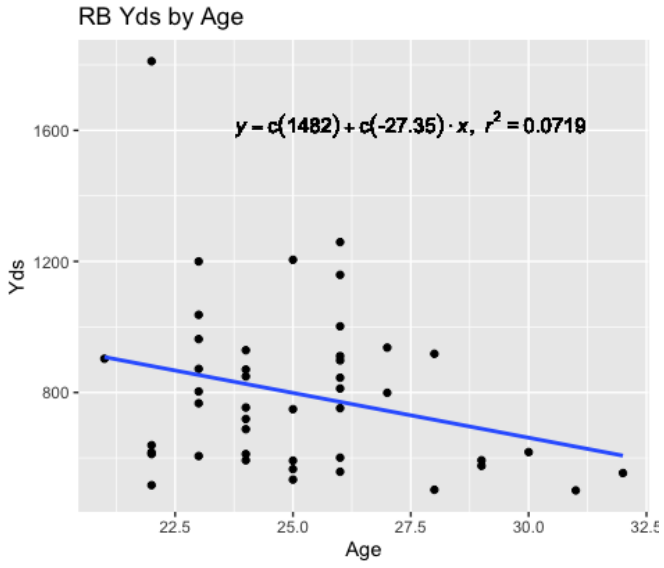
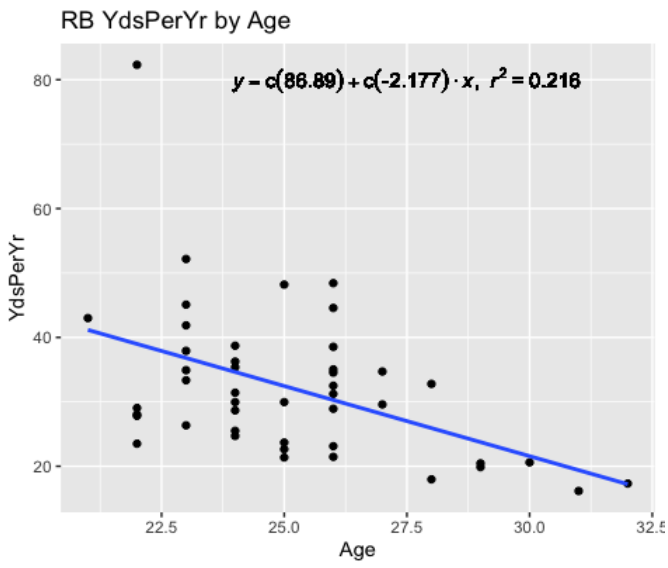
Does Age Correlate With Success in the NFL?
By: Jacob Andrews

Looking at the NFL landscape today, we see great success from rookies and 42-year-old veterans and everywhere in between, but recently there seem to be some trends for success in certain positions. For instance, when looking at the running back position, the average career length is the lowest of all positions at 2.57 years, compared to a 3.3 year league average. Teams have responded to this by not wanting to invest in their running backs, often finding other options through the draft and free agency. Some recent "casualties" of the position were Todd Gurley of the Rams and Leonard Fournette of the Jaguars, who were successful but not retained. Seeing this recent trend in the NFL made me curious if the stats would show the same story when comparing across ages/positions/yards per player. In my study, I researched from the NFL 2021 season the 4 skill positions on the offensive side of the ball (QB, RB, WR, TE), analyzing yards per player, and generated two separate graphs - yards by age, and yards per year by age - looking for comparisons. Evaluating these graphs, I will be using data to see if there is any synchrony of trends with age and performance in the NFL by position.

