

Class06

Neva Olliffe (PID: A69026930)

Establishing sample data:

```
# 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)

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

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>”

```
grade <- function(student) {

  # Convert na scores to 0 in the gradebook
  na_vector <- is.na(student)
  student[na_vector] <- 0

  # Drop this student's lowest score, then average their remaining scores
  student <- student[-which.min(student)]
  return(mean(student))

}
```

Q2.

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

```
# Compute the mean scores of each student, then store it in a new column
mean_scores <- apply(gradebook_df, 1, grade)
gradebook_df["mean"] <- mean_scores

gradebook_df[which.max(mean_scores),]
```

```
      hw1 hw2 hw3 hw4 hw5 mean
student-18  91  NA 100  87 100 94.5
```

Student 18 is the highest scoring.

Q3.

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

```
# Replace all NAs with zeros, then average the scores for each hw
num_gradebook <- gradebook_df[is.na(gradebook_df)] <- 0
hw_avg <- apply(gradebook_df[, -6], 2, mean)

# Find the lowest scoring homework
which.min(hw_avg)
```

```
hw2
2
```

The lowest scoring homework is hw2.

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)?

```
# Compute the correlation of every hw grade with average grade
hw_cor <- cor(x = gradebook_df[, -6], y = gradebook_df[,6])

# Find the highest correlation
which.max(hw_cor)
```

```
[1] 5
```

Score on HW5 had the highest correlation with final average grade.