

DAV 5400 Fall 2019 Project 3 (Due Sunday Nov 10: 100 Points)

Analyzing Chess Tournament Results

In this assignment, you're given a text file ("tournamentinfo.txt") with chess tournament results where the information has some structure. Your job is to create a Jupyter Notebook that generates a .CSV file with the following information for all of the chess players:

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

For the first player shown in the file excerpt below, the information would be:

Gary Hua, ON, 6.0, 1794, 1605

His "Average Pre Tournament Chess Rating of Opponents" score of 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.

For each player we are provided with the total points they've won during the tournament and details on the results of their seven rounds of play. For each round we are given the unique ID of their opponent (an integer value) and an indicator of whether they won ('W'), lost ('L'), achieved a draw ('D'), had a bye for that round ('B'), or were unable to compete ('U').

If you have questions about the meaning of the remainder of the data or the results, please post them in the weekly 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 work in small groups of no more than three (3) people for this assignment.** ** Save all of your work for this project within a single Jupyter Notebook and upload it to your online DAV5400 GitHub directory. Be sure to save your Notebook using the following nomenclature : **first initial_last name_project3**" (e.g., J_Smith_project3). **Small groups should identify all group members at the start of the Jupyter Notebook and each team member should submit their own copy of the team's work within Canvas.**

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 ON	GARY HUA 15445895 / R: 1794 ->1817	6.0 N:2	W 39 W	W 21 B	W 18 W	W 14 B	W 7 W	D 12 B	D 4 W
2 MI	DAKSHESH DARURI 14598900 / R: 1553 ->1663	6.0 N:2	W 63 B	W 58 W	L 4 B	W 17 B	W 16 W	W 20 B	W 7 B
3 MI	ADITYA BAJAJ 14959604 / R: 1384 ->1640	6.0 N:2	L 8 W	W 61 B	W 25 W	W 21 B	W 11 B	W 13 W	W 12 B
4 MI	PATRICK H SCHILLING 12616049 / R: 1716 ->1744	5.5 N:2	W 23 W	D 28 B	W 2 W	D 26 B	W 5 B	D 19 B	W 1 B
5 MI	HANSHI ZUO 14601533 / R: 1655 ->1690	5.5 N:2	W 45 B	W 37 W	D 12 B	D 13 W	D 4 B	W 14 W	W 17 B

Grading rubric: you'll receive (up to 70) points if you successfully write the player name and total points into a pandas DataFrame, then into the .CSV file. You'll receive (up to 90) points if you also successfully process the information from the second line for each player: state and pre-tournament rating. To get the full 100 points on the assignment, you will also need to successfully calculate and process the average pre-tournament rating for each player's opponents.

As a reminder, this assignment is due no later than 11.59pm on Sunday Nov 10.