Introduction to Data 413/613

Hamid Semiyari

Acknowlegemt

I would like to express my sincere gratitude to Drs. James Dickense, David Gerard, and Richard Ressler for their insightful feedback, unwavering encouragement, and dedication to excellence have significantly contributed to the success of this course. I am truly fortunate to have had the privilege of working with such esteemed professionals who have played a pivotal role in shaping and enhancing the outcome of this course.

Data 413/613 Topics

- The State of the art computational tools for data science. This course builds on the R/tidyverse programming skills developed in the collection, organization, analysis, interpretation, and presentation of data. Topics include version control, web scraping, date manipulation, vectorized operations, web application development with R Shiny, R package development, and introductions to SQL and python.

Update R, RStudio, and your R Packages

Update Your Version of R.

- Check your version of R when you start RStudio by looking at the Console pane or entering `R.Version`

```
R.Version()
```

```
$platform
[1] "x86_64-w64-mingw32"

$arch
[1] "x86_64"

$os
[1] "mingw32"

$crt
[1] "ucrt"

$system
[1] "x86_64, mingw32"

$status
[1] ""

$major
[1] "4"

$minor
[1] "3.2"

$year
[1] "2023"

$month
[1] "10"

$day
[1] "31"

$`svn rev`
[1] "85441"

$language
[1] "R"

$version.string
[1] "R version 4.3.2 (2023-10-31 ucrt)"

$nickname
```

```
[1] "Eye Holes"
```

- It should be at least R version "Eye Holes" released on 2023/10/31.

```
R.Version()$nickname
```

```
[1] "Eye Holes"
```

The released date was:

```
R.Version()$version.string
```

```
[1] "R version 4.3.2 (2023-10-31 ucrt)"
```

- If not, go to The Comprehensive R Archive Network, CRAN and get the latest version for your OS.
  - Windows: To install R on Windows, click the "Download R for Windows" link. Then click the "base" link. Next, click the first link at the top of the new page. This link should say something like "Download R-4.3.2 for Windows." The link downloads an installer program, which installs the most up-to-date version of R for Windows. Run this program and step through the installation wizard that appears. The wizard will install R into your program files folders and place a shortcut in your Start menu. Note that you'll need to have all of the appropriate administration privileges to install new software on your machine.
  - Mac: To install R on a Mac, click
    - For Apple silicon (M1/M2) Macs: R-4.3.2-arm64.pkg
    - For older Intel Macs: R-4.3.2-x86_64.pkg An installer will download to guide you through the installation process, which is very easy. The installer lets you customize your installation, but the defaults will be suitable for most users.
- If you had to update to a major new release
- **Question: What is the difference between "Arm64", "x86", and "x64"?**
  - Arm64 is used in SmartPhone, Apple Computers and some Servers.
  - X86 released in 1978 and its creator was "Intel". 32-Bit Architecture and a 32 bit system can handle a maximum of 2^{32} of RAM, which comes out to be 4GB of RAM.
  - X64 released in 2000 and its creator was "AMED". 64-Bit Architecture and 64 bit system can handle a maximum of 2^{64} of RAM, which comes out to be 16 Exabytes (EB) of RAM which is two levels up from Terabyte (TB).

Therefore x64 allows the CPU to store more data and access it faster.

Update Your Version of RStudio

- Check your version of RStudio using `Help - About RStudio`.
- Your version should be at least 2023.09.1 Build 494 or higher.
- If not, go to RStudio Desktop to download the latest version
  - Install and restart RStudio
- **RStudio does not pick upthe updated version of R automatically**, you need to do it manually.
  - In RStudio, Go to `Tools` then `Global Options`, then click `Change` under `R version` and from the dropdown menu select the one version you need.

