CSE 3505
Foundation of Data Analytics

LAB-03

Name: **Preyash** Date: 26-Aug-2022

Registration Number: **20BPS1022**

Part A

Working with matrices in R

1. Represent the height in cm information of a team of 12 players as a matrix of dimension 4x3 in row major form.
2. Access the height at row 3 and column 2.
3. Display all the heights in row 2.
4. Display all the heights in column 3.
5. Extract the heights in all rows but only in column 1 and 3.
6. Find the transpose of the matrix.
7. Four more players got added to the team. Update the matrix to reflect the heights of the players.
8. Append three more players' height in the matrix.

Codes:

partA.R* | partB.R | partC.R | FDA_lab2c.R

Source on Save | Run

```r
#######----- Q1--------#########
team <- matrix(c(171,169,154,175,180,192,172,165,165,170,188,176),nrow=4, ncol=3, byrow=TRUE)
print(team)

#######----- Q2--------#########
print("Access the height at 3rd row and 2nd column:")
print(team[3,2])


#######----- Q3--------#########
print("All the heights in row 2:")
print(team[2,c(1:3)])
#or
print(team[2,])


#######----- Q4--------#########
print("All the heights in column 3:")
print(team[,3])


#######----- Q5--------#########
print("Height of all the row but only of col 1 and col 3:")
print(team[,c(1,3)])


#######----- Q6--------#########
print("Transpose :")
print(t(team))

S
#######----- Q7--------#########
newMembers <- c(180, 176, 165, 154)
team <- cbind(team, newMembers)
print(team)


#######----- Q8--------#########
newMembers2 <- c(150,161, 173, 177)
team <- rbind(team, newMembers2)
print(team)
```

Output:

```
> ######----- Q1--------#########
> team <- matrix(c(171,169,154,175,180,192,172,165,165,170,188,176),nrow=4, ncol=3, byrow=TRUE)
> print(team)
     [,1] [,2] [,3]
[1,]  171  169  154
[2,]  175  180  192
[3,]  172  165  165
[4,]  170  188  176
>
> ######----- Q2--------#########
> print("Access the height at 3rd row and 2nd column:")
[1] "Access the height at 3rd row and 2nd column:"
> print(team[3,2])
[1] 165
>
>
> ######----- Q3--------#########
> print("All the heights in row 2:")
[1] "All the heights in row 2:"
> print(team[2,c(1:3)])
[1] 175 180 192
> #or
> print(team[2,])
[1] 175 180 192
>
>
> ######----- Q4--------#########
> print("All the heights in column 3:")
[1] "All the heights in column 3:"
> print(team[,3])
[1] 154 192 165 176
>
>
> ######----- Q5--------#########
> print("Height of all the row but only of col 1 and col 3:")
[1] "Height of all the row but only of col 1 and col 3:"
> print(team[,c(1,3)])
     [,1] [,2]
[1,]  171  154
[2,]  175  192
[3,]  172  165
[4,]  170  176
>
```

```
> ######----- Q5--------#########
> print("Height of all the row but only of col 1 and col 3:")
[1] "Height of all the row but only of col 1 and col 3:"
> print(team[,c(1,3)])
     [,1] [,2]
[1,]  171  154
[2,]  175  192
[3,]  172  165
[4,]  170  176
>
>
> ######----- Q6--------#########
> print("Transpose :")
[1] "Transpose :"
> print(t(team))
     [,1] [,2] [,3] [,4]
[1,]  171  175  172  170
[2,]  169  180  165  188
[3,]  154  192  165  176
>
>
> ######----- Q7--------#########
> newMembers <- c(180, 176, 165, 154)
> team <- cbind(team, newMembers)
> print(team)
                  newMembers
[1,] 171 169 154        180
[2,] 175 180 192        176
[3,] 172 165 165        165
[4,] 170 188 176        154
>
>
> ######----- Q8--------#########
> newMembers2 <- c(150,161, 173, 177)
> team <- rbind(team, newMembers2)
> print(team)
                        newMembers
            171 169 154        180
            175 180 192        176
            172 165 165        165
            170 188 176        154
newMembers2 150 161 173        177
```

Part B

Creating data frames in R

A college has conducted technical events for the students. It maintains the name of the participant and the score obtained in different events.

1. Create a data frame by considering 5 students and 4 events. Each event has a maximum score of 10. If a student participates in an event, its entry contains the score value and 0 otherwise.
2. View the contents of the data frame.
3. Find the total score of each participant.
4. Append a column to include the total score of the participants and view the data frame.
5. Find the maximum score and display the name of the participant who scored it.
6. Compute the average score of each events and append it as a new row in the data frame.
7. Store the details in a comma separated values (csv) file. Also suppress the row numbers.

