Analysis

11/18/2019

The purpose of this document is to make use of the data stored within the `FIDE` data folder.

Before beginning, we need to import several packages that help run the code below.

```
library(tidyverse)
library(data.table)
library(xts)
library(dygraphs)
library(foreach)
library(doParallel)
library(zoo)
library(stringi)
library(cowplot)
library(gganimate)
library(lubridate)
```

`tidyverse` imports dplyr, tidyr and several other useful packages for data wrangling and manipulation.

`data.table` is used purely for speeding up imports.

Access the data

The data is stored in a path we have to manipulate to get to.

```
## [1] "APR01.csv" "APR02.csv" "APR03.csv" "APR04.csv" "APR05.csv" "APR06.csv"
```

Create functions to rename datasets

Below, I define a few functions that help us rename the datasets eventually

Get the month number of all of the files in the dataset

Below, I rename create the variables names when they are imported in memory.

```
## [1] "2001-04-01" "2002-04-01" "2003-04-01" "2004-04-01" "2005-04-01"
## [6] "2006-04-01"
```

We can see that the datasets correspond to dates on which the is recorded.

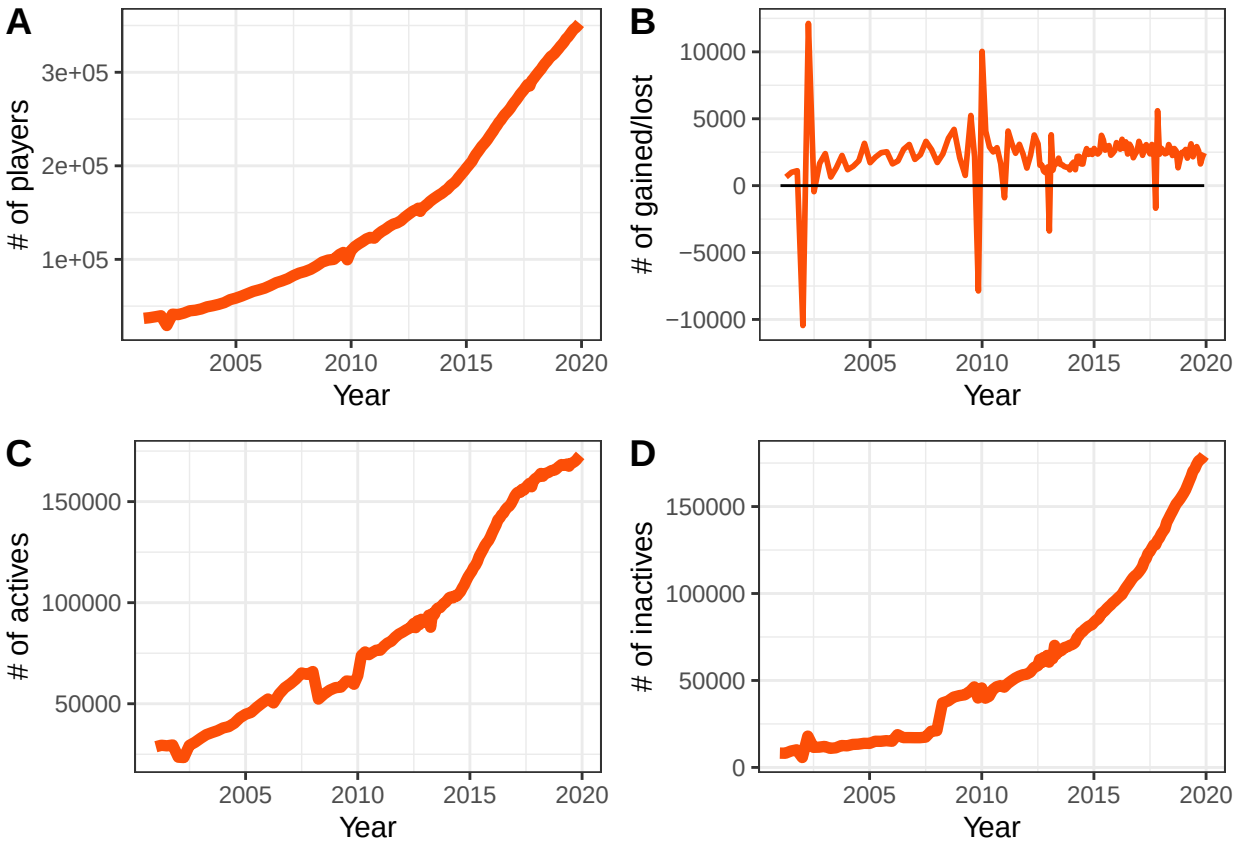
Import a sample of datasets

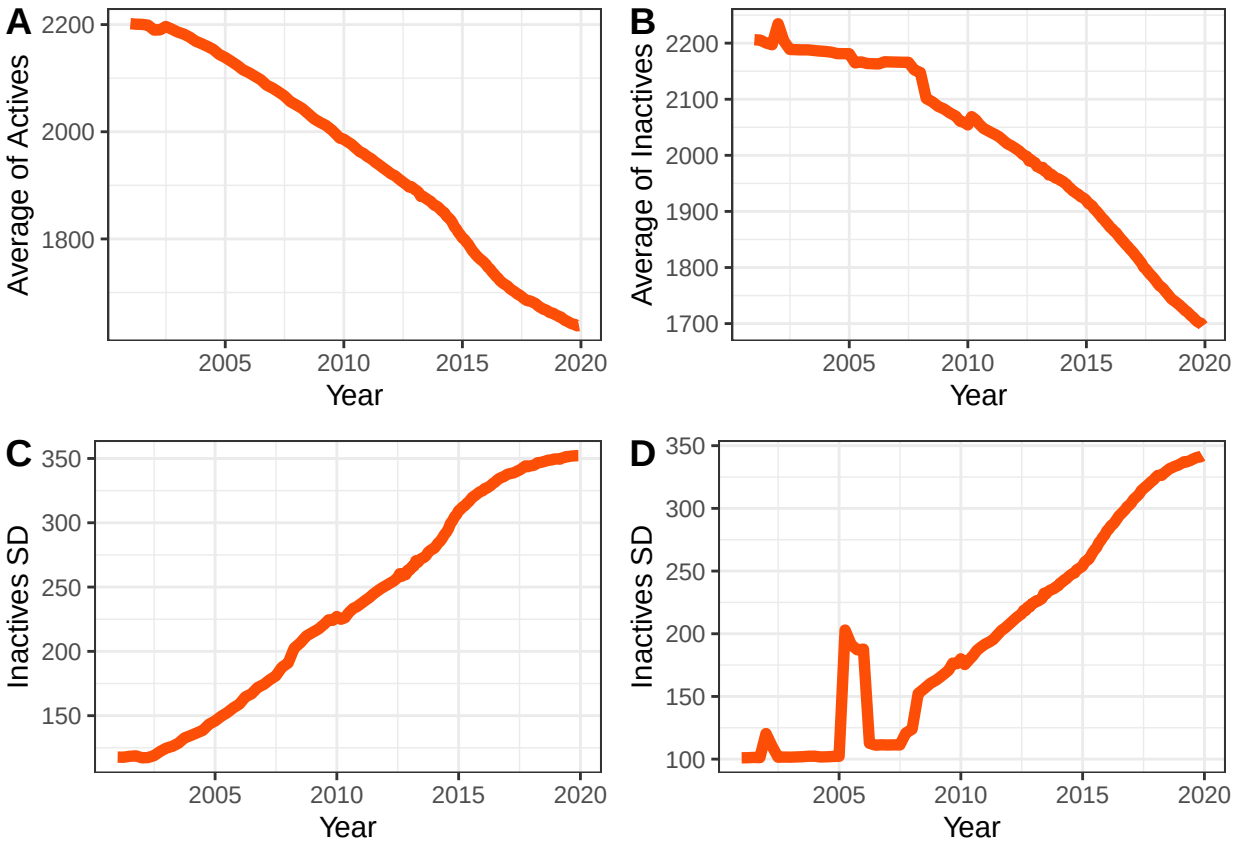
Due to the data being quite large, your RAM may be used heavily.

