



Getting Started With R

Installation and Configuration of R for the
APSY510/511 Statistics Courses and
Other BCD classes/workshops

B. Dudek
8-28-17 version

What is R?

- An integrated programming language for
 - data handling
 - statistical analysis
 - graphics
- Open source (free) software
- A fast growing choice among statisticians, researcher, and educators
 - <http://www.nytimes.com/2009/01/07/technology/business-computing/07program.html>

R, the Software, Finds Fans in Data Analysts - NYTimes.com - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.nytimes.com/2009/01/07/technology/business-computing/07program.html

Most Visited Dynamic Drive Forums Dudek Autism Genetic... University at Albany L... ILLiad Logon Your online guide to b... hmbirds : Hudson-Mo...

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

The New York Times Search All NYTimes.com

Business Computing

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

Search Technology **Inside Technology** Bits Personal Tech »
 Internet Start-Ups Business Computing Companies Blog » Cellphones, Cameras, Computers and more

Innovation. High-Tech. Jobs. GLOBALFOUNDRIES
 Learn more about the future of innovation in America. **AMD** The future is fusion

Data Analysts Captivated by R's Power



Stuart Issett for The New York Times

R first appeared in 1996, when the statistics professors Robert Gentleman, left, and Ross Ihaka released the code as a free software package.

By **ASHLEE VANCE**
 Published: January 6, 2009

To some people R is just the 18th letter of the alphabet. To others, it's the rating on racy movies, a measure of an attic's insulation or what pirates in movies say.

R is also the name of a popular

More Articles in Technology »

Customize your headlines.
 The New York Times **nytimes.com** | **LinkedIn**
 The New York Times and LinkedIn have teamed up to give you customized industry news. Automatically.
 Click here to activate

See how you can accelerate application deployment time with our storage solutions to help your company go further, faster.
 Watch Our Webcast

NetApp

Done zotero

What is R?

- A powerful core program that can do statistical analysis at levels ranging from simple introductory methods to advanced modern modeling and multivariate methods
- An object oriented programming language that has tremendous flexibility
- A program that gives the analyst complete control over all aspects of the analysis

What is R?

- The core program has extensive capabilities in statistics/graphics
- Add-on programs (called packages) provide tremendous versatility and extend the capabilities
 - over 8900 packages now available (as of 8-24-2016)

What is R?

- A multi-platform program
 - Unix
 - Linux
 - Macs
 - Windows PCs
- R can interact well with other software
 - for example, add-ins for MS Excel permit direct sharing of data and functions

What is R?

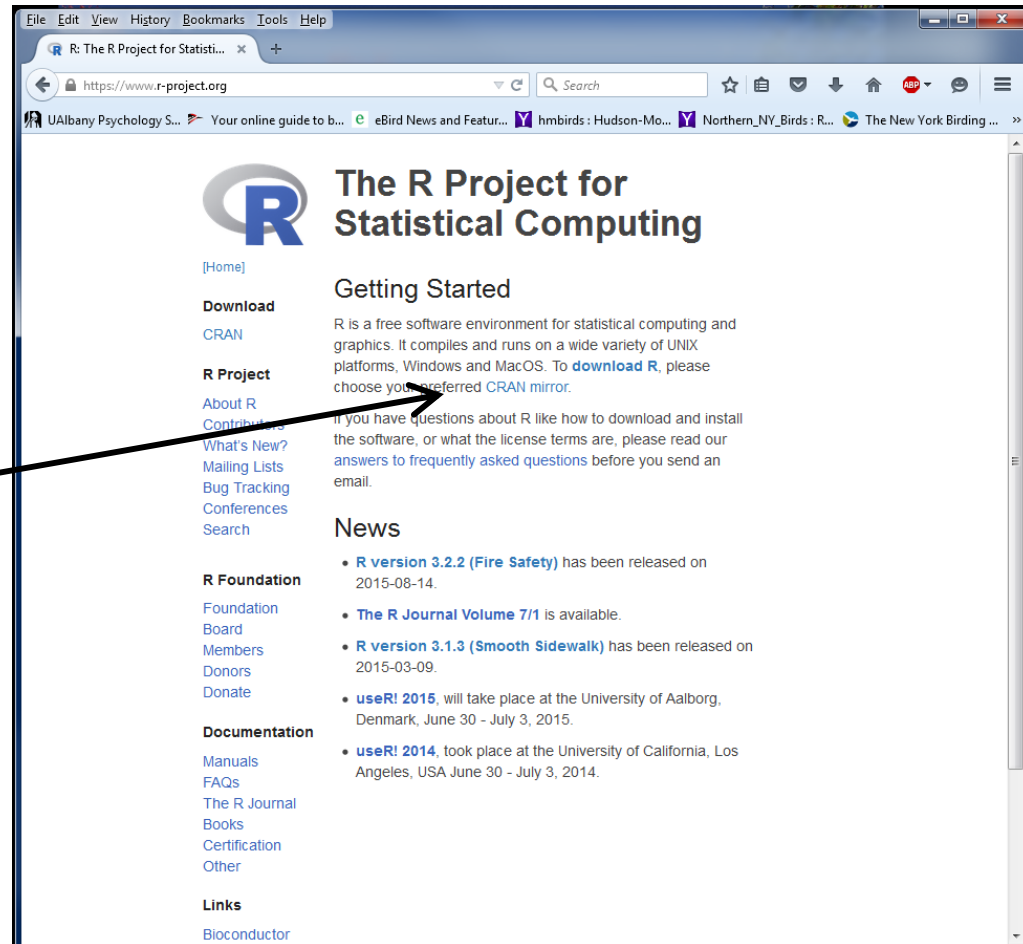
- A difficult program to learn?
 - if computer skills are limited to the menuing type of environment found in Microsoft software, perhaps
 - but ample help/tutorials are available
- A program that is worth the effort to learn since modern commercial statistical software is so expensive
- Did I say that R is free?

