**NFL Trained Data Analysis**

|  |  |  |  |
| --- | --- | --- | --- |
| Archana Kumari B.A | Adnan Elahi | Srinivas Edulakanti | Mrudul Gaurang Choksi |
| 424-393-6105 | 310-873-8636 | 424-386-1785 | 773-543-0213 |
| Department of Information systems | Department of Information systems | Department of Information systems | Department of Information systems |
| California state University | California state University | California state University | California state University |
| Los Angeles | Los Angeles | Los Angeles | Los Angeles |
| abajine@calstatela.edu | aelahi@calstatela.edu | sedulak@calstatela.edu | mchoksi@calstatela.edu |

**Abstract**

Data and analytics have been part of the sports industry from many years. However, it is only recently advanced techniques like Hadoop, hive and many tools are utilized for facilitating the operations of sports franchises. While part of the reason is related with ability to collect more filtered data an equal important factor for the team and the associations to analyze about the playing turf. Different association and people associated with the teams, sports clubs and other stakeholders in the industry invest in analytics. We are excited to analyze an NFL game as the datasets and variables are available to examine the effects of playing the game on synthetic turf versus natural turf to understand how the player moves on the field and the factors that results in low extremity injuries. These files have data about injuries, player plays and player movement during the game.

**1. Introduction**

NFL is a professional American football league is one of the four major professional sports in North America and is 17-week regular season. If we have to analyze football, we need to understand the sport first. Two teams compete with each other who take turns to pass the ball on the field in to endzone and other team would stop from passing the ball to endzone.

NFL team consists of 11 members each comprising of offensive and defensive on the field and every player from either team would try to gain the ball and run to reach the endzone .The battle between the team would involve lot of planning and blocking the opposition team from passing the ball to the endzone. NFL Coaches make over $10 million a year.

A black and red text

Description automatically generated

Figure 1. Player position

**2. Players Positions**

In Gridron football there are 11 players from each side. Purpose of game is to move the ball to the opposition’s end zone that is either the person holding the ball should run or pass the ball to the other teammates until the ball is moved to end zone. Player positions on the field are Defensive Back, Line Backers, Wide receiver.

**2.1 Defensive Backs**

Defensive Backs take positions somewhat back from the line of scrimmage, they are distinguished from the defensive line players and line backers and who take positions directly behind or close to the line of scrimmage.

**2.2 Linebackers**

Linebackers are members of defensive team and they are approximately 3 to 5 yards behind the line of scrimmage behind the defensive lineman, therefore back up the line. Linebackers generally align themselves be snapped by standing upright in a two-point stance.

**2.3 Wide receiver**

This is an offensive position in gridiron football, and it is a key player. They get their name because they are split out wide farthest away from the rest of the team. Wide receivers are fastest players on the field.

**3. Data Analysis**

Let us concentrate on Datasets and what variables provide to examine the effects of playing on Synthetic turfs versus natural turfs can have on player movements and factors that contribute to lower extremity injuries. There are columns like field type, stadium type, play type in which we have information related to Natural turfs and synthetic turfs. Stadium type has information related to indoor, outdoor and dome and playtype has information like what kind of actions a player makes during the course of the game like pass, punt etc . All these actions would also result in injuries. We have done analysis with the help of Hive QL and Tableau.

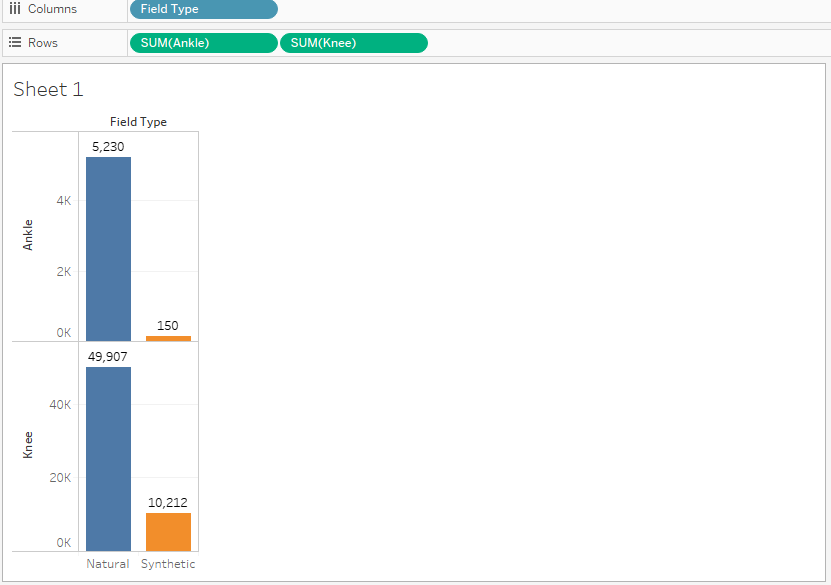


Figure 2. Field vs. injuries

In Figure 2. we analyzed the NFL data to find out two important attributes. One being the field type and the other being the injuries caused in defensive back position type. We can see the number of injuries on natural field are more when compare to Synthetic turf. Injuries to knee are more than injuries to Ankle on Natural turf .

