# First Draft

## Title

*"Separating Good from Great: Analyzing the Impact of Separation on an Individual NFL Receiver Success"*

## Author and Date

Dylan Johnson

April 7th, 2023

## Abstract

In this study, I aim to investigate whether a Wide Receiver's (WR) Separation (SEP) is a predictor of their success in the National Football League (NFL). Using regression analysis, exploratory analysis, and time series analysis, I examined the relationship between Separation and other variables such as Catch Percentage (CTCH%), Yards (YDS), and Targets (TAR). The results of this study were surprising, as I found some interesting correlation between a WR's Separation and their success in the NFL. This finding was further supported by insightful graphical representations that made the results easy to interpret. The implications of this study are significant, as it suggests that Separation is a critical factor in the success of a WR in the NFL. Therefore, players can narrow their focus during the offseason as well as adjust during the game. In addition, coaches and teams may benefit from focusing on this factor when scouting and developing their WRs. In conclusion, this study adds to the understanding of the factors that contribute to the success of a WR in the NFL and recommends future

research to explore the impact of other factors such as quarterback play and game strategy.

## Introduction

The National Football League (NFL) is an extremely popular League that continues to grow in both views and performance. In 2021, the NFL made a record-high $17.19 billion in revenue (Kolmar, 2023). The average value of an NFL team franchise as of 2022 is about $4.5 billion (Kolmar, 2023, Mintel, 2022). According to the Mintel Consumer Data report on US football (2022), the NFL is the leader of America's most popular sport and shows "no signs of slowing down." In 2021, fans spent 370 billion minutes consuming NFL content (McCarthy, 2022). As America's favorite sport continues to grow, each franchise is striving to reach the forefront from hiring new head coaches to trading players and draft picks. Less than 2% of college football players get drafted into the NFL (Cesconetto, 2021). As teams strive for a winning record and Super Bowl title, undrafted players must prove their worth and ability to make it as a professional football player in this highly competitive industry. Current research is limited and breeds mixed results on what measurements or factors correlate with the success of an athlete. The aim of this paper is to analyze the recent Separation Rates of NFL receivers by considering the other factors that may generate "success" for wide receivers in the National Football League. In doing so, we have to be aware of some issues. Initially we have to accept the fact that NFL success is and can be defined by numerous different factors. One analysis by Connor Kings defined success for NFL receivers as "...determined by having better career average statistics in yards, attempts,

touchdowns, and yards per attempt" (King, p.20). In my analysis, I intend to look more closely at more recent data that may give insight to coaches, players, and fans about success in the NFL today.

## Background

The NFL has been using advanced analytics to gain a competitive edge for years. Next Gen Stats, a division of the NFL, collects and analyzes data on player performance, including Separation. Separation is a key metric for wide receivers, as it measures the amount of space, they have to make a catch. A receiver with a high Separation is more likely to catch a pass, as they have more space to work with. However, it is unclear whether Separation is directly linked to a receiver's success in the NFL. Much of this unclarity is linked to the lack of a collective definition of "success" in the NFL. In the context of football, success can be defined in different ways, depending on the viewers and the purpose of analysis. However, many statistics that could potentially be related to success for a wide receiver include Targets, Receptions, Reception Yards, Catch %, and much more. I intend on clarifying the relationships and connections between Separation and the variables considered to represent success.

## Lit Review

As the NFL continues to dominate the entertainment and sporting industries, franchises are looking to recruit the best talent. A multitude of variables influence the productivity of various positions in the NFL. Though a wide range of variables have

been looked at, there still is limited research in what factors correlate with more productive NFL careers. This is likely due in part to the observation that there are limited studies of relevant research, but they each examine a multitude of factors. It may also be due to the mixed results and conclusions that are even contradictory to each other when comparing some studies. Some studies suggest there are early variables correlated with the productivity of players, while other studies found no correlation in variables. However, the current research has yet to agree on a variable that consistently correlates with productivity. Several studies have examined the relationship between Separation and other variables such as Catch %, Yards, and Targets. In a study by Tim B Martens, he found that a receiver's separation was positively correlated with their Catch %, and negatively correlated with Receiving Yards. This means that receivers with more separation tend to catch the ball at a higher rate yet gain fewer yards per catch. Although this study relates to my analysis, Martens left Tight End's in the data which produces much different results. Also, I have combined the past five seasons of NFL data -as opposed to just one season- to achieve more information as well as to get more accurate results. Lastly, not only do I go further into the individual statistics, but I also evaluate individual players and their history of success.

## Ethical Considerations:

This analysis was conducted using publicly available data from the NFL's Next Gen Stats website. No personal information was collected or used in this analysis. To ensure that I had permission to use this dataset, I followed the NFL's data sharing policy, which required registering an account and agreeing to the terms of use. I will

also keep the guidelines in mind to ensure that all the data used in this study are anonymized and do not violate any privacy laws or ethical considerations.

