

Data-607 Week-4 Project1

Student Name : Sachid Deshmukh

Date : 09/23/2018

- GitHub Location for rmd file
 - GitHub Location for pdf file
 - RPub's location of published file
 - Output file location file
-

1] Library Initialization

```
library(stringr)
```

```
## Warning: package 'stringr' was built under R version 3.4.3
```

2] Read Data

```
col.names <- c("PairNum", "PalyerName", "TotPts", "Round1", "Round2", "Round3", "Round4", "Round5", "Round6", "Round7", "Round8", "Round9", "Round10", "Round11", "Round12", "Round13", "Round14", "Round15", "Round16", "Round17", "Round18", "Round19", "Round20", "Round21", "Round22", "Round23", "Round24", "Round25", "Round26", "Round27", "Round28", "Round29", "Round30", "Round31", "Round32", "Round33", "Round34", "Round35", "Round36", "Round37", "Round38", "Round39", "Round40", "Round41", "Round42", "Round43", "Round44", "Round45", "Round46", "Round47", "Round48", "Round49", "Round50", "Round51", "Round52", "Round53", "Round54", "Round55", "Round56", "Round57", "Round58", "Round59", "Round60", "Round61", "Round62", "Round63", "Round64", "Round65", "Round66", "Round67", "Round68", "Round69", "Round70", "Round71", "Round72", "Round73", "Round74", "Round75", "Round76", "Round77", "Round78", "Round79", "Round80", "Round81", "Round82", "Round83", "Round84", "Round85", "Round86", "Round87", "Round88", "Round89", "Round90", "Round91", "Round92", "Round93", "Round94", "Round95", "Round96", "Round97", "Round98", "Round99", "Round100")
player.data <- read.table("https://github.com/mlforsachid/MSDSQ1/blob/master/Data607/Week4/tournamentin")
```

```
## Warning in scan(file = file, what = what, sep = sep, quote = quote, dec =  
## dec, : EOF within quoted string
```

3] Remove white spaces

```
player.data <- subset(player.data, !PalyerName == "")  
player.data = as.data.frame(sapply(player.data, str_trim))
```

4] Add State and players pre rating columns

```
player.tmp = player.data[seq(2, nrow(player.data), 2), ]  
player.data = player.data[seq(1, nrow(player.data), 2), ]  
player.data$PlayerState = player.tmp$PairNum  
player.data$PlayerPreRanking = str_trim(str_extract(player.tmp$PalyerName, "[[:blank:]]{1}[[:digit:]]{3}"))
```

5] Transform Opponent players

```

TransformPoints = function(x)
{
  str_trim(str_extract(x, "[[:space:]]+[[:digit:]]{1,2}"))
}
player.data[4:10] = sapply(player.data[4:10], TransformPoints)

```

6] Calculate opponent players pre-rating

```

for (i in 1:nrow(player.data)) {
  for(j in 4:10) {
    player.data[i,j] <- player.data[player.data$PairNum == player.data[i,j],12][1]
  }
}

```

7] Calculate players avg rating

```

player.data[4:10] = sapply(player.data[4:10], function(x) as.numeric(x))
player.data$AvgRating <- round(rowMeans(player.data[, c(4:10)]), na.rm = TRUE), 1)

```

8] Create final output

```

player.data.out = subset(player.data, TRUE, c(PalyerName, PlayerState, TotPts, PlayerPreRanking, AvgRating))
write.csv(player.data.out, file = "output.csv", row.names = FALSE)

```