Chris Peña
4/22/2020
SI 330

Final Project Report

Motivation:

Every fantasy football analyst has their own method of ranking players. Their methods are usually fairly secretive, and I believe there is a lot of group-think around the fantasy football world. That being said, I wanted to derive my own tool to help make an independent analysis on players.

This could be a rabbithole of analysis, so I decided to stick to current wide receivers considering it's the biggest group of players for fantasy football. I want to see if their previous year production in receiving yards correlates to their average draft position (ADP) for the upcoming year.

I figured you'd also like a break from reading COVID-19 related projects. Presuming a good portion of the class worked with COVID-19, I wanted to minimize my chances of having similar data / projects as others to avoid the suspicion of collaboration. Lastly, I'm a big fan of fantasy football.

Data Sources:

The two sources used were Fantasy Pros and Fantasy Footballers. Fantasy Pros' projected rankings were used mostly because their rankings resemble other draft rankings I've seen, and because they have a handy option to rank by scoring system (point per reception, half-point per reception, and standard). Fantasy Footballers has a neat and simple table containing the stats for the previous year. Not only is it easy to read on their site, it makes for simple web scraping.

Fantasy Pros offers free .csv downloads of their ADP content, which made that search simple. I opted for web scraping Fantasy Footballers after I had spent more time than I'd like to admit searching for an API that was both free and offered what I was looking for. In addition, their stats table contained lots of other valuable information that could be used for analysis, such as receptions, touchdown totals, games played, and receptions per game.

For this project, I'm only concerned with their yardage totals for last season. While injuries and off-the-field issues occur, and is something that's debated amongst analysts to justify a player's rank, at the end of each football week all that matters (in fantasy) is total points. You can't produce points if you're not on the field. If you're not on the field, you shouldn't be ranked higher than those who are. For this reason I believe ignoring last season's injuries and

suspensions, that is, not using yards per game as a variable, is the best way to determine average draft position.

Fantasy Pros Rankings: https://www.fantasypros.com/nfl/rankings/wr-cheatsheets.php

Fantasy Footballers 2019 Stats: https://fantasyfootballers.org/wr-wide-receiver-nfl-stats/

Data Processing:

The data processing was pretty straight forward. I started with the gathering of last year's receiving yards leaders from Fantasy Footballers. Another reason I opted for their data was the ability to view the receiving yards leaders amongst only the wide receivers. This saved the step of removing tight ends and running backs from the yardage leaders data. I pulled the statistics table using Beautiful Soup into one variable, and isolated the wide receiver names in a separate list. Next was importing the Fantasy Pros rankings list from the .csv file. This was saved as a Pandas DataFrame, and I followed that import by making the Beautiful Soup object also a DataFrame. I added a respective rank column to each DataFrame to compare later.

- A simple problem to address was whether to add the yard and ADP rank column to each DataFrame, or rename the index.
    - I opted to add a new column thinking it would be easier to graph.

- Adding the rank column (or renaming the index) would have labeled the first player element '0' instead of '1'. This might have resulted in that data point being graphed on an axis instead of the graph plane.
    - I created a new row dummy DataFrame, and concatenated it to both DataFrames. After merging the two, I dropped the first dummy row to start the ranks at '1'.
    - https://www.geeksforgeeks.org/add-a-row-at-top-in-pandas-dataframe/

- Saving the player names from the Beautiful Soup object also pulled the player's team name, as both were associated with the table's 'a' tag.
    - I re-saved the name list, but removed every other element with '[0::2]' found on Stackoverflow
    - https://stackoverflow.com/questions/17777482/how-to-remove-every-other-element-of-an-array-in-python-the-inverse-of-np-repe