# Data & Methods:

## Sub-section on data:

The data for this study was obtained from Next Gen Stats, an NFL statistical database that provides in-depth analysis of player performance using advanced metrics. Specifically, I utilized data on the receiving statistics of Wide Receivers (WRs) from the 2018 to 2022 NFL seasons. This dataset included a total of 23,281 observations of individual WR performances, including the following variables:

- **Player Name:** The name of the WR

- **Team**: The NFL team the player was playing for

- **Games** Played: The total number of games the player appeared in during the season.

- **Targets**: The total number of times a player was targeted by the quarterback

- **Receptions**: The total number of receptions made by the player

- **Catch Percentage (CTCH%):** The percentage of receptions made by the player relative to the total number of targets.

- **Yards**: The total number of receiving yards gained by the player

- **Yards per Reception (Y/R)**: The average number of yards gained per reception made by the player.

- **Touchdowns**: The total number of touchdown receptions made by the player

- **Longest Reception**: The longest distance a player caught a pass for during the season.

- **Air Yards:** The total number of yards gained by the ball while in the air before being caught by the player.

- **Yards After Catch (YAC):** The total number of yards gained by the player after catching the ball.

- **Average Separation:** The average distance in yards measured between the player and the nearest defender at the time of catch or incompletion.

To obtain a comprehensive view of the relationship between Separation and the other variables, I analyzed data from all WRs who played in the NFL from 2018 to 2022 and excluded the Tight End position. The dataset was then preprocessed to remove any outliers or irrelevant observations that could affect the accuracy of the analysis. This dataset served as the foundation for the statistical analysis conducted in my study, which aimed to explore the relationship between a WR's Separation and their success in the NFL.

## Sub-section on methods: Regression Analysis

I conducted a regression analysis to examine the relationship between Separation and the other variables. I first removed Tight Ends from the dataset in order to focus deeply on the wide receivers of the NFL. To analyze the relationship between separation rates and success in the NFL, I have decided to use regression analysis. This will help me build a predictive model that identifies the relationship between separation rates and various performance metrics, while controlling for other factors that may influence receiver success, such as the quality of the quarterback and opposing

6

defense. My choice of method is consistent with established research practices in sports analytics. For instance, previous research has shown that regression analysis is commonly used to model the relationship between performance metrics in sports (Grier & Altman, 2018), while hypothesis testing is often used to examine the significance of differences in performance between different groups of athletes (Bock, 2013). In my regression, I made Separation the dependent variable, and the rest of the variables were independent variables.

## Regression Analysis Results

I started by testing the regression between 'SEP' and all other statistics. What I found was that the relationships between variables were actually very surprising. In this case, the F-statistic of my entire model is 84.36 with a p-value of 1.02e- 108, indicating that the model is statistically significant. There was also a lot of variation in the results and in Figure 1 and you can see the basic relationship plots of each variable. Based on the results from the first test, the relationship between 'SEP' and the following variables are displayed in Figure 1.
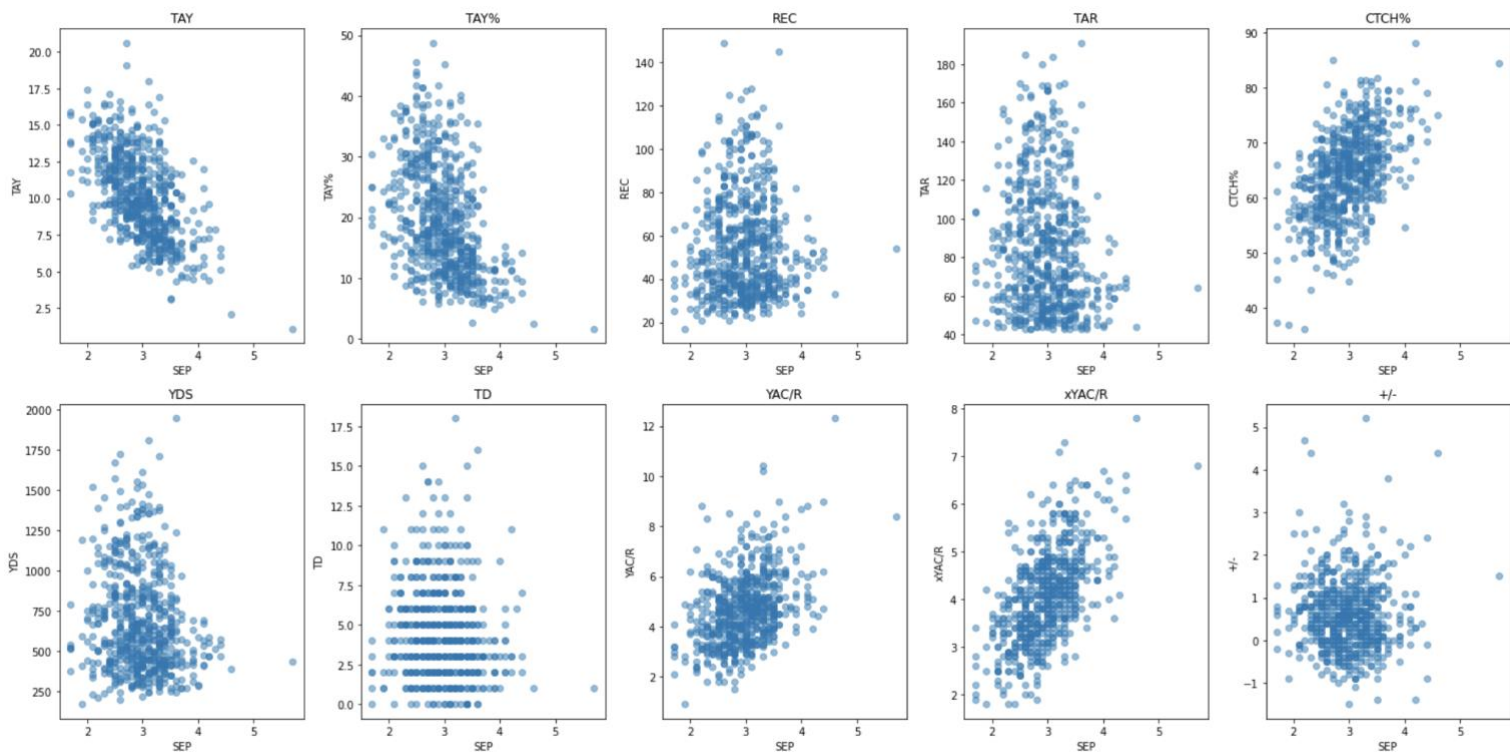
*Figure 1: In this plot chart, TAY, TAY%, REC, TD, YAC/R, xYAC/R, and +/- are not statistically significant predictors of SEP, as their p-values are greater than 0.05. However, TAY%, REC, TD, YAC/R, xYAC/R, and +/- may still have some predictive power in combination with other variables in the model.*

