

Analysis

December 31, 2019

The purpose of this document is to visually analyze all of the FIDE data files collected in the previous step's folder.

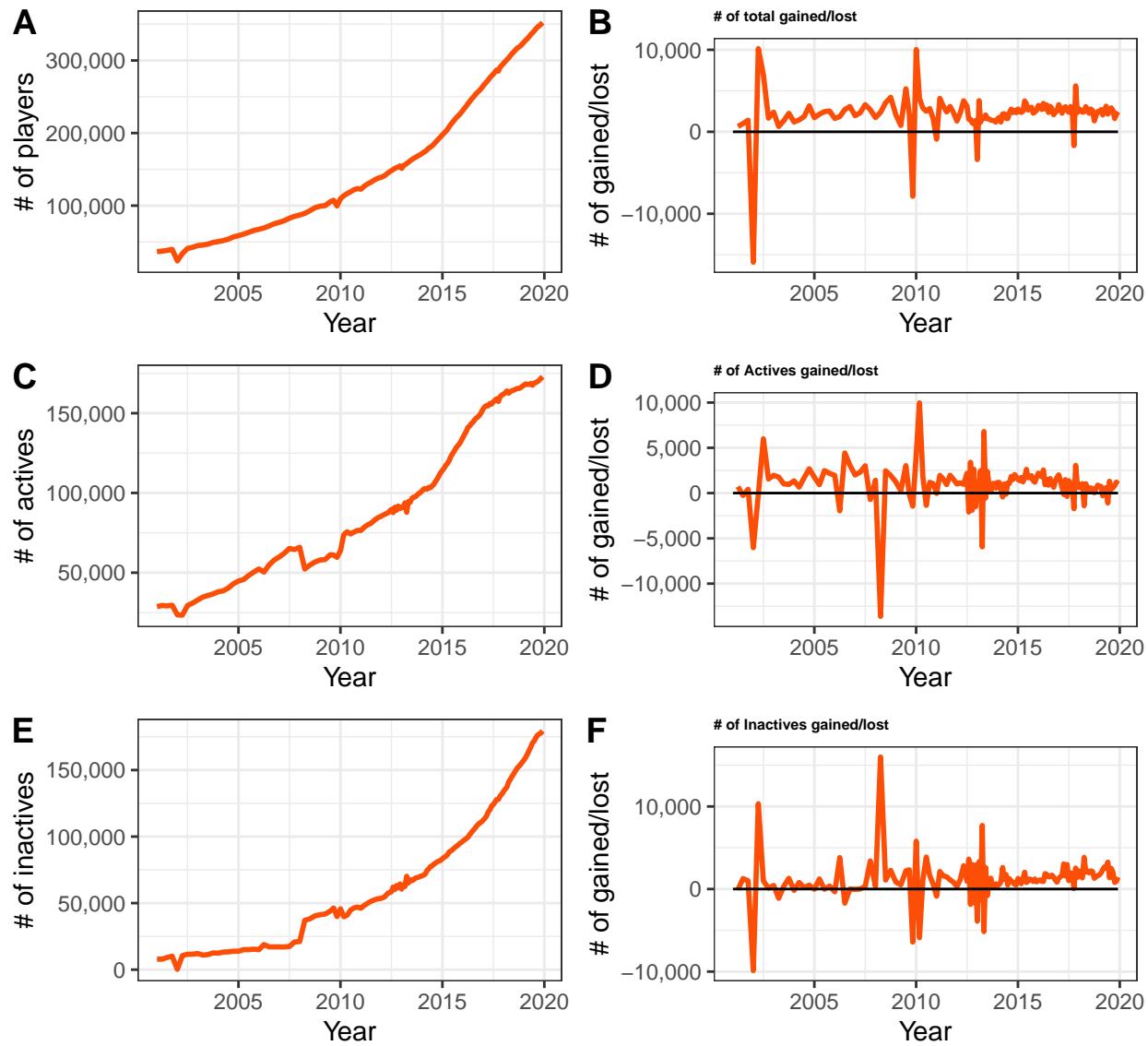
Irregular values by year

| Date | # of irregular values |
|----------|-----------------------|
| 2002.249 | 7435 |
| 2002.003 | 5455 |
| 2001.003 | 306 |
| 2001.249 | 305 |
| 2001.497 | 304 |
| 2005.497 | 244 |
| 2005.003 | 194 |
| 2004.751 | 170 |
| 2004.500 | 136 |
| 2004.251 | 114 |

