Derrick Dollesin

Project Overview

Your task is to undertake a project that demonstrates your analytical skills and creative thinking within the realm of sports analytics. The scope of the project is intentionally broad to allow you the freedom to explore an area that interests you the most. Here are a few ideas to get your thoughts flowing:

- Develop a machine learning model to predict the outcome of a specific game or series of games
- Conduct a detailed data analysis on a set of sports data, showcasing your ability to derive insightful conclusions and trends
- Create a compelling visualizations that tell a story or reveal hidden patterns within sports data.

Guidelines:

- Time Commitment: The project should not require more than 3-5 hours of your time, reflecting your ability to generate impactful work efficiently.
- Submission Format: Please submit your project in a format that best represents your work. This could be a code notebook, a detailed report, a presentation, or any other format you feel showcases your project effectively.
- Evaluation Criteria: We are looking for creativity, analytical depth, clarity of presentation, and , most importantly, your passion for sports analytics.

About:

I was never really interested in sports growing up. I played competitive golf, but never really got into basketball or football or any other mainstream sports. When I transferred to UCSD I was roomed with avid football watchers and was looped into watching games. The sport became much more interesting to me the more I watched, and I became curious about some components of the sport. The one thing I was most curious about was why quarterbacks don't run with the ball more often. Assuming you have a good defense, just take off with the ball and you'll score more touchdowns. This project stems from this curiosity, and my journey to answer my own question.

Summary:

Passing and rushing are two of the most essential skills in American football. The better a quarterback can pass and rush; the more yards, touchdowns, and wins they will accomplish. Star quarterback Lamar Jackson demonstrated the importance of having strong rushing skills in the 2023 season, while other players like Jared Goff and Brock Purdy showed how strong passing skills can dominate in the NFL. But what is more important?

Analyzing data from the 2023 NFL season shows that not only is passing the more important of the two; greater ability to pass the ball will result in more yards, touchdowns and wins overall.

Fig 1. Example Quarterback and Team Data

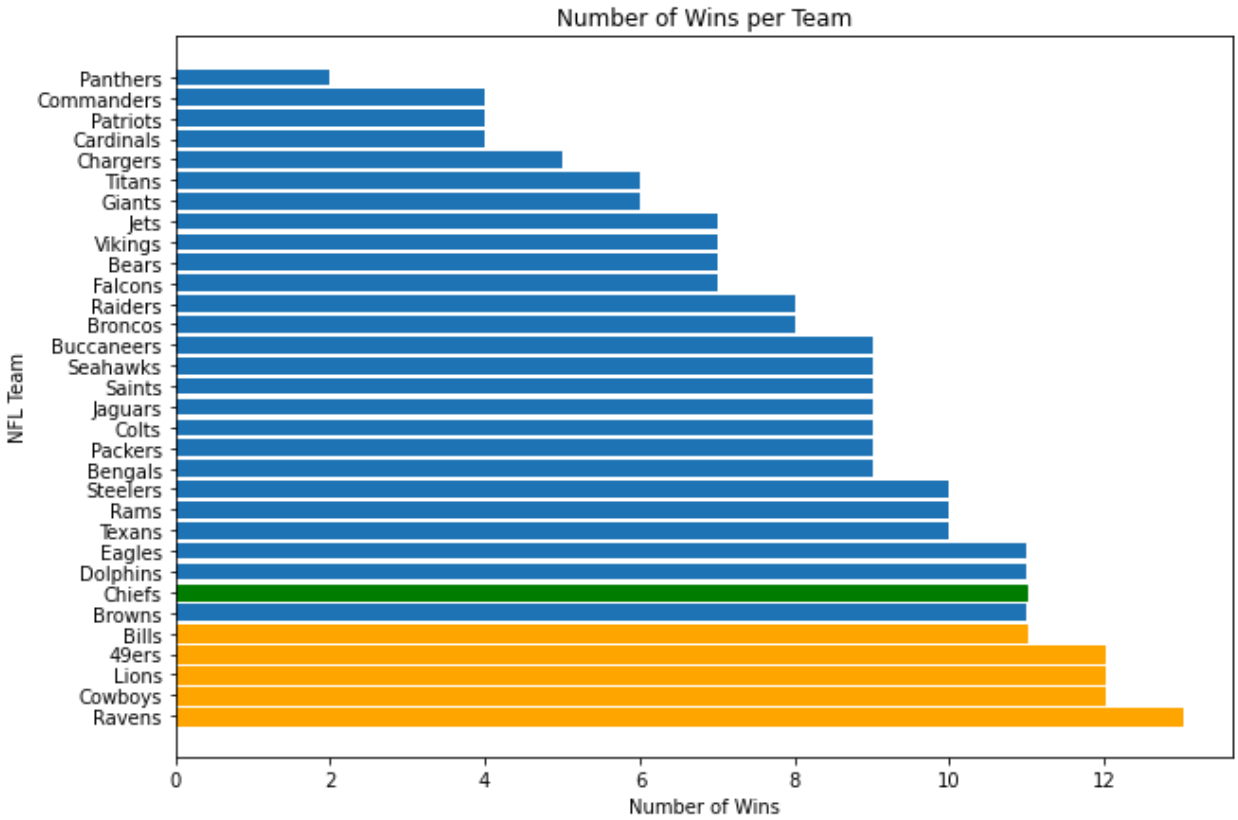
	Player	Team	Pass Yds	Yds/Att	Att_x	Cmp	Cmp %	TD_x	INT	Rate	...	SckY	Rush Yds	Att_y	TD_y	20+_y	40+_y	Lng_y	Rush 1st	Rush 1st%	Rush FUM
0	Lamar Jackson	Ravens	3678.0	8.0	457	307.0	67.2	24	7.0	102.7	...	218.0	821	148	5	4	0	30	48	32.4	4
1	Josh Allen	Bills	4306.0	7.4	579	385.0	66.5	29	18.0	92.2	...	152.0	524	111	15	1	0	23	57	51.4	5
2	Joe Burrow	Bengals	2309.0	6.3	365	244.0	66.8	15	6.0	91.0	...	180.0	88	31	0	1	0	20	9	29.0	0
3	C.J. Stroud	Texans	4108.0	8.2	499	319.0	63.9	23	5.0	100.8	...	331.0	167	39	3	0	0	16	18	46.2	2
4	Anthony Richardson	Colts	577.0	6.9	84	50.0	59.5	3	1.0	87.2	...	29.0	136	25	4	1	0	23	8	32.0	2

