Installation of R with RStudio and Packages

Overview

This document describes the installation procedure of *Microsoft's Open* R, which is an enhanced version of the standard open-source version[1] of R CRAN). For this course we prefer to use *Microsoft's Open* R, because it has computational advantages over the standard open-source version. However, *Microsoft's Open* R, usually is one version number behind the standard R.

The main enhancements of Microsoft's Open R are:

1. some internal code has been substantially optimized to utilized internal functions of your computer's specific operation system and central processing units (CPU); and
2. it uses Intel's optimized *Mathematical Kernel Library* (MKL), which makes use of all computing cores on your computer's CPU. For instance, Intel's I7 chips have 4 cores.
   Therefore, MKL allows specific computations to be executed in parallel by multithreading, which can substantially speed up specific numerically extensive tasks.

In case you want to install the standard version of R follow the guidelines described in *Chapter 1: Getting R* by Lander[2].

Since both standard R and *Microsoft Open* R just implement a rudimentary user interfaces, it is highly recommended to install the integrated development shell RStudio. It provides tools that go far beyond simple code editing of R scripts (these are a collection of R commands), such as, browsing generated graphics, exploring data datasets, writing marked-up documents that incorporate the results of your R data analyses, or other programming languages.

The installation of RStudio as well as some of its features are discussed in *Chapter 2: The R Environment* by Lander. Note that we will not use RStudio's advanced functionalities such as version control or reporting tools, compiling our own packages etc., however, an example of SHINY, an interactive graphical interface to your own procedures, will be shown in class.

The functionality of either standard R or *Microsoft Open* R can be extended by installing additional packages. These package bundle items such as [i] collections of specialized functions as well as their online documentation, [ii] sample datasets, [iii] scripts with sample code on how to use the functions or datasets and [iv] potentially the marked-up documentation of full-fledged data analysis projects. Each package is installed into its specific "library" directory on your computer's hard drive.
We will use several additional packages, which are not part of R's standard installation, throughout this semester. To learn more about packages see Lander *Chapter 3: R Packages*.

---

[1] The most recent standard version of R can be downloaded from www.r-project.org . This site also provides links to the Comprehensive R Archive Network (CRAN), the R journal etc., and the source code of R and all its libraries on CRAN.

[2] The book by Jared P. Lander (2014). *R for Everyone Advanced Analytics and Graphics* can be viewed online from UTD's Library website.

Installation Steps

Important note for Mac OS X users: some packages only install properly under Mac's OS X if you first install X11 (also known as XQuartz) on your Mac computer (see https://support.apple.com/en-us/HT201341) before you proceed with the installation steps of .

1. Install the current version[3] of *Microsoft Open* from https://mran.microsoft.com/download/. Click on the installation link for your specific operating system and accept the all default installation options.
2. Install Studio from https://www.rstudio.com/products/rstudio/download/. Select ***RStudio Desktop (free license)***. Select the appropriate ***Installer*** version for your operating system, click on the installation link and select all default installation options.
   If you want faster access to Studio on your computer then copy and paste a link to Studio from the ***Start*** folder ***RStudio*** onto your desktop.

   You should now have the following icon on your desktop:

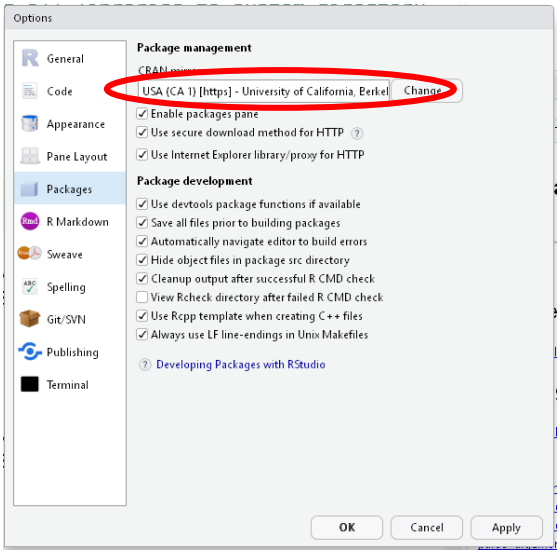
3. ***Optionally*** you may want to install the RTOOLS at this step. You find the latest stable version at https://cran.r-project.org/bin/windows/Rtools/ . Click on it and accept the default values during the installation process. The RTOOLS allow you to build your own packages and they are needed to compile a few packages during their installation from CRAN.
4. Open Studio. Unless you have older versions of on your computer *Microsoft Open* (MRO) will be your current version of .
5. Download the script **`installPackages.R`** from **ELEARNING**'s folder ***Working with R***. It automates the installation of over 1,100 packages from a CRAN server.
6. In Studio select File ▶ Open File… to open the script **`installPackages.R`** in Studio's script editor. The editor window should show the script:

```
Source on Save | Run | Source
 1 ##
 2 ## Objectives: Loads list "ipnew" of more than 1100 package and installs these packages
 3 ## Note: Depending on internet connection the installation may take _several_ hours
 4 ##
 5
 6 ## load dataframe from URL
 7 load(url("http://thinkr.spatialfiltering.com/Downloads/Packages.txt"))
 8 ipold <- installed.packages()[,1]                    # get names of existing packages
 9
10 ##
11 ## install packages not present in current installation
12 ##
13
14 ## Send all packages to system directory (currently commented out)
15 #.libPaths("C:/Program Files/Microsoft/R Open/R-3.5.3/library")
16
17 for (i in setdiff(ipnew,ipold)){
18   install.packages(i, dependencies=TRUE)
19 }
20
21 ## Install private course package
22 install.packages("http://thinkr.spatialfiltering.com/Downloads/DallasTracts_0.1.2.tar.gz",
23                  repos=NULL)
24
25 cat("# of installed packages: ", length(installed.packages()[,1]))
26
```

[3] As of August 2019 the most recent version of *Microsoft Open* is 3.5.3 call "Great Truth".

7. Click in the right upper corner on **Source** (or Ctrl-Shift-Enter) to execute this script.
   Alert: Depending on the speed of your internet connection and the time of day, it may take an hour or longer to execute this script and install all the requested packages.
   Once the command prompt "**>**" in RStudio's Console window appears, the installation of all packages has been completed.
8. The Console should show that slightly more than 1,100 packages were installed on your computer.

Note: there may be a few warning messages that specific packages could not be installed. You can install these packages later, if we need them in this course. However, should you receive a lot of warning and error messages during the installation of the packages, then most likely the internet connection to the default CRAN server is not stable. You can select an alternative CRAN server in the RStudio from the **TOOLS ▶ GLOBAL OPTION…** menu by selecting the **PACKAGES** tab and picking another server from CRAN's mirror list:



Once another server has been selected please repeat step 6 in RStudio. Hopefully, the installation of the packages proceeds now smoothly.

User Packages: We will also make use packages that were written by GISC graduate students in collaboration with the course instructor. These packages have not been posted on CRAN but they can be installed at the R command prompt. For instance, the package **`DallasTracts_0.2.2.tar.gz`** can be installed with:

```
> install.packages("http://thinkr.spatialfiltering.com/Downloads/DallasTracts_0.2.2.tar.gz",
                   repos=NULL)
```

The option **`repos=NULL`** indicates that the package is not on the CRAN server but can be downloaded directly from **`http://thinkr.spatialfiltering.com/Downloads/`** . Note that the download path with the package name need to be enclosed in quotation marks because a string variable is expected as input.

In order to make a package accessible to your current R session, you need to call the function **`library(`*`LibName`*`)`** . For example use **`> library(DallasTracts)`** at the command prompt.