

Class 6: R functions - Assignment

Minjun Kang (PID: A69042800)

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>” [3pts]

```
## Function
grade <- function(scores) {
  # 1. Omitting NA
  scores_no_na <- na.omit(scores)
  if (length(scores_no_na) < length(scores)) {
    # If there is NA
    final_scores <- scores_no_na
  } else {
    # If there is no NA
    final_scores <- scores_no_na[-which.min(scores_no_na)]
  }

  # Calculating mean
  mean(final_scores)
}
```

```
## Example
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)

grade(student1)
```

[1] 100

```
grade(student2)
```

```
[1] 91
```

```
grade(student3)
```

```
[1] 90
```

Description: If the student has NA, then we may calculate the mean for the scores which are not NA. IF the student has no NA, then we may omit the lowest score and then calculate the mean.

Q2

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

```
# Read csv file
grades <- read.csv("https://tinyurl.com/gradeinput")

# check it
head(grades)
```

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

```
# grade calculation
std_grades <- apply(grades[,-1], 1, grade)
std_best <- which.max(std_grades) # Find the index for the best student
print(grades$X[std_best]) # Print the best student
```

```
[1] "student-18"
```

Q3

From your analysis of the gradebook, which homework was toughest on students (i.e. obtained the lowest scores overall)?

```
# Mean values per each homework
hw_means <- colMeans(grades[,-1], na.rm = TRUE)

# toughest hw
hw_tough <- which.min(hw_means)

# print
names(hw_means)[hw_tough]
```

```
[1] "hw3"
```

Answer: HW3

Q4

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]

```
# Find correlations for each homework
hw_corr <- sapply(grades[,-1], function(hw) cor(hw, std_grades, use="complete.obs"))

# Return best homework for maximum correlation
best_predictor <- which.max(hw_corr)
print(names(hw_corr)[best_predictor])
```

```
[1] "hw2"
```

Answer: HW2