Update Your R Packages

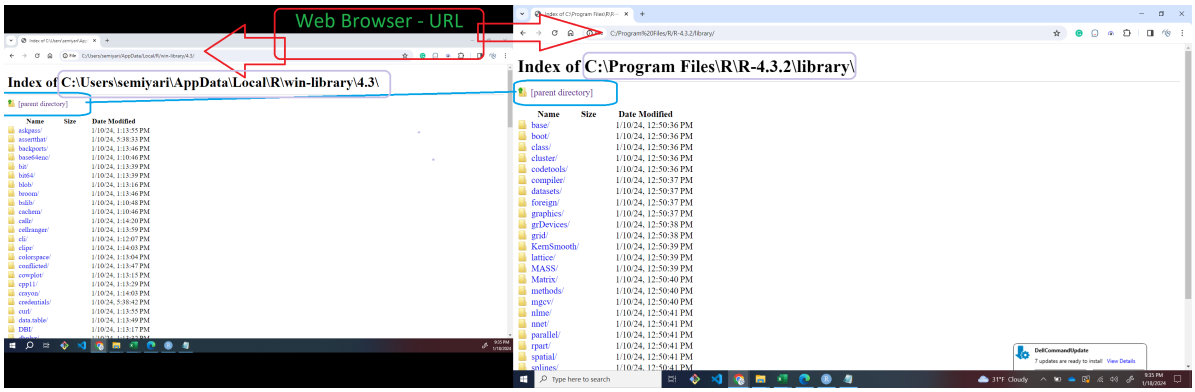
- Select the `Packages` pane in RStudio and click on `Update`.
- You can `Select All` or choose individual packages to update. Be sure to check any tidyverse packages.
- If asked "Do you want to install from sources the packages which need compilation?" enter `No`.
- Restart R.
- **Question: How to Restart R?**
  - In RStudio, Go to `Session` then click `Restart R`.
    - Short cut: Window use "CTRL + SHIFT + F10" and for Mac use "Command + SHIFT + F10"
- **Question: Where R packages are installed?**
  - R packages are installed in a directory called "library". The command `.libpaths()` can be used to get to the library.
- **Note:** There are two default locations for libraries
- Under "C:/Program Files/"
- Under "C:/Users/" for specific users
- **Note:** When we install a new version of R the default is to keep the old version too. If you don't need the old version then you should uninstall it. *Go to* ***`Setting`*** *on your local machine to uninstall it.* You also need to manually delete corresponding library folder of old version of R.

- **Question: How manually delete corresponding library folder of old version of R?**
  - The easiest way is to type `.libpaths()` in your console.

```
.libPaths()
```

```
[1] "C:/Users/semiyari/AppData/Local/R/win-library/4.3"
[2] "C:/Program Files/R/R-4.3.2/library"
```

The paste each default location into URL click on the folder "[parent directory]" and delete the old version of R folder.



If You Are Using a Mac ···

Update XQuartz

- Since we use RStudio (or R) from th terminal, we must have a copy of XQuartz installed. There are many applications that are not designed for Mac, and `XQuartz` allows us to use them.
- **Question: Do i need to install "xquartz" on Mac?**
  - On Mac OS X if you use RStudio or R from the terminal, you must have a copy of XQuartz, the X11 window manager, installed. This is no longer a default install since Mac OS X 10.8. You need to install XQuartz, before installing the nat package.
  - **Question: What is nat package](https://cran.r-project.org/web/packages/nat/vignettes**

    * nat is an R package and therefore runs on Mac/Linux/Windows. The only pre-requisite for most functionality is a recent version of R (>=3.1.0 recommended). NeuroAnatomy Toolbox (nat) enables analysis and visualisation of 3D biological image data, especially traced neurons.

Installing Homebrew

- It is recommended to install `Homebrew`, a general package manager that allows you to install a wide variety of developer-type software not generally available in the Mac App Store.
  - To install `Homebrew`, see Homebrew Installation and install in the default location.
  - This could take a few minutes.
- **Question: Do I need to install "Homebrew" on Mac?**
  - Homebrew is a package manager for macOS (and Linux, too). For more info click here