

1. The project I am pursuing is to determine the improvement or regression of NBA players during the 2017-2018 NBA season and determine if they deserve an upgrade or downgrade on their NBA 2K19 rating. NBA 2K19 has already come out but an easily-accessible database was unavailable so at the end I plan on comparing my conclusions to the actual updated rankings in the new video game. This proposal is inspired by my love for the NBA and for the NBA2K video game series.
2. To accomplish this I will be using and merging the following datasets, the first from a link on operationsports.com/forums/ and the other from basketball-reference.com

| Name | Size | Location | Format | Access method |
|-----------------------------|---------|---|--------|----------------------------|
| NBA 2k18 All Players Master | 1.20 MB | https://www.dropbox.com/s/js2hqcd152mq4kh/NBA%202k18%20All%20Players%20Master.xls?dl=0 | Csv | Excel file on local device |
| 2017-18 NBA data | 1.96 MB | https://www.basketball-reference.com/leagues/NBA_2018_player_game.html | html | Html file on local device. |

3. The 2k18 dataset contains several duplicates as well as players who are not active in the league, removing either before or during the merge will be necessary. In the 2017-18 statistical dataset, there are also duplicates as it has different entries for a player if they played on multiple teams during the season. These duplicates will also have to be taken care of before merging. During the merging, the merge will occur on player names as this is the best identifier, the merge will return a new dataframe with the player's stats and 2k ratings. This would allow me to compare their ratings from the video game and their real-world statistics to see if they deserve to be upgraded or downgraded in the next video game. I will likely have to write a function to define a upgrade and downgrade statistically and even how big of an upgrade/downgrade should occur. From the function, I could write a new column to show the decided upgrades and downgrades.
4. I'm not 100% set on the visualization I will use but as of now I would like to use either a bar chart of the resulting column I create to show the numbers of upgrades and downgrades. Or perhaps a scatter plot to show the size (numerically) of the upgrade/downgrade for each player based on their performance.