

So the plan is basically to analyze all Loyola Ramblers games since Fall 2016 (when I first enrolled). The binary response variable is outcome (win or loss) and the explanatory variables can include location (home, away, neutral), opponent, conference of opponent, opponent ownership (public vs. private), day of week, start time, attendance, points scored/allowed, duration (in sets, innings, whether there was overtime, etc.), ongoing win/loss streak, national poll ranking (if applicable), whether or not I attended, etc. I will analyze games from the seven "head-to-head" sports (both men's and women's basketball, soccer, and volleyball, plus softball). I'm still figuring out how to code draws for soccer, I think I may exclude them or count them as 0.5. Also, since games are ongoing, I think I may exclude currently ongoing seasons (men's volleyball and softball) so the data are not influenced.

I plan to do analyses on how much the pandemic affected performance (performance during first full season during the pandemic vs. before), comparing seasons (e.g., before and after coaching changes), home vs. road performance (chi-square?), whether a team made a postseason tournament and/or won a conference championship, etc.

Since the project is rather open-ended, I found it hard to narrow myself down to one dataset, so when I saw binary outcome variable wins and losses were the first thing I thought of. I also saw that datasets could be created so this was what I thought of first. I didn't want to go back to the Chicago Data Portal because I thought that might be repetitive and this is something I have been very passionate about so I'm hoping this will work.

I don't have the full dataset available because there is no single source for this, but all of the data are publicly available on the official website (loyolaramblers.com). I will deep dive into the website and pull the data from every game to create the dataset (I visit the website a lot so I trust all the data are available). I estimate there are roughly 1,100 games since Fall 2016. Also, I will be working by myself.