*Figure 2*

| Dependent Variable | Relationship |
|---|---|
| TAY | Negative |
| TAY% | Negative |
| REC | Negative |
| TAR | Positive |
| CTCH% | Positive |
| YDS | Negative |
| TD | Positive |
| YAC/R | Positive |
| xYAC/R | Negative |
| +/- | Negative |

All of these regression results will help me further my analysis, but I wanted to point out a few of the results that will directly impact my next move. I wanted to check the regression for each of these dependent variables individually, in order to see how much each variable is affected by Separation. I was able to do this by focusing variables such as that R-squared values; which indicate the proportion of the variance in 'SEP', the P-values; which indicate the statistical significance of my variables, the F-statistic; which tests whether the independent variables as a group are significant predictors, the Standard Error; which provide information about the precision of the estimates, and the coefficients; which gives me the relationships between the variables.

## Separation and Catch Percentage

First, I looked into the Catch Percentage (CTCH%) variable. The regression analysis shows that there is a statistically significant positive relationship between Catch % and Separation. The coefficient of determination (R-squared) is 0.26, which means that 26% of the variation in Catch % can be explained by the variation in Separation. The p-value for the slope coefficient is less than 0.05, indicating that the relationship between Catch % and Separation is significant. The intercept term of 41.14 indicates the expected value of Catch % when Separation is equal to 0. The regression line has a

positive slope of 7.87, which means that as Separation increases, Catch % also tends to increase. The omnibus test and Jarque-Bera test are not statistically significant, indicating that the residuals are normally distributed. Overall, these results suggest that Separation is an important predictor of Catch %, and players with higher separation tend to have a higher catch percentage.

## Separation and Targets

In the Regression Results, the coefficient for Targets is 0.0093 with a standard error of 0.004, and a p-value of 0.033. This indicates that there is a statistically significant positive relationship between the target variable "SEP" and the predictor variable "TAR" in this regression model. The coefficient of 0.0093 suggests that on average, an increase of one target is associated with an increase of 0.0093 points in SEP (the dependent variable). This coefficient can be interpreted as the marginal effect of Targets on Separation, holding all other variables constant. This result is important because it suggests that the more a player is targeted in the passing game, the more likely they are to contribute to their team's scoring efficiency (as measured by Separation). This could have implications for team strategy and player selection, as coaches and scouts may want to focus on players who are more heavily involved in the passing game.