What is R?

- Much more than a Statistics program
- It is one of the central tools for modern data science
- Flexible capabilities for both the scientific researcher (like us) and the high level data scientist/programmer/coder

How do I obtain/install R?

Download the program by choosing a mirror site of CRAN



R Project Home Page: <http://www.r-project.org/>

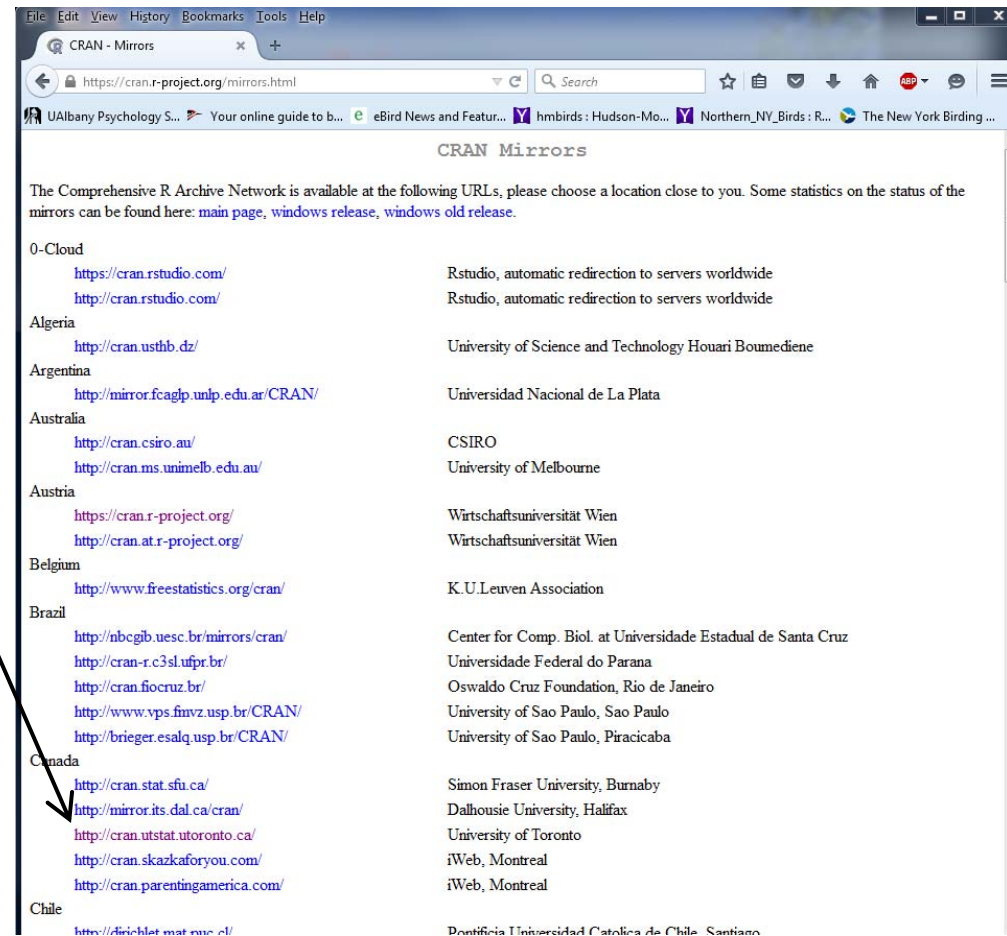
R Installation and Configuration for
APSY510/511 and B. Dudek workshops

CRAN is an Archive Network for R

I usually use the Toronto Site

Note that as of this writing (Aug 2016), R repositories are switching to secure server capability. Only R version 3.2.2 and later can use https sites. Right now, cran mirrors are a mix of http and https sites. Best to use https going forward.