5 rows × 26 columns

	NFL Team	W	L	T	PCT	PF	PA	Net Pts	Home	Road	...	TD_y	INT	Rate	1st	1st%	20+_y	40+_y	Lng_y	Sck	SckY
0	Ravens	13	4	0	0.765	483	280	203	6 - 3 - 0	7 - 1 - 0	...	27	7	102.5	180	36.4	52	9	80T	41	246
1	Cowboys	12	5	0	0.706	509	315	194	8 - 0 - 0	4 - 5 - 0	...	36	10	104.6	229	37.3	64	7	92T	40	263
2	Lions	12	5	0	0.706	461	395	66	6 - 2 - 0	6 - 3 - 0	...	30	12	98.1	228	37.6	70	9	70T	31	205
3	49ers	12	5	0	0.706	491	298	193	5 - 3 - 0	7 - 2 - 0	...	33	12	110.2	207	42.2	75	15	76T	34	193
4	Bills	11	6	0	0.647	451	311	140	7 - 2 - 0	4 - 4 - 0	...	29	18	92.2	199	34.4	49	9	81T	24	152

5 rows × 50 columns

Fig 2. Number of Wins per Team



In fig 2 above, the top five teams with the most wins in the last season are highlighted in orange, while the Super Bowl winners are highlighted in green. This is to get a better visual understanding in later graphs of how the team performed overall based on their total rushing and passing yards.

(Note: in fig 3, total rush yards come from data other than just the quarterback.)

