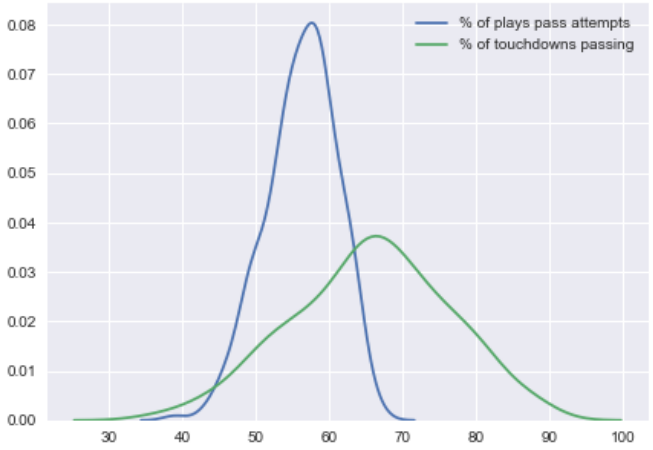
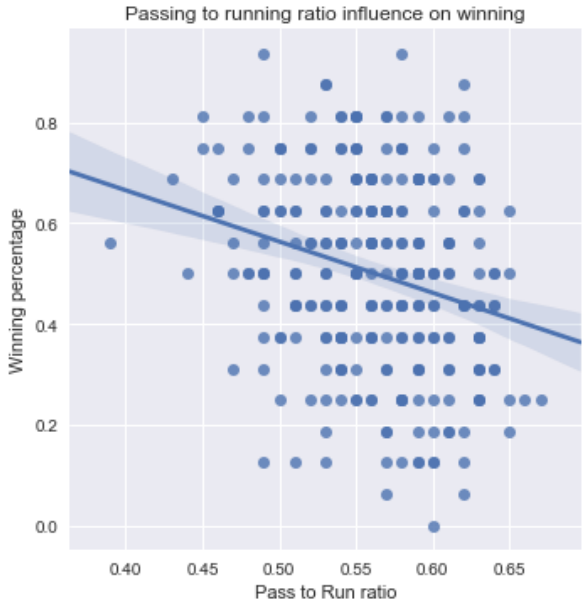
NFL Matchups - Exploratory Analysis

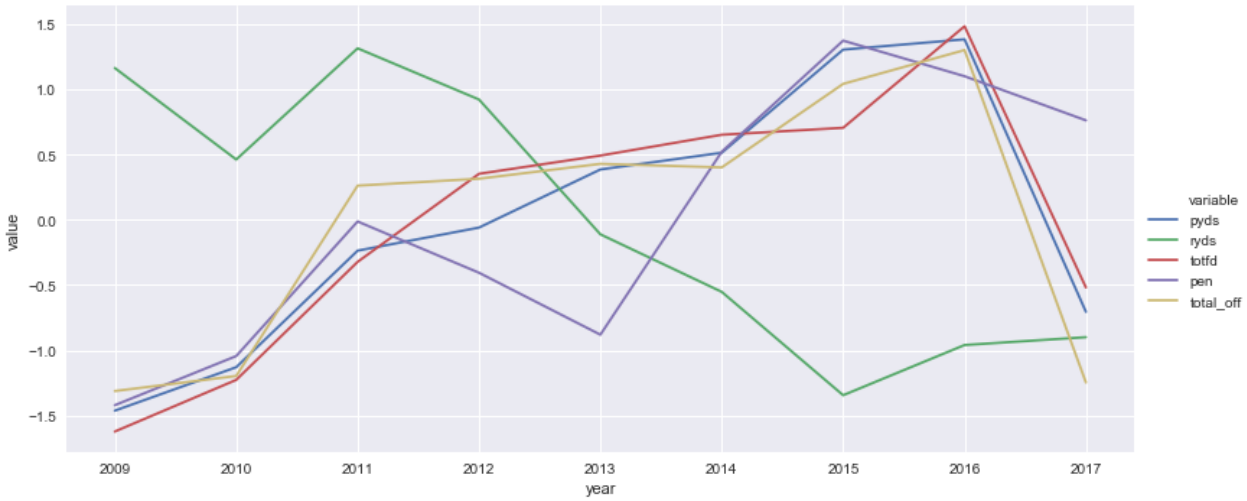
The first topic analyzed was the team play composition. We looked at the distribution of pass to run play ratios. As well as pass to run ratios for scoring, quarterbacks, and running backs. As for play calls passing was firmly the dominate play call accounting for about 57% of plays and even higher for plays that scored. I hypothesized that overall passing is a more effective method of getting yards and scoring so teams that are good at passing pass often and are able to win while teams that are bad at passing are forced to rely on a run game, an overall less effective strategy.