If your initial list shows only https sites, then use the USA (MI1) site.



<http://cran.r-project.org/mirrors.html>

Which Platform?

- We will only use the PC/Windows platform in the 510/511 class
- But you can install onto which ever platform you use on your own computer
 - virtually all functions are the same

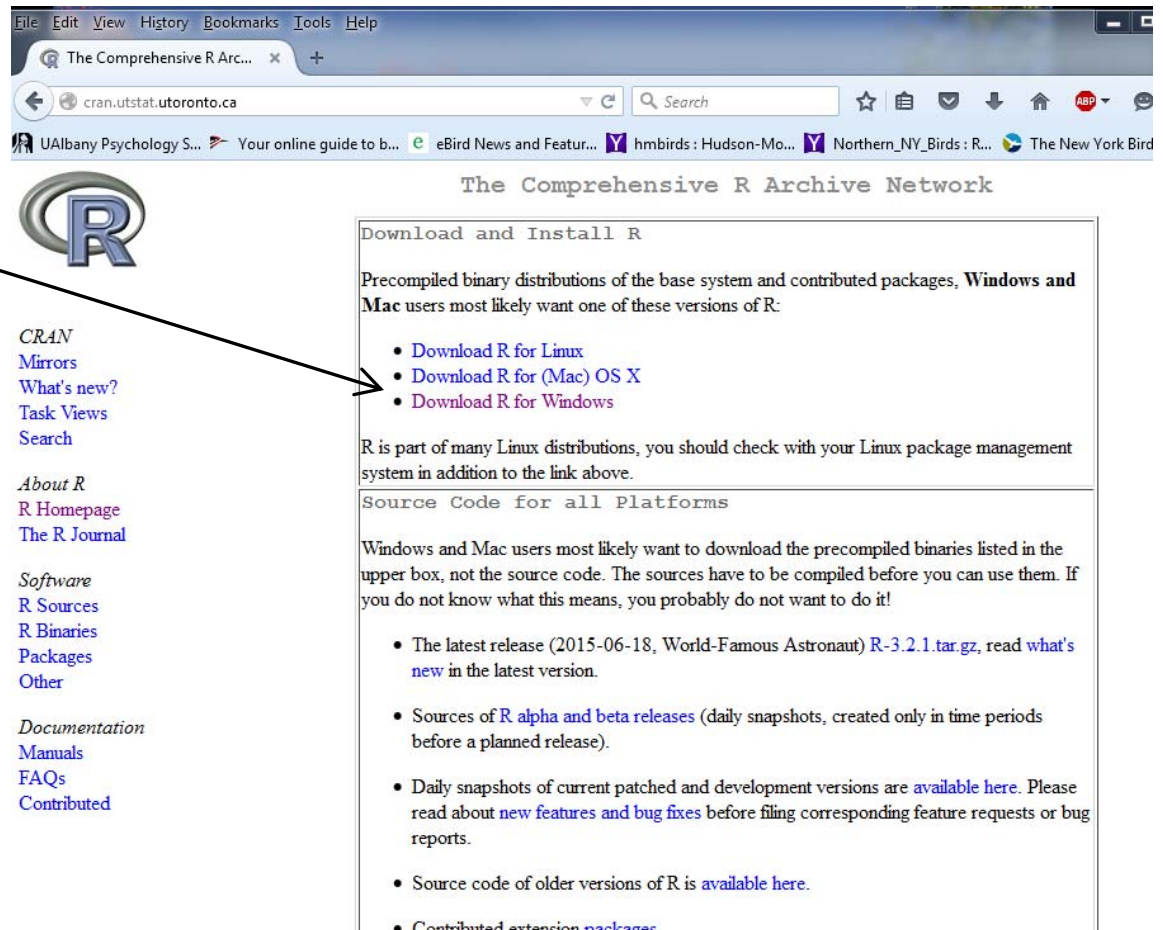
Which Platform?