Fig 3. Total Rush Yards per Team

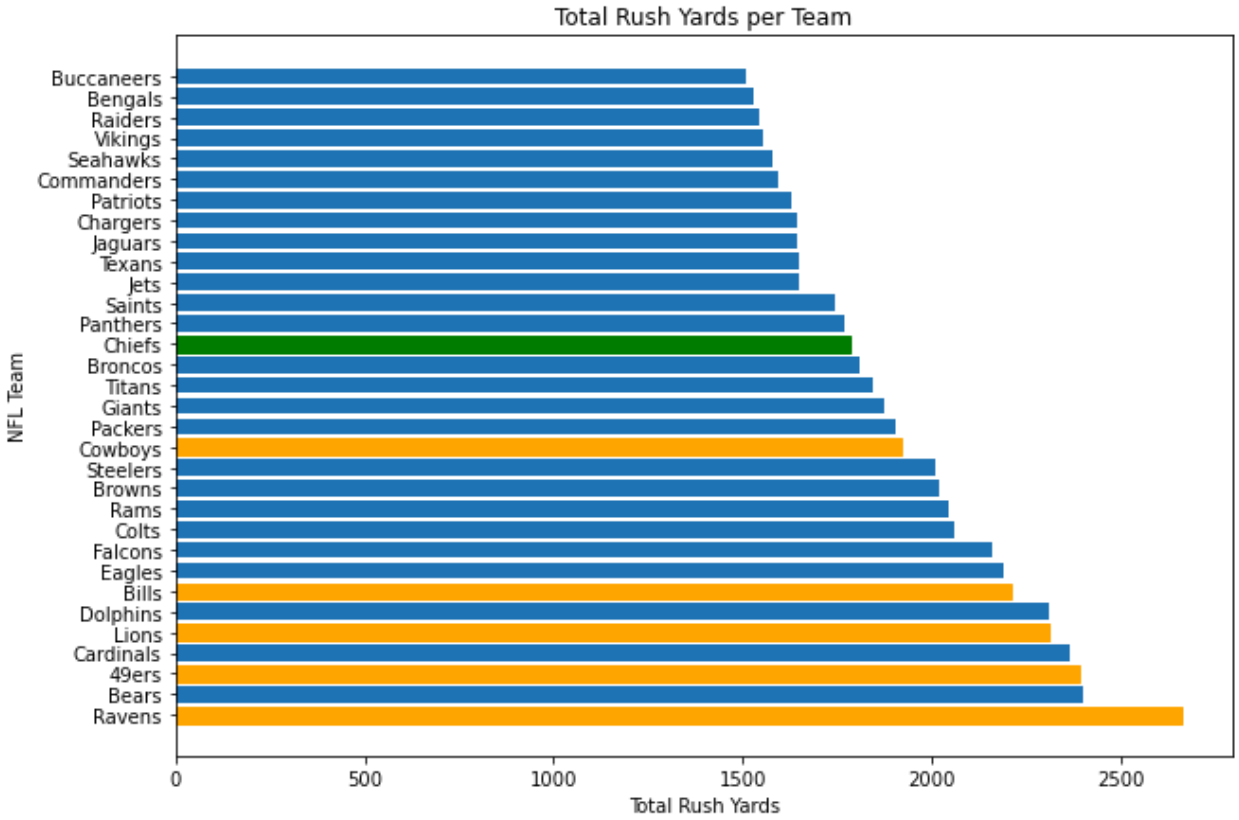
