

CS 133 - Introduction to Computational and Data Science

Instructor: Renzhi Cao
Computer Science
Department
Pacific Lutheran University
Spring 2017



Reference book

- *R Programming for Data Science. By Roger Peng. ISBN-10: 1365056821, April 20, 2016.*

Homework

- *Read the book from Page 1 to Page 12.*

Background of R

- *What is R?*
- *A dialect of S, S is a language that was developed by John Chambers and others at the old Bell Telephone Laboratories, originally part of AT&T Corp*

Background of R

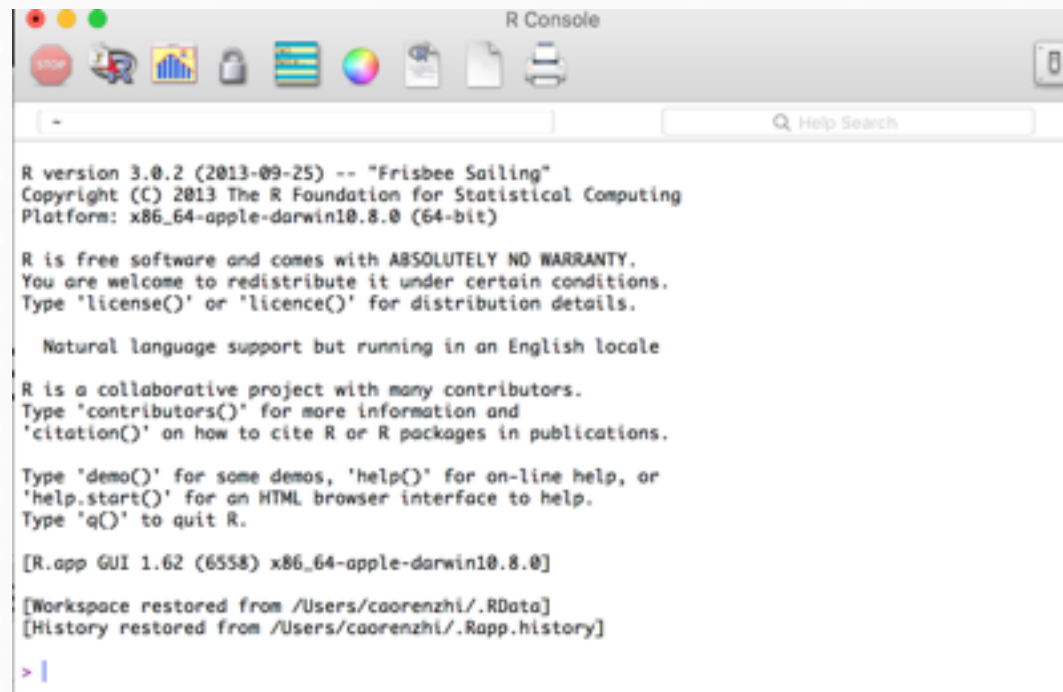
What are the basic features of R?

- *R is free and open source.*
- *R runs on most standard operating systems.*
- *R has frequent releases.*
- *R has sophisticated graphic capabilities.*
- *R is both useful for interactive work and a powerful programming language for developing new tools.*

