Intro2R Workshop 2022 Fall

Bin Chen

2022-09-22

Table of Contents

Welcome to the “Introduction To R” workshop! This workshop is for scholars who are interested in data analysis and visualization with R. In this workshop, we will talk about (1) what is R and RStudio, (2) R basics before diving into coding such as working directory, file path, function, package, and import/export etc. and (3) data classes and structures in R.

Recommended Materials:

- An Intro to R: https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf
- R for Data Science: https://r4ds.had.co.nz/introduction.html

R and RStudio

R is a programming language for statistical computing and graphics.

RStudio is an integrated development environment for R and Python. RStudio combines a source code editor, build automation tools and a debugger.

R vs Python

R: Statistical computing and graphics.

Python: Machine learning, deep learning, etc.

Both languages have a lot in common and can do most common tasks.

R for J&M scholars

Statistics: Fast, Flexible, Visualization.

Collect Data: Scraping webpages, Using API.

Network Analysis: Descriptive (e.g., igraph), Inference (e.g., ergm)

Text Analysis: Classification (e.g., Topic Modeling), Semantic Network

Walkthrough R Studio

Install R: https://www.r-project.org/

Install RStudio: https://www.rstudio.com/products/rstudio/download/

RStudio window: *Console*, *Environment*, *Files*.

R Projects and RMD

Use *RProject* folder to manage different projects. Starting from R project to access your codes and data is a good habit, and it also makes your life way easier! (e.g., wd issues)

RMarkdown is a file format for making dynamic documents with R (HTML, PDF, PPT, etc.)

Let’s create a new R project and an RMD file.

WD (Working Directory) and Path

Get current WD

```
getwd()
```

```
## [1] "C:/Users/chenb/OneDrive - The University of Texas at
Austin/Documents/R/22fall_intro2r_workshop"
```

Notice the reverse "/" from Windows path:

Windows Path: "C:- The University of Texas at Austin\22fall_intro2r_workshop"

Set a different WD

Suppose you want to access a file stored under "/data" and now you want to change your WD.

```
setwd("C:/Users/chenb/OneDrive - The University of Texas at
Austin/Documents/R/22fall_intro2r_workshop/data")

getwd()

## [1] "C:/Users/chenb/OneDrive - The University of Texas at
Austin/Documents/R/22fall_intro2r_workshop/data"
```

Note: You don't do this in practice because you can just specify the path of file (current WD will be your start point). Tips: Use interface to (re)set WD.

Package and Function

R packages are extensions to the R statistical programming language.

Install a package

```
install.packages("praise")
```

Load a package

```
library(praise)
#require(praise)
```

Use ? + command/package name to get help, e.g., C:/Users/chenb/OneDrive - The University of Texas at Austin/Documents/R/win-library/4.1/praise/help/praise

```
?praise
```

Take a try

```
praise()

## [1] "You are exquisite!"
```

One more

```
praise()

## [1] "You are impressive!"
```

Data Types and Structure

Date Types

Numeric and Integer

```
x <- 1
class(x) # you can also check [environment] to see types of objects

## [1] "numeric"

y = 1L # "<-" and "=" are same
class(y)

## [1] "integer"
```

Note. sometimes you see “double” in R, which is basically same as “numeric”

Create a vector

```
z <- c(1, 2, 3, 4, 5)
```

Character and Factor

```
a <- "1"
class(a)

## [1] "character"

b <- "high"
class(b)

## [1] "character"

c <- c("high", "high", "low", "high", "low")
class(c)

## [1] "character"
```

Sometimes we want to convert character to factor

```
c1 <- as.factor(c)
class(c1)

## [1] "factor"
```

Then you can use high, low to obtain factor levels

```
levels(c1)

## [1] "high" "low"
```

Date

```
d0 <- "09/20/2022"
class(d0)
```

```
## [1] "character"

d1 <- as.Date(d0, "%m/%d/%y")
class(d1)

## [1] "Date"

d <- c("09/20/2022", "09/21/2022", "09/22/2022", "09/24/2022", "09/20/2022")
```

Logical Data

```
L = T
class(L)

## [1] "logical"

L1 = FALSE
class(L1)

## [1] "logical"

L == L1 #you use "==" for equal to, rather than "="

## [1] FALSE
```

Data Structure

List

```
mix <- list(66.6, "high", TRUE)
class(mix)

## [1] "list"
```

Dataframe

```
df <- data.frame(z, c, d1)

df

##   z    c         d1
## 1 1 high 2020-09-20
## 2 2 high 2020-09-20
## 3 3  low 2020-09-20
## 4 4 high 2020-09-20
## 5 5  low 2020-09-20
```

Useful links

- https://www.google.com/
- https://stackoverflow.com/
- https://www.reddit.com/r/RStudio/

```
praise()
```

```
## [1] "You are super-excellent!"
```