Which NFL Statistics Best Predict Team Wins? A Multi-Year Analysis

Team Members:
Jonathan Irwanto
Fischer Harrison
Levi Robertson
Aaryan Thacker

Our Topic, Motivation, and Hypotheses

We are looking to see which stats can predict the wins of different NFL teams over the course of the season. We are comparing matching offensive and defensive statistics.

Our motivation comes from wanting to better understand the ways teams win in football and what winning strategies could be. This could also help coaching and front offices decide what the best priorities should be when creating a roster.

H_0: The {statistic measured} doesn't statistically contribute to the amount of wins that a team obtains in any given NFL regular season from 2014 - 2024

H_a: The {statistic measured} statistically contributes to the amount of wins that a team obtains in any given NFL regular season from 2014 - 2024

The {statistic measured} changes for each statistic we measured. This means there are 6 null hypotheses for each year.

$\alpha=0.05$

What We Are Measuring

Pro football reference: Randomly 4 select regular seasons 2014-2024.

The selected seasons: 2021, 2022, 2024, 2019

The statistics, found r and r^2 when matched with total games won-

- Passing Yards/Allowed Per Game
- Rushing Yards/Allowed Per Game
- Points Scored/Allowed Per Game

As well as-

- Number Of Wins
- Team Name
- Year

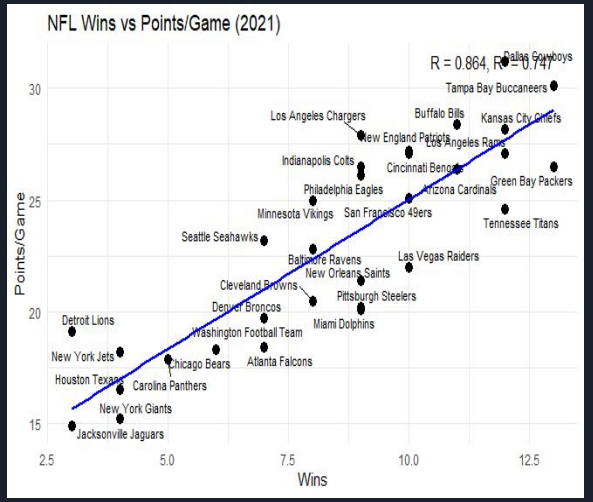
We used simple linear regressions for each statistic to find the p-values.

2021 Graphs and Statistics

Offense Stats:

- Points per game had an R value of .864, P<0.001
- Rushing yards had R value of 0.242 P = 0.31
- Passing Yards per game had an R value of of 0.693, P=0.42
- Overall we found that only one data point produced statistically

  significant data



Defense Stats:

- Passing Yards Allowed had an R value of -0.1111, P =0 .09
- Rushing Yards Allowed has an R value of -0.486, P =0.27
- Points Allowed has an R Value of -0.625, P = <0 .001

Comparing the defence stats to the offensive stat it showed that for the year 2021 offense stats of Passing Yards a game and Points per game was more correlated with wins than the defensive counter parts. However, rushing yards allowed had a larger R value than the offensive counter part

2022 Graphs and Statistics

Rushing Yards - r:.085 r^2 :.007

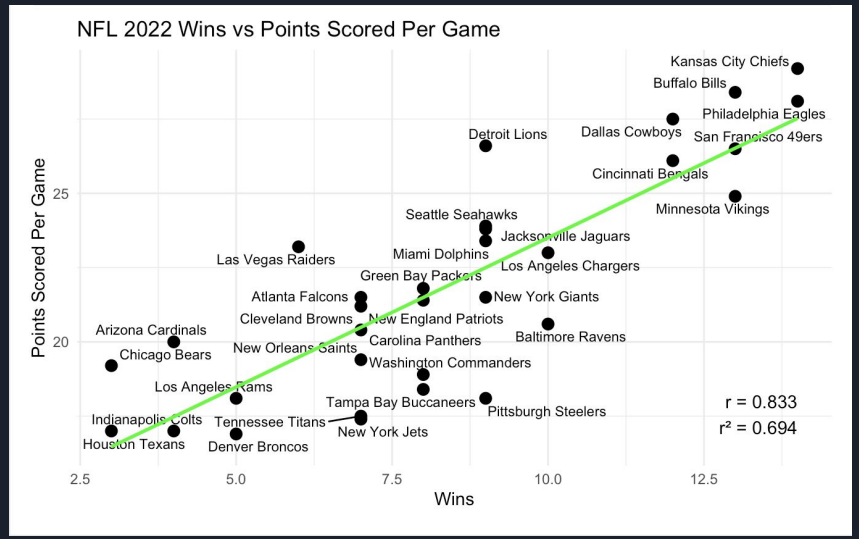
Passing Yards - r:.632 r^2:.399

Points Scored - r:.833 r^2:.694

Rushing Allowed - r:-.369 r^2:.136, more impactful

Passing Allowed - r:.015 r^2:0, less impactful

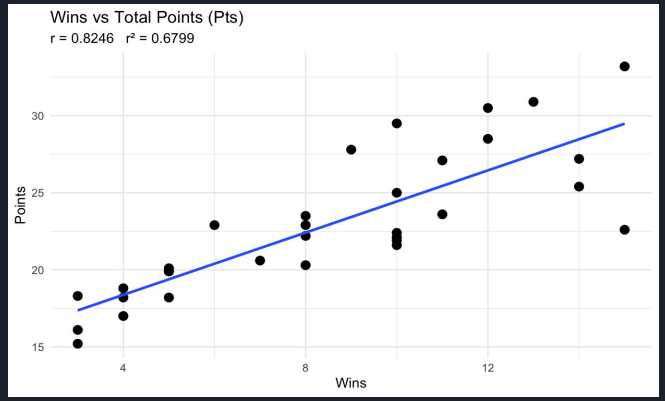
Points Allowed - r:-.546 r^2:.298, less impactful

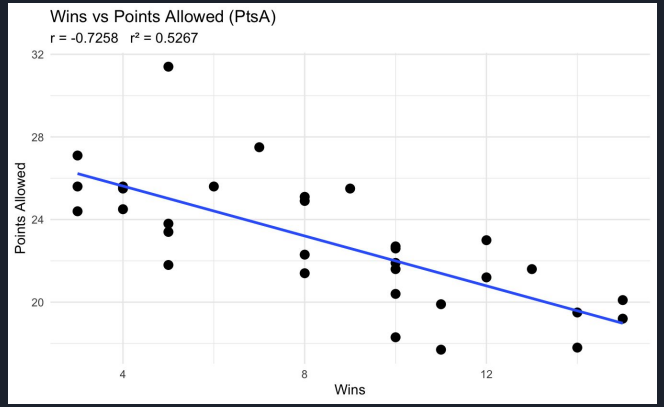


Besides points and points allowed, the most impactful statistics were passing yards and rushing yards allowed for offense and defense respectively. However, the only statistically significant results were Points allowed and points for, with p values less than our significance level

2024 Graphs and Statistics

Total Points and Total Points Allowed were easily the two most significant statistics in 2024 when determining whether or not a team had a winning or losing season with r values of r=0.8246 and r=-0.7258, respectively. These graphs both have p-values < 0.01, with the other statistics having p-values greater than our alpha value, making them statistically insignificant.

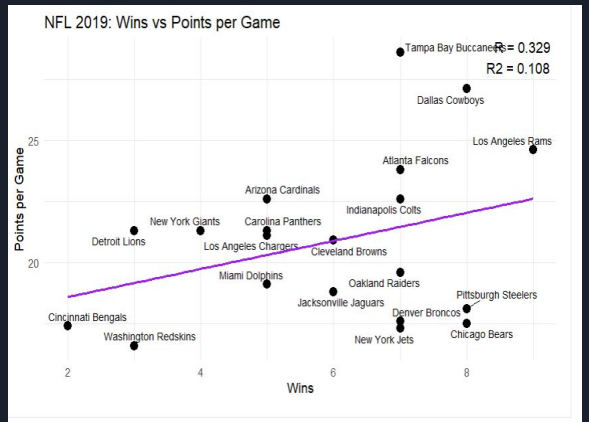




2019 Graphs and Statistics

In 2019, Points Allowed per Game was the strongest predictor of whether an NFL team ended the season with a winning record. This statistic produced the highest absolute R value and the largest R^2, indicating the strongest linear relationship between any of the six variables tested and total wins. On the offensive side, Passing Yards/Game and Rushing Yards/Game both showed very weak correlations, with low R and R^2 values and statistically insignificant p-values. Points Scored/Game had a moderate positive relationship with wins, but remained less predictive than defensive scoring.

Our hypothesis tests support these findings: while none of the variables achieved statistical significance at $\alpha = 0.05$, Points Allowed/Game still showed the clearest linear relationship with wins. By contrast, all yardage-based metrics had p-values well above the alpha value and extremely low R^2 values, demonstrating limited predictive power.

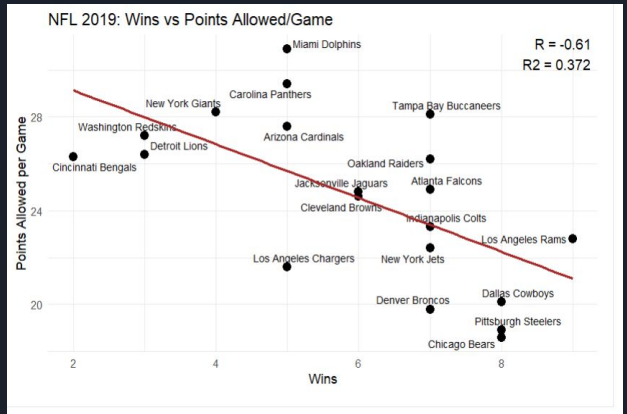


Wins vs Points Scored

r = 0.329, r^2 = 0.108 p = 0.1564

Wins vs Points Allowed

r = -0.61, r^2 = 0.372. p = 0.00428



Conclusions:

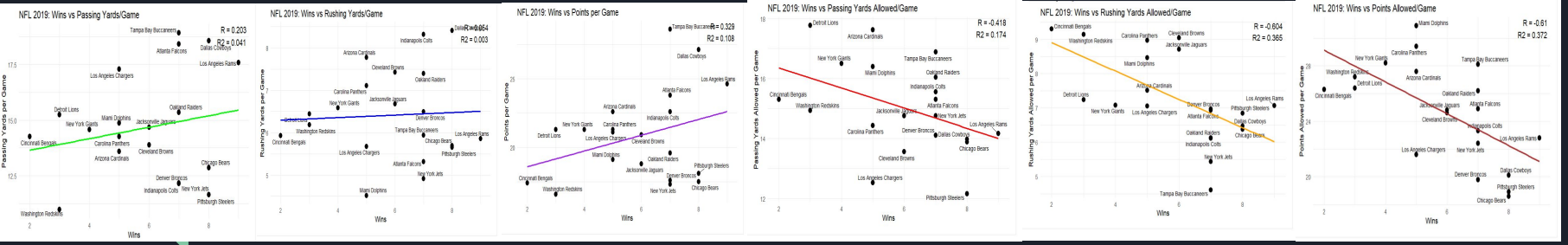
Offense:

- Highest R value for each graph was Points/Game
- Excluding Points per game the most important stat was Passing Yards for 3 out of 4 graphs

Defense:

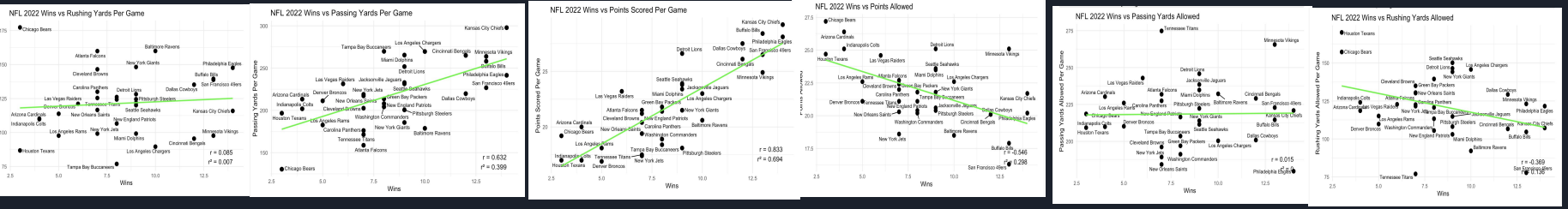
- Highest R value for each year was Points Allowed/Game
- However not far off is Rushing Yards Allowed/Game being the next best stat for all graphs
- Passing Yards Allowed/Game had almost no correlation with win.

After collecting and analyzing the data, our null hypothesis was rejected because Points and Points Allowed were statistically significant

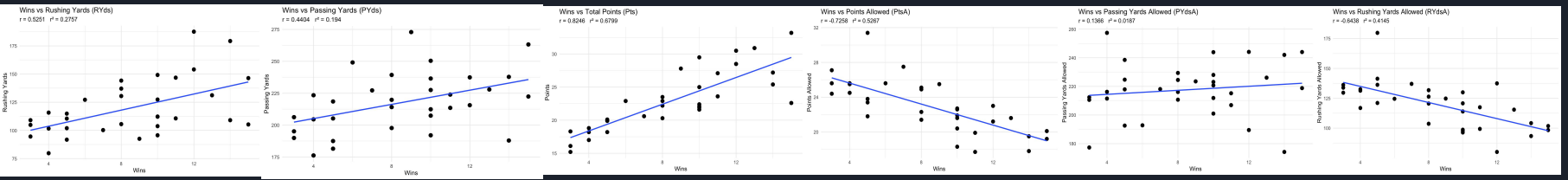


2019

2021



2022



2024