Installation of R

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

- 1. *Windows*
- 2. *Mac*



```
R Console
-
Help Search

R version 3.0.2 (2013-09-25) -- "Frisbee Sailing"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin10.8.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

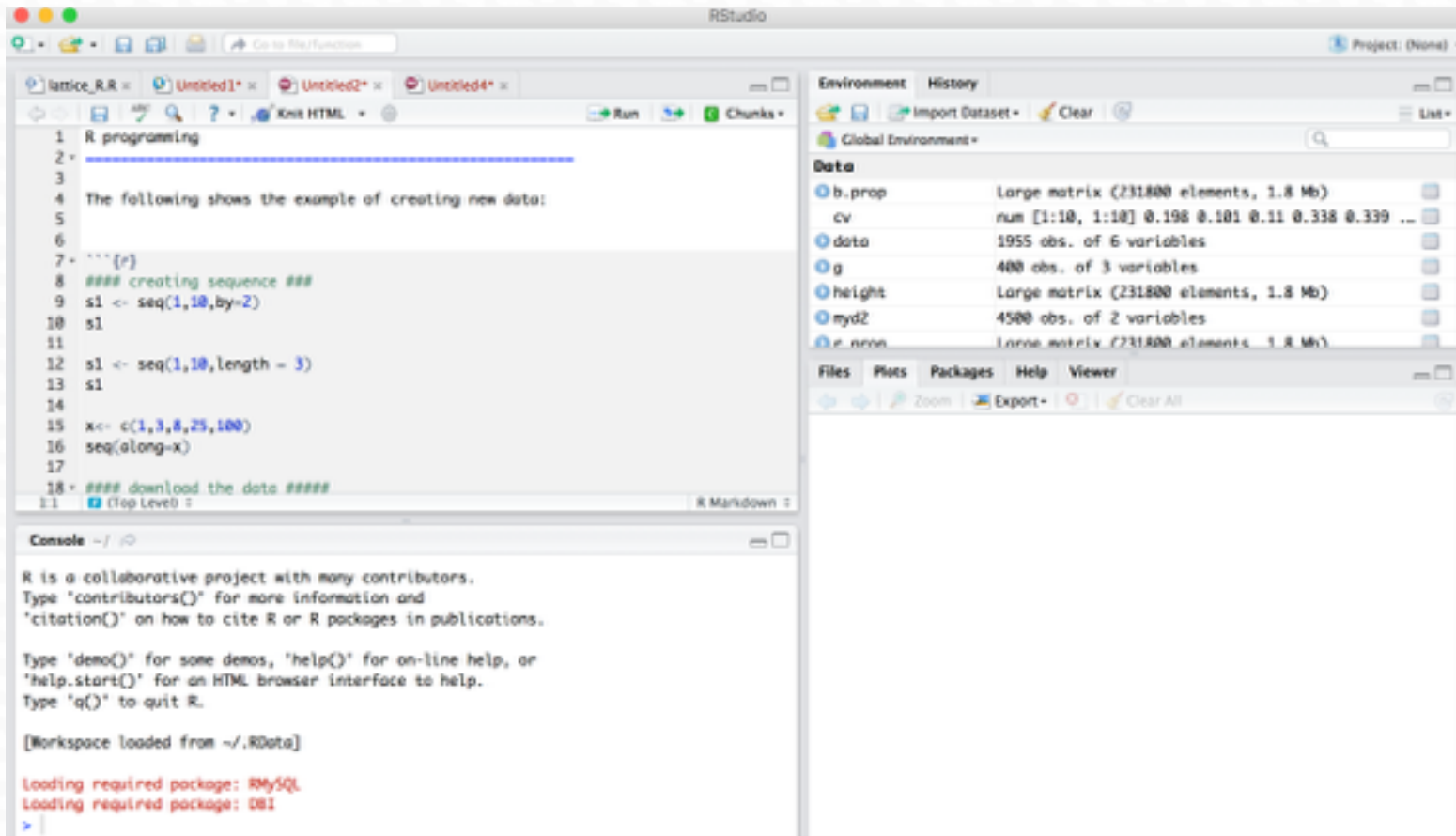
[R.app GUI 1.62 (6558) x86_64-apple-darwin10.8.0]

[Workspace restored from /Users/caorenzhi/.RData]
[History restored from /Users/caorenzhi/.Rapp.history]

> |
```

Useful IDE for R only: Rstudio

<https://www.rstudio.com>



The screenshot displays the RStudio IDE interface. The main editor window shows an R script with the following code:

```
1 R programming
2
3
4 The following shows the example of creating new data:
5
6
7- ```{r}
8 #### creating sequence ###
9 s1 <- seq(1,10,by=2)
10 s1
11
12 s1 <- seq(1,10,length = 3)
13 s1
14
15 x<- c(1,3,8,25,100)
16 seq(along=x)
17
18- #### download the data #####
19- (Top Level)
```

The console window at the bottom left shows the following output:

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

Loading required package: RMySQL
Loading required package: DBI
>
```

The environment pane on the right side of the interface shows the following data objects:

Object	Description
b.prop	Large matrix (231800 elements, 1.8 Mb)
cv	num [1:10, 1:10] 0.198 0.101 0.11 0.338 0.339 ...
data	1955 obs. of 6 variables
g	400 obs. of 3 variables
height	Large matrix (231800 elements, 1.8 Mb)
mydZ	4500 obs. of 2 variables
r.non	Large matrix (231800 elements, 1.8 Mb)

Demo of R

- *Setting work directory and edit R code*
- *Demo: `ls()`, `dir()`, `getwd()`, `setwd()` in windows*
- *Use Atom to edit R code in Windows*

R console input

- *`<-` as assignment operator*
- *`#` indicate comment*

Get familiar with the R environment