香 港 大 學
THE UNIVERSITY OF HONG KONG

FITE7410 Financial Fraud Analytics
R Language Tutorial

Zejia Liu

Agenda

- What is R?
- Setting up the R environment
- R Studio
- R packages

R Language

- R is a language and environment for statistical computing and graphics.
- It provides a wide variety of statistical and graphical techniques and is highly extensible.
- It is widely used among statisticians and data miners.

Why Use R?

- Free and open source
- Analytical support
- Extensive functionality

Steps

- Installing R
- Installing RStudio
- (Optional) Installing Rtools (Windows) / R build tools (Mac)

Installing R

Download website: https://cloud.r-project.org/

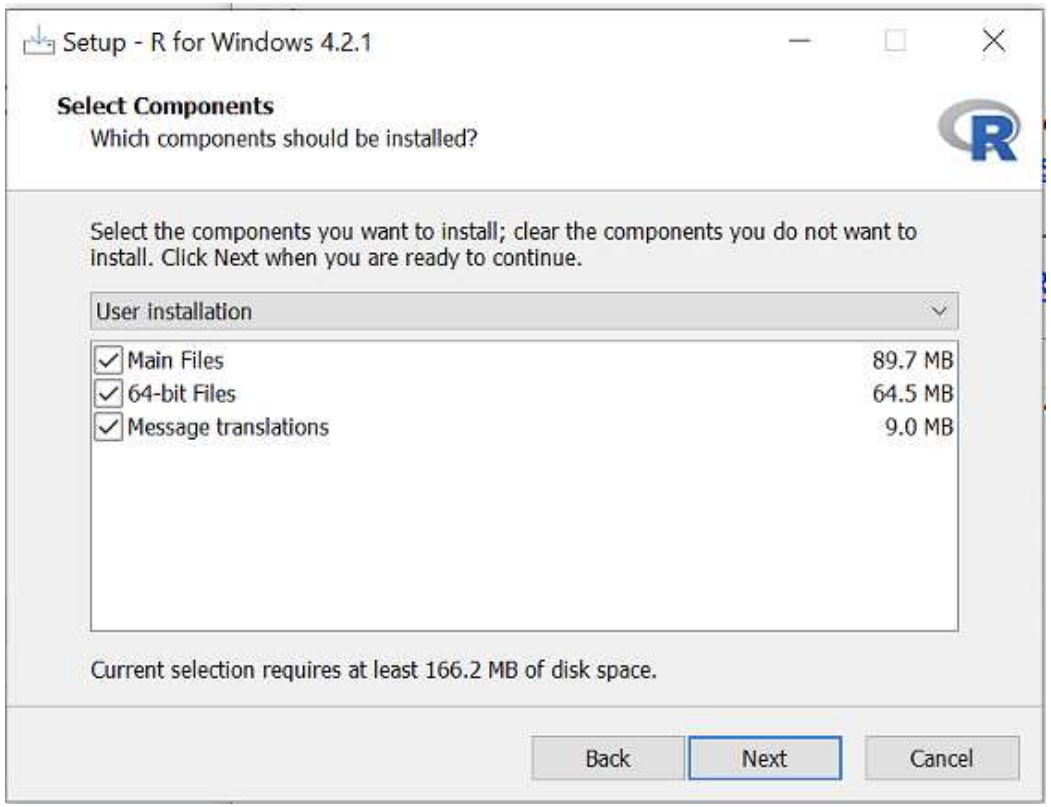