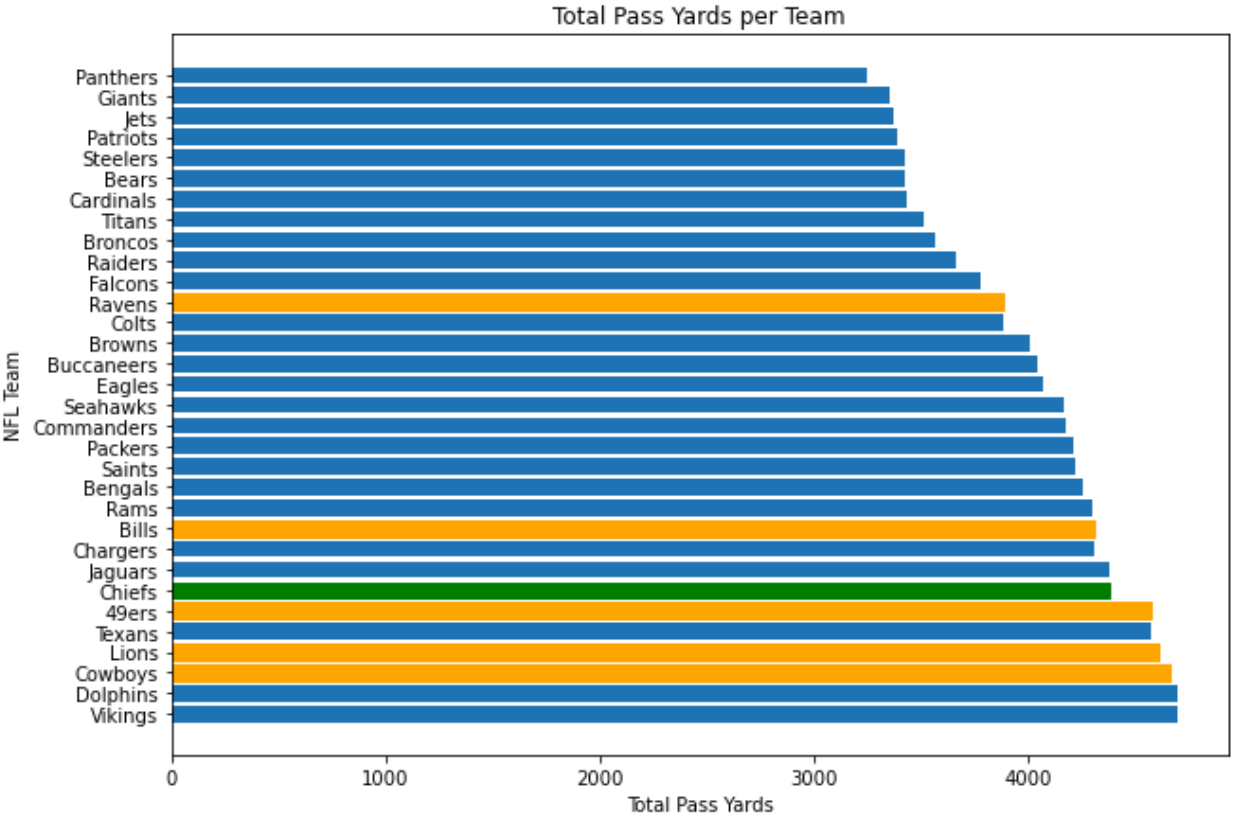