In the OLS regression results presented, the variable "YDS" has a negative coefficient of -0.001 and a significant p-value of 0.000, indicating that there is a significant negative relationship between yards gained and the dependent variable Separation. In other words, as the number of yards gained decreases, the probability of

a positive increase in Separation. This variable is an important metric in football, as it reflects the ability of the player to gain ground and move the ball down the field. Therefore, it is expected that players who gain more yards would be evaluated more positively by their coaches. However, there is a negative relationship between "YDS" and "SEP" in this regression, so why is this important? Well, considering the Yards After Catch displayed a positive correlation, I am not that surprised to see this negative relationship on yards prior to the play. It is safe to assume that when Quarterbacks need to get a first down, they will throw the ball to the receiver who is most likely to catch the ball, even if the yardage is shorter. A receiver who can get more Separation from the defense, as well as catch the ball more consistently can and should be more successful than receivers who may run deep routes but do not catch as well.

The regression analysis proved very helpful in the sense that I am able to see each dependent variable's contribution to my analysis. Evidently, CTCH%, YDS, and TAR as they have a statistically significant coefficient with the outcome variable and do not appear to have high multicollinearity with other variables in the model. My model also provided evidence that receivers have a greater chance of catching the football when they have more Separation. In addition, NFL receivers tend to gain Targets as they increase their Separation. They also tend to lose yards as they increase their Separation, which will prove to be more relevant than I had originally predicted. All three of these variables are significant enough that I will carry them with me to my next task in my analysis.

## Methods: Exploratory Analysis:  All WRs in Dataset

Exploratory data analysis is the process of analyzing and visualizing data to better understand its characteristics, identify patterns, and formulate hypotheses. This process helps identify patterns and relationships within the data that may not be immediately obvious, and it helps identify outliers or anomalies within the data that may be errors or have special significance. It can help ensure the data is clean and can be trusted. To start my exploratory analysis, I sorted the remaining wide receiver data by each of the significant variables. This was my initial plan in order to identify possible outliers that may stray my data. To avoid the effects of the initial outliers, I calculated the median of Separation, Catch Percentage, Receiving Yards, and Targets for the dataset. I then used these values as a benchmark to determine whether a player's performance was above or below average. Overall, exploratory data analysis is an important step in my analysis as it can provide insights for furthering my study.

## Results of Exploratory Analysis: All WRs in Dataset

Based on my initial exploratory analysis, WR Rondale Moore has the highest Separation rate among all receivers in the dataset. Additionally, he also has the highest Catch %, which indicates a positive relationship between Separation and Catch % for Moore, as my model predicted. Figures 3 and 4 show these results by displaying the top 10 receivers in Separation and Catch %.
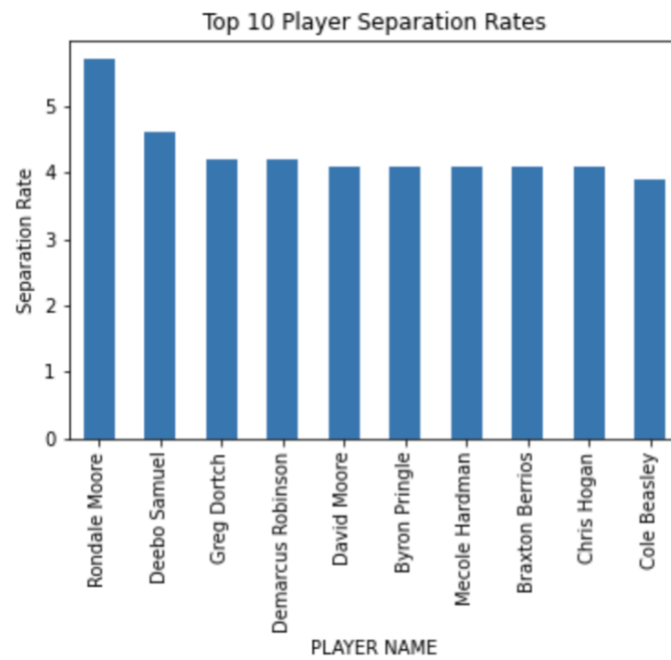
**Figure 3**



