Project1

Benson

2/27/2022

Data Import

using read.csv to read the data from txt file.

```
chess_DS <- read.csv(file="https://raw.githubusercontent.com/Benson90/project1/main/project1.txt",head =
#test reading file
chess_DS[1:9,]
```

```
## [1] " Pair | Player Name                     |Total|Round|Round|Round|Round|Round|Round|Round| "
## [2] " Num  | USCF ID / Rtg (Pre->Post)       | Pts |  1  |  2  |  3  |  4  |  5  |  6  |  7  | "
## [3] "-----------------------------------------------------------------------------------------"
## [4] "    1 | GARY HUA                        |6.0  |W  39|W  21|W  18|W  14|W   7|D  12|D   4|"
## [5] "   ON | 15445895 / R: 1794   ->1817     |N:2  |W    |B    |W    |B    |W    |B    |W    |"
## [6] "-----------------------------------------------------------------------------------------"
## [7] "    2 | DAKSHESH DARURI                 |6.0  |W  63|W  58|L   4|W  17|W  16|W  20|W   7|"
## [8] "   MI | 14598900 / R: 1553   ->1663     |N:2  |B    |W    |B    |W    |B    |W    |B    |"
## [9] "-----------------------------------------------------------------------------------------"
```

Retrive Information

- Seprate the 1st row data and 2nd row data into 2 dataframe.
- Transform the data into attributes.
- Create a dataframe with Player's Name, Player's State, Total Number of Points, Player's Pre-Rating, and Average Pre Chess Rating of Opponents

```
chess_DS <- chess_DS[-c(1:3),]

firstrowdata <- chess_DS[seq(1, length(chess_DS),3)]

#test name
firstrowdata[1:6]
```

```
## [1] "    1 | GARY HUA                       |6.0  |W  39|W  21|W  18|W  14|W   7|D  12|D   4|"
## [2] "    2 | DAKSHESH DARURI                |6.0  |W  63|W  58|L   4|W  17|W  16|W  20|W   7|"
## [3] "    3 | ADITYA BAJAJ                   |6.0  |L   8|W  61|W  25|W  21|W  11|W  13|W  12|"
## [4] "    4 | PATRICK H SCHILLING            |5.5  |W  23|D  28|W   2|W  26|D   5|W  19|D   1|"
## [5] "    5 | HANSHI ZUO                     |5.5  |W  45|W  37|D  12|D  13|D   4|W  14|W  17|"
## [6] "    6 | HANSEN SONG                    |5.0  |W  34|D  29|L  11|W  35|D  10|W  27|W  21|"
```

```
secondrowdata <- chess_DS[seq(2, length(chess_DS),3)]

#test status
secondrowdata[1:6]
```

```
## [1] "   ON | 15445895 / R: 1794   ->1817     |N:2  |W    |B    |W    |B    |W    |B    |W    |"
## [2] "   MI | 14598900 / R: 1553   ->1663     |N:2  |B    |W    |B    |W    |B    |W    |B    |"
## [3] "   MI | 14959604 / R: 1384   ->1640     |N:2  |W    |B    |W    |B    |W    |B    |W    |"
## [4] "   MI | 12616049 / R: 1716   ->1744     |N:2  |W    |B    |W    |B    |W    |B    |B    |"
## [5] "   MI | 14601533 / R: 1655   ->1690     |N:2  |B    |W    |B    |W    |B    |W    |B    |"
## [6] "   OH | 15055204 / R: 1686   ->1687     |N:3  |W    |B    |W    |B    |B    |W    |B    |"
```

```
#attributes for new dataframe
Players_Name <- str_trim(str_extract(firstrowdata,"(\\w+\\s){2,3}"))
Players_State <- str_extract(secondrowdata,"\\w+")
Total_Number_of_Points <- as.numeric(str_extract(firstrowdata, "\\d+\\.\\d+"))
Players_preRating <- as.integer(str_extract(str_extract(secondrowdata, "[^\\d]\\d{3,4}[^\\d]"),"\\d+"))


#calculate average rating
Pair <- as.integer(str_extract(firstrowdata, "\\d+"))
Opponent <- str_extract_all(str_extract_all(firstrowdata, "\\d+\\|"),"\\d+")
```

```
## Warning in stri_extract_all_regex(string, pattern, simplify = simplify, :
## argument is not an atomic vector; coercing
```

```
Average_preChess_Rating <- length(firstrowdata)
for(i in 1:length(firstrowdata)){
  Average_preChess_Rating[i] <- round(mean(Players_preRating[as.numeric(unlist(Opponent[Pair[i]]))]),di
}

#new dataframe

NewDF <- data.frame(Players_Name,Players_State,Total_Number_of_Points,Players_preRating,Average_preChess
NewDF
```

```
##                Players_Name Players_State Total_Number_of_Points
## 1                  GARY HUA            ON                     6.0
## 2            DAKSHESH DARURI            MI                     6.0
## 3               ADITYA BAJAJ            MI                     6.0
## 4        PATRICK H SCHILLING            MI                     5.5
## 5                 HANSHI ZUO            MI                     5.5
## 6                HANSEN SONG            OH                     5.0
## 7          GARY DEE SWATHELL            MI                     5.0
## 8           EZEKIEL HOUGHTON            MI                     5.0
## 9                STEFANO LEE            ON                     5.0
## 10                 ANVIT RAO            MI                     5.0
## 11       CAMERON WILLIAM MC             MI                     4.5
## 12            KENNETH J TACK            MI                     4.5
## 13         TORRANCE HENRY JR            MI                     4.5
## 14              BRADLEY SHAW            MI                     4.5
## 15   ZACHARY JAMES HOUGHTON             MI                     4.5
```

```
