An Analysis of Winning in College Football

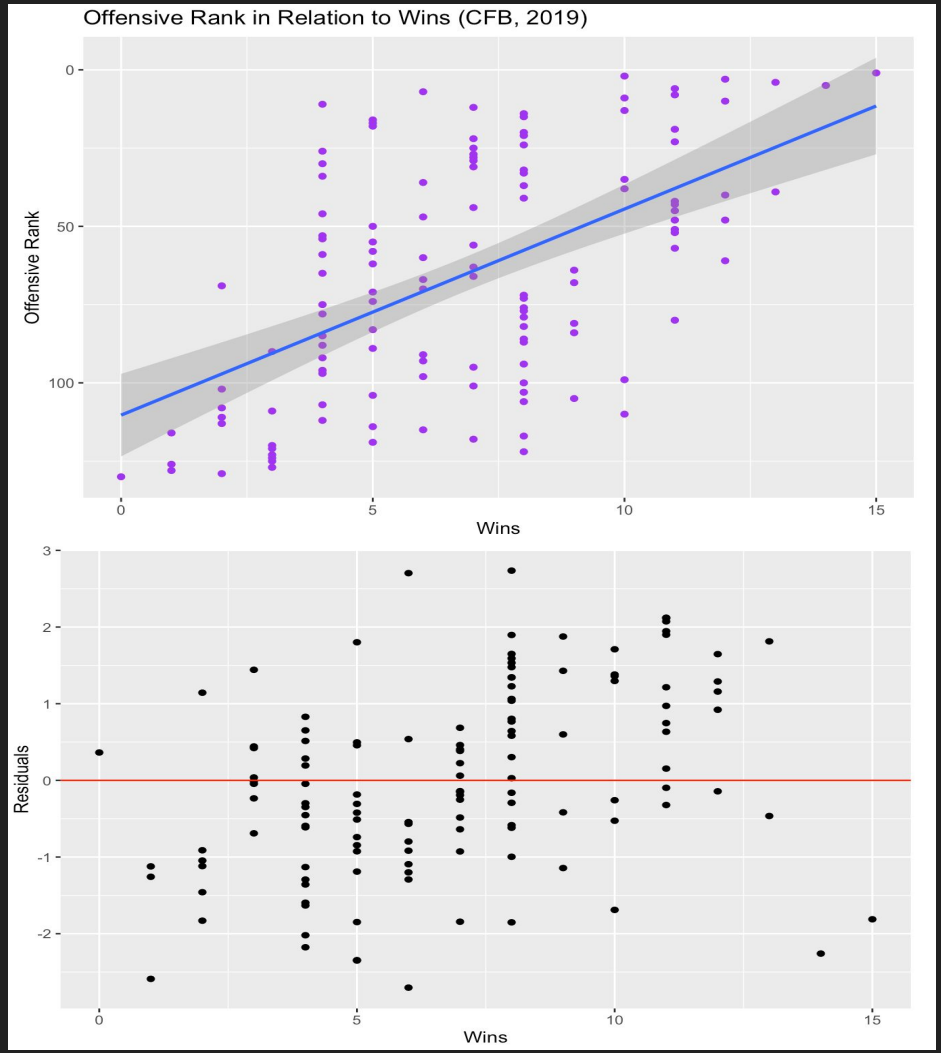
Joey Rheaume, Rodrigo Lara, Samuel Zhang, Steven Orbain

Goal

- Main question
  - What aspects of a team are most important in winning college football games?
- Goal
  - Observe the numerous range of statistics to determine which ones are most important to winning games and to qualitatively determine what aspect of football (offense, defense, special teams, etc.) mattered most
- Procedure
  - Observe statistics based on different aspects
    - Create plots and models
  - Qualitatively interpret illustrations and models

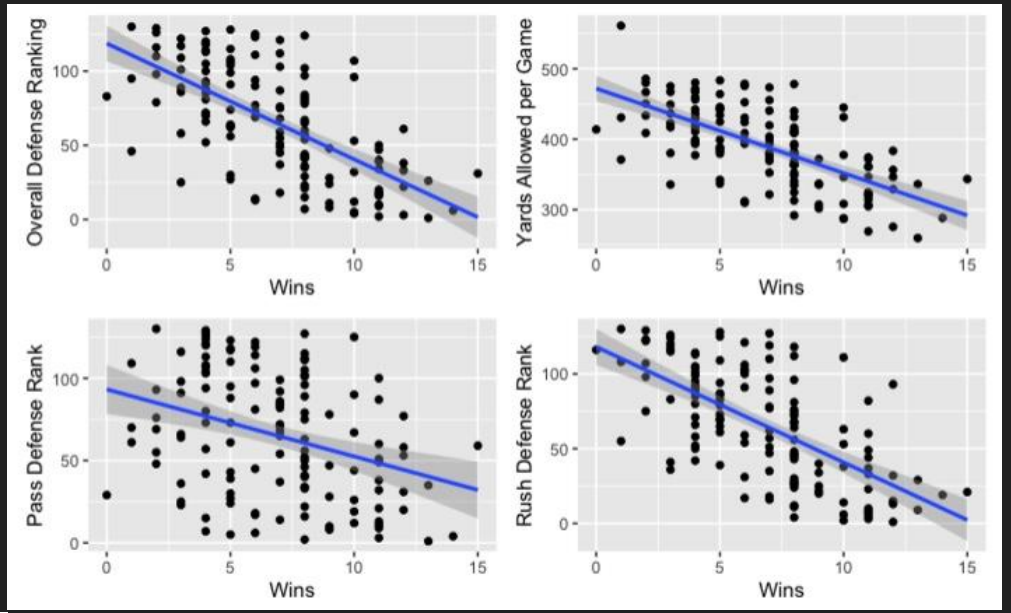
Results (Offense)

- Top plot
  - Purple dots represent each team, blue line represents the best-fit
  - Showcases a strong linear relationship between the offensive rank and wins
- Bottom plot
  - Residuals of regression model
  - Equal dispersion above and below line proves appropriate model for data
  - Based on offensive yards, TDs, rank, plays, first downs, scoring rank, & turnover margin
  - 84% adj-R^2



Results (Defense)

- Clear linear correlation between all plots
- All graphs are based on rankings so closer to zero the better and farthest right have most wins
- Not all teams play same level of competition so leads to not so high r-squared values but values around 42 are significant enough to represent a correlation in this situation

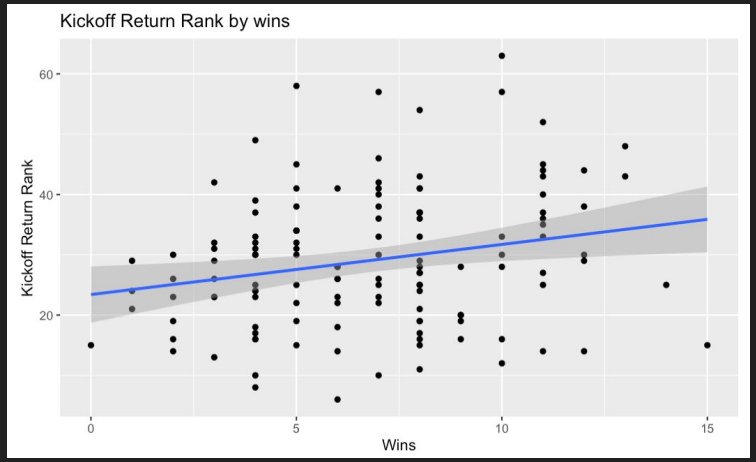


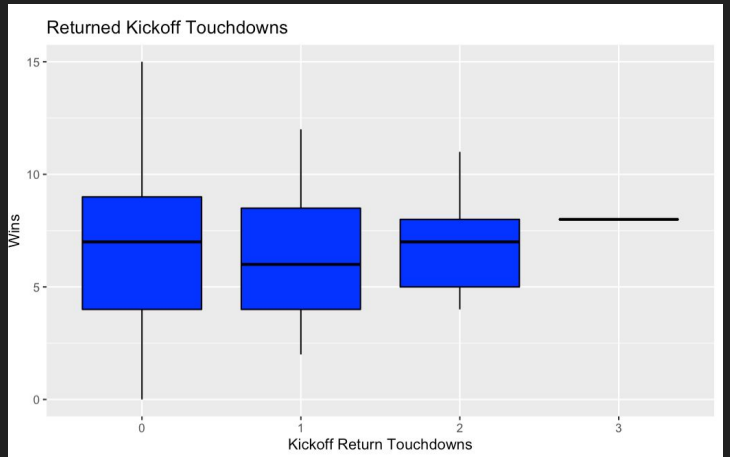
Stat <chr>	R.Squared <dbl>
Defensive Rank	0.4219717
Yards per Game Allowed	0.4289294
Pass Defense Rank	0.1146840
Rush Defense Rank	0.4120047

Plots represent Overall Defensive Ranking (Top Left), Yards Allowed per Game (Top Right) , Pass Defense Rank (Bottom Left) and Rush Defense Rank (Bottom Right) vs the number of wins for a team

Results (Special Teams)

- Plot relation between Kickoff Return Rank and Wins.
- Dots represent each teams.
- The greater the rank the greater the wins.



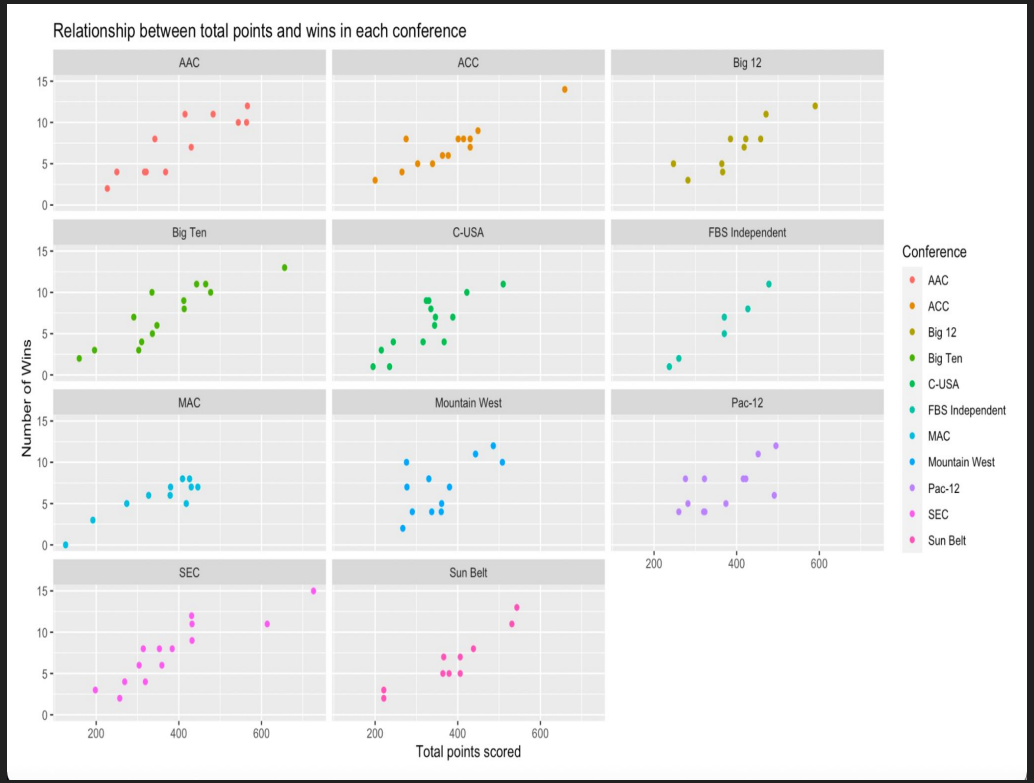


- Boxplot relation between Returned Kickoff Touchdowns and Wins.
- Surprisingly the team with most wins of the season didn’t make a single RKT
- Doesn’t have much of an impact on victory probability.

Relationship between points scored and wins in each conference

- Analysis through conference perspective.
- The plots show a positive relationship between total points scored and wins, except conference Mountain West.
- Especially San Diego State university in Mountain West, they won 10 games out of 13 games by only scored 276 points which is relatively low compared to other teams that also won 10 games.

Team <chr>	Win <int>	Total.Points <int>	Def.Rank <int>
Hawaii	10	508	96
Iowa	10	335	12
Louisiana Tech	10	422	53
San Diego St.	10	276	5
SMU	10	544	107
UCF	10	564	32
Wisconsin	10	477	4



Conclusion

Which aspect mattered most?

- Overall offense and defense performance seems to be the most important aspects to win games, meanwhile, special teams doesn’t seem to have that big of an impact on winning probability.

Which linear models showcased the best correlation with winning?

- Offense and defense linear models seem to showcase the best correlation with winning, there must be an equally good balance between them to achieve a victory in games.

Are there any possible improvements to be made?

- Plotting techniques are limited and data is limited to one year

References

https://www.kaggle.com/jeffgallini/college-football-team-stats-2019

STT 180 Slides