To test this claim I generated a scatter plot to test the correlation between passing and winning and believed that high passing percentage teams would win more. However, the opposite turned out to be true. In fact, the correlation was negative and statistically significant. I then changed the assumptions to the following in order to fit the data. Passing plays are more common because there are more situations where running isn't viable. Also, teams that are winning will run the ball during the end of the game so winning causes a higher run percentage instead of the other way around.

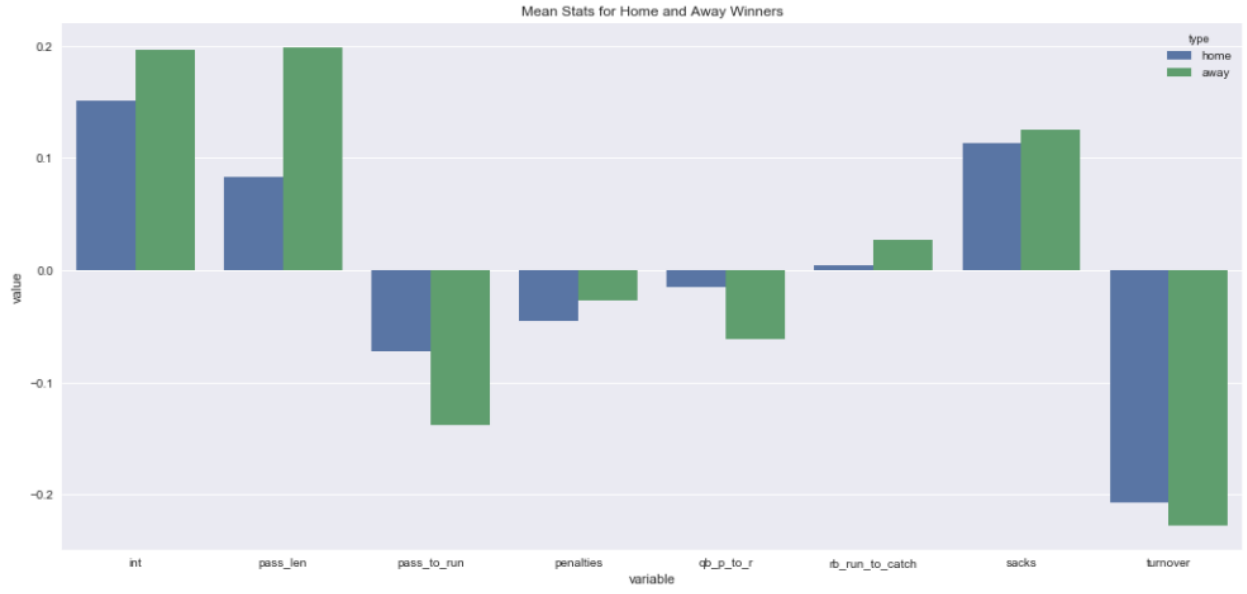


Next, I decided to look at the trends of league wide statistics over the past eight years in order to gain insight how the league has changed. My observations and explanations are then given.



The first observation is the strong correlation between passing yards, first downs, and total offensive yards. As shown earlier passing is the dominant play call so as the yardage for passing fluctuates first downs and total yardage should change with while being relatively independent of rushing yardage. Another thing to note is the increase of penalties (purple). I believe this is due to the NFL being under pressure for preventing player concussions by the media and fans.

I then wanted to look at the mean of normalized statistics for teams that won games and compare both home and away. That is, for teams that won what was there average statistics compared to all teams. By doing this I can hypothesize what traits are key to a winning team and if winning takes a different type of team depending on whether the team is home or away.



For this chart it appears that turnovers are a critical part of winning. The most extreme statistics for teams that won are a high amount of intercepting your opponent and a very low number of

turnovers. Also, low passing percentages are seen for winning teams which align with our previous observations.

One interesting observation of the above chart is that teams that win while away have similar statistics to teams that win while at home except they are more extreme. Because of this I proposed that winning while away is in general harder and therefor the statistics of teams that won while away were more pronounced in these positive ways than the teams winning at home had to be. In order to further investigate this claim, I calculated the difference in winning percentage of home vs away teams. Home teams are nearly 15% more likely to win than an away team. By conducting a hypothesis test I determined the chance of this occurring due to luck is less than 1%, showing that home field advantage is genuine effect.