

chess

Yadu

February 13, 2016

Download the file first and then read the txt file into R.

```
library(RCurl)
```

```
## Loading required package: bitops
```

```
library('stringr')
url = "https://raw.githubusercontent.com/cyadusha/tournamentinfo/master/tournamentinfo.txt"
x = getURL(url)
tournamentinfo = read.delim(file = textConnection(x), header = TRUE)
```

```
nrow(tournamentinfo)
```

```
## [1] 195
```

We know that we desire a sequence that has a length of 64 values. The txt file that is inputted into R has 195 rows. This does not imply that there are 195 players. Therefore, we input the following command in order for us to know which rows should be subsetted from the data.

```
subsetrows = c(seq(from = 4, to = 193, by = 3))
subsetrows
```

```
## [1] 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52
## [18] 55 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100 103
## [35] 106 109 112 115 118 121 124 127 130 133 136 139 142 145 148 151 154
## [52] 157 160 163 166 169 172 175 178 181 184 187 190 193
```

This is where we obtain the player names, total points, and the opponents' identifications.

```
t1 = tournamentinfo[c(subsetrows[1:64]), ]
names = str_trim(unlist(str_extract(t1, "[[:alpha:]]+ ?[[:alpha:]]+ [[:alpha:]]+")))
totalpoints = str_trim(unlist(str_extract(t1, "[[:digit:]]+[[:digit:]]")))
location = str_locate(t1[1], "39|W")
opponents = str_sub(t1, location[1])
opponents = unlist(str_replace_all(opponents, "[[:alpha:]]+", ""))
opponents = unlist(str_replace_all(opponents, "[[:space:]]+", ""))
opponents = str_sub(opponents, 1, -2)
opponents = unlist(str_split(opponents, "[|]"))
opponents = c(as.numeric(unlist(str_extract(opponents, "[[:digit:]]+"))))
```

The rows where the state names and the preratings are stored are one row below those of the player's names.

```
subsetrows2 = subsetrows + 1
subsetrows2
```

```
## [1] 5 8 11 14 17 20 23 26 29 32 35 38 41 44 47 50 53
## [18] 56 59 62 65 68 71 74 77 80 83 86 89 92 95 98 101 104
## [35] 107 110 113 116 119 122 125 128 131 134 137 140 143 146 149 152 155
## [52] 158 161 164 167 170 173 176 179 182 185 188 191 194
```

This is where we obtain the state names and the preratings.

```
t2 = tournamentinfo[c(subsetrows2[1:64]), ]
states = str_trim(unlist(str_extract(t2, "[:alpha:]+[:alpha:]")))
prerating = str_trim(unlist(str_extract(t2, "R: ?.[[:digit:]]{3}|R: ?.[[:digit:]]{4}")))
prerating = str_sub(prerating, start = 4, end = 7)
prerating = as.numeric(prerating)
prerating = data.frame(prerating)
```

We have to set up the identification numbers of the opponents as numeric values.

```
opponents = data.frame(matrix(as.numeric(opponents), nrow = 64, ncol = 7, byrow = T))
```

Next, we match the identification number of each opponent given in each round with the prerating of the opponent given in the row below the name.

```
opponents1 = data.frame(V1 = prerating$prerating[opponents$X1])
opponents2 = data.frame(V2 = prerating$prerating[opponents$X2])
opponents3 = data.frame(V3 = prerating$prerating[opponents$X3])
opponents4 = data.frame(V4 = prerating$prerating[opponents$X4])
opponents5 = data.frame(V5 = prerating$prerating[opponents$X5])
opponents6 = data.frame(V6 = prerating$prerating[opponents$X6])
opponents7 = data.frame(V7 = prerating$prerating[opponents$X7])
opponents = data.frame(opponents1, opponents2, opponents3, opponents4, opponents5, opponents6, opponents7)
```

Finally we compute the average opponent rating for each player. For some players, the value in some of their rounds is actually null. For those players we will just average the given values.

```
averageopponentrating = round(rowMeans(opponents, na.rm = T), 0)
```

We collect all of the columns computed above into a single data frame of 5 columns and 64 rows.

```
tournamentdata = data.frame(names, states, totalpoints, prerating, averageopponentrating)
colnames(tournamentdata) = c("Name", "State", "Total Number of Points", "Pre-Rating", "Average Pre-Rating")
```

Now we write the data into a .csv file.

```
write.table(tournamentdata, file = "tournamentdata.csv", sep = ",", row.names = FALSE)
```

Now we make sure the data was read into a .csv file.

```
tournamentdata = read.csv(file = "/Users/chittampalliyashaswini/Desktop/Yadu/tournamentdata.csv", sep =
```