DATA607_LAB_1_Rajan

Krishna Rajan

2/24/2018

```
##Read the data table from text file store in GitHub

data <- read.table("https://raw.githubusercontent.com/RAJANCUNYSPSDATA607/DATA607/master/DATA%20607%20LA
head(data,5)
```

```
##                                                                                                 V1
## 1  -----------------------------------------------------------------------------------------------
## 2  Pair | Player Name                     |Total|Round|Round|Round|Round|Round|Round|Round|
## 3  Num  | USCF ID / Rtg (Pre->Post)       | Pts |  1  |  2  |  3  |  4  |  5  |  6  |  7  |
## 4  -----------------------------------------------------------------------------------------------
## 5     1 | GARY HUA                        |6.0  |W  39|W  21|W  18|W  14|W   7|D  12|D   4|
```

```
## remove dashes and seperates date using "|" as demarcation to split data into indiviual columns

delim <- read.delim("https://raw.githubusercontent.com/RAJANCUNYSPSDATA607/DATA607/master/DATA%20607%20I
head(delim,5)
```

```
##                                                                                                 V1
## 1                                                                                                1
## 2                                                                                               ON
## 3 -----------------------------------------------------------------------------------------------
## 4                                                                                                2
## 5                                                                                               MI
##                                 V2    V3    V4    V5    V6    V7    V8
## 1  GARY HUA                        6.0   W  39 W  21 W  18 W  14 W   7
## 2  15445895 / R: 1794   ->1817     N:2   W     B     W     B     W
## 3
## 4  DAKSHESH DARURI                 6.0   W  63 W  58 L   4 W  17 W  16
## 5  14598900 / R: 1553   ->1663     N:2   B     W     B     W     B
##       V9   V10 V11
## 1 D  12 D   4  NA
## 2 B     W      NA
## 3              NA
## 4 W  20 W   7  NA
## 5 W     B      NA
```

```
##now time to read data into a frame by extracting information like name, state, points, rating, & avera

chess <- data.frame(Name = delim$V2[c(TRUE, FALSE, FALSE)], State = delim$V1[c(FALSE, TRUE, FALSE)], Po
head(chess,5)
```

```
##                              Name  State Points Rating Opponents
## 1  GARY HUA                    ON     6.0      0         0
## 2  DAKSHESH DARURI             MI     6.0      0         0
## 3  ADITYA BAJAJ                MI     6.0      0         0
## 4  PATRICK H SCHILLING         MI     5.5      0         0
## 5  HANSHI ZUO                  MI     5.5      0         0
```

```
## Next get the pre rating for each player using the stringr function
library(stringr)
Rating <- str_extract(delim$V2[c(FALSE, TRUE, FALSE)], "R:[:blank:]+[:digit:]+")
chess$Rating <- as.numeric(str_extract(Rating, "[:digit:]+"))
head(chess,5)
```

```
##                                Name  State Points Rating Opponents
## 1  GARY HUA                          ON     6.0   1794         0
## 2  DAKSHESH DARURI                   MI     6.0   1553         0
## 3  ADITYA BAJAJ                      MI     6.0   1384         0
## 4  PATRICK H SCHILLING               MI     5.5   1716         0
## 5  HANSHI ZUO                        MI     5.5   1655         0
```

```
##create a rating table for each player vased on his oponents based on 7 rounds of play
Rate <- data.frame(playnum = str_trim(delim$V1[c(TRUE, FALSE, FALSE)]), rate = chess$Rating, R1 = as.nu
head(Rate,5)
```

```
##   playnum rate R1 R2 R3 R4 R5 R6 R7 rateAvg
## 1       1 1794 39 21 18 14  7 12  4       0
## 2       2 1553 63 58  4 17 16 20  7       0
## 3       3 1384  8 61 25 21 11 13 12       0
## 4       4 1716 23 28  2 26  5 19  1       0
## 5       5 1655 45 37 12 13  4 14 17       0
```

```
##Replace player number with pre ranking number in each round for all 64 players.