- The next few slides show Windows (7) installation
- Other OS's should be similar
- Later slides summarize some things for Mac OS installations.
- Mac OS users should look through the windows install slides to get a sense of the relevant steps

R is updated every few months

The Windows version is a full version

Choose the windows version



The screenshot shows the CRAN website with the R logo and navigation links. The main content area is titled 'The Comprehensive R Archive Network' and contains a section 'Download and Install R'. This section provides instructions for downloading precompiled binary distributions for Windows and Mac. A list of links is provided: 'Download R for Linux', 'Download R for (Mac) OS X', and 'Download R for Windows'. An arrow points from the text 'Choose the windows version' to the 'Download R for Windows' link. Below this, there is a section 'Source Code for all Platforms' which explains that Windows and Mac users should download precompiled binaries rather than source code.

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 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

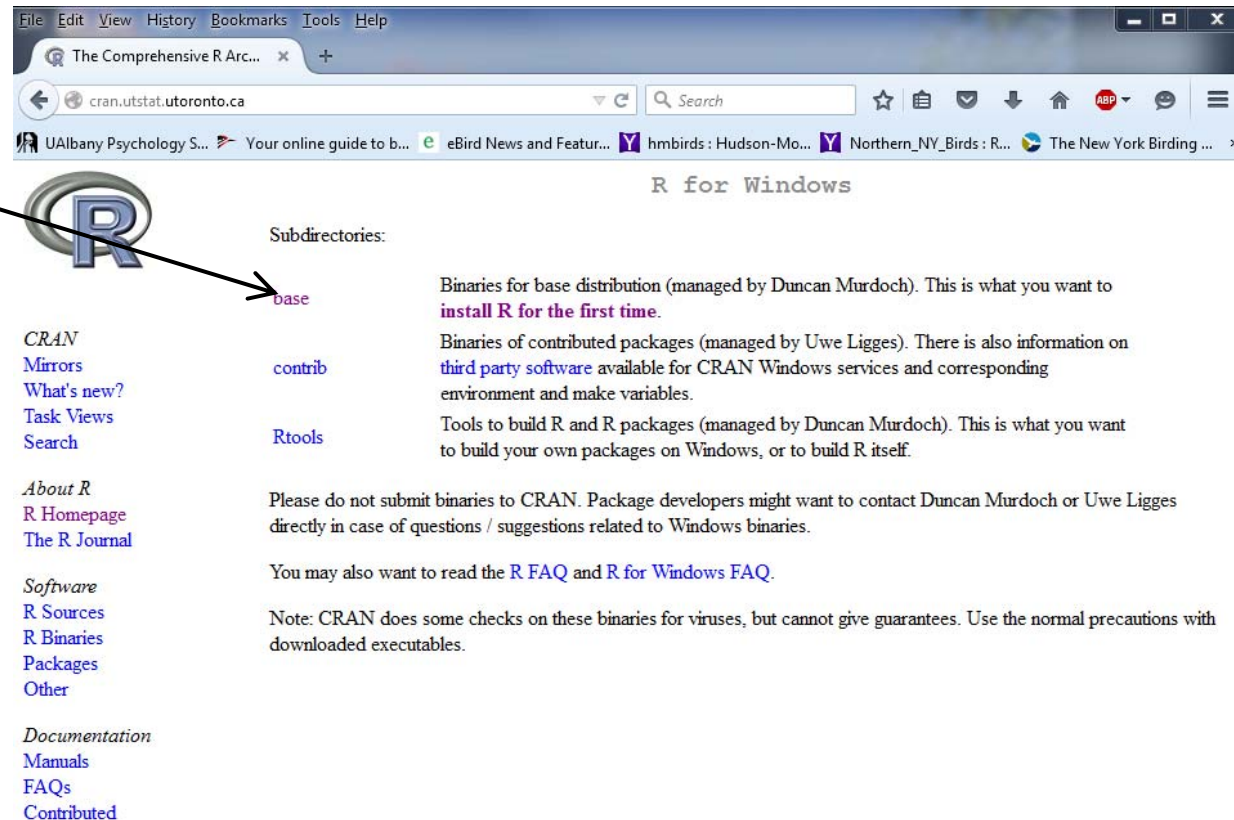
Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2015-06-18, World-Famous Astronaut) [R-3.2.1.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](#)

<http://cran.utstat.utoronto.ca>

The Base R Program can be downloaded as one install file here

install the base



The screenshot shows a web browser window displaying the CRAN website for Windows binaries. The address bar shows cran.utoronto.ca. The page title is "R for Windows". The main content area is titled "Subdirectories:" and lists three options: "base", "contrib", and "Rtools". The "base" subdirectory is highlighted with a pink background and a blue arrow pointing to it from a callout box on the left. The "base" description states: "Binaries for base distribution (managed by Duncan Murdoch). This is what you want to **install R for the first time.**". The "contrib" description states: "Binaries of contributed packages (managed by Uwe Ligges). There is also information on **third party software** available for CRAN Windows services and corresponding environment and make variables." The "Rtools" description states: "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." Below the subdirectories, there is a note: "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." and a link to "R FAQ and R for Windows FAQ". At the bottom, there is a note: "Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables." The left sidebar contains 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".