Codes:

partA.R | partB.R* | partC.R

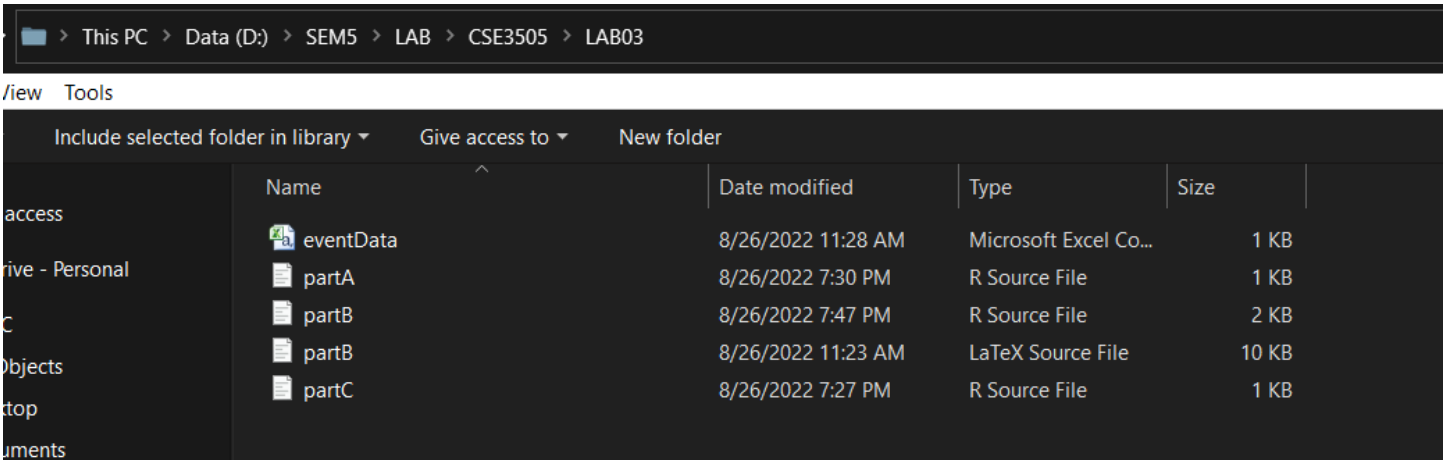
Source on Save | Run | Source

```r
#Q1
#event dataframe
event.data <-data.frame(
  student_id = c(1:5),
  student_name = c("Preyash", "Michelle", "Rose", "Nafeesa", "Edward"),
  technovit=c(10, 7 , 0 , 8 , 6),
  vibrance=c(10,3,4,0,4),
  riveria=c(8,10,3,5,0),
  DjSpidey =c(10, 6,7 ,8,9)
)

#Q2
#view contents
print(event.data)

#Q3
#total score of each participant
col_list <- c("technovit", "vibrance", "riveria", "DjSpidey")
totalScore = apply(event.data[,col_list], 1, sum)
print(totalScore)

#Q4
#appending a column to show total score
event.data <- cbind.data.frame(event.data, totalScore)
print(event.data)

#Q5
#getting the highest scorer
highestScore = which.max(event.data$totalScore)
highestScorer = event.data$student_name[highestScore]
print(paste("The highest scorer is:", highestScorer))

#Q6
avgEventScore <-c(0,0,colMeans(event.data[, c(3:7)]))
print(avgEventScore)
event.data <- rbind.data.frame(event.data, avgEventScore)
print(event.data)

#Q7
write.csv(event.data,"D:\\SEM5\\LAB\\CSE3505\\LAB03\\eventData.csv", row.names = TRUE)
```