As we can see from the table above, most of irregular values in the files come from early on (2001 - 2005) rather than the latest files.

I'll look to address many of the values in the early datasets eventually. For now though, over 99.9% of the data is interpretable.

Total player count over time



The four charts above reveal the total count of FIDE's members over time. Chart A shows a smooth gradual increasing growth curve among total players. This may lead you to believe that more chess players are playing tournaments, but a player's **activity** is a better metric to go by.

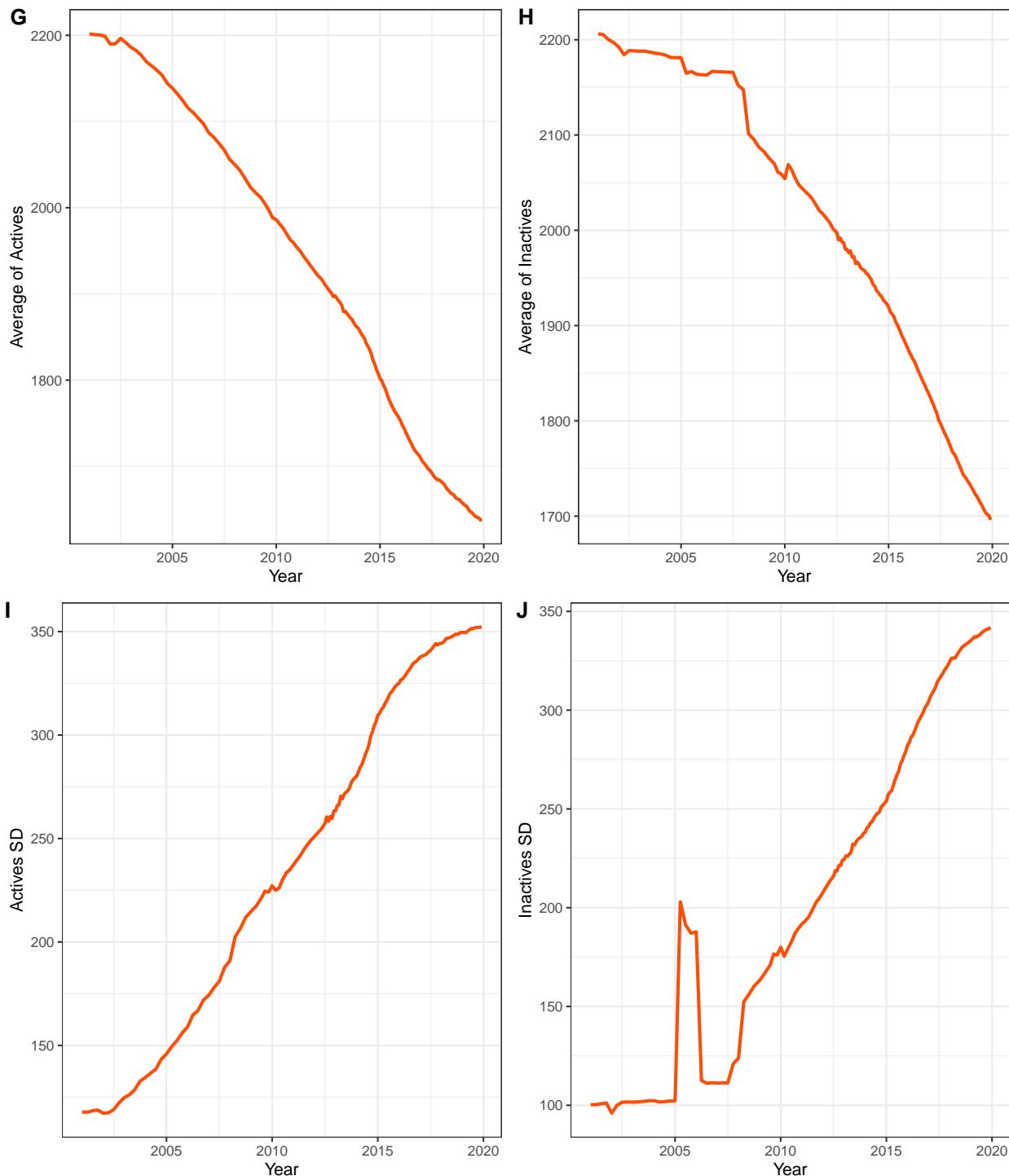
Activity is defined by if a given FIDE player had played a rated FIDE game within the past 12 months. If we take this into account, charts C and E show how the total active & inactive player count increase over time. Both charts show a fairly linear trend over time, but from 2007 to 2010 in each graph, there was a noticeable drop off in the active player base and increase in the inactive player base. I have my doubts on if there is faulty data here because chart A shows no irregularity during that time.

