ANALYSIS

A game plan for the perfect fantasy football draft

(Paul Sancya, Godofredo A. Vásquez, Rick Scuteri/AP)

Fantasy Football ADP

Jackson Savio
CSCI 3656
Numerical Computation
Dr. Osita Onyejekwe

ABSTRACT:

This project explores the application of numerical computation techniques to optimize personal draft strategies in fantasy football. Utilizing a comprehensive fantasy football dataset from the years 2017-2021 that contains player performance metrics from each year, I developed a systematic approach to generate a customized draft order tailored to match real-world draft projections created by professionals. The main priority of this project is to develop a concise and accurate projection of fantasy football rankings in order to improve on existing models. Numerical computation methods used include polynomial regression, data cleaning, weighted scoring and interpolation, each utilized in order to manipulate the previous year's fantasy points in order to better estimate their projected fantasy points for future years. After utilizing these methods, the draft projections were mostly consistent with future ADP projections designed by professionals with small alterations due to each having different ways to manipulate the data. The main differences between this project and real-world projections are rooted strictly in qualitative data. This project demonstrates the practical utility of numerical methods in real world model building, bridging the gap between statistical modeling and strategy games.

Table of Contents

Introduction:

My interest in exploring Fantasy Football draft projections in this project stems directly from my long standing personal interest with the game and a strong belief that the draft plays a pivotal role in overall team performance. While partially rooted in data analysis, current day draft projections vary widely in accuracy depending on the source of the projection. These inconsistencies introduce uncertainty and can significantly affect outcomes for fantasy managers. My goal is to create a definitive example of a strictly analytical fantasy draft. This project, therefore, attempts to hone in on what indicators exist to improve projections for future years through strictly data manipulation and analytics, avoiding any personal bias on what makes for a good fantasy football player. By focusing on historical data of previous years, I gain access to complete datasets that are well-suited for manipulation and evaluation.. This specific scope of years, 2017 to 2021, was selected because 2017 marked my initial involvement with fantasy football, making it a logical starting point for the beginning of analysis.

Background:

Fantasy football is a recreational game in which players act as virtual team managers for their own NFL team, drafting players to assemble a roster that earns points based on each player's real-world performance. The game originated in 1962 when a group of friends created an on-paper fantasy league that manually tracked player statistics with a personal scoring system created by the owner of the league. As the game grew in popularity, a standard ruleset and format was developed to allow for consistency between leagues and to increase overall fairness. In a standard league, each team drafts players from all offensive positions and a single defense. The draft itself is crucial for creating your team, as besides trading in between teams this process essentially determines the winner of the league. This structure creates a need for solid draft performance if a player wants to compete to win the league and is consistently the primary decider of fantasy performance.

Since this game was created based upon a real-world sport, there are plenty of errors that create discrepancies in data collection and could cause incorrect analysis for pre-season draft making. Injuries are inherently random and will skew data to lead to increased variance in ADP if a player was injured in the previous season and did not have any data to view (“Average Draft Position (ADP).”). Personal player decisions can be made that skew data such as contract holdouts, random game absences, or ejections in game. How this should be explored in data collection is an important thing to consider in order to improve draft projections.

Data Collection and Concerns:

This data was collected independently from a source online and separately implemented into CSV files with full statistics for each player that recorded playtime on the season. Each CSV file is focused on one specific year and does not include any qualitative data. An example of the data set is pictured below, including many rows including yardage, touchdowns and overall fantasy points (Dominguez):

Player	Tm	Pos	Age	G	GS	Tgt	Rec	PassingYds	PassingTD	PassingAtt	RushingYds	RushingTD	RushingAtt	ReceivingYds	ReceivingTD	FantasyPoints
Derrick Henry	TEN	RB	26	16	16	31	19	0	0	0	2027	17	378	114	0	314
Alvin Kamara	NOR	RB	25	15	10	107	83	0	0	0	932	16	187	756	5	295
Dalvin Cook	MIN	RB	25	14	14	54	44	0	0	0	1557	16	312	361	1	294
Travis Kelce	KAN	TE	31	15	15	145	105	4	0	2	0	0	0	1416	11	208
Davante Adams	GNB	WR	28	14	14	149	115	0	0	0	0	0	0	1374	18	243

To improve and clean up this data set in order to create the best draft predictions, specific concerns must be handled that could hinder data manipulation:

- ➢ How will different positions be considered for data manipulation?

This specific analysis, moreover, will focus specifically on the FLEX positions: running back, wide receiver, and tight end. This decision was made as these are the most crucial positions in fantasy drafting as they have more than one spot to fill on the roster. Conventionally during the draft, non-FLEX positions typically have less draft priority than FLEX positions (Staff). Although certain quarterbacks may carry high ADP, their overall positional depth and replaceability reduce their value compared to FLEX players, which is not considered under the dataset. Due to the high variance from the position due to non-data driven reasons, non-FLEX players will be exempt from data manipulation for my analytical analysis.

- ➢ What new variables should I introduce to improve accuracy?

Through rigorous testing and manipulation, I decided to introduce both the total yards and total touchdowns variables as each touchdown and yard type for FLEX players are considered equally for both rushing and receiving positions. This decision was made in order to more easily compare these players in polynomial regression and linear interpolation.

- ➢ How should injuries and holdouts be handled to improve draft accuracy?

Since mid-season injuries seriously impact end of season statistics, which encompasses the entirety of these datasets, we must solve this issue before moving forward with methodology. My solution to this issue was quite simple. Take the average fantasy score per game by dividing the ‘FantasyPoints’ and ‘G’ columns and multiplying by the total number of games per season -

sixteen. This both improves consistency by removing injury time and also accounts for any potential real-world complications that result in a player missing a game. To facilitate this adjustment, I would also not consider players with fewer than three games played to avoid However, this decision was not made without doubts, as this could cause injury-prone players to be drafted higher than they should have because the model does not consider how often the player was injured in their careers or other variations. After thoughtful consideration on whether I should just eliminate all injured players from consideration in this dataset, I decided for more accurate draft predictions, I should implement the average game solution.

After manipulating these alterations, the change was immediately noticeable. Here is the original data set compared to the new altered version:

0	Todd Gurley	383.30
8	Russell Wilson	343.92
1	Le'Veon Bell	341.60
3	Alvin Kamara	312.40
5	DeAndre Hopkins	309.80

Original (2017)

Todd Gurley	2232.533333	408.853333
Le'Veon Bell	2075.733333	364.373333
Antonio Brown	1752.000000	352.342857
DeAndre Hopkins	1469.866667	330.453333
Ezekiel Elliott	2003.200000	325.120000

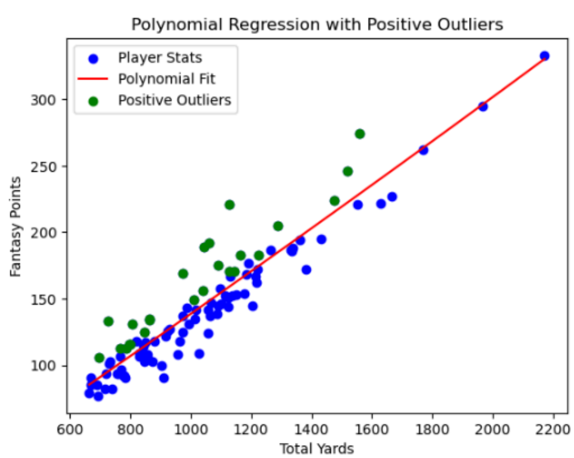
Altered (2017)

The results are immediately noticeable and explains why these results were reached. Antonio Brown was injured for part of the season, and was necessary to get a better view of performance. Similarly, Alvin Kamara scored his point total over the entire season while other players scored similarly while missing games, such as DeAndre Hopkins. Russell Wilson was also removed from the dataset from the original to the altered. If going specifically off of previous season results, Russell Wilson would have the third ADP. Instead, he was actually placed forty-eighth on the average draft board. This should be a clear example of why quarterbacks are difficult to rank analytically due to their average value.

Methodology:

The core objective of this analysis is to develop a personalized draft order for fantasy football by identifying statistical indicators that predict high performance among FLEX-position players. The methodology is divided into four primary steps: data preparation, polynomial regression modeling, interpolation to define expected results, and small adjustments based on other factors that affect ADP. Since we already discussed the data preparation, we will begin with the polynomial regression modeling.

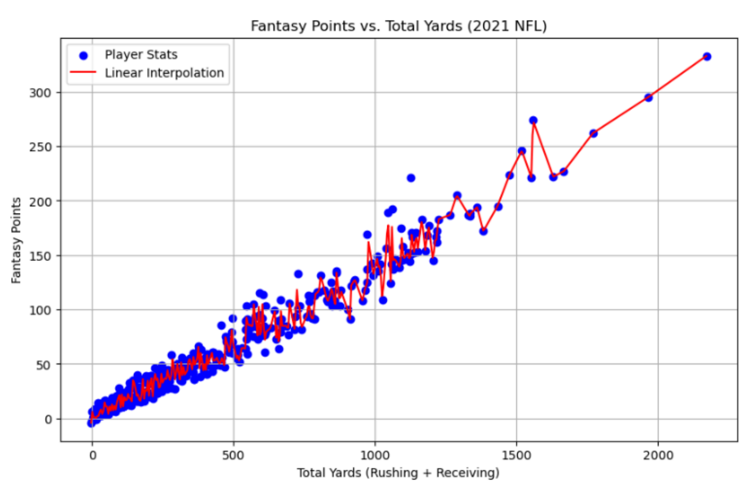
Our use of the polynomial regression model was to observe specific tendencies we can look at to adapt our data set to better implement our model. Our 2021 Data set was utilized in order to create a polynomial regression line that acted as a rough interpolation in order to view outliers. With a threshold value of one standard deviation off the polynomial regression line, we were able to view outliers and reach conclusions based upon this:



Pictured 2021 Fantasy Statistics

What does this tell us? These players are generating more fantasy points than their total yardage alone would predict, based on the model. In fantasy football terms, this could be considered both valuable for a player to be scoring very well or as an indicator that they are performing well above the average standards for this level. Thus, we want to normalize this data to help reduce the impact of outliers. This adopts a precedent set by most professional draft projections to assume regressions in performance from historic outlier years (Staff).

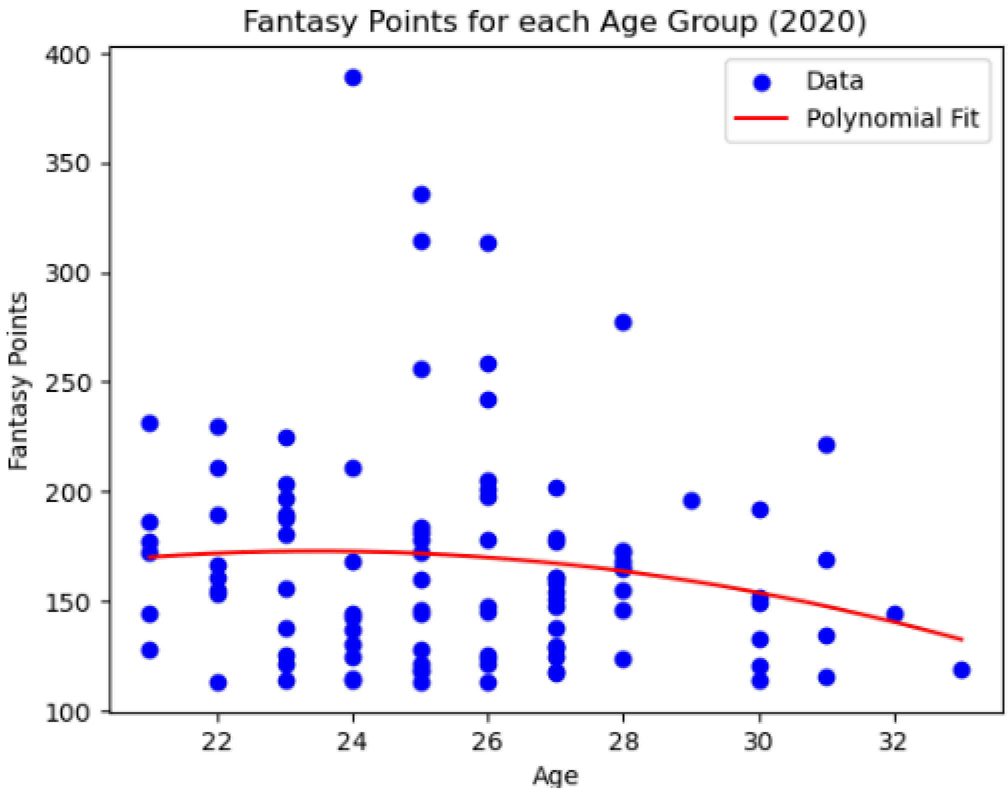
Building on this analysis, the dataset was further manipulated using linear interpolation to compare 2021 player performance trends with data from previous seasons. Specifically, I used the interp1D function from the SciPy library to construct a linear model relating total yards to overall fantasy score. Total yards were selected as the independent variable due to their strong and direct correlation with fantasy points, as well as their relative consistency compared to other performance metrics like touchdowns.



Pictured 2021 Fantasy Statistics

From this point, we are able to manipulate the data in previous years by comparing it to the interpolation of 2021. This is good information we can use to predict the draft position of future players based on previous performance. For the purposes of this test, rookies will be exempt from ADP due to not having previous information on performance. By comparing previous years to the current year, we are able to take the residuals mentioned above and remove them from consideration. To normalize this data, I made a decision to reduce or increase the total fantasy points by a standard deviation of the residual created by the interpolation. This helps reduce the impact of statistical outliers and improves the data collection to a more stable level.

Following the interpolation step, the dataset was further refined using two additional factors: positional importance and age-based metrics. What does each of these refer to in terms of real-world importance for fantasy football drafting? Positional importance refers to the relative value of positions based on the quality of their players in comparison to other positions. For instance, if only ten running backs appear in the top fifty players while forty wide receivers do, the running backs are considered more valuable due to their limited availability and higher demand. Age metrics, on the other hand, account for the typical performance trajectory of players as they age, recognizing that high age is typically a negative indicator for performance in real-world physical sports. Since these are slightly more complicated than just looking directly at statistical data, I had to get creative to solve these issues. To solve the issue of aging players, I decided to implement a polynomial regression of age to see if there existed a pattern of decreasing quality as players aged:



I repeated this process over multiple years and the same pattern repeated itself: the players improved or stagnated in performance until they were around twenty-six years of age, and then steadily decreased in value with a sharp drop off after around nine years. From this pattern, I decided to implement a linear decay starting at twenty-six years of age. Since this appears to be a very minimal amount, we will simply use exponential decay after the average age that players start to regress in performance across each year. Since this is a very small input, we will have a very low initial decay value, set at 0.01.

The final change we will make to the average draft position is by implementing a multiplier based on the commonality of these players capped at lowering player value by five percent to avoid massive changes in ADP. The code for this is displayed below:

```
position_counts = df_2020['Pos'].value_counts()

relative_representation = (position_counts / 100).sort_values(ascending=False)

positional_value = 1 / (relative_representation / relative_representation.mean())

min_val = positional_value.min()
max_val = positional_value.max()

scaled = (positional_value - min_val) / (max_val - min_val)
positional_value = 0.95 + scaled * (1.00 - 0.95)
```

Overall, the model should demonstrate that specific alterations can effectively improve draft decision-making by identifying undervalued and overvalued players. By determining what deviations to make that will improve the model, the adjusted rankings should be able to provide a solid draft strategy that should be very similar to professional draft predictions.

Conclusions and Results:

To evaluate the effectiveness of the model developed in this project, I can compare the projected draft rankings generated through our methodology to the real world ADP of the fantasy football players. This comparison serves as a conclusive test to see how this model compares to professional iterations of the model by directly looking at the largest differences between my model's ADP and the actual ADP. The divergence of the draft strategies that we implemented should indicate whether our development and training of the model was successful.

For the sake of only looking at the most important sections of the fantasy football draft, we will only be observing discrepancies between the top twenty-five players in each draft. The drafts from the real-world will include both rookie players and non-FLEX players which will be ignored in the comparison.

Player	Age
Christian McCaffrey	24
Dalvin Cook	25
Alvin Kamara	25
Derrick Henry	26
Davante Adams	28
Nick Chubb	25
Tyreek Hill	26
Jonathan Taylor	21
Aaron Jones	26
James Robinson	22
David Montgomery	23
Travis Kelce	31
Josh Jacobs	22
Joe Mixon	24
A.J. Brown	23
Calvin Ridley	26
Myles Gaskin	23
Chris Carson	26
Stefon Diggs	27
Will Fuller	26
Antonio Gibson	22
Miles Sanders	23
Justin Jefferson	21
D.K. Metcalf	23
David Johnson	29

My Model (2021)

Christian McCaffrey SF (6) O
Dalvin Cook
Derrick Henry BAL (8)
Alvin Kamara NO (6)
Ezekiel Elliott
Nick Chubb
Aaron Jones Sr. MIN (7)
Davante Adams LAR (11)
Jonathan Taylor IND (14)
Saquon Barkley PHI (14)
Tyreek Hill MIA (14)
Travis Kelce KC (12)
Patrick Mahomes II KC (12)
Najee Harris LAC (7)
Austin Ekeler WAS (9)
Stefon Diggs NE (14)
Antonio Gibson NE (14)
DK Metcalf PIT (7)
Joe Mixon HOU (10)
DeAndre Hopkins BAL (8)
Calvin Ridley TEN (13)
Clyde Edwards-Helaire NO (6)
A.J. Brown PHI (14)
Darren Waller
Justin Jefferson MIN (7)

Real World Model("Average Draft Position (ADP).")

The most immediate comparison that can be made is the complete removal of Ezekiel Elliot from my ADP board. Why did this occur? In the previous year, he was injured and had very low performance on an average basis. My model is skewed toward very consistent performances, however the real-world model predicted Ezekiel Elliot to have a bounce-back year and improve as a result. Interestingly, in the real-world, Ezekiel Elliot continued to regress coming off of injury, and was sharing time with another player on his team as the starter ("Ezekiel Elliott - Los Angeles Chargers Running Back."). Certain players on my draft board have large variance in comparison to the other draft board, including some serious issues that happen when you do not consider any real-world occurrences from year to year. Will Fuller and David Johnson both switched their teams in free agency, and thus lowered their average draft value on the actual draft board. This assessment from the actual draft projections was accurate, as they performed significantly worse than my projections. Although this was a predictable change for the draft, through strictly numerical computation methods and the study of this data set, it was significantly limited from the jump for this project. For the most part, besides some noticeable discrepancies, the rest of the players were generally accurate besides small ADP differences of under five positions. For the rest of the years, comparisons will be shown in the Appendices section of this paper.

The most notable difference for all four of my years I observed the ADP in was Todd Gurley being dropped from #1 in 2019 on my model to #10 on the actual model. This is by far the greatest decrease in value from my projections to the actual average draft position. While this would make zero sense speaking strictly analytically due to him having an incredible performance the previous year, coming in third in OPOY voting, this drop in value is likely due to reports in the fantasy football drafting process that he had serious knee issues that could hinder his ability to play during the season (Patra). This is likely the greatest hindrance to a strictly analytical draft analysis and the greatest flaw to my model, because it does not consider real-world issues such as injuries if they do not occur during the football season, and does not consider likelihood to get reinjured. Barring these limitations and issues, this model has very few discrepancies that cannot be explained by the alterations made intentionally by numerical computation techniques.

This project explored the use of numerical computation to improve fantasy football draft strategies by developing an analytic model created by concepts we learned during class. Through specifically analytical means, my model attempted to improve upon current-day models by implementing techniques such as interpolation and polynomial regression. Through creation of the model, I identified consistent inefficiencies in player valuations which allowed us to improve the model. While my model was not flawless and had some issues based on limitations of my scope, it did notice glaring flaws of some players that were left unnoticed by current models . The methodology outlined here can serve as a foundation for further refinement of my model that could potentially be improved upon to match or exceed opinion-based analysis. The future of this project could be extended to include data from more years to improve interpolation over the years, or to include more data to manipulate, such as team and other free agency indicators that impacted performance as mentioned above. This project was a great lens into the process of using numerical computation techniques to improve upon a model and create a unique piece of technology.

Code:

Github Link with iPYNB file + CSV files + README.md:

https://github.com/js678/numerical-computation-final

References:

"Average Draft Position (ADP)." *FantasyPros*, www.fantasypros.com/nfl/adp. Accessed 4 May 2025.

Dominguez, Ben. "Fantasy-CSV-Data." *GitHub*, github.com/bendominguez0111/fantasy-csv-data/tree/master. Accessed 4 May 2025.

"Ezekiel Elliott - Los Angeles Chargers Running Back." *ESPN*, ESPN Internet Ventures, www.espn.com/nfl/player/_/id/3051392/ezekiel-elliott. Accessed 4 May 2025.

Patra, Kevin. "Gurley's Trainer Confirms 'arthritic Component' to Knee." *NFL.Com*, NFL, 28 Feb. 2023, www.nfl.com/news/gurley-s-trainer-confirms-arthritic-component-to-knee-0ap3000001034072.

Staff, Fantasy. "2025 NFL Draft: Fantasy Football Analysis of Rounds 1-3." *ESPN*, ESPN Internet Ventures, www.espn.com/fantasy/football/story/_/id/44786360/2025-nfl-draft-fantasy-football-projections. Accessed 4 May 2025.

Appendices:

2020:

Unnamed: 0		
0	Christian McCaffrey	424.609352
6	Michael Thomas	321.304905
5	Dalvin Cook	291.272277
14	Chris Godwin	272.228701
3	Aaron Jones	270.035725
4	Ezekiel Elliott	269.785460
2	Derrick Henry	269.505590
9	Austin Ekeler	267.203763
20	Saquon Barkley	264.242352
40	Alvin Kamara	243.321022
32	Mike Evans	236.881901
28	Leonard Fournette	236.312650
26	Julio Jones	235.140109
34	DeAndre Hopkins	234.866802
73	Davante Adams	231.754389
25	Cooper Kupp	224.368157
15	George Kittle	220.283786
8	Nick Chubb	215.761071
36	Keenan Allen	213.838347
7	Travis Kelce	211.659789
38	Allen Robinson	209.841613
57	D.J. Moore	209.650384
72	Tyreek Hill	208.345125
17	Chris Carson	206.899169
54	Michael Gallup	204.895871

Christian McCaffrey SF (11)
Saquon Barkley PHI (9)
Ezekiel Elliott
Derrick Henry BAL (7)
Dalvin Cook
Michael Thomas
Alvin Kamara NO (6)
Nick Chubb
Clyde Edwards-Helaire NO (6)
Joe Mixon HOU (8)
Davante Adams LAR (9)
Josh Jacobs GB (5)
Tyreek Hill MIA (7)
Patrick Mahomes II KC (10)
Aaron Jones Sr. MIN (7)
Julio Jones
Lamar Jackson BAL (7)
Kenyan Drake
Miles Sanders DAL (10)

2019:

0	Todd Gurley	369.576224
1	Saquon Barkley	346.781178
2	Christian McCaffrey	344.986994
3	Alvin Kamara	327.965944
12	Melvin Gordon	312.571236
6	Ezekiel Elliott	308.313287
13	James Conner	297.367003
9	Davante Adams	296.136037
22	Kareem Hunt	292.682831
5	Tyreek Hill	281.750382
10	DeAndre Hopkins	281.292889
8	Antonio Brown	280.802524
18	Michael Thomas	267.190077
11	Julio Jones	266.111551
40	Odell Beckham	256.728130
21	JuJu Smith-Schuster	255.361886
7	Travis Kelce	250.343941
17	Adam Thielen	249.983367
23	Joe Mixon	240.414498
16	Zach Ertz	239.120651
15	Mike Evans	237.939055
32	Stefon Diggs	235.606497
27	James White	229.516669
131	Cooper Kupp	224.583282
14	George Kittle	223.037782

PLAYER Team (Bye)
Saquon Barkley PHI (10)
Christian McCaffrey SF (4)
Alvin Kamara NO (9)
Ezekiel Elliott
DeAndre Hopkins BAL (8)
James Conner ARI (12)
Le'Veon Bell
Davante Adams LAR (9)
David Johnson
Nick Chubb
Todd Gurley II
Patrick Mahomes II KC (12)
Michael Thomas
Julio Jones
Dalvin Cook
Tyreek Hill MIA (5)
JuJu Smith-Schuster KC (12)
Odell Beckham Jr.

2018

0	Todd Gurley	367.449427
1	Le'Veon Bell	318.135561
6	Antonio Brown	293.318542
17	Ezekiel Elliott	289.596247
5	DeAndre Hopkins	284.700965
3	Alvin Kamara	277.208510
2	Kareem Hunt	260.457796
217	Odell Beckham	252.295974
12	Leonard Fournette	248.891114
4	Melvin Gordon	248.522765
13	Keenan Allen	235.554231
7	Mark Ingram	229.514687
202	Dalvin Cook	227.735473
10	Rob Gronkowski	220.831476
24	Michael Thomas	219.206591
33	Jarvis Landry	218.436270
14	Tyreek Hill	218.191970
9	LeSean McCoy	214.155224
30	Davante Adams	213.061714
11	Travis Kelce	212.632740
16	Julio Jones	204.587228
25	Larry Fitzgerald	199.154345
104	Chris Thompson	198.693926
43	Christian McCaffrey	197.563263
20	Zach Ertz	197.542015

Todd Gurley II
Le'Veon Bell o
David Johnson
Ezekiel Elliott
Antonio Brown
Saquon Barkley PHI (9)
Alvin Kamara NO (6)
Leonard Fournette
Kareem Hunt KC (12)
DeAndre Hopkins BAL (10)
Melvin Gordon III
Odell Beckham Jr. o
Dalvin Cook
Julio Jones
Christian McCaffrey SF (11)
Michael Thomas
Devonta Freeman
Davante Adams LAR (12)