

# class06 submission

Jaewon Kim

2024-01-25

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>”

```
#Final answer
```

```
#Function grade()
grade <- function(x) {                                #take student vector x as input
  x[is.na(x)] <- 0                                     #part e)
  mean(x[-which.min(x)])                             #part d)
}
```

```
#Now read the online gradebook (csv file)
url <- "https://tinyurl.com/gradeinput"
gradebook <- read.csv(url, 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

```
#Apply function to gradebook
results <- apply(gradebook, 1, grade)
```

```
results
```

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	

WORK BELOW—————

```
#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)
```

```
student1 #Check student's entire grade
```

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

```
which.min(student1) # a) Print index of lowest grade of student's grade
```

```
[1] 8
```

```
student1[which.min(student1)] #b) Print score of corresponding index
```

```
[1] 90
```

```
student1[-which.min(student1)] #c) Remove lowest score from vector
```

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

```
mean(student1[-which.min(student1)]) #d) Average grade after dropping score
```

```
[1] 100
```

```
student2[is.na(student2)] <- 0 #e) Assign NA as zero. For instance, student 2  
student2
```

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

```
grade <- function(x) { #Call the function grade, x as input student vector  
  x[is.na(x)] <- 0 #part e)  
  mean(x[-which.min(x)]) #part d)  
}
```

```
grade(student1) #Test the function
```

```
[1] 100
```

```
grade(student2)
```

```
[1] 91
```

```
grade(student3)
```

```
[1] 12.85714
```

```
url <- "https://tinyurl.com/gradeinput" #read the online gradebook (csv file)  
gradebook <- read.csv(url, 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

```
results <- apply(gradebook, 1, grade) #Apply function to gradebook
results
```

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	

Q2. Using your grade() function and the supplied gradebook, Who is the top scoring student overall in the gradebook?

student 18 has highest overall grade.

```
which.max(results) #Find maximum in gradebook
```

```
student-18
18
```

Q3. From your analysis of the gradebook, which homework was toughest on students (i.e. obtained the lowest scores overall?)

Homework 2 was most difficult for students.

```
which.min(apply(gradebook, 2, sum, na.rm=T)) #Add scores of all students
```

```
hw2
2
```

```
difficulty <- apply(gradebook, 2, grade) #Evaluate difficulty by total scores
difficulty
```

hw1	hw2	hw3	hw4	hw5
89.36842	76.63158	81.21053	89.63158	83.42105

```
which.min(difficulty) #Find lowest total score
```

hw2  
2

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)?

Homework 5 showed highest correlation coefficient of 0.633, compared to average grade of each assignments. Therefore, homework 5 is most predictive of overall score.

```
masked <- gradebook
masked[is.na(masked)] <- 0 #make all NA to zero

cor(masked$hw1, results) #find correaltion between hw1 and other Homeworks.
```

```
[1] 0.4250204
```

```
cor(masked$hw2, results)
```

```
[1] 0.176778
```

```
cor(masked$hw3, results)
```

```
[1] 0.3042561
```

```
cor(masked$hw4, results)
```

```
[1] 0.3810884
```

```
cor(masked$hw5, results)
```

```
[1] 0.6325982
```

```
apply(masked, 2, cor, results) #apply cor using resultand masked's col.
```

	hw1	hw2	hw3	hw4	hw5
	0.4250204	0.1767780	0.3042561	0.3810884	0.6325982

Q5. Make sure you save your Quarto document and can click the “Render” (or Rmark-down”Knit”) button to generate a PDF format report without errors. Finally, submit your PDF to gradescope.

Yes.