This dip may be due to the economic crash during that time, but I need to do more exploration on this topic before making any definitive statements. (Exploring the dip by country is future work)

Charts B, D and F show the number of players gained and lost over time.

The most relevant of the 3 graphs is B which shows several instances where total player counts dropped off. I genuinely don't if the data is faulty because of my doing or if FIDE is providing incomplete datasets.

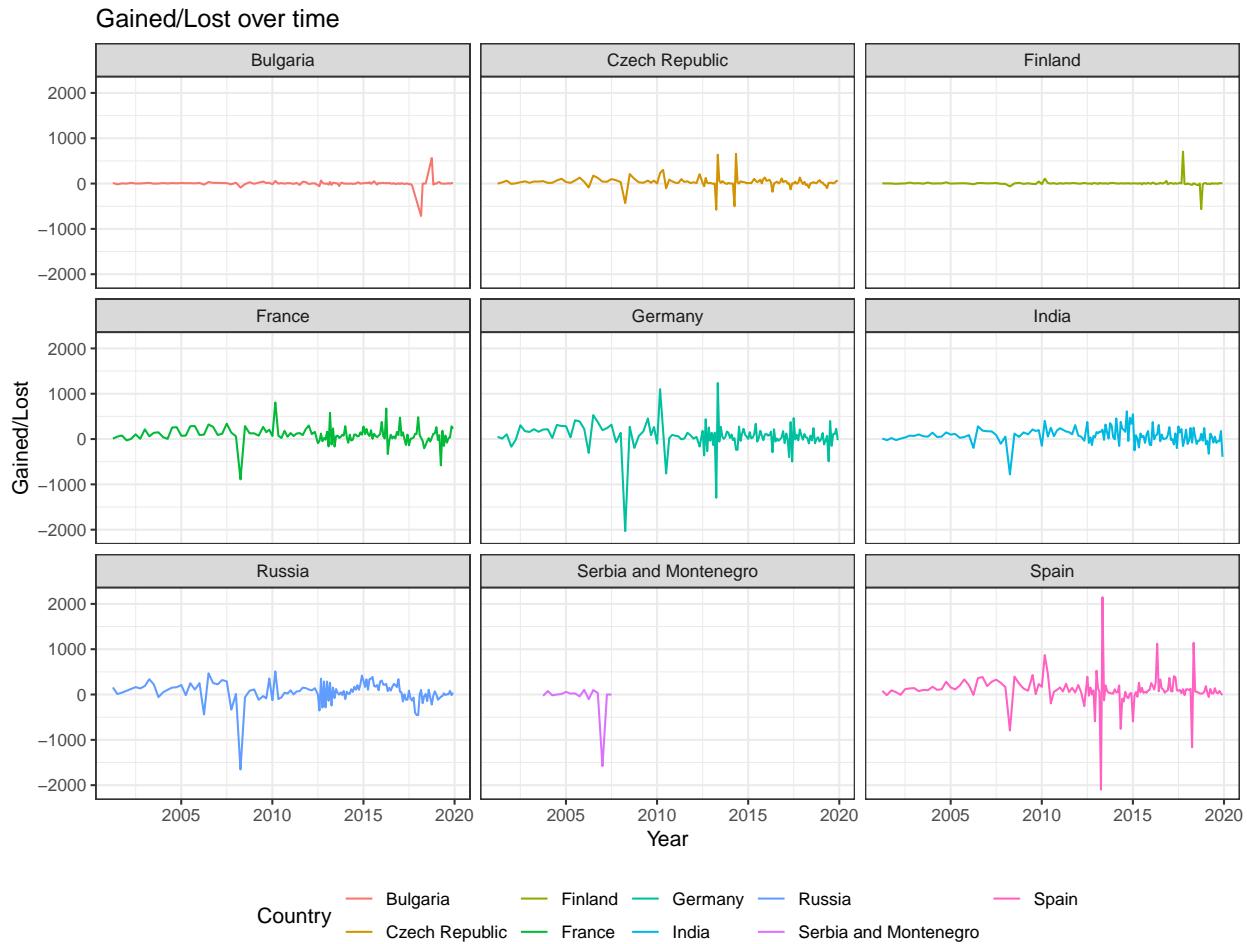
Rating stability over time (misleading)