## 16             MIKE NIKITIN           MI                 4.0
## 17       RONALD GRZEGORCZYK           MI                 4.0
## 18            DAVID SUNDEEN           MI                 4.0
## 19             DIPANKAR ROY           MI                 4.0
## 20              JASON ZHENG           MI                 4.0
## 21            DINH DANG BUI           ON                 4.0
## 22         EUGENE L MCCLURE           MI                 4.0
## 23                 ALAN BUI           ON                 4.0
## 24        MICHAEL R ALDRICH           MI                 4.0
## 25         LOREN SCHWIEBERT           MI                 3.5
## 26                  MAX ZHU           ON                 3.5
## 27           GAURAV GIDWANI           MI                 3.5
## 28              SOFIA ADINA           MI                 3.5
## 29         CHIEDOZIE OKORIE           MI                 3.5
## 30       GEORGE AVERY JONES           ON                 3.5
## 31             RISHI SHETTY           MI                 3.5
## 32    JOSHUA PHILIP MATHEWS           ON                 3.5
## 33                 JADE GE            MI                 3.5
## 34   MICHAEL JEFFERY THOMAS           MI                 3.5
## 35         JOSHUA DAVID LEE           MI                 3.5
## 36            SIDDHARTH JHA           MI                 3.5
## 37     AMIYATOSH PWNANANDAM           MI                 3.5
## 38                BRIAN LIU           MI                 3.0
## 39            JOEL R HENDON           MI                 3.0
## 40             FOREST ZHANG           MI                 3.0
## 41      KYLE WILLIAM MURPHY           MI                 3.0
## 42                 JARED GE           MI                 3.0
## 43        ROBERT GLEN VASEY           MI                 3.0
## 44       JUSTIN D SCHILLING           MI                 3.0
## 45                DEREK YAN           MI                 3.0
## 46 JACOB ALEXANDER LAVALLEY           MI                 3.0
## 47              ERIC WRIGHT           MI                 2.5
## 48             DANIEL KHAIN           MI                 2.5
## 49         MICHAEL J MARTIN           MI                 2.5
## 50               SHIVAM JHA           MI                 2.5
## 51           TEJAS AYYAGARI           MI                 2.5
## 52                ETHAN GUO           MI                 2.5
## 53            JOSE C YBARRA           MI                 2.0
## 54              LARRY HODGE           MI                 2.0
## 55                ALEX KONG           MI                 2.0
## 56             MARISA RICCI           MI                 2.0
## 57               MICHAEL LU           MI                 2.0
## 58             VIRAJ MOHILE           MI                 2.0
## 59                SEAN M MC           MI                 2.0
## 60              JULIA SHEN            MI                 1.5
## 61            JEZZEL FARKAS           ON                 1.5
## 62            ASHWIN BALAJI           MI                 1.0
## 63     THOMAS JOSEPH HOSMER           MI                 1.0
## 64                  BEN LI            MI                 1.0
##    Players_preRating Average_preChess_Rating
## 1               1794                    1605
## 2               1553                    1469
## 3               1384                    1564
## 4               1716                    1574
```

```
## 5              1655                1501
## 6              1686                1519
## 7              1649                1372
## 8              1641                1468
## 9              1411                1523
## 10             1365                1554
## 11             1712                1468
## 12             1663                1506
## 13             1666                1498
## 14             1610                1515
## 15             1220                1484
## 16             1604                1386
## 17             1629                1499
## 18             1600                1480
## 19             1564                1426
## 20             1595                1411
## 21             1563                1470
## 22             1555                1300
## 23             1363                1214
## 24             1229                1357
## 25             1745                1363
## 26             1579                1507
## 27             1552                1222
## 28             1507                1522
## 29             1602                1314
## 30             1522                1144
## 31             1494                1260
## 32             1441                1379
## 33             1449                1277
## 34             1399                1375
## 35             1438                1150
## 36             1355                1388
## 37              980                1385
## 38             1423                1539
## 39             1436                1430
## 40             1348                1391
## 41             1403                1248
## 42             1332                1150
## 43             1283                1107
## 44             1199                1327
## 45             1242                1152
## 46              377                1358
## 47             1362                1392
## 48             1382                1356
## 49             1291                1286
## 50             1056                1296
## 51             1011                1356
## 52              935                1495
## 53             1393                1345
## 54             1270                1206
## 55             1186                1406
## 56             1153                1414
## 57             1092                1363
## 58              917                1391
```

```
## 59              853                 1319
## 60              967                 1330
## 61              955                 1327
## 62             1530                 1186
## 63             1175                 1350
## 64             1163                 1263
```