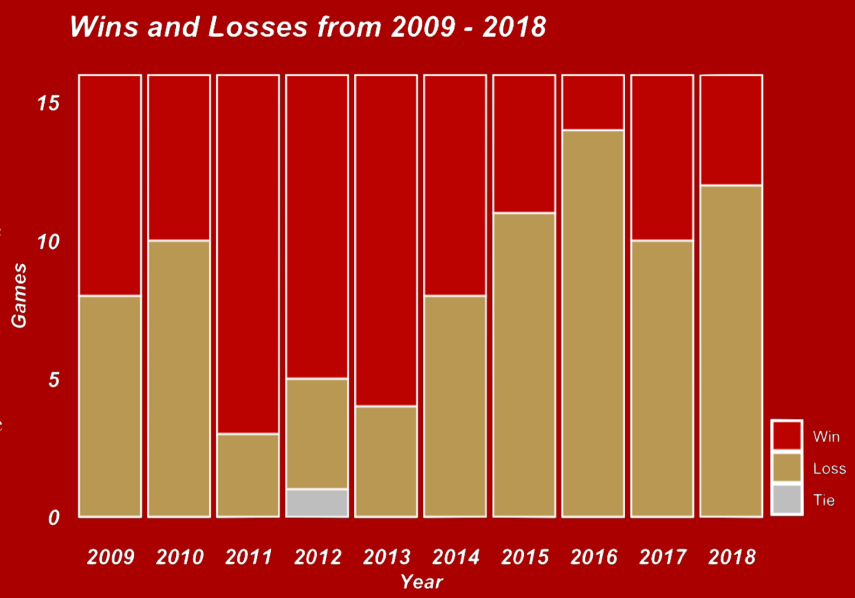
Data Mining for Gold

Winning 49ers football games with data-driven decisions

Tim Hulak IST-719: Information Visualization

About the Team and Audience

The San Francisco 49ers are a professional American NFL Football team who play in the NFC West Division and NFC Conference. Strategy in American football is more important than perhaps the casual viewer may realize. Football is sometimes referred to as a "game of inches", which was taken from the quote “Football is a game of inches and inches make the champion”, credited to Hall of Fame coach Vince Lombardi. Therefore, coaches and scouts who are looking for a data-driven edge would be the ideal audience for this poster.



About the Data

The dataset contains 449,371 observations across 255 variables. Each variable represents information about a given play during an NFL football game. Each record represents a specific play during a given game. The data contain information on plays from the 2009- 2018 season. In order to single out the 49ers, a subset of the data was taken and split into two data frames: 49ers on offense and 49ers on defense.

How Have Penalties Impacted the Team?

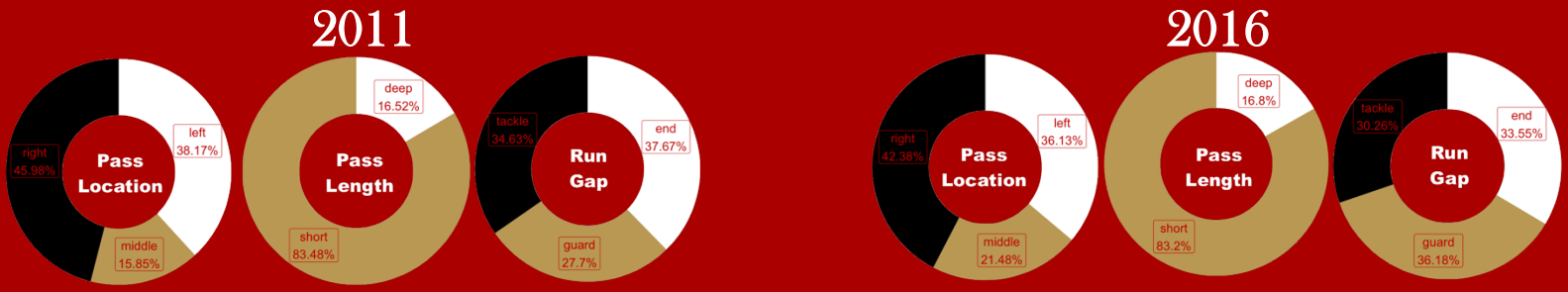
In total, just from the most frequently commented penalties, the 49ers lost 4360 yards (50.44% of total yards lost), which is over 40 times the length of the football field.

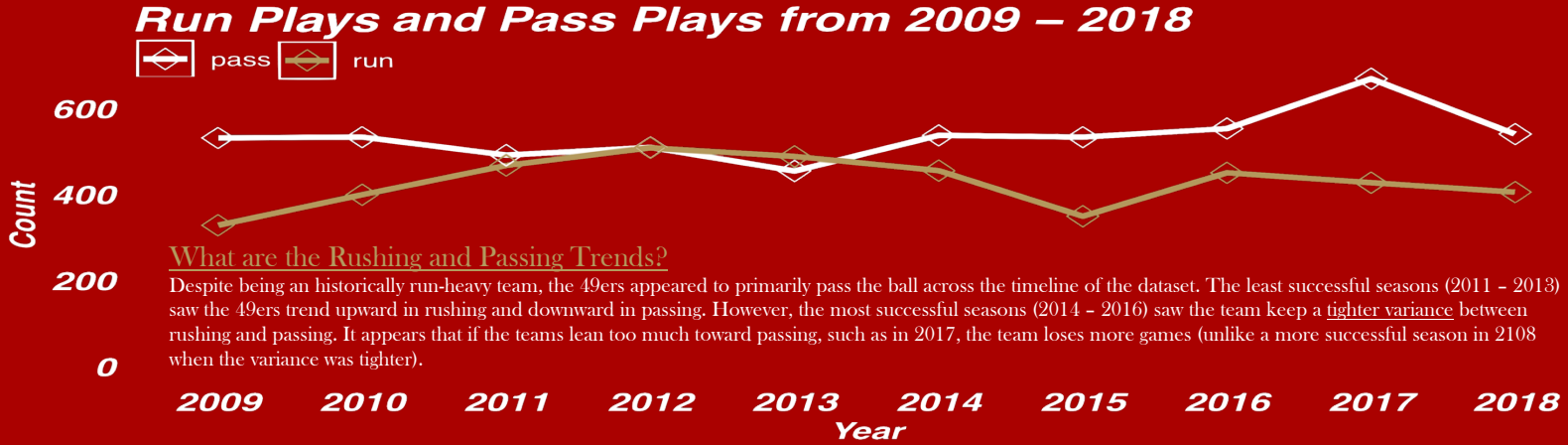
- 198 False Starts (976 total yards at 5 yards per penalty)
- 149 Offensive Holding (1425 total yards at an average of 9.6 yards per penalty).

- 66 Defensive Offsides (330 total yards at 5 yards per penalty)
- 65 Defensive Pass Interference (1052 total yards at an average of 16.1 yards per penalty)
- 56 Unnecessary Roughness (722 total yards at an average of 13.8 yards per penalty)
- 53 Defensive Holding (254 yards at an average of 4.8 yards per penalty).

Do Pass Location, Pass Length, or Run Gap Effect Team Success?

When comparing the least successful season against the most successful season, it appears that these metrics can play a slight role in team success. The team appears to throw short passes over 83% of the time in both seasons. Therefore, Pass Length may have little to no effect. However, the team passed to the middle of the field 22% of the time and had more success in 2016 as opposed to 16% in 2011. In addition, the 49ers ran 36% of the time to the guard side of the offensive line in their more successful season rather than 27% of the time in their least successful season.





What are the Rushing and Passing Trends?

Despite being an historically run-heavy team, the 49ers appeared to primarily pass the ball across the timeline of the dataset. The least successful seasons (2011 – 2013) saw the 49ers trend upward in rushing and downward in passing. However, the most successful seasons (2014 – 2016) saw the team keep a tighter variance between rushing and passing. It appears that if the teams lean too much toward passing, such as in 2017, the team loses more games (unlike a more successful season in 2108 when the variance was tighter).