Top 10 Player Separation Rates

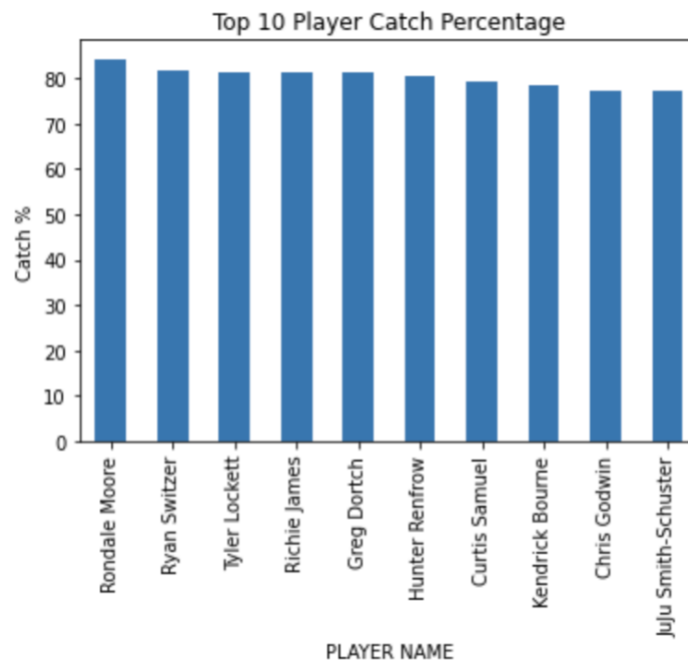**Figure 4**



Top 10 Player Catch Percentage

*Figure 3 and 4: Rondale Moore has the highest Separation as well as Catch Percentage in our dataset. Greg Dortch has the third highest Separation and the fifth highest Catch Percentage.*

These findings support my model's results and suggest that Separation and Catch % have a significant and positive relationship. Specifically, the fact that Rondale Moore has the highest Separation rate and highest Catch % out of anyone, indicates that creating separation from defenders likely increases the likelihood of making a catch. Moreover, Rondale Moore's high Separation rate and Catch % also suggest that Separation is a significant predictor of Catch % for wide receivers in football. If Separation was not an important predictor, I would expect to see many other receivers with similarly high Catch % despite having lower Separation rates. Overall, the fact that Rondale Moore has the highest Separation rate and highest Catch %, it validates my model's prediction that Separation and Catch % have a significant and positive relationship. These results highlight the importance of Separation as a predictor of Catch % for wide receivers in football and can be useful for developing strategies to improve a receiver's performance.  In addition to Moore, we see that receiver Greg Dortch is the third highest in Separation, as well as the fifth highest in Catch Percentage. There are over 475 receivers on the list, so the fact that these two receivers both make the top 10 of both lists is very helpful and very very significant for my analysis.

## Methods: Exploratory Analysis: Receivers who have met All Requirements

I then filtered the dataset to see how many wide receivers had above-average statistics in all of the categories: Separation, Catch Percentage, Receiving Yards, and Targets.

Scatter Plot of Separation and Targets


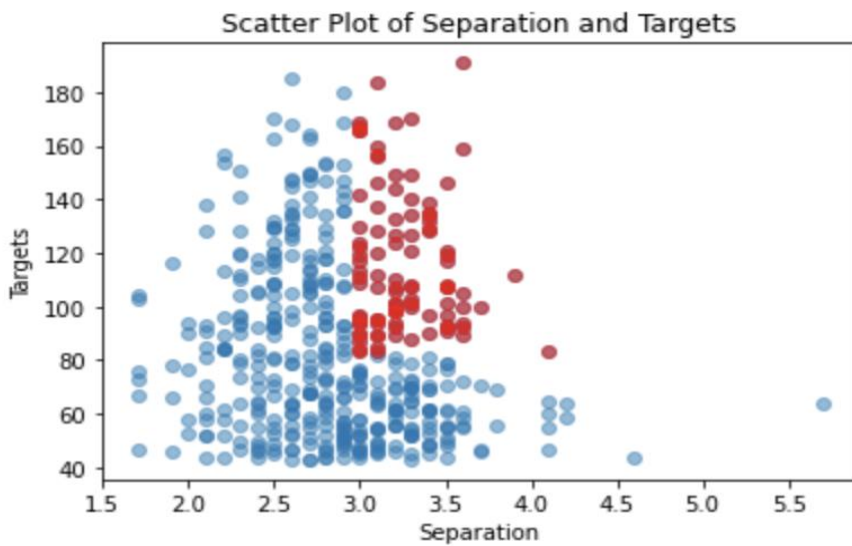Scatter Plot of Separation and Catch Percentage

*Figure 5, 6, 7: These scatter plots show the reduction of data after I compared the dataset to the total medians on each variable. The Red dots indicate the receivers are moving on with us in the analysis. The reduction will allow us to specialize and get more accurate results.*