#write dataframe to folder

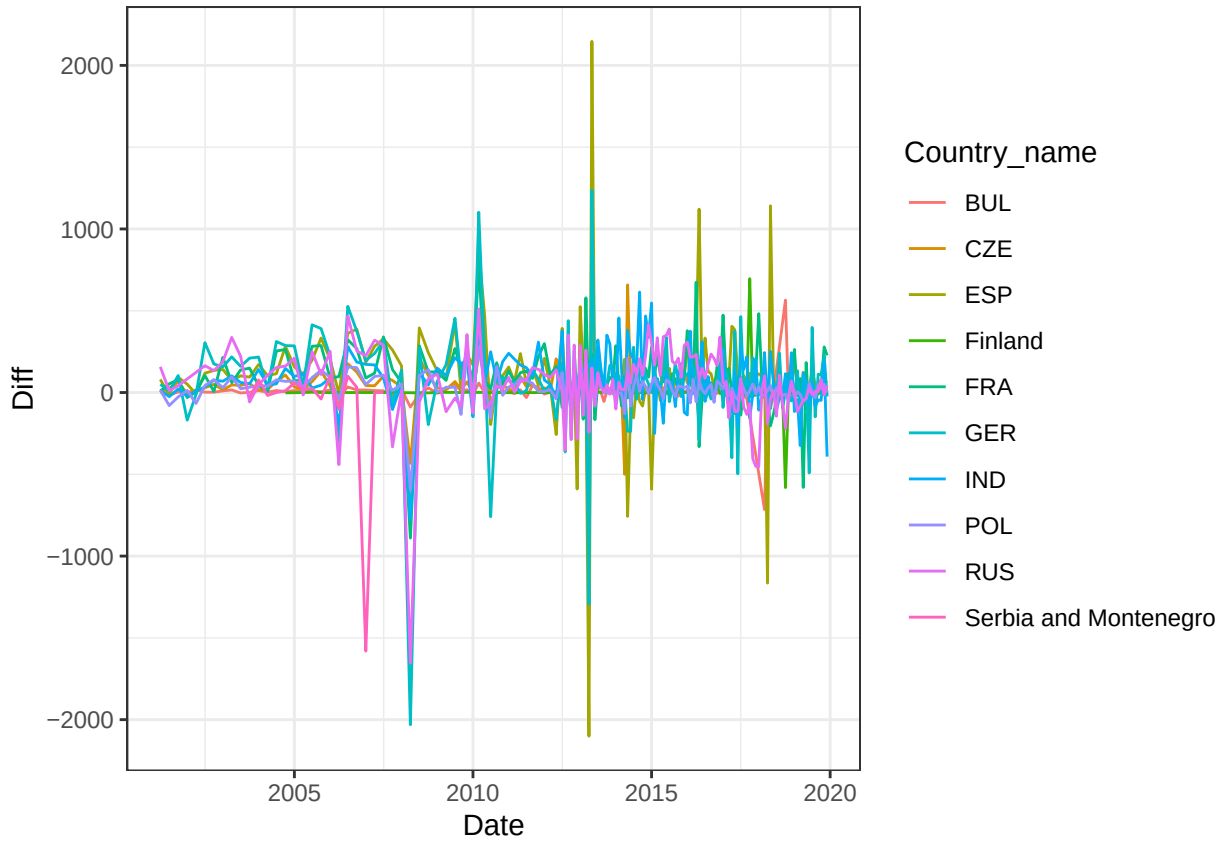
```
# fwrite(FIDE, "FIDE.csv")
# FIDE <- fread("FIDE.csv", data.table = FALSE)
```

Now that we have the data in a more clean, usable format, it's time we analyzed it.

