



**Northeastern University**  
College of Engineering



# **IE 6600 Computation and Visualization for Analytics**

*Basics of R*

**Zhenyuan Lu**

### R Programming:

R for Data Science (open license)

Cookbook for R (open license)

Text Mining with R (open license)

R for Everyone

(NEU library provided free online access)

### Documentation:

R Markdown (open license)

### R Visualization:

ggplot2 for static viz

Plotly for interactive viz

Shiny for web app and interactive viz

### Additional materials:

Advanced R (open license)

R Packages (open license)

R Cheatsheets List (very useful)

# 1. Introduction to R *Why*

People who using R:

Karl Pearson



**Statisticians** with  
advanced machine  
learning training



**Programmers** familiar  
with other languages

**Who** are not necessarily  
trained in advanced data  
analysis but are tired of  
using Excel



# 1. Introduction to R *Wide Variety Of Fields Can Be Applied*

Banks  
Ecology  
Genetics  
Hospitals  
Insurance  
Tech startups  
Food startups  
Pharmaceuticals  
Aid organizations  
Online advertising  
Political campaigns  
Real estate developers  
International development



# 1. Introduction to R *Other Visualization Tools*

R

Tableau

Python

Excel

Adobe Creative Suite

Powerpoint

D3.js

ArcGIS

Stata

RAW

Plot.ly

Echarts

3ds Max

Maya

Processing

HTML/CSS

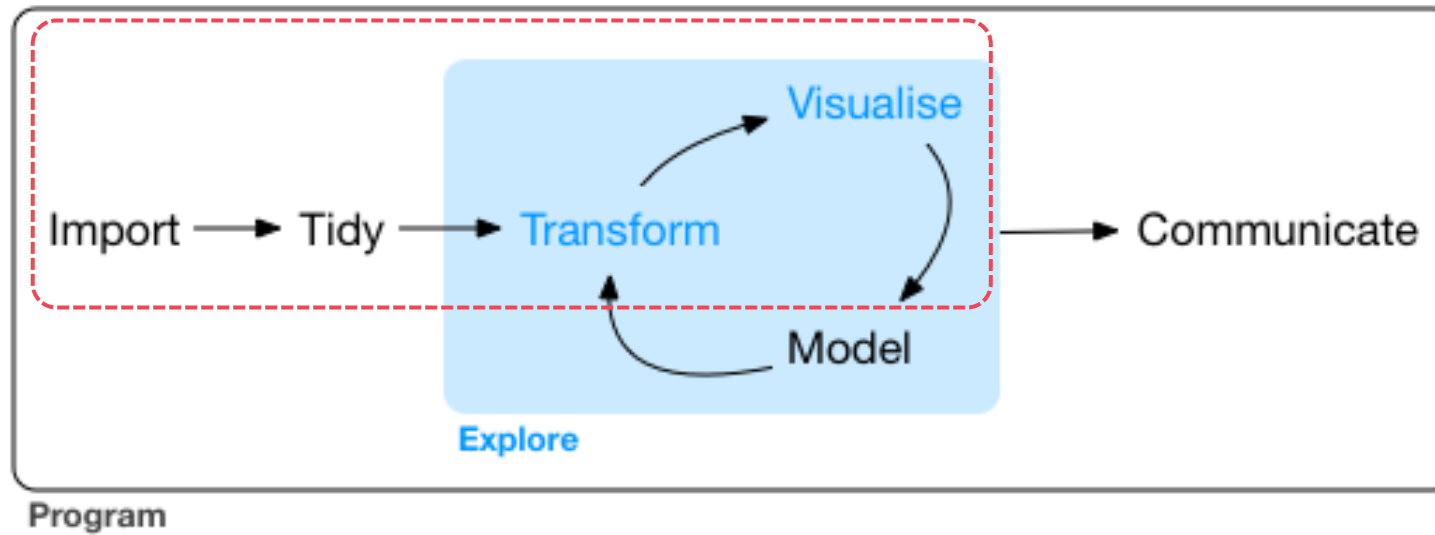
Coolors

MagicaVoxel

JS

Illustration

# 1. Introduction to R



Wickham, Hadley, and Garrett Grolemund. *R For Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, 2017.

# 1. Introduction to R *Installation of R*

After you click the following link:

<https://cran.r-project.org/>

You will see as the following figure.