Scatter Plot of Separation and Yards Per Catch
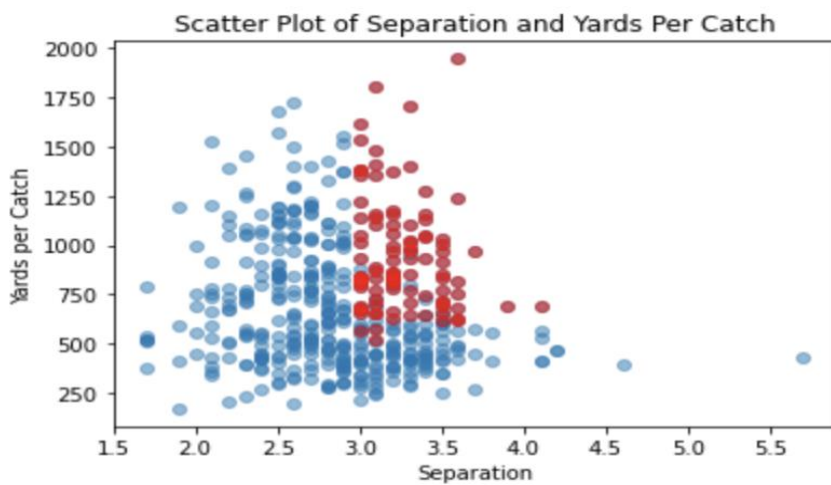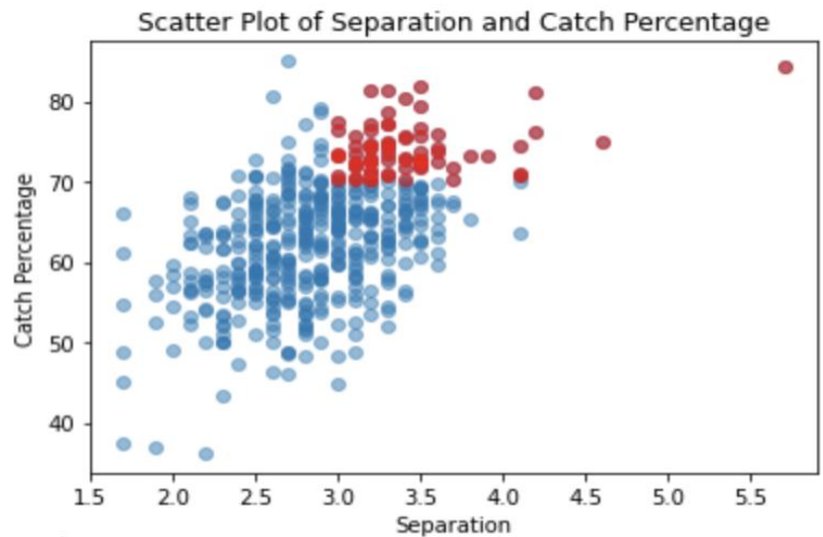
This becomes my final breakdown of the dataset, looking how many receivers had above-average statistics in each of these categories. Therefore, I am looking at the receivers who have excelled more than most in every category. For this method, I could consider that "success" in the NFL can be related to being above average in all four of our significantly related statistics: Separation, Catch Percentage, Receiving Yards, and Targets. Before I can move on in my analysis, I have to make sure to remove the duplicates in this new data because I was looking to see how many individual players have lined up with my model. Exploring the data through filtering and selecting specific variables is a method that not only helped in my progression in my analysis, but I was also able to check my assumptions from my model, such as whether or not an increase in a "successful" wide receivers Separation rate leads to a better Catch %.

## Results of Exploratory Analysis: Receivers who have met All Requirements

The results came back with 68 receivers meeting all of the requirements. As I mentioned in my methods section, I planned to proceed with the 68 players who met all the requirements because I am looking for Separation's relation to an NFL Receiver's success. When I removed the duplicates, the number of receivers above average in every category dropped to 42. That means that over a third of the "successful" players in this dataset prove my Model true. All of these receivers are considered "successful" in the season recorded and this number proves that receivers with more Separation get less yards on average but get thrown to more often and have a higher probability of catching the ball.

## Methods: Time Series Analysis: TOP 42

To prove the capability and success of these wide receivers, I wanted to first see how these receivers' statistics compare to the original data set with 478 receivers.

*Figure 8*

| Variable | Successful Wide Receiver Dataset | Initial Dataset | |
| --- | --- | --- | --- |
| | Median | Median | DIFF |
| SEP | 3.20 | 2.9 | +0.3 |
| CTCH% | 69.57 | 64 | +5.57 |
| YDS | 850.50 | 656 | +194.5 |
| TAR | 107.00 | 82 | +25 |

*Figure 8: We can see that every statistic in the new dataset has increased due to the focus on such successful receivers. We have to acknowledge that the data in the new dataset is not exactly lined up with our model's results. But a NFL's receivers' "success" has a complicated definition and too many factors go into it to consider one exact sentence. Therefore, the data is shaped and ready to move on.*

Now that I have a dataset compiled of 42 successful NFL Receivers, I want to iterate through each group and calculate the difference for each year. For this analysis, I will do a Time Series Analysis, which is a statistical method used to analyze and

interpret sequential data points collected over time. By analyzing the data over time,

trends can be detected, and forecasts can be made, which can be valuable in predicting

future performance. In the context of NFL receivers, time series analysis can be a

powerful tool for understanding how a receiver's performance has changed over time.

By examining their statistics over multiple seasons, patterns in their performance can be

identified, such as whether they have improved, declined, or remained consistent in key

performance metrics such as catch percentage or yards gained per reception. By

understanding how these factors influence performance over time, coaches, players,

and analysts can make better-informed decisions when it comes to player recruiting,

development, drafting, and game planning.