for (i in 1:64) {
        Rate$R1[i] <- Rate$rate[Rate$R1[i]]
        Rate$R2[i] <- Rate$rate[Rate$R2[i]]
        Rate$R3[i] <- Rate$rate[Rate$R3[i]]
        Rate$R4[i] <- Rate$rate[Rate$R4[i]]
        Rate$R5[i] <- Rate$rate[Rate$R5[i]]
        Rate$R6[i] <- Rate$rate[Rate$R6[i]]
        Rate$R7[i] <- Rate$rate[Rate$R7[i]]
}
head(Rate,5)
```

```
##   playnum rate   R1   R2   R3   R4   R5   R6   R7 rateAvg
## 1       1 1794 1436 1563 1600 1610 1649 1663 1716       0
## 2       2 1553 1175  917 1716 1629 1604 1595 1649       0
## 3       3 1384 1641  955 1745 1563 1712 1666 1663       0
## 4       4 1716 1363 1507 1553 1579 1655 1564 1794       0
## 5       5 1655 1242  980 1663 1666 1716 1610 1629       0
```

```
## Calculate average ranking using for loop and using function "rowMeans"
for (i in 1:64) {
        Rate$rateAvg[i] <- rowMeans(Rate[i, 3:9], na.rm = TRUE)
}
head(Rate,5)
```

```
##   playnum rate   R1   R2   R3   R4   R5   R6   R7  rateAvg
## 1       1 1794 1436 1563 1600 1610 1649 1663 1716 1605.286
## 2       2 1553 1175  917 1716 1629 1604 1595 1649 1469.286
## 3       3 1384 1641  955 1745 1563 1712 1666 1663 1563.571
## 4       4 1716 1363 1507 1553 1579 1655 1564 1794 1573.571
## 5       5 1655 1242  980 1663 1666 1716 1610 1629 1500.857
```

```r
## Put the final average ranking value in a data frame
chess$Opponents <- round(Rate$rateAvg, digits = 0)
chess
```

```
##                                 Name  State Points Rating Opponents
## 1   GARY HUA                          ON     6.0   1794      1605
## 2   DAKSHESH DARURI                   MI     6.0   1553      1469
## 3   ADITYA BAJAJ                      MI     6.0   1384      1564
## 4   PATRICK H SCHILLING               MI     5.5   1716      1574
## 5   HANSHI ZUO                        MI     5.5   1655      1501
## 6   HANSEN SONG                       OH     5.0   1686      1519
## 7   GARY DEE SWATHELL                 MI     5.0   1649      1372
## 8   EZEKIEL HOUGHTON                  MI     5.0   1641      1468
## 9   STEFANO LEE                       ON     5.0   1411      1523
## 10  ANVIT RAO                         MI     5.0   1365      1554
## 11  CAMERON WILLIAM MC LEMAN          MI     4.5   1712      1468
## 12  KENNETH J TACK                    MI     4.5   1663      1506
## 13  TORRANCE HENRY JR                 MI     4.5   1666      1498
## 14  BRADLEY SHAW                      MI     4.5   1610      1515
## 15  ZACHARY JAMES HOUGHTON            MI     4.5   1220      1484
## 16  MIKE NIKITIN                      MI     4.0   1604      1386
## 17  RONALD GRZEGORCZYK                MI     4.0   1629      1499
## 18  DAVID SUNDEEN                     MI     4.0   1600      1480
## 19  DIPANKAR ROY                      MI     4.0   1564      1426
## 20  JASON ZHENG                       MI     4.0   1595      1411
## 21  DINH DANG BUI                     ON     4.0   1563      1470
## 22  EUGENE L MCCLURE                  MI     4.0   1555      1300
## 23  ALAN BUI                          ON     4.0   1363      1214
## 24  MICHAEL R ALDRICH                 MI     4.0   1229      1357
## 25  LOREN SCHWIEBERT                  MI     3.5   1745      1363
## 26  MAX ZHU                           ON     3.5   1579      1507
## 27  GAURAV GIDWANI                    MI     3.5   1552      1222
## 28  SOFIA ADINA STANESCU-BELLU        MI     3.5   1507      1522
## 29  CHIEDOZIE OKORIE                  MI     3.5   1602      1314
## 30  GEORGE AVERY JONES                ON     3.5   1522      1144
## 31  RISHI SHETTY                      MI     3.5   1494      1260
## 32  JOSHUA PHILIP MATHEWS             ON     3.5   1441      1379
## 33  JADE GE                           MI     3.5   1449      1277
## 34  MICHAEL JEFFERY THOMAS            MI     3.5   1399      1375
## 35  JOSHUA DAVID LEE                  MI     3.5   1438      1150
## 36  SIDDHARTH JHA                     MI     3.5   1355      1388
## 37  AMIYATOSH PWNANANDAM              MI     3.5    980      1385
## 38  BRIAN LIU                         MI     3.0   1423      1539
## 39  JOEL R HENDON                     MI     3.0   1436      1430
## 40  FOREST ZHANG                      MI     3.0   1348      1391
## 41  KYLE WILLIAM MURPHY               MI     3.0   1403      1248
## 42  JARED GE                          MI     3.0   1332      1150
## 43  ROBERT GLEN VASEY                 MI     3.0   1283      1107
## 44  JUSTIN D SCHILLING                MI     3.0   1199      1327
## 45  DEREK YAN                         MI     3.0   1242      1152
## 46  JACOB ALEXANDER LAVALLEY          MI     3.0    377      1358
## 47  ERIC WRIGHT                       MI     2.5   1362      1392
## 48  DANIEL KHAIN                      MI     2.5   1382      1356
## 49  MICHAEL J MARTIN                  MI     2.5   1291      1286
```

```
## 50  SHIVAM JHA                          MI     2.5   1056      1296
## 51  TEJAS AYYAGARI                      MI     2.5   1011      1356
## 52  ETHAN GUO                           MI     2.5    935      1495
## 53  JOSE C YBARRA                       MI     2.0   1393      1345
## 54  LARRY HODGE                         MI     2.0   1270      1206
## 55  ALEX KONG                           MI     2.0   1186      1406
## 56  MARISA RICCI                        MI     2.0   1153      1414
## 57  MICHAEL LU                          MI     2.0   1092      1363
## 58  VIRAJ MOHILE                        MI     2.0    917      1391
## 59  SEAN M MC CORMICK                   MI     2.0    853      1319
## 60  JULIA SHEN                          MI     1.5    967      1330
## 61  JEZZEL FARKAS                       ON     1.5    955      1327
## 62  ASHWIN BALAJI                       MI     1.0   1530      1186
## 63  THOMAS JOSEPH HOSMER                MI     1.0   1175      1350
## 64  BEN LI                              MI     1.0   1163      1263
```

```
## Write the output file to .CSV
write.csv(chess,file="/Users/rajans/Desktop/CUNY/Data Acquition & Management/DATA-607-Project-1/project
```