Click the proper link and start the installation



## 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.



# 1. Introduction to R *Installation of R*

For Mac User:

Make sure you are downloading the correct version, otherwise it will cause errors

## Latest release:

[R-4.1.2.pkg](#) (notarized and signed)  
SHA1-hash: 61d3909bc070f7fb86c5a2bd67209fda9408faaa  
(ca. 87MB)

**R 4.1.2** binary for macOS 10.13 (**High Sierra**) and higher, **Intel 64-bit** build, signed and notarized package. Contains R 4.1.2 framework, R.app GUI 1.77 in 64-bit for Intel Macs, Tcl/Tk 8.6.6 X11 libraries and Texinfo 6.7. The latter two components are optional and can be omitted when choosing "custom install", they are 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 macOS to a new major version.

This release supports Intel Macs, but it is also known to work using Rosetta2 on M1-based Macs. For native Apple silicon arm64 binary see below.

**Important:** this release uses Xcode 12.4 and GNU Fortran 8.2. If you wish to compile R packages from sources, you may need to download GNU Fortran 8.2 - see the [tools](#) directory.

[R-4.1.2-arm64.pkg](#) (notarized and signed)  
SHA1-hash: 69e8845ffa134c822d4bdcf458220e841a9eeaa5  
(ca. 87MB)

**R 4.1.2** binary for macOS 11 (**Big Sur**) and higher, **Apple silicon arm64** build, signed and notarized package. Contains R 4.1.2 framework, R.app GUI 1.77 for Apple silicon Macs (M1 and higher), Tcl/Tk 8.6.11 X11 libraries and Texinfo 6.7.  
**Important: this version does NOT work on older Intel-based Macs.**

Note: the use of X11 (including `tcltk`) requires [XQuartz](#). Always re-install XQuartz when upgrading your macOS to a new major version.

This release uses Xcode 12.4 and experimental GNU Fortran 11 arm64 fork. If you wish to compile R packages from sources, you may need to download GNU Fortran for arm64 from <https://mac.R-project.org/libs-arm64>. Any external libraries and tools are expected to live in `/opt/R/arm64` to not conflict with Intel-based software and this build will not use `/usr/local` to avoid such conflicts.



# 1. Introduction to R *Installation of R Studio*



RStudio is a set of integrated tools designed to help you be more productive with R.

After you click the following link:

<https://www.rstudio.com/products/rstudio/download/>

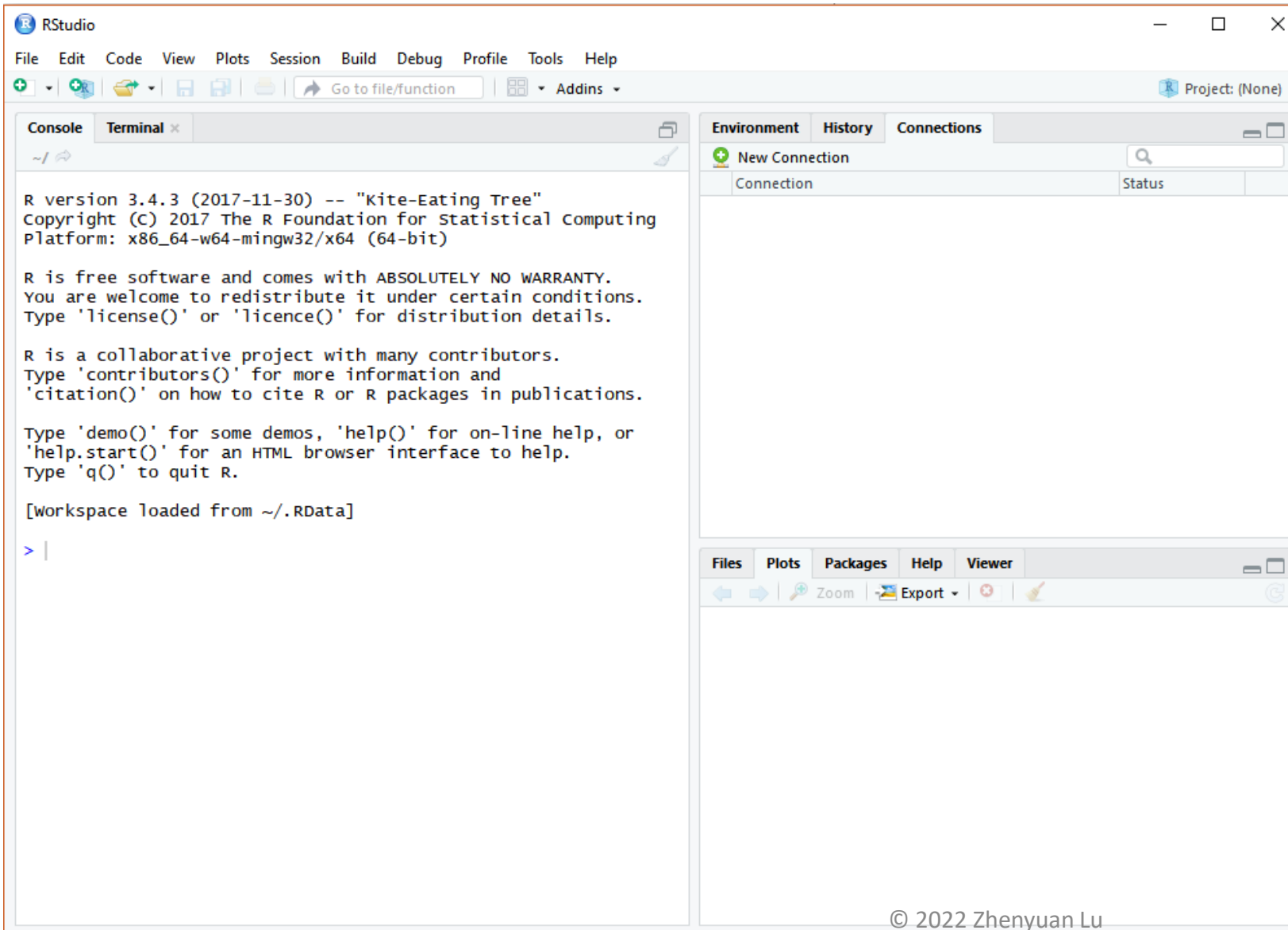
You will see as the following figure.

