

Progress Report 1

Ali

Introduction

In this report I will provide a summary of the work I have done on ranking sports teams. I decided to start with making algorithms for chess data. I found a source that publishes data on all official FIDE (International Chess Federation) games weekly. The name of the source is The Week in Chess. That database uses PGNs (Portable Game Notation) and I have been working with them by transforming into .csv file format. Therefore, I have access to the most current data.

In order to keep the report clean, I will not be displaying much of the coding part here.

Creating a Dataset

Since the initial file format of the database is .pgn, I have to transform it into .csv. First, I downloaded three .pgn files from The Week in Chess website (<https://theweekinchess.com/twic>) containing data on the three weeks of chess games played in June 2025. Second, I wrote some code to create a .csv dataset from .pgn files (I used ChatGPT to assist me with coding).

Here are 10 rows of the resulting dataset:

```
fide_games <- read.csv("fide_games_weekly.csv")
head(fide_games, 10)
```

		Event	Date	White	Black	Result
1	13th Norway Chess 2025	2025-06-03	Erigaisi,Arjun	Caruana,F	1-0	
2	13th Norway Chess 2025	2025-06-03	Wei Yi	Carlsen,M	1/2-1/2	
3	13th Norway Chess 2025	2025-06-03	Nakamura,Hi	Gukesh,D	1-0	
4	13th Norway Chess 2025	2025-06-05	Nakamura,Hi	Erigaisi,Arjun	1/2-1/2	
5	13th Norway Chess 2025	2025-06-05	Gukesh,D	Wei Yi	1-0	
6	13th Norway Chess 2025	2025-06-05	Carlsen,M	Caruana,F	1-0	

7	13th Norway Chess 2025	2025-06-06	Erigaissi,Arjun	Carlsen,M	1/2-1/2
8	13th Norway Chess 2025	2025-06-06	Caruana,F	Gukesh,D	1-0
9	13th Norway Chess 2025	2025-06-06	Wei Yi	Nakamura,Hi	1/2-1/2
10	13th Norway Armageddon	2025-06-03	Wei Yi	Carlsen,M	1-0

	WhiteElo	BlackElo
1	2782	2776
2	2758	2837
3	2804	2787
4	2804	2782
5	2787	2758
6	2837	2776
7	2782	2837
8	2776	2787
9	2758	2804
10	2758	2837

So far I have been working with the three weeks of data. Once I make the most accurate (to my ability) algorithm for ranking chess players, I will be able to add more weeks of data without an issue.

First Algorithms

I used the code given for ranking football teams as a template for my algorithms. Since there are three possible game outcomes, I encoded them as +1/0/-1. Then, I built the X matrix, added sum-to-zero constraint, calculated ratings using least squares. Finally, I saved the resulted ratings in a new data frame, also adding the “Rank” variable.

First Results

I like to brainstorm and make notes before and after doing some work. That’s my initial comment on the results:

“I am happy with the first results because I can observe some correlation between the rankings I made and the official FIDE rankings. However, there are still many outliers and inaccurate ratings. I think I should account for Elo and different events to make more accurate rankings. After that, I should try working with a 52-week dataset instead of a 3-week dataset. (It might take hours to run a ranking algorithm for a 52-week dataset).”

Here are the first 20 rows in the resulted dataset:

```
fide_rankings <- read.csv("chess_rankings_3weeks.csv")
head(fide_rankings, 20)
```

	Player	Rating	Rank
1	Silva Cerda,Mark Valentin	3.133185	1
2	Minguell Soler,Joan	2.455160	2
3	Serrano Batova,Vicent	2.321823	3
4	Sole Pijuan,Ferran	2.315067	4
5	Namay Caceres,S	2.238848	5
6	Dubov,Daniil	2.176203	6
7	Sriniaiyer,Sudarshan	2.153066	7
8	Lyu,Taifeng (Matt)	2.088820	8
9	Destic,Elena	2.047504	9
10	Zhang,Hongya	1.988588	10
11	Aryan,C	1.948555	11
12	Carlsen,M	1.939103	12
13	Medghoul,S	1.927331	13
14	Stearman,Josiah	1.921986	14
15	Berro,Mahdi	1.900397	15
16	Nesterov,Arseniy	1.892054	16
17	Edwardss,Dylan	1.877689	17
18	Zwirs,Nico	1.877689	18
19	Valiyev,Shahin	1.802556	19
20	Maslovskiy,Vladimir	1.783875	20

Although I was able to instantly spot a few highest ranked players by FIDE ratings, there was still much room for improvement. I noticed that the majority of outliers consisted of players who played a small number of games. Therefore, I decided to filter out players with fewer than 10 games played.