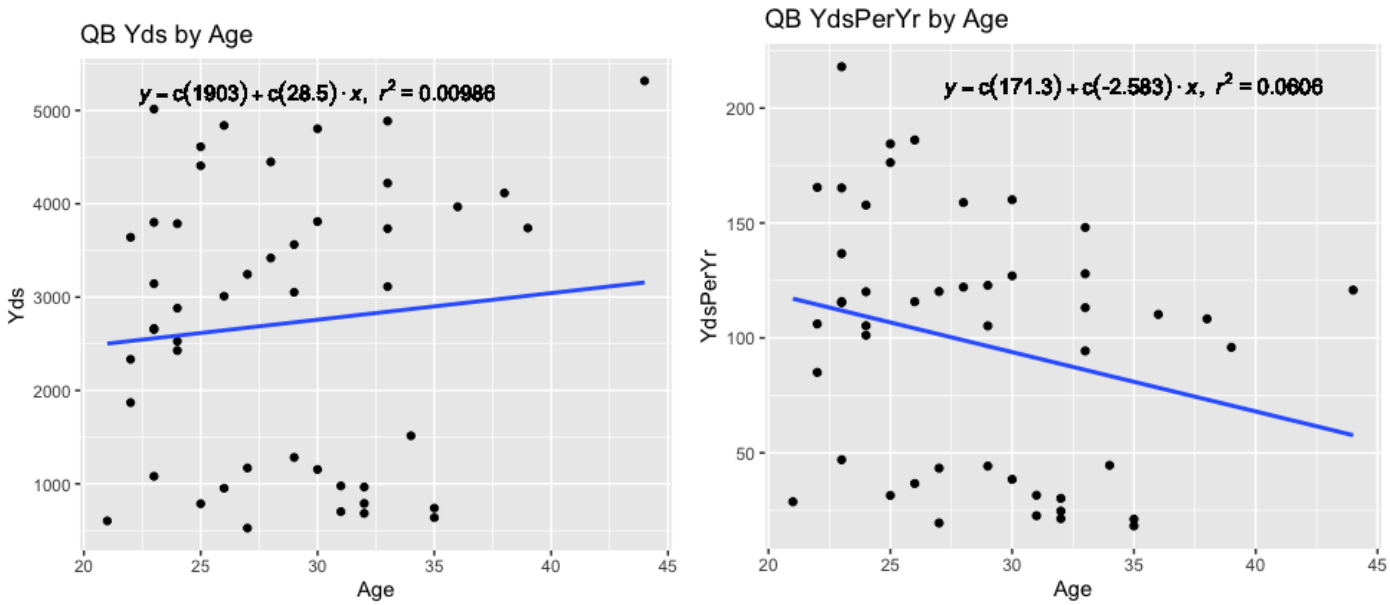
In my analysis of the data, I excluded any players under certain thresholds of yards for each position group in order to get rid of players that could've skewed the data with numbers below statistical variance. The thresholds I set in making my datasets: QBs and RBs over 500 yards, WRs and TEs over 200 yards. The QBs and RB's cutoff reflects their time on the field and productivity, which is typically higher because of player rotations and play-calling. By analyzing the YdsPerYr statistic, my main goal was to make a statistic that combined Yards and Years and brought the data closer together on the graph to see a better line of best fit.