Choose the first version, which is totally enough for our class application and practice



|  |                                       |                                       |  |   |
|--|---------------------------------------|---------------------------------------|--|---|
| RStudio Desktop<br>Open Source License | RStudio Desktop<br>Commercial License | RStudio Server<br>Open Source License | RStudio Server Pro<br>Commercial License | RStudio Server Pro +<br>RStudio Connect<br>Commercial License |
| FREE                                   | \$995 per year                        | FREE                                  | \$9,995 per year                         | \$29,995 per<br>year  |
| <b>DOWNLOAD</b>                        | <b>BUY</b>                            | <b>DOWNLOAD</b>                       | <b>DOWNLOAD</b>                          | <b>TALK</b>   |
| <a href="#">Learn More</a>             | <a href="#">Learn More</a>            | <a href="#">Learn More</a>            | <a href="#">Learn More</a>               | <a href="#">Learn More</a>                                    |

# 1. Introduction to R *R Studio*



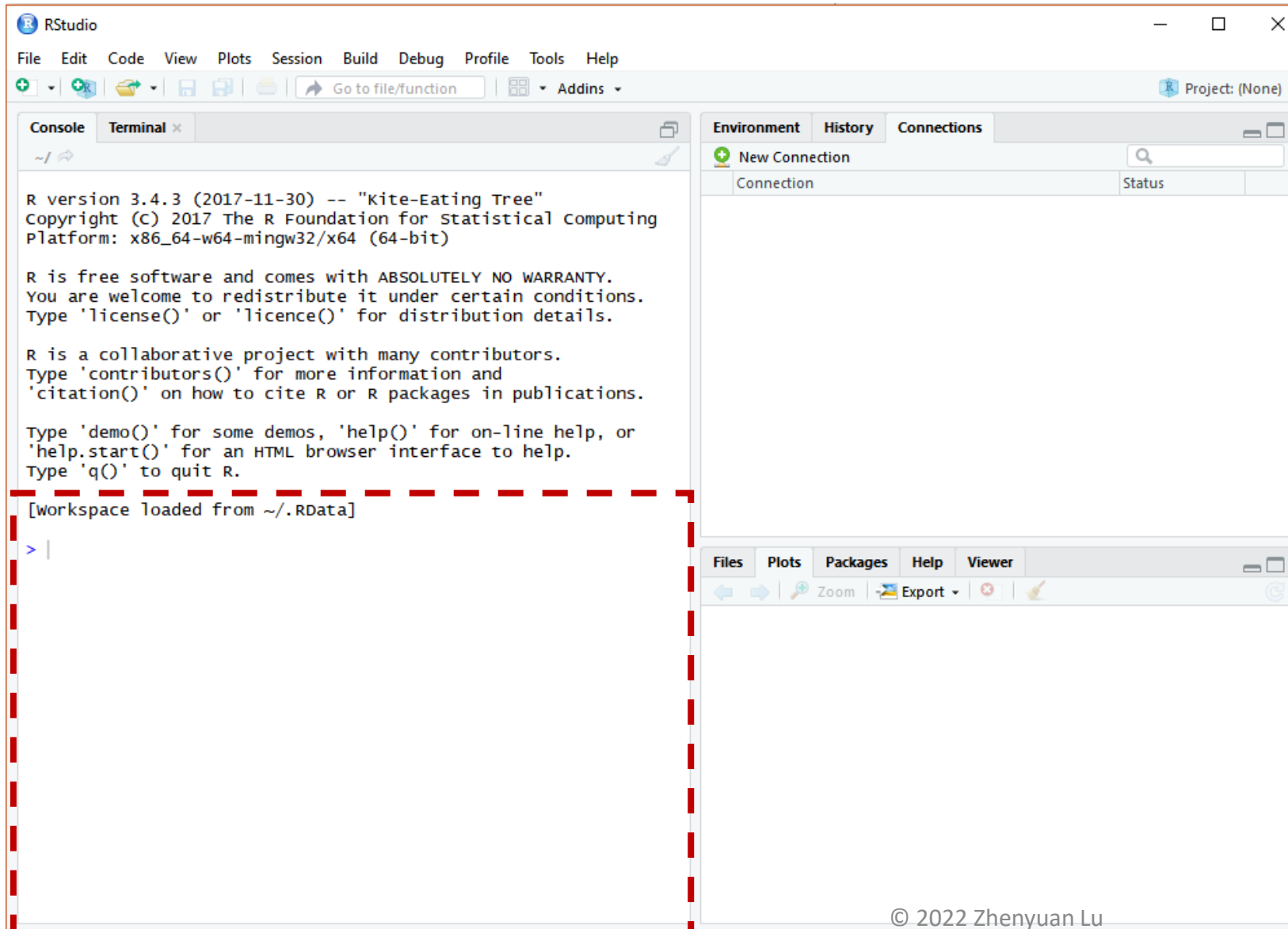
If you have *successfully* installed the R studio, you will see the same R studio console as the left figure showing after launching on the RStudio (OS: Windows 10)

