IntroMarkDown

Viraj @DataTasteMaker

March 19, 2016

Contents

Introduction to Markdown through Class Assignment

Using the Basic of R, we can see how to create R-Markdown file. This is useful to document your code for future reference. It is intersting and also very easy to use.

Remove/clear the memory

Always clear the R-Environment and plots history before starting with any new code/script. This helps freeing up the memory which otherwise is unnecessarily blocked my the R-objects.

```
## null device
##           1
```

Load libraries if any

Always have the list of libraries to be included at the start of your script. Also if possible do make a reference as to why you are including/loading this library.

```
library(psych)  # for the use of "describe()" function to get the summary statistics.
```

---

1. Create a vector of length 12. Print the 3rd and 7th element

```
vec12 <- seq(1,12)
vec12
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

```
vec12[c(3,7)]
```

```
## [1] 3 7
```

2. Create a vector of lenth 21 stating with the value 3.4 and ending with value 9.6.

```
vec2 <- seq(3.4,9.6, length.out = 32 )
vec2
```

```
##  [1] 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0 5.2 5.4 5.6 5.8 6.0 6.2 6.4 6.6
## [18] 6.8 7.0 7.2 7.4 7.6 7.8 8.0 8.2 8.4 8.6 8.8 9.0 9.2 9.4 9.6
```

What is the position of value 8.0?

```
which(vec2 == 8.0)
```

```
## [1] 24
```

3. Create the 4x4 matrix

```
m <- matrix(c(1,3,4,2,6,2,12,9,3,7,8,9,2,1,8,0), byrow = T, nrow = 4)
m
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1    3    4    2
## [2,]    6    2   12    9
## [3,]    3    7    8    9
## [4,]    2    1    8    0
```

** 3a. Return 2nd column of the matrix**

```
m[,2]
```

```
## [1] 3 2 7 1
```

3b. Change the element in the 2nd row and 2nd col to 3 Element in 2nd row and 2nd col – Before

```
m[2,2]
```

```
## [1] 2
```

Element in 2nd row and 2nd col – After

```
m[2,2] <- 3
m[2,2]
```

```
## [1] 3
```

4. Randomly sample and form a vector of length = 200 and mean = 4 and std.dev =2.

Use set.seed(100)

```
set.seed(100)
r200 <- rnorm(200, mean = 4, sd = 2)
head(r200)
```

```
## [1] 2.995615 4.263062 3.842166 5.773570 4.233943 4.637260
```

```
tail(r200)
```

```
## [1] 7.790552 3.140018 7.151094 4.323882 1.829094 5.153875
```

4a. Total Values less than 2.

```
sum(r200 < 2)
```

```
## [1] 25
```

4b. Total values between 3 and 5 (excluding 3 and 5)

```
sum(r200 > 3 & r200 <5)
```

```
## [1] 98
```

4c. The mean and std. dev of the data

```
mean(r200)
```

```
## [1] 4.014053
```

```
sd(r200)
```

```
## [1] 1.826132
```

5a Read the day1_data.csv file

```
day1data <- read.csv("./datasets/IntroMarkdown_Data.csv", header = T)
head(day1data)
```

```
##          X1       X2 X3       X4 X5
## 1 3.3315342 1.004398  3 5.894764  0
## 2 0.7459214 1.923868  3 5.505580  1
## 3 4.1737246 1.706009  3 5.639135  1
## 4 3.4343659 1.098579 NA 4.508899  1
## 5 3.4063718 1.451969  4 5.199390  1
## 6 3.7275906 1.058451  4 5.449157  0
```

Dimension of the data file read in

```
dim(day1data)
```

```
## [1] 230   5
```

*__Mean, Std. dev., Min and Max of column 4 i.e. X4__

```
mean(day1data$X4, na.rm = T)
```

```
## [1] 4.92087
```

```
sd(day1data$X4, na.rm = T)
```

```
## [1] 1.028961
```

```
min(day1data$X4, na.rm = T)
```

```
## [1] 2.091993
```

```
max(day1data$X4, na.rm = T)
```

```
## [1] 7.929153
```

** Alternate way - using describe from *psych* library**

```
describe(day1data$X4, na.rm = T)
```

```
##   vars   n mean   sd median trimmed  mad  min  max range  skew kurtosis
## 1    1 229 4.92 1.03   4.94    4.94 0.96 2.09 7.93  5.84 -0.13     0.08
##     se
## 1 0.07
```