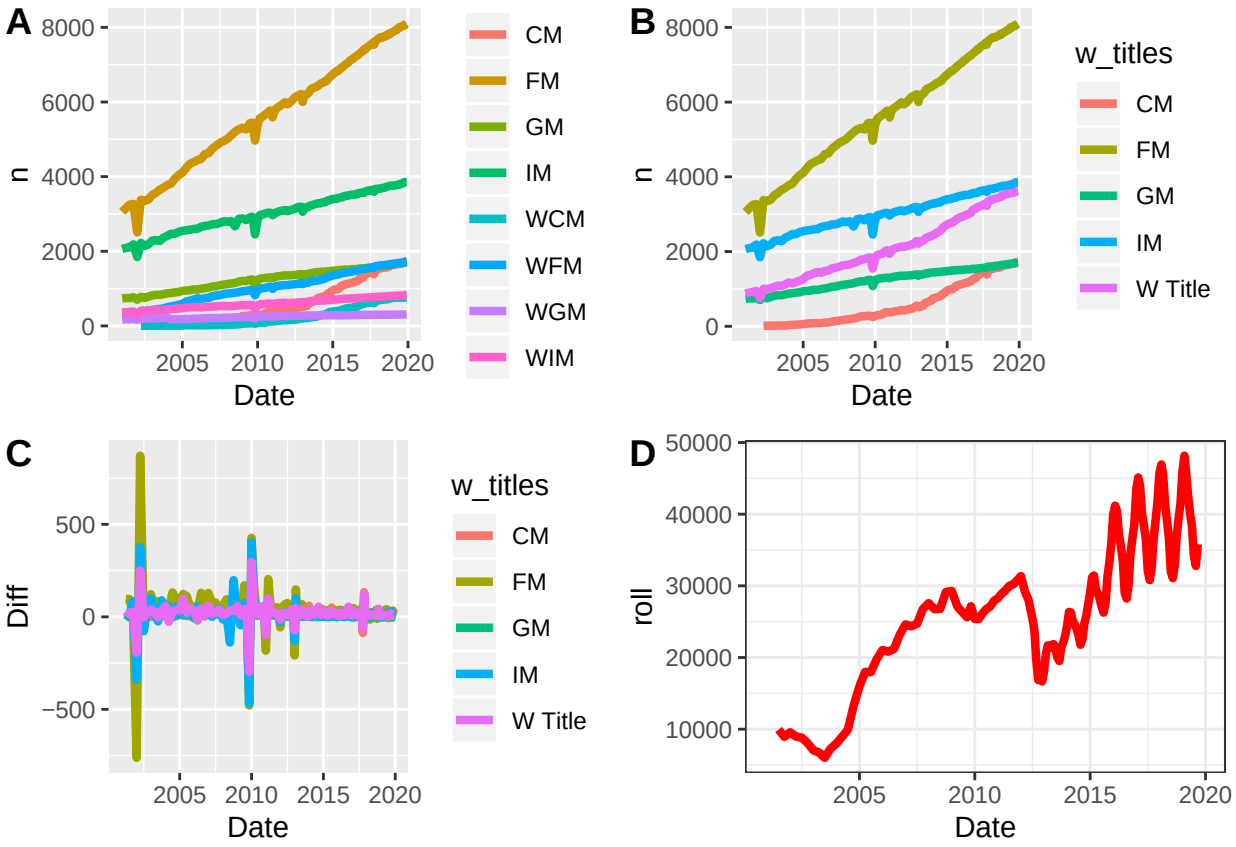


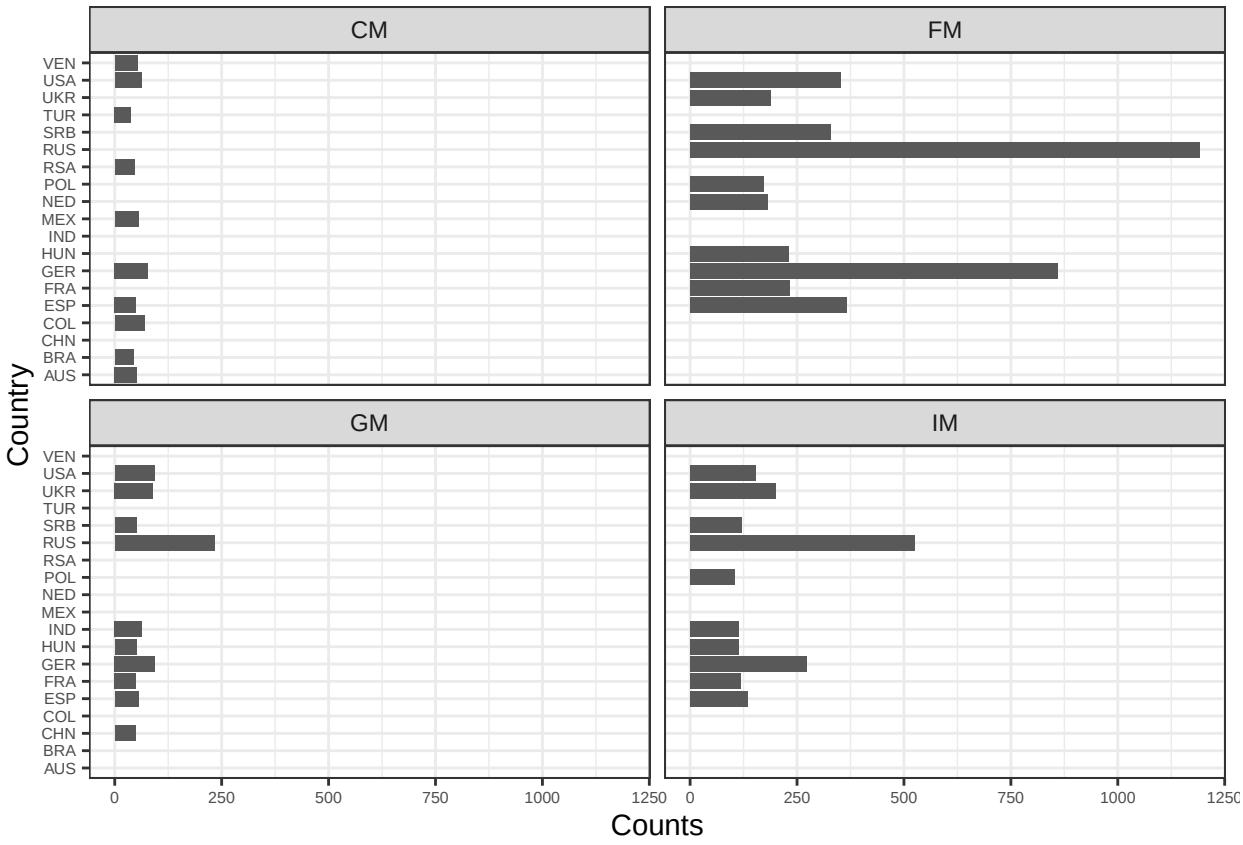


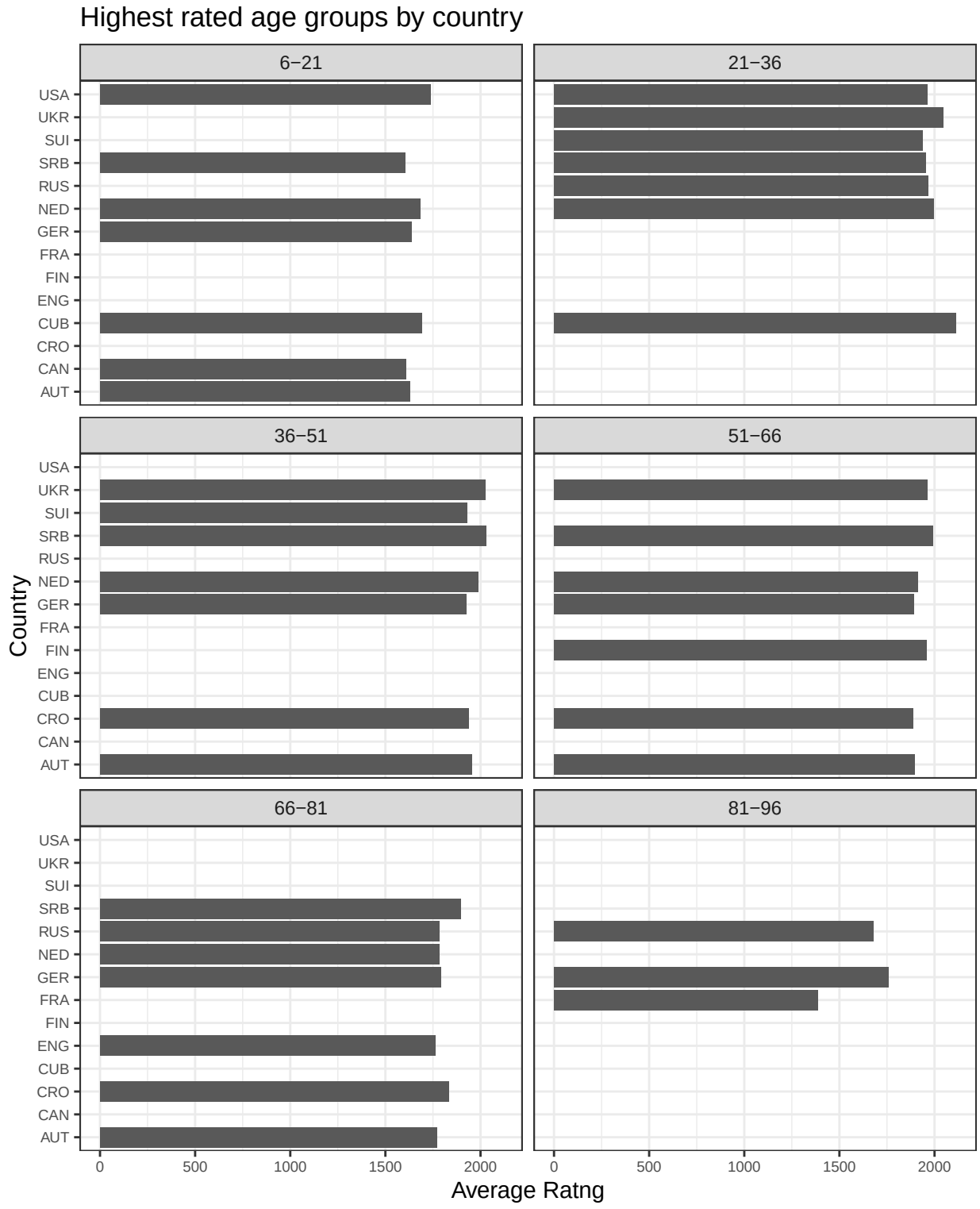

```
## Warning in (if (.Platform$OS.type == "unix")
## system else shell)(paste0("(", : '(~/GitHub/FIDE/
## Chess Scripts/FIDE Country Codes/FIDE_codes.csv) > C:
## \Users\Anuj\AppData\Local\Temp\RtmpEr1FFE\file55087970977' execution failed
## with error code 1

## Warning in