

Project 1

In this project, you're given a text file with chess tournament results where the information has some structure. Your job is to create an R Markdown file that generates a .CSV file (that could for example be imported into a SQL database) with the following information for all of the players:

Player's Name, Player's State, Total Number of Points, Player's Pre-Rating, and Average Pre Chess Rating of Opponents

For the first player, the information would be:

Gary Hua, ON, 6.0, 1794, 1605

1605 was calculated by using the pre-tournament opponents' ratings of 1436, 1563, 1600, 1610, 1649, 1663, 1716, and dividing by the total number of games played.

If you have questions about the meaning of the data or the results, please post them on the discussion forum. Data science, like chess, is a game of back and forth...

The chess rating system (invented by a Minnesota statistician named Arpad Elo) has been used in many other contexts, including assessing relative strength of employment candidates by human resource departments.

You may substitute another text file (or set of text files, or data scraped from web pages) of similar or greater complexity, and create your own assignment and solution. You may work in a small team. All of your code should be in an R markdown file (and published to rpubs.com); with your data accessible for the person running the script.

Excerpt from text file:

Pair Num	Player Name USCF ID / Rtg (Pre->Post)	Total Pts	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7
1	GARY HUA ON 15445895 / R: 1794 ->1817	6.0	W 39	W 21	W 18	W 14	W 7	D 12	D 4
2	DAKSHESH DARURI MI 14598900 / R: 1553 ->1663	6.0	W 63	W 58	L 4	W 17	W 16	W 20	W 7
3	ADITYA BAJAJ MI 14959604 / R: 1384 ->1640	6.0	L 8	W 61	W 25	W 21	W 11	W 13	W 12
4	PATRICK H SCHILLING MI 12616049 / R: 1716 ->1744	5.5	W 23	D 28	W 2	W 26	D 5	W 19	D 1
5	HANSHI ZUO MI 14601533 / R: 1655 ->1690	5.5	W 45	W 37	D 12	D 13	D 4	W 14	W 17