Each chart above shows how the Active and Inactive player's average rating and standard deviation have progressed over time. For the most part, it is a meaningless metric because FIDE has brought in more **lower** rated chess players into the player pool over time.

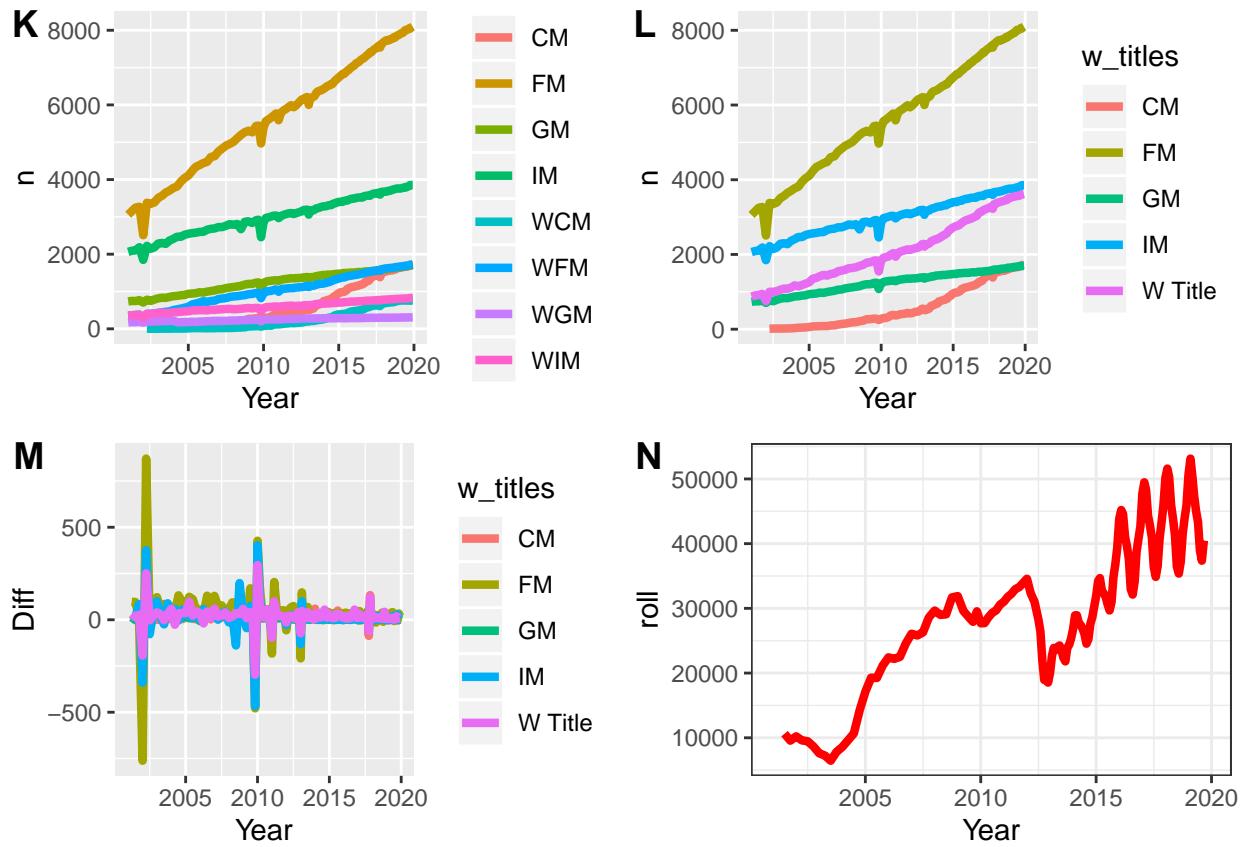
This explains the slow decrease in average rating and increase in standard deviation over time.