fread("~/GitHub/FIDE/Chess Scripts/
## FIDE Country Codes/FIDE_codes.csv", : File 'C:
## \Users\Anuj\AppData\Local\Temp\RtmpEr1FFE\file55087970977' has size 0.
## Returning a NULL data.table.
```



#Titled player count over time


A
8000
6000
4000
2000
0
n
2005
2010
2015
2020
Date
CM
FM
GM
IM
WCM
WFM
WGM
WIM
B
8000
6000
4000
2000
0
n
2005
2010
2015
2020
Date
w_titles
CM
FM
GM
IM
W Title
C
500
0
−500
Diff
2005
2010
2015
2020
Date
w_titles
CM
FM
GM
IM
W Title
D
50000
40000
30000
20000
10000
roll
2005
2010
2015
2020
Date


CM
FM
GM
IM
Country
VEN
USA
UKR
TUR
SRB
RUS
RSA
POL
NED
MEX
IND
HUN
GER
FRA
ESP
COL
CHN
BRA
AUS
0
250
500
750
1000
1250
Counts


Highest rated age groups by country
6−21
21−36
36−51
51−66
66−81
81−96
USA
UKR
SUI
SRB
RUS
NED
GER
FRA
FIN
ENG
CUB
CRO
CAN
AUT
Country
0
500
1000
1500
2000
Average Ratng