

Spring 2018

# Predictive Analytics for Fantasy Football: Predicting Player Performance Across the NFL

Jack W. Porter

*University of New Hampshire, Durham, jwp2002@wildcats.unh.edu*

Follow this and additional works at: <https://scholars.unh.edu/honors>



Part of the [Business Analytics Commons](#)

---

## Recommended Citation

Porter, Jack W., "Predictive Analytics for Fantasy Football: Predicting Player Performance Across the NFL" (2018). *Honors Theses and Capstones*. 406.

<https://scholars.unh.edu/honors/406>

This Senior Honors Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Honors Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

Undergraduate Honors Thesis

**Predictive Analytics for Fantasy Football:  
Predicting Player Performance Across the NFL**

By

Jack W. Porter

Peter T. Paul College of Business and Economics  
University of New Hampshire

Advisor:

Dr. Ali Hojjat

Assistant Professor of Decision Sciences

May 2018

## Table of Contents

1	Introduction .....	1
2	Literature Review .....	2
2.1	Fantasy Sports Evolution and New Ideas .....	2
2.2	Legality of Fantasy Sports .....	4
2.3	Analysis and Prediction of Player Performance .....	6
3	Data .....	8
4	Our ARIMA-based Methodology .....	9
5	Results .....	10
5.1	Analysis of Player Performance Metrics .....	10
5.2	Player Rankings – By Position .....	16
5.3	Player Rankings – Overall Top 100.....	17
5.4	Forecast Validation and Accuracy .....	19
6	Limitations/Future Work .....	19
	References.....	21
	Appendix: R Codes.....	23

# 1 Introduction

Fantasy sports are games where participants take on the role of a coach or general manager. Participants build their own teams through a drafting process and compete against the teams other fantasy owners compile. Winners and losers are determined by the real-life statistics of the professional athletes that make up fantasy teams. In recent years, this notion of fantasy sports has become increasingly popular. The majority of this popularity growth can be attributed to technological advancements. The birth of fantasy sports was roughly 40 years ago, where people recorded their own statistics, scores, and league standings on paper (Shipman, 2001) Participants were scarce, but today over 30 million individuals actively participate in fantasy sports. “The web has made the barrier to entry to play [fantasy sports] drop to practically nothing” and it requires a lot less work to participate (Benderoff, 2007). With technology creating easier-to-operate fantasy sports leagues popular, entertainment and technology companies jumped in and began offering fantasy sports applications. Currently, the three major companies involved in fantasy sports (in terms of market share) are Yahoo Inc., ESPN Inc., and CBS Corporation (Peters, 2017). Additionally, massive companies like DraftKings and FanDuel were created for the sole purpose of fantasy sports. Fantasy sports revenues total billions of dollars a year, and that is not including league entry fees, premium cable/internet packages participants buy to watch every game, and the cost of detailed information (Benderoff, 2007).

Fantasy sports began in the 1960s with baseball. Today anything from golf to MMA fighting is available for fantasy sports users to experience of all fantasy sports, football tends to be the most popular. In 2017, 35.6% of fantasy sports participants played football, which is over 20% more than the second biggest segment of the market, baseball (Peters, 2017). This is because there are only 16 games over a 17-week season. It is easy to check/change line-ups, keep up with scores, there is more time (a week) between games and all the games are on the same day.

Fantasy sports have become very important to fans, some argue that fantasy has caused people to become fans of players, rather than teams. Even individuals who are not sports fans are participating because they enjoy the analytical aspect. Some even referring to fantasy football as a “numbers geek world” (Wood, 2012). Fantasy games are available for everyone; participants, advertising companies, major corporations, and start-ups.

The goal of this research is to develop a quantitative method of ranking and listing players in terms of performance. These rankings can then be used to evaluate players prior to and during a

fantasy football draft. To produce these rankings, we develop a methodology for forecasting the performance of each individual player (on different metrics) for the upcoming season (16 games) and use these forecasts to estimate player fantasy football scores for the 2018 season.

More specifically, this work answers the following:

1. In what order should players be drafted in a 2018 fantasy football draft and why?
2. Which players can be expected to perform the best at their given position (Quarterback, Running back, Wide Receiver, Kicker, Team Defense) in 2018, and which players should we expect to perform poorly?

We used a dataset from *ArmchairAnalysis.com*, and used an ARIMA methodology to forecast the fantasy scores of each player one season into the future. The forecast was based on historical individual performance metrics. Our validation revealed that this approach produces forecasts that are, on average, only 3.53% off from the actual fantasy scores of NFL players.

## **2 Literature Review**

Several studies in the literature and ample coverage in the media have been devoted to fantasy sports. Intriguingly, the literature pertains to three distinct subtopics involving fantasy sporting games. The first is the evolution of fantasy sports, discussing where it began and their thoughts of where it should go in the future. The second is the legality of it. In the many years of fantasy sports, never was there a thought of its similar structure to illegal sports gambling. That changed in the early 2000s, and still has an impact on fantasy application providers to this day. The third is the analysis and prediction of player performance, which is directly tied to fantasy scoring, i.e., winning/losing.

### **2.1 Fantasy Sports Evolution and New Ideas**

In the early 2000s fantasy sports began to skyrocket. There were more options and providers of fantasy applications than ever before. This level of growth continued to climb. In the mid to late 2000s, instantaneous communications technology began to allow users to play fantasy sports in real time from almost anywhere (Plimi, 2006). Smartphones were becoming increasingly popular and on-the-go internet access meant people could check, change, even draft line-ups wherever

they were. The mobilization of technology assisted fantasy growth astronomically. It made live drafting and participating in multiple leagues simpler. “As the internet has grown, so has the business [of fantasy sports] ... and that’s no fantasy” (Benderoff, 2007). With this increased simplicity, Plimi (2006) constructed a method for selecting players for a fantasy team prior to and during a fantasy draft. The method included selecting a player, assigning a logical rule to that player defining one or more conditions governing if/when that player will be drafted. This process would then repeat itself throughout the draft. The use of logical conditions in this model was its shortcoming. The pre-draft analysis of players is highly general when draft status is based solely on one or more logical conditions. Especially in this particular model where there was a lack of statistical analysis.

Many believed that to succeed at fantasy sports you had to be a real-life sports fanatic. That belief changed in 2010, when a man named John Rozek, a soft-spoken accountant from Chicago, was named the best Fantasy Football player in the world (McNamara, 2010). Rozek did not attribute his success to being an avid football fan. Instead it was detailed that he was an intellectual who understood how to analyze something and put the pieces together; all that was necessary was moderate knowledge of football. McNamara (2010) detailed that it is not necessarily about who you draft, but who your opponent does not draft. Explaining that to be good at fantasy sports, participants must take advantage of opponent’s mistakes on draft day. This notion of analytics, problem solving, and the finer parts of fantasy sports changed the way people played the game. Especially when Rozek “earned \$25,000 for sitting around in shorts and flip-flops, sipping beer out of a Packers pint glass and pretending to be an NFL general manager.” Talk about a fantasy (McNamara, 2010).

Another recent alteration of fantasy sports, especially in football, are the multitude of scoring options. Originally, fantasy football was scored the same way for all leagues and relatively the same across all fantasy platforms. Today, there are many aspects of scoring. Standard Scoring method, which is the original scoring method (simply no points for player receptions). Within the Standard Scoring method, league participants can even set the number of decimal places they desire. Selection usually ranges from 0-2 decimal places. In addition to standard scoring there is Points Per Reception (PPR), meaning that each catch by a player counts as a certain number of points (usual PPR is 1 or 0.5 points). Besides player scoring, participants can enter one day, one

week, or full season match-ups if they so wish (Shipman, 2001) All these options and differentiators were not available until the late 2000s. Additionally, Shipman (2001) expresses in detail how technology changed the game of fantasy sports:

- Most of the scoring and analysis used to be calculated on paper, by the league commissioner. Now, it is all done via computer programming.
- Player statistics had to be collected and consolidated manually, whereas they are now available on all fantasy platforms.
- Communication that was required in leagues before was astronomical. Now you do not often know the people against whom you are playing.
- The level of commitment and time necessary to participate in fantasy sports used to be much greater.

Lastly, a strong point about the positive domino effect caused by fantasy sports was explained. Fantasy sports blend gamesmanship and spectatorship. The success of fantasy impacts not only the companies directly providing fantasy services, but also the original real sports themselves. Because fantasy teams are comprised of players from multiple real-life teams participants are motivated to watch more games outside of their local team. Fantasy sports are causing sports spectators to become more immersed and engaged in the actual sport itself (Shipman, 2001).

## **2.2 Legality of Fantasy Sports**

Since the boom of fantasy sports in the early 2000s, many have questioned if fantasy sports should be considered illegal sports gambling. It is important to recall that using fantasy sports services were not always free as they are today. When fantasy games were becoming popular, big platform providers would charge a fee to use their service. The winner of the league would walk away with most of those fees, most of the time in the form of prizes (Greenfield, 2003). Now, this is no longer the case. Today, fantasy service providers like ESPN and Yahoo Inc. make their revenues from selling advertisement space on the fantasy service platform. There are still some pay-to-play services, the major ones being DraftKings and FanDuel which charge fees to enter contests and have payouts for a certain threshold of participants. These companies started in the 2010s and are relatively new to the market and have been the center of the talk around illegal gambling concerns.

Greenfield (2003) both refutes and confirms that fantasy sports should be considered gambling. One side likes to describe fantasy sports as hobbies, entertainment, a skill, even a way to alleviate the urge to gamble. People have even claimed “[fantasy sports are] saving me from gambling. You put up X amount of dollars and that’s all you can lose. It’s the best thing since sliced bread”. Others say that fantasy sports expose the youth to gambling, because you don’t have to be an adult to play. Illinois Council on Problem and Compulsive Gambling explained there is no legal grey area. Fantasy sports are wagers, and whenever someone bets anything on the outcome of an event it is wagering, which is illegal involving sports (Greenfield, 2003). There are obviously stances on the argument, but in 2003 there still had not been one prosecution involving fantasy sports.

In 2006, the first substantial lawsuit was filed by Charles Humphrey against a plethora of online fantasy sports providers. These providers included Viacom, CBS, and ESPN (Moorman, 2008). Humphrey claimed that online fantasy providers were engaged in a gambling enterprise that was illegal by state wagering and gambling laws. This was during the time where participants paid providers a fee to use their services and the fees would be distributed to the winner of a league. The lawsuit sought to recover the fees Humphrey paid to the fantasy service providers under *qui tam* laws, which “allow individuals to recoup losses sustained while gambling” and challenged these providers under multiple state statutes. Eventually it was deemed that online fantasy sports leagues are not gambling, and that Humphrey would not be recouped for his losses. According to Moorman (2008) the court could differentiate entry fees and bets/wager and say that the conditionality of the entry fee and the guarantee of prizes/payouts constitute when a contest is gambling or not.

Most recently, the new craze of DraftKings and FanDuel and their “daily fantasy” platforms caused a lot of turmoil in courtrooms. The idea behind their one-day fantasy is a quick contest, that brings in high volumes of participants, with buy-ins and pay-outs for winners, very similar to traditional fantasy play. But, in 2015 daily fantasy sports were constituted as illegal gambling in the state of New York, and both DraftKings and FanDuel were issued cease-and-desist notices (Purdum and Rovell, 2015). New York Attorney General Eric Schneiderman demanded that these two industry giants stop taking wagers/bets from the people of New York. Because of this case other states took notice and began to rethink the legality of daily fantasy sports. Both companies based their businesses around the Unlawful Gambling Enforcement Act of 2006,



defending that they were not accepting wagers based on a contest of chance, explaining that winning and losing depends on a plethora of elements and not pure chance (Purdum and Rovell, 2015). The argument is that if players use all information that is available, perform the analysis, become thinkers, and put pieces together, they can improve their skills. Therefore, improve their chances of success.

### **2.3 Analysis and Prediction of Player Performance**

Participants of fantasy sports realize the importance of accurate projections and player analysis. Deriving key statistics, performance indicators, and information is as precious as gold in fantasy sports (Wood, 2012). Over time, the ability of technology and fantasy participants to predict future player performance has improved in both information and accuracy. Morris et al (2006) developed a model to predict fantasy player performance based on player Performance Prediction Index (PPI), which is made up of all players performance predictions. These PPIs were then updated periodically to match up with gameplay, update weekly in the fantasy football (a game every week). The calculation of PPIs uses historical individual player performance information as well as game play information (e.g., whether it is a home or away game) in a least-squares regression. This study used a total of 14 explanatory input variables in the least-squares regression, most of which pertained to game play rather than individual performance metrics, making the accuracy of player performance projections very questionable. Nevertheless, the PPIs calculated were then utilized to draft fantasy sports teams in fantasy football.

Besides drafting, there are many further important decisions one must make to succeed in fantasy sports. To address some of the challenging post-draft decisions, Ware and Webb (2008) developed a model to evaluate when to add a player to a line-up and when to drop another. The model encompasses both historical and current statistics to predict how many points a given fantasy team will score over a full season. This methodology is then applied to an alter roster, in which the participant may add a new player or change the starting line-up, and a full season score is derived for the altered roster. With the season totals calculated, the difference between the original and altered rosters represents the value of the line-up change, showing the fantasy owner the impact of the addition or subtraction of a player on the team's overall season points. "To successfully manage a fantasy football team, an owner must constantly examine different players

and evaluate how they can contribute to the stats earned by the [overall] fantasy team” (Ware and Webb, 2008). This study also details the steps fantasy managers should follow to use this model successfully:

1. Examine the rankings of all players available (not on a fantasy owners team).
2. Examine the rankings of players on your team and in the starting line-up.
3. Change the starting line-up around to optimize the rankings and overall score of the line-up/team.

The “core [of fantasy football] is to estimate how much a given player will score in each game and over the course of an entire season” (Optimized Financial Systems, 2014). If a fantasy owner can predict future scores, then they obviously put themselves at a great advantage over the competition. A recent study by Optimized Financial Systems (2014) used weekly statistics for all players in the league from the year 2000 through the 2013 NFL season to predict fantasy football scores for Quarterbacks using Vectorial ARMA models. The specific method used was a Python script that migrated and organized the data into a SQLite database. Once the data was cleaned and migrated due to the amount of data points for many of the Quarterbacks, the model had to shift focus. It would analyze quarterback scores over time against specific opponents, then given the performance against each team the QBs could be ranked. To accomplish this a Postulate ARMA script was run in MATLAB to develop base models for each team. After that, additional variables such as sacks, interceptions, and other statistical metrics were added into the model to improve it, yet none were deemed statistically significant. With the base models for each team in hand, individual QB data was added to the model to forecast points and rankings. The results showed that many of the weekly forecasted scores were quite close to the QBs actual average weekly score, whereas the forecasted rankings were not close to the actual rankings. This mismatch shows fantasy owners do not always draft in the order of highest ranked player, even though this ranking of QBs provides fantasy owners a relative scale for comparison prior to a fantasy draft.

In this paper, we use Autoregressive Integrated Moving average (ARIMA) models for forecasting players’ fantasy points for the upcoming season. A reader not familiar with the ARMA modeling approach may refer to Hyndman and Athanasopoulos (2018) for an introduction to this methodology.

There are still voids in the research, discussions, and ideas in the literature that can be built upon. Many individuals research the cases between fantasy providers and those who bring them to court on speculation of illegal gambling. There are frequent attempts to build the next fantasy sports platform, or a game that all consumers will want to experience. Information is gold and possessing the best information is highly important to fantasy sports success. Why not try to provide the best information available? Develop new models to analyze not just QBs or one position, but all positions for upcoming drafts. The ability to predict the future and predict it well is something that millions of fantasy sports diehards would pay top dollar to get their hands on. Proving both consumers and fantasy sports providers with optimal draft rankings and player scoring projections on an annual basis can be used throughout the industry.

In the following two sections we elaborate on our dataset, our use of ARIMA models for forecasting player performance and future fantasy scores, and the R script we developed for this purpose.

### **3 Data**

The dataset used in this study was the NFL data for research, obtained from the website *ArmchairAnalysis.com*. It contains 18 years of complete information about every NFL game/player/play, totaling 4,700 games, 10,000+ players, and 780,000+ plays.

This study focused on the last 18 seasons for:

- Offensive players, for which we used the following metrics:
  - Pass Attempts
  - Pass Completions
  - Pass Yards
  - Quarterback Interceptions (INTs)
  - Touchdown Passes (TD Pass)
  - Rushing Attempts
  - Rushing Yards
  - Rushing Touchdowns (Rush TD)
  - Wide Receiver Targets

- Receptions
- Receiving Yards (Rec. Yards)
- Receiving Touchdowns (Rec. TD)
- Fumbles Lost
- Kickers, for which we used:
  - Points After Touchdowns (PAT)
  - Short Field Goals (Less than 40 yards)
  - Medium Field Goals (40-49 Yards)
  - Long Field Goals (50+ Yards)
- Defense, for which the data provided us with:
  - Defensive Fantasy Points Per Team

The original data came in a multitude of separate tables, which needed to be queried, filtered, and related to one another using Power Query in Microsoft Excel. The purpose of this was to gather only the relevant fantasy football performance metrics for each player in a clean table structure format for time series analysis.

## **4 Our ARIMA-based Methodology**

All players that were examined in this study had played in at least one game in 2017 and played in at least 15 games throughout their career. The threshold of 15 games was chosen to ensure a given player is valuable enough to be selected in a fantasy football draft and that he has a reliable number of data points to be analyzed. If a player had played in the required number of games, his historical performance throughout his career was used to project fantasy scoring metrics (listed in section 4.1) one season (16 games) into the future. Players were then ranked based on their calculated fantasy football scores for the 2018 season. These players were ranked by position and overall.

For this study, we assume each fantasy football scoring metric to be an independent univariate time series, meaning correlation between performance metrics was ignored in this research study.

We developed an R Script to iterate through each performance metric for each individual player and fit the best univariate ARIMA model to predict that specific performance metric of that player in the upcoming season (16 games into the future). With the performance projections

in hand, the R Script then calculated the total fantasy football score for each offensive player for the upcoming season, using a Points Per Reception (PPR) fantasy scoring formula:

$$\begin{aligned} \textit{Offense Fantasy Points} = & (0.4 \times \textit{Pass Yards}) + (-2 \times \textit{INTs}) + (4 \times \textit{TD Pass}) \\ & + (0.1 \times \textit{Rush Yards}) + (6 \times \textit{Rush TD}) + (1 \times \textit{Reception}) \\ & + (0.1 \times \textit{Rec. Yards}) + (6 \times \textit{Rec TD}) + (-2 \times \textit{Fumble Lost}) \end{aligned}$$

Kickers require standard scoring method (as they do not play offense):

$$\textit{Kicker Fantasy Points} = (1 \times \textit{PAT}) + (3 \times \textit{Short FG}) + (4 \times \textit{Medium FG}) + (5 \times \textit{Long FG})$$

Defensive fantasy points were the only metric being projected for defenses, they required no formula. The outputs from the R Script were then made into their own comma separated value (CSV) files. The output files were queried once again in Power Query to include names of all players.

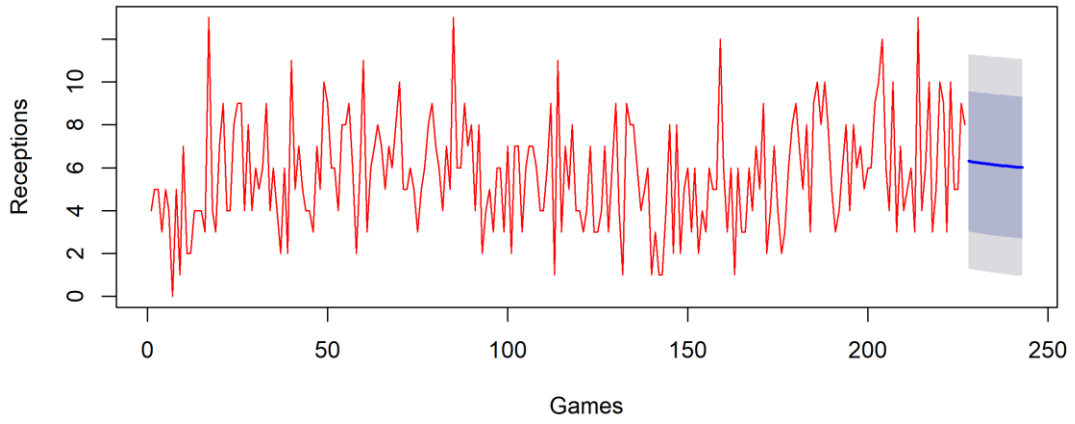
## 5 Results

The results were broken down by position as well as an overall list of the top fantasy football players available in 2018. In addition to fantasy scores, it was important to analyze the projections for each scoring metric forecasted in the ARIMA model. The following charts exhibit the scoring metric for a few sample players over time as well as the one season (16 game) projection period.

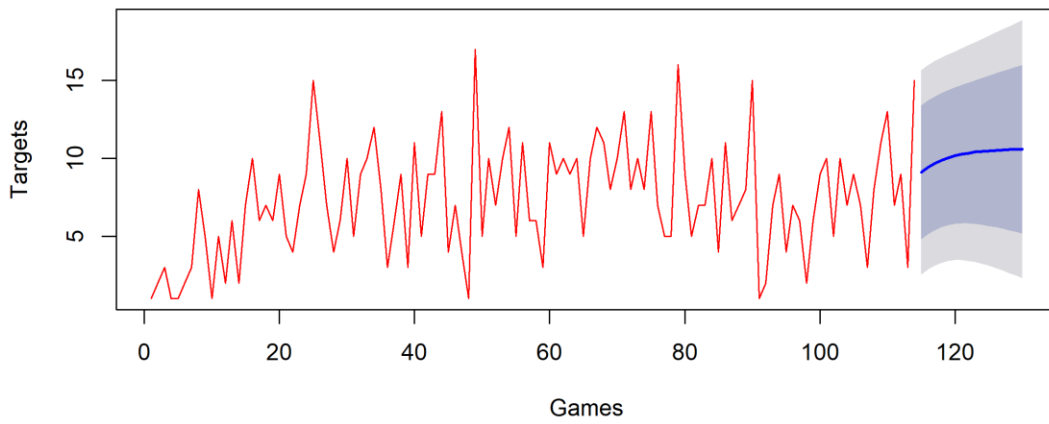
### 5.1 Analysis of Player Performance Metrics

The results for players of various offensive positions and metrics are as follows:

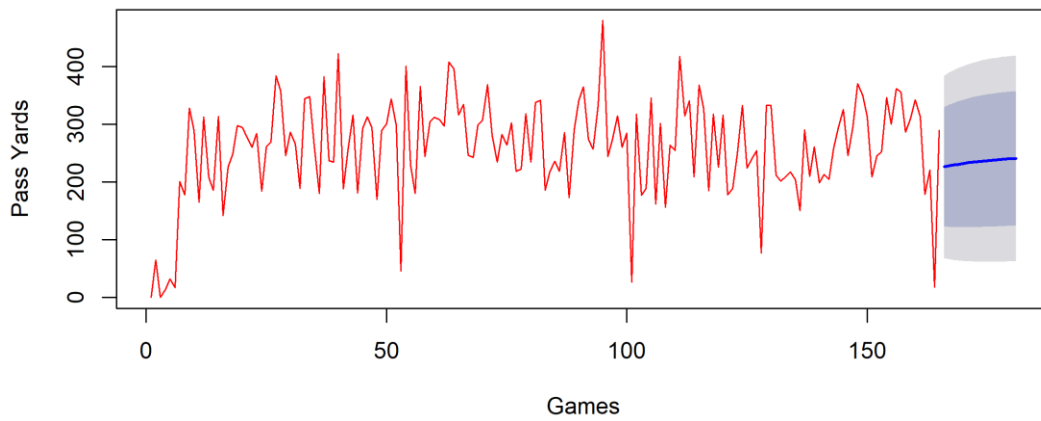
**Larry Fitzgerald Receptions Per Game**



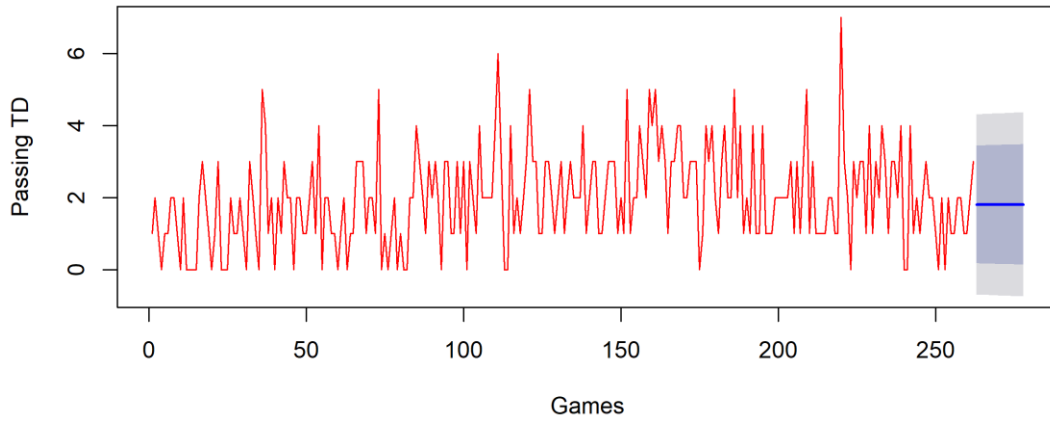
**Rob Gronkowski Targets Per Game**



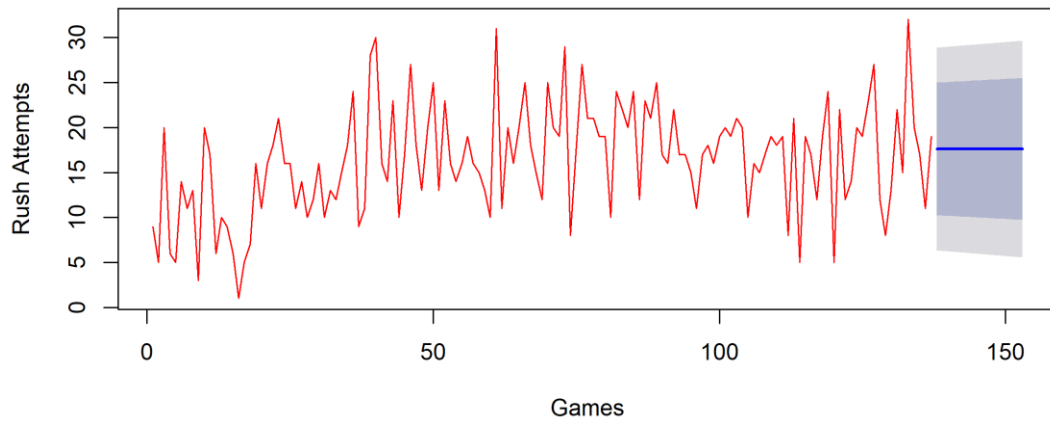
**Aaron Rodgers Passing Yards Per Game**



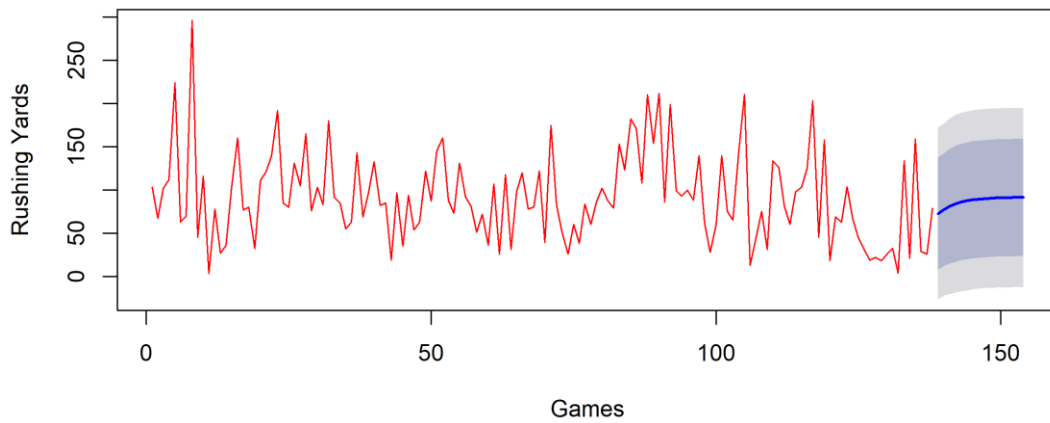
**Drew Brees Passing TD Per Game**



**LeSean McCoy Rush Attempts Per Game**

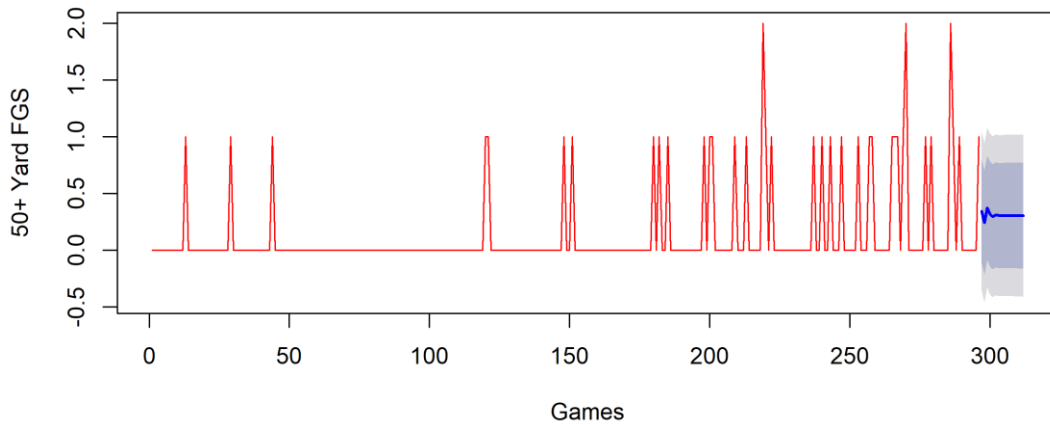


**Adrian Peterson Rushing Yards Per Game**

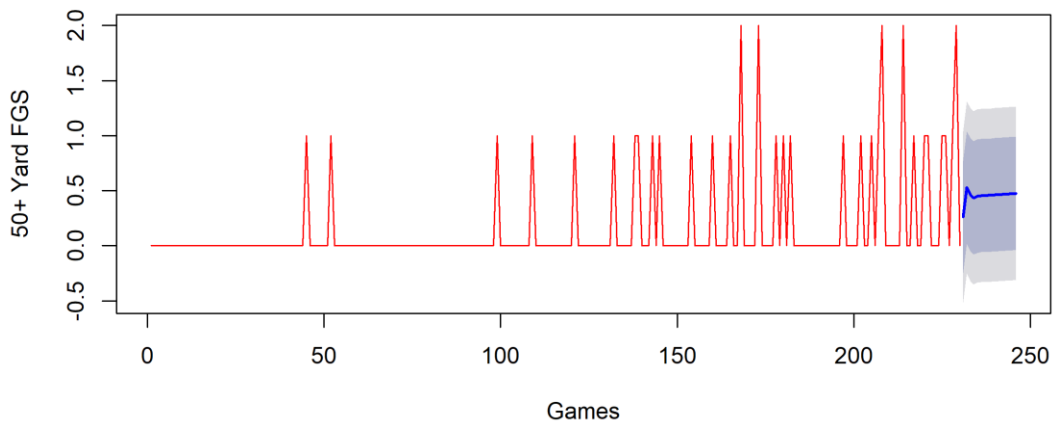


The four ways of scoring as a kicker are also very important, kickers that make the most long-distance field goals are the most desirable, fantasy participants look for kickers with high frequencies of field goals, as well as for kickers who play indoors (weather is less of a factor). Analysis and predictions of the frequency of long-distance field goals for popular fantasy kickers are as follows:

**Adam Vinetieri 50+ Yard FGs Per Game**

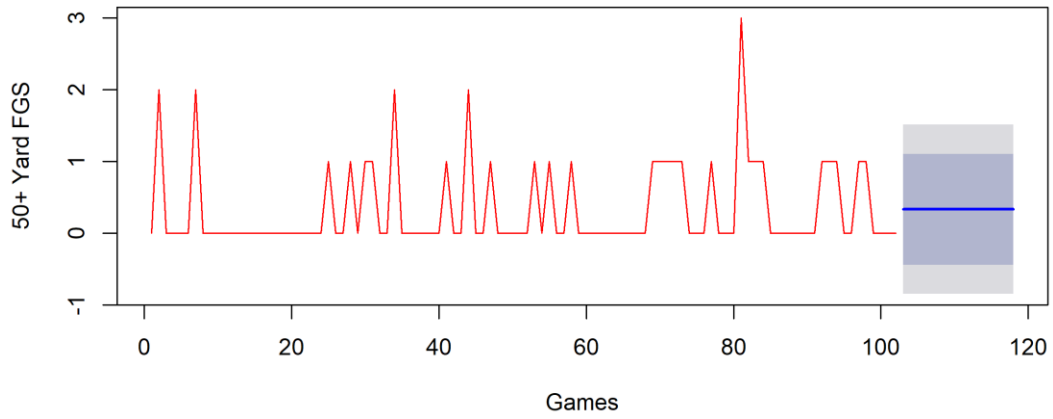


**Matt Bryant 50+ Yard FGs Per Game**

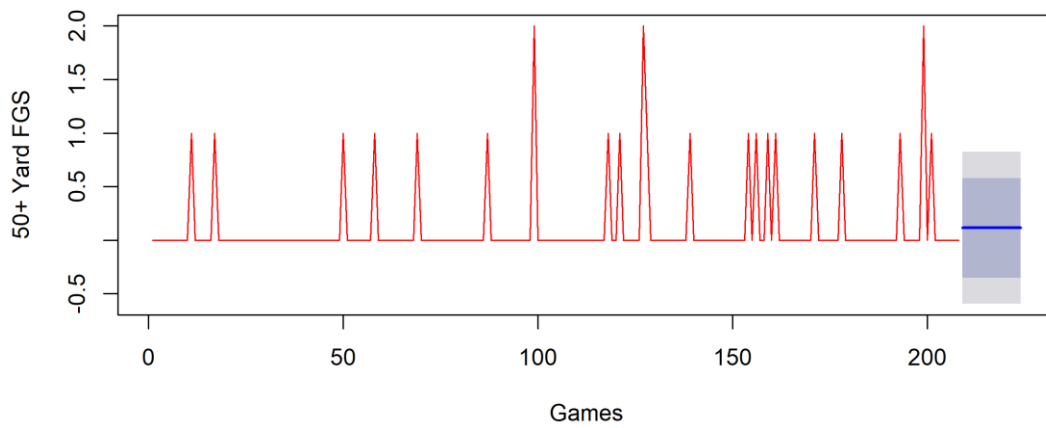




**Justin Tucker 50+ Yard FGs Per Game**

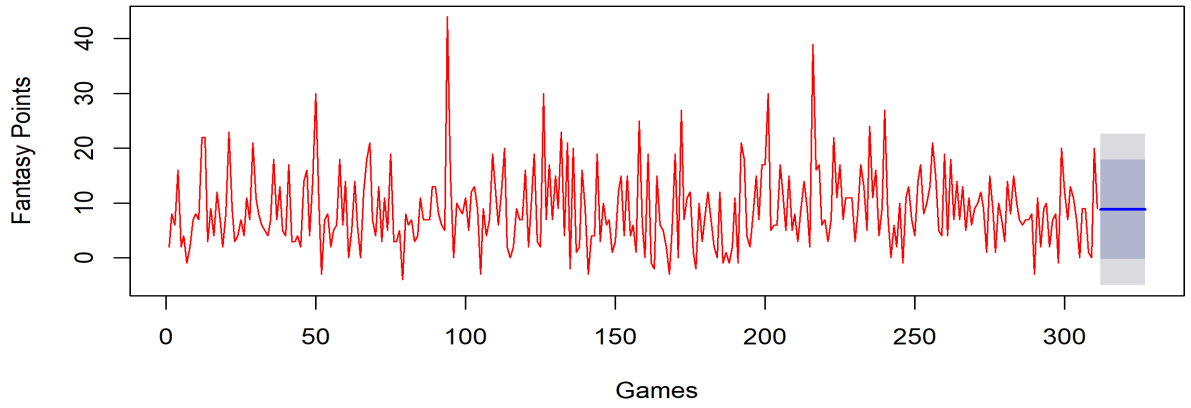


**Stephen Gostkowski 50+ Yard FGs Per Game**

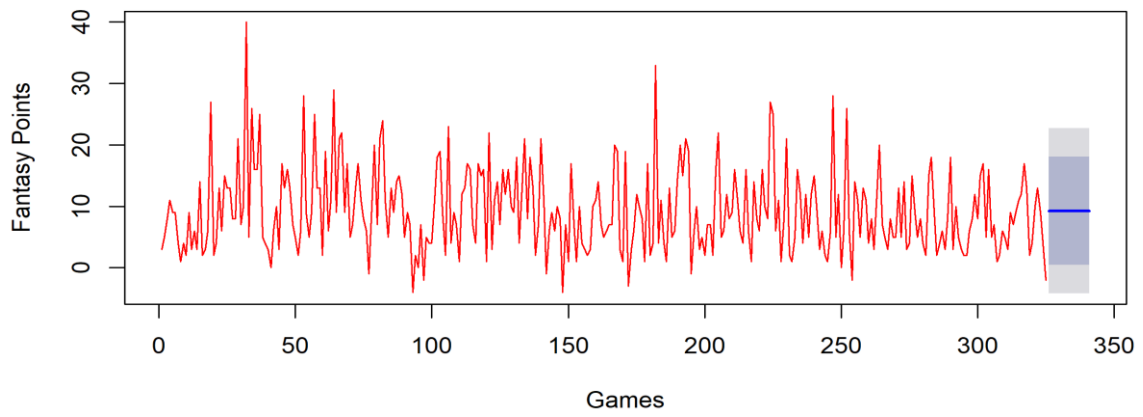


Fantasy defenses can make or break a participant's week. First, they do not score many points, and second, they are the most likely player in a fantasy line-up to post negative points in each week. Analysis and predictions of popular fantasy football defenses are as follows:

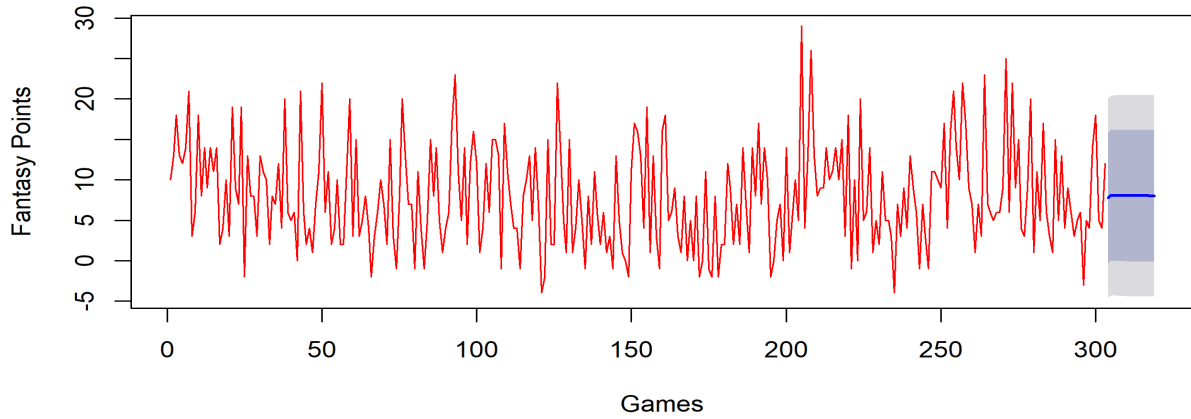
**Seattle Fantasy Points**



**New England Fantasy Points**



**Denver Fantasy Points**



## 5.2 Player Rankings – By Position

One of the keys of this research and analysis is to see which players can be expected to perform well at their given position in the 2018 season. If a player can be expected to perform well, then it would be in the best interest of a fantasy football participant to select them early on in a fantasy draft.

We break down players into the top five for each position, offensive players within the top five thresholds for each position should receive serious consideration when drafting players in early fantasy draft rounds (rounds 1-3). The top five ranking for each position in 2018, according to our model, are as follows:

### Quarterbacks:

First Name	Last Name	POS	Pass Att.	Completions	Pass Yards	INTs	Pass TD	Fumbles	2018 Fant. Points
Carson	Wentz	QB	582.05	353.52	3905.10	11.59	44.60	3.47	330.10
Tom	Brady	QB	614.51	400.37	4848.17	9.04	34.24	1.90	322.75
Aaron	Rodgers	QB	473.27	324.49	3765.74	8.53	38.80	2.83	322.62
Cam	Newton	QB	502.76	294.48	3704.83	13.93	23.20	2.90	316.70
Russell	Wilson	QB	526.73	322.22	3821.74	9.93	26.81	2.67	304.75

### Running Backs:

First Name	Last Name	Rush Att.	Rush Yards	Rush TD	Targets	REC	REC Yards	REC TD	Fumbles	2018 Fant. Points
Todd	Gurley	284.44	1213.95	10.31	121.86	74.22	1096.80	0.00	1.42	364.31
Le'Veon	Bell	317.58	1394.42	8.97	104.30	80.32	680.69	1.94	0.97	351.35
Ezekiel	Elliott	360.00	1686.15	13.54	48.62	36.31	375.38	1.85	0.00	334.77
David	Johnson	213.16	598.21	11.43	116.13	96.00	849.74	3.66	2.29	326.74
Alvin	Kamara	125.33	705.78	8.00	98.67	77.33	798.22	5.33	0.00	307.73

### Tight Ends:

First Name	Last Name	POS	Targets	REC	REC Yards	REC TD	Fumbles	2018 Fant. Points
Rob	Gronkowski	TE	163.03	79.27	1308.09	12.35	0.00	284.19
Travis	Kelce	TE	118.81	85.67	977.55	5.49	0.12	216.14
Jordan	Reed	TE	114.72	85.74	885.43	6.94	1.21	213.52
Zach	Ertz	TE	123.36	89.13	952.09	4.72	0.62	211.42
Vance	McDonald	TE	152.04	87.26	1038.63	2.61	0.00	206.79

### Wide Receivers:

First Name	Last Name	Rush Att.	Targets	REC	REC Yards	REC TD	Fumbles	2018 Fant. Points
Antonio	Brown	0.70	195.91	112.09	1943.02	15.04	0.77	397.01
Odell	Beckham Jr.	4.52	169.67	105.67	1480.00	12.67	0.00	331.50
Julio	Jones	2.04	166.73	100.94	1536.00	7.61	0.62	301.09
Tyreek	Hill	21.82	127.25	71.27	1537.29	6.51	0.00	281.72
A.J.	Green	0.01	150.55	88.22	1374.69	8.84	0.91	276.89

### Kickers:

First Name	Last Name	PAT	FG Short (< 40 Yrds.)	FG Medium (40-49 Yrds.)	FG Long (50+ Yrds.)	2018 Fantasy Points
Jake	Elliott	44.07	16	13.18	5.65	173.01
Matt	Bryant	41.17	18.57	6.75	7.26	160.17
Chris	Boswell	35.59	19.59	13.13	1.89	156.32
Stephen	Gostkowski	51.58	19.23	9.28	1.85	155.63
Justin	Tucker	35.76	17.98	9.57	5.33	154.64

### Defenses:

TEAM	2018 Fant. Pts.
Jacksonville Jaguars	153.42
New England Patriots	148.18
Philadelphia Eagles	146.60
Baltimore Ravens	145.80
Seattle Seahawks	141.27

### 5.3 Player Rankings – Overall Top 100

The second major question of this research is to determine the order in which players should be drafted. The forecasted list of top 100 players available in a 2018 fantasy football draft provides fantasy participants the players that should come off the draft board throughout the duration of the draft (depending on a ten or twelve team league), also showing where value can be found later in the draft after the early rounds are over.

Rank	First Name	Last Name	Position	2018 Fantasy Points
1	Antonio	Brown	WR	397.01
2	Todd	Gurley	RB	364.31
3	Le'Veon	Bell	RB	351.35
4	Ezekiel	Elliott	RB	334.77
5	Odell	Beckham Jr.	WR	331.50
6	Carson	Wentz	QB	330.10
7	David	Johnson	RB	326.74
8	Tom	Brady	QB	322.75
9	Aaron	Rodgers	QB	322.62
10	Cam	Newton	QB	316.70
11	Alvin	Kamara	RB	307.73
12	Russell	Wilson	QB	304.75
13	Julio	Jones	WR	301.09
14	Orleans	Darkwa	RB	298.36
15	Jakeem	Grant	WR	295.40
16	Ben	Roethlisberger	QB	290.13
17	Chris	Godwin	WR	289.60
18	Alex	Smith	QB	284.99
19	Rob	Gronkowski	TE	284.19
20	Matthew	Stafford	QB	284.17
21	Leonard	Fournette	RB	283.80
22	Tyreek	Hill	WR	281.72
23	Kareem	Hunt	RB	278.32
24	Drew	Brees	QB	277.24
25	Dak	Prescott	QB	277.05
26	A.J.	Green	WR	276.89
27	Michael	Thomas	WR	274.81
28	Philip	Rivers	QB	273.35
29	Davante	Adams	WR	271.26
30	LeSean	McCoy	RB	270.22
31	Keenan	Allen	WR	263.59
32	Marcus	Mariota	QB	261.37
33	DeAndre	Hopkins	WR	258.38
34	Jameis	Winston	QB	253.02
35	Matt	Ryan	QB	252.95
36	Blake	Bortles	QB	252.44
37	Mike	Evans	WR	251.72
38	Kirk	Cousins	QB	251.28
39	Martavis	Bryant	WR	250.67
40	Carson	Palmer	QB	249.95
41	Larry	Fitzgerald	WR	249.63
42	DeMarco	Murray	RB	246.96
43	Melvin	Gordon	RB	245.21
44	Andy	Dalton	QB	241.24
45	Josh	Gordon	WR	238.24
46	Derek	Carr	QB	238.15
47	Jarvis	Landry	WR	234.77
48	Tyrod	Taylor	QB	233.87
49	Marcus	Murphy	RB	233.60
50	Demaryius	Thomas	WR	233.52

Rank	First Name	Last Name	Position	2018 Fantasy Points
51	Adrian	Peterson	RB	233.03
52	Dez	Bryant	WR	232.29
53	Marshawn	Lynch	RB	230.86
54	Golden	Tate	WR	230.29
55	Allen	Robinson	WR	230.25
56	Ty	Hilton	WR	229.49
57	Robert	Woods	WR	228.57
58	Brandon	Marshall	WR	228.55
59	Christian	McCaffrey	RB	228.05
60	Doug	Baldwin	WR	224.23
61	Jay	Cutler	QB	223.84
62	Brandin	Cooks	WR	223.77
63	Devonta	Freeman	RB	223.10
64	Mark	Ingram	RB	222.52
65	Stefon	Diggs	WR	221.65
66	Alshon	Jeffery	WR	219.82
67	Joe	Flacco	QB	218.13
68	Carlos	Hyde	RB	217.47
69	Travis	Kelce	TE	216.14
70	Eli	Manning	QB	216.01
71	Jared	Goff	QB	215.50
72	Kenyan	Drake	RB	215.47
73	Sam	Bradford	QB	214.80
74	Frank	Gore	RB	214.74
75	Matt	Forte	RB	213.88
76	Jordan	Reed	TE	213.52
77	Jordan	Howard	RB	213.50
78	Josh	McCown	QB	213.01
79	Zach	Ertz	TE	211.42
80	Jamaal	Williams	RB	211.41
81	Jeremy	Maclin	WR	210.14
82	Case	Keenum	QB	206.81
83	Vance	McDonald	TE	206.79
84	Ryan	Fitzpatrick	QB	205.02
85	Michael	Crabtree	WR	204.70
86	DeSean	Jackson	WR	204.33
87	Nick	Foles	QB	201.97
88	Amari	Cooper	WR	201.97
89	Randall	Cobb	WR	200.53
90	Marvin	Jones	WR	200.31
91	Ted	Ginn	WR	199.33
92	Brock	Osweiler	QB	198.91
93	Cooper	Kupp	WR	198.70
94	Juju	Smith-Schuster	WR	197.55
95	Jacoby	Brissett	QB	196.73
96	Kelvin	Benjamin	WR	195.20
97	Pierre	Garcon	WR	195.05
98	Mike	Glennon	QB	194.13
99	Teddy	Bridgewater	QB	193.32
100	Sterling	Shepard	WR	191.83

## 5.4 Forecast Validation and Accuracy

To ensure the validity of the 2018 fantasy football scores and rankings, the method was repeated using all data for each player up until 2016, rather than 2017, producing fantasy scores for 2017. Fantasy score projections for 2017 were then analyzed against players actual 2017 fantasy scores (excluding first year players). The absolute percentage error between 2017 forecasts and actuals was derived for each player. The absolute percentage errors were then averaged leaving the Mean Absolute Percentage Error (MAPE).

**The MAPE between the 2017 forecasted and actual fantasy scores was 3.53%.** Most of the error is believed to be attributed to injuries that players sustain causing them to miss plays, games, even seasons. Considering injuries were not accounted for in this model, the fantasy score forecasts are substantially accurate.

## 6 Limitations/Future Work

Even though our approach shows a small average percentage error of nearly 3.5%, our model has a few key limitations which leaves the door open for future developments.

First, in the majority of fantasy football leagues, if a kicker were to miss a field goal or a point after touchdown (PAT) then negative points would be applied to that kicker's fantasy score. Missed field goals and PATs were absent from the dataset used in this model. The forecasting accuracy of kicker scores may be improved by considering misses in the model.

Second, injury is a large part of fantasy football, and many players do get hurt during the season. Some for longer stretches of time than others. Our model did not account for the probability of injury to a player, which is extremely difficult to determine.

Finally, our ARIMA model assumed that a player's various performance metrics are not correlated. Considering correlations between player performance metrics, (e.g., completions and passing yards) using a multivariate time series method may improve the accuracy of future projections, but this would require an advanced graduate-level understanding of timeseries modeling.

Our methodology can indeed be applied across all fantasy team sports. It would be beneficial to adapt this model to analyze baseball, basketball, hockey, and the various other fantasy sports

available to consumers. In addition to expanding to other sports, developing the ability to analyze performance on a game-to-game basis could drastically assist a fantasy owner beyond the initial draft period, to win his/her match-up in any given week. With fantasy sports drafts being annual events, this model could be updated to provide fantasy football owners with the information they need every year.

## References

- Benderoff, E. (2007). Real-world Advances Alter Fantasy Sport. Retrieved April 29, 2018, from [http://articles.chicagotribune.com/2007-02-26/business/0702260015\\_1\\_fantasy-sites-fantasy-sports-fantasy-football](http://articles.chicagotribune.com/2007-02-26/business/0702260015_1_fantasy-sites-fantasy-sports-fantasy-football)
- Greenfield, J. (2003). Hooked on Fantasy. Retrieved April 30, 2018, from [http://articles.chicagotribune.com/2003-09-12/news/0309130246\\_1\\_fantasy-leagues-fantasy-football-problem-gambling](http://articles.chicagotribune.com/2003-09-12/news/0309130246_1_fantasy-leagues-fantasy-football-problem-gambling)
- Hyndman, R. J., & Athanasopoulos, G. (2014). *Forecasting: principles and practice*. 2<sup>nd</sup> ed. OTexts. Retrieved May 14, 2018, from <https://otexts.org/fpp2/>
- McNamara, C. (2010). Chicago Accountant is World's Top-ranked Fantasy Football Player. Retrieved April 29, 2018, from [http://articles.chicagotribune.com/2010-09-05/sports/ct-spt-0906-fantasy-football--20100905\\_1\\_fantasy-football-football-background-leagues](http://articles.chicagotribune.com/2010-09-05/sports/ct-spt-0906-fantasy-football--20100905_1_fantasy-football-football-background-leagues)
- Moorman, A. M. (2008). Fantasy Sports Leagues Challenged as Illegal Gambling. *Sports Marketing Quarterly*, 17(4), 232-234. Retrieved April 29, 2018, from <https://search.proquest.com/openview/c1886498495cd76fd87612d78f463f02/1?pq-origsite=gscholar&cbl=28711>.
- Morris, D., Krieg, D., Willmarth, D., & Kievatt, D. (2006). System and Method for Predicting Performance of Fantasy Athletes. *U.S. Patent Application No. 11/058,079*. Retrieved April 30, 2018, from <https://patents.google.com/patent/US20060183548A1/en>.
- Optimized Financial Systems (2014). Fantasy Football Prediction Using Vectorial ARMA Models. Retrieved April 28, 2018, from <https://www.optimizedfinancialsystems.com/fantasy-football-prediction-using-vectorial-arma-models/>.
- Peters, I. (2017). Fantasy Sports Services in the US. *IBIS World*. Retrieved April 28, 2018, from <http://clients1.ibisworld.com.libproxy.unh.edu/reports/us/industry/default.aspx?entid=4577>
- Plimi, J. (2006). Fantasy Sports League Pre-Draft Logic Method. *U.S. Patent Application No. 11/208,112*. Retrieved May 30, 2018, from <https://patents.google.com/patent/US20060040719A1/en>.



- Purdum, D., & Rovell, D. (2015). N.Y. AG Declares DraftKings, FanDuel are Illegal Gambling, Not Fantasy. Retrieved April 30, 2018, from [http://www.espn.com/chalk/story/\\_/id/14100780/new-york-attorney-general-declares-daily-fantasy-sports-gambling](http://www.espn.com/chalk/story/_/id/14100780/new-york-attorney-general-declares-daily-fantasy-sports-gambling)
- Shipman, F. M. (2001). Blending the real and virtual: Activity and spectatorship in fantasy sports. *In Proceedings from DAC'01: The Conference on Digital Arts and Culture*. Vol. 8, p. 2007. Retrieved April 30, 2018, from <http://www.csd.tamu.edu/~shipman/papers/dac01.pdf>
- Ware, S. B., & Webb, B. A. (2008). Evaluation of Fantasy Players. *U.S. Patent Application No. 11/606,717*. Retrieved April 29, 2018, from <https://patents.google.com/patent/US20080125228A1/en>.
- Wood, M. (2012). Read Our Coverage of Fantasy Football. *Learn Fantasy Football*. Retrieved April 30, 2018, from [http://articles.chicagotribune.com/2012-07-31/news/chi-read-our-coverage-of-fantasy-football-learn-fantasy-football-20120731\\_1\\_fantasy-football-fantasy-league-fantasy-sports](http://articles.chicagotribune.com/2012-07-31/news/chi-read-our-coverage-of-fantasy-football-learn-fantasy-football-20120731_1_fantasy-football-fantasy-league-fantasy-sports)

## Appendix: R Codes

---

```
require(forecast)
# Read offense Table
offense <- read.csv("Clean_Offense.csv", header=TRUE)
players <- read.csv("Clean_players.csv", header=TRUE)

#Table of forecasted values for each offensive metric for each player
nFactors <- 13
nPlayers <- 409
Forecast_offense <- matrix(, nrow = nPlayers, ncol = nFactors+2)

colnames(Forecast_offense) <- c("player","pa","pc","py","ints","tdp","ra","ry","tdr","trg","rec","recy","tdrec","fum1","FF")
FF_weights <- c(0,0,0.04,-2,4,0,0.1,6,0,1,0.1,6,-2)

for (player in 1:nPlayers) {
  #Filtering by each individual player and sorting by season and week
  playerID <- as.character(players[player,1])
  Forecast_offense[player,1] <- playerID

  player_data <- subset(offense,player == playerID) #CHANGE
  sort_player_data <-player_data[order(player_data$GAME.seas,player_data$GAME.wk),]

  for (factor in 1:nFactors) {
    #Fining the best ARIMA model for the given scoring metric
    ARIMAFit = auto.arima(sort_player_data[[7+factor]]) #CHANGE TO DOUBLE BRACKET COLUMN NUMBER
    #Using the best ARIMA model to predict values for the given offensive metric out 16 periods
    pred <- forecast(ARIMAFit,h=16)

    Forecast_offense[player,1+factor] <- sum(pred$mean)
  }
  #add line of code to calculate FF score for each player and write it in column 15 of the table (have to add another column to table in line 11)
  player_factors <- as.double(Forecast_offense[player,2:(nFactors+1)])
  FF <- crossprod(FF_weights,player_factors)
  Forecast_offense[player,nFactors+2] <- FF
}
#add line of code to save "forecast_offense" into a csv file, should call it the same as the table here
write.csv(Forecast_offense,"Forecast_offense.csv")
```

```
require(forecast)
# Read kicking Table
kicking <- read.csv("Clean_kickscore.csv", header=TRUE)
players <- read.csv("Clean_kickers.csv", header=TRUE)

#Table of forecasted values for each kicking metric for each player
nFactors <- 4
nPlayers <- 37
Forecast_kickers <- matrix(, nrow = nPlayers, ncol = nFactors+2)

colnames(Forecast_kickers) <- c("PLAYER","PAT","FGS","FGM","FGL","FFP")
FF_weights <- c(1,3,4,5)

for (player in 1:nPlayers) {
  #Filtering by each individual player and sorting by season and week
  playerID <- as.character(players[player,1])
  Forecast_kickers[player,1] <- playerID

  player_data <- subset(kicking,player == playerID) #CHANGE
  sort_player_data <-player_data[order(player_data$GAME.seas,player_data$GAME.wk),]

  for (factor in 1:nFactors) {
    #Fining the best ARIMA model for the given scoring metric
    ARIMAFit = auto.arima(sort_player_data[[6+factor]])
    #Using the best ARIMA model to predict values for the given kicking metric out 16 periods
    pred <- forecast(ARIMAFit,h=16)

    Forecast_kickers[player,1+factor] <- sum(pred$mean)
  }
  #add line of code to calculate FF score for each player and write it in the table (have to add another column to table in line 11)
  player_factors <- as.double(Forecast_kickers[player,2:(nFactors+1)])
  FF <- crossprod(FF_weights,player_factors)
  Forecast_kickers[player,nFactors+2] <- FF
}
#add line of code to save "forecast_kickers" into a csv file, should call it the same as the table here
write.csv(Forecast_kickers,"Forecast_kickers.csv")
```

```

require(forecast)
# Read defense Table
defense <- read.csv("Clean_Defense.csv", header=TRUE)

#team <- read.csv("")
players <- read.csv("Clean_Teams.csv", header=TRUE)
|

#Table of forecasted values for each defensive metric for each team
nFactors <- 1
nTeams <- 32
Forecast_defense <- matrix(nrow = nTeams, ncol = nFactors+1)

colnames(Forecast_defense) <- c("TEAM", "DEF Fant. Pts.")

FF_weights <- c(1)

for (team in 1:nTeams) {

  #Filtering by each individual team and sorting by season and week
  this_team <- as.character(players[team,1])
  Forecast_defense[team,1] <- this_team

  player_data <- subset(defense, tname == this_team) #CHANGE
  sort_player_data <- player_data[order(player_data$GAME.seas, player_data$GAME.wk),]

  for (factor in 1:nFactors) {

    #Finding the best ARIMA model for the given scoring metric
    ARIMAFit = auto.arima(sort_player_data[[3+factor]])
    #Using the best ARIMA model to predict values for the given defensive points out 16 periods
    pred <- forecast(ARIMAFit, h=16)

    Forecast_defense[team, 1+factor] <- sum(pred$mean)

  }
  #add line of code to calculate FF score for each player and write it in the table (have to add another column to table in line 11)

  player_factors <- as.double(Forecast_defense[team, 2:(nFactors+1)])
  FF <- crossprod(FF_weights, player_factors)
  Forecast_defense[team, nFactors+1] <- FF
}
#add line of code to save "forecast_defense" into a csv file, should call it the same as the table here
write.csv(Forecast_defense, "Forecast_defense.csv")

require(forecast)
# Read Offense 2017 Table
offense <- read.csv("Test_Offense.csv", header=TRUE)

players <- read.csv("Clean_players.csv", header=TRUE)
#Table of forecasted values for each offensive metric for each player
nFactors <- 13
nPlayers <- 409
Forecast_offense_Test <- matrix(, nrow = nPlayers, ncol = nFactors+2)

colnames(Forecast_offense_Test) <- c("player", "pa", "pc", "py", "ints", "tdp", "ra", "ry", "tdr", "trg", "rec", "recy", "tdrec", "fum1", "FF")
FF_weights <- c(0,0,0.04,-2,4,0,0.1,6,0,1,0.1,6,-2)

for (player in 1:nPlayers) {
  #Filtering by each individual player and sorting by season and week
  playerID <- as.character(players[player,1])
  Forecast_offense_Test[player,1] <- playerID

  player_data <- subset(offense, player == playerID) #CHANGE
  sort_player_data <- player_data[order(player_data$GAME.seas, player_data$GAME.wk),]

  #for (factor in 1:nFactors) {

    #Finding the best ARIMA model for the given scoring metric
    ARIMAFit = auto.arima(sort_player_data[[7+factor]]) #CHANGE TO DOUBLE BRACKET COLUMN NUMBER
    #Using the best ARIMA model to predict values for the given offensive metric out 16 periods
    pred <- forecast(ARIMAFit, h=16)

    Forecast_offense_Test[player, 1+factor] <- sum(pred$mean)

  }

  #}

  #add line of code to calculate FF score for each player and write it in column 15 of the table (have to add another column to table in line 11)

  player_factors <- as.double(Forecast_offense_Test[player, 2:(nFactors+1)])
  FF <- crossprod(FF_weights, player_factors)
  Forecast_offense_Test[player, nFactors+2] <- FF
}
#add line of code to save "forecast_offense" into a csv file, should call it the same as the table here
write.csv(Forecast_offense_Test, "Forecast_offense_Test2017.csv")

```