**Executive Summary**

*This section provides an overview to the project. It should briefly touch on the motivation, data question, data to be used, along with any known assumptions and challenges.*

While the internet is replete with fantasy football research resources, there are very few developed for fantasy basketball. This is in part because fantasy basketball is typically played in a *category* format, which is more strategically flexible—and therefore more analytically complex—than *points* formats that characterize fantasy football. I therefore want to create a fantasy basketball research tool that draws from up-to-date player game logs in order to analyze player contributions to category fantasy leagues. The primary challenges are connecting to updateable data and normalizing the statistical categories in fantasy basketball so that player value can be compared.

**Motivation**

*Here you will go into more detail about why you have chosen this project.*

I play a lot of fantasy basketball—and fantasy sports in general. Sports in general function as a sort of statistical petri dish, and sports enthusiasts often distill somewhat ridiculous amounts of data on a regular basis. The brilliance of fantasy sports lies in giving the sports enthusiast an outlet for their sports knowledge via competitive decision-making. Within this context, an impressive range of political and economic scenarios arise (for instance, negotiating league rules, negotiating trades, and weighing team need against overall player value in a draft), which lend to a richer game experience that goes beyond the brutish, beer-and-brat-bro culture typically associated with fandom.

The structure of fantasy basketball category leagues especially interests me here, over against fantasy football leagues or fantasy basketball points leagues. Fantasy basketball leagues are typically played in a 9-category format, where the categories are:

1. Field goal percentage (non-free-throw shots made / non-free-throw shots attempted)
2. Free throw percentage (free throws made / free throws attempted)
3. Points scored
4. Three-pointers made
5. Rebounds
6. Assists
7. Steals
8. Blocks
9. Turnovers (you want the fewer of these)

In the typical 9-category league, one drafts a team of 13 or 14 players (with certain positional requirements based off a typical basketball roster construction). Each week, those players collect statistics in the categories listed above, while an opponent’s roster also acquires statistics in those categories. If your team has better statistics in 5 or more categories at the end of the week, then your team wins; otherwise, you lose or tie (though ties are rare, since they require tying in a specific statistical category and then splitting the remaining categories).

This structure means that the fantasy-basketball player has a lot of freedom in building a winning roster. While a roster full of famous high scorers is often desirable, it’s not the only way to win. One could build a team that regularly wins rebounds, assists, steals, blocks, and turnovers, and that might get you to a championship. For similar reasons, trades in category leagues become more a negotiation of which categories are important to each player—a “trash and treasure” conversation—rather than simply two people guessing about future value.

All of this means, however, that more sophisticated tools are needed for comparing player values and category contributions. In fantasy football, you can say, “This guy averages 15 fantasy points per week, and that guy averages 13 fantasy points per week,” and barring some weighing of positional value, the decision-making isn’t extensive. In fantasy basketball, on the other hand, one has to determine whether one player’s 1.6 steals per game are more valuable to your team than another player’s 8.5 rebounds per game. And how does one make an overall determination across categories so that they can say, “Player X is better than Player Y”? Beyond that, if one wants to add a player who can contribute in one or more specific categories, how can they focus just on those categories to determine player value while ignoring (“punting”) the other categories?

My project therefore attempts to provide a dashboard research tool for 9-category fantasy basketball. There are very few tools that currently offer these resources, and the ones that do are primarily in list/grid format without much graphical insight. Ideally, this project will serve as a prototype for a future app or website.

**Data Question**

*Present your question. Feel free to include any research/articles that are relevant or show where others have attempted to answer this question.*

Assuming a 9-category fantasy basketball playing format, which players have the best overall value for the entire season, the last month, and the last week? Which players offer the most value for the category or categories that a user is most interested in—again, over the full season, the last month, and the last week?

[Basketball Monster](https://basketballmonster.com/playerrankings.aspx)

[Hashtag Basketball](https://hashtagbasketball.com/fantasy-basketball-rankings)

**Minimum Viable Product (MVP)**

*Define your MVP. This should be a description of what your final capstone will look like, including visualizations, how the analysis will be presented, who the intended audience is, etc.*

The intended audience is the 9-category fantasy-basketball player who cares enough to research player values rather than just pick their favorite players. This capstone should offer the user a dashboard or collection of dashboards (like a Tableau Storyboard) with which they can research player trends and values across the 9 categories. This dashboard will normalize player statistics in the various categories by determining each player’s respective z-score contribution to each category (on average).

Each category will have its own dashboard with the same structure: two graphs on the top half and a player list/grid on the bottom half. The list/grid on the bottom half will sort (descending) by the dashboard category but will list all statistics and overall value for those players listed. If the dashboard is focused on assists, then the grid will sort by top assist players, but it will also show those players’ contributions in the other 8 categories. In the top left, the user will see a bar chart depicting assist averages for selectable groups of players for comparison purposes. On the top right, the user will see a time-series chart depicting selectable players’ assists over a time range. The user will be able to filter/slice this whole dashboard by time range.

Again, the idea is to develop either one dashboard like this for each category or otherwise add some feature to the dashboard that allows the user to switch the category of interest. This would be like a slicer that works on variables. Finally, I would have an option for looking at overall player values aggregated by average z-score over all categories.

Finally, if possible (though this is no longer *minimum* viable product), I would like to build in a category “punt” option to change player values by disregarding certain categories. This feature is important for anyone who plays category fantasy basketball. Among the most important examples in Giannis Antetokounmpo, who, when evaluated over all categories, is the #17 fantasy player this year; however, if you disregard his abysmal free throw percentage, he is #2.

**Schedule (through 6/29/21)**

1. Get the Data (5/22/21)
2. Clean & Explore the Data (5/29/21)
3. Create Presentation of your Analysis (6/12/21)

* Should be a presentation, but could include a Jupyter Notebook or dashboard in Excel, Tableau, or PowerBI

1. Internal demos (6/26/21)
2. Demo Day!! (6/29/21)

**Data Sources**

*Document the data you use and the source of that data*

[NBA.com Traditional Player Stats](https://www.nba.com/stats/players/traditional/?sort=PTS&dir=-1)

[NBA Data Retriever](https://chrome.google.com/webstore/detail/nba-data-retriever/cibebblabkdibhnidfnipfnjkfbcmeha?hl=en)

[Basketball Reference](https://www.basketball-reference.com/)

[nbastatR](https://www.rdocumentation.org/packages/nbastatR/versions/0.1.110202031)

**Known Issues and Challenges**

*Explain any anticipated challenges with your project, and your plan for managing them. Be sure to include:*

* *If you need to request data or an api key*
* *Based on your data sources, known data cleaning steps*

Basketball Reference looks pretty web-scraping-averse, but I have at least three other avenues for collecting the data. nbastatR is my fallback option. I have already used it for personal fantasy basketball analysis in R; however, I would have to get the data in R and then transfer it over to Python. That might keep the dashboard from automatically updating with current stats the way that I want it to.

One of the challenges that I need to think through is normalizing percentage categories. As an example, a player who averages 90% free throws on 1 free throw per game should not have as high of a free throw value as a player who averages 84% on 6 free throws per game. The player who shoots more free throws affects the fantasy team more overall. This means that calculating a z-score for these percentage categories depends upon both percentage and volume.