Download and Install R

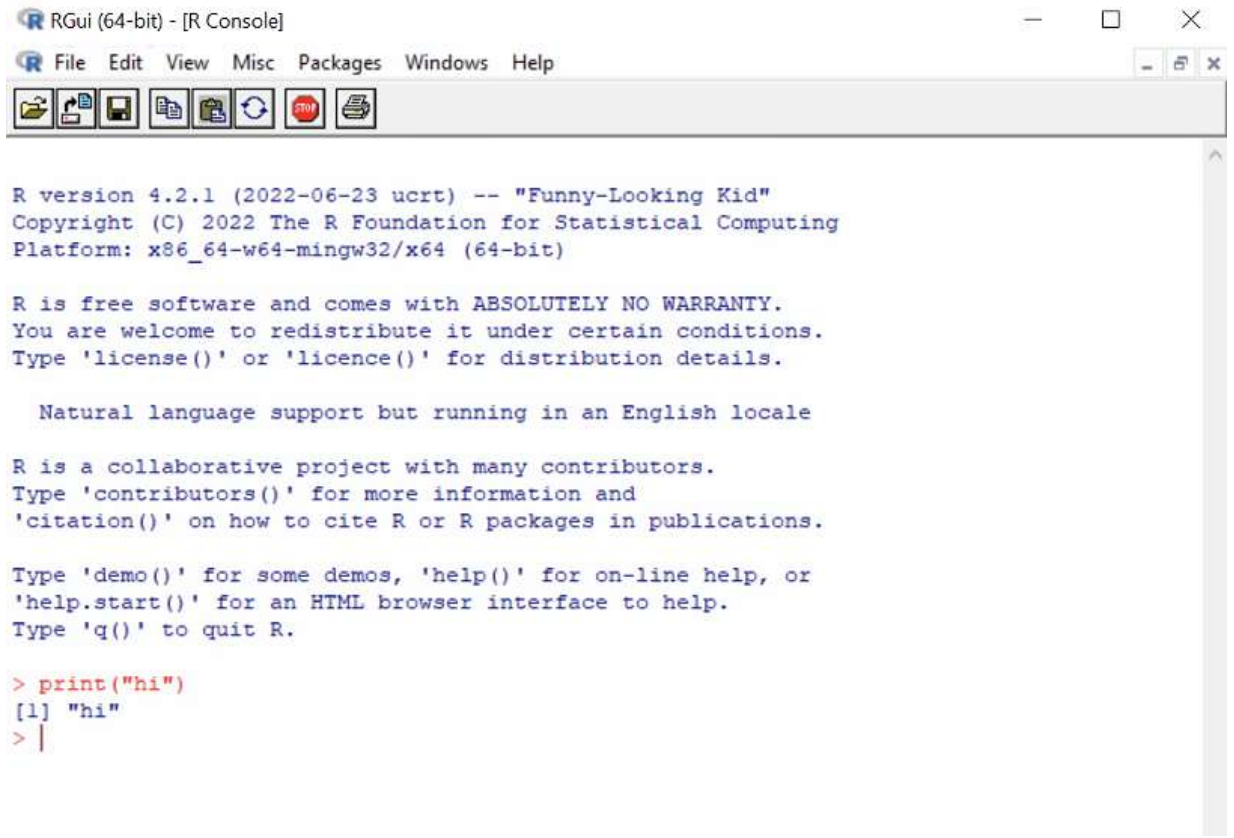
Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- Download R for Linux (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

Setting up the R environment

RGUI





Note: Please remind setting up your environment variables.

RStudio

The official IDE: https://posit.co/download/rstudio-desktop/

Note: Install R before installing RStudio.

- https://docs.posit.co/previous-versions/rstudio/

(Optional) Installing Rtools (Windows) / R build tools (Mac)

Tools to build R and R packages.

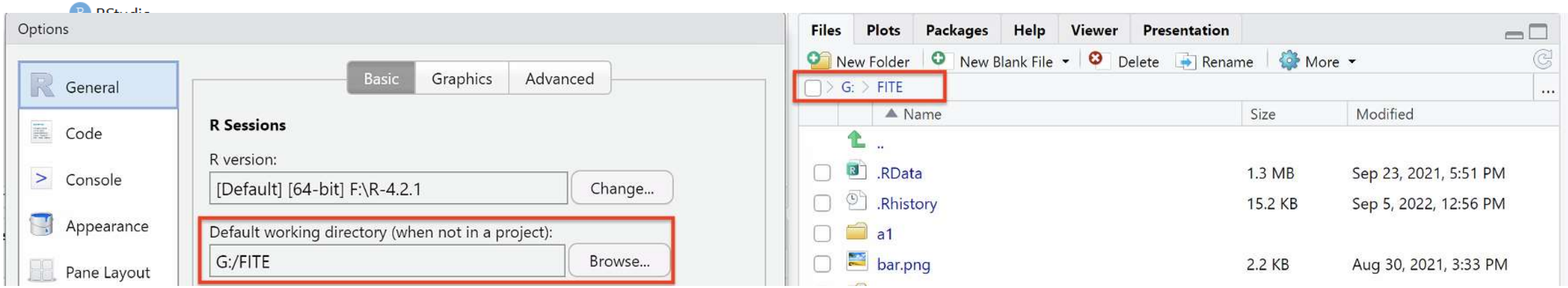
- Rtools (Windows): https://cran.rstudio.com/bin/windows/Rtools/
- R build tools (Mac): Xcode and the associated command line tools.

