**March Madness Results**

To predict upsets in the tournament, I need a dataset containing the results of previous tournaments. Thankfully, using a quick google dataset search, there was a dataset containing the results of all March Madness tournaments from 1985 to 2021. The file for this dataset is titled **data\_cleaned.csv**. I can filter the data to look for the upsets and see whether there is a trend within years, or it is simply just a school problem i.e. one team seems to be getting upset at the most inopportune times.

**Box Scores for January 11, 2023**

In the file titled **Jan 11, 2023 Box Scores.csv**, I looked at all the games that were played that day and got some relevant data like points, rebounds, and assists - the basic stats. This data was scraped via ESPN box scores in the only code file given in this step. I am looking to see why a team won or lost besides looking at who scored the most points, hence the scraping for rebounds and assists as a start, and hopefully find some correlations between points and other stats. An extra file titled **American\_v\_Army** showcased how to extract data from one game on that day. That creation corresponds to the largely commented out code using triple single quotes.

**2021-2022 Regular Season Stats**

This looks at the total number of points, rebounds, assists, etc. across the whole season. I will use this in hopes to 1, predict the results of the 2022 March Madness and 2, using the results of the 2022 March Madness tournament to hopefully correctly predict the upcoming March Madness tournament (since that is what the model is supposed to do, or rather, get the upsets since it would probably be unethical to try to predict the whole tournament). The dataset was from [sports-reference.com/cbb/seasons/men/2022-school-stats.html](http://sports-reference.com/cbb/seasons/men/2022-school-stats.html), and I simply copied the entire text. All credit goes to them of course. The file is titled **2021-22\_reg\_season\_totals.csv**.