I start my time series analysis by iterating through each group using a for loop.

This analyzes the change in performance over time for each NFL receiver by calculating

the difference in their statistics between each year. This hopefully will allow us to see if

a player's performance is improving or declining over time and can help identify trends

or patterns in their performance. Next, I created a new column 'Increase' which

indicates if the player increased in all three categories and decreased in yards. This will

allow me to see if any receiver from the dataset any receivers have who have changed

their statistics in line with my significant variables.

**Results: Time Series Analysis: TOP 42**

*Figure 9*



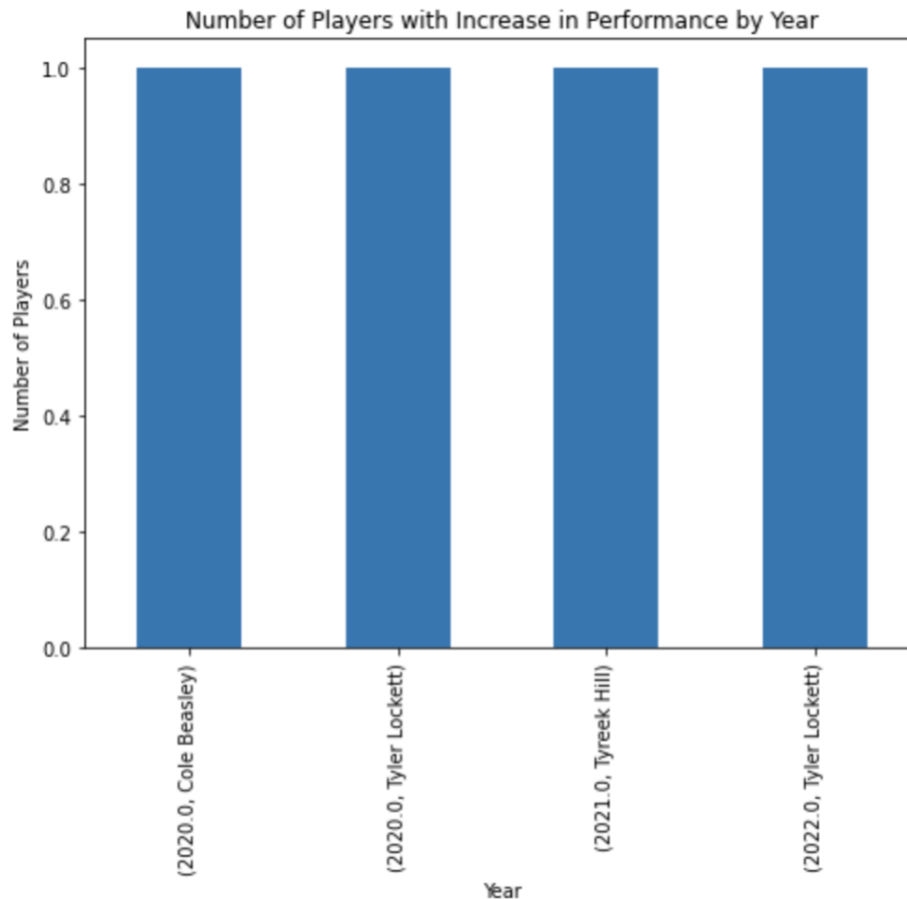Number of Players with Increase in Performance by Year

*Figure 9: The blue lines indicate that the NFL receiver has experienced the same relationships in statistics as our model. Three NFL receivers, Cole Beasley, Tyler Lockett, Tyreek Hill have increased in Separation, CTCH%, and TAR while decreasing their YAR.*

The data presented in the graph shows that three NFL receivers, namely Cole Beasley, Tyler Lockett, and Tyreek Hill, have shown a consistent trend of increase in Separation, CTCH%, and TAR while decreasing their YAR. It is interesting to note that Tyler Lockett has done it twice in the past 5 years, which could indicate his adaptability

and resilience in a highly competitive league. Moreover, the three receivers mentioned are considered to be very successful players in the league, with Tyler Lockett and Tyreek Hill continuing to dominate the defense. This further supports the idea that an increase in receiver Separation may be correlated with success in the NFL. It is worth noting that while the data suggests a correlation, it does not necessarily indicate causation. There could be various factors at play that contribute to a receiver's success, including but not limited to their physical abilities, experience, team dynamics, and game strategy. Nonetheless, the consistency of the trend observed in the three players over time could provide valuable insights for football analysts and researchers. Further investigation into the factors that influence a receiver's Separation and their relationship with success in the NFL could lead to more accurate predictions and better player development strategies.

## Justin Jefferson

Next, I will look into the individual growth of Justin Jefferson -the leader in Targets, Receptions, and Yards in 2022- over the past year, to see how my regression analysis lines up with the best in the NFL. The results I gathered show the changes in four statistical categories for Justin Jefferson between the 2021 and 2022 NFL seasons. Here's a breakdown of what each of those results stated:

- Justin Jefferson had 17 more targets in 2022 than he did in 2021.

- Jefferson's separation increased by 17 yards on average in 2022 compared to 2021.

- Jefferson's catch percentage increased by 4.9 percentage points in 2022 compared to 2021.

- Jefferson gained 193 more receiving yards in 2022 than he did in 2021.

The increase in Justin Jefferson's separation likely made it easier for him to catch the ball, which is supported by the increase in catch percentage. The increase in targets may have also contributed to the increase in yards gained, as more opportunities to catch the ball means more opportunities to gain yards. The gray area is that Separation has a positive relationship with Targets, but on the other hand, Target has an opposite relationship with YDS than Separation. Further studies could concentrate on Targets and Yards as opposed to Separation in order to unveil that relationship. Overall, these results suggest that Jefferson's performance improved in 2022 compared to 2021, and that his ability to create separation likely played a role in this improvement. This highlights the importance of separation as a factor in a receiver's ability to catch the ball and gain yards and suggests that improving separation may be a key area of focus for receivers looking to improve their performance.

## Major Consideration

One major consideration I have to mention is that the Quarterback position is very important when it comes to receiver statistics in the NFL. Whether or not the receiver gets targeted is solely in the hands of the quarterbacks. So, this authority may weaken or even strengthen our model's reliability if a receivers QB changes the next season. It all depends on how much that quarterback trusts the receiver. I would argue

that Catch Percentage should be the favorite statistic for quarterbacks looking for a go-to receiver. Dynasty Analytics analyst Martens talks about one of the best receivers in the NFL, Stephon Diggs; "He has a quarterback who trusts that he will come down with the ball when thrown his way" (Martens, 2022). This is true in the sense that quarterbacks and receivers are both trying to score. So, the receiver with the best chance of catching the ball is most likely the guy that should and would be thrown to, which is not possible without increasing that receiver's Target statistic. Therefore, I promote the future analysis on quarterbacks' effect on receiver's vs the effects they control themselves. This could be a great analysis to succeed.

## Conclusion

In this study, I analyzed whether a Wide Receiver's Separation is a predictor of their success in the National Football League. With the first regression analysis, I found out that my model provided evidence that receivers have a greater chance of catching the football when they have more Separation, receivers tend to gain Targets as they increase their Separation, and they also tend to lose yards as they increase their Separation. In my exploratory analysis, we learned that Rondale Moore has the highest Separation rate and highest Catch % in our dataset, and that Greg Dortch also was in the top 5 of both tests. These findings both validate my model's prediction that Separation and Catch % have a significant and positive relationship. In my time series analysis, I found that three NFL receivers, Cole Beasley, Tyler Lockett, Tyreek Hill have increased in Separation, CTCH%, and TAR while decreasing their YAR. This directly aligned with the model's assumption that Separation is a valuable statistic that is more

important in achieving NFL success than prior studies have found. Lastly, we looked

into the game's current "most successful" receiver, Justin Jefferson. His Separation

went up in 2022 and so did his Catch Percentage, and Targets. This highlights the

importance of separation as a factor in a receiver's ability to catch the ball and suggests

that improving separation may be a key area of focus for receivers looking to improve

their involvement and performance. In conclusion, this study adds to the understanding

of the factors that contribute to the success of a WR in the NFL and recommends future

research to explore the impact of other factors such as quarterback play and game

strategy.

# References and Appendices

Bock, T. M. (2013). Applied statistics in sport research. Journal of Sport Psychology in Action, 4(3), 135-146.

Brown, D. (2021). How Separation Impacts Receptions and Yards Per Route Run for NFL Wide Receivers (Undergraduate thesis). John H. Fisher College.

Dynasty Nerds. (2021, July 26). WR Separation: A Little Goes a Long Way. Dynasty Nerds. https://www.dynastynerds.com/wr-separation-a-little-goes-a-long-way/

Grier, S. A., & Altman, D. G. (2018). Statistics applied to the clinical and health sciences (5th ed.). Elsevier.

Ourand, J. (2022, January 13). NFL to Add Media to Boost Streaming Content in 2022. Front Office Sports. https://frontofficesports.com/nfl-media-expansion-2022/

Mintel Group Ltd. (2022). Football - US. Mintel Group Ltd. https://store.mintel.com/report/football-us-2022

NFL Next Gen Stats. (2022). Receiving - 2022 Regular Season - All Players - Yards. https://nextgenstats.nfl.com/stats/receiving/2022/REG/all#yards

Zippia. (n.d.). NFL Demographics and Financials: The Complete Analysis. Zippia. https://www.zippia.com/advice/nfl-demographics-financials/

Link to GitHub Repo with Data, Code, and README:
https://github.com/DylanJohnson02/da-401-johnson