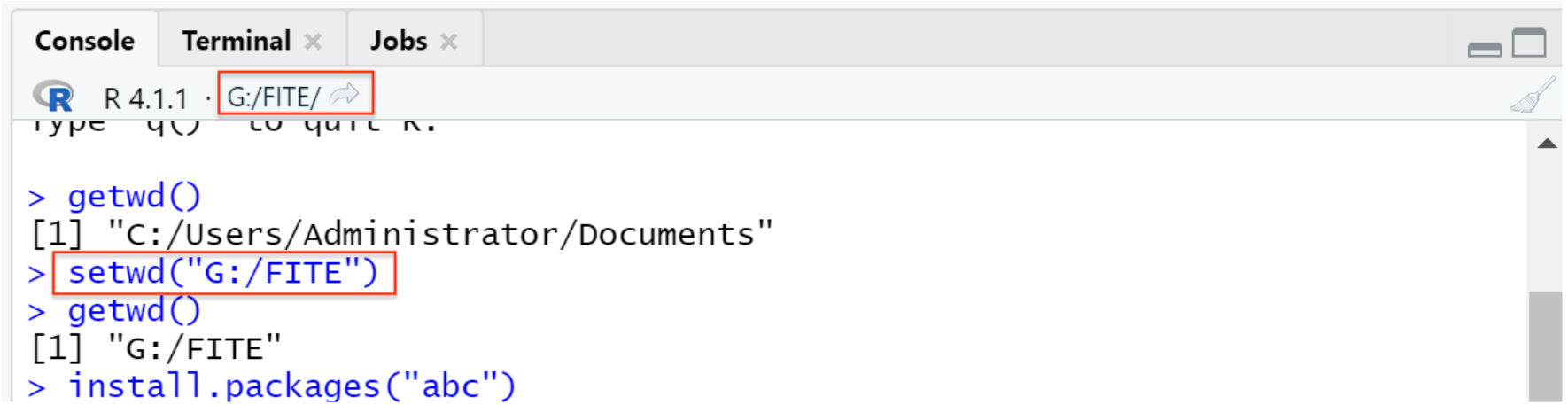
- https://ohdsi.github.io/Hades/rSetup.html#Installing_R_build_tools

RStudio

Working Directory

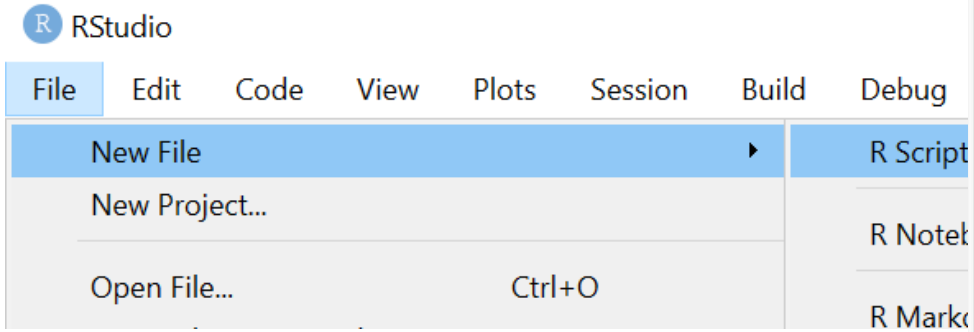
- Tools --> Global Options --> General --> Default working directory --> restart RStudio

