Fantasy Football Draft Computer

September 4th, 2023

Jason A Hill

**Executive Summary:** 

I am using what I know and can expand on spreadsheets to create a data-driven Fantasy Football draft strategy for my family's Fantasy Football league.

### Introduction:

I recently joined my family's fantasy football league. Last year I joined the same league, but I was unable to draft my team due to a schedule conflict. This year I have the same problem, and I wanted to give my auto draft a data-driven strategy.

I decided to create a data project to build me a viable draft list so the auto-draft could be as effective as possible. There are two phases to this project.

Phase one is to create the draft order before draft day on August 29th. The second phase will track the progress of the resulting draft through the fantasy season and compare the result with a control team. Since the project started on August 22nd, I had limited time to complete phase one, but it was useful to see how I could work with a deadline like this.

#### **Data Collection:**

My initial methodology was to use last year's NFL player statistics with their average draft rankings. I assumed this information would be fairly easy to gather, but it turned out to be a bit of both easy and more difficult data to capture.

All of the data was sourced from sports and fantasy-related websites. In some cases, I copy and paste the data into Google Sheets directly; in other cases, I web scrape the data using a simple HTML code to copy the data from tables embedded in the webpage.

My league is hosted on Yahoo Sports, and our scoring system is set to the "half PPR" scoring system. The league has 12 teams, and the roster is set to: QB, WR, WR, RB, RB, TE, W/R/T, K, DEF, BN, BN, BN, BN, BN, BN, IR, IR.

All the sources of data contain statistics and measurements based on past and potential performance in the upcoming NFL season. These statistics and measurements are then combined and extrapolated into categories to compute a final rating.

These categories are weighted to:

Player Ranking: 25%

VORP: 25%

Risk Score: 25% Player Flex: 25%

Player rank is based on a composite score between the ESPN overall rankings and the betiq.teamrankings.com total point projections for the top 300 players based on the half-PPR fantasy scoring system. I ended up not using the ESPN ranking in the computation of the ranking number and just using the potential points ranking based on the betIQ data, but I did compare the ranking I got with the ESPN list to make sure there wasn't too much deviation.

Value added over replacement (VORP) is valuable for assessing a player's relative value compared to replacement-level players at their position. This was a metric that was new to me. I had to read more about it to understand it's significance. But it seems to be almost as important as potential points.

The risk score mainly has to do with the potential of a player missing playing time for various reasons. I decided to use a website for the injury metric, draftsharks.com, and composite that against the player's age to complete my risk metric.

Player flex score gives a score based on the utility of a player starting in multiple positions in a fantasy week lineup. More weight is given to WRs, RBs, and TEs since there are six out of nine spots in the lineup, and any one of them can be an additional RB, WR, or TE. This could mean a lot of points either way, depending on which players you draft and are available.

My sources for the data are the following websites:

ESPN <a href="https://fantasy.espn.com/football/players/projections">https://fantasy.espn.com/football/players/projections</a> (Scraped)

BetIQ <a href="https://betiq.teamrankings.com/fantasy-football/projections/half-ppr/">https://betiq.teamrankings.com/fantasy-football/projections/half-ppr/</a> (Scraped)

Draft Sharks Injury Predictor: <a href="https://www.draftsharks.com/injury-predictor/">https://www.draftsharks.com/injury-predictor/</a> (Scraped)

Google.com: (I googled a lot of player data directly from front page results, like age and position.)

## **Data Cleaning and Processing:**

My data was gathered from a combination of web site tables and manually entering information from Google searches.

The "Player Ranking" consisted of the ESPN Fantasy Rankings by sports writer Tristan H. Cockcroft. This was scraped from the first table on the web page. The main ranking score was calculated by using BetlQ's rankings of potential point totals from the list. This data was also scraped from a table on the points totals page. After the data was saved in a spreadsheet, it needed to be copied into a new list for sorting purposes.

The data for the VORP score was copied from the player data sheet that was scraped from the BetIQ data sheet. I created another table for the lowest potential score by position and cross-referenced the two lists to make a score for each player. Lastly, I sorted the list by the final VORP score, from highest to lowest.

The flex score was done manually by giving weights to each player by position. Using a list of players from a master list, I assigned a score weight to each based on their position and value to a fantasy starting roster.

The risk score was mostly taken from the Draft Sharks website. They have a complete score for the probability of injury. (cite?) I used this list and cross-referenced it with the player's age. These two scores together made up the final risk score.

