

# Week 5: R Functions Lab

Chantal Rabay

2/7/2022

Here we will write a function to grade some student homework.

Simple input example: vector of student scores

```
# 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)
```

Question 1. 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.

```
student1
```

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

The regular average will be returned by the ‘`mean()`’ function.

```
mean(student1)
```

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

Finding the position of the minimum value in a vector using ‘`which.min()`’

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

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

```
student1[8]
```

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

Returning the minimum value.

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

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

To get everything but the min value.

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

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

Take the mean.

```
#First solution
```

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

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

Changing NA values to zero. Using the 'is.na()' function.

```
student2
```

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

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

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

```
x <- student2
```

```
x
```

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

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

```
x
```

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

```
mean(x)
```

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

Find the average score for student 3

```
x <- student3
```

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

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

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

## Making the function

Using the working snippet from above I will make a function

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

Example using function:

```
grade(student1)
```

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

```
grade(student2)
```

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

```
grade(student3)
```

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

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>” [3pts]

```
## Calculate average scores for a vector of homework scores
## Dropping the lowest single score. Missing values will be treated as zero scores
## @param x Numeric vector of homework scores
##
## @return Average score
## @export
##
## @examples
##   student <- c(100, NA, 90, 80)
##   grade(student)
grade <- function(x) {
  # This Maps NA missing homework values to zero
  # Missing homeworks score zero
  x[is.na(x)] <- 0
  # We exclude the lowest score homework
  mean( x[ -which.min(x)])
}
```

###Grading the whole class using the gradebook:

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

```
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
```

Question 2. Using your grade() function and the supplied gradebook, Who is the top scoring student overall in the gradebook? [3pts]

```
results <- apply(gradebook,1,grade)
which.max(results)
```

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

Question 3. From your analysis of the gradebook, which homework was toughest on students (i.e. obtained the lowest scores overall? [2pts]

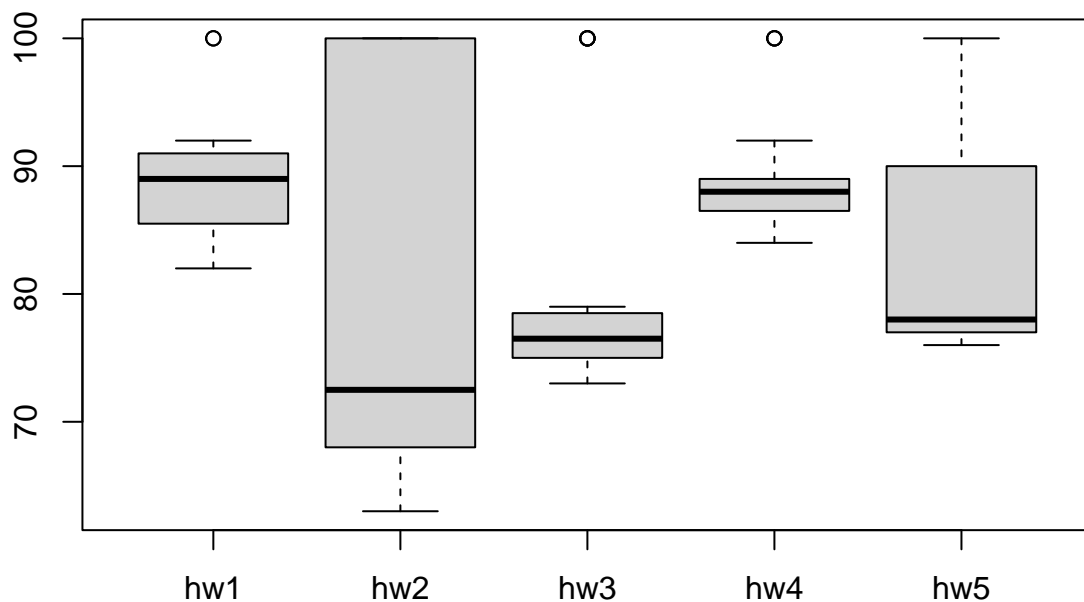
Here we want to calculate a summary stat for each column of the gradebook. Which stat should we use.

```
#Let's try average
hw.ave <- (apply(gradebook, 2, mean, na.rm=TRUE))
which.min(hw.ave)
```

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

Plot data to see and compare medians.

```
boxplot(gradebook)
```



> Question 3 Answer **Homework two is the most difficult question.**

Question 4. Optional Extension: From your analysis of the gradebook, which homework was most predictive of overall score (i.e. highest correlation with average grade score)? [1pt]

```
gradebook[is.na(gradebook)] <- 0
cor(results, gradebook$hw5)
```

```
## [1] 0.6325982
```

```
apply(gradebook, 2, cor, x=results)
```

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

HW 2 was least predictive, HW 5 was most predictive of overall score.

Question 5. Make sure you save your Rmarkdown document and can click the “Knit” button to generate a PDF format report without errors. Finally, submit your PDF to gradescope. [1pt]