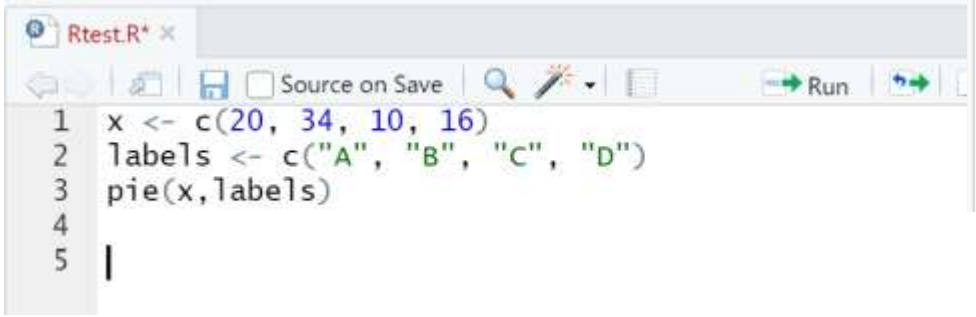


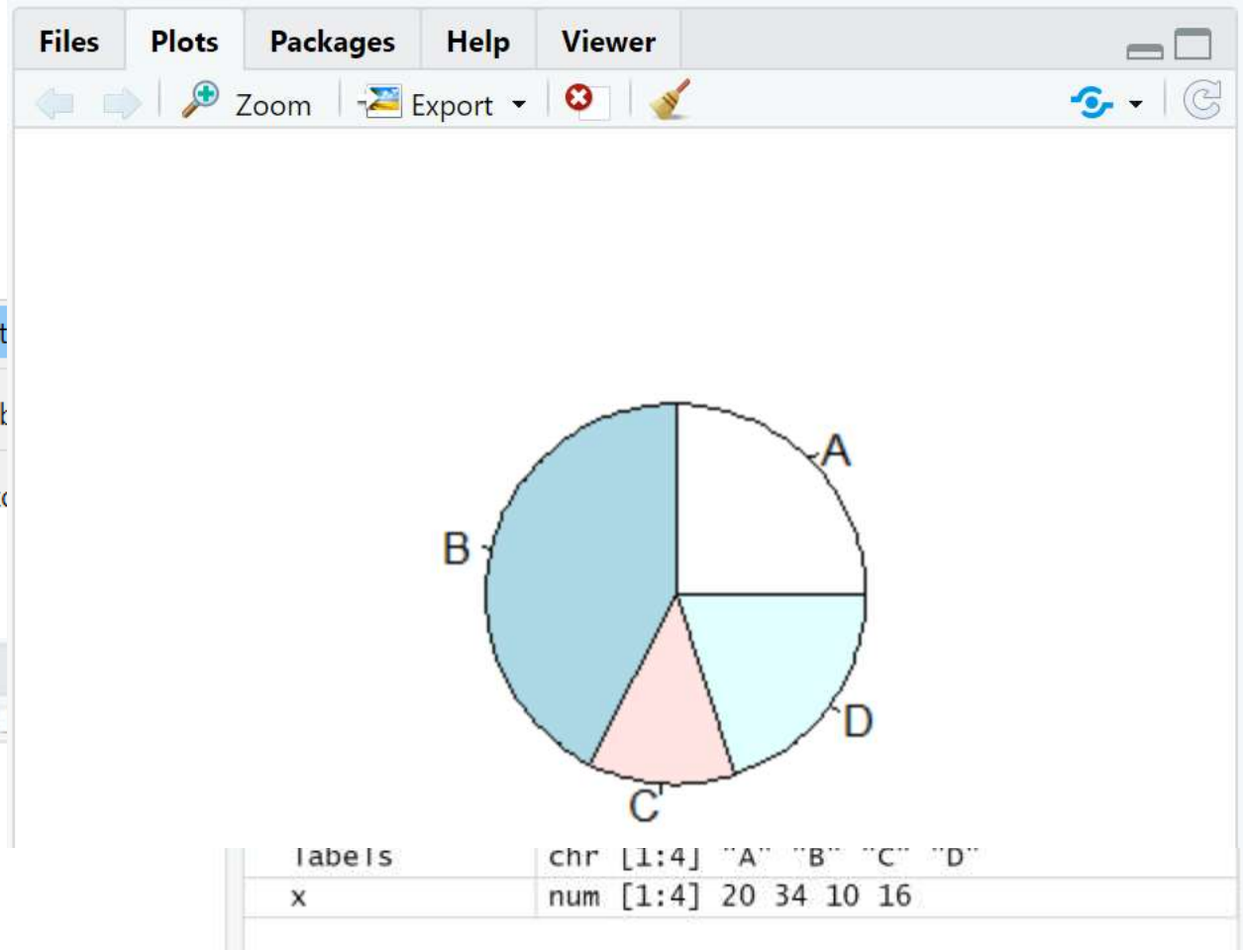


RStudio

R Script



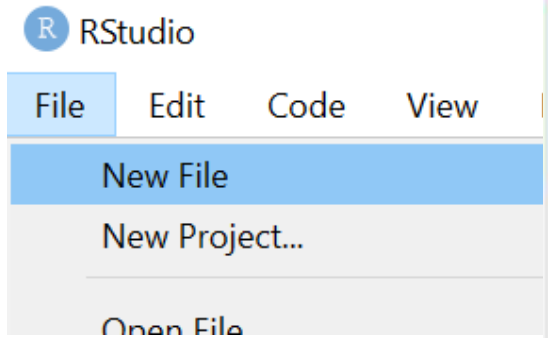


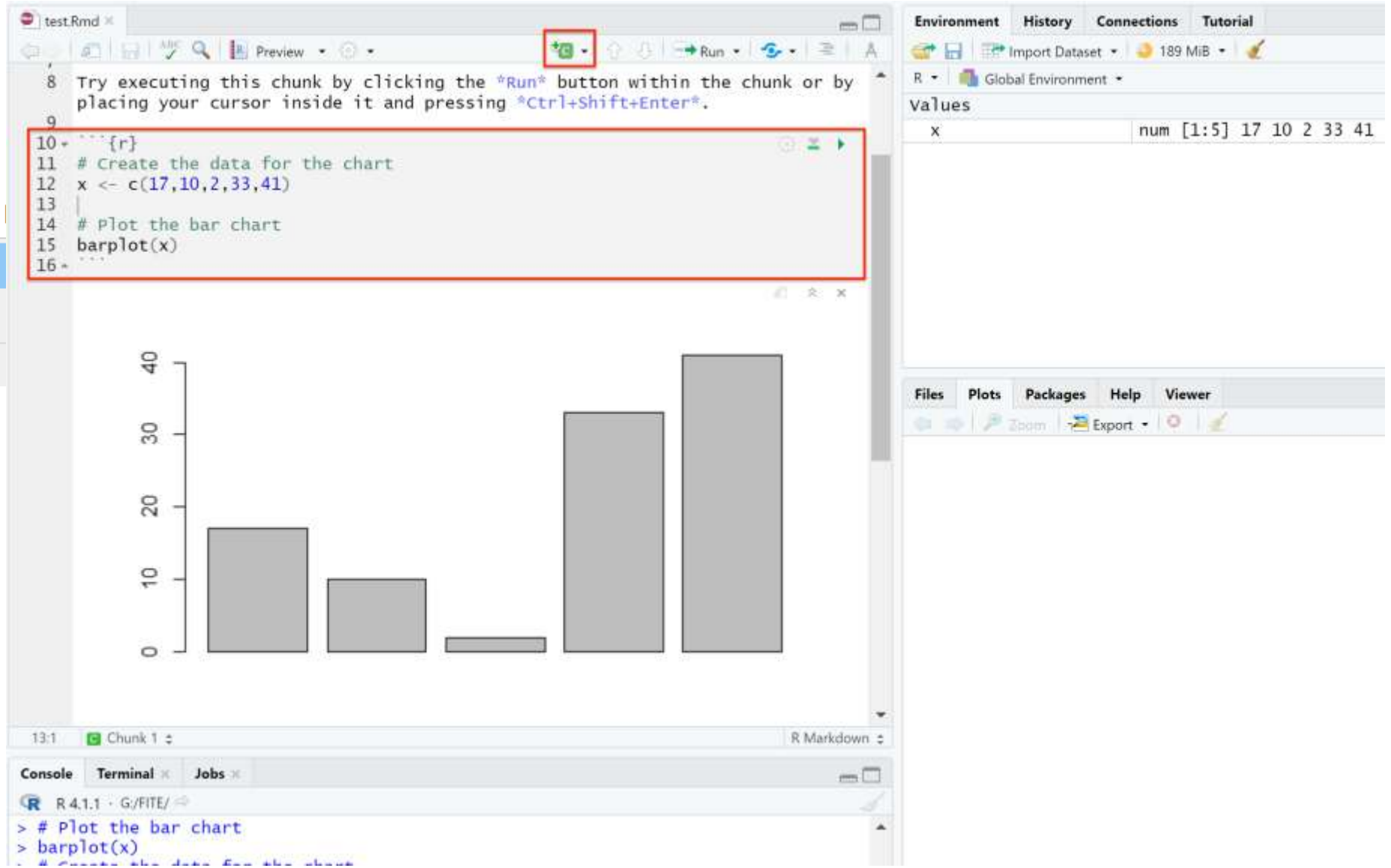


Note:
Press “Crtl +Enter” to run the code at source panel line by line.
Click “Source” to run the code at source panel.

RStudio

R Notebook





R Project

- Project --> New Project --> Choose a directory

What are R Packages?

- R packages are a collection of R functions, complied code and sample data.
- Load packages to implement different functions.

Packages

- Available CRAN Packages: https://cran.r-project.org/web/packages/available_packages_by_name.html
- Others: Github, etc.

