Lecture6 Class Work Set

Forrest Wang

Table of contents

Function Basics

All functions in R consist of at least 3 things: name, input argument, body

> Q1

#IMPORTANT! #THIS IS THE OBJECTIVE: I've got a vector called student 1, and I want to find the average and drop the lowest score

Example input vectors to start with

student1 <- c(100, 100, 100, 100, 100, 100, 100, 90) student2 <- c(100, NA, 90, 90, 90, 90, 97, 80) student3 <- c(90, NA, NA, NA, NA, NA, NA, NA)

```r
# Example input vectors to start with
student1 <- c(100, 100, 100, 100, 100, 100, 100, 90)
student2 <- c(100, NA, 90, 90, 90, 90, 97, 80)
student3 <- c(90, NA, NA, NA, NA, NA, NA, NA)
```

I can calculate the average by using 'mean()' function

```
which.min(student1)
```

```
[1] 8
```

```
student1
```

```
[1] 100 100 100 100 100 100 100  90
```

#MY PROTOTPYE: I can get the same vector without the 8th element

```
p <- which.min(student1)
q <- student1[-p]
mean(q)
```

```
[1] 100
```

#PROF Working PROTOTPYE:

```
mean(student1[-which.min(student1)])
```

```
[1] 100
```

#Student 2 # "is.na() -> 0" basically goes down your data set, and at every point it asks, "is it NA?". If the point isn't NA, it moves on. If the point is NA, it assigns it as "NA"

```
student2[is.na(student2)] <- 0
mean( student2[-which.min(student2)])
```

```
[1] 91
```

#Student 3 #Is.na is like a find and replace

```
student3[is.na(student3)] <- 0
mean( student3[-which.min(student3)])
```

```
[1] 12.85714
```

#FUNCTION MAKING!

```
x <- student1
x[ is.na(x) ] <- 0
mean( x[ -which.min(x) ] )
```

```
[1] 100
```

#Now turn it into a function

```
grade <- function(x) {
  x[ is.na(x) ] <- 0
  mean( x[ -which.min(x) ] )
}
```

```
grade(student1)
```

```
[1] 100
```

Q2

```
url <- "https://tinyurl.com/gradeinput"
gradebook <- read.csv(url, row.names = 1)
```

#H’ave a wee look lad at th’ first six rows

```
head(gradebook)
```

```
          hw1 hw2 hw3 hw4 hw5
student-1 100  73 100  88  79
student-2  85  64  78  89  78
student-3  83  69  77 100  77
student-4  88  NA  73 100  76
student-5  88 100  75  86  79
student-6  89  78 100  89  77
```

#How to “apply function”

```r
results <- apply(gradebook, 1, grade)
```

#Who is the top scoring student?

```r
which.max(results)
```

```
student-18 
        18 
```

```r
results[ which.max(results) ]
```

```
student-18 
      94.5 
```

```r
max(results)
```

```
[1] 94.5
```

Q3

#Use apply:

```r
which.min(apply(gradebook, 2, sum, na.rm=TRUE))
```

```
hw2 
  2 
```

Q4

#We’re looking for correlation, cor

```r
mask <- gradebook
mask [ is.na(mask) ] <- 0

cor(mask$hw5, results)
```

```
[1] 0.6325982
```

#˜˜˜0.6 means a slightly stronger correlation

```r
apply(mask, 2, cor, y=results)
```

```
      hw1       hw2       hw3       hw4       hw5 
0.4250204 0.1767780 0.3042561 0.3810884 0.6325982 
```