Fig 4. Total Pass Yards per Team



Looking at the graphs, the Ravens had some interesting data. Although they had the most wins and total rush yards in the 2023 season, they ranked 21st when it came to total passing yards. The Cowboys who had tied for the second most wins of the season, ranked 14th in total rushing yards, and ranked 3rd in total passing yards, essentially swapping places in total pass yards and total rush yards. The other three teams in the top five most wins of the season placed somewhere in between the top 10. But what does this tell us?

Fig 5. Scatter Plot of Rush Yards vs. Wins
Correlation: 0.409
95% Confidence Interval for Correlation: [0.071, 0.663]

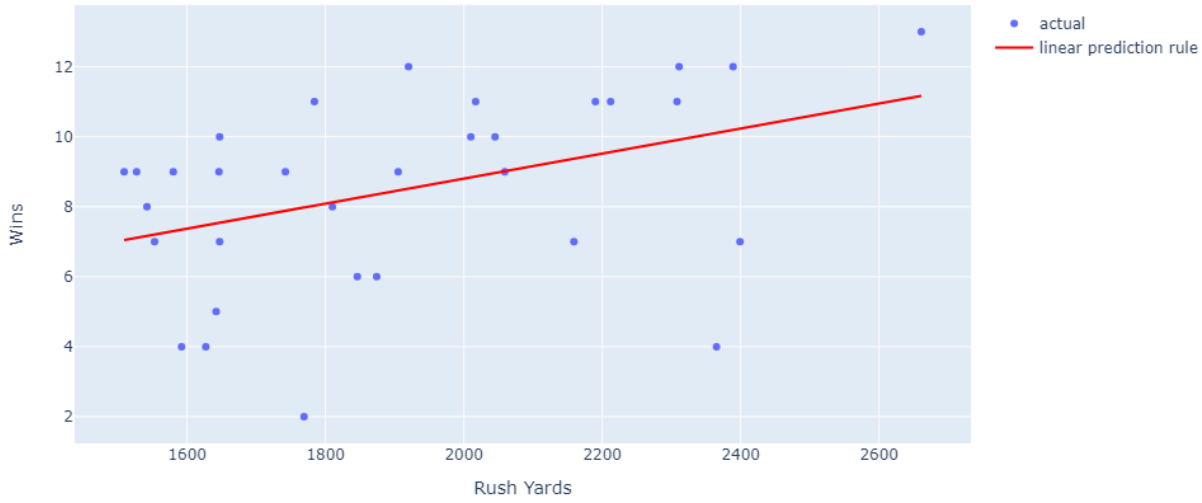
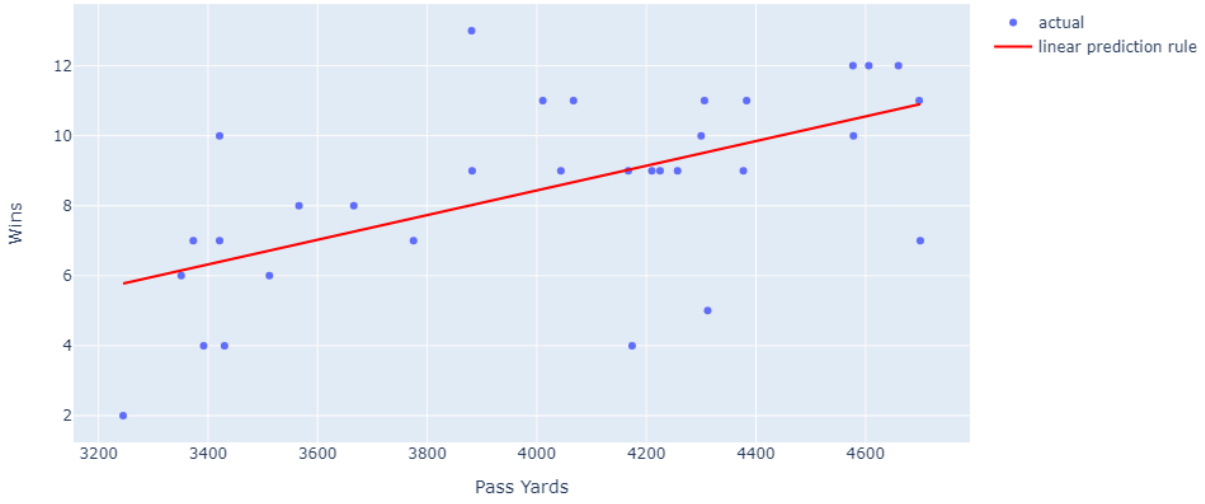


Fig 6. Scatter Plot of Pass Yards vs. Wins
Correlation: 0.592
95% Confidence Interval for Correlation [0.307, 0.78]



When analyzing the teams rushing vs. wins and passing vs. wins, there appears to be a linear correlation between the the two comparisons. When calculating these correlations, it's found that total passing yards has a higher correlation to number of wins than rushing, with a much narrower confidence interval as well. Should quarterbacks prioritize throwing the ball more from this data? Not necessarily...

