Class06

Jiawei Xu

```
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)
```

```
#for grading the student1, who does not have NA
which.min(student1)
```

```
[1] 8
```

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

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

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

```
[1] 100
```

```
# for students that have NA scores, NA needs to be removed first.
student2_temp <- student2
student2_temp[is.na(student2_temp)] <- 0   # assign 0 to NA
student2_temp
```

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

```
mean(student2_temp[-which.min(student2_temp)])
```

```
[1] 91
```

```r
student3_temp <- student3
student3_temp[is.na(student3_temp)] <- 0
student3_temp
```

```
[1] 90  0  0  0  0  0  0  0
```

```r
mean(student3_temp[-which.min(student3_temp)])
```

```
[1] 12.85714
```

```r
# write r function grade()
grade <- function(student_grade){
  student_grade_temp <- student_grade
  student_grade_temp <- as.numeric(student_grade_temp)
  student_grade_temp[is.na(student_grade_temp)] <- 0
  mean(student_grade_temp[-which.min(student_grade_temp)])
}
```

```r
# grade() fucntion test
grade(student1)
```

```
[1] 100
```

```r
grade(student2)
```

```
[1] 91
```

```r
grade(student3)
```

```
[1] 12.85714
```

```r
# input csv file
scores <-  read.csv("https://tinyurl.com/gradeinput", row.names = 1)
```

```
  result <- apply(scores, 1, grade)
  result
```

```
 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 Who is the top scoring student overall in the gradebook?
  which.max(result)
```

```
student-18
        18
```

```
  # student 18 is the top scoring student
```

```
  # Q3. From your analysis of the gradebook, which homework was toughest on students
  # (i.e. obtained the lowest scores overall?
  apply(scores, 2, mean, na.rm = T)
```

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

```
  which.min(apply(scores, 2, mean, na.rm = T))
```

```
hw3
  3
```

```
  # HW3 was toughest on students
```

```
  #Q4. From your analysis of the gradebook, which homework was most predictive of overall
  # score (i.e. highest correlation with average grade score)?
  mask_Q4 <- scores
  mask_Q4[is.na(mask_Q4)] <- 0
```

```r
cor(result, mask_Q4)
```

```
          hw1      hw2       hw3       hw4       hw5
[1,] 0.4250204 0.176778 0.3042561 0.3810884 0.6325982
```

```r
apply(mask_Q4, 2, cor, y = result)
```

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

```r
# HW5 was the most predictive
```