



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

What is R?

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.



What is R?

Why Use R?

- Free and open source
- Analytical support
- Extensive functionality

Steps

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

Setting up the R environment

Installing R

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

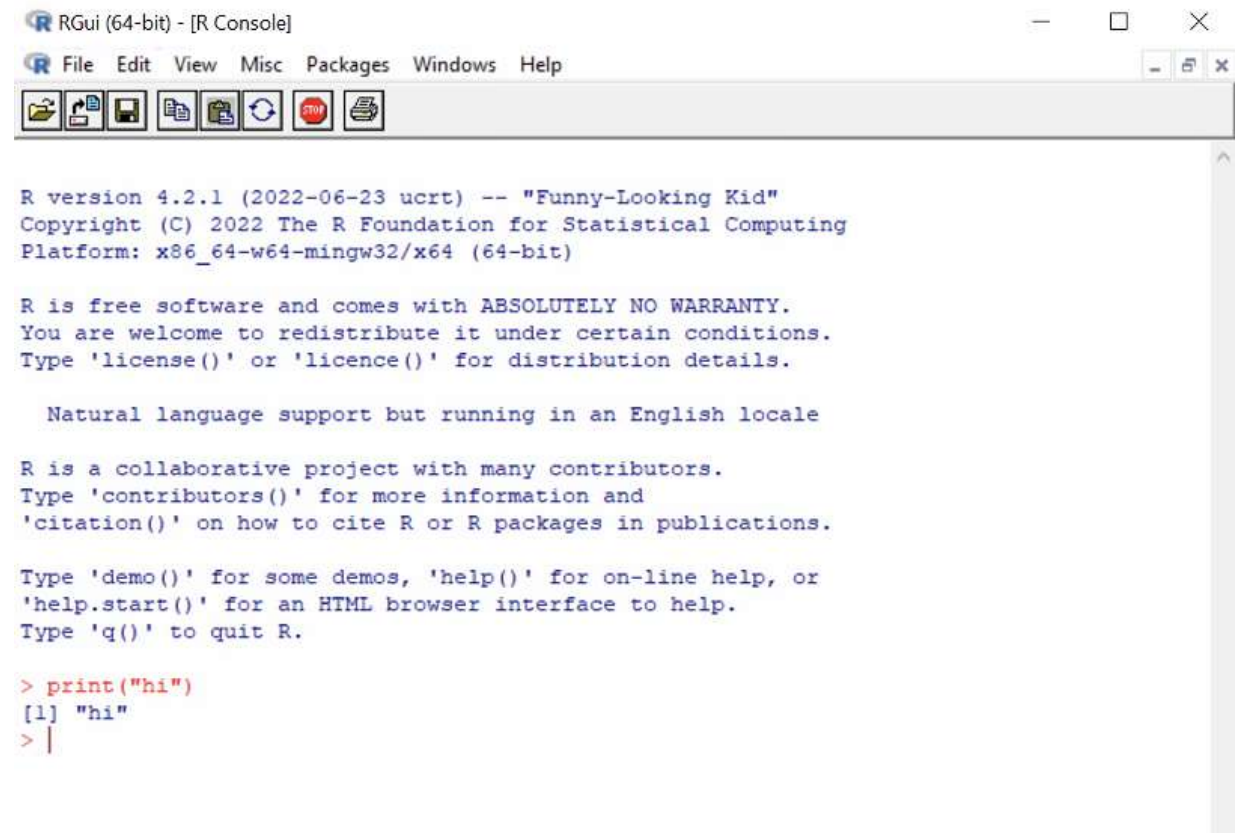
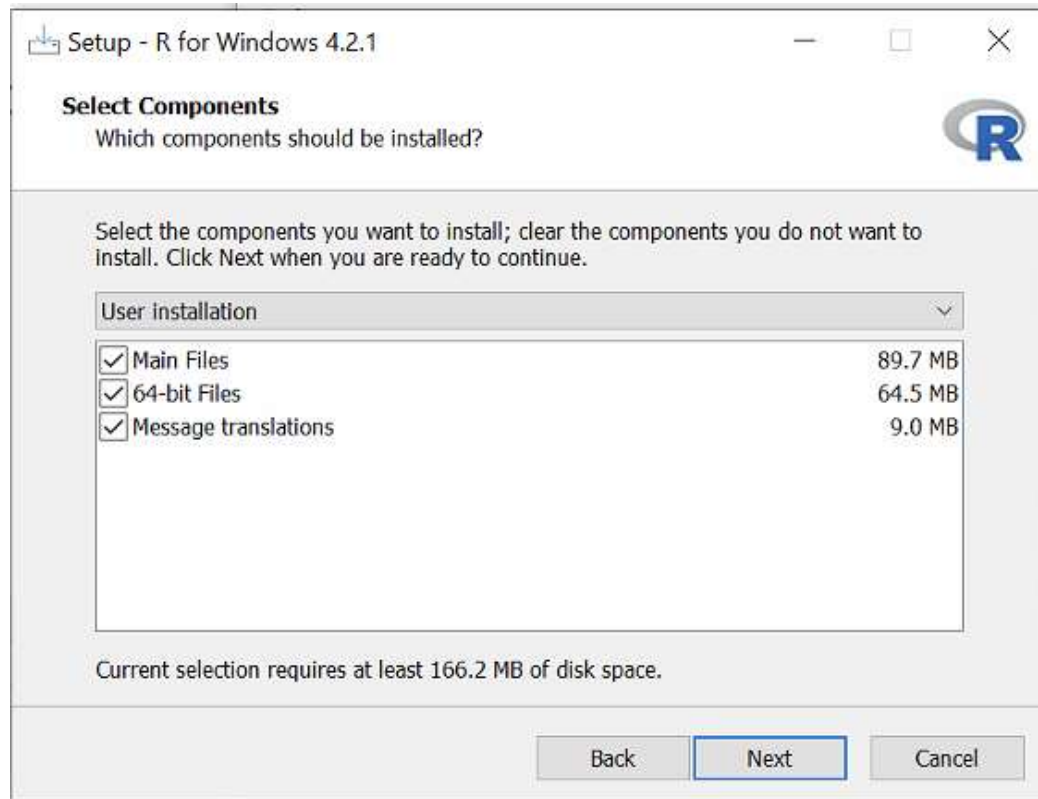
Download and Install R