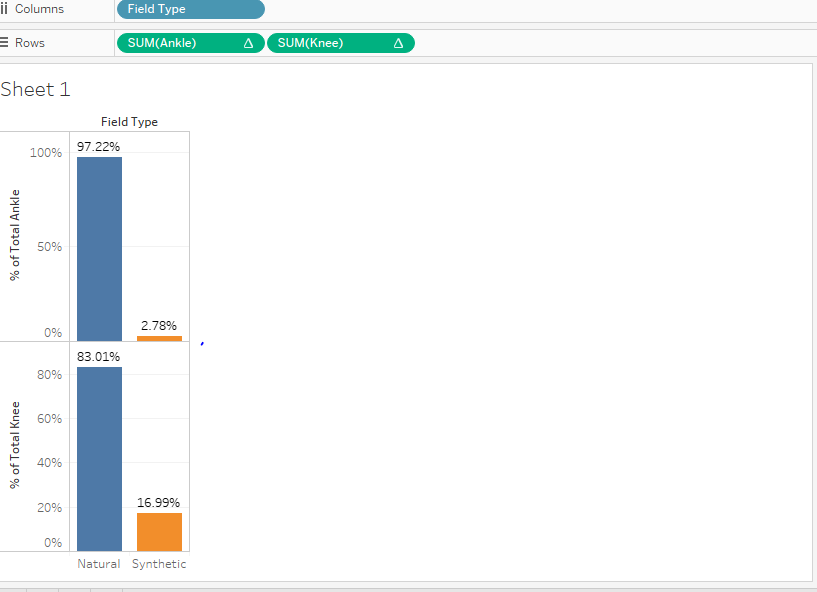


Figure 3. Field vs. injuries (%)

In figure 3. Percentage of injury to Ankle is 97% on Natural turf which is quite high when compared to percent of injury to Ankle on Synthetic turf. Percentage of injury to knee is 83% more on natural turf compared to synthetic turf. If you observe the percentage of knee injuries on synthetic turf are higher than the percentage of Ankle injuries on same turf.

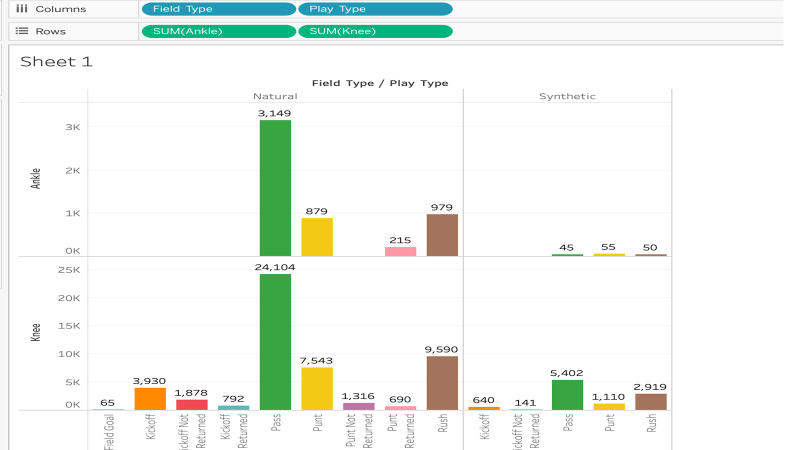


Figure 4. Playtype vs. injuries

We have taken field type, Play type and the injury to knee and ankle to analyze the number of injuries in particular playtype on natural and synthetic turf . From the above figure 4. we can clearly see that injuries to players have caused by play type “pass” are higher on natural turf. While on synthetic turf the number of injuries to knee are more while pass playtype. Injuries to body part ,Ankle, are higher on synthetic turf while play type is Punt.



Figure 5. Weather vs. injuries

In figure 5. visualization we have fieldtype, weather and the number of injury fields. Reason for using these attributes is to analyze the maximum temperature and minimum temperatures and the type of weather results in a greater number of injuries on both Turfs. As you can see on natural turf temperature can extremely vary on hot and cold weather whereas on synthetic turf there is not much temperature difference. Synthetic turf keeps a constant temperature in hot and cold weather. There are also less injuries to the knee and ankle on hot and cold weather.

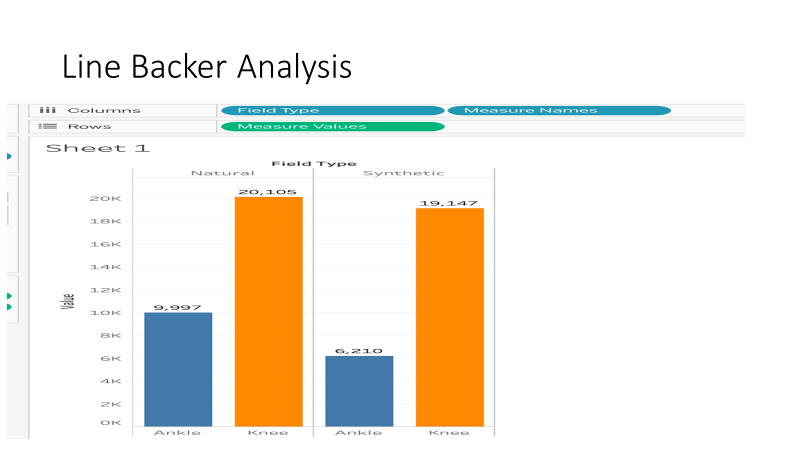


Figure 6. Linebacker Analysis

We have analysis for the linebackers on both field types. On the overall we could see the number of injuries on the natural field type are higher as usual. Two important things we analyzed from the above visualization is there is not much gap between injury in both field types. The other being when you compare the