R is up to Ver 3.4.1 now, and possibly higher by the time you read this. For windows, both 32 bit and 64 bit versions are installed if the machine is running a 64 bit operating system. Apple products should all be 64 bit. If you are in a 32 bit windows machine, the R Installation should detect that.

Download file to local PC (see next pg)

Read Instructions

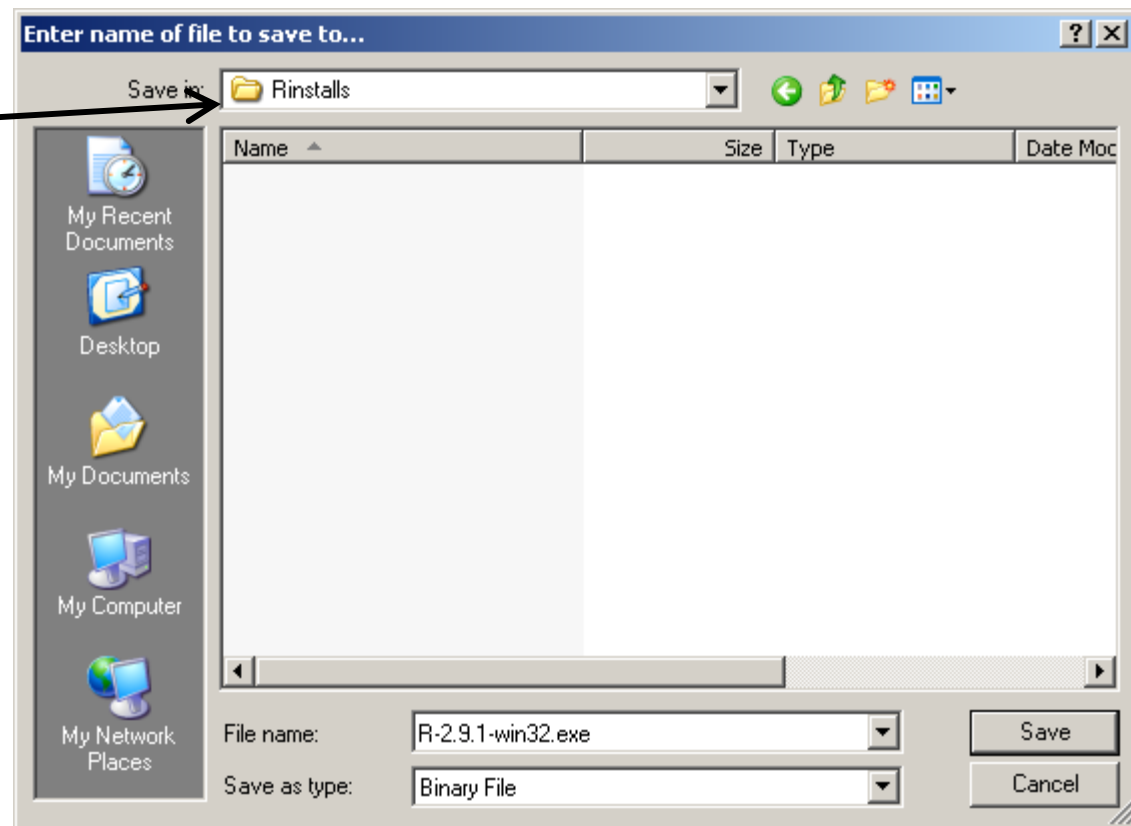


Make sure the mirror site you chose provides R version 3.4.1 for the Fall semester 510 course to match what is in the classroom!