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.

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 212.78 MB | [SHA-256: BCF6B866](#) | Version: 2023.06.2+561 |
Released: 2023-08-30

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

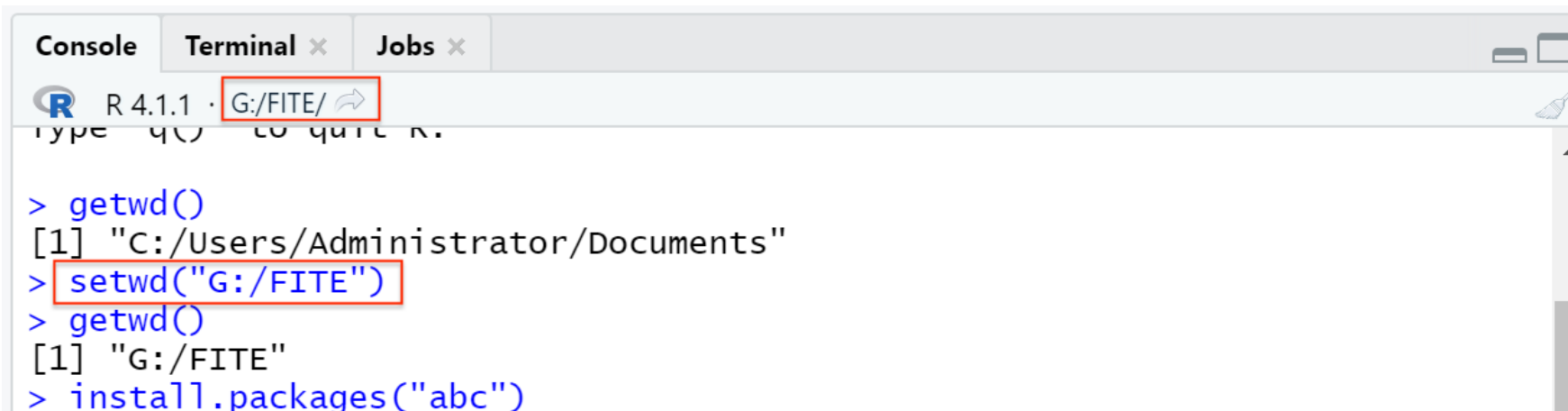
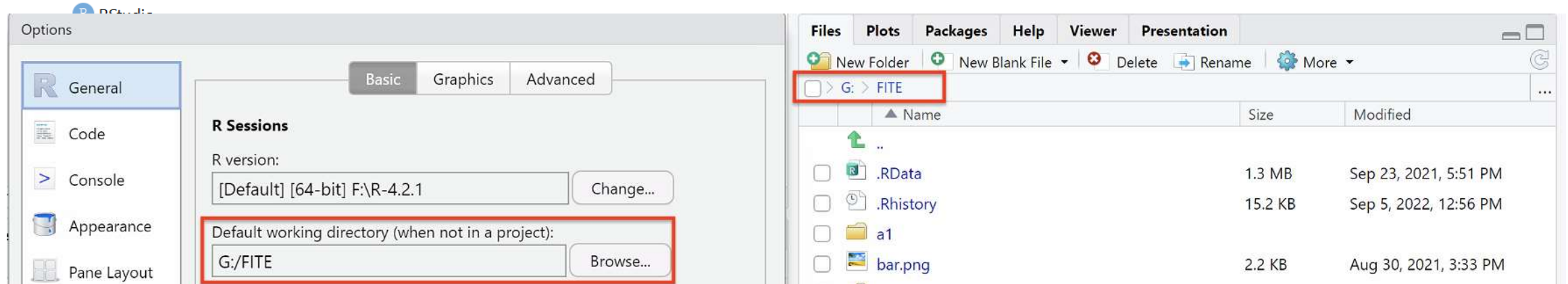
(Optional) Installing Rtools (Windows) / R build tools (Mac)

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.

