Introduction to Scientific Computing for Biologists

ISCB20.09 - Introduction to R

Md. Jubayer Hossain
https://jhossain.me/
jubayer@hdrobd.org

Founder
Health Data Research Organization
Lead Organizer
Scientific Computing for Biologists

01 February 2021

Section-2: Getting Started

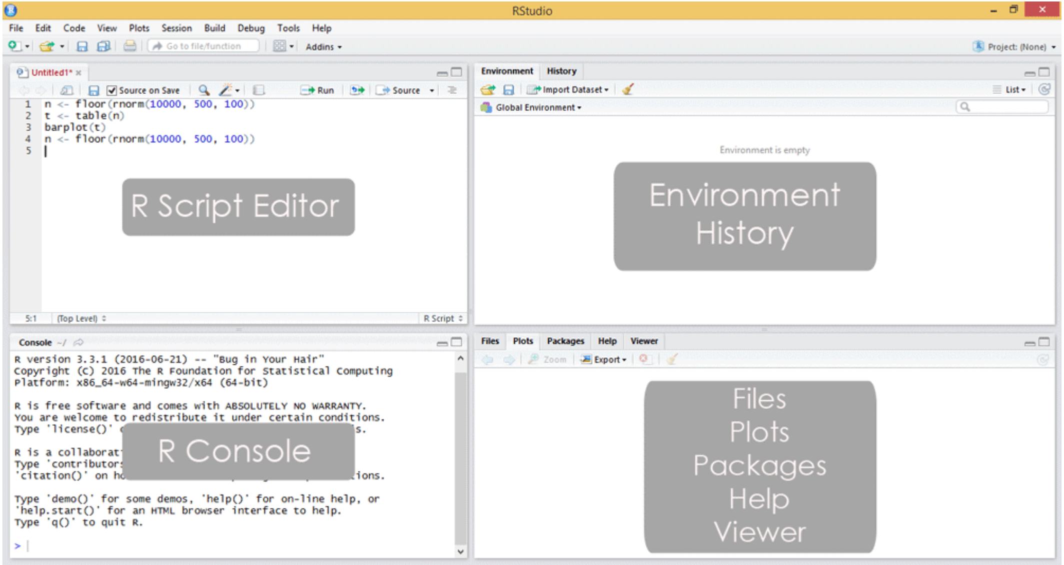
Installing R

- Installing R on Windows - https://cran.r-project.org/bin/windows/base/
- Installing R on Linux - https://cran.r-project.org/bin/linux/ubuntu/README.html
- Installing R on Mac - https://cran.r-project.org/bin/macosx/

Installing RStudio

- RStudio is an Integrated Development Environment(IDE) available for R that is built by RStudio.
- Download and Install https://rstudio.com/products/rstudio/download/#download

The RStudio Interface



```
n <- floor(rnorm(10000, 500, 100))
t <- table(n)
barplot(t)
n <- floor(rnorm(10000, 500, 100))
```

```
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' o                                   s.

R is a collaborat
Type 'contributor
'citation()' on ho                                   ions.

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.

>
```

Section-3: Variables and Reserved Keywords

Variables(Identifiers) in R

- Variables are used to **store data**, whose value can be changed according to our need.

Variables(Identifiers) in R

- Variables are used to **store data**, whose value can be changed according to our need.
- A variable is a name given to a memory location, which is used to store values in a computer program.

Variables(Identifiers) in R

- Variables are used to **store data**, whose value can be changed according to our need.
- A variable is a name given to a memory location, which is used to store values in a computer program.
- Variables in R programming can be used to store numbers (real and complex), words, matrices, and even tables.

Variables(Identifiers) in R

- Variables are used to **store data**, whose value can be changed according to our need.
- A variable is a name given to a memory location, which is used to store values in a computer program.
- Variables in R programming can be used to store numbers (real and complex), words, matrices, and even tables.
- R is a dynamically programmed language which means that unlike other programming languages, we do not have to declare the data type of a variable before we can use it in our program.

Variables(Identifiers) in R

- Variables are used to **store data**, whose value can be changed according to our need.
- A variable is a name given to a memory location, which is used to store values in a computer program.
- Variables in R programming can be used to store numbers (real and complex), words, matrices, and even tables.
- R is a dynamically programmed language which means that unlike other programming languages, we do not have to declare the data type of a variable before we can use it in our program.
- Unique name given to variable (function and objects as well) is **identifier**.

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
- It should not start with a number (e.g:- 2x)

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
- It should not start with a number (e.g:- 2x)
- It should not start with a dot followed by a number (e.g:- .2x)

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
- It should not start with a number (e.g:- 2x)
- It should not start with a dot followed by a number (e.g:- .2x)
- It should not start with an underscore (e.g:- _x)

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
- It should not start with a number (e.g:- 2x)
- It should not start with a dot followed by a number (e.g:- .2x)
- It should not start with an underscore (e.g:- _x)
- Reserved words in R cannot be used as identifiers(e.g: TRUE, FALSE)

Rules for writing Identifiers in R

- Identifiers can be a combination of letters, digits, period (.) and underscore (_).
- It must start with a letter or a period. If it starts with a period, it cannot be followed by a digit.
- It should not start with a number (e.g:- 2x)
- It should not start with a dot followed by a number (e.g:- .2x)
- It should not start with an underscore (e.g:- _x)
- Reserved words in R cannot be used as identifiers(e.g: TRUE, FALSE)

Basically, there are 5 naming conventions

- alllowercase: e.g. `myname`
- period.separated: e.g. `new.name`
- underscore_separated: e.g. `my_name`
- lowerCamelCase: e.g. `myName`
- UpperCamelCase: e.g. `MyName`

Creating Variables

Using equal(=) operator

```
x = 10
```

Using leftward(<-) operator

```
y <- 15
```

Using rightward(->) operator

```
z <- 20
```

Reserved Keaywords in R

- Don't use any reserved keyword as variable name. List all of reserved words in R by using (?Reserved).
- The reserved words in R's parser are - if else repeat while function for in next break - TRUE FALSE NULL Inf NaN NA NA_integer_ NA_real_ NA_complex_ NA_character_

Entering Input

At the R prompt/console we type expressions.

```
num <- 10
```

The <- symbol is the **assignment** operator.
The grammar of the language determines whether an expression is complete or not.

Evaluation

When a complete expression is entered at the R console, it is evaluated and the result of evaluated expression is returned. The result may be auto-printed.

```r
x <- 10
x
```

```
## [1] 10
```

```r
x <- 10
print(x)
```

```
## [1] 10
```

```r
x <- 10
cat(x)
```

```
## 10
```

Comments in R

The # character indicates a comment.

```
x <- 10 # This is a comment
```

Anything to the right of the # (inlcuding the # itself) is ignored.
This is only comment character in R.
R does not support multi-line comments or comment block.