Programming for Data Science in R

Efstathios (Stathis) D. Gennatas, MBBS AICSM PhD

2020-10-20

Contents

Welcome

This is the online book for the new UCSF Biostat 213, Fall 2020. It is being updated weekly.

It was created using bookdown[1] (**?**)

EDG, San Francisco, CA, September 2020

[1] `https://CRAN.R-project.org/package=bookdown`

Preface

Throughout this book you will see boxes with R code followed by its output, if any. The code (or input) is decorated with a teal border on the left to separate it from its output, like in the following example:

```
x <- rnorm(200)
x[1:20]
```

```
 [1] -0.105246533 -1.123473372 -2.083177469  0.589137998 -0.045634427
 [6] -0.035330051 -2.207060073  1.161668977 -0.942975317  0.818967929
[11]  0.685416119  0.818576854 -0.008722038 -2.347228798  1.319719902
[16]  0.727747357 -1.272729163  0.082654474 -0.633194696 -1.141172476
```

Notice that R adds numbers in brackets in the beginning of each row. This happens when R prints the contents of a vector. The number is the integer index of the first element in that row. Therefore, the first one is always `[1]` and the number of the subsequent rows depends on how many elements fit in each line. If the output is a single element, it will still have `[1]` in front of it.

Also notice that if we enclose the assignment operation of a variable in parentheses, this prints the resulting value of the variable. Therefore, this:

```
(y <- 4)
```

```
[1] 4
```

is equivalent to:

```
y <- 4
y
```

```
[1] 4
```

Note that if you mouse over the input code box, a clickable "Copy to clipboard" appears on the top right of the box allowing you to copy paste into an R session or file.