



# 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

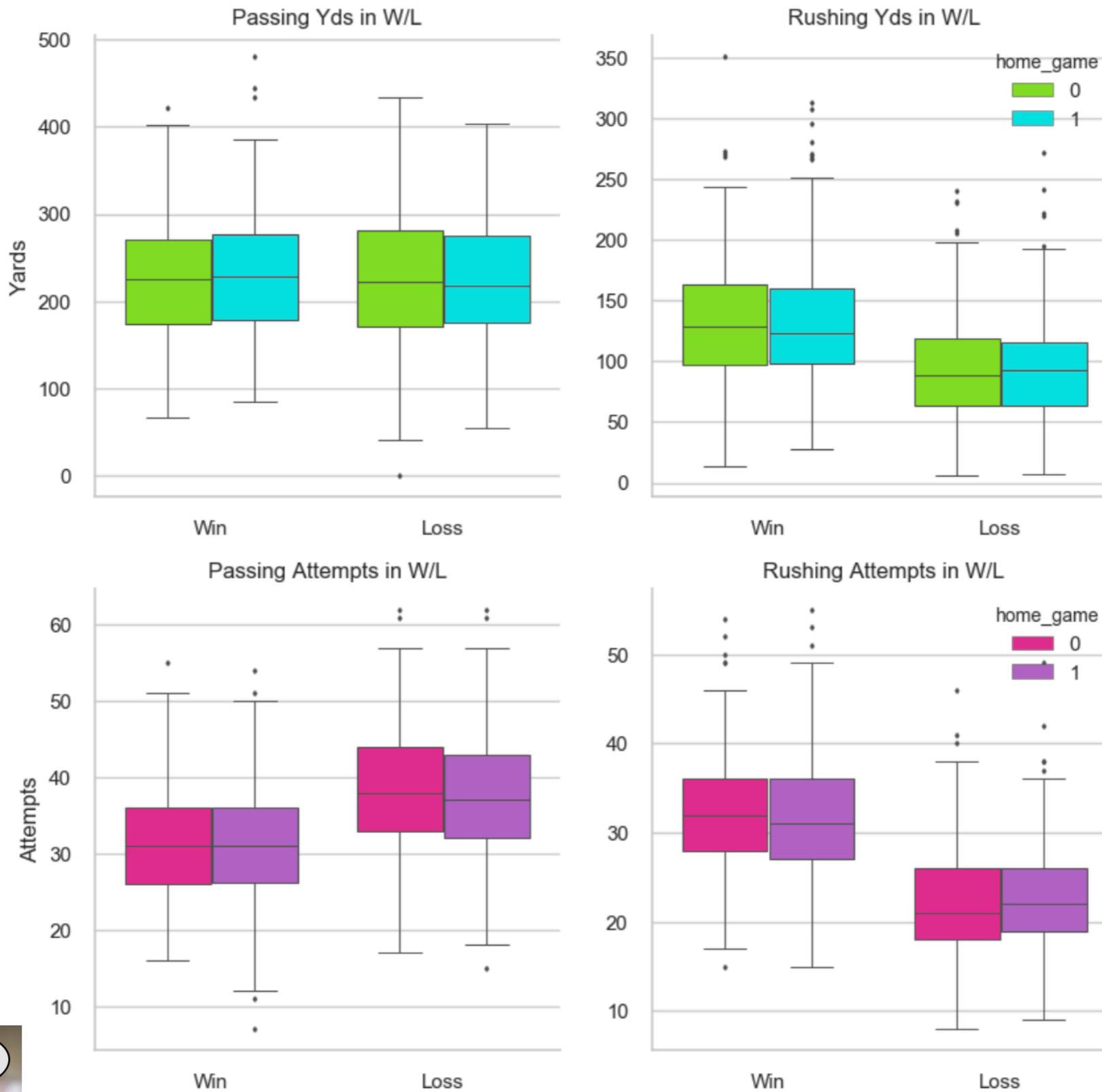
- Regular season data will be split into a training and test group at 8:2 ratio and be trained on various classification models
- Top 10 features will be gathered from the most accurate model
- The trained model will be tested against both the regular season test data and the playoff data

# DATA

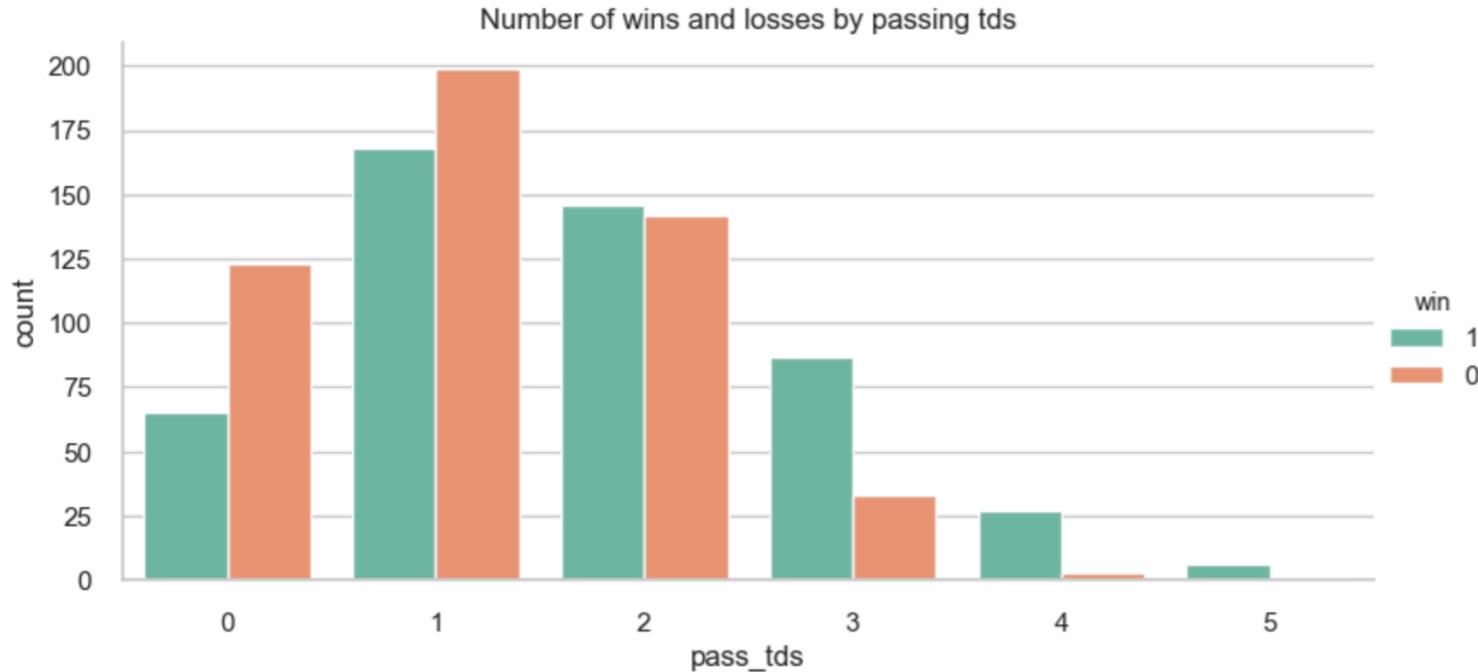


- Source: [www.pro-football-reference.com](http://www.pro-football-reference.com)
- Gathered offensive data from 1,000 random regular season games and 110 playoff games from 2009-2018 seasons

# Rushing and Passing in Wins vs Losses

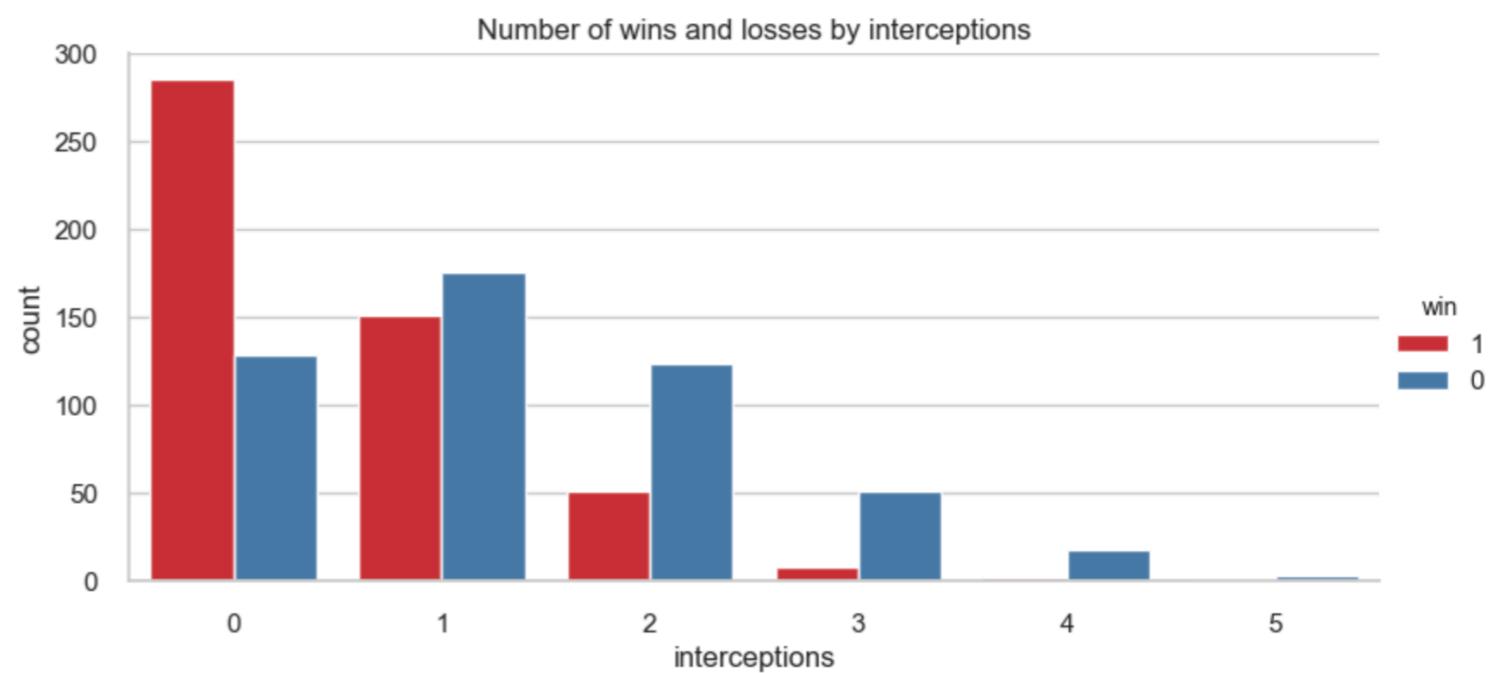


# Quarterbacks: Hero or villain?

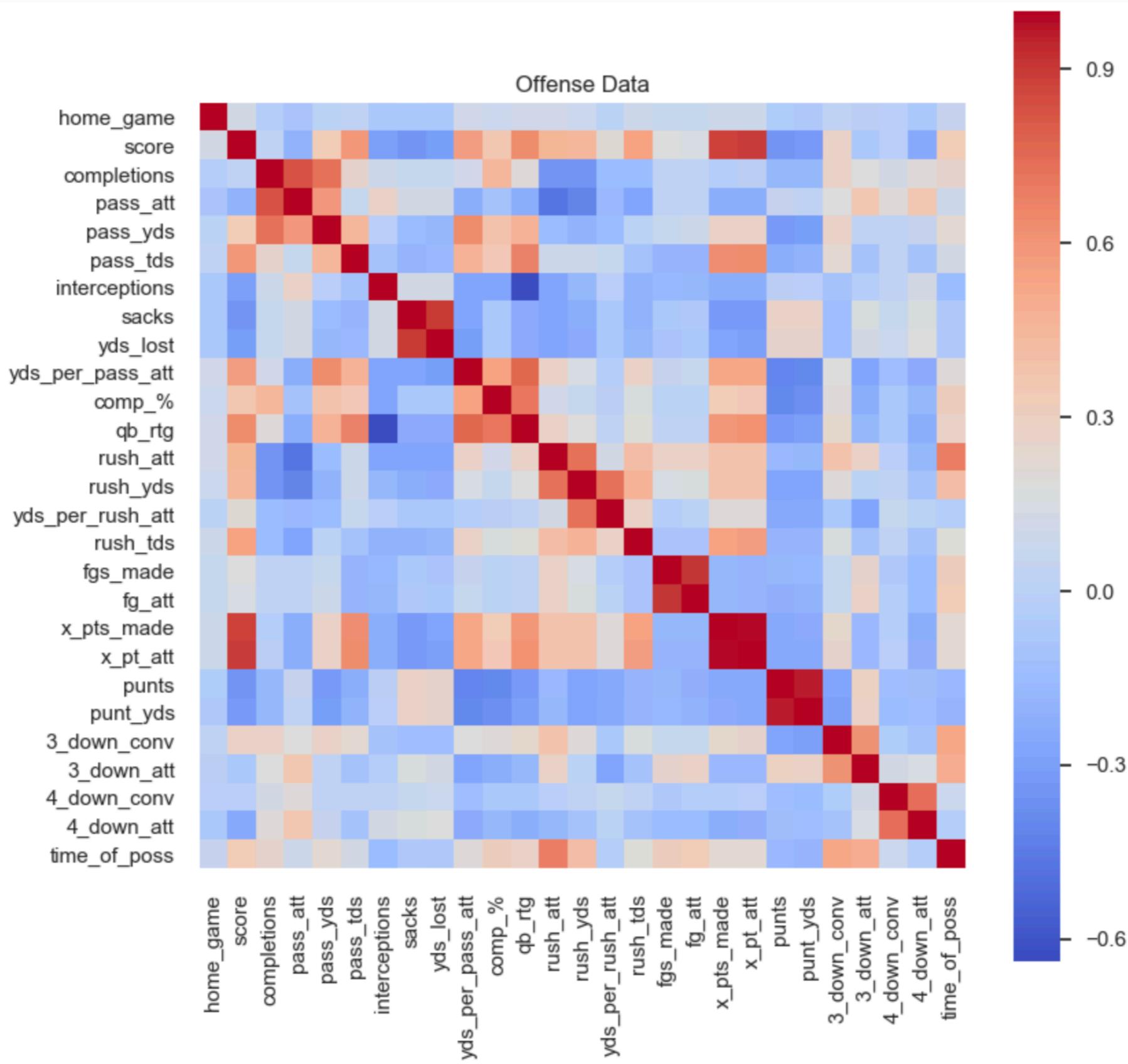


- Win-to-loss ratio goes up as the number of passing touchdowns go up

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



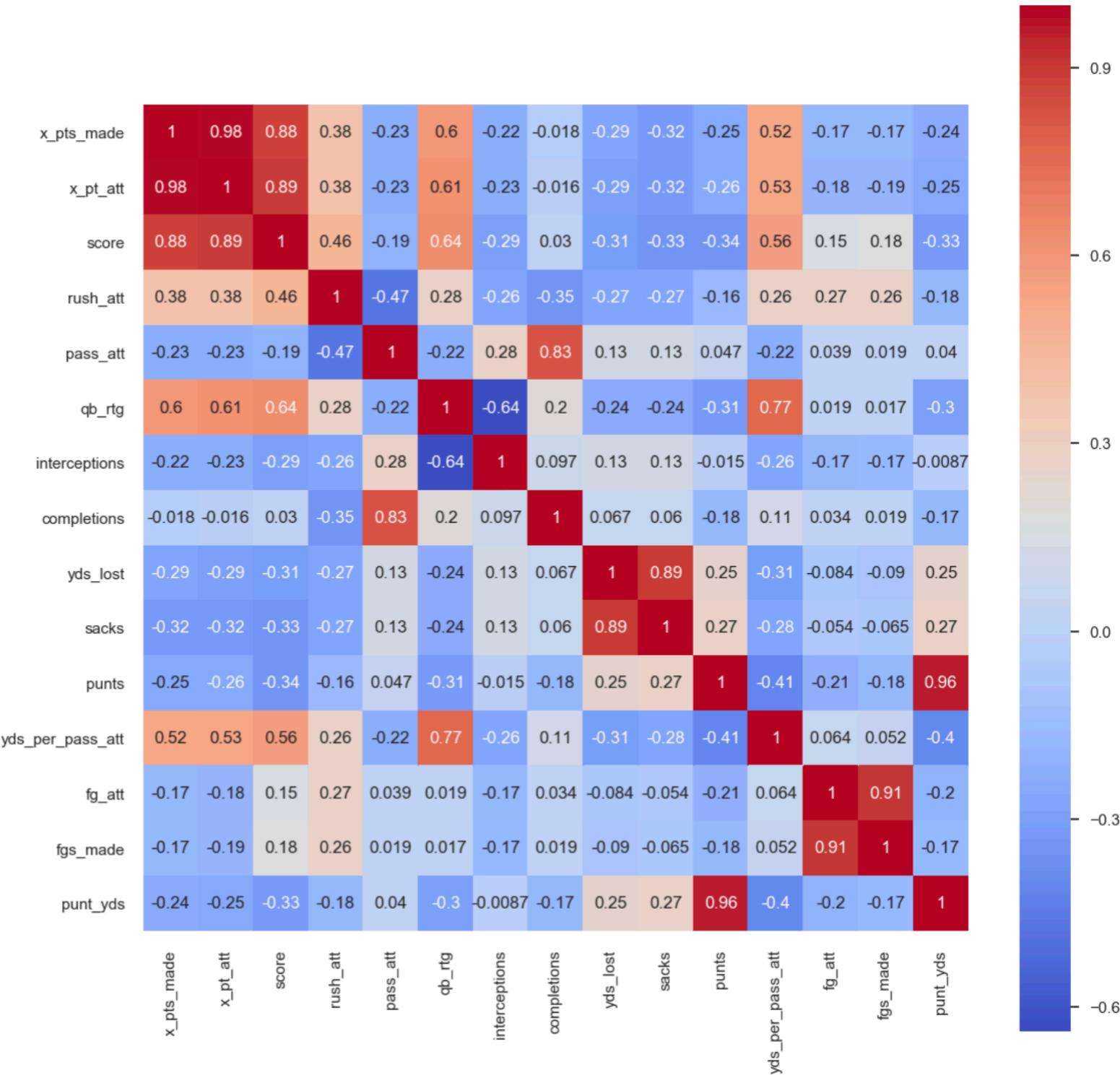
# CORRELATION MATRIX



# 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.