In this class, all the practice, homework, and project will be processing on RStudio.

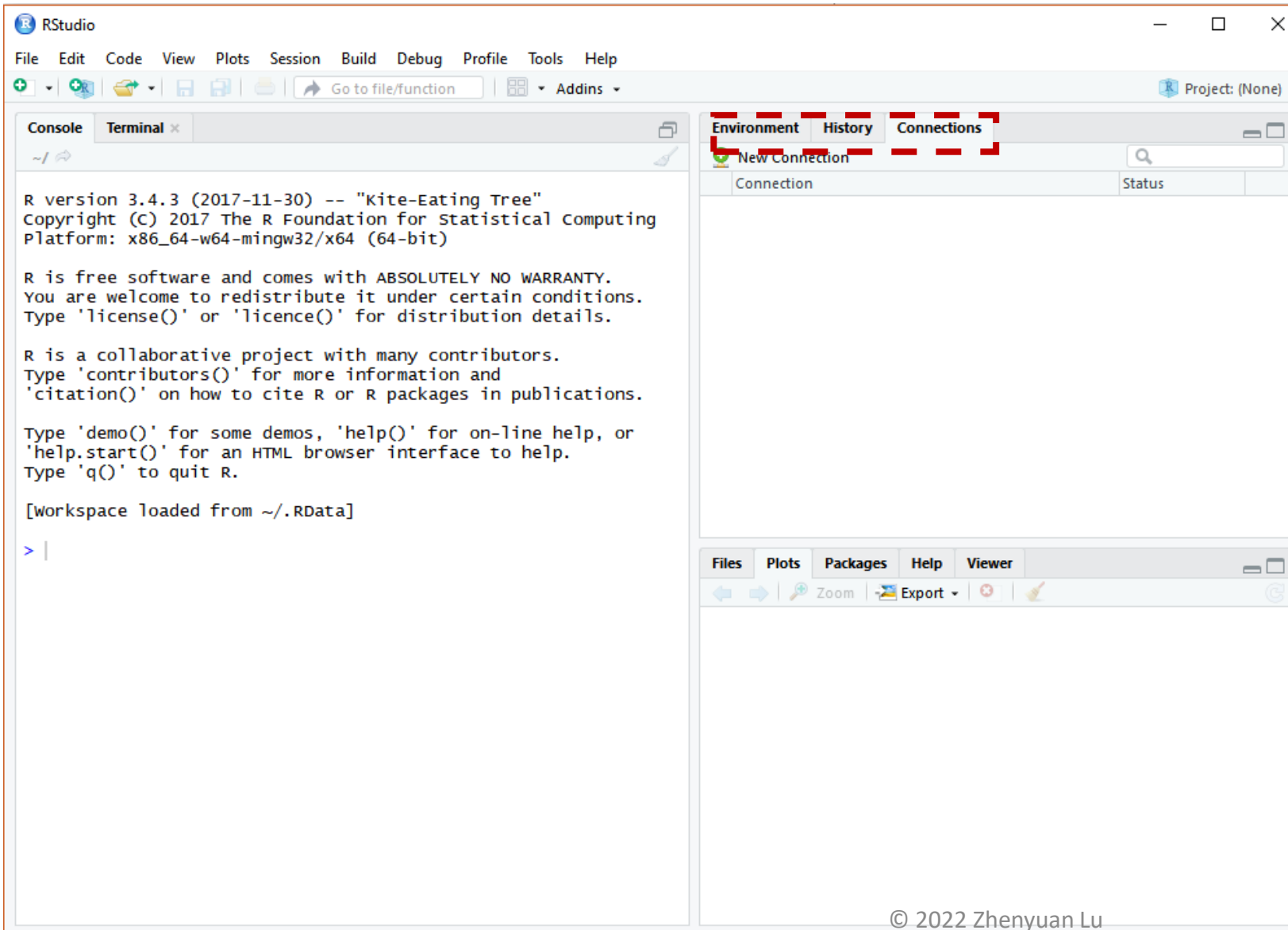
# 1. Introduction to R *R Studio*