Available CRAN Packages By Name

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
AATtools	Reliability and Scoring Routines for the Approach–Avoidance Task
ABACUS	Apps Based Activities for Communicating and Understanding Statistics
abbreviate	Readable String Abbreviation
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcADM	Fit Accumulated Damage Models and Estimate Reliability using ABC
ABCanalysis	Computed ABC Analysis
abclass	Angle–Based Large–Margin Classifiers
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests
abcrlda	Asymptotically Bias–Corrected Regularized Linear Discriminant Analysis
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abdiv	Alpha and Beta Diversity Measures
abe	Augmented Backward Elimination
abess	Fast Best Subset Selection
abglasso	Adaptive Bayesian Graphical Lasso
ABHgenotypeR	Easy Visualization of ABH Genotypes
abind	Combine Multidimensional Arrays
abjData	Databases Used Routinely by the Brazilian Jurimetrics Association
abjutils	Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association
abmR	Agent–Based Models in R
abn	Modelling Multivariate Data with Additive Bayesian Networks

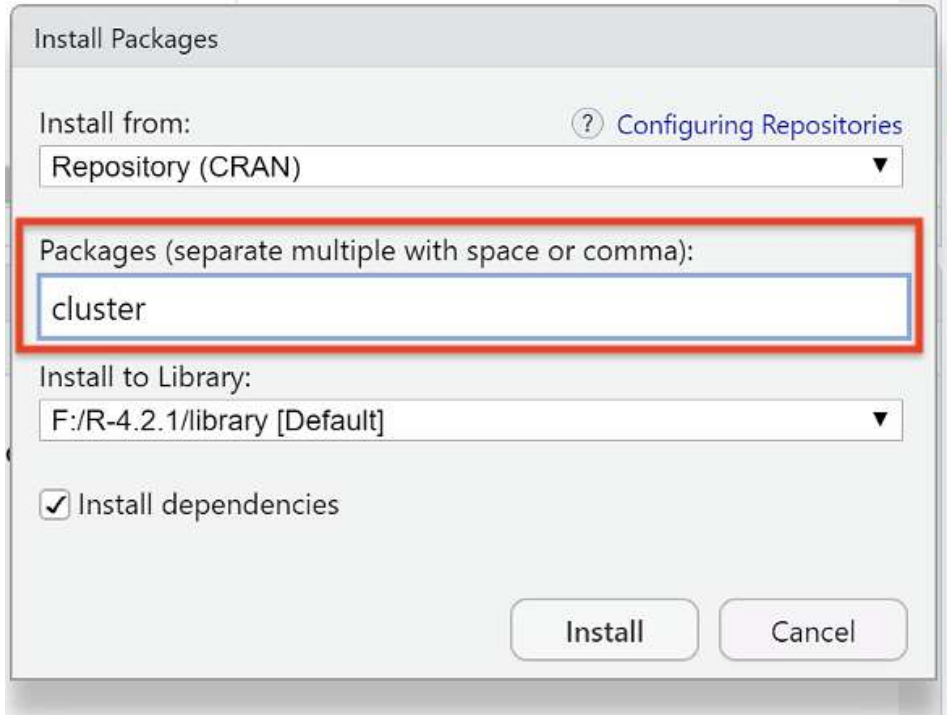