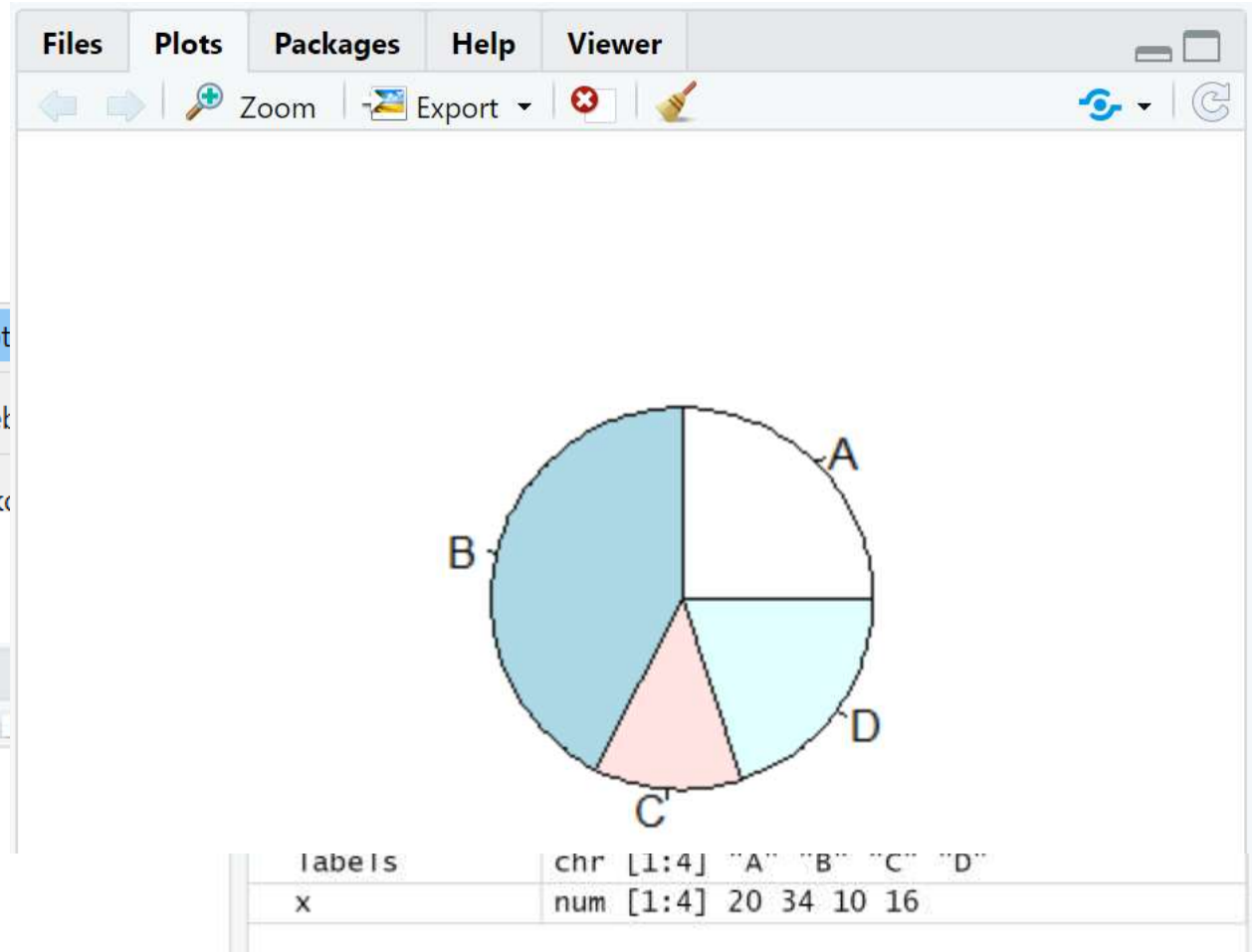
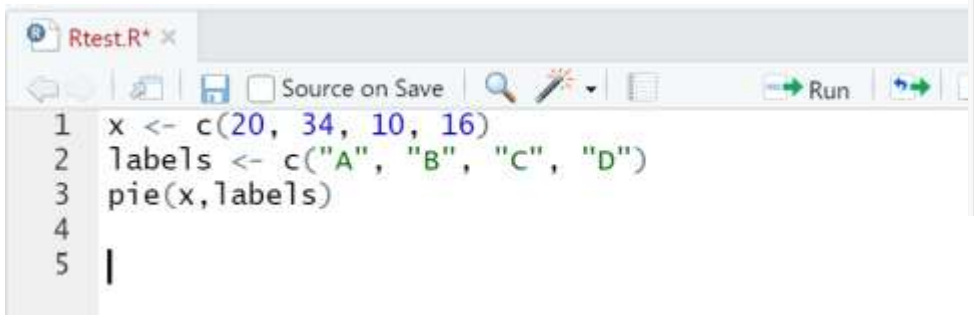
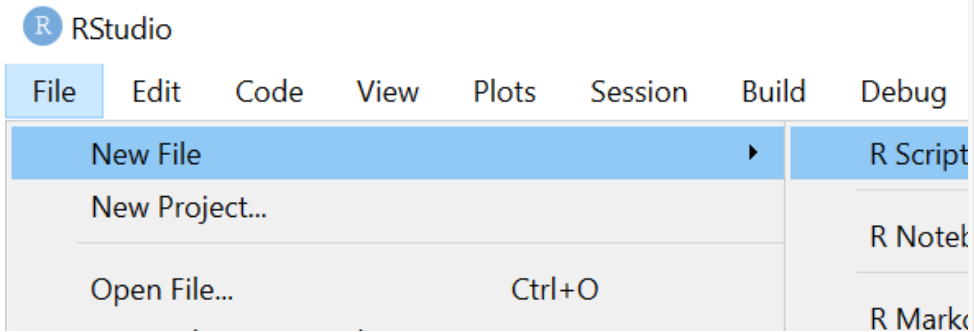
Working Directory

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



RStudio

R Script



Note:

Press "Ctrl +Enter" to run the code at source panel line by line.