The rectangle area is for scripts

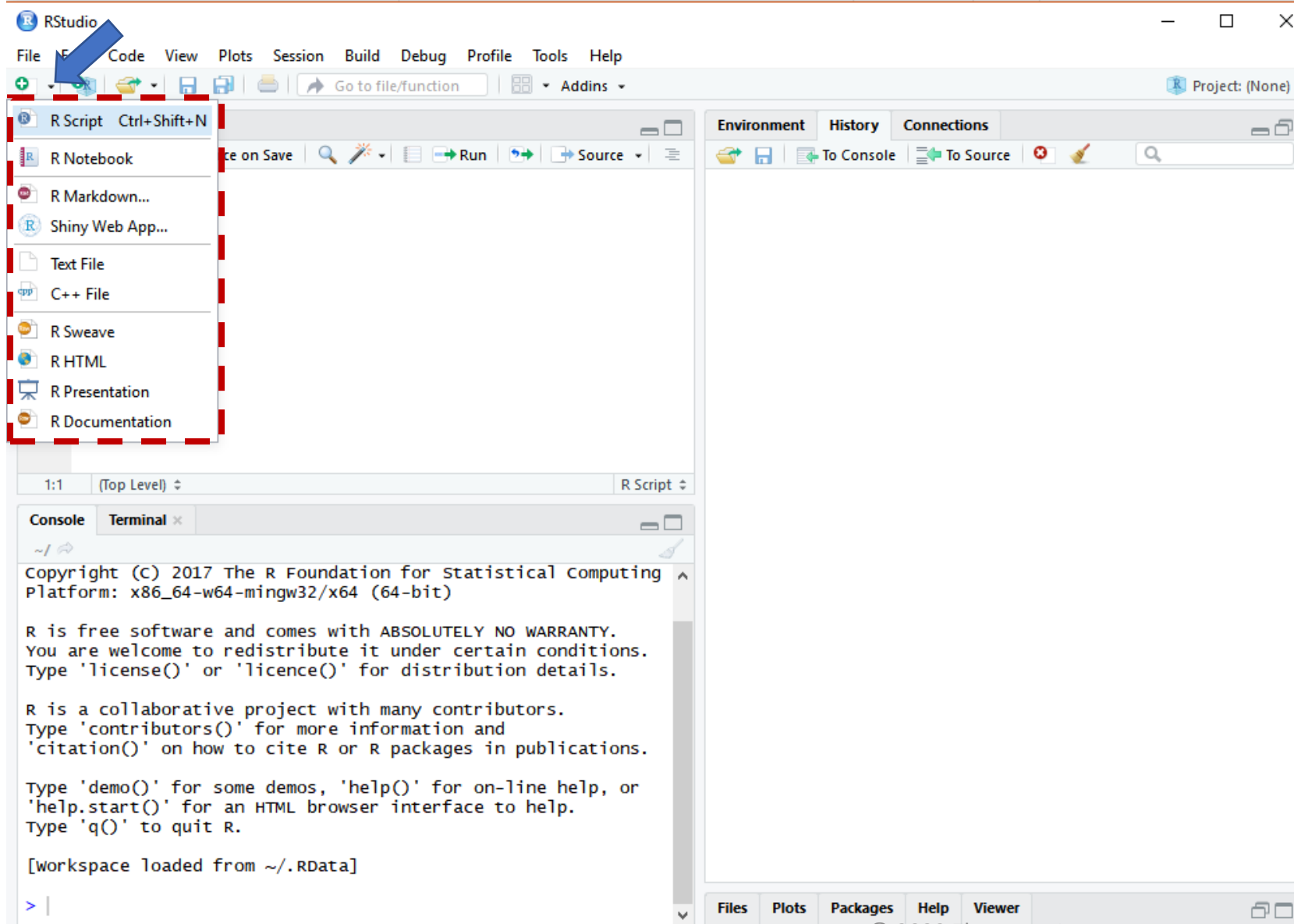


# 1. Introduction to R *R Studio, Environment, History, and Connections*