Save the executable to your local machine

Save the installer executable to a folder you can remember

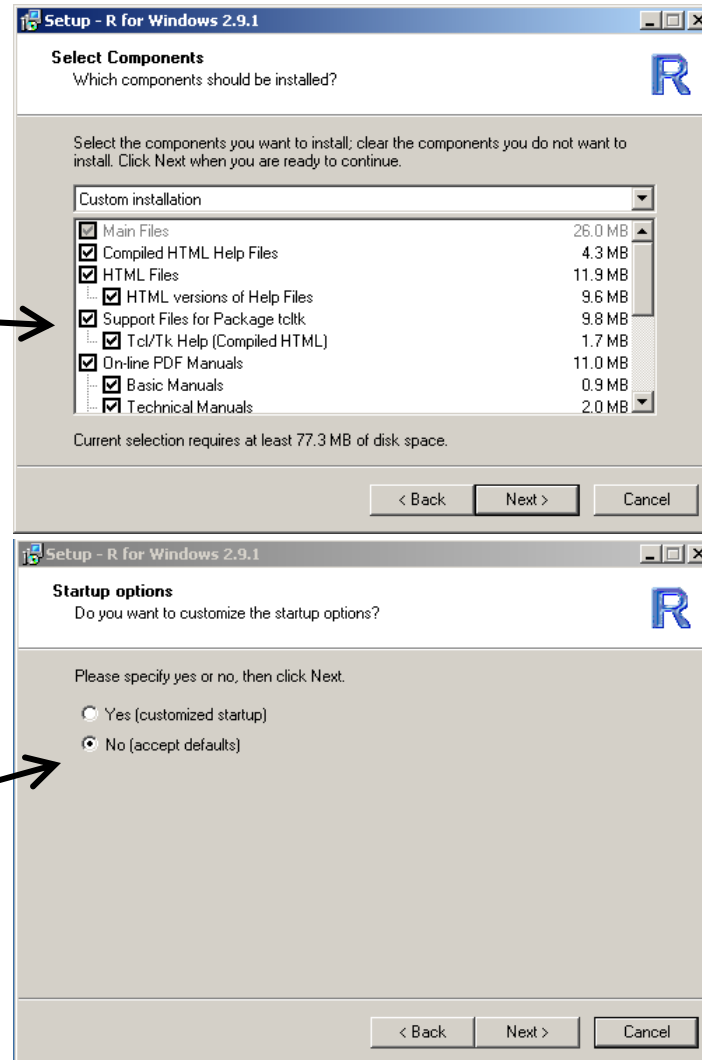
Then, run the installer.



Run the install program you saved

Install all components

Use the default configuration



When the installation is finished...

- An R icon should be on the desktop
 - run it
- If the icon is not on the desktop, run R from the start menu

A few pointers for Mac OS Installation

- As of version R 3.2.2, R requires Mac OS X version 10.9 (Mavericks) or higher.
- R 3.2.1 supports Mac OS X version 10.6 (Snow Leopard) or higher.
- It is strongly recommended that you update the OS and install the latest version of R

General steps for Mac OS Installation

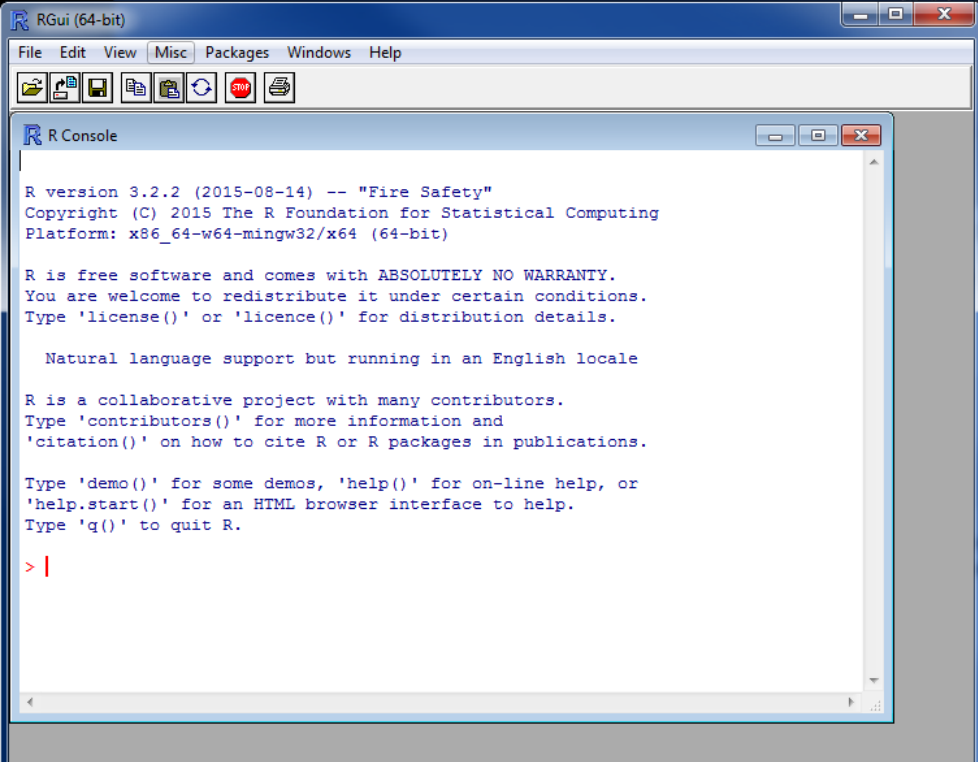
1. Go to <https://www.r-project.org/>
2. In order to download the R installer, choose a preferred CRAN mirror.
 - For the https sites, I recommend USA (MI1), which is <https://cran.mtu.edu/>
 - For http sites, I typically use the Univ Toronto Statistics Dept. site, which is <http://cran.utstat.utoronto.ca/>
3. Download the newest version of R for (Mac) OS X
 - Currently the R-3.4.1.pkg
4. Open the installer and proceed following instructions
5. Ensure that the R program is available via Applications within Finder or on your desktop dock