Install R Packages

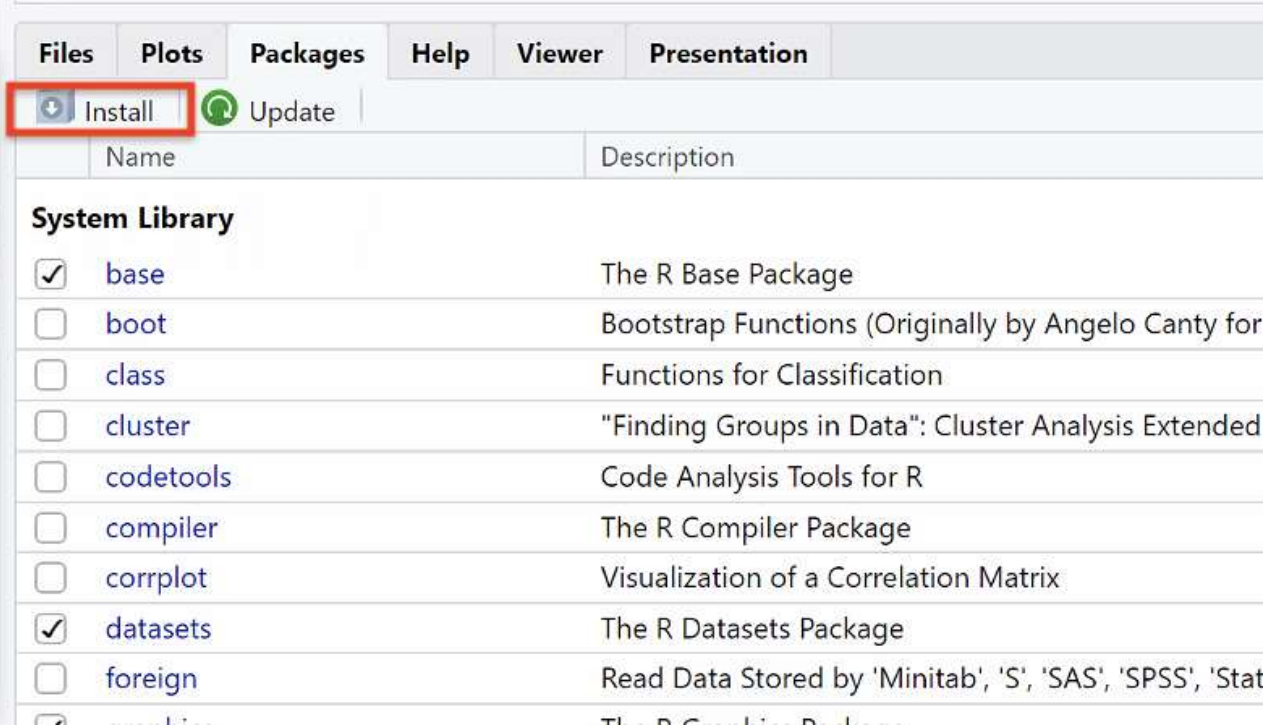
- Packages Installation Command: install.packages('package_name')

```
> install.packages("corrplot")
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/corrplot_0.92.zip'
Content type 'application/zip' length 3844927 bytes (3.7 MB)
downloaded 3.7 MB

package ‘corrplot’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Administrator\AppData\Local\Temp\RtmpEFXd64\downloaded_packages
```

Install R Packages

```
> installed.packages()
             Package      LibPath                Version   Priority
base         "base"       "F:/R-4.2.1/library"   "4.2.1"   "base"
```





CRAN Archive

- CRAN Archive: Find archived packages/older versions.
- https://cran.r-project.org/src/contrib/Archive/package_name/
- Installation Command: install.packages('devtools')

  devtools::install_version("package_name",version="package_version")

```
> install.packages('DMwR')
Warning in install.packages :
  package ‘DMwR’ is not available for this version of R

A version of this package for your version of R might be available elsewhere,
see the ideas at
https://cran.r-project.org/doc/manuals/r-patched/R-admin.html#Installing-packages
```