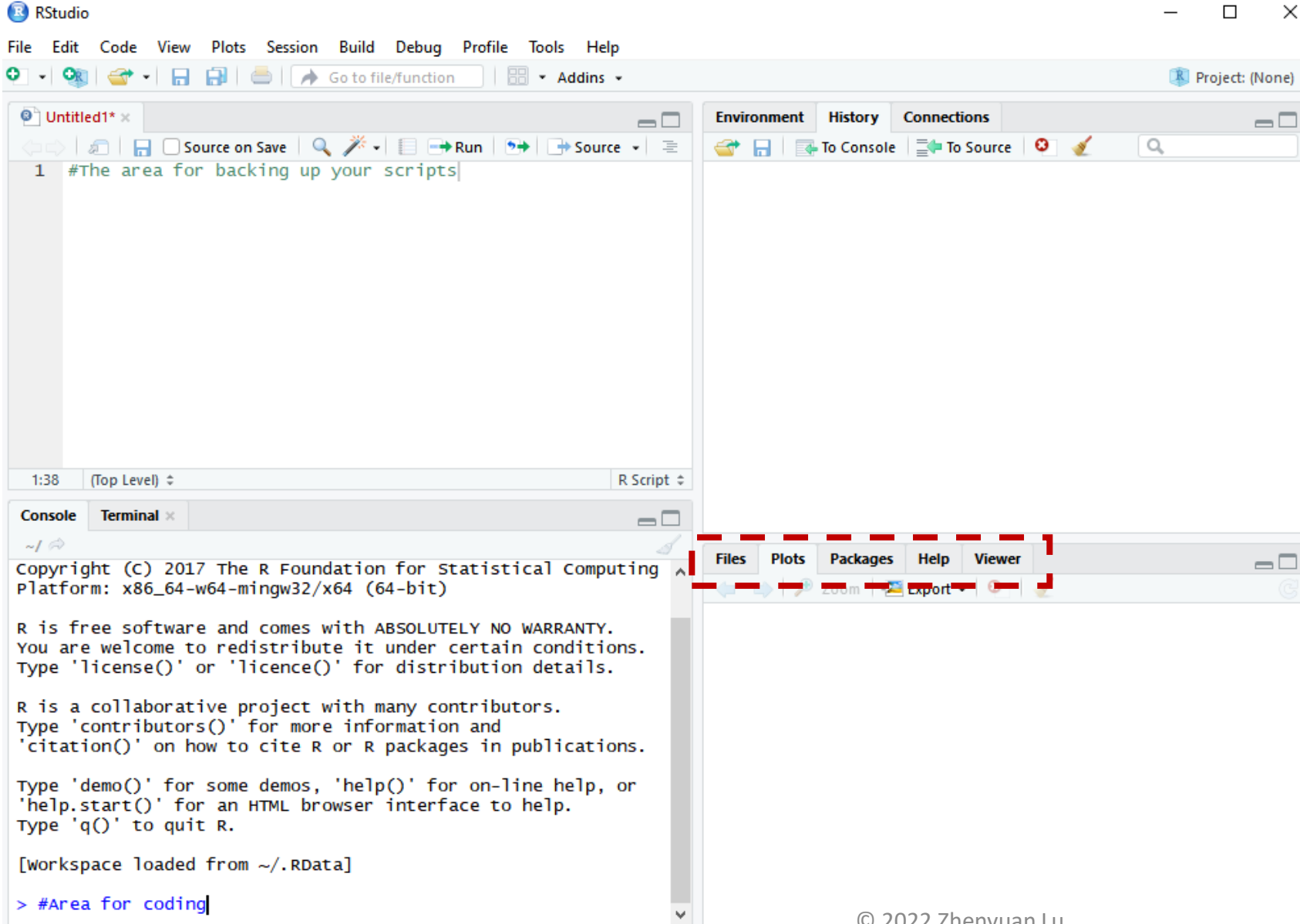
The three tabs on the top right are environment which contains all the data frame, variables, etc., history, and connections.