Fig 7. Number of Rushing Yards per Starting Quarterback

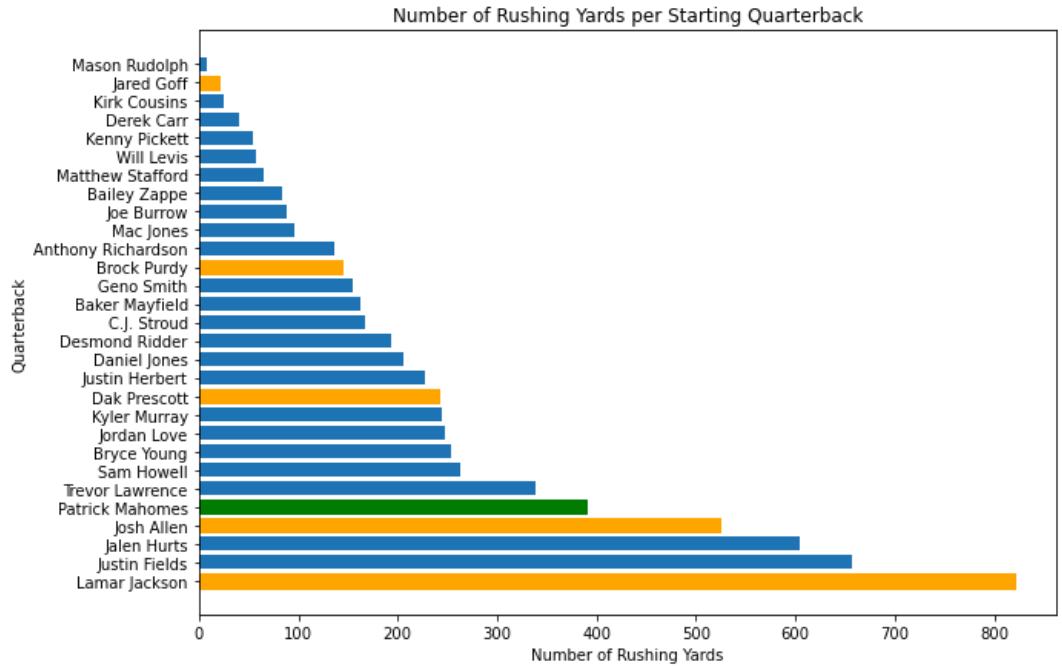


Fig 8. Number of Passing Yards per Starting Quarterback

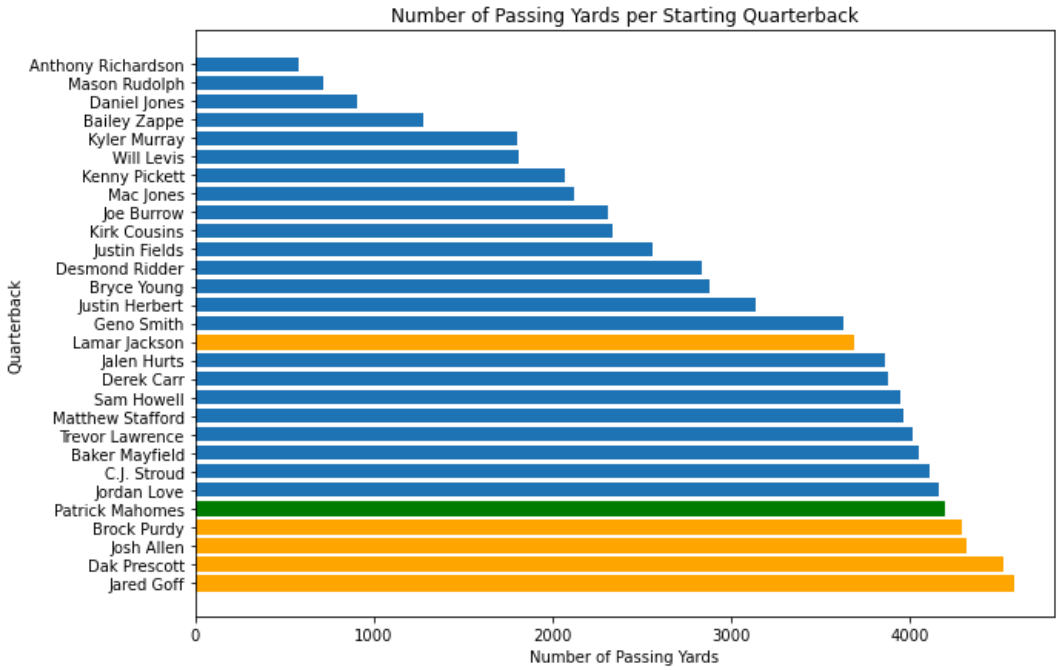


Fig 7 and 8 show quarterback data on their number of passing yards and number of rushing yards. When comparing the Ravens (who had the most wins in season) quarterback Lamar Jackson, and Lions (who tied for the second most wins in season) quarterback Jared Goff, Jackson had by far the most rushing yards with 821 yards while Goff had a measly 21 yards. However in terms of passing, Goff had 4575 yards while Jackson had 3678 yards. Although Jackson far out-ran most of his competition and Goff out-passed his, football comes down to getting the ball into the endzone.

Fig 9. Scatter Plot of Rushing Attempts vs. Touchdowns
Correlation: 0.7333
CI: [0.516, 0.862]

