# Statistical Analysis & Report

## NFL Draft Overview

* Looking at the overall data, on the graphs, a significant majority of players drafted between 2016 – 2019 did not sign second contracts (77.9%) with their respective team that drafted them showing challenges of the teams selecting long-term talent.
* During that same period, as one would suspect, there is a descending line of draftees that signed an extension with the highest point starting in the first round of this period descending to the seventh round.
* We describe a team’s “hit rate” as signing their own drafted player to a contract extension.

## Player Overview

* Between offense and defense, the highest number of resigned players came from the wide receivers and cornerbacks respectively.
  + The number of wide receivers extended in this period is about 25 players, compared to about 19 players select
  + What is interesting to note in these findings is that running backs are not that far behind wide receivers at all in the number. Considering today’s environment where running backs are reportedly considered high values in the early rounds.

## Team Overview

* Across the measured years, we continue to see a growing trend of teams resigning their own players, based on the numbers.
* In that time span, the top half of the league tends to extend about 20% or more of the time of their respective draftees.
  + Some “hit rates” go to a third (33%) to a little over
* Of the teams with the highest hit rates, their extensions made to their draftees were mainly around cornerbacks and a tie between running backs and wide receivers, following the trend mentioned earlier in the player overview.

## Winning Percentage vs. Hit Rate

* There are arguments both for and against whether a high hit rate ultimately determines a higher winning percentage.
  + For example, the team with the highest hit rate, Indianapolis (48.8%), is just barely underneath the 50% split between wins and losses. The counter to this metric is that teams above the 50% win rate, like Miami, New Orleans, and Dallas have all drafted players they wanted to keep in house and thus their results have proven to improve the team’s standing within the league.
* Considering the opposite side of the graph, there are a fair number of teams landing at or below a 20% hit rate on their players and very few teams near the 50% split on wins.
  + Most notably, the New York Jets are showing an abysmal hit rate less than 10% and seeing a very poor 32.5%-win percentage.
  + Kansas City is an extreme outlier in this area. While their hit rate in this period is less than 20%, they have the **highest** win-rate of all the teams in the win percentage time span (nearly 76%)!
* While the r-value of the graph is 0.1150, there is a notable trend upwards of the hit rate and a team’s win percentage that cannot be ignored. With the exception of one team above a 50%-win percentage, what might explain the lack of improvement from other teams who do have a higher hit rate in the top half of the league could have something to do with coaching, management, or other factors.

Overall Report:

It can be determined based on the data that most teams who have higher hit-rates with resigning their own players (about the top half of the league) do not always see higher win rates which can equate to better output on the field, such as playoff appearances and Super Bowls.

Of the players that do get resigned, majority of them tend to come from the offensive side of the ball, with running backs (RBs) and wide receivers (WRs) holding the top two spots of the chart, vs. on the defensive side of the ball, cornerbacks were at the top. While Quarterbacks are widely considered one of the most important positions in American Football, there are not as many drafted and used compared to these other offensive positions which have more than one rotated in during on a play-by-play basis.

Some of the struggles and possible caveats that we can take from the data is that these are public libraries and are not from the original source (NFL). It is possible that some of the information may not be the most accurate. In addition, some naming conventions used in curating the DataFrames may need to be adapted to more recognizable formats to not confuse the audience listening to this information. This is an important part in the process that our team understood during retrieving and formatting the data to answer our questions.