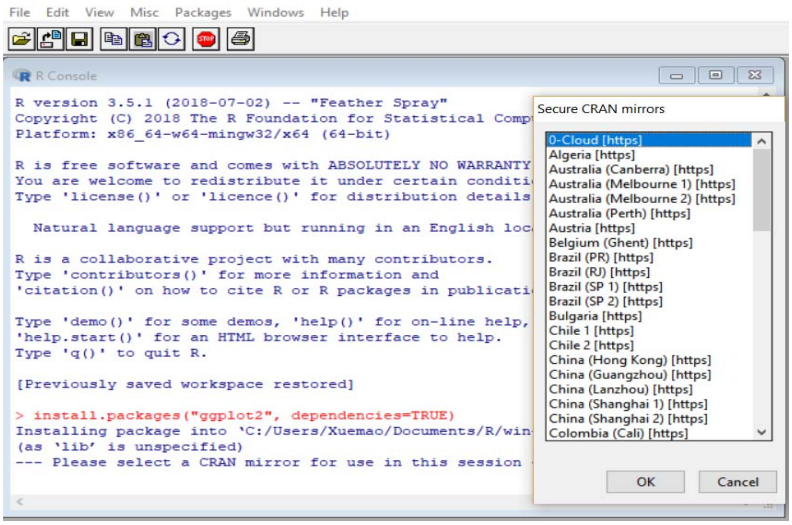
then install R by double click the downloaded file. Accept all the defaults during installation.

Step 2: Start R as administrator (Windows users need to right-click the R icon and click the corresponding item); the following is the R interface. You can enter commands one at a time at the command prompt (>) or run a set of commands from a source file. Keep in mind that R is case-sensitive.

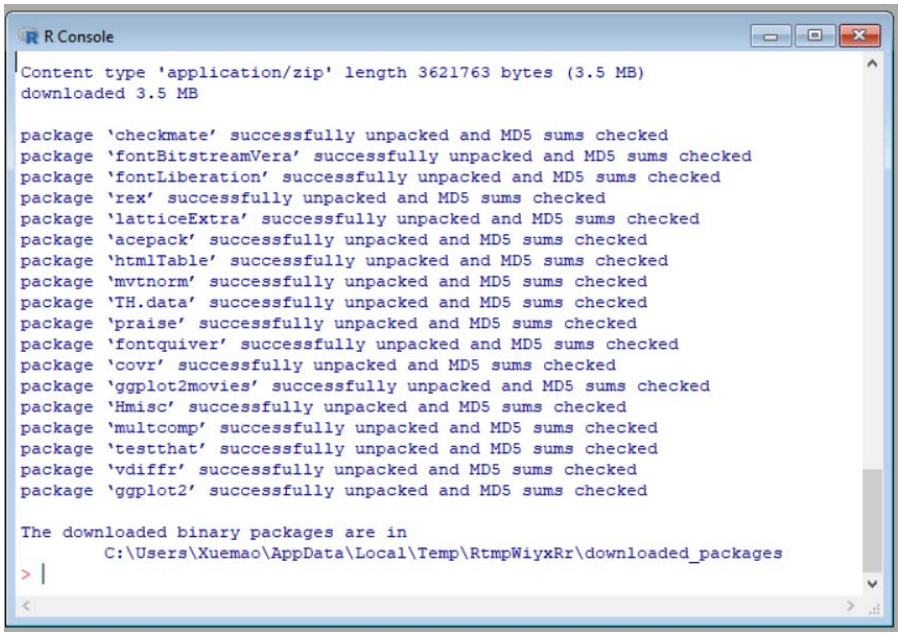


To install an R package, for example ggplot2, you can run the following command after the command prompt (>):

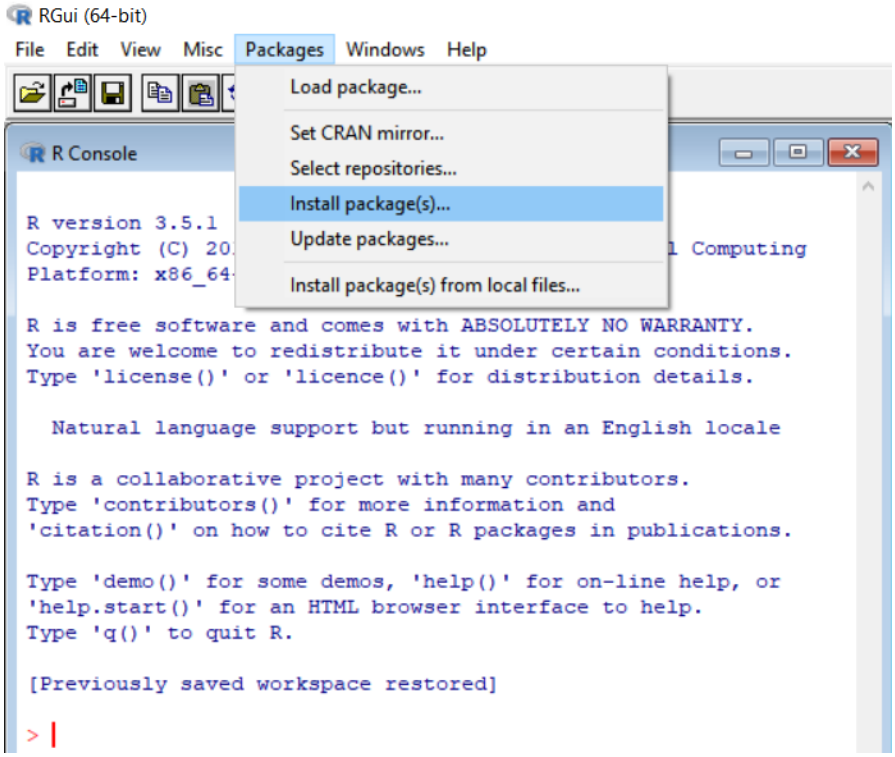
```
install.packages("ggplot2", dependencies=TRUE)
```



The option `dependencies=TRUE` lets you install uninstalled packages which these packages depend on/link to/import/suggest (and so on recursively). Then you choose any mirror and click "OK" to continue.