General steps for Mac OS Installation

- Current versions of Mac OS may not have native x11 capability, unlike previous versions.
- Mac users should always have the most up to date version of XQuartz
- See the Mac installation page at:
- <https://cran.r-project.org/bin/macosx/>
- Another link that provides a nice outline of Mac installation steps can be found at <http://rud.is/b/2015/10/20/installing-r-on-os-x/>
 - This page also describes the XQuartz installation issue
 - Also see the notes on R Commander installation on Mac a few pages down in this document

The R GUI (screen captures from Windows)

- The Main R interface is called RGUI
 - but it is not really a full graphic interface
 - At the R Console prompt you write code/functions
 - Later we will work in an integrated development environment called RStudio



The screenshot shows the R GUI (64-bit) window. The title bar reads "R GUI (64-bit)". The menu bar includes "File", "Edit", "View", "Misc", "Packages", "Windows", and "Help". Below the menu bar is a toolbar with icons for file operations and execution. The main area is the "R Console" window, which displays the following text:

```
R version 3.2.2 (2015-08-14) -- "Fire Safety"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

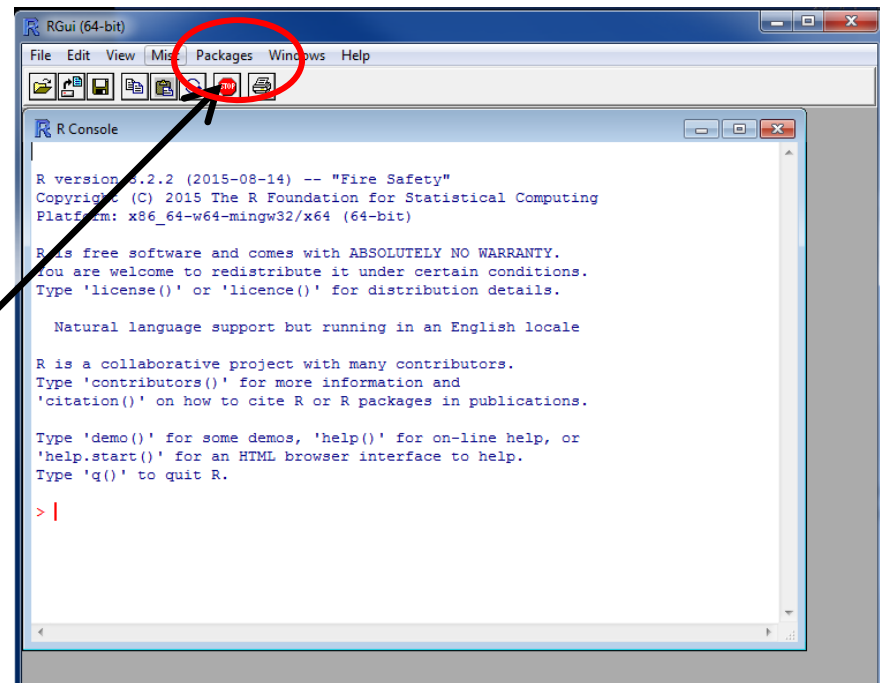
R Programming

- Addressed in other documents at other times in the course or workshop

You will need to install some add on “packages”

Either use this menu, or “source” a syntax file that I provide.

Recommended packages are discussed below



Some Packages are already Installed with the base program

- base
- boot
- class
- cluster
- codetools
- datasets
- foreign
- graphics
- grDevices
- grid
- KernSmooth
- lattice
- MASS
- Matrix
- methods
- mgcv
- nlme
- nnet
- rpart
- spatial
- splines
- stats
- stats4
- survival
- tcltk
- tools
- utils

Additional Packages Suggested for APSY510/511

- A partial list of these is available via the “Learning R” link on the Dudek Class Web Pages: <http://www.albany.edu/psy/bcd/dudekclass.htm>
- You can install them via the menus as indicated above,
 - OR,
- Preferably, you can run (“source”) the script file provided (bcd_recommended_510.R)
 - OR,
- R determines which packages it has available on the local installation by scanning the sub-folders in its library folder. So, you can install additional libraries either from the program as indicated above, or by copying folders of already-installed libraries from someone else (such as BCD). This may not work well across differing platforms.
