lab6

Zijing

2022-10-15

```
#load homework data from gradebook
homework <- read.csv("student_homework.csv")
#Data from the gradebook
View(homework)
```

Question 1

```
grade <- function(hw){
  student_mean <- c()
  homework_mean <- c()
  for (i in 1:nrow(hw)){    #iterate through students
    student <- unlist(hw[i,2:ncol(hw)])
    if (any(is.na(student))){   #check if NA exists
      stumean <- mean(student,na.rm=TRUE)
    }
    else{
      stumean <- (sum(student)-min(student))/(ncol(hw)-1)
    }
    student_mean <- c(student_mean,stumean)     #adding overall score of the student to the list of stu
    }
  for(i in 2:ncol(hw)){    #iterate through homeworks
    hwn <- hw[,i]
    if (any(is.na(hwn))){      #check if NA exists
      hwmean <- mean(hwn,na.rm=TRUE)
    }
    else{
      hwmean <- (sum(hwn)-min(hwn))/(nrow(hw)-1)
    }
      homework_mean <- c(homework_mean,hwmean)    #adding overall score of the homework to the list of
  }
  names(student_mean) <- hw$X    #naming the student_mean vector
  names(homework_mean) <- colnames(hw)[2:ncol(hw)]    #naming the question_mean vector
  return(list(student_mean,homework_mean))
}
```

```
mean <- grade(homework)
student_mean <- mean[[1]]
homework_mean <- mean[[2]]
```

```
View(homework_mean)
View(student_mean)
```

Question 2

```
student_mean[which.max(student_mean)]
```

```
## student-18
##       94.5
```

The top scoring student in the grade book is student 18, with a mean score (lowest score excluded) of 94.5.

Question 3

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

```
##      hw2
## 80.88889
```

Homework 2 is the toughest, with a mean score (lowest score excluded), with a mean score of 80.9

Question 4

```
correlation <- function(student_mean,hw){
  cor_vector=c()
  for(i in 2:ncol(hw)){    #iterate through homeworks
      hwn <- hw[,i]
      cor <- cor.test(hwn,student_mean,"two.sided","pearson")   #pearson correlation between student ov
      cor_vector=c(cor_vector,cor)
  }
  max=colnames(hw)[which.max(unlist(cor_vector))+1]
  return(max)
}
```

```
correlation(unname(student_mean),homework)
```

```
## Warning in which.max(unlist(cor_vector)): NAs introduced by coercion
```

```
## [1] "hw2"
```

Homework 2 was most predictive of overall score.