Click "Source" to run the code at source panel.

RStudio

R Notebook

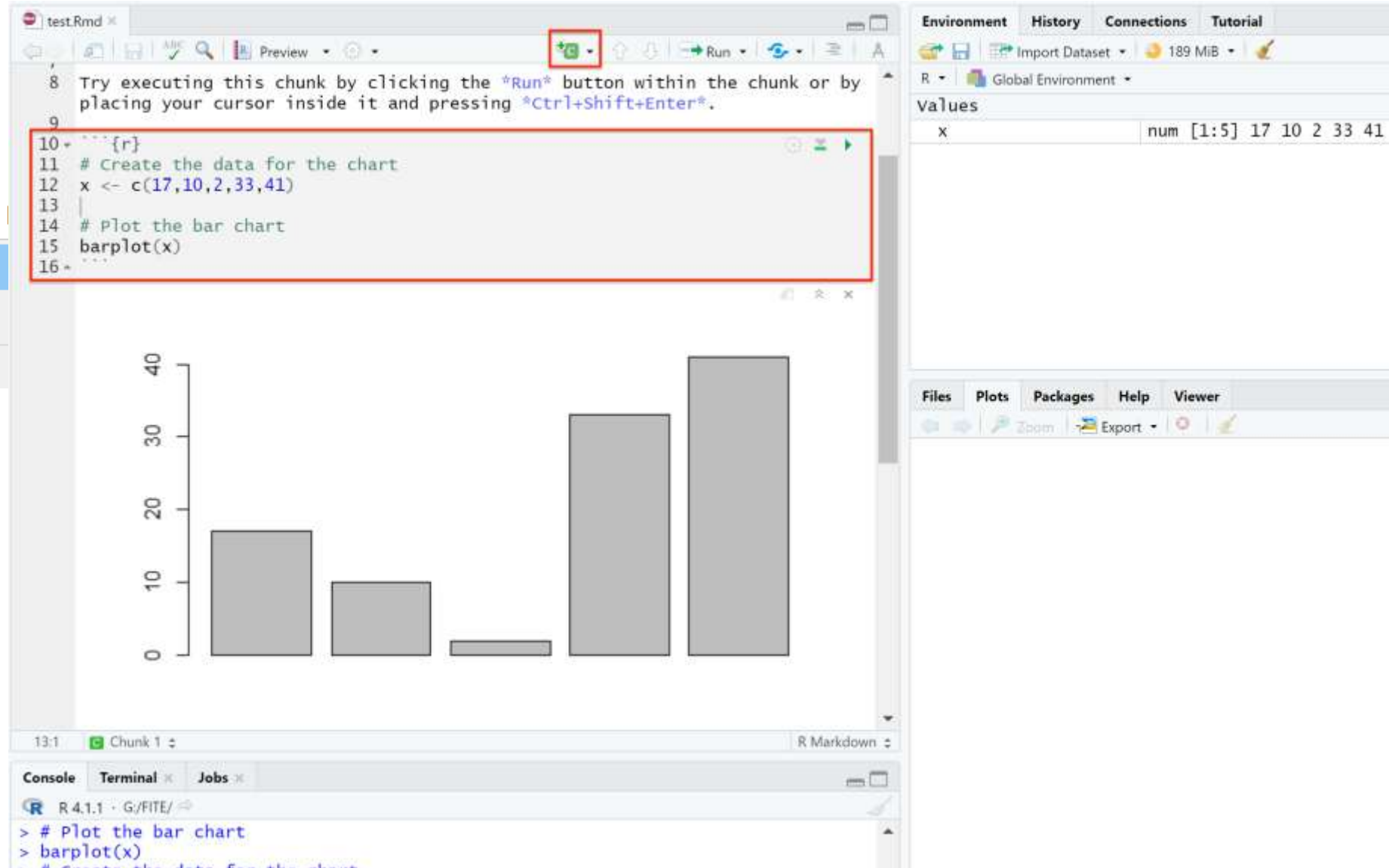
R RStudio

File Edit Code View

New File

New Project...

