
NFL
NFL
A Deeper Dive Into
NFL Data

The Data We Explored

- Utilizing the NFL Data from Project 1: NFL Teams data, NFL stadiums/weather data and NFL betting data showing over/under line and game spread.
- NFL Injury data: python library containing injury data including season, player, position and status before the game from 2009 to 2022
- NFL Penalties data: detailing penalties for each game showing the quarter, player and time during the game from 2009 to 2022.
- NFL Draft data: python library including draft picks and rounds for each NFL team from 2009-2022.

Unlocking Insights: A Journey into the NFL Data

- Using the injury data set, we want to determine how a team's overall record could be impacted by injuries to key players on their team. Which injuries had the highest impact on a team's outcome?

- With access to the NFL penalty data set, we want to see how penalties impact the overall outcome of a game. How many games were decided by a late penalty?

- Using the draft picks data set and seasonal nfl data set showing player career stats, we have an exploratory interactive chart to examine the differences between players efficiency based on position and their corresponding draft rounds.

NFL Injury Predictions

Analytic approach for predicting injuries

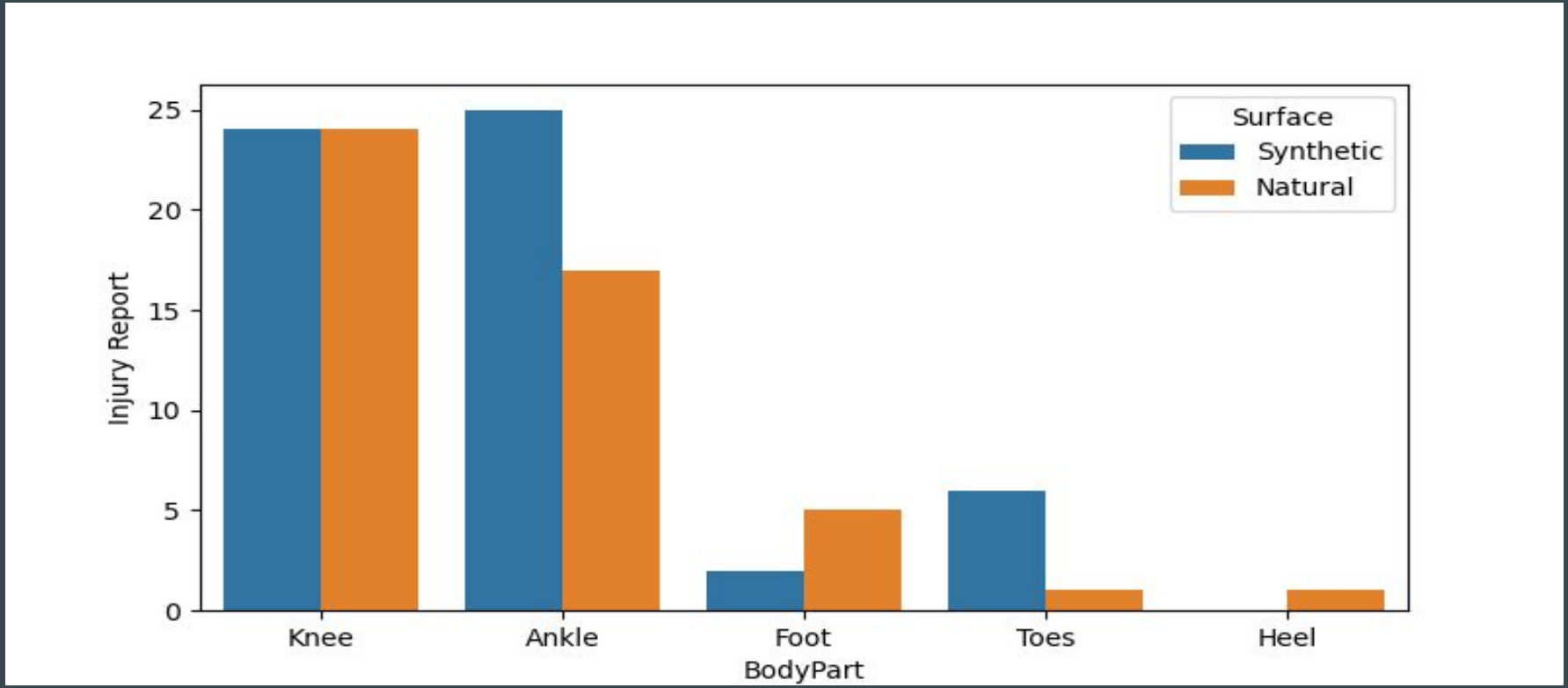
RIP Chubb

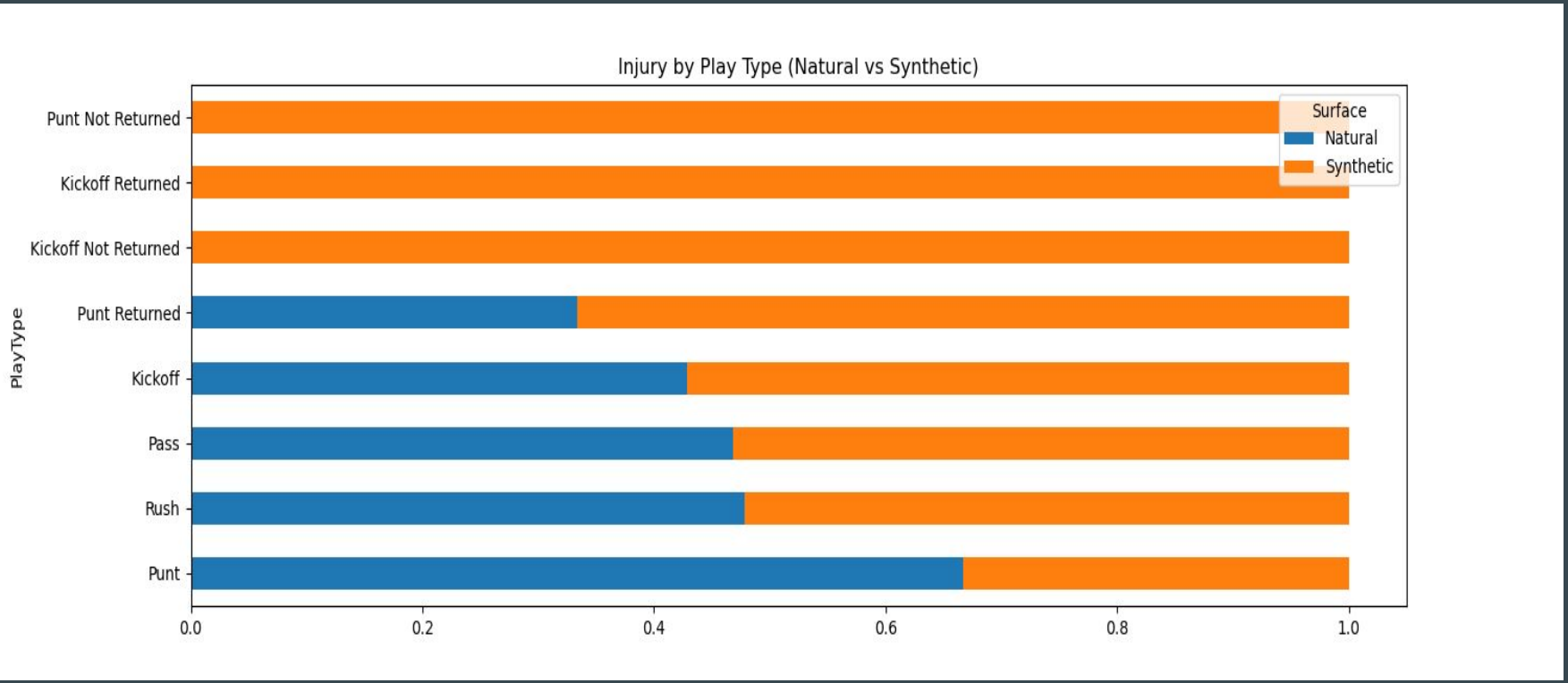
Injury Analysis Overview

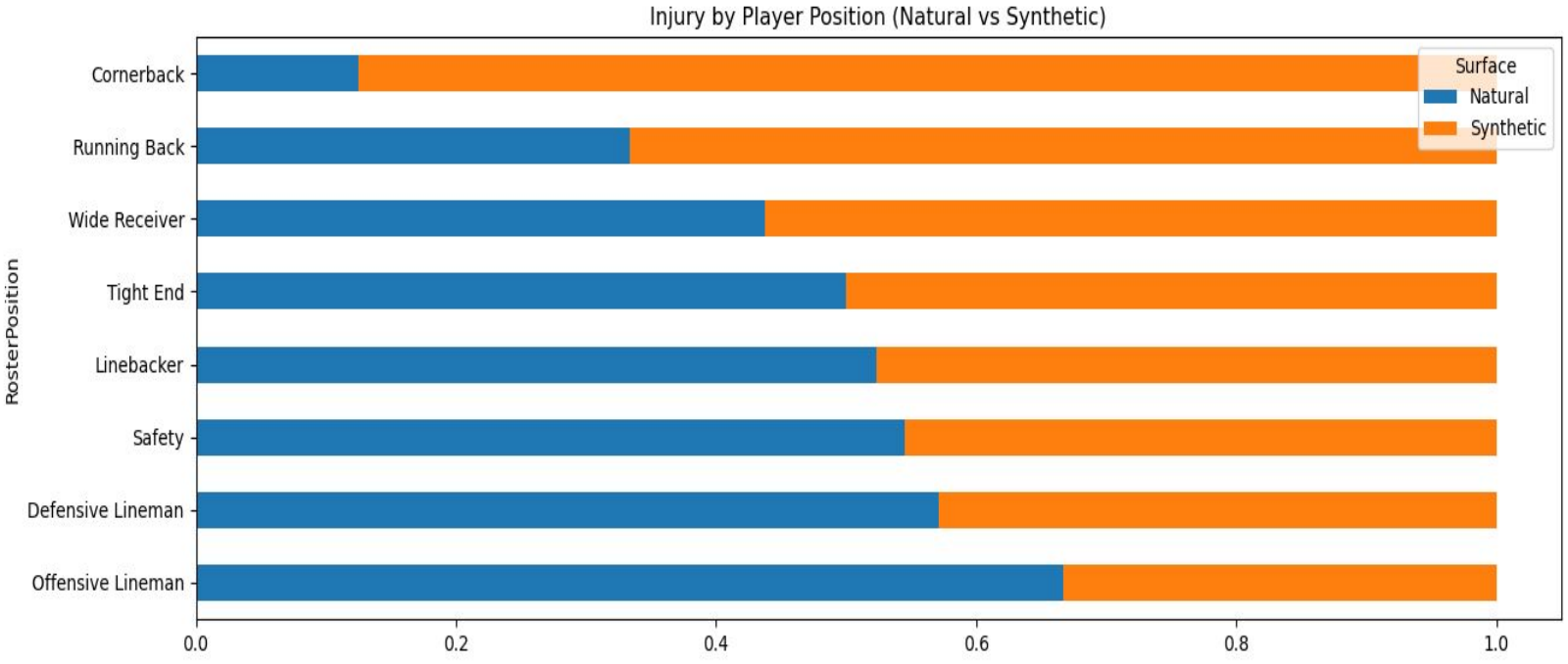
Data Factors in Considerations:

- Effects that playing on synthetic vs natural turf
- Factors such as stadiums, weather, play type and player position
- 250 complete player in-game histories from two subsequent NFL regular seasons.

