R Programming

By

Rakhi Singh

What is R

- R is a programming language and software environment for statistical computing. It is used by statisticians and data miners for development and analysis.

S and S-Plus

- S is statistical programming language developed by John Chambers, Rick and Allen by calling FORTRAN subroutines in 1976 while in Bell Labs.
- Later on the system was written in C and resemble the system we have today.
- Two modern implementation of S are R and S-PLUS.
- R is a part of GNU free software project.
- S-PLUS is a commercial implementation of S programming language sold by TIBCO software.

History of R

- Created in New Zealand by Statistics professor Robert Gentleman and Ross in 1991 at University of Auckland.
- In 1995 they used GNU General public License to make R free software.
- In 2000 R version 1.0.0 is released.
- R version 3.0.2 was released in Dec 2013.
- R is very similar to S, making it easy for S-PLUS user to switch over.

Features of R

- R is a tool for
- Data Manipulation
- Modeling and Computation
- Data Visualization

Modeling & Computation

- Statistical Analysis
- Regression Analysis
- Classification
- Numerical Simulation
- Principal component Analysis

Data Manipulation

- Connecting to different data sources
- Slicing and dicing data
- Modifying data
- Cleaning data
- Filling data

Data Visualization

- Graphical view of data
- Composing Statistical graphs
- Visualizing fit of models

SAS vs R

- R is free, SAS costs a lot of money
- R is open source, SAS is proprietary
- R has newest statistics method much more quickly than SAS.
- R is object oriented but SAS is table based
- R is better for build your own graphics but SAS is not so good. But they are improving.

Design of R system

- It has 2 conceptual parts
- The base R system that you download from CRAN
- It has packages like utils, stats, datasets, graphics, grDevices, grid, methods, tools, parallel, compiler, splines, tcltk, stats4.
- Other than that there are around 4000 packages on CRAN developed by users and programmers and many more are present on people's website.

Obtaining R

- Go to Google.
- Search for Cran R
- Click on Cran.r-project.org
- You will find three links
    - Download R for Linux
    - Download R for Mac
    - Download R for Windows
- Click on required link

Installing R on windows

- Next it will ask subdirectory. Select **base**
- Then click on Download R 3.3.1 for windows.
- Once download done, double click the setup file.
- It will ask for language. Select English
- Select the installation directory.
- In components select all of them
- In startup options select No.
- In display mode select SDI
- In Help Style select Plain text
- In Internet access select Standard.
- Check all boxes in Select Additional Tasks.
- After this finish installation. Open R through startup.

Setting your working directory

- All files which R is reading or writting are stored in working directory.
- > getwd()
- [1] “c:/Users/rdpeng/Documents”
- Goto file, select option change dir.
- Now select the directory from explorer to make it working directory.
- > dir()
- This will give the list of all files in your working directory.
- Later on you can save all your files/function in this directory

Objects

- All the things we encounter and manipulate in R are objects.
- R has 5 basic atomic classes of objects. They are character, real number, integers, complex numbers and logicals.

R as a calculator

- Simple math

```
> 2+2
4
```

- Storing results in variables

```
> x<- 2+2 #<- is a syntax for “=” assignment
> x^2
16
```

Atomic classes

- R has five basic or atomic classes of objects
- Character
- Numeric
- Integer
- Complex
- Logical

Atomic classes

- x <- 5 #<- is an assignment operator, x is numeric
- is.numeric(x) returns TRUE
- x <-'5' # here x is of character data type
- is.character(x) returns TRUE
- x = as.integer(3) # interger data type
- is.integer(y) returns TRUE
- A = 5+2i # complex, Re(a) real part of a, Im(a) imaginary part of a.
- z <-TRUE # logical
- is.logical(z) returns TRUE
- y <- c(1,2,3,4,5) # here y is a vector,
- is.vector(y) returns TRUE

R as a calculator

- Please try following operations on command prompt.
- > 1+1
- >1+10+12+1234
- >sqrt(16)
- >log10(100)
- >10*3
- >pi
- >100*1.15*pi
- >floor(1.9)
- >ceiling(1.9)
- >round(1.9)
- >round(1.4)
- >abs(1.9)
- > 23%%5
- >23%/%5

Operators

- Airthmatic :
- Addition +
- Subtraction -
- Multiplication *
- Division /
- Power ^
- Modulo %%,
- Integer Division %/%

- Comparison :
- < lesser than
- > greater than
- <= lesser than or equal to
- >= greater than or equal to
- == equal
- != different

Operators

- Logical :
- !x   logical NOT.
- X & Y logical AND.
- X && Y identical.
- X | Y   logical OR.
- X || Y identical.
- Xor(x, y) exclusive OR.

Attributes

- All objects in R have two intrinsic attributes: mode and length.
- There are four modes – numeric, character, complex and logical
- Length is the number of elements in the object.
- To display the mode and length of an object, one can use the functions mode and length, respectively.
- >X <- 1
- > mode(x)
- [1] “numeric”
- > length(x)
- [1] 1
- A <- “test”; compare <- TRUE; z <- 1i
- > mode(A); mode(compare) ; mode(z)
- ## test this command

Exponents

```
> X <- 5/0
> X  ## print x
[1] Inf
> exp(x)
[1] Inf
> Exp(-x)
0
> x -x
[1] NAN
```

Questions