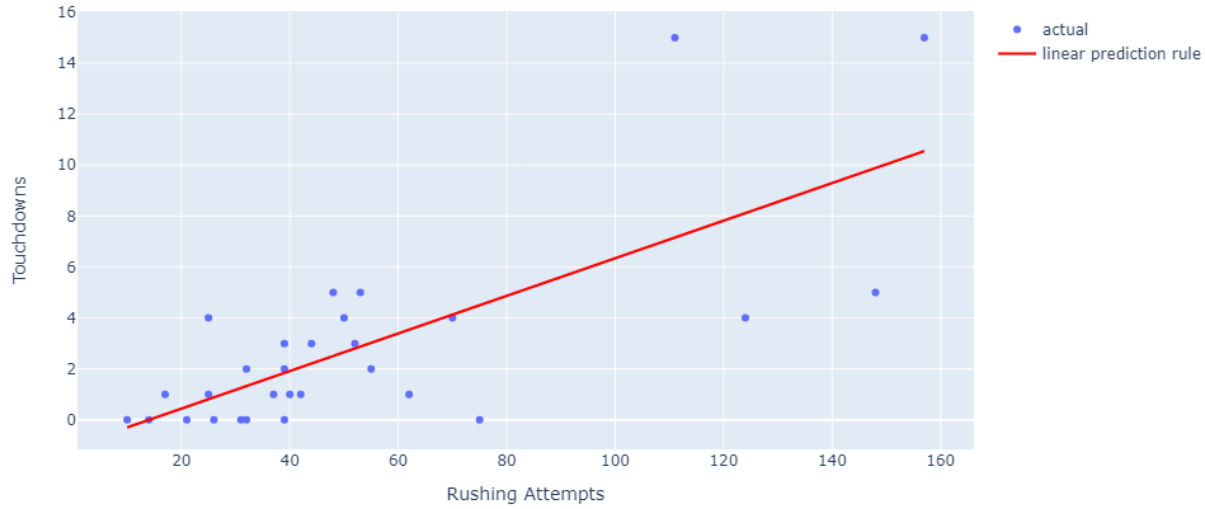
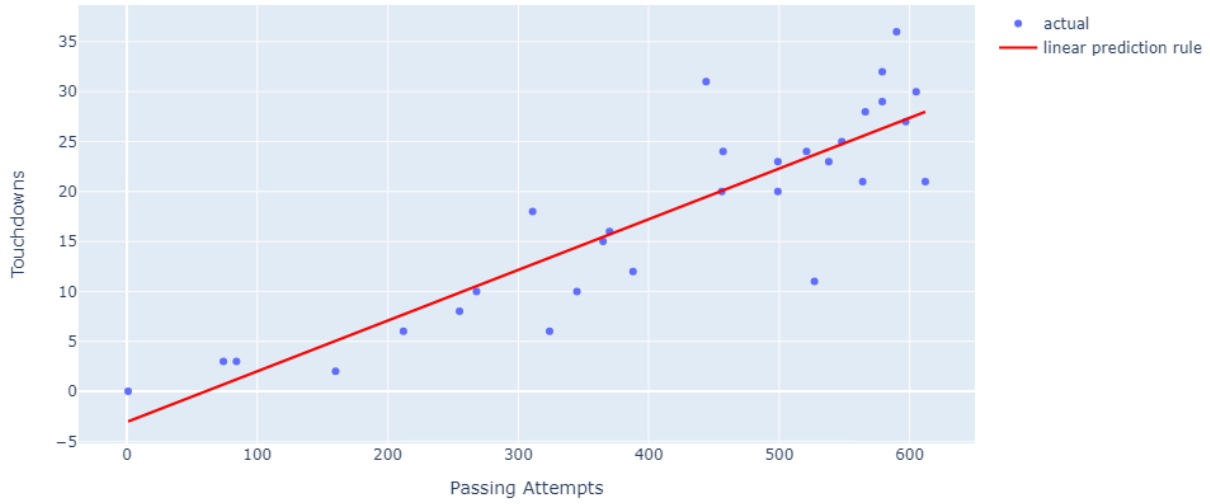


Fig 10. Scatter Plot of Passing Attempts vs. Touchdowns
Correlation: 0.879
CI: [0.764, 0.94]



Analyzing each quarterback in their rushing attempts vs. touchdowns and passing attempts vs. touchdowns shows how passing has a much stronger correlation with touchdowns than rushing, while also having a much narrower confidence interval. Analyzing rushing attempts vs. 20+ and passing attempts vs. 20+ similarly shows that their is a much stronger correlation when passing the ball than rushing and also with a narrower confidence interval.

Fig 11. Scatter Plot of Rush Attempts vs. 20+ Yards
Correlation: 0.609
CI: [0.33, 0.79]

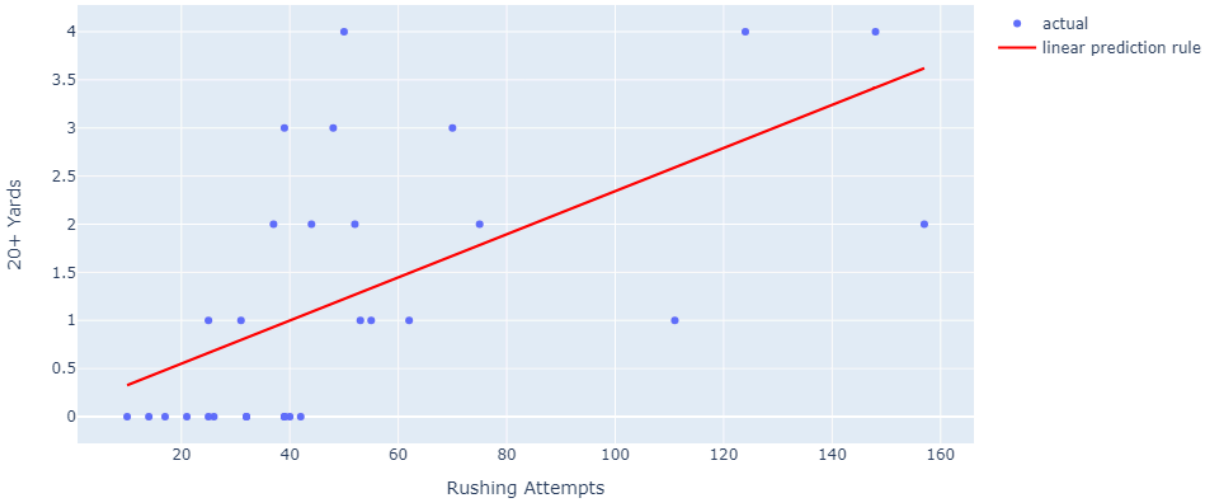
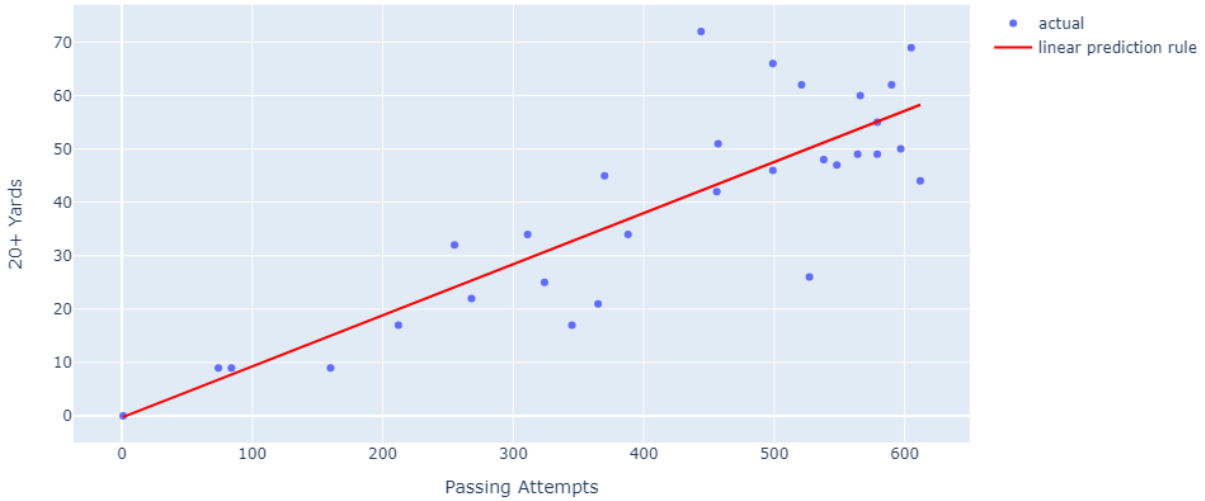


Fig 12. Scatter Plot of Pass Attempts vs. 20+ Yards
Correlation: 0.844
CI: [0.702, 0.922]



Although Lamar Jackson led his team to the most wins in the 2023 season while having the most rushing yards and being mediocre in passing yards, the data suggests the importance of passing over rushing as a quarterback in the NFL. Passing the ball results in higher numbers of touchdowns and gained yards, while also being higherly correlated with wins per season. However, as Lamar Jackson has demonstrated, rushing can make up for a lack of passing. In other words, Ryan McCrary from with the First Pick stated, "It is still more important for a quarterback to be able to throw the ball at a high level, but recently we've seen quarterbacks who aren't great passers lead efficient offenses primarily because they are good rushers." Finding the sweetspot for both rushing and passing with a greater emphasis on passing will lead quarterbacks to greater success in their career.