# A Statistical Analysis of Fantasy Football Data

## Fantasy Statistics of QB, WR, and RB from 1999-2019


Mark Haley†
Department of Computer Science
University of Colorado
Boulder, CO USA
mark.haley@colorado.edu

Varunjit Srinivas
Department of Computer Science
University of Colorado
Boulder, CO USA
varunjit.srinivas@colorado.edu



## ABSTRACT

Fantasy football is a popular American pastime that involves drafting players from NFL teams and getting points based off of their statistical performance in that week's game. As we have moved into the internet age, the game has evolved to the point where knowledge of the sport of football is now second to statistical knowledge when it comes to predicting fantasy football success. Platforms that host fantasy leagues often predict each player's points for a given week based off of their performances previously but this metric fails to account for extenuating circumstances such as fatigue and opposing team strength. Using an automatically updating dataset keeping track of fantasy statistics, we were able to discover several useful maxims to help aid fantasy managers in strategizing a successful season.

Firstly, we determined how players are affected by fatigue over the course of a season. It was determined that offensive players who experience more contact see increases in performance over the years while players who experience less contact are unaffected by fatigue over the course of the season. Also, across all positions, sophomore slumps were observed to occur regardless of position. Also, running backs experienced steady declines as they entered their veteran years while other positions became highly variable.

Finally, a k-means clustering analysis was completed to help determine what constitutes good vs middling vs disappointing seasons. In this, players were categorized into four clusters (strong SOS and high scoring, strong SOS and low scoring, weak SOS and strong scoring, and finally, weak SOS and weak storing). This allows every season by every player to be categorized. Taken together, an accurate painting of a player's total fantasy career can be extrapolated by how many appearances they make in each cluster over the course of their career.



## CCS CONCEPTS

• Linear Regression Analysis • Grid Based Clustering •

## KEYWORDS

Quarterback, Wide Receiver, Running Back, Fantasy Football, In season fatigue, Long-term career fatigue, SOS (strength of schedule)


## 1 Introduction

Currently, the system for estimating a player's fantasy performance is fully qualitative. Our research aims to take metrics such as fatigue, opposing defensive strength, as well as previous performance to help determine some heuristics that can be applied to help fantasy football players make data-driven decisions about their weekly performance. Our analysis will occur both on weekly and annual scales in order to help gain the full understanding of the factors that can affect a player's performance.

To ensure the data collected was relevant we restricted our research to players who appeared in 80% or more of the games in a given season. Our seasons were organized from weeks 1-15 from years 1999-2019. The first trait that we hoped to give insight on is fatigue and how it affects a fantasy owners draft strategy as well as their lineup decisions. Football as a game is incredibly violent and injury-prone, understanding how in season fatigue affects performance can help a fantasy owner determine what players to sit, trade, or start on any given week. Furthermore, understanding how age affects performance over the course of a career can help fantasy managers make decisions on what players to draft. Finally, the second factor we hope to dig into is classification. Through classification of every player's individual seasons, we can understand their abilities to perform against different levels of competition, further allowing owners to make informed draft day decisions based off of past performance.

Please note, the positions of TE (Tight End), D/ST (Defense/Special Teams), and K (Kicker), were not analyzed for this position as TE's serve the same purpose as WR and would thus yield redundant data while D/ST and K lack enough variance and quite frankly effect the game so little that there is relatively little point in studying them

## 2 Pre-Processing

Our Data is hosted on a github that automatically updates with the most recent statistics available. In its raw form, our data consists of many .csv files. For weekly stats, there is a csv that corresponds to every week of every season and for annual stats there is a csv that corresponds to every season. In order to perform sophisticated analysis on this data, we knew we would have to aggregate data from our hundreds of csv files into a few large dataframes that contain data for all weeks of every year. Python code was run to create 8 dataframes from which we could then mine, analyze, and report. Our resulting data took the form of the following dataframes: weekly statistics for QB, weekly statistics for RB, weekly statistics for WR, weekly statistics for all positions, annual statistics for QB, annual statistics for WR, annual statistics for RB, and annual statistics for all positions. Preprocessing this data was an arduous process as we had to take data from over 100 discrete files and compile them into understandably simple dataframes to ensure scalable research down the line as well as a clean database where structure and summary statistics are easily understood.

## 3 Early Exploration of Data - Annual Stats

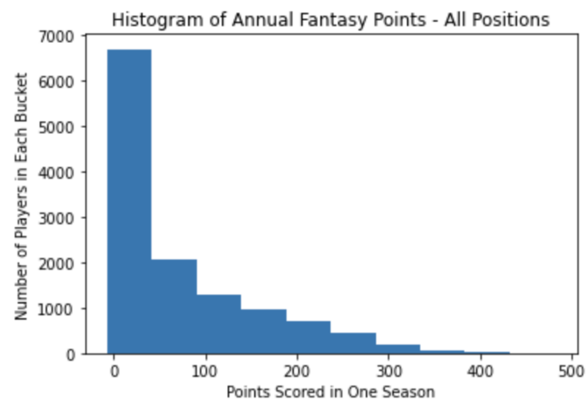



**Figure 3.1**

The histogram from figure 1 highlights the earliest and most crucial exploration of our data. This histogram highlights an already critical and unique finding, the fact that a large majority of NFL players generate less than 50 fantasy points in a given statistic. Looking ahead, this could potentially mean that our data is going to be heavily right skewed. As we move forward, we might try to further clean our data to only include players who perform above a certain threshold in order to better mimic a real-life fantasy football environment where underperforming players are left out of lineups in favor of better candidates.

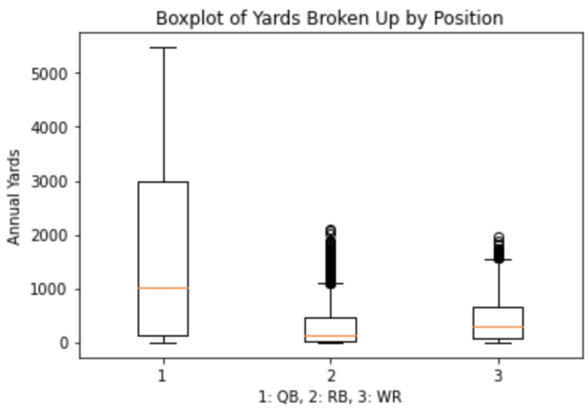



**Figure 3.2**

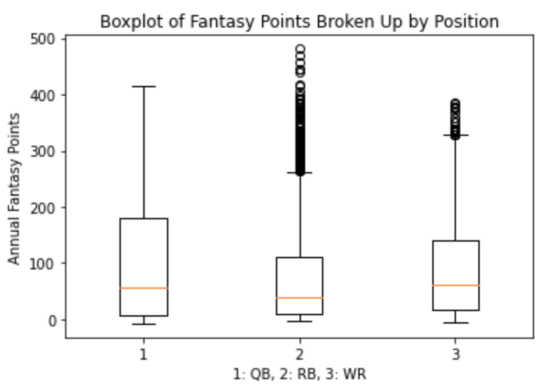



**Figure 3.3**

The figures above highlight another interesting finding of our data, the baseline summary statistics for annual quarterback, running back, and wide receivers. Specifically, the first box plot highlights the mean, IQR, and outliers for fantasy points while the second box plot displays the same statistics but for yards gained. From this data we gained several key insights that are not only interesting in and of themselves but also offer key guiding points for our further study. Firstly, quarterbacks generate more yards and points than the other positions (interesting anecdote, modern fantasy leagues know this and skew this by making quarterbacks earn points on a wider scale than the other positions). Also, we learned that it is relatively common for wide receivers and running backs to have an outlier season, where they are outside the range of 1.5x our IQR, whereas for quarterbacks an outlier season was never observed. Taken together, these statistics show that at least when observed an a season by season scale, quarterbacks will very rarely experience a truly historically good season in relation to others of the same position

## 4 Further Trimming Of Annual Data

As evidenced by the histogram in section 3 of this report, our data was heavily right skewed, with our distribution of fantasy points not following a normal distribution. In order to remedy this and ensure our data is robust and interesting, the data was trimmed to only count players who started 12 or more games in a season, which corresponds to 80% of all games played in a given season.

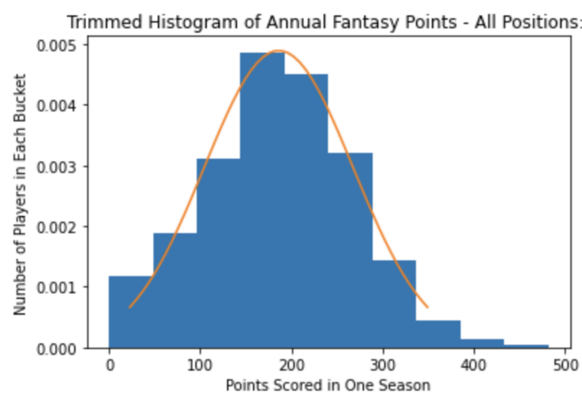



**Figure 4.1**

As evidenced by the improved histogram above, trimming players who started less than 80% of all games resulted in a much improved distribution of fantasy point values. Take note of the orange line in the graph which corresponds to the probability density function of this distribution. It clearly shows a normal distribution of values, which is ideal for analysis of our sort. Also note, this new figure is a density histogram which further normalizes our data range.

## 5 Exploratory Analysis - Weekly Stats

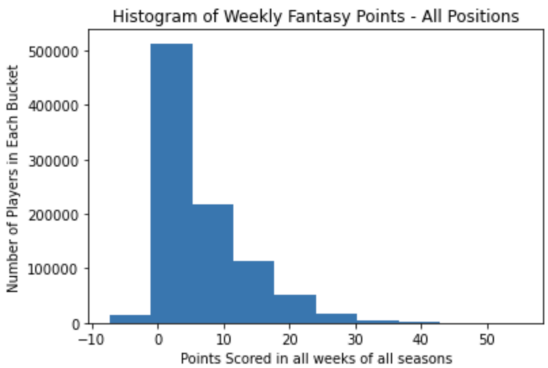



**Figure 5.1**

This histogram corroborates Section 3's histogram quite well, in showing that the vast majority of players, regardless of position, score less than 10 points in a week. We can use this to support the claim that the majority of football players will not get more than 50 points within a given year / season. This histogram also shows that our data, like the annual data, is skewed right. Like how Section 4 cleaned up the data from Section 3, we will need to add a section to where our data is cleaned to only include players who perform above a certain threshold in order to better mimic a real-life fantasy football environment where underperforming players are left out of lineups in favor of better candidates.

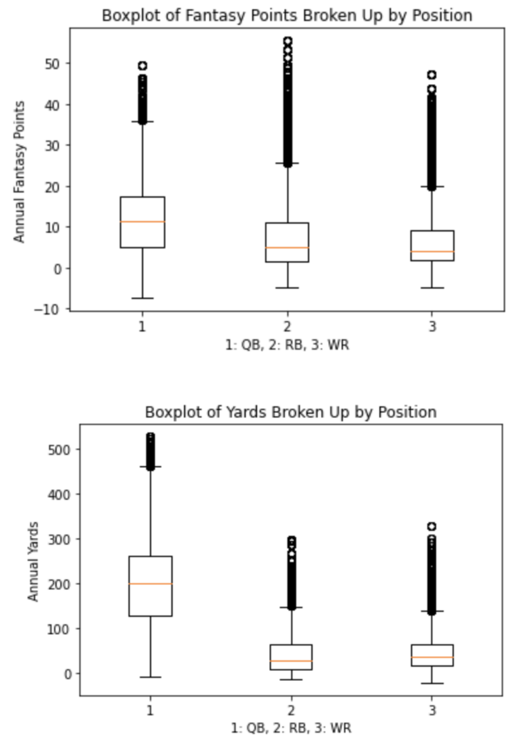
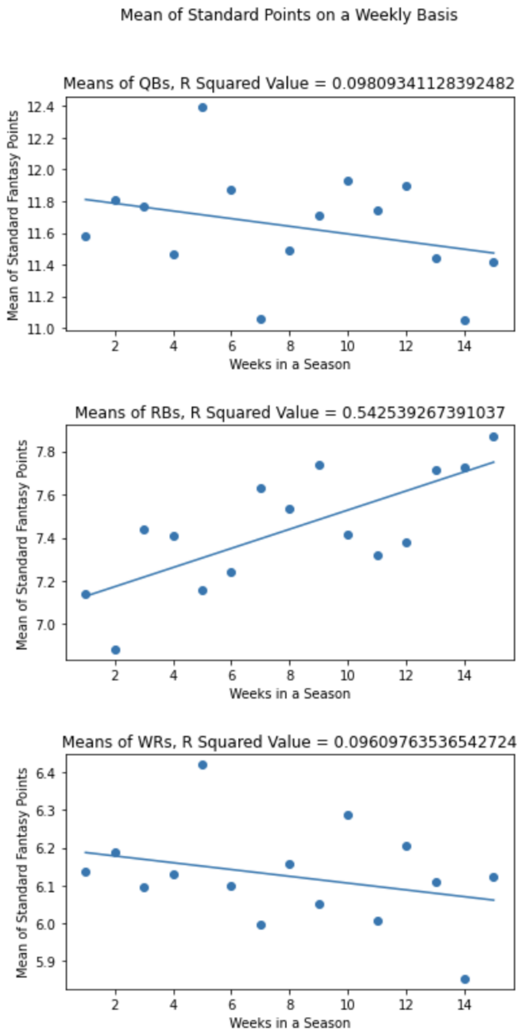



**Figure 5.2, Figure 5.3**

Again, our findings for Fantasy Points Broken Up by Position and Yards Broken Up by Position for the positions of Quarterback, Running Back, and Wide Receiver on a weekly basis corroborate our Section 3, which is where we look at these findings on a yearly basis. Just like the boxplots in Section 3, these boxplots show the mean, IQR, and outliers for Fantasy Points Broken Up by Position and Yards Broken Up by Position for each position we are focusing on. The key insights from weekly statistics are also very similar, being that the Quarterbacks gain the most points weekly, and run the most yards weekly. One key difference to note is that the number of weekly outlines for points gained by Running Backs is substantially higher than that of a yearly basis.

## 6 Analysis of Weekly Point Means




```
QBs R Squared Value =  0.09809341128392482
RBs R Squared Value =  0.542539267391037
WRs R Squared Value =  0.09609763536542724
```

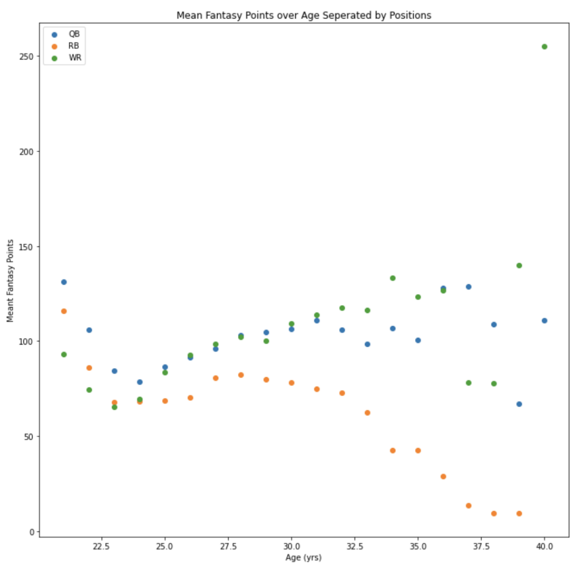
**Figure 6.1, Figure 6.2, Figure 6.3**

After calculating and plotting the means of each position on a weekly basis, from each week of every year from 1999 to 2019, we used a python library to calculate the linear regressions of the mean data for each position. Doing so would allow us to see if there is an overall increase or decrease in points gained throughout an average season of football for each position.

At first glance, the data for the means appears to be all over the place, however that is not the case for one of the three positions. Looking at the mean data of points for Running Backs, we can see that there is an upward trend of points gained throughout the season. The $R^2$ value of this data corroborates that, with a value of approximately 0.54 We can explain this gradual increase in mean points per week for Running Backs by understanding what this position does in a game, and how it gains standard fantasy points in a game. Usually in a game, a Running Back receives a handoff from the quarterback and charges for a rushing play, or they push up against the defense so they can catch the ball for a normal play. Doing these actions in a game will get them fantasy points, regardless if the actual play of the team was successful or not. Also as a season goes on, defensive players get worn out from playing their all in every game, which then makes it easier for running backs to gain more points and out maneuver them. Considering both a weakening defense and the fact that running backs have a consistent flow of points, we can now understand why there is an upward trend in the Running Back mean data, since the average running back would gain more points a week as the season progresses.

It seems that there is no correlation between the weekly means of points per week for Quarterbacks, meaning that we cannot conclude that points gained per week increase or decrease throughout the season for Quarterbacks. The R squared value for the mean of the points gained by Quarterbacks points to the same indecisive conclusion, since the $R^2$value is approximately 0.9809 which is a very low $R^2$value, which shows little to no correlation. This is a similar case for the Wide Receiver Position. The graph of the mean data of points for Wide Receiver, visually, follows a similar pattern to that of the graph of the mean data of points for Quarterback. It also has a similar $R^2$ value, being approximately 0.9609. With this analysis, we can also say that there is o correlation between the weekly means of points per week for Wide Receiver.

## 7 Analysis of Performance by Age



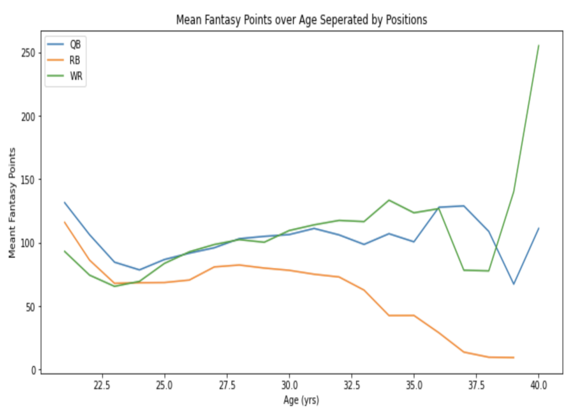




**Figure 7.1, Figure 7.2**

When looking at performance separated by age, there are several interesting observations to be extrapolated. The first and most pertinent observation is that across all positions, there is an observable dip in performance for a 1-2 year period immediately following a players rookie year before eventually recovering and stabilizing when a player reaches their physical peak around the ages of 26-30. This phenomenon is often qualitatively observed and referred to by football analysts, coaches, and fans as the "Sophomore Slump" and refers to an oft-occurring season or two of growing pains and reduced performance when compared against their rookie season. There are several explanations for the "Sophomore Slump", but for fantasy players this knowledge is very useful. When drawing projections for a draft before the season. Players should count on second or third year players to have a marked decrease in performance when compared to their earliest seasons. This knowledge could be used by fantasy managers to pinpoint potentially overvalued players in their draft to avoid in favor of more reliable players in a different age group.

Reliability itself is another important discovery to be taken away from this figure. In particular, between the ages of 26 and 33, players of all positions enter a period of lower variability from season to season. This occurs as players mature into veterans as they enter the years of their physical peak. It is known that the male brain does not finish maturing until around the age of 25, which is a possible explanation for the marked decrease in variability observed after this milestone. As players age past their prime, performance becomes more erratic year-to-year and becomes harder to predict. This can most likely be attributed to increasing mental mastery of the game, while physical prowess degrades as players enter middle age. These competing factors affect all players differently, and explain why it is harder to predict future performance of veteran players. Fantasy managers can leverage this knowledge by assuming that past performance of players between 26 and 33 is indicative of future performance. Furthermore this knowledge can be used to assess risk vs reward of certain players who age into veteran status.

Of Note, should be the relatively stable decline in performance observed in the running back position from their career peak which occurs around age 28. This is a statistical representation of yet another qualitative observation made by many football analysts and coaches, which is that running backs have much shorter windows of career success compared to wide receivers and quarterbacks. This is easily explained by the fact that running backs experience a much higher degree of contact and physical trauma when compared to the other offensive positions, shortening their time period of peak success markedly. Fantasy owners can leverage this knowledge by assuming running backs will see diminishing returns each season, allowing them to better predict what players to avoid on draft night.

## 8 Strength of Schedule Cluster Analysis

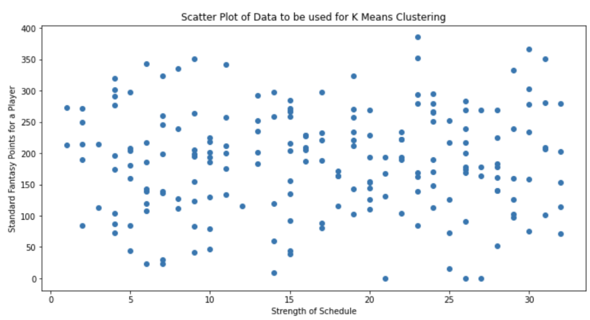
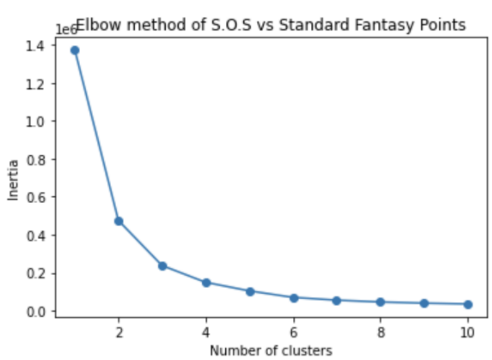



**Figure 8.1**



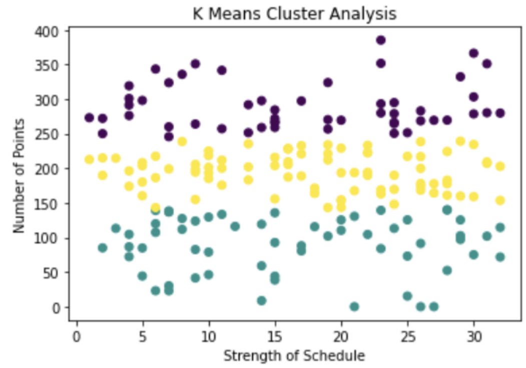




**Figure 8.2, Figure 8.3**

For our k-means cluster analysis. Our results were surprising, to say the least. We expected to see classes that were roughly structured around Strength of Schedule, where harder SOS corresponded to lower scores. What we saw instead, was clusters instead based around score and score alone, with three clusters for high performance, medium performance, and poor performance.

In fact, after further consultation, it was determined that there may indeed not even be clusters in this data set. At first this was a troublesome discovery, but ended up still yielding interesting findings. The first of which was that scoring is independent of strength of schedule. Furthermore, it also yielded a good metric for classifying high vs medium vs low performance regardless of strength of schedule.

Nevertheless, this information is still useful for fantasy owners, as many often use strength of schedule and players with strong schedules often fall to later in a draft. Fantasy owners should avoid letting SOS play into their decisions as they have little influence on a player's performance.

The data for this section consisted of a size 200 random sample from our weekly data set. Each player's annual performance was then cross referenced with their corresponding Strength of Strength that was hosted on a different data set. This was an arduous task as there have been several Location, and name changes that franchises undergone, and our code had to handle this complexity.

As mentioned above, although our results were far from what was expected, they were nevertheless enlightening. Especially the independence we discovered between strength of schedule and performance. Note, this analysis was performed on a annual basis as we lacked data for week-to-week strength. This is another point that could be used for further study later.

## 9 Discussion

This project set out to define some heuristics that fantasy football owners could use to aid in their draft day decision making as well as their lineup changes over the course of the season. A significant challenge faced by the researchers involved determining what heuristics should be examined. In the end, fatigue was chosen as the earliest metric as it approached the topic of fantasy prediction in a unique way, by attempting to notice trends of stability and variability over the course of a career and season. This was thought to be a good metric as knowledge of this fact can be combined with existing predictions to allow players to more accurately pinpoint scenarios where past performance can accurately predict future performance. Aggregating this data worked extremely well, moreso on the yearly data as some truly interesting data was observed in this scope, in particular the discovery of the sophomore slump, an observation which can drastically affect a fantasy owners thought process. Future work in this space could revolve around examining individual stats, such as receptions, and touchdowns, to gain further insight into the real-life football metrics that affect fantasy scores.

A k-means cluster analysis was also performed to help cluster player performance based off of their strength of schedule. Once the clusters were determined, these clusters were then applied to the players to see how their seasons stack up against other players. This method worked well and the clusters were easily determined. Further research could be completed on this topic by doing an analysis on every player, and seeing which cluster each season belongs to for them. This could then be analyzed to determine trends over the course of their career, as well as to help determine more advanced predictions. (ie. this player tends to have seasons of success against strong opponents, I shouldn’t fear drafting this player despite their strong strength of schedule).

## 10 Conclusion

Overall, we successfully were able to determine several heuristics that fantasy owners can use in addition to the predictions already provided by fantasy-league providers to make informed roster decisions both week to week, and at the beginning of the season. When the correct heuristics are used within the proper scope, we know knowledge of our findings can offer a significant tactical advantage compared to the average fantasy owner. In addition to what rules we proved, we also had several hypothesized rules that we proved were not valid through our analysis.

Even from our pre-processing, valuable data was gained. When we first created a histogram of points over year, we were shocked at how the majority of players had seasons of very low point scoring. We attributed this to the fact that the majority of players are either backups, or experience a shortened season due to extended periods of injury. To combat this and ensure our data was robust and representative of reality, we pruned out players who started less than 12 games. Not only did this result in our annual data forming a perfect normal distribution, but it also mirrors the reality of fantasy ownership, where players who are backups, or known to be experiencing a period of lengthy injury are either waived (cut), or not even drafted at all.

This then led us into our analysis of week-to-week performance within the scope of a single season. In order to be able to show in-season points progression on our raw data set, which combined the data from over 20 seasons, the mean of each week's points for a player was taken, before a linear regression was performed to help determine any trends in our data from week to week. The results we found were extremely enlightening, to say the least. Quarterbacks experience roughly no noticeable change in performance from week to week. Intuitively, this makes sense as quarterbacks experience both less contact than the other offensive positions as well as less physical exertion during a game. This is not to say that quarterbacks don't experience fatigue, but rather that the fatigue they experience is better managed by medical staff than other positions. Our analysis of points for the Wide Receiver position yielded similar results. Differing from the previous positions however, A noticeable performance improvement trend was observed for running backs as the season progressed. Our R^2 value was indicative of a strong trend, and the graph of our data was observed to make the result of our regression seem appropriate.

We then performed an analysis of annual data for players, this time with a player's age acting as a metric for their 'annual' performance. We hypothesized that annual performance would decrease as players aged as it is known that as humans get older. Our findings did prove that players reach a peak somewhere between the ages of 25 and 30. After roughly age 30, a noticeable decrease in performance occurs from this point until retirement. The most interesting part of our analysis however, was found early in players' careers. We observed a noticeable decline in players' performance in their 2nd and 3rd seasons that eventually recovers after one to two years. This was interesting as it provided a statistical representation of a phenomenon that is already qualitatively observed in football that is often referred to as the 'Sophomore Slump'. Furthermore, this 'sophomore slump' was observed across all positions studied. Finally, one notices from our figures that Wide Receivers saw a giant jump near the age of 40. This was caused by an outlier performance by Baltimore Ravens WR Steve Smith Sr. at age 40 in which he had a record-breaking performance in his last season. Because this was an outlier case, we cannot say that the wide receiver position experiences strong final seasons.

Our final bit of analysis of this project involved completing a k-means clustering analysis of annual performance relative to Strength of Schedule. Strength of schedule is a metric that ranks the overall strength of a player's opposition. We were expecting to see 2-3 clusters that highlighted reduced performance as strength of schedule increased. What we observed was that Strength of schedule has little to no effect on performance. We determined this through a qualitative look at the result of our cluster analysis, in which good seasons, mid seasons, and bad seasons were clustered exclusively by performance with no change to the cluster as strength of schedule increased. This was both a disappointment, as it was a heuristic that we couldn't report, as well as an interesting finding.

Overall, we summarized our findings into a small set of heuristics that fantasy owners can use to make informed roster decisions. They are as follows:

**Advanced Fantasy Football Heuristics For Owners**

1. Expect QB and WR performance to be relatively consistent week to week. Furthermore, expect better performance from the RB position as the season goes on.

2. Expect significantly reduced performance for 2nd and 3rd year players but also expect a recovery that last until a peak around age 30, before eventually seeing a sustained but slow degradation for every season following. This heuristic applies to all positions.
3. Do not let opposing strength play into your weekly roster decisions, as opposing strength does not play into performance at all.

If fantasy owners follow these heuristics, they will be able to more accurately predict players performance and be able to effectively maximize their total points and put themselves in the best position to win.

This was an exciting project that offers a much greater understanding of fantasy football and also offered many points for further study. Fantasy Football is a complex and competitive game, and we feel that knowledge of our findings can allow fantasy owners to have a decisive competitive advantage.

## ACKNOWLEDGMENTS

CSCI 4502 Faculty, National Football League

## REFERENCES

[1] Goode, Hugh. "Data Driven Fantasy Football Lineup Optimization - Part 1." *Nycdatascience.com*, NYC Data ScienceAcademy, https://nycdatascience.com/blog/student-works/data-driven-fantasy-football-lineup-optimization-part-1/.Goode, Hugh. "Data Driven Fantasy Football Lineup Optimization - Part 1." *Nycdatascience.com*, NYC Data Science Academy, https://nycdatascience.com/blog/student-works/data

[2] Rodriguez, Carroll, et al. "Identifying Undervalued Players in Fantasy Football." *Smu.edu*, 2019, https://scholar.smu.edu/cgi/viewcontent.cgi?article=1092&context=datasciencereview.

[3] Jermann, Michael. "Identifying over Valued Players in Fantasy Football - Stanford University." *Stanford.edu*, http://cs229.stanford.edu/proj2013/Jermann-IdentifyingOverValuedPlayersinFantasyFootball.pdf.

## Appendix

"On my honor, as a University of Colorado Boulder student, I have neither given nor received unauthorized assistance."

-Mark Haley

-Varunjit Srinivas

Mark's individual contributions consisted of completing the code regarding annual statistics. Mark also completed some of the pre-processing of the data, and developed the method for aggregating the hundreds of unique csv files into a cohesive dataframe ripe for future analysis. Mark also focused on writing the paper sections regarding annual statistics as well as the k means clustering analysis.

Varunjit's individual contributions consisted of completing some preprocessing code as well as all code portions regarding weekly data and k-means clustering. He also helped write the analysis on individual statistics. He developed the method for completing k means clustering. Varunjit also worked on all formatting for this project (Slides, Report, Code). Varunjit and Mark equally contributed to the slide presentations.