CRAN Archive

- https://cran.r-project.org/src/contrib/Archive/DMwR/

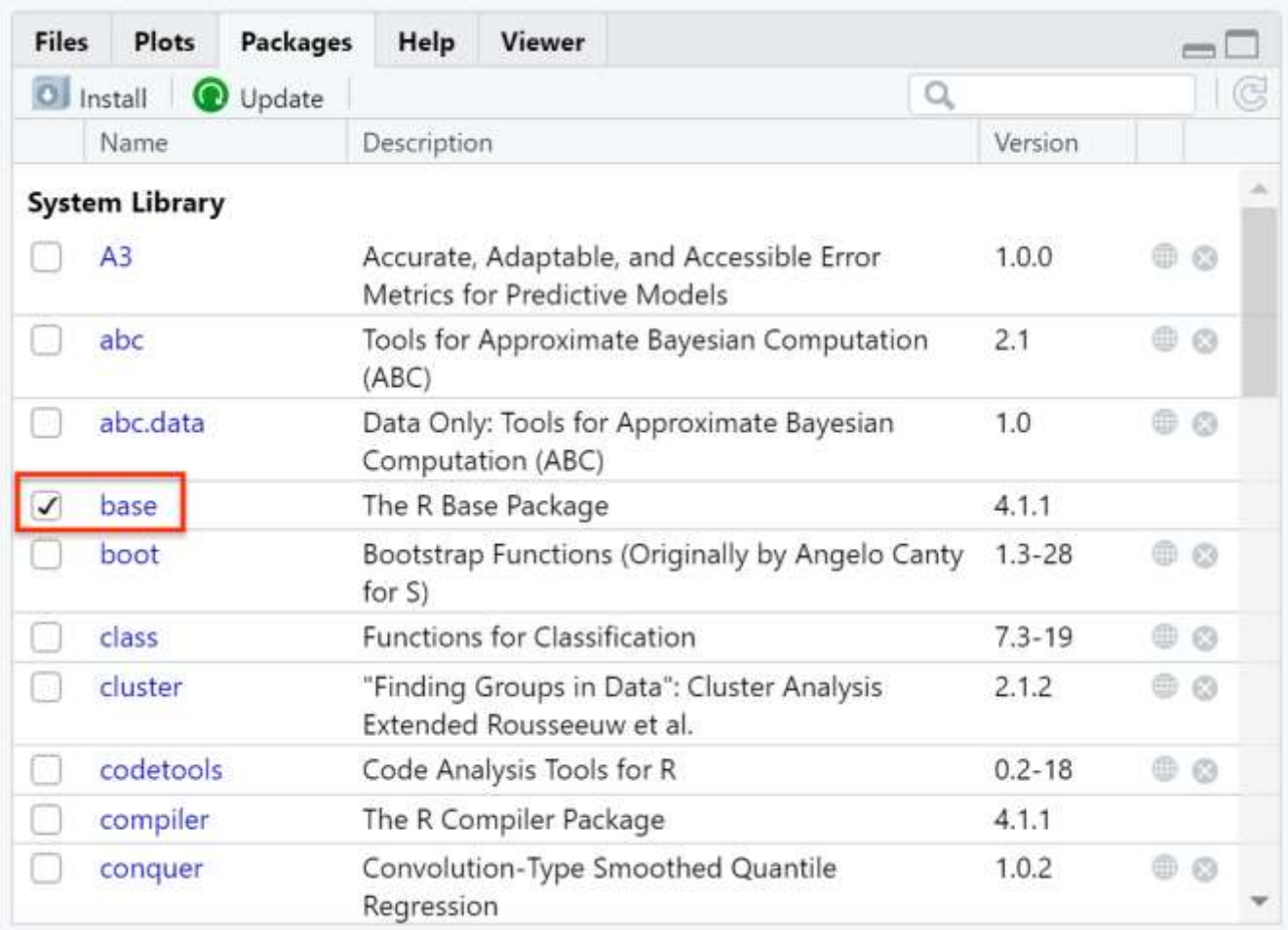

```
> install.packages('devtools')


> devtools::install_version("DMwR",version="0.4.1")
Downloading package from url: https://cran.rstudio.com//src/contrib/Archive/DMwR/DMwR_0.4.1.tar.gz
Installing 10 packages: bitops, caTools, gtools, zoo, xts, gplots, TTR, ROCR, abind, quantmod
trying URL 'https://cran.rstudio.com/src/contrib/bitops_1.0-7.tar.gz'
Content type 'application/x-gzip' length 10809 bytes (10 KB)
downloaded 10 KB

trying URL 'https://cran.rstudio.com/src/contrib/caTools_1.18.2.tar.gz'
Content type 'application/x-gzip' length 64960 bytes (63 KB)
downloaded 63 KB

trying URL 'https://cran.rstudio.com/src/contrib/gtools_3.9.3.tar.gz'
```

Load R Packages

- Command: library(package_name)

```
> library(A3)
Loading required package: xtable
Loading required package: pbapply
> detach("package:A3", unload = TRUE)
```

library(lattice,grid,DMwR)



References

- “R: What Is R?” The R Project for Statistical Computing, www.r-project.org/about.html.
- Neo, Benedict. “R for Data Science | Towards Data Science.” Medium, 14 Oct. 2020, towardsdatascience.com/how-to-learn-r-for-data-science-3a7c8326f969.
- “R Tutorial.” Tutorialspoint, www.tutorialspoint.com/r/index.htm.
- “R Tutorial.” GeeksforGeeks, www.geeksforgeeks.org/r-tutorial.
- “Setting up the R Environment.” HADES, ohdsi.github.io/Hades/rSetup.html.
- “R Packages: A Beginner’s Guide.” DataCamp Community, www.datacamp.com/community/tutorials/r-packages-guide#what.
- “How Do I Install a Package That Has Been Archived from CRAN?” Stack Overflow, stackoverflow.com/questions/24194409/how-do-i-install-a-package-that-has-been-archived-from-cran.