- More package installation info would probably be proved in class or in workshops.

A few tips on Managing the R Installation

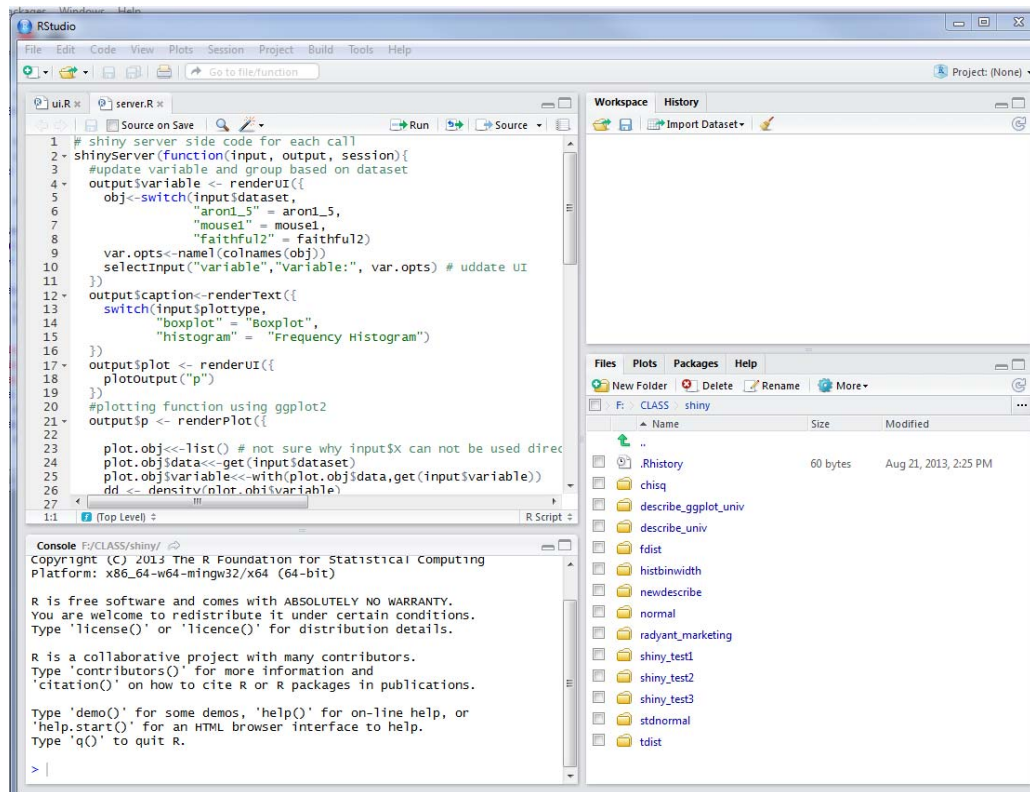
- R versions are upgraded every few months.
 - Upgrades should be done by uninstalling your current version and then installing the new
 - old collections of installed libraries will not be deleted and can usually be used in their same location. It is important to keep package versions up to date with the `update.packages()` function.
 - **A better approach to package installation in a new installation is to use the `bcd_recommend_510.R` script described above. But you will need to also install other packages that you might have added on your own**
- You can also update your currently installed libraries to more recent releases by using the “update packages” option in the “Packages” menu in RGUI.
- All analyses done in R should be recorded as having been done by the specific versions of the base program and libraries that you use. This information should go into methods sections of your papers.

A few tips on Managing the R Installation

- Since much R programming is done by writing functions/code, a good text editor is useful
- the Script Editor in R is ok, but alternatives should be considered
- **I strongly recommend Notepad++ (do a google search for it)**
- **We will emphasize learning R Studio (open source IDE for R), an important and developing part of the R programmer's skillset**

R Studio

- <http://www.rstudio.com/>



Info for MAC users

- A few differences in the installation procedure from those in this doc will become apparent. None are major and you should be able to sort them out with the text instructions from the two slides above
- One important difference may be the creation of the capability to run the package called R Commander. MAC OS will need an X11 or XQuartz capability. Before installing R Commander, see the relevant section of the following instructions:
 - <http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/installation-notes.html>

Learning R

- Many tutorials are available on the web.
- See the main link on the Dudek class web pages for pointers to those tutorials. You can find it in the left-hand menu bar for extensive help on working through the initial learning curve in R
 - <http://www.albany.edu/psy/bcd/dudekclass.htm>