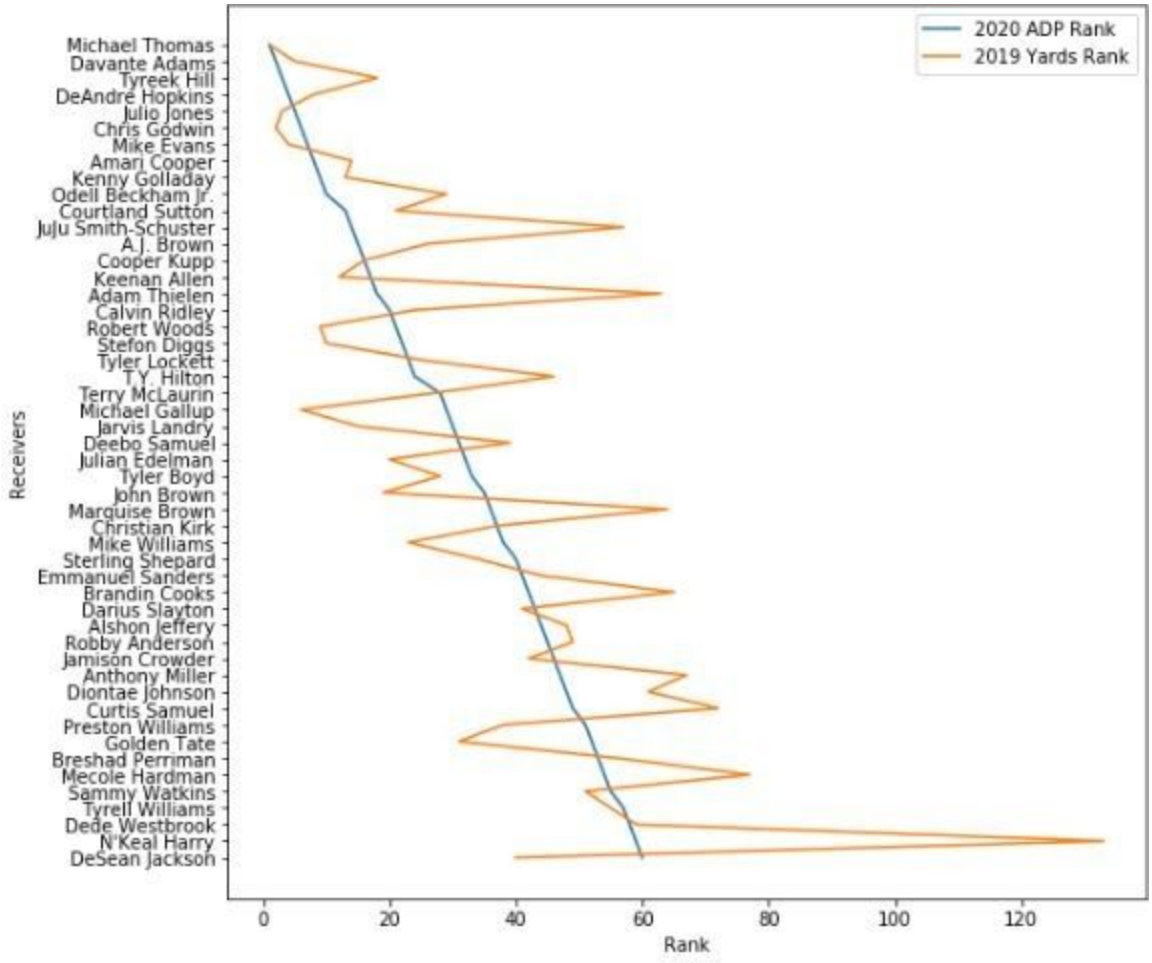
Analysis and Visualization:

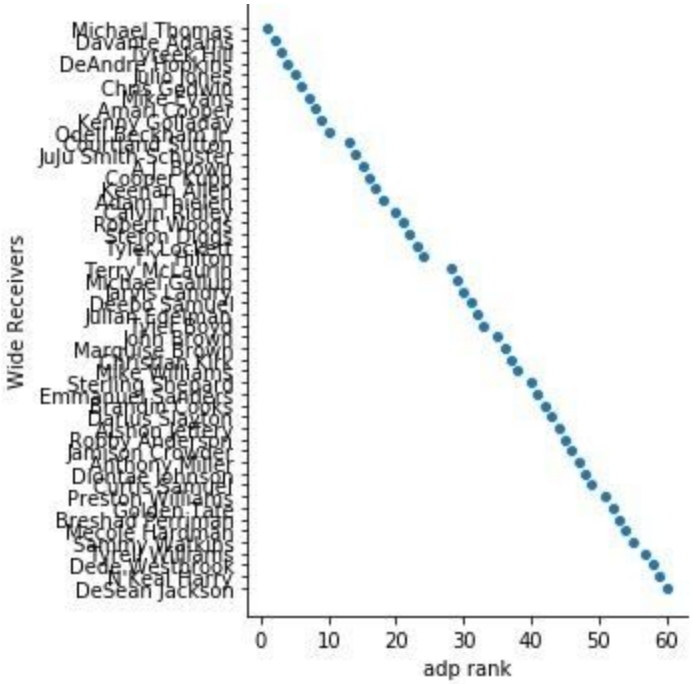
I hypothesized average draft position ranks would largely be based on previous year yardage totals. I thought this because total yards is the big statistic football analysts concern

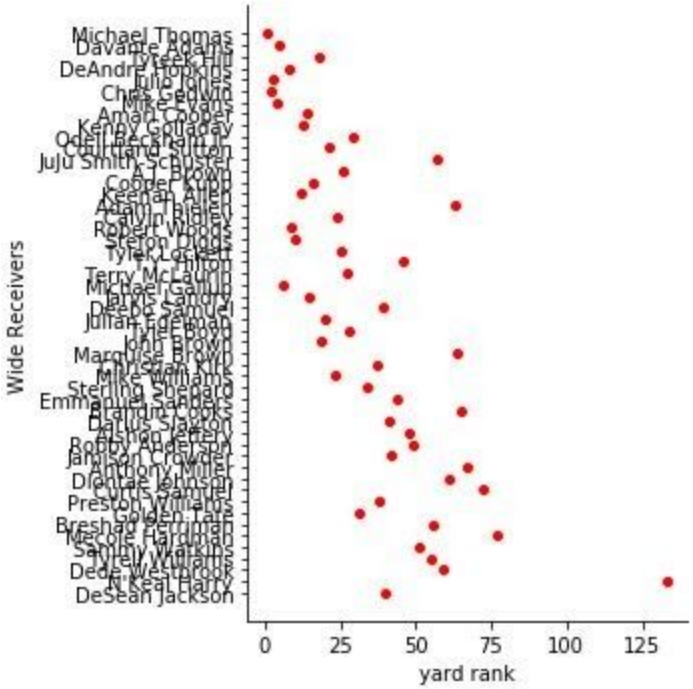
themselves with, along with total touchdowns. Or at least, it's usually the first stat they mention or give the most weight to. This leads me to believe total yards is the biggest contributor to next season's ADP.

I was surprised to see there wasn't as much of a correlation between the two as I might have thought. From the variance we can suggest that certain players are riskier than others, and some might be better value picks (better performance than projected). Players with high/low ADP/yards might be riskier (yardage line to the right of the ADP line), and I might avoid those players in a fantasy football draft. Players such as Davante Adams, Tyreek Hill, Odell Beckham Jr., Adam Thielen, T.Y. Hilton, and Mecole Hardmon.

Players with low/high ADP/yards (yardage line to the left of the ADP line) are now players that I'd like to further analyze for a fantasy football draft. It can help determine both which earlier round WRs to take, as well as which later round WRs to target as 'value' picks. Such players include Julio Jones, Chris Godwin, Robert Woods, Michael Gallup, John Brown, Mike Williams, Preston Williams, and Golden Tate.




Wide Receivers
Michael Thomas
DeSean Jackson
0
10
20
30
40
50
60
adp rank


Wide Receivers
Michael Thomas
DeSean Jackson
0
25
50
75
100
125
yard rank