

An Intro to R for Non-Programmers

William F. Lamberti ¹

George Mason University

Feb. 28, 2017

¹Student in MS Statistical Science and
PhD Computational Sciences and Informatics

Outline

Introduction

- Course

- About Me

R

- What is R

- Text Editor

- .R Files

Using R

- Objects and Types

- Functions

- The Directory

- .RData Files

- Packages in R

Other Resources

Introduction

- ▶ 1 hour course (7:30 PM - 8:30 PM)
- ▶ For those who want to know what R is even if you don't know how to program
- ▶ Will reference RGalleon.com pages for additional information
- ▶ Will *not* go over the "why" of statistics (simply do not have the time)
- ▶ If you like this course, please let the GMU's Student ASA Chapter know
- ▶ If you have any questions, please feel free to ask at any point



About Me

- ▶ From New Jersey
- ▶ Undergrad at The College of New Jersey
- ▶ Major: Mathematical Statistics
- ▶ Minors:
 - ▶ Political Science
 - ▶ Actuarial Science and Risk Management
- ▶ Likes include running, reading, video games, cooking, hiking



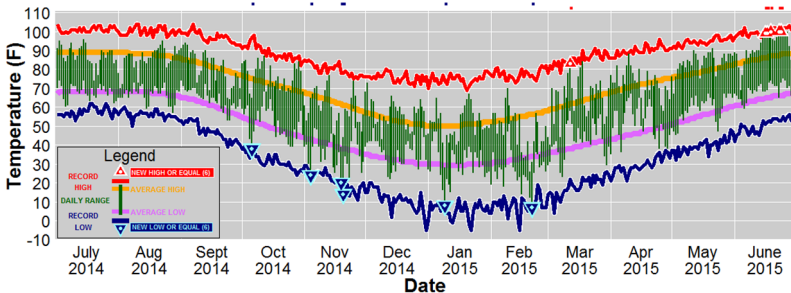
What is R

- ▶ R is free statistical software to analyze data
- ▶ Functions
 - ▶ Visualization
 - ▶ Statistical Analysis
 - ▶ Machine Learning ("A-I")
- ▶ Requires ability to "code" (Much more approachable than other languages such as Java)
- ▶ Used in research, practical applications, academia, government, and private sector
- ▶ See <http://www.rgalleon.com/topics/learning-r/about-r/>



Example: Visualize Weather Data²

Charolette, North Carolina, Temperatures from July 2014 to June 2015 and Historical Temperatures from 1962 to 2014



²<http://www.rgalleon.com/fivethirtyeight-graphic-redesign/>

Downloading R

- ▶ Go to <https://www.r-project.org/>
- ▶ Pick place to download
- ▶ Follow fairly typical download process
- ▶ See <http://www.rgalleon.com/topics/learning-r/downloading-r-and-other-tools/> for more information and video

Text Editor

- ▶ You will need a text editor to write code
- ▶ Programs like Microsoft Words have "invisible" lines of code to present the word document as you see it in the program
- ▶ Implies that you need a text editor
- ▶
 - ▶ Notepad++ (Windows/Mac with Wine)
 - ▶ Notepad (Windows)
 - ▶ R Studio (Windows/Mac/Linux)
 - ▶ Notepadqq (Ubuntu/Lubuntu)



.R Files

- ▶ R has its own file type just like Microsoft Word has its own file type (i.e. .doc or .docx)
- ▶ .R files are used to save scripts and custom functions
- ▶ The importance of these files will not be emphasized today, but it is important as you learn more about R

Objects and Types

- ▶ Kinds of Objects

- ▶ Vectors
- ▶ Matrices

- ▶ Kinds of Types

- ▶ Numeric

- ▶ Used for number data
- ▶ Can use for matrix algebra

- ▶ Character

- ▶ Used for word data
- ▶ Use " "

- ▶ Logical

- ▶ Used for "Yes" or "No" questions
- ▶ Used for two categories (i.e. Male or Female, Normal or Cancer)