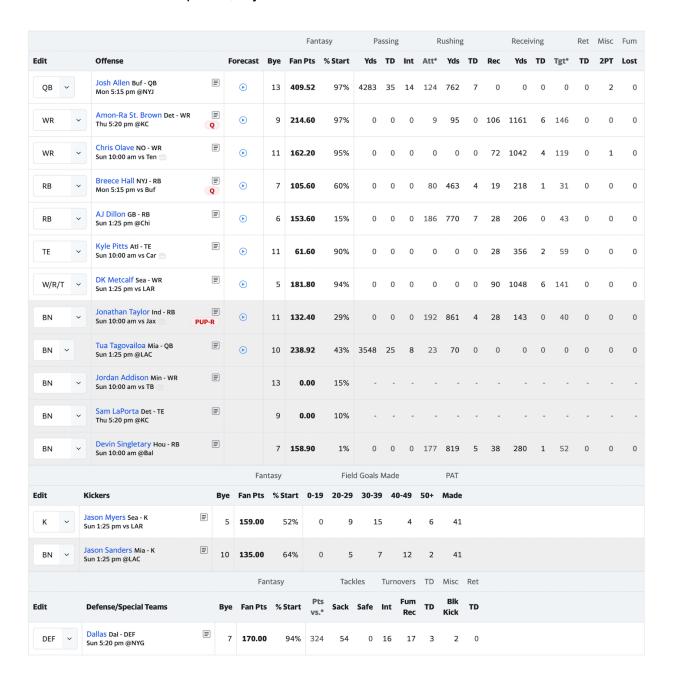
Combining the player lists was the most difficult, as players names were spelled differently and some disagreements on age, position, and injury status made it necessary to sort the list many times. I finally had to create codes for each player to solve many of the tabulation errors. I used vlookup, regenextract, and substitute in Google Sheets to help sort my player name data. I also created new "clean" versions of the list to make the final tabulations easier to sort.

Because some of the data was not directly available, such as player age, I had to Google it and manually enter it for the whole set. I would prefer not to have to do that and rather use some tool for the extraction of this data.

# **Exploratory Data Analysis (EDA):**

After completing the tabulation of the results, I ended up with a draft board. Yahoo Sports allows you to enter a player by ranking for the auto draft feature, so it was an easily transferable thing to go from my spread sheet to the draft list.

After the draft was completed, my team:



## Shortly after the draft, I made one transaction:

Dropped WR Trylon Burkes, added RB Devin Singletary, and needed another running back since Johnathan Taylor was on the PUP list for the first few games.

My draft was graded by Yahoo at A-, and I am projected to finish third in the league by record, so initially the selections seem to be generally favorable.

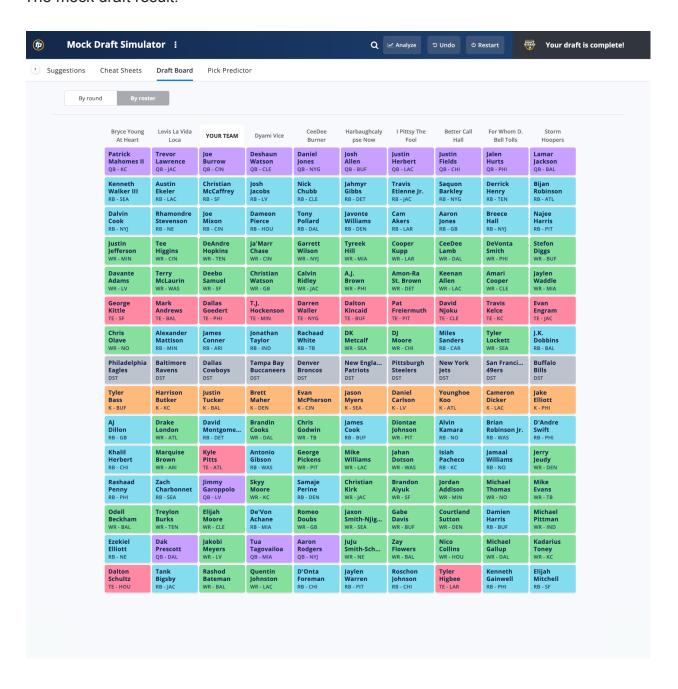
# "Jason's Argonaughts."

League Draft Grades	
JES3 Steelers	A.
OC Quakes	A <sup>+</sup>
Jason's Argonaughts	A <sup>-</sup>
ARS Wildcats	В
The A Rays	C.
Low End Theory	D,
The More You Joe	D.
AJ C's Amazing Team	D
Mother of Drag-routes	D
Sudbury Sour Patch Kids	F

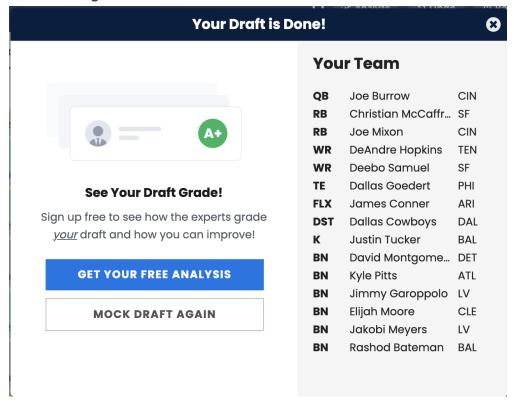
Before I start tracking the progress of this drafted team, I felt it necessary to add a control to use to help compare the results.

I used <a href="https://www.fantasypros.com/">https://www.fantasypros.com/</a> mock draft system to create the "mock." team. I used the same draft board I made for my league, but I drafted in real time against CPU players in a mock 10-team draft.

#### The mock draft result:



The resulting Mock team:



I will add a section on the data sheet for tracking fantasy points.

At the conclusion of the fantasy year, I will do a final report on the results of both teams.