Output:

```
> #Q1
> #event dataframe
> event.data <-data.frame(
+   student_id = c(1:5),
+   student_name = c("Preyash", "Michelle", "Rose", "Nafeesa", "Edward"),
+   technovit=c(10, 7 , 0 , 8 , 6),
+   vibrance=c(10,3,4,0,4),
+   riveria=c(8,10,3,5,0),
+   DjSpidey =c(10, 6,7 ,8,9)
+ )
>
> #Q2
> #view contents
> print(event.data)
  student_id student_name technovit vibrance riveria DjSpidey
1          1      Preyash        10       10       8       10
2          2     Michelle         7        3      10        6
3          3         Rose         0        4       3        7
4          4      Nafeesa         8        0       5        8
5          5       Edward         6        4       0        9
>
> #Q3
> #total score of each participant
> col_list <- c("technovit", "vibrance", "riveria", "DjSpidey")
> totalScore = apply(event.data[,col_list], 1, sum)
> print(totalScore)
[1] 38 26 14 21 19
>
> #Q4
> #appending a column to show total score
> event.data <- cbind.data.frame(event.data, totalScore)
> print(event.data)
  student_id student_name technovit vibrance riveria DjSpidey totalScore
1          1      Preyash        10       10       8       10         38
2          2     Michelle         7        3      10        6         26
3          3         Rose         0        4       3        7         14
4          4      Nafeesa         8        0       5        8         21
5          5       Edward         6        4       0        9         19
>
```

```
> #Q5
> #getting the highest scorer
> highestScore = which.max(event.data$totalScore)
> highestScorer = event.data$student_name[highestScore]
> print(paste("The highest scorer is:", highestScorer))
[1] "The highest scorer is: Preyash"
>
> #Q6
> avgEventScore <-c(0,0,colMeans(event.data[, c(3:7)]))
> print(avgEventScore)
                  technovit   vibrance    riveria   DjSpidey totalScore
       0.0        0.0        6.2        4.2        5.2        8.0       23.6
> event.data <- rbind.data.frame(event.data, avgEventScore)
> print(event.data)
  student_id student_name technovit vibrance riveria DjSpidey totalScore
1          1      Preyash      10.0     10.0     8.0       10       38.0
2          2     Michelle       7.0      3.0    10.0        6       26.0
3          3         Rose       0.0      4.0     3.0        7       14.0
4          4      Nafeesa       8.0      0.0     5.0        8       21.0
5          5       Edward       6.0      4.0     0.0        9       19.0
6          0            0       6.2      4.2     5.2        8       23.6
>
>
> #Q7
> write.csv(event.data,"D:\\SEM5\\LAB\\CSE3505\\LAB03\\eventData.csv", row.names = TRUE)
```



Part C

Indexing and Slicing data frames

1. Read the content of 'Events.csv' in a data frame and view it.
2. Access the scores of participants in event2 using the column name.
3. Use index number to retrieve the same data.
4. Extract the score of third participant in event3.
5. Extract the scores of the first and second participant in all the events.
6. Display the names and total scores of all participants.
7. Obtain the names whose total score is above its average.

Codes:

partA.R | partB.R | partC.R* | FDA_lab2c.R

Source on Save

```r
#Q1
getwd()
setwd("D:\\SEM5\\LAB\\CSE3505\\LAB03")
event_data <- read.csv("eventData.csv", header = TRUE, sep = ",")
event_data

#Q2
print (event_data$vibrance)
#here vibrance is the event2

#Q3
#Use index number to retrieve the same data
print(event_data[,4])

```

```r
#Q4
print(event_data$riveria[3])
#here riveria is the event3

#Q5
print(event_data[c(1,2),c(3,4,5,6,7)])

#Q6
print(event_data[c(1,2,3,4,5), c(3,8)])

#Q7
avgOfScore = mean(event_data$totalScore)
avgOfScore
 for (x in 1:5){
  newRecord[x] = (event_data$totalScore[x])
}
for(x in 1:5){
  if(newRecord[x]>avgOfScore)
    print(event_data$student_name[x])
}
```

Output:

```
> #Q1
> getwd()
[1] "D:/SEM5/LAB/CSE3505/LAB03"
> setwd("D:\\SEM5\\LAB\\CSE3505\\LAB03")
> event_data <- read.csv("eventData.csv", header = TRUE, sep = ",")
> event_data
  X student_id student_name technovit vibrance riveria DjSpidey totalScore
1 1          1      Preyash      10.0     10.0     8.0       10       38.0
2 2          2     Michelle       7.0      3.0    10.0        6       26.0
3 3          3         Rose       0.0      4.0     3.0        7       14.0
4 4          4      Nafeesa       8.0      0.0     5.0        8       21.0
5 5          5       Edward       6.0      4.0     0.0        9       19.0
6 6          0            0       6.2      4.2     5.2        8       23.6
>
> #Q2
> print (event_data$vibrance)
[1] 10.0  3.0  4.0  0.0  4.0  4.2
> #here vibrance is the event2
>
> #Q3
> #Use index number to retrieve the same data
> print(event_data[,4])
[1] 10.0  7.0  0.0  8.0  6.0  6.2
>
> #Q4
> print(event_data$riveria[3])
[1] 3
> #here riveria is the event3
>
> #Q5
> print(event_data[c(1,2),c(3,4,5,6,7)])
  student_name technovit vibrance riveria DjSpidey
1      Preyash        10       10       8       10
2     Michelle         7        3      10        6
>
```

```
> #Q6
> print(event_data[c(1,2,3,4,5), c(3,8)])
  student_name totalScore
1      Preyash         38
2     Michelle         26
3         Rose         14
4      Nafeesa         21
5       Edward         19
>
> #Q7
> avgOfScore = mean(event_data$totalScore)
> avgOfScore
[1] 23.6
>  for (x in 1:5){
+   newRecord[x] = (event_data$totalScore[x])
+ }
> for(x in 1:5){
+   if(newRecord[x]>avgOfScore)
+     print(event_data$student_name[x])
+ }
[1] "Preyash"
[1] "Michelle"
```