Open File



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](#)

[AATtools](#)

[ABACUS](#)

[abbreviate](#)

[abbyyR](#)

[abc](#)

[abc.data](#)

[ABC.RAP](#)

[abcADM](#)

[ABCanalysis](#)

[abclass](#)

[ABCOptim](#)

[ABCp2](#)

[abcrf](#)

[abcrlda](#)

[abctools](#)

[abd](#)

[abdiv](#)

[abe](#)

[abess](#)

[abglasso](#)

[ABHgenotypeR](#)

[abind](#)

[abjData](#)

[abjutils](#)

[abmR](#)

[ahn](#)

Accurate, Adaptable, and Accessible Error Metrics for Predictive Models

Reliability and Scoring Routines for the Approach–Avoidance Task

Apps Based Activities for Communicating and Understanding Statistics

Readable String Abbreviation

Access to Abbyy Optical Character Recognition (OCR) API

Tools for Approximate Bayesian Computation (ABC)

Data Only: Tools for Approximate Bayesian Computation (ABC)

Array Based CpG Region Analysis Pipeline

Fit Accumulated Damage Models and Estimate Reliability using ABC

Computed ABC Analysis

Angle–Based Large–Margin Classifiers

Implementation of Artificial Bee Colony (ABC) Optimization

Approximate Bayesian Computational Model for Estimating P2

Approximate Bayesian Computation via Random Forests

Asymptotically Bias–Corrected Regularized Linear Discriminant Analysis

Tools for ABC Analyses

The Analysis of Biological Data

Alpha and Beta Diversity Measures

Augmented Backward Elimination

Fast Best Subset Selection

Adaptive Bayesian Graphical Lasso

Easy Visualization of ABH Genotypes

Combine Multidimensional Arrays

Databases Used Routinely by the Brazilian Jurimetrics Association

Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association

Agent–Based Models in R

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

R Packages

Install R Packages

```
> installed.packages()
```

	Package	LibPath	Version	Priority
base	"base"	"F:/R-4.2.1/library"	"4.2.1"	"base"

Install Packages

Install from: [? Configuring Repositories](#)

Repository (CRAN)

Packages (separate multiple with space or comma):

cluster

Install to Library:

F:/R-4.2.1/library [Default]

☒ Install dependencies

Install Cancel

Files

Plots

Packages

Help

Viewer

Presentation

Install

Update

Name	Description
System Library	
<input checked="" type="checkbox"/> base	The R Base Package
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for
<input type="checkbox"/> class	Functions for Classification
<input type="checkbox"/> cluster	"Finding Groups in Data": Cluster Analysis Extended
<input type="checkbox"/> codetools	Code Analysis Tools for R
<input type="checkbox"/> compiler	The R Compiler Package
<input type="checkbox"/> corrplot	Visualization of a Correlation Matrix
<input checked="" type="checkbox"/> datasets	The R Datasets Package
<input type="checkbox"/> foreign	Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stat