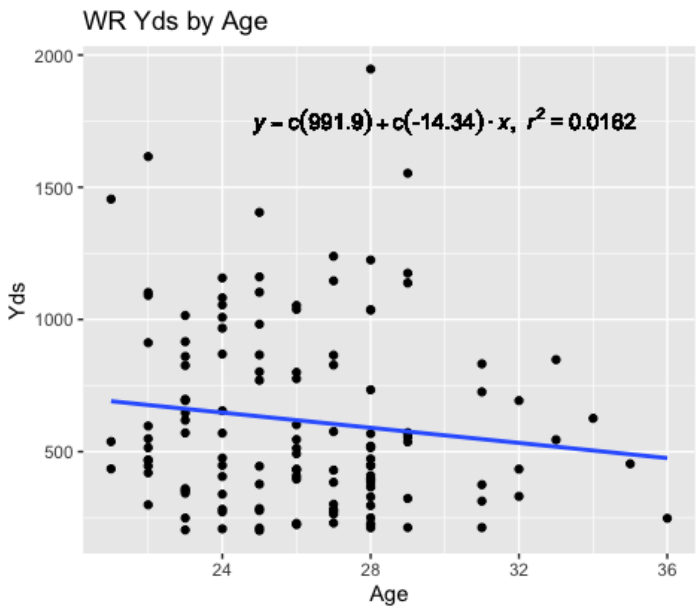


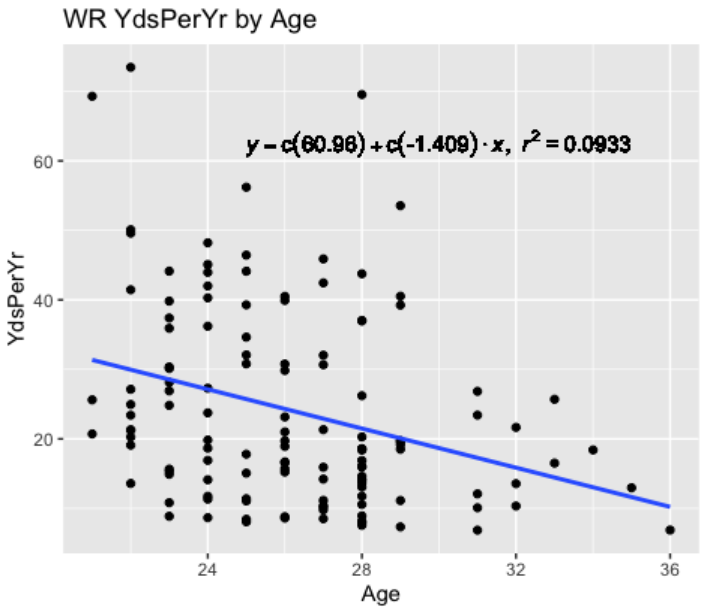
My main inspiration for looking at the data was the running back position. As you can see from the significant negative slopes in both graphs, most running back production is in the early parts of players' careers. Between the ages 22 and 26, we see the largest amounts of production, with a big fall-off after the fact where none of the 5 oldest RBs top 650 yards. There

is an outlier with Jonathan Taylor at 22 years old that brings the production of young players up by a lot, but even without him, the strong correlation would still be there. When looking at the R2 values in both graphs, we can see that the YdsPerYr graph has a way higher value with its data points way closer to the regression line. This only strengthens our argument and confirms a strong negative relationship between age and production at the running back position.

