Class6: R functions

Alice Lai (PID:A16799081)

Functions are how we get work done in R. We call functions to do everything from reading data to doing analysis and outputing plots and results.

All functions in R have at least 3 things: - a **name** (you get to pick this) - input **arguments** (there can be only one or loads) - the **body** (where the work gets done, these codes are between the curly brackets)

1. A first silly function

```
x <- 10
y <- 10
x + y
```

```
[1] 20
```

```
add <- function(x) {
  y <- 10
  x + y
}

add(1)
```

```
[1] 11
```

```
add <- function(x, y) {
  x + y
}
```

```r
add(10, 10)
```

```
[1] 20
```

```r
add(10, 100)
```

```
[1] 110
```

2. grade() function

trials

```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)

mean(student1)
```

```
[1] 98.75
```

```r
mean(student2, na.rm=TRUE)
```

```
[1] 91
```

```r
mean(student3, na.rm=TRUE)
```

```
[1] 90
```

```r
min(student1)
```

```
[1] 90
```

```
min(student2)
```

[1] NA

```
min(student3)
```

[1] NA

```
#which.min() returns the location where the minimal value occurs.
which.min(student1)
```

[1] 8

```
which.min(student2)
```

[1] 8

```
which.min(student3)
```

[1] 1

```
minLocation <- which.min(student1)
newGrades <- student1[-minLocation]
meanGrade <- mean(newGrades, na.rm = TRUE)
meanGrade
```

[1] 100

Q1. my working function!

```
grade <- function(grades) {
  #find and replace all the NA to 0
  grades[is.na(grades)] <- 0
```

```
    #find the location of the min value
    minLocation <- which.min(grades)
    #create a new vector that omits the min value
    newGrades <- grades[-minLocation]
    #calculate the mean of the new vector
    meanGrade <- mean(newGrades)
    meanGrade
  }

  grade(student1)
```

[1] 100

```
  grade(student2)
```

[1] 91

```
  grade(student3)
```

[1] 12.85714

Q1. applying function to url gradebook

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

  # array = gradebook, margin = row (1 for row, 2 for col), function = grade
  grades <- apply(gradebook, 1, grade)
```

Q2. top scoring student

```
  which.max(grades)
```

student-18
       18

```
max(grades)
```

[1] 94.5

Q3. hardest homework

```
apply(gradebook, 2, mean, na.rm=TRUE)
```

```
     hw1      hw2      hw3      hw4      hw5
89.00000 80.88889 80.80000 89.63158 83.42105
```

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

hw3
  3

Q4. highest correlation with average grade score

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

```
           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   0  73 100  76
student-5   88 100  75  86  79
student-6   89  78 100  89  77
student-7   89 100  74  87 100
student-8   89 100  76  86 100
student-9   86 100  77  88  77
student-10  89  72  79   0  76
student-11  82  66  78  84 100
student-12 100  70  75  92 100
```

```
student-13  89 100  76 100   80
student-14  85 100  77  89   76
student-15  85  65  76  89    0
student-16  92 100  74  89   77
student-17  88  63 100  86   78
student-18  91   0 100  87  100
student-19  91  68  75  86   79
student-20  91  68  76  88   76
```

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

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