

Week_5_R_function_lab_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>” [3pts]

Example input vectors to start with

```
# 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
```

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

```
mean(student1)
```

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

```
min(student1)
```

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

To find the position of the smallest value use the `which.min()`

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

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

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

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

to get everything except the min value

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

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

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

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

```
student2
```

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

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

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

```
mean(student2[-2])
```

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

```
student3
```

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

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

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

```
mean(student3)
```

```
## [1] NA
```

Change the NA values to zero use `is.na()` function to identify NA value in the vector

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

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

```
student2[is.na(student2)]
```

```
## [1] NA
```

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

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

```
mean(x)
```

```
## [1] NA
```

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

```
## [1] NA
```

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

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

```
which.min(x)
```

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

Making the function

```
## Title  
##  
## @param x  
##  
## @return  
## @export  
##  
## @examples  
grade <- function(x){  
  # map NA missing homework values to zero  
  # missing homework scores zero  
  x[is.na(x)] <- 0  
  # print(x)  
  # print(mean(x))  
  # excluding the lowest value score homework  
  print(mean(x[-which.min(x)]))  
}
```

```
grade(student3)
```

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

```
grade(student2)
```

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

```
grade(student1)
```

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

```
gradebook = read.csv(file = "C:/Users/User/Bioinformatics/BGGN-213/Week_5/R_week5/class06/student_homework.csv")
# data
# row.names(data)
```

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

```
## [1] 91.75
## [1] 82.5
## [1] 84.25
## [1] 84.25
## [1] 88.25
## [1] 89
## [1] 94
## [1] 93.75
## [1] 87.75
## [1] 79
## [1] 86
## [1] 91.75
## [1] 92.25
## [1] 87.75
## [1] 78.75
## [1] 89.5
## [1] 88
## [1] 94.5
## [1] 82.75
## [1] 82.75
```

```
## 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
```

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

```
## [1] 91.75
## [1] 82.5
## [1] 84.25
## [1] 84.25
## [1] 88.25
## [1] 89
## [1] 94
## [1] 93.75
## [1] 87.75
## [1] 79
## [1] 86
## [1] 91.75
## [1] 92.25
## [1] 87.75
## [1] 78.75
## [1] 89.5
```

```
## [1] 88
## [1] 94.5
## [1] 82.75
## [1] 82.75
```

```
sort(results)
```

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

```
which.max(results)
```

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

```
hw.ave<-apply(gradebook,2,mean, na.rm=TRUE)
which.min(hw.ave)
```

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

```
hw.ave
```

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

```
hw.med <- apply(gradebook, 2, median, na.rm=T)
which.min(hw.med)
```

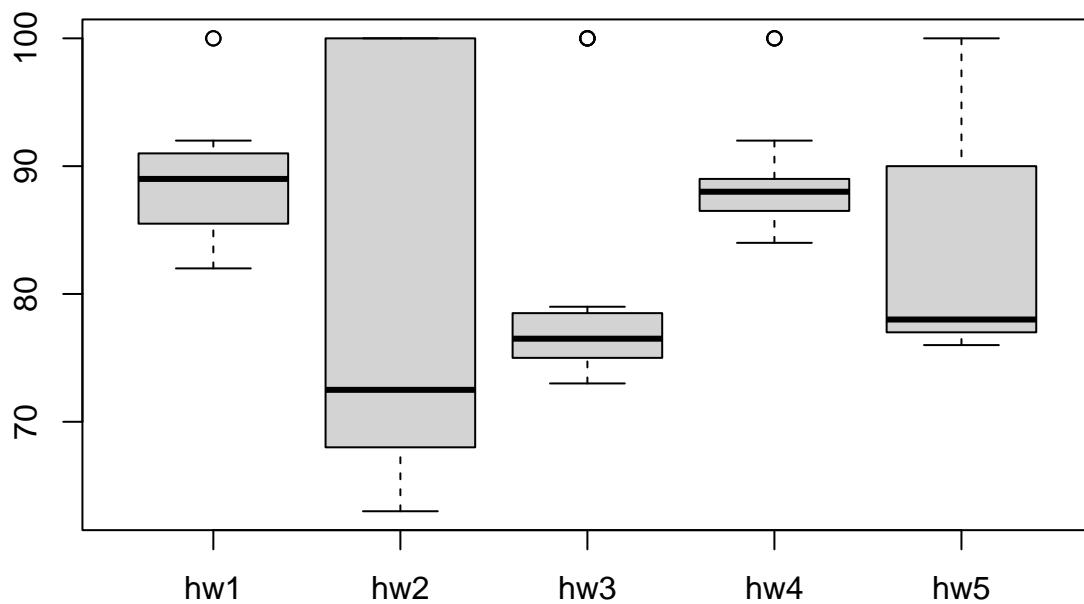
```
## hw2
##    2
```

```
hw.med
```

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

Plot the data using `boxplot()`

```
boxplot(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
```

```
cor(results,gradebook)
```

```
##           hw1 hw2           hw3 hw4 hw5
## [1,] 0.4250204 NA 0.3042561 NA NA
```

```
gradebook[is.na(gradebook)]<-0
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   0  73 100  76
## student-5  88 100  75  86  79
## student-6  89  78 100  89  77
```

```
## student-7 89 100 74 87 100
## student-8 89 100 76 86 100
## student-9 86 100 77 88 77
## student-10 89 72 79 0 76
## student-11 82 66 78 84 100
## student-12 100 70 75 92 100
## student-13 89 100 76 100 80
## student-14 85 100 77 89 76
## student-15 85 65 76 89 0
## student-16 92 100 74 89 77
## student-17 88 63 100 86 78
## student-18 91 0 100 87 100
## student-19 91 68 75 86 79
## student-20 91 68 76 88 76
```

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

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

Q2. Using your `grade()` function and the supplied gradebook, Who is the top scoring student overall in the gradebook? [**3pts**]

Q3. From your analysis of the gradebook, which homework was toughest on students (i.e. obtained the lowest scores overall)? [**2pts**]

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**]

Q5. Make sure you save your Rmarkdown document and can click the “**Knit**” button to generate a PDF for a report without errors. Finally, submit your PDF to gradescope. [**1pt**]