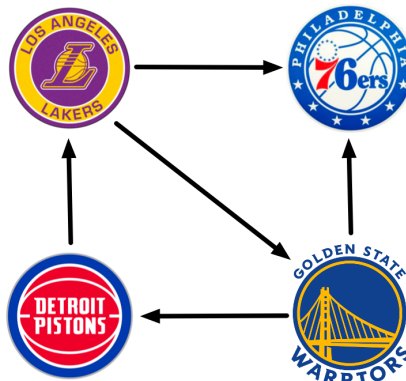


### Session 3: Tackling Your Ranking

Please submit your completed work before 7 AM EDT on Thursday, July 29.

This homework will have you 1) working with the Colley and Massey algorithms, 2) learning about the Elo ranking algorithm, 3) reflecting on the impact of bad data, and 4) coming up with a way to compare the Massey and Colley methods when run on a given dataset. Note, you will need to download codes from Session 3 to complete some of the problems.

1. Watch *The Elo Rating System for Chess and Beyond* at <https://www.youtube.com/watch?v=AsYfbmp0To0>. In 2-3 paragraphs describe what you found most interesting about the Elo method and how it compares to the Massey and Colley methods.
2. Read *If Your Data Is Bad, Your Machine Learning Tools Are Useless* at <https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless>. In 2-3 paragraphs describe what you found most interesting about this article. If you have thoughts about how it relates to your research, I encourage you to share them.
3. Let the directed graph below represent a series of games between (going in clockwise order) Los Angeles Lakers, Philadelphia Sixers, Golden State Warriors, and Detroit Pistons. An arrow from team  $i$  to team  $j$  represents one game between  $i$  and  $j$  in which team  $i$  beat  $j$ . For instance, the Warriors beat the Pistons.



- (a) Write the linear system  $A\mathbf{r} = \mathbf{b}$  below that when solved will calculate the Colley ratings for this network of games. The matrix  $A$  should have 4 rows and 4 columns. Assign the Lakers (L), Sixers (S), Warriors (W) and Pistons (P) to rows 1, 2, 3, and 4, respectively. So, you will fill in a system of the following form:

$$\begin{pmatrix} & & & \\ & & & \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} L \\ S \\ W \\ P \end{pmatrix} = \begin{pmatrix} 3/2 \\ 0 \\ 3/2 \\ 1 \end{pmatrix}$$

- (b) Using computer code, solve this system. Write the rating of each team and then list the teams and their rating beginning with the team ranked first and ending with the last place team.

4. The files attached on Schoology to this assignment `2021games.txt` and `2021teams.txt` contain games played in 2020-21 NBA regular season. Run the Massey and Colley methods to get rankings of all the NBA teams. In your write-up, list the top 10 teams in each ranking along with their ratings (which are decimals and not integers, which are the rankings). Ranking research often involves comparing rankings. Come up with a way to compare these rankings. They differ. You could do this with exposition and write about the differences. You could do this visually and create a data visualization that details their differences. You could come up with a metric that explains their differences numerically. Compare these rankings. Be sure that the reader can understand what you are doing. Note, this can but does not need to involve any discussion of basketball or the NBA teams.
5. What are your current thoughts regarding your research? After seeing Colley and Massey and already having seen PageRank, what comes to mind? Do you have questions that will help you think about research into a ranking question? If what, what are your questions?

---

**Submitting your work** - All your responses should be contained in one document saved as a PDF. Upload the PDF and your programming files to Schoology.