Here are the first 20 rows in the filtered dataset:

```
fide_rankings_filtered <- read.csv("chess_rankings_filtered.csv")
head(fide_rankings_filtered, 20)
```

	Player	Rating	Rank	GamesPlayed
1	Carlsen,M	1.939103	12	38
2	Rostovtsev,Dmitry	1.783875	21	11
3	Abdusattorov,Nodirbek	1.724174	27	35
4	Firouzja,Alireza	1.691195	33	71
5	Nakamura,Hi	1.620964	42	87

6	Woodward,Andy	1.607011	44	40
7	Wei Yi	1.589027	45	12
8	Niemann,Hans Moke	1.588503	46	153
9	Aronian,L	1.577216	49	65
10	Erigaisi,Arjun	1.572610	51	65
11	Puranik,A	1.554412	53	11
12	Sevian,Samuel	1.518326	57	32
13	Andreikin,D	1.516246	59	20
14	Giri,A	1.505594	60	82
15	Sindarov,Javokhir	1.501122	61	35
16	Movahed,Sina	1.500922	62	21
17	Tahbaz,Arash	1.477165	71	11
18	Nepomniachtchi,I	1.475138	72	46
19	Xiong,Jeffery	1.461710	73	72
20	Salem,AR	1.457514	74	11

The results in the filtered dataset were great, I could recognize 15 players out of 20. However, filtering out players by the number of games played has its disadvantages. For example, Daniil Dubov was ranked the 6th with 5 games played in the unfiltered dataset. Although he is indeed one of the highest ranked chess players in the world, filtering out by the number of games played excluded him completely from our rankings. Therefore, I should keep working on improving the algorithms

Weighted Least Squares

I have carefully analyzed the code for weighted rankings in football and used it as a template for my code. My goal was to weigh games based on how closely players were matched. The coding part was quite complicated and might require revision in future.

Nevertheless, here are the first 20 rows in the weighted rankings dataset:

```
weighted_rankings <- read.csv("chess_weighted_rankings.csv")
head(weighted_rankings, 20)
```

	Player	Rating	Rnk
1	Silva Cerda,Mark Valentin	3.133185	1
2	Minguell Soler,Joan	2.455160	2
3	Serrano Batova,Vicent	2.321823	3
4	Sole Pijuan,Ferran	2.315067	4
5	Namay Caceres,S	2.238848	5
6	Dubov,Daniil	2.176203	6

7	Sriniaiyer,Sudarshan	2.153066	7
8	Lyu,Taifeng (Matt)	2.088820	8
9	Destic,Elena	2.047504	9
10	Zhang,Hongya	1.988588	10
11	Aryan,C	1.948555	11
12	Carlsen,M	1.939103	12
13	Medghoul,S	1.927331	13
14	Stearman,Josiah	1.921986	14
15	Berro,Mahdi	1.900397	15
16	Nesterov,Arseniy	1.892054	16
17	Edwardss,Dylan	1.877689	17
18	Zwirs,Nico	1.877689	18
19	Valiyev,Shahin	1.802556	19
20	Maslovskiy,Vladimir	1.783875	20

The results were unexpected because I couldn't recognize most of the highest rated players in my dataset. I think it happened because I still do not fully account for the number of games played. I have an idea of giving each game a weight based on the total number of games played by both players, and that's what I am currently working on.

Filtering Games for Weighted Rankings

Although I was not happy with the first results of my algorithm for weighted rankings, I wanted to see how it would perform if I filtered out games where at least one player played less than 5 games. The initial total amount of chess games over three weeks of data was 25265. After applying filtering, the number of games decreased to 22429. Since I wrote the algorithm for weighted rankings as a function, I just ran the function on the new filtered dataset.

Here are the first 20 rows in the filtered weighted rankings dataset:

```
weighted_rankings_filtered <- read.csv("chess_weighted_rankings_filt.csv")
head(weighted_rankings_filtered, 20)
```

	Player	Rating	Rnk
1	Dubov,Daniil	1.879787	1
2	Zwirs,Nico	1.809888	2
3	Carlsen,M	1.770886	3
4	Nesterov,Arseniy	1.693184	4
5	Wei Yi	1.639697	5
6	Rostovtsev,Dmitry	1.583120	6
7	Maslovskiy,Vladimir	1.583120	7
8	Stopa,Ja	1.577783	8

9	Abdusattorov, Nodirbek	1.545570	9
10	Fedoseev, Vl3	1.534668	10
11	Bauer, Ch	1.521772	11
12	Decuigniere, Tom	1.521772	12
13	Firouzja, Alireza	1.509723	13
14	Sankalp, Gupta	1.444380	14
15	Nakamura, Hi	1.441517	15
16	Woodward, Andy	1.432023	16
17	Bu Xiangzhi	1.424400	17
18	Niemann, Hans Moke	1.424163	18
19	Fedorchuk, S	1.420308	19
20	Aronian, L	1.403354	20

The results are pretty good. Filtering helped a lot and now we can observe some correlation between our results and official FIDE ratings. However, I do not like the idea of using filtering because we excluded almost 3.000 observations from our dataset.

Ideas

The idea of weighting each game based on the total number of games played by both players sounds the most rational at the moment. If the algorithm works satisfactorily, I will test it on a 52-week dataset.