Project 1

Ka Man Chan

September 20, 2016

## Introduction:

### The goal of this project is extracting the info from the tournamentinfo text file, doing the some mapping with Pair Numbers and calculating the Average of Pre Chess Rating of Opponents based on the Opponents' Pair Numbers in each round.

library(stringr)

### Read the txt file through File Dialog Box

result = tryCatch({ TEXTFILE = file.choose() }, warning = function(war) { return("") }, error = function(err) { return("") }, finally = { })

if(result!=""){ conn = file(TEXTFILE,open="r") line = readLines(conn) MyPNum=""

### Create Table 1 to store Player's Pair Number, Name and Point

### Create Table 2 to store Player's Pair Number, State and PreRate

### Create Table 3 to store Player's Pair Number, Opponents' Pair Numbers

table1 =data.frame("PairNum" = character(), "Name"=character(), "Point"=character(), stringsAsFactors=FALSE) table2 =data.frame("PairNum" = character(), "State"=character(), "PreRate"=character(), stringsAsFactors=FALSE) table3 =data.frame("OriginalPairNum" = character(), "PairNum"=character(), stringsAsFactors=FALSE)

for (i in 5:length(line)){

## Skip the row with long dashes  
   
 if(line[i] != "-----------------------------------------------------------------------------------------"){  
   
 ## Split the columns with | as delimiters  
   
 MyColumn=unlist(strsplit(line[i], "[|]"))  
   
 if(suppressWarnings(!is.na(as.numeric(str\_trim(MyColumn[1]))))){  
   
 MyPNum=str\_trim(MyColumn[1])  
   
 #Round 1-7  
   
 for(j in 4:10){  
 for(p in str\_extract\_all(MyColumn[j],"[[:digit:]]\*")) {  
 for(i in p){  
 if(i!=""){  
 ##Insert Play's PairNumber and Opponents' Pair Numbers to table 3  
 table3[nrow(table3) + 1, ] = c(str\_trim(MyColumn[1]), i)  
 }  
 }  
 }  
 }  
 ## Insert a record with Pair Number, Name and Point to Table 1  
 table1[nrow(table1) + 1, ] = c(str\_trim(MyColumn[1]), str\_trim(MyColumn[2]),str\_trim(MyColumn[3]))  
 }  
 else{  
   
 ##Search for Starting position of PreRate  
 y = str\_locate(pattern ='R:', MyColumn[2])  
   
 ##Insert a record with Pair Number, State and PreRate to Table 2  
 table2[nrow(table2) + 1, ] = c(MyPNum, str\_trim(MyColumn[1]),str\_trim(str\_sub(MyColumn[2],y[1]+2,y[1]+6)))  
 MyPNum=""  
 }  
 }

}

### Close the file

close(conn)

### merge data frames 2 and 3 by PairNum

table4=merge(table2,table3,by="PairNum")

library(plyr) ### Calculate the Average of Pre Chess Rating of Opponents table5=ddply(table4, .(OriginalPairNum), summarize, Average=round(mean(as.integer(PreRate)),0)) ### Rename the key table5=rename(table5, c("OriginalPairNum"="PairNum"))

### merge three data frames Table 1,2 and 5 by PairNum

table6=Reduce(function(x, y) merge(x, y, all=TRUE), list(table1, table2, table5))

### Transform Char to Numeric fields (Pair Number, PreRate and Point)

table6=transform(table6, PairNum = as.numeric(PairNum)) table6=transform(table6, PreRate = as.numeric(PreRate)) table6=transform(table6, Point = as.numeric(Point))

### Order the records by Pair Number accendingly

table6=table6[with(table6, order(PairNum)), ]

### Create a table with the required fields (Name, State, Point, PreRate and Average of Pre Chess Rating of Opponents)

table7 = table6[c("Name", "State", "Point","PreRate","Average")]

### Print the output table

print(table7,row.names = FALSE)

### Export the output table as tournamentinfo.csv file

TEXTFILE = str\_replace(TEXTFILE, ".txt",".csv") print(paste("You can find your output file in ",TEXTFILE ))

write.table(table7, file = TEXTFILE, append = FALSE, quote = FALSE, sep = ',', row.names = FALSE,col.names = FALSE)

## Summary:

### It is an interesting project for me to practice extracting and transforming data through R. I could save some times if I used regular expression proficiently. Because of the time contraint, I used a dummy way to get the PreRate. Transforms are not needed if I configured the columns as numeric in temp tables at the beginning but it did not work even I declared the fields as integer and inserted the records with as.numeric column.

}