# 1. Introduction to R *R Studio, Saving Scripts*



For saving your codes, you may want to click, the arrow closed to the green +, then click R script (key ctrl+shift+N). This is the most convenient way for backing up your coding projects.

# 1. Introduction to R *R Studio, others*



In this area, you may check the files, the plots, packages you are working with, the help of R, and viewer.



**Northeastern University**  
College of Engineering



# **IE 6600 Computation and Visualization for Analytics**

*Basics of R*

**Zhenyuan Lu**



Python for Data Science:

Python Data Science Handbook (open license)

Python for Data Analysis (2<sup>nd</sup> edition on  
Amazon; 1<sup>st</sup> edition on NEU library)

# 1. Introduction to R *R Packages*

A **package** is essentially a library of **prewritten code** designed to accomplish some task or a collection of tasks. It is the most powerful tool in the R.

In this class we will be more focused on the package of **ggplot2**, which is most efficient package for visualization. Though there are still some very robust and well-maintained packages out there, such as **dplyr**, **readr**, **caret**, **tibbles**, or **knitr**.

Package lists:

[https://cran.r-project.org/web/packages/available\\_packages\\_by\\_name.html](https://cran.r-project.org/web/packages/available_packages_by_name.html)

# 1. Introduction to R *Installation of R Packages*

The fastest way to do is typing the following scripts in the Rstudio.

```
>install.packages("ggplot2")
```

After running...

```
> install.packages("ggplot2")
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.4/ggplot2_3.0.0.zip'
Content type 'application/zip' length 3145849 bytes (3.0 MB)
downloaded 3.0 MB

package 'ggplot2' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
```

**NOTE:** the package name is case sensitive, which means Ggplot2 is different from ggplot2

Package lists:

[https://cran.r-project.org/web/packages/available\\_packages\\_by\\_name.html](https://cran.r-project.org/web/packages/available_packages_by_name.html)

# 1. Introduction to R *Installation of R Packages*

The fastest way to do is typing the following scripts in the Rstudio.

```
>install.packages("ggplot2")
```



There are 3 ways to get help document.

1. Press **F1** (for laptop may click **Fn+F1**, or "Other function key"+F1) and then the help tab will pop up on the right bottom of the RStudio.

2. Type `help(install.packages)`

3. Type `?install.packages`



# 1. Introduction to R *Un/Loading R Packages*

## Loading R packages (Two ways)

```
>library(ggplot2)
```

```
> library(ggplot2)
warning message:
package 'ggplot2' was built under R version 3.4.4
> |
```

```
>require(ggplot2)
```

```
> require(ggplot2)
Loading required package: ggplot2
warning message:
package 'ggplot2' was built under R version 3.4.4
> |
```

---

## Unloading R packages

```
>detach("package:ggplot2")
```

NOTE0: require is designed for use inside functions returning FALSE.

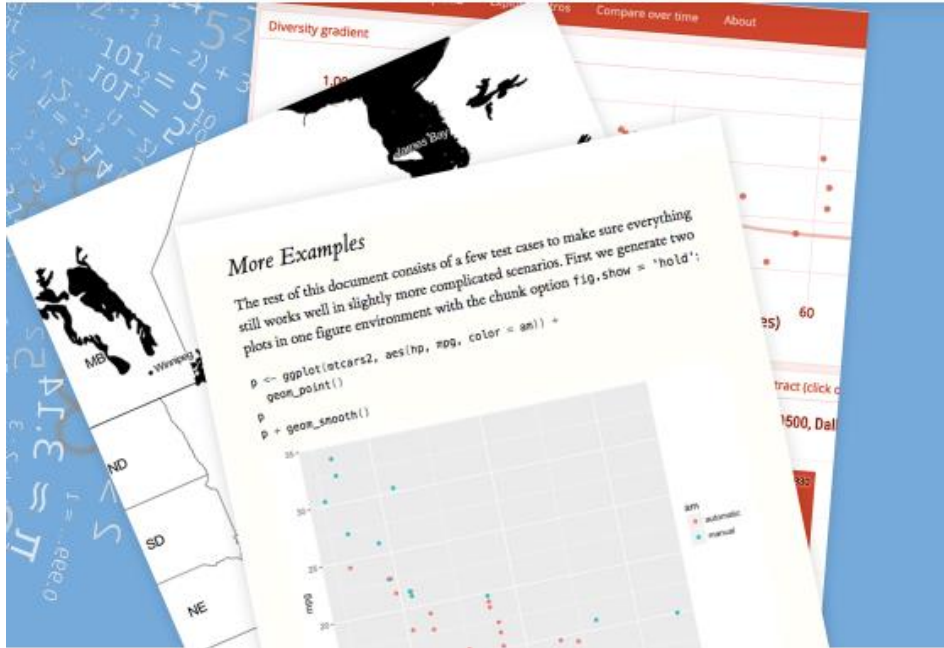
NOTE1: We usually ignore the warning message...