Another way to install a package is to use the R menu shown below.



For Mac users, if you do not have the latest software, you should first update your Mac software and secondly install XQuartz.app (XQuartz, https://xquartz.macosforge.org/) if you do not have X11.app because installation of lots of R packages requires XQuartz to be installed. If you update your macOS version, you should re-install R (and perhaps XQuartz). For more information, see https://cran.r-project.org/doc/manuals/r-release/R-admin.html.

To use the package ggplot2, type the following after the command prompt (>):

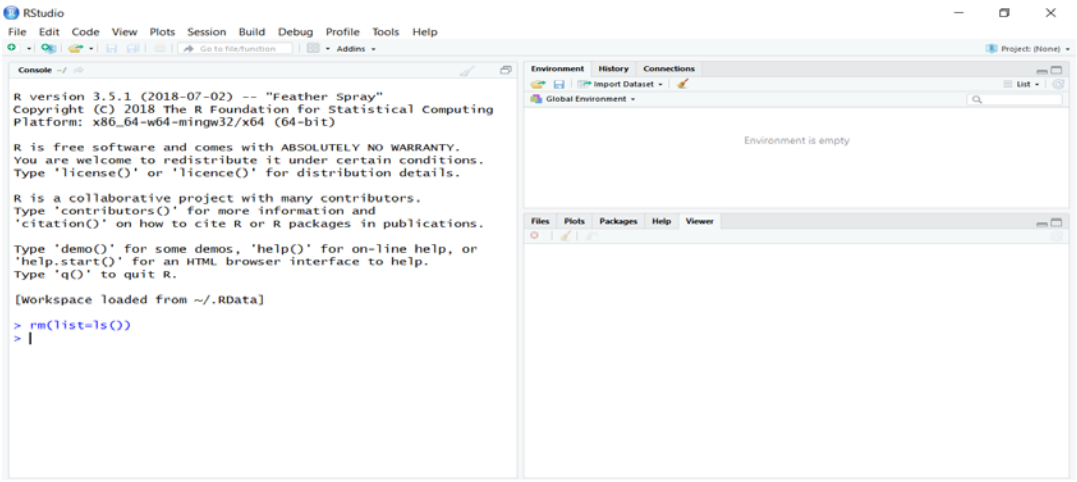
```
library(ggplot2)
```

If you have installation problems, see the guide to installation and administration for R: https://cran.r-project.org/doc/manuals/r-release/R-admin.html

Part II: Introduction to RStudio

RStudio (https://www.rstudio.com) is a popular IDE (Integrated Development Environment) for R programming. RStudio makes R easier to use. It is a powerful editor for R coding and debugging. It is also a powerful generator for HTML, PDF, dynamic documents and slide shows. Download and install it from https://www.rstudio.com/products/rstudio/download/. Again, choose the appropriate installer depending on your operating system.

RStudio includes a code editor, debugging & visualization tools. When you first start up RStudio, it will likely look something like the following figure.



The area on the left is an interactive console which is just the R console.

The two panes on the right become useful as you create and run your code.

- The "Environment" tab shows your R Environment – what objects are loaded into your session at the moment. If you're new to programming, don't worry; this will make more sense once you start coding.
- The "History" tab shows your command history. So if you typed 3+5 into the console, if you go to the History tab, you should see that in the history tab. You can select one or more lines in the history tab and then click the "To Console"" button at the top of the pane to send the line(s) back to the console, or the "To Source" option to send them to the top-left script pane. You can search the history pane as well (you should see a search box at the upper right).

The lower right pane has several different, useful tabs.

- The "Files" tab, similar to Windows File Explorer or Mac Finder. Although not quite as robust as those, this area is convenient for quickly renaming or deleting files, opening files, or changing your working directory.
- The "Plots" tab is where you can view graphs and other data visualizations you create in R.
- The "Packages" tab shows what packages are 1) available for you to use and 2) actually loaded into your working session. Anything listed is on your system; anything with a check mark to the left of the name is currently loaded in memory.
- The "Help" tab is where you can view help files for functions and packages. You'll likely be using that a lot, no matter how expert you become in R. If you click the home button on the help tab (it's the house icon), you'll see links to a lot of general R and RStudio information – some of it for beginners, some considerably more advanced. But you can also ask for more specific help in the R console for functions and packages, and search for help by keyword.
- The "Viewer" pane can be used to view local web content.

RStudio is what's known as an IDE, or Integrated Development Environment. One common feature of IDEs is projects. In RStudio, opening a project automatically sets you up in the project's working directory, making it easier to find files stored in that directory. You can create a new project by going to File > New Project. You can create a new project by going to File > New Project. You'll be given three choices: New Directory (for a brand new project with nothing in it), Existing Project (to create a project from an existing directory that might already have files in it), and Version Control.