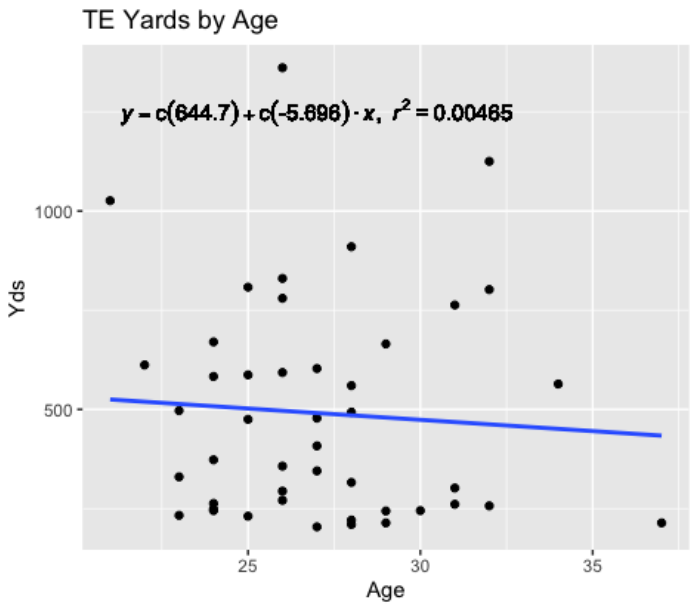


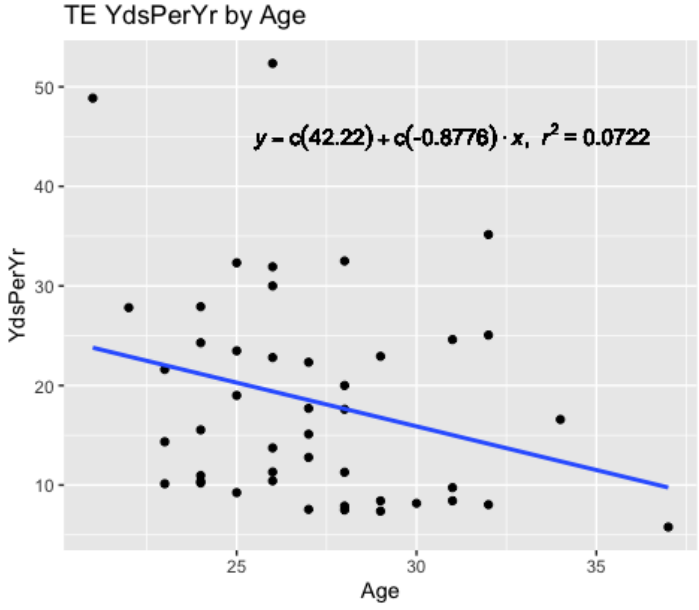
The next position group I collected data to analyze was Quarterbacks. Looking at this data, it is far from what we saw with the running backs. As we can see in the graphs, there is not a consistent trend and the data points are all placed sporadically. The distribution of the data like this leads to a very low R2 value, which tells us that age does not have as much of a factor on a quarterback's production. Taking into account how there are star quarterbacks at every age and phase of their career, this conclusion does make sense. The data also makes sense because of how much the league protects its QBs with a focus on roughing the passer, which enables them to have a long career without major injuries and maintain consistent performance.





The third position group, wide receivers, shows a correlation much like what we saw with the running backs. In the graphs above, we can see that there is a consistent negative slope, and just like the running backs, it gets steeper and more concentrated when looking at the YdsPerYr statistic. While the R2 values and slopes are not nearly as big as the running back values, it still shows there is a significant negative relationship between age and production for wide receivers. This makes sense when looking at the NFL today because of all of the young WRs that are already stars like Jefferson and Chase, but there still are some star older WRs like Adams and Allen that somewhat balance it out.





The last position group I looked at was the tight ends. Much like the what the WRs were to the RBs, the TEs are to the WRs. When looking at the graphs, you can see the consistent negative slope on both of them along with the increase in R2 when moving to YdsPerYr. The smaller values for the slope and R2 compared to the WR and RB values show a smaller correlation between age and production for TEs, but there is definitely still a negative relationship there. The smaller correlation for TEs makes sense, as there are many great young ones like Waller and Pitts, but the older ones like Kelce and Ertz still produce enough to stop a steep negative slope like the WRs and RBs.

In conclusion, from this analysis, we have learned that most skill positions on offense have some sort of negative relationship between age and production, with the exception of quarterbacks. As I predicted in the introductory paragraph, the running backs have the strongest negative relationship compared to the others. Wide receivers are the second strongest and tight ends are the third, which makes sense as we talked about above. If I were to go back and do this experiment again, I think I would try and get a dataset that had multiple seasons in it to get more data points along with finding a specialized statistic for each position that better conveys production instead of just yards. In the end, though, we found out that age does correlate with success at some positions in the NFL and teams have begun to notice the trends and make moves accordingly. It will be interesting to see if this trend continues and if it gets stronger at certain positions in the future!

Sources for my Research

Datasets: https://www.pro-football-reference.com/years/2021/
Linear Model Equation for R script:
https://www.roelpeters.be/how-to-add-a-regression-equation-and-r-squared-in-ggplot2/
Running back durability statistics:
https://tfclarkfitnessmagazine.com/why-running-backs-dont-last-long-in-the-nfl/
https://www.statista.com/statistics/240102/average-player-career-length-in-the-national-football-league/