NOTE2: You may need to load the packages again if you relaunch RStudio

NOTE3: It will automatically detach all the global packages you loaded after closing Rstudio, then follow NOTE2

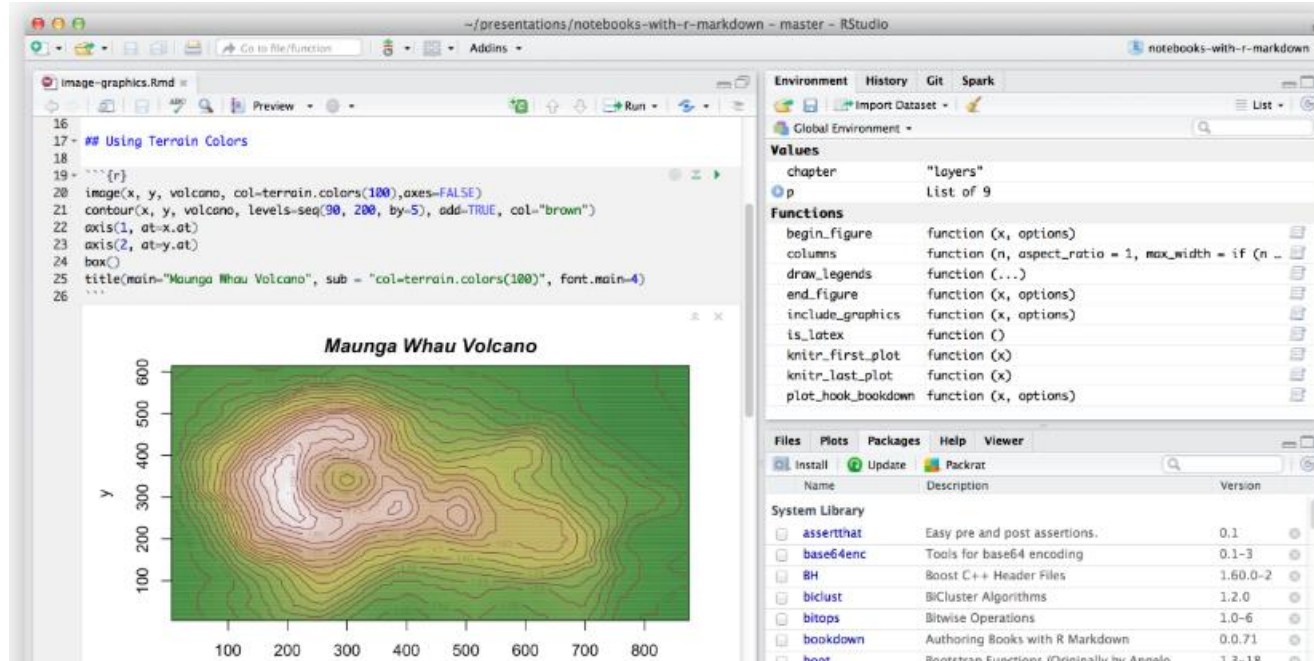


# 1. Introduction to R *R* Markdown



Documentation for analyze, share, and reproduce

# 1. Introduction to R *R Markdown*



R Markdown documents are fully reproducible. Use a productive [notebook interface](#) to weave together narrative text and code to produce elegantly formatted output. Use [multiple languages](#) including R, Python, and SQL.



# 1. Introduction to R *R* Markdown

R Markdown supports dozens of static and dynamic output formats including [HTML](#), [PDF](#), [MS Word](#), [Beamer](#), [HTML5 slides](#), [Tufte-style handouts](#), [books](#), [dashboards](#), [shiny applications](#), [scientific articles](#), [websites](#), and more.



# 1. Introduction to R *R Markdown*

```
# Install from CRAN
install.packages('rmarkdown')

# Or if you want to test the development version,
# install from GitHub
if (!requireNamespace("devtools"))
  install.packages('devtools')
devtools::install_github('rstudio/rmarkdown')
```

If you have any R markdown or other common issues please check file on Piazza: Q&A IE6600 Documentation

# 1. Introduction to R *R Markdown*

If you want to generate PDF output, you will need to install LaTeX. For R Markdown users who have not installed LaTeX before, we recommend that you install TinyTeX

```
install.packages("tinytex")  
tinytex::install_tinytex() # install TinyTeX
```

If you have any R markdown or other common issues please check file on Piazza: Q&A IE6600 Documentation