class06

Xinlong Wan

Q1.

Grade is a function to determine an overall grade from a vector of student homework assignment scores dropping the lowest single score.

It's unfair to drop all the NA submission, so we'll make all NA to zero.

```r
#' Calculate the mean grade from a vector of student homework assignment scores dropping t
#' Missing values will be treated as zero.
#' @param hw_grades A numeric vector of homework scores
#'
#' @return Average score
#' @export
#'
#' @examples
#'   student1 <- c(100, 100, 100, 100, 100, 100, 100, 90)
#'   grade(student1)
#'
grade <- function(hw_grades) {

  # make all the missing submission 0 by masking
  hw_grades[is.na(hw_grades)]<-0
  # drop the lowest numerical hw grade
  mean(hw_grades[-which.min(hw_grades)])

}
```

Use the function on the whole class. CSV format: "https://tinyurl.com/gradeinput"

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

apply(gradebook, 1, grade)
```

```
 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.

```r
results <- apply(gradebook, 1, grade)
# highest scoring student
which.max(results)
```

```
student-18
        18
```

```r
# or
sort(results, decreasing = TRUE)
```

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

Q3.

```r
# hardest homework (lowest score hw)

# by means
which.min(apply(gradebook, 2, mean, na.rm=TRUE))
```
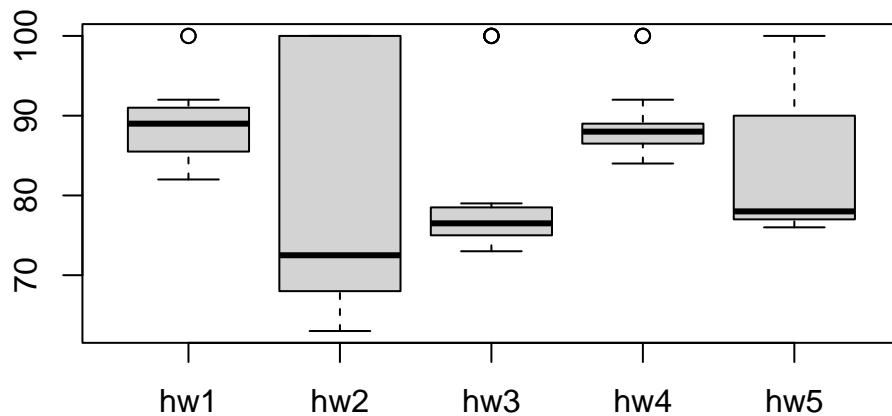
```
hw3
  3
```

```r
# by median
which.min(apply(gradebook, 2, median, na.rm=TRUE))
```

```
hw2
  2
```

Check the data distribution to determine which one is more appropriate

```r
boxplot(gradebook)
```


Looks like the hw2 is the hardest hw.

Q4.

```r
# mask the gradebook na
masked.gradebook <- gradebook

masked.gradebook[is.na(masked.gradebook)] <- 0

# calculate correlation

which.max(apply(masked.gradebook, 2, cor, x=results))
```

```
hw5
  5
```