Does Turf has impact on injury based on play type and player positions?




Injury by Play Type (Natural vs Synthetic)
PlayType
Punt Not Returned
Kickoff Returned
Kickoff Not Returned
Punt Returned
Kickoff
Pass
Rush
Punt
0.0
0.2
0.4
0.6
0.8
1.0
Surface
Natural
Synthetic


Injury by Player Position (Natural vs Synthetic)
RosterPosition
Cornerback
Running Back
Wide Receiver
Tight End
Linebacker
Safety
Defensive Lineman
Offensive Lineman
0.0
0.2
0.4
0.6
0.8
1.0
Surface
Natural
Synthetic

Questions we've ask ourselves:

- ***--What is likely the risk of a player's injury based on turf type?***
- ***--Is the Predictions of Injuries Accurate?***
- ***--Does the importance of grass type impact injuries?***

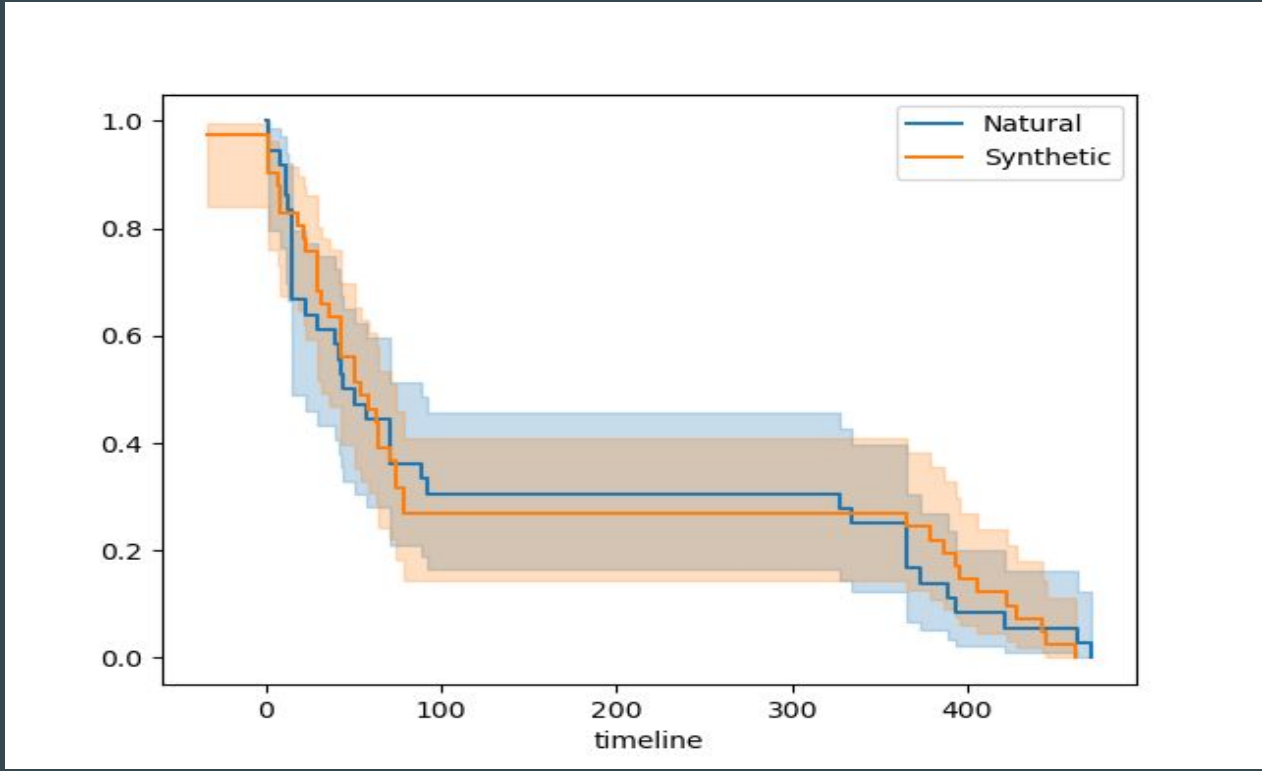
The Solution..... Data Hazard Model

- ***Kaplan-Meier Estimate (Curve)***
- ***Confusion Matrix (Logistic Regression)***
- ***Random Forest Classifier***

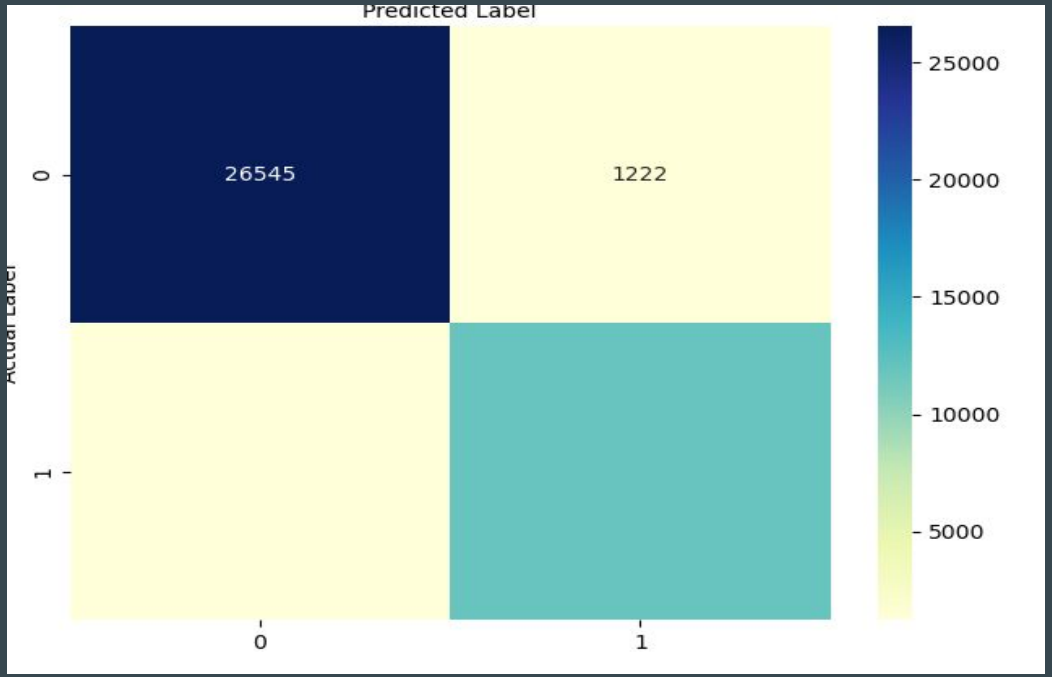
Kaplan Meier Curve

Question: What is likely the risk of a player's injury based on turf type?

Answer: *both curves indicated with both significant risk of injuries*



Logistic Regression...also known as confusion matrix



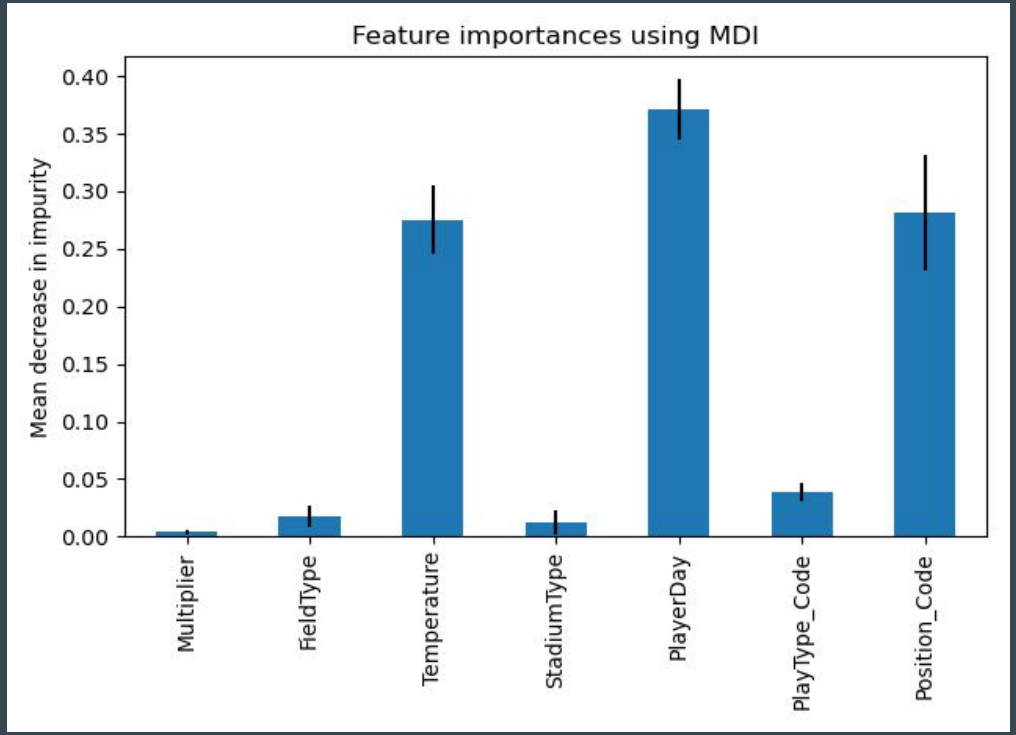
Question: is the predictions of the injuries accurate….?

Answer: *No…. the model suggested that the prediction model is inaccurate*

Random Forest Classifier

Question: Does the importance of the grass type affect the injury?

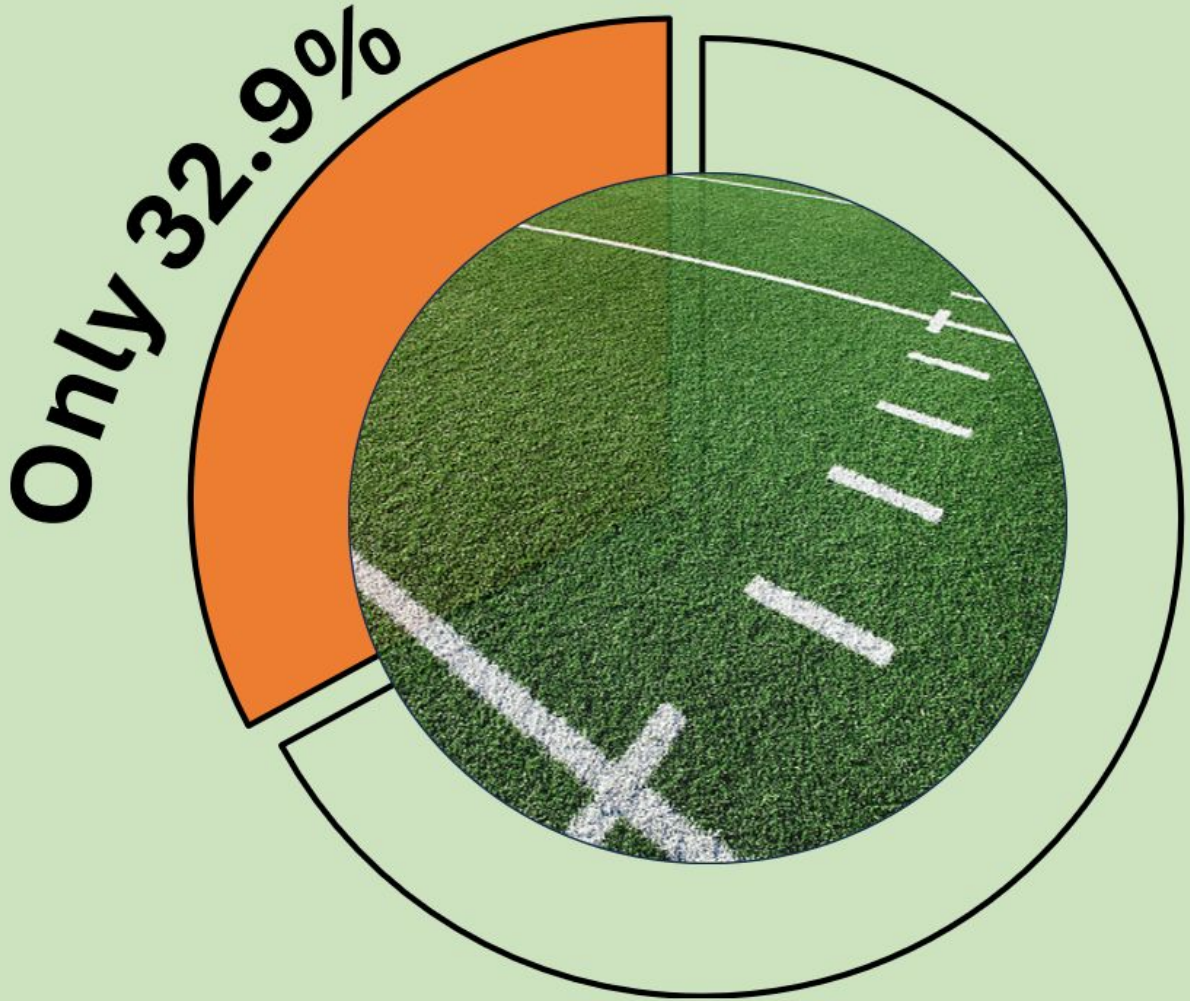
Answer: *No….shows that grass types does not impact players injuries*



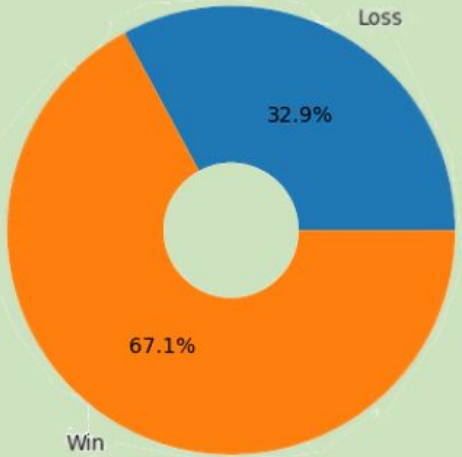
NFL PENALTY DATA

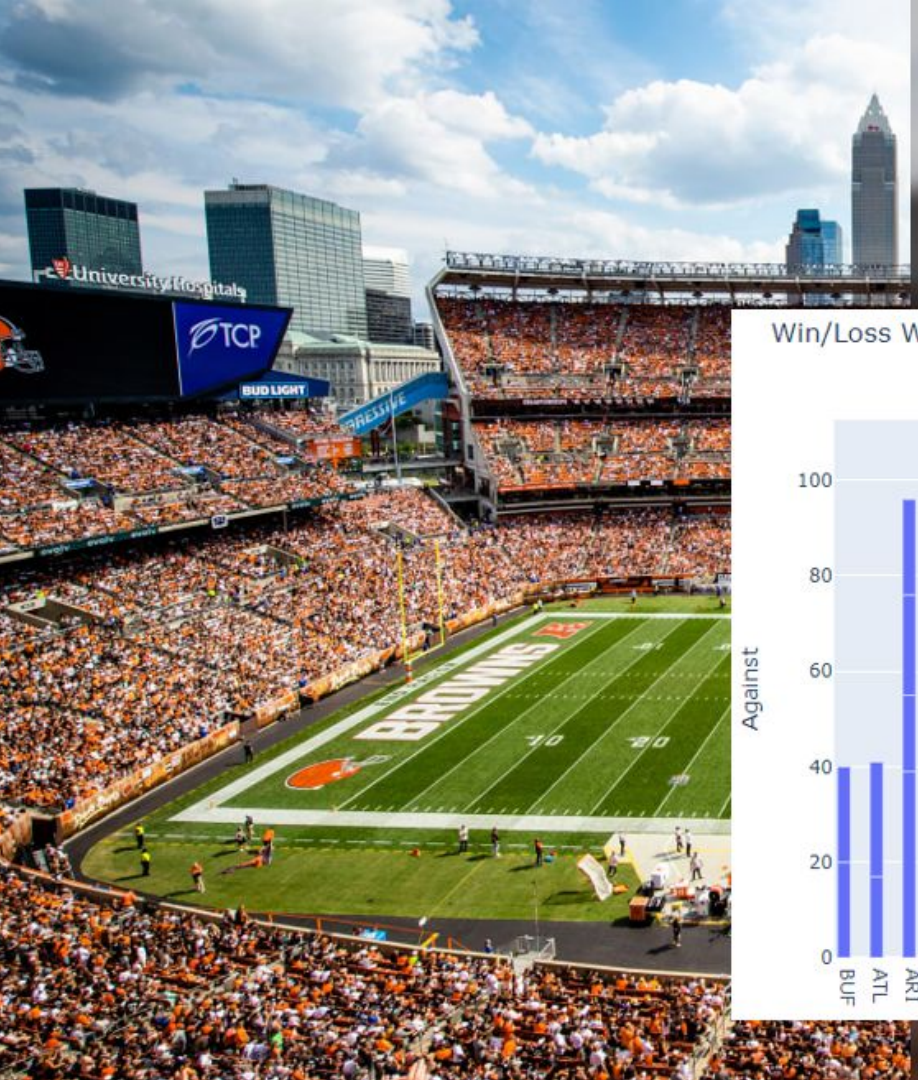
Looking at penalties that were not offset, or declined in the remaining 2 minutes of the game, and over 15 total penalty yards, only by offense or defense (no special teams)

Did these high penalty yards affect the outcome



Only 32.9% of games were lost by teams who had 15 or more total penalty yards in the last two minutes of the game





The Denver Broncos had the highest overall total penalty yards for games lost

The New York Jets had the highest overall total penalty yards for games won

The New York Jets had the highest penalty yards in 1 game @ 50 yards (13 yd DPI, 24 yd DPI, 13 yd DPI) week 5 of the 2019 season against the Dallas Cowboys.

Despite these penalties, the Jets won the game 24-22.

The Jets were up 24-16 prior to the 2-minute warning.

NFL Draft Data Interactive Chart

Features:

- Sort by Minimum Number of games played
- Sort by Season year or range of seasons
- Sort by number of pro bowls received
- Sort by Position
- Filter Y-axis by passing, rushing, receiving efficiencies or PPR fantasy points

Details:

- 4173 total data points
- 960 total Unique Players
- Looking only at Qb's, RB's, WR's and TE's
- Seasons 2009 to 2023
- Regular Season Stats

Results from Exploratory Analysis

- Early round draft picks tend to have more games played than later round picks
- Early round draft picks tend to be more inconsistent when it comes to efficiency numbers but generally show a higher average efficiency than later round picks
- QB's taken in the first round show a significantly higher range of efficiency and 5th round QB's appear to be very inefficient when it comes to contributing to their team
- Looking at RB's and WR's.. the round they were selected in did not appear to significantly impact their efficiency suggesting it would be smart to draft a RB or WR any round and they could show potential.
- The best time to draft a TE appears to be the second or third round.

Questions?

...