



WINNING

A look into the current offensive trends in the NFL

HYPOTHESIS

- Recent trend in the NFL has been to throw as often as possible, and rush once in a while just to keep the defense honest.
- I believe a more balanced attack leads to more wins.
- *Goal: Discover the top 10 attributes of an offense that gives the best chances of winning.*

METHOD

- Regular season data will be split into a training and test group at 8:2 ratio and be trained on various classification models
- Top 10 features will be gathered from the most accurate model
- The trained model will be tested against both the regular season test data and the playoff data

DATA



- Source: www.pro-football-reference.com
- Gathered offensive data from 1,000 random regular season games and 110 playoff games from 2009-2018 seasons

DATA IN DETAIL

[Game Logs](#) [Share & more ▾](#) [Glossary](#)

Week	Day	Date	OT	Opp	Score		Passing								Rushing				Scoring				Punting			Downs					
					Tm	Opp	Cmp	Att	Yds	TD	Int	Sk	Yds	Y/A	Cmp%	Rate	Att	Yds	Y/A	TD	FGM	FGA	XPM	XPA	Pnt	Yds	3DConv	3DAtt	4DConv	4DAtt	ToP
1	Mon	September 10	L	New York Jets	17	48	29	52	300	1	5	0	0	5.8	55.8	39.4	15	39	2.6	0	1	3	2	2	3	152	3	13	0	2	26:53
2	Sun	September 16	L	@ San Francisco 49ers	27	30	34	53	329	3	0	2	18	6.2	64.2	100.3	18	98	5.4	0	2	2	3	3	5	211	7	16	0	1	30:26
3	Sun	September 23	W	New England Patriots	26	10	27	36	255	2	1	1	7	7.1	75	101.0	33	159	4.8	0	4	4	2	2	2	90	7	14	0	0	39:15
4	Sun	September 30	L	@ Dallas Cowboys	24	26	24	30	286	2	0	3	21	9.5	80	128.6	20	96	4.8	1	1	1	3	3	4	171	6	11	0	0	29:55
5	Sun	October 7	W	Green Bay Packers	31	23	14	26	170	2	0	3	13	6.5	53.8	99.8	28	94	3.4	2	1	2	4	4	5	228	5	13	0	0	26:40
7	Sun	October 21	W	@ Miami Dolphins	32	21	18	22	209	2	0	1	8	9.5	81.8	136.6	35	248	7.1	1	4	4	2	2	1	61	2	8	1	1	33:39
8	Sun	October 28	L	Seattle Seahawks	14	28	27	40	297	2	1	3	13	7.4	67.5	95.5	13	34	2.6	0	0	0	2	2	4	181	6	10	0	0	25:05
9	Sun	November 4	L	@ Minnesota Vikings	9	24	25	36	143	0	0	10	56	4.0	69.4	76.5	24	66	2.8	0	3	3	0	0	5	222	4	15	3	3	36:45
10	Sun	November 11	L	@ Chicago Bears	22	34	25	42	229	2	2	6	45	5.5	59.5	70.4	24	76	3.2	1	1	1	1	1	4	156	4	15	4	5	32:00
11	Sun	November 18	W	Carolina Panthers	20	19	23	37	215	1	0	1	5	5.8	62.2	87.1	26	94	3.6	1	2	2	2	2	6	281	4	12	0	0	31:40
12	Thu	November 22	L	Chicago Bears	16	23	28	38	222	0	2	2	14	5.8	73.7	65.9	24	111	4.6	2	1	1	1	1	5	227	5	14	1	2	32:46
13	Sun	December 2	L	Los Angeles Rams	16	30	20	34	208	1	1	4	37	6.1	58.8	74.1	26	102	3.9	0	3	3	1	1	6	286	2	12	0	0	31:44
14	Sun	December 9	W	@ Arizona Cardinals	17	3	15	23	96	0	0	1	5	4.2	65.2	73.8	31	122	3.9	1	1	1	2	2	7	317	4	12	0	0	28:16
15	Sun	December 16	L	@ Buffalo Bills	13	14	22	29	208	1	0	0	0	7.2	75.9	106.7	26	105	4.0	1	0	1	1	1	7	317	5	13	0	0	28:54
16	Sun	December 23	L	Minnesota Vikings	9	27	23	43	137	0	0	3	24	3.2	53.5	59.9	23	86	3.7	0	3	3	0	0	6	249	4	16	1	3	31:28
17	Sun	December 30	W	@ Green Bay Packers	31	0	21	33	272	3	0	1	2	8.2	63.6	119.8	38	130	3.4	1	1	1	4	4	4	161	7	14	1	1	36:54

- Target Variable: Win
- Predictor Variables (27 in total):

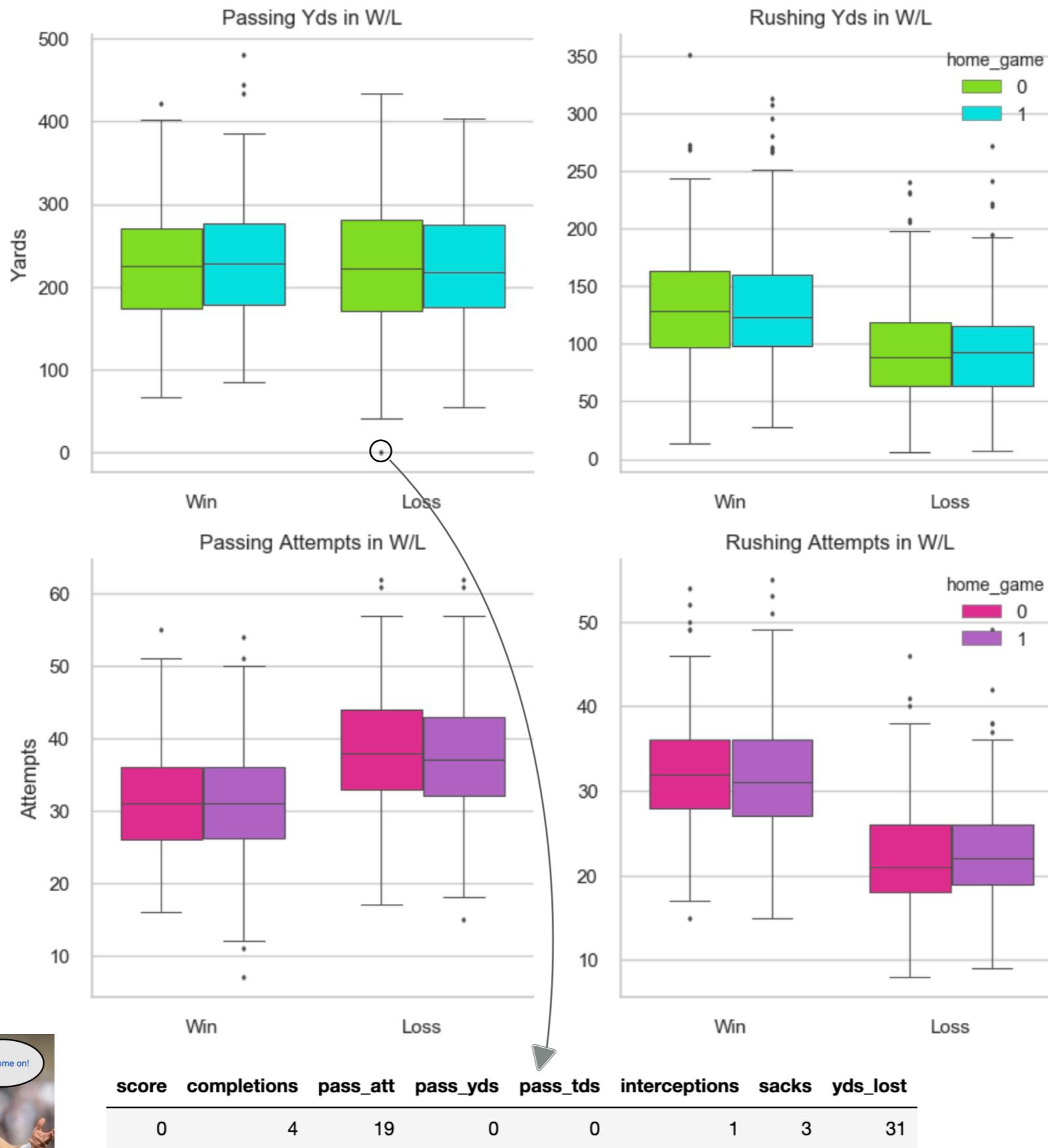
home game
passing yards
yards lost
rushing att
fgs made
punts
4th down conversions

score
passing touchdowns
yds/passing attempt
rush yards
fg attempts
punt yards
4th down attempts

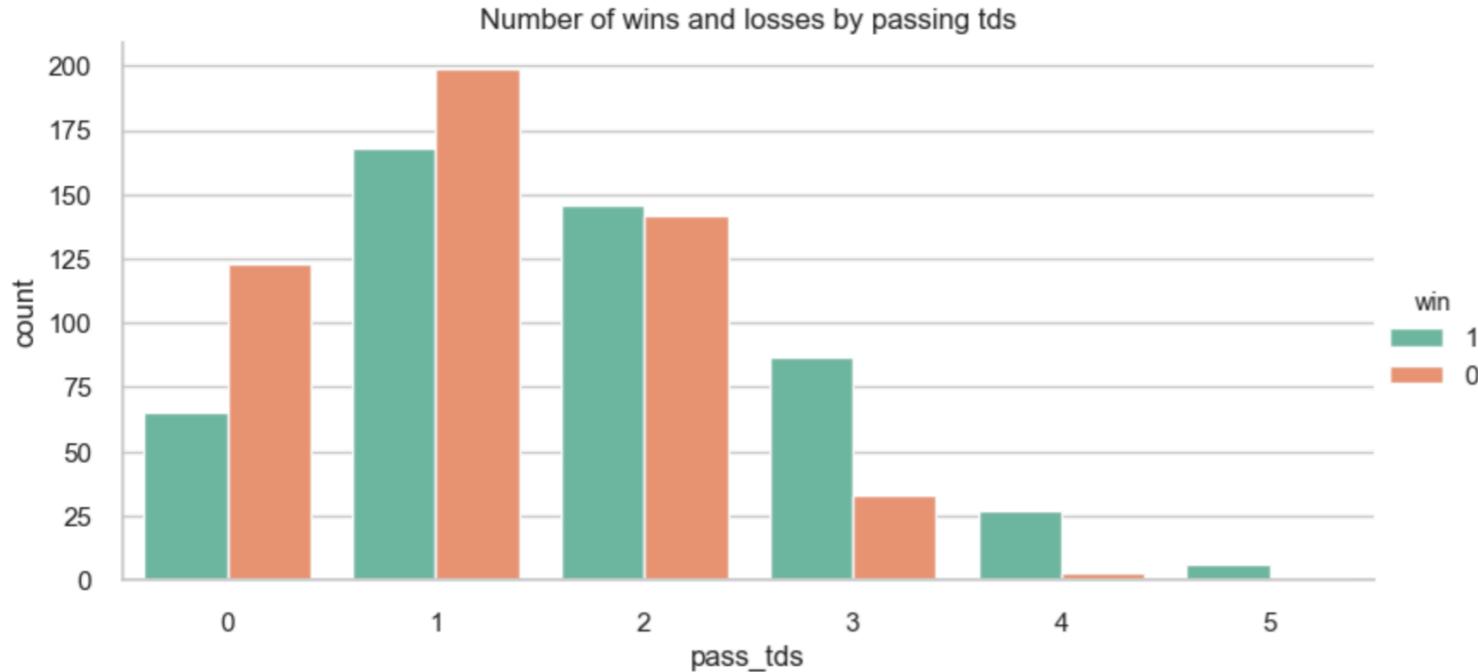
completions
interceptions
comp %
yds/rush attempt
xtra pts made
3rd down conversions
time of poss

pass attempts
sacks
QB rating
rushing tds
xtra pt attempts
3rd down attempts

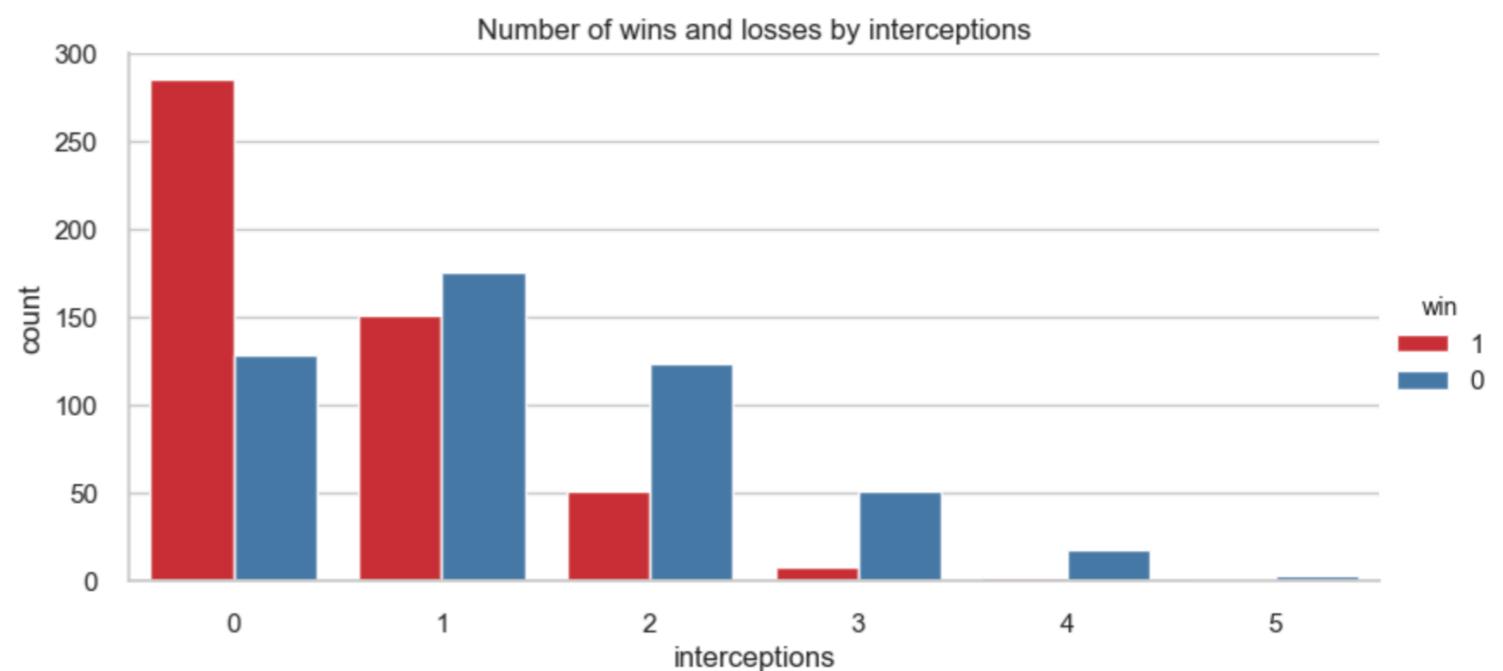
Rushing and Passing in Wins vs Losses



Quarterbacks: Hero or villain?

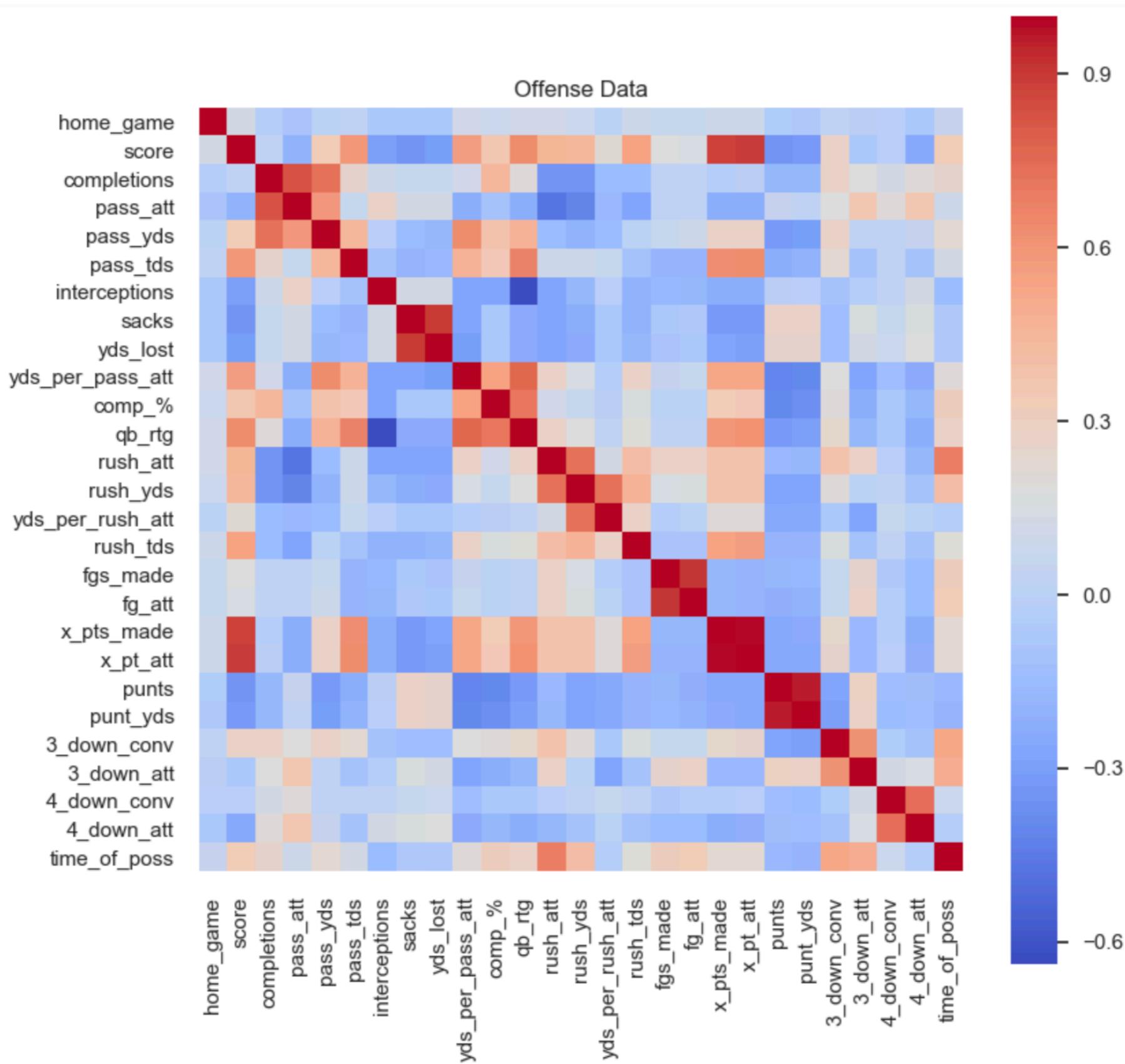


- Win-to-loss ratio goes up as the number of passing touchdowns go up



- Win-to-loss ratio goes down as the number of interceptions go up

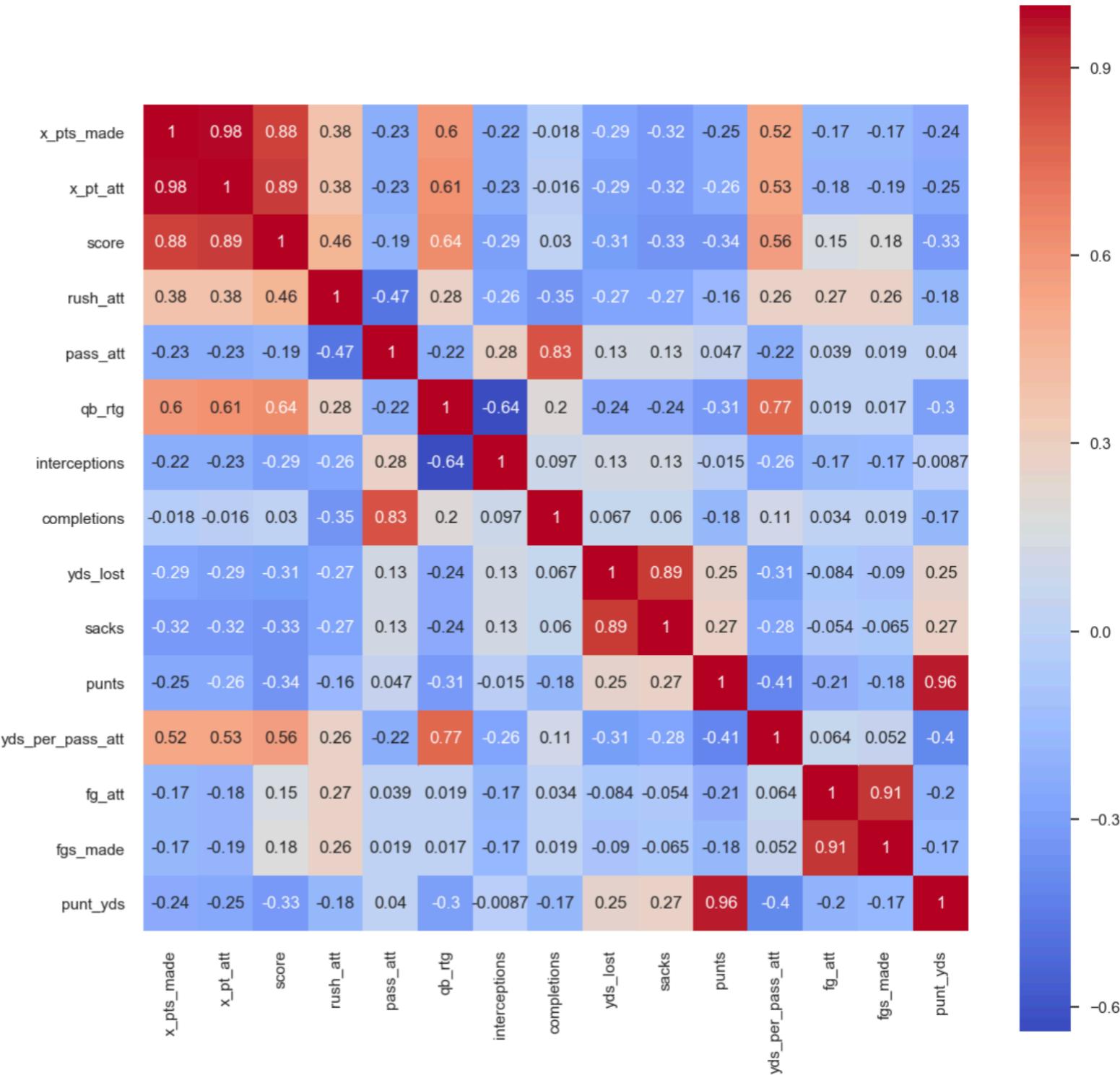
CORRELATION MATRIX



REDUNDANT FEATURES

Dropped Features (coef > 0.6)

- score
- x_pt_att (extra point attempt)
- pass_att (pass attempts)
- qb_rtg (quarterback rating)
- yds_lost (yards lost to sacks)
- punts
- fg_att (field goal attempts)



MODELS

	Model	Time	Training Accuracy	Test Accuracy	FPR	False positives	FNR	False negatives
0	naive-bayes	0.0027	0.7386	0.8000	0.2268	22	0.1748	18
1	logistic	0.7770	0.8600	0.8900	0.1134	11	0.1068	11
2	knn	0.0017	0.7063	0.6700	0.3402	33	0.3204	33
3	svm	8.3713	0.8524	0.8700	0.1546	15	0.1068	11
4	randomforest	0.1550	0.8425	0.8550	0.1649	16	0.1262	13
5	gradientboost	0.3463	0.8412	0.8400	0.1546	15	0.1650	17

MVF: MOST VALUABLE FEATURES

4_down_att	-0.604
fgs_made	0.544
x_pts_made	0.527
interceptions	-0.442
pass_tds	0.334
home_game	0.287
4_down_conv	0.249
sacks	-0.206
yds_per_pass_att	0.198
rush_att	0.155

POST SEASON VS. REGULAR SEASON

	Accuracy	False Positive Rate	False Positive Number	False Negative Rate	False Negative Number
Regular Season	89.00%	11.34%	11 / 103	11.34%	11 / 97
Post Season	84.55%	14.89%	7 / 47	15.87%	10 / 63

CONCLUSION

- Quarterbacks are still important
 - Maximize on passing touchdowns, yards per passing attempt and minimize on interceptions
 - Protect the quarterback to keep sacks to a minimum
- Kickers are also important
 - Make field goal and extra point attempts at a high rate
- Home field advantage is very much a thing
- Think very hard about going for it on 4th down
- My initial hypothesis is disproven
- Ideal purpose for model is for analysis after the fact in preparation for the next game.
- Since football is unique in that it essentially has two squads (offense and defense) that rarely share personnel, a separate defensive model is also needed.
- For playoffs, since the competition is more stiff hence acceptable margin of error is much slimmer. Additionally, tactics also change, so a bespoke model just for the playoffs should work much better.