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/atools_3.9.3.tar.gz'
```

Index of /src/contrib/Archive/DMwR

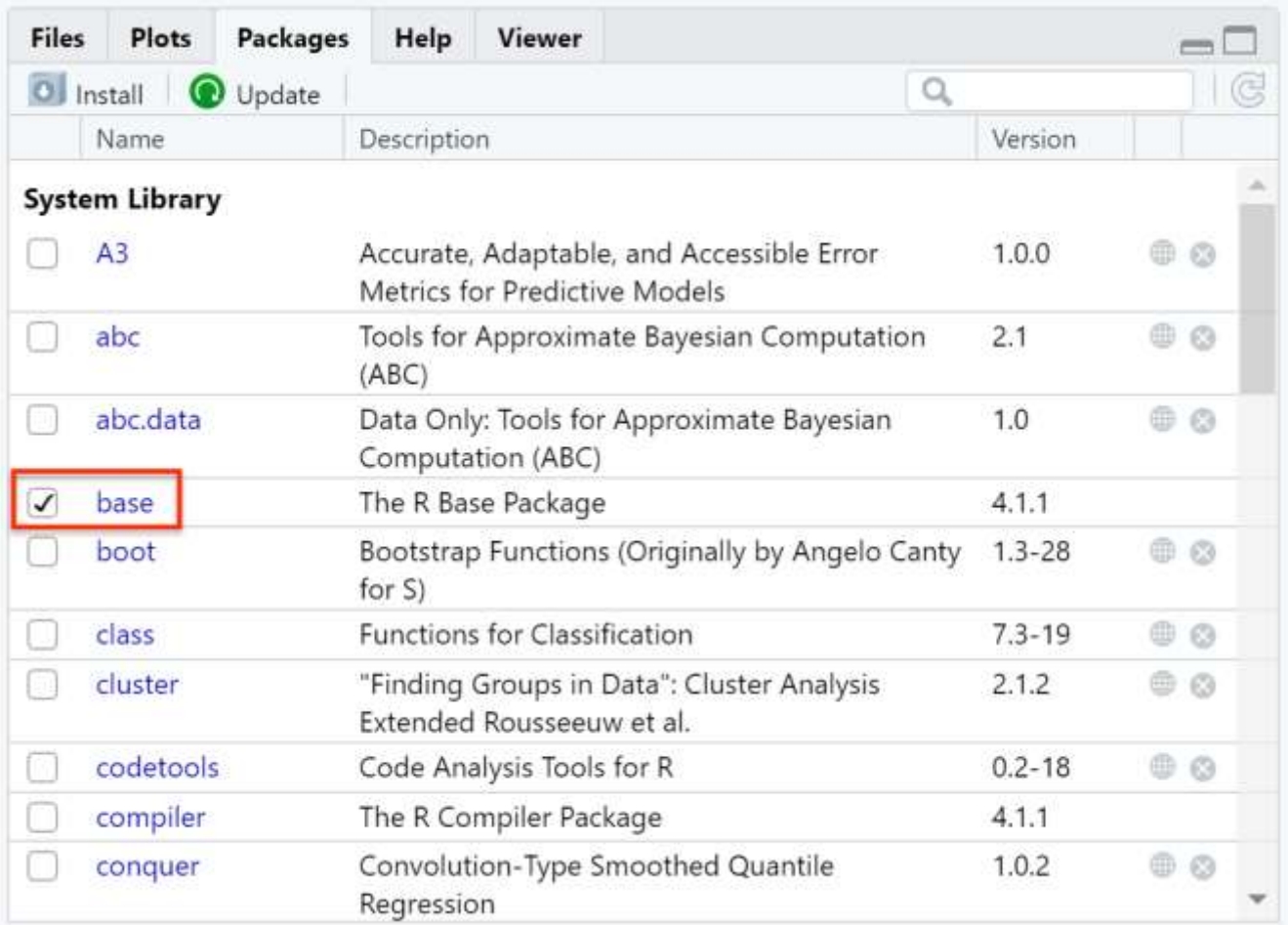
Name	Last modified	Size	Description
 Parent Directory		-	
 DMwR_0.1.0.tar.gz	2010-11-26 17:36	2.7M	
 DMwR_0.2.0.tar.gz	2011-03-17 20:49	1.8M	
 DMwR_0.2.1.tar.gz	2011-04-15 07:15	1.8M	
		1.8M	
		1.8M	
		2.2M	
		2.1M	
		1.8M	
		1.9M	

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)
```



Files	Plots	Packages	Help	Viewer
Install		Update		
Name	Description	Version		
System Library				
<input type="checkbox"/> A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models	1.0.0		
<input type="checkbox"/> abc	Tools for Approximate Bayesian Computation (ABC)	2.1		
<input type="checkbox"/> abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)	1.0		
<input checked="" type="checkbox"/> base	The R Base Package	4.1.1		
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-28		
<input type="checkbox"/> class	Functions for Classification	7.3-19		
<input type="checkbox"/> cluster	"Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.	2.1.2		
<input type="checkbox"/> codetools	Code Analysis Tools for R	0.2-18		
<input type="checkbox"/> compiler	The R Compiler Package	4.1.1		
<input type="checkbox"/> conquer	Convolution-Type Smoothed Quantile Regression	1.0.2		

References

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.