Which countries have seen the greatest changes in player counts?



Explanation:

Titled player count over time



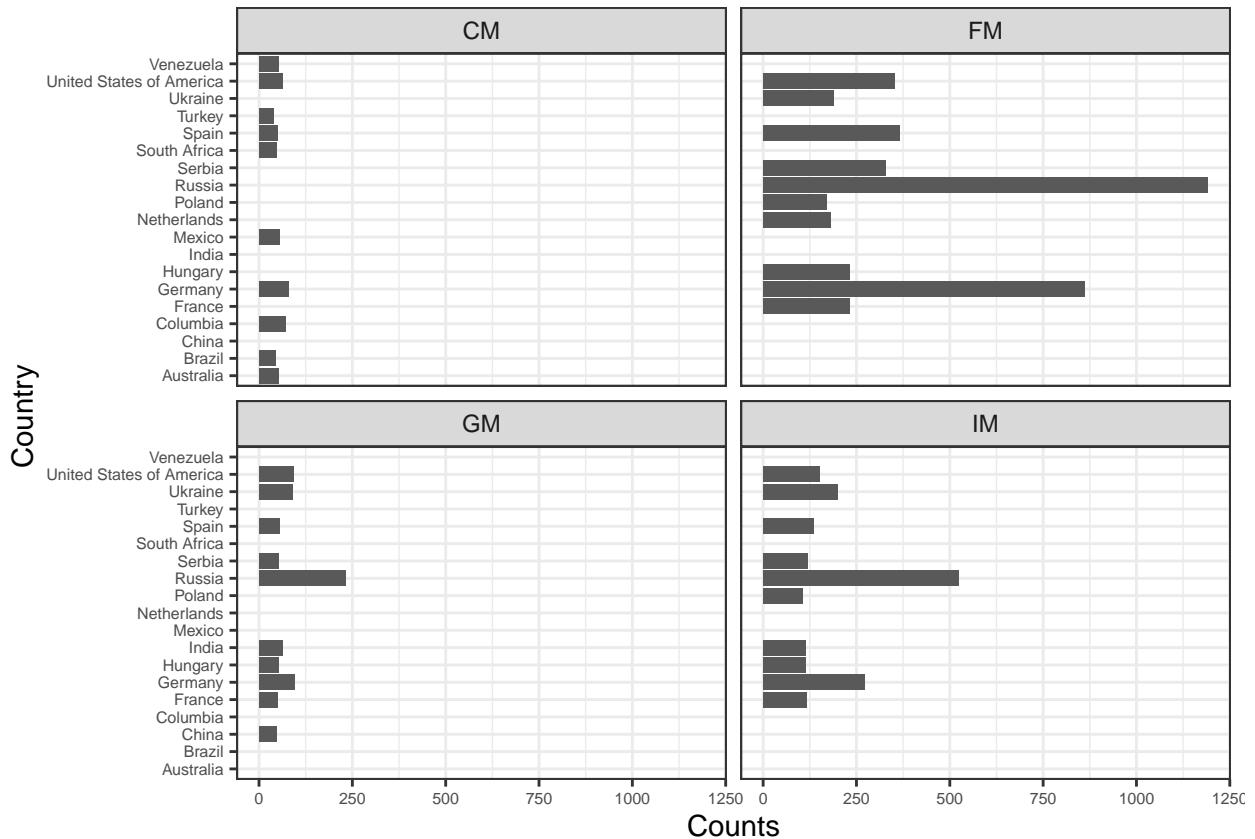
A:

B:

C:

D:

Titled player counts by country



CM:

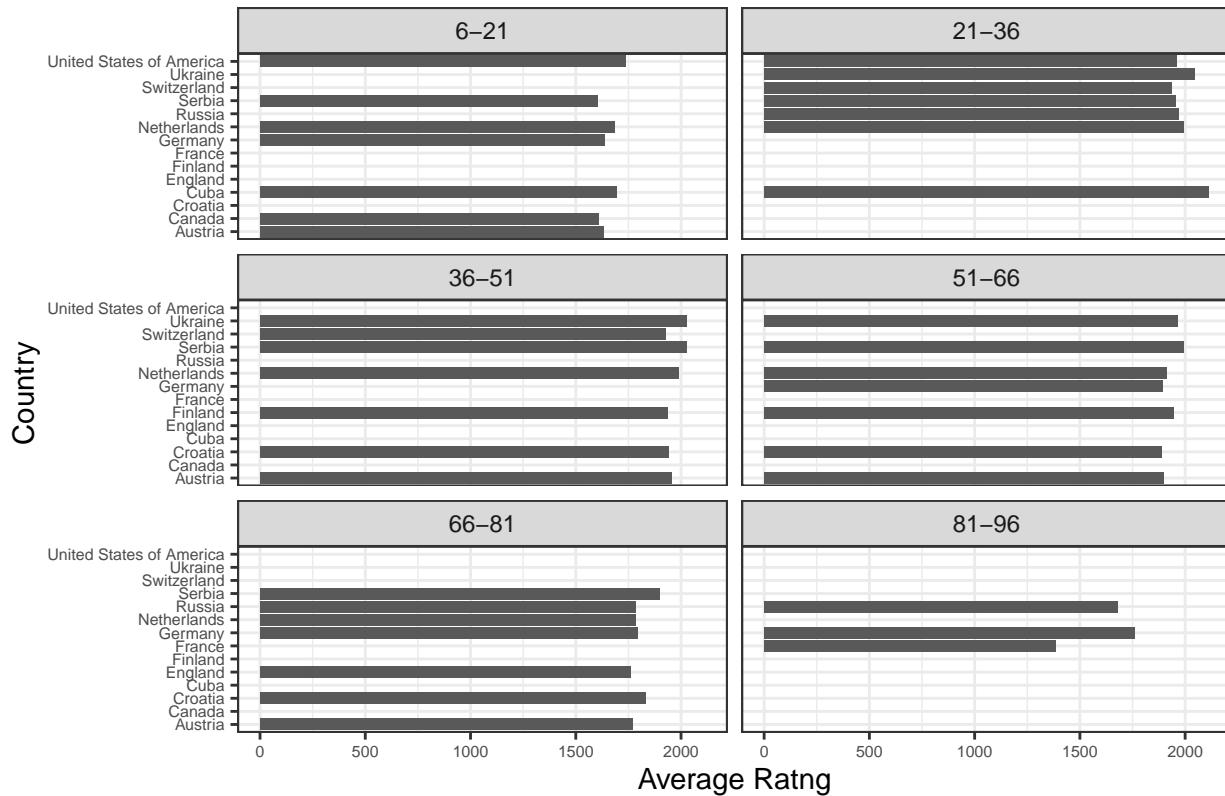
FM:

GM:

IM:

Strongest countries by age group

Highest rated age groups by country



6-21:

21-26:

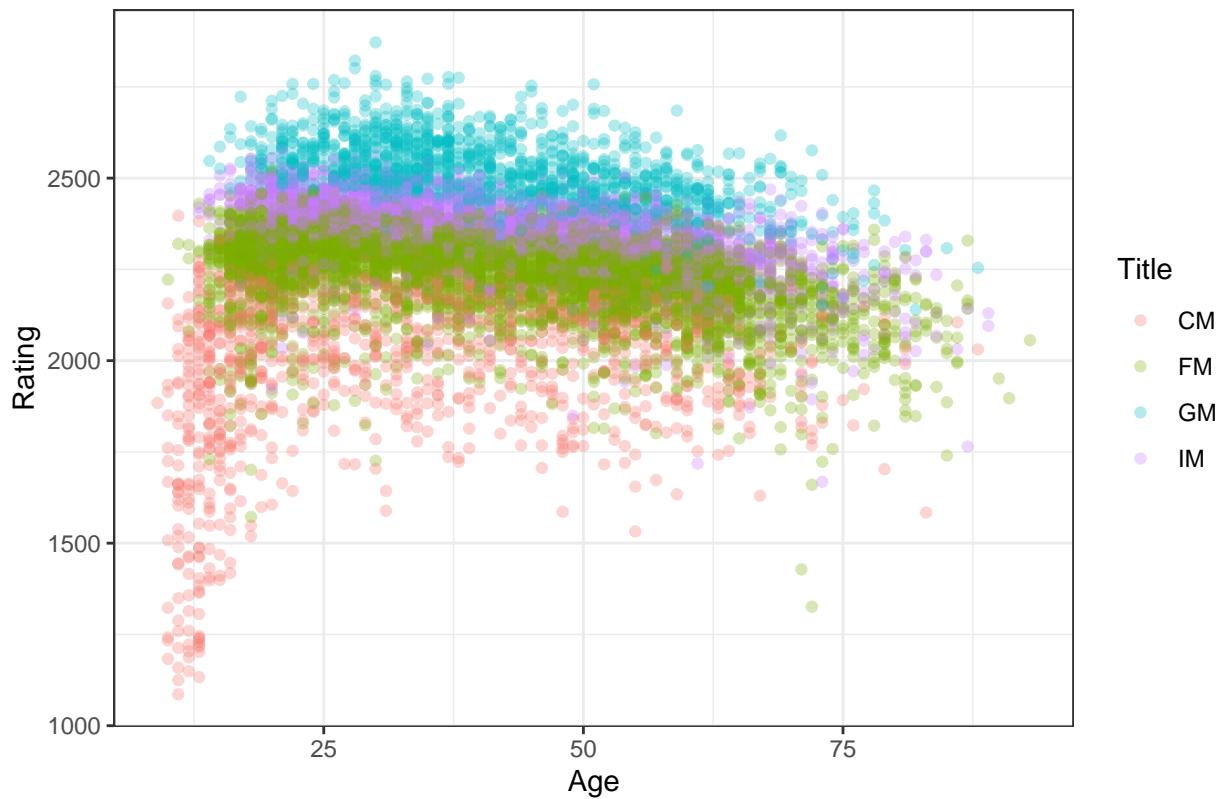
36-51:

51-66:

66-81:

81-96:

Age vs Rating of different titles for December 2019



Explanation: