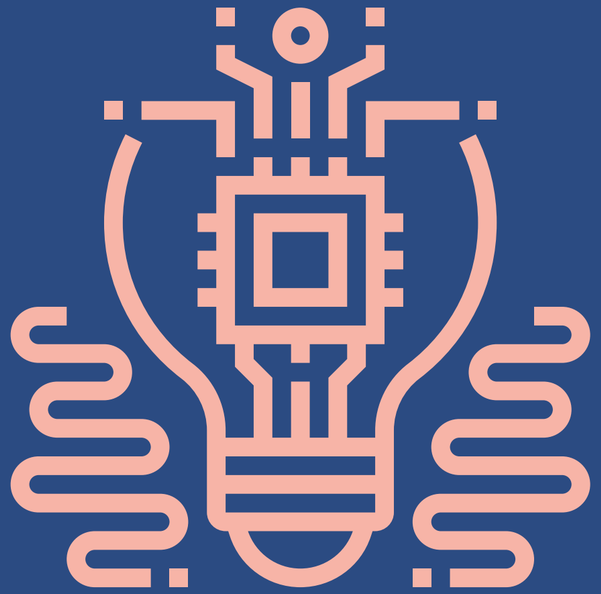


Game of Inches

Applying machine learning techniques to gain that extra edge in the NFL.

Motivation

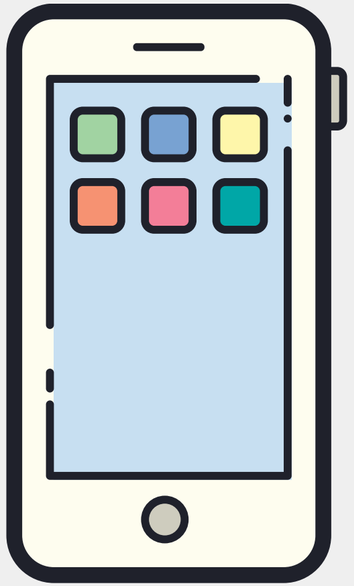
- Develop a machine learning model that business owners can use to gauge how close their team is to making the playoffs.
- Identify which variables contribute the most to winning.
- Create a tool that can aide in player acquisition to target individuals that can boost team's area of weakness.

Financials

Making the playoffs can be a lucrative financial boom for teams. While teams only gain $1-2 million dollars per additional game, it's the TV exposure that brings in the true revenue.

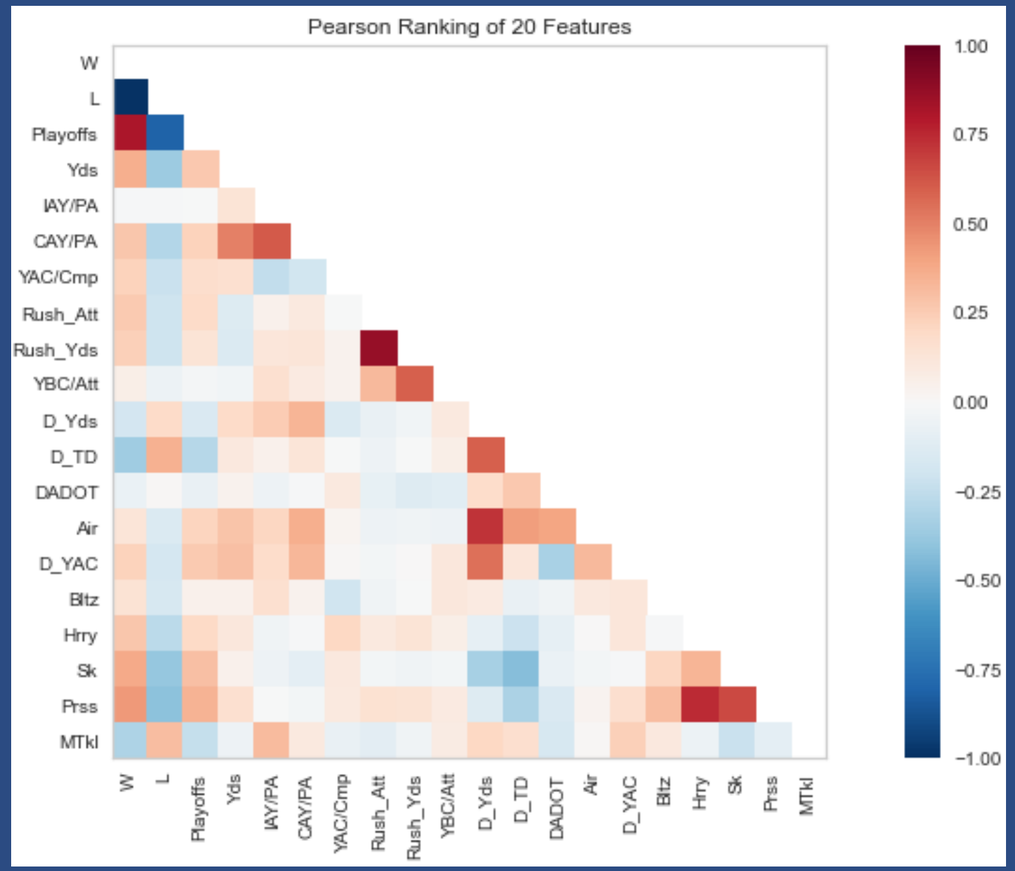
Methods...

- Compiled datasets from 2018-2021 from pro-football-reference.
- Removed team names due to relocation.
- Added conference affiliation and win-loss record.



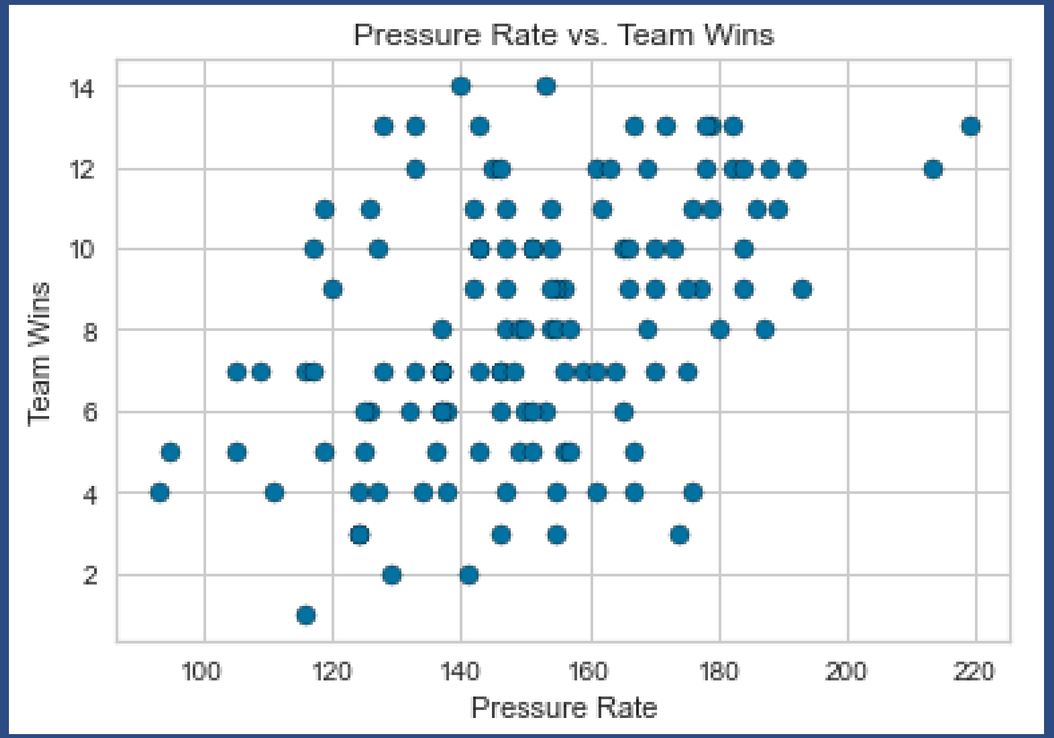
Model Selection

- Linear regression model due to evidence of correlated variables.
- Two individual models constructed, one for AFC and one for NFC.
- Determine correlation coefficient of models.



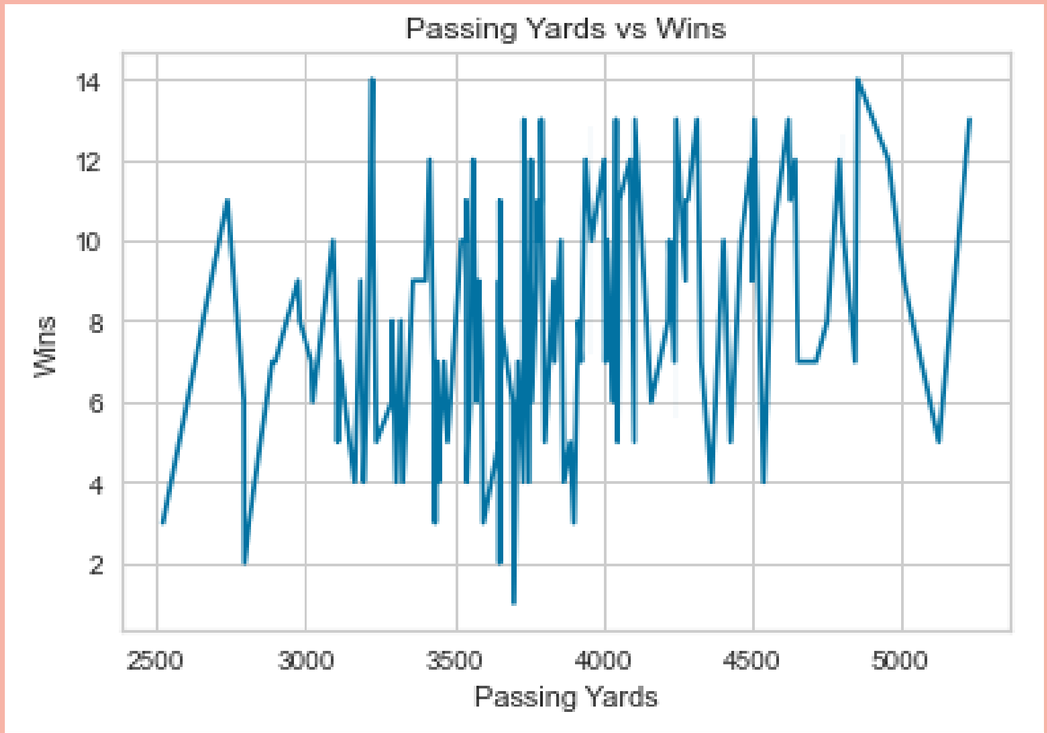
Model Selection, cont...

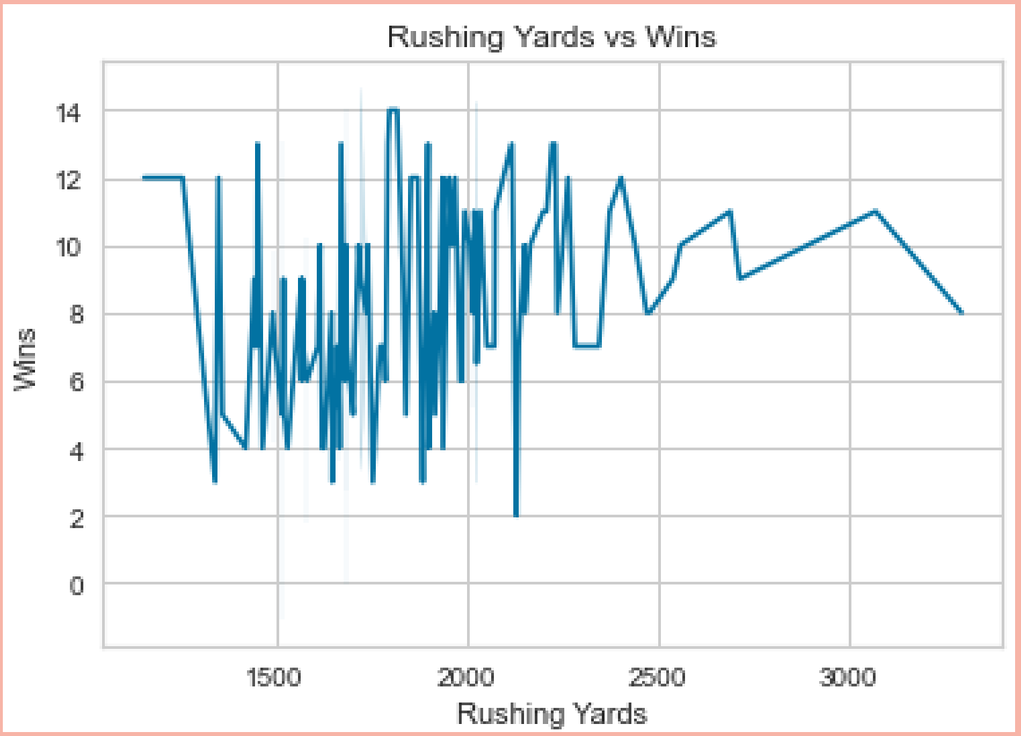
- From Pearson's Correlation Coefficient chart, we're able to see the pressure rate has a positive correlation to wins.
- Scatterplot confirms what the Pearson Correlation Coefficient identified.



Passing vs. Rushing

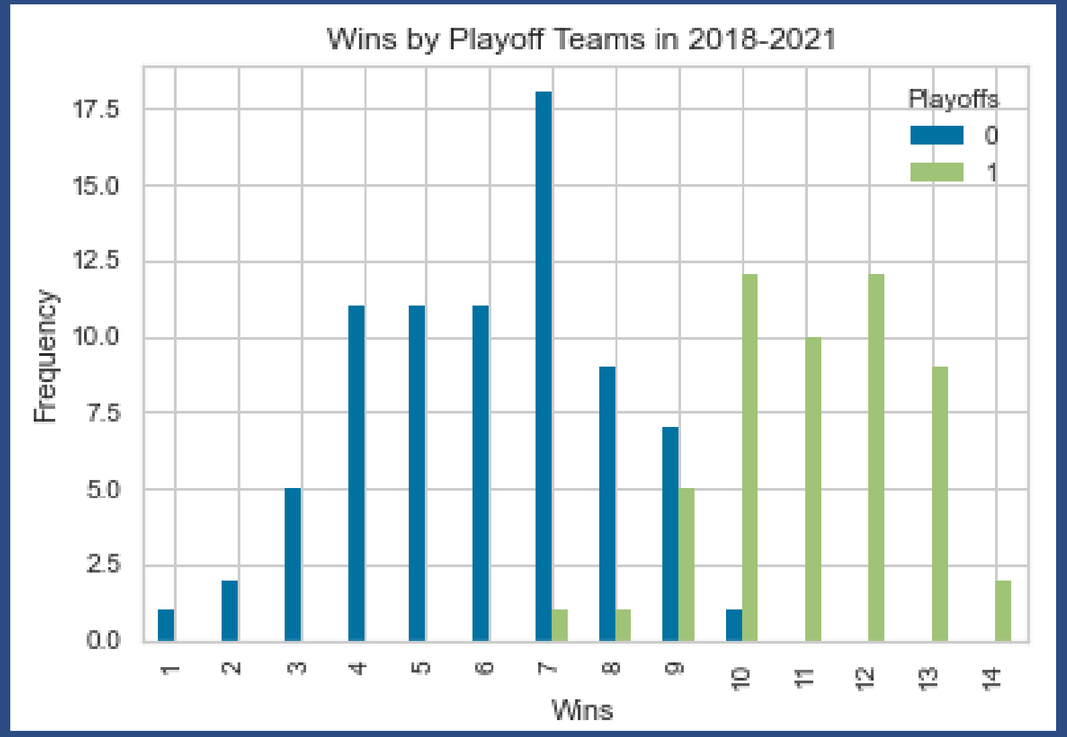
HAS THE RUNNING GAME LOST ITS IMPACT?





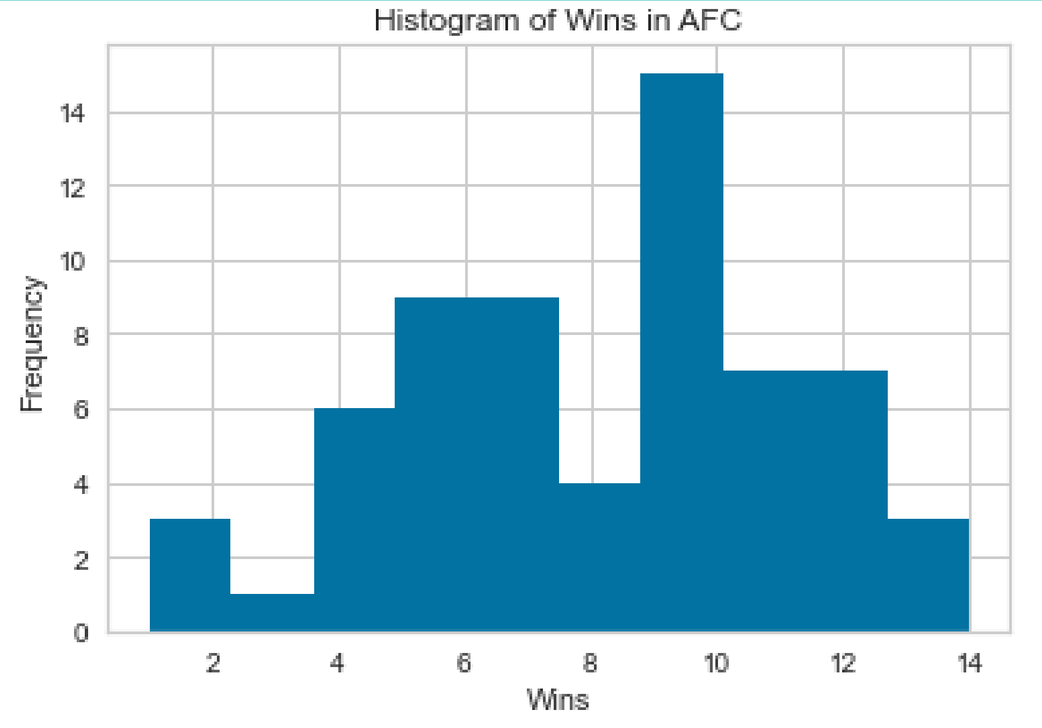
What does it take?

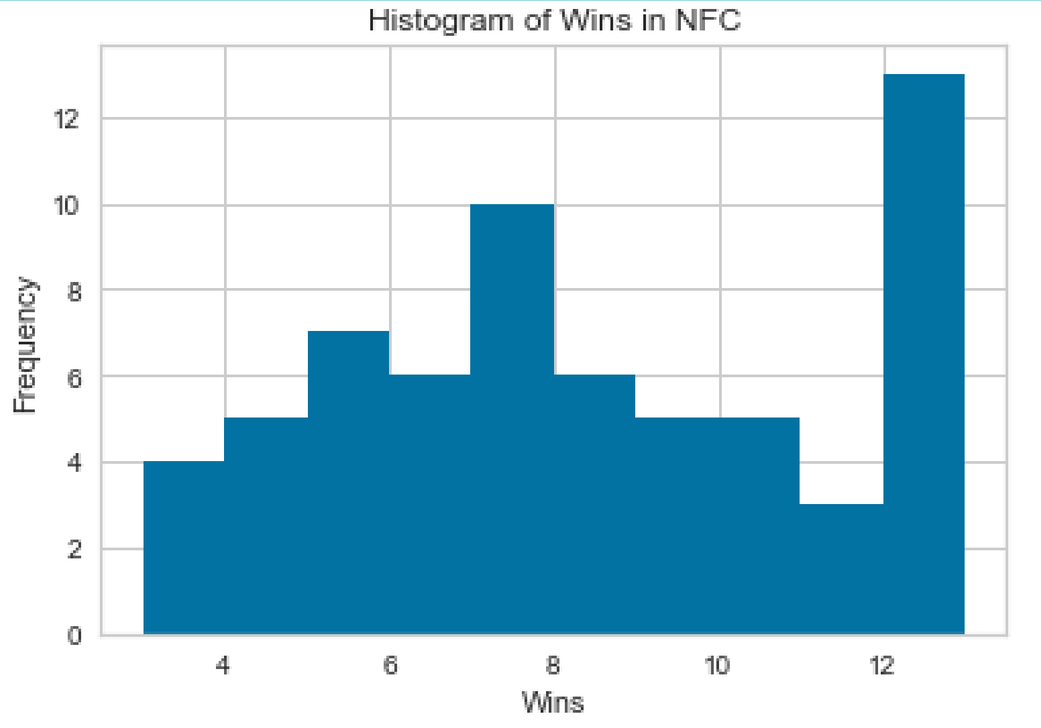
FREQUENCY OF PLAYOFF TEAMS BY WIN TOTAL



AFC vs. NFC

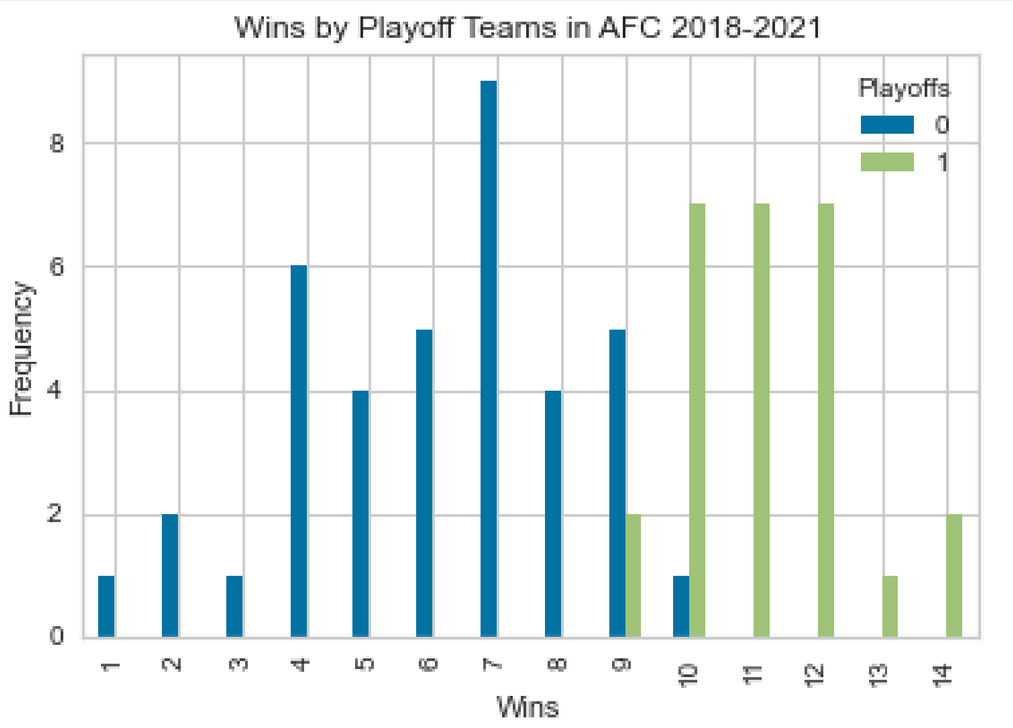
Frequency of win totals per conference

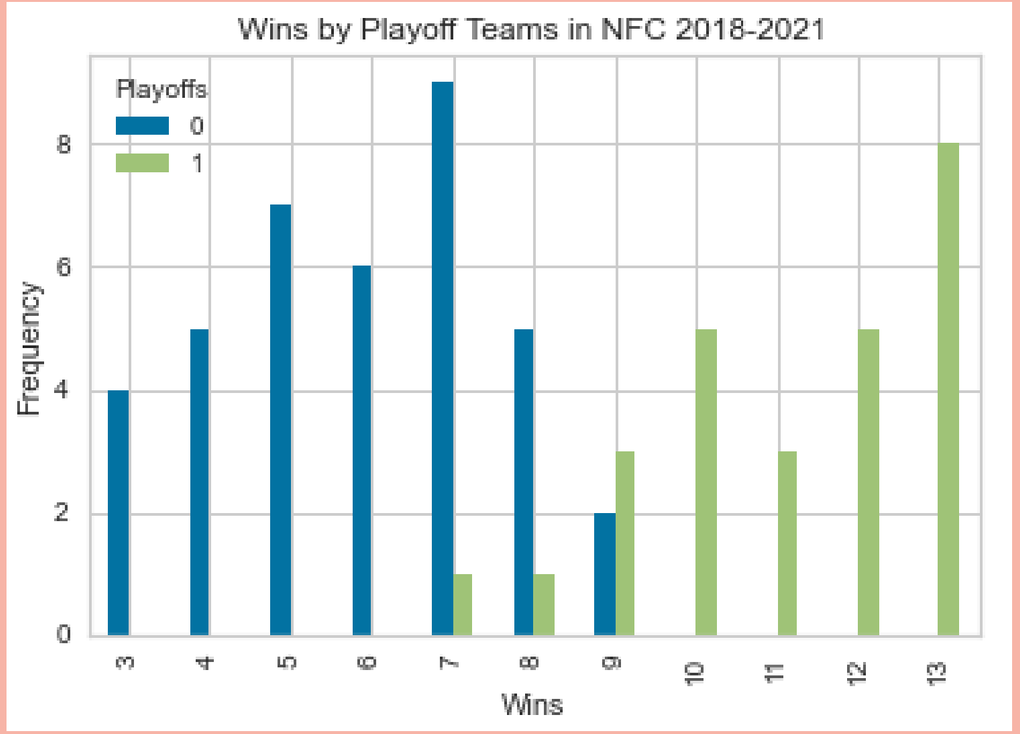




AFC vs. NFC, cont...

BAR FOR PLAYOFF TEAMS PER CONFERENCE





AFC vs. NFC, cont...

MEAN ATTRIBUTES PER VARIABLE BETWEEN PLAYOFF VS. NON-PLAYOFF TEAMS

AFC

Playoffs	W	L	Cmp	Att	Yds	IAY	IAY/PA	CAY	CAY/Cmp	CAY/PA	YAC	YAC/Cmp
0	6.105263	10.052632	350.052632	556.236842	3584.289474	4413.289474	7.934211	2029.210526	5.784211	3.644737	1825.973684	5.223684
1	11.153846	5.076923	367.653846	560.730769	3997.653846	4521.384615	8.103846	2252.000000	6.184615	4.053846	1957.076923	5.311538

NFC

Playoffs	W	L	Cmp	Att	Yds	IAY	IAY/PA	CAY	CAY/Cmp	CAY/PA	YAC	YAC/Cmp
0	5.894737	10.184211	367.631579	577.342105	3807.078947	4630.710526	8.002632	2255.894737	6.118421	3.886842	1851.552632	5.042105
1	11.115385	5.153846	378.423077	561.500000	4014.884615	4394.692308	7.819231	2224.423077	5.907692	3.973077	2024.115385	5.369231

AFC vs. NFC, cont...

MEAN ATTRIBUTES PER VARIABLE BETWEEN PLAYOFF VS. NON-PLAYOFF TEAMS

AFC Rushing

Rush_Att	Rush_Yds	1D	YBC	YBC/Att	YAC.1	YAC/Att	BrkTkl	Att/Br
426.657895	1866.552632	104.052632	1021.842105	2.381579	844.710526	1.965789	27.684211	17.092105
440.653846	1916.538462	112.307692	1060.307692	2.376923	856.230769	1.934615	30.000000	15.784615

NFC Rushing

Rush_Att	Rush_Yds	1D	YBC	YBC/Att	YAC.1	YAC/Att	BrkTkl	Att/Br
415.447368	1804.947368	99.552632	1017.315789	2.447368	787.631579	1.900000	28.210526	16.436842
441.807692	1947.000000	114.961538	1080.807692	2.423077	866.192308	1.969231	29.230769	16.646154

Model Metrics

PERFORMANCE SCORES

Coefficient of Determination of models

AFC model had a coefficient of determination score of .68 whereas the NFC model had a score of -.13. This tells me that the data from the AFC was much more correlated.

Root mean squared error

AFC model had a root mean squared error score of 1.85 whereas the NFC model had a root mean squared error of 2.02.

Actionable Insights

WHAT DOES THIS DATA TELL US?

Pass Rushers are appropriately valued

Putting pressure on the quarterback was found to have the strongest correlation to wins.

Rushing still matters

Even the league has moved to a passing league, teams that are able to rush the ball effectively end up with more wins.

Conferences are similar

There weren't many differences between the two conferences. The AFC had greater parity based on average wins, but the NFC had a higher ceiling.

Q & A...

- Which variable is correlated the most to wins?
- Which variable is correlated the most to losses?
- Is there a difference between the AFC/NFC models?
- Is there a difference between the divisions within each conference?
- Why were these the variables that were chosen for the model?
- What is the biggest takeaway from this project?
- Has COVID impacted this data in anyway?
- Is there something you would change if you did this project again?
- What was one thing that surprised you with the data?
- Which NFL team utilizes analytics the most in today's game and how has that translated to on field success?

References

- Badenhausen, K. (2022, January 14). NFL playoffs generate no financial windfall for individual teams. Sportico.com. Retrieved April 3, 2022, from https://www.sportico.com/leagues/football/2022/nfl-playoffs-generate-no-financial-windfall-1234658372/
- 2018 NFL Opposition & Defensive Statistics. Pro. (n.d.). Retrieved April 3, 2022, from https://www.pro-football-reference.com/years/2018/opp.htm#all_advanced_defense
- NFL. Spotrac.com. (n.d.). Retrieved April 3, 2022, from https://www.spotrac.com/nfl/