Example Code: Numerical

```
1 #Numerical vector and matrix
2
3 vector1<-c(1,2,3,4)
4 vector1 #displaying the vector
5
6 matrix1<- matrix(c(1,2,3,4), nrow=2)
7 matrix1 #displaying the matrix
8
9 sum(vector1)
```

Example Code: Character

```
1 #Character vector and matrix
2
3 vector2<-c("one","two","3")
4 vector2 #displaying the vector
5
6 matrix2<- matrix(c("1","2","3","4"), nrow=2)
7 matrix2 #displaying the matrix
8
9 sum(vector2)
```

Example Code: Logical

```
1 #Logical vector
2
3 vector3<- c(TRUE, FALSE, TRUE, FALSE, FALSE)
4 vector3 #displaying vector3
```

Functions

- ▶ Used to perform complicated calculations with ease
 - ▶ Sum - `sum()`
 - ▶ Mean - `mean()`
 - ▶ Standard deviation - `sd()` or `sqrt(var())`
 - ▶ etc.
- ▶ Used to do various other actions such as load data into R - `load()`
- ▶ Can write your own functions as well

Example Code: Functions

```
1 #Functions Code Example
2
3 x<- c(-0.6, -2.3, -0.4)
4
5 ?mean
6 mean(x) # calculate the mean
7
8 ?sd
9 sd(x) #calculate the stand deviation
```

The Directory

- ▶ The Directory is where you are working
- ▶ If you want to work with data, it needs to be the directory
- ▶ You can change your working directory easily
- ▶ Think of the directory as your desk; if your pen is not at your desk, you cannot write. You either need to move a pencil to your desk, or go to a desk with a pen.

Example Code: The Directory

```
1 #The Directory Code
2
3 dir()
```

.RData Files

- ▶ R can read many different file types such as .CSV, .TXT
- ▶ R has its own native file type called .RData
- ▶ R does have the ability to convert .CSV, .TXT, and other files, but we will not cover that today

Example Code: .RData Files

```
1 #Using .RData Files
2
3 x<- c(-0.6, -2.3, -0.4)
4
5 save(x, file="MyData.RData")
6 dir()
7
8 ls() #check objects in workspace
9 rm(list=ls()) #clears workspace
10 ls()
11 mean(x)
12
13 load(file="MyData.RData")
14 ls()
15 mean(x)
```

Packages in R

- ▶ Used to extend functionality of R
- ▶ Need to be loaded into R every time you want to use them
- ▶ See <http://www.rgalleon.com/topics/learning-r/packages-in-r/>

Example

- ▶ With the following data, calculate the sum: 1, 2, 3, 4, 5
- ▶ Answer: 15
- ▶ You have 5 minutes

Example Code

```
1 #Creating data
2 x<-c(1,2,3,4,5)
3
4 x
5
6 sum(x)
```

Exercise

- ▶ Download R and setup R
- ▶ Go to <http://http://www.rgalleon.com/talks/george-mason-university-february-28-2017/>
- ▶ Download the exercise files associated with this lecture
- ▶ **Goal: Calculate the mean and standard deviation from the data.**
- ▶ Bonus: Create a histogram of the data.
- ▶ You have 15 minutes. I will be available if you have any questions.
- ▶ We will go over the answer together after 15 minutes have passed

Example Code

```
1  dir()
2
3  #Loading data
4  load("EX_1data.RData")
5
6  ls()
7
8  length(data)
9
10 data
11
12 mean(data)
13 sd(data)
14
15 hist(data) #histogram
16
17 windows() #makes 2nd space for 2nd graphic
18 hist(data, main="Histogram of Exercise 1 Data
    ", col='blue')
```


Other Resources

- ▶ <https://www.rgalleon.com/>
- ▶ <https://www.Lynda.com>
- ▶ <https://www.codeschool.com/courses/try-r>
- ▶ <http://www.statmethods.net/>
- ▶ Stack Overflow
- ▶ <https://cran.r-project.org/doc/manuals/R-intro.html>

Outline

Introduction

- Course

- About Me

R

- What is R

- Text Editor

- .R Files

Using R

- Objects and Types

- Functions

- The Directory

- .RData Files

- Packages in R

Other Resources