

Class06

Samuel Do (PID:A15803613)

2/3/2022

Exploring R functions

#Grade calculations

```
#Input factors
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
```

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

```
## [1] 91
```

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

```
## [1] 90
```

#Determining function is.na()

```
student2
```

```
## [1] 100 NA 90 90 90 90 97 80
```

```
is.na(student2)
```

```
## [1] FALSE TRUE FALSE FALSE FALSE FALSE FALSE
```

#Use is.na() to get NA values and convert to zero

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

```
## [1] 100 0 90 90 90 90 97 80
```

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

```
## [1] 90 0 0 0 0 0 0 0
```

```
#Use this logical to calculate now calculate grade
```

```
mean(student1)
```

```
## [1] 98.75
```

```
mean(student2)
```

```
## [1] 79.625
```

```
mean(student3)
```

```
## [1] 11.25
```

```
#Using min()/which.min() # min() finds lowest value # which.min() finds location of lowest value in vector
```

```
min(student1)
```

```
## [1] 90
```

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

```
## [1] 8
```

```
#Using this function to remove lowest grade
```

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

```
## [1] 100 100 100 100 100 100 100
```

```
student2[-which.min(student2)]
```

```
## [1] 100 90 90 90 90 97 80
```

```
student3[-which.min(student3)]
```

```
## [1] 90 0 0 0 0 0 0
```

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

```
## [1] 100
```

```
mean(student2[-which.min(student2)])
```

```
## [1] 91
```

```
mean(student3[-which.min(student3)])
```

```
## [1] 12.85714
```

Make function `grade()` to do this

All R functions have 3 things

Name(`grade`); Input Arguments(`scores`); body(`does work`)

```
grade <- function(x, na.rm = FALSE) {x[is.na(x)] <- 0  
  mean(x[-which.min(x)]) }  
grade(student1)
```

```
## [1] 100
```

```
grade(student2)
```

```
## [1] 91
```

```
grade(student3)
```

```
## [1] 12.85714
```

Now use `grade()` on data set for assignment

[Q1] Write a function `grade()` to determine an overall grade from a vector of student homework

assignment scores dropping the lowest single score. If a student misses a homework (i.e. has an NA value) this can be used as a score to be potentially dropped. Your final function should be adequately explained with code comments and be able to work on an example class gradebook such as this one in CSV format: “<https://tinyurl.com/gradeinput>”

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

#Now use apply() function to apply grade() to whole class

```
apply(gradebook, 1, grade)
```

```
## student-1 student-2 student-3 student-4 student-5 student-6 student-7
##      91.75      82.50      84.25      84.25      88.25      89.00      94.00
## student-8 student-9 student-10 student-11 student-12 student-13 student-14
##      93.75      87.75      79.00      86.00      91.75      92.25      87.75
## student-15 student-16 student-17 student-18 student-19 student-20
##      78.75      89.50      88.00      94.50      82.75      82.75
```

```
scores <- apply(gradebook, 1, grade)
```

[Q2] Who is the top scoring student overall in the gradebook?

```
which.max(scores)
```

```
## student-18
##           18
```

```
max(scores)
```

```
## [1] 94.5
```

Student 18 has the highest score with a score of 94.5

[Q3] Which homework was toughest on students?

```
# apply() new function to get mean/median HW scores
HWmean <- apply(gradebook, 2, mean, na.rm = TRUE)
HWmean
```

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

```
HWmedian <- apply(gradebook, 2, median, na.rm = TRUE)
HWmedian
```

```
## hw1 hw2 hw3 hw4 hw5
## 89.0 72.5 76.5 88.0 78.0
```

```
# Find minimum mean/median
which.min(HWmean)
```

```
## hw3
## 3
```

```
min(HWmean)
```

```
## [1] 80.8
```

```
which.min(HWmedian)
```

```
## hw2
## 2
```

```
min(HWmedian)
```

```
## [1] 72.5
```

HW2 was the toughest since it had the lowest median score of **72.5**, but **HW3** could be considered the toughest since it had the lowest mean score of **80.8**

[Q4] Which HW correlates most with the average student score?

```
gradebook2 <- gradebook
gradebook2[is.na(gradebook2)] <- 0
apply(gradebook2, 2, cor, scores)
```

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

HW5 has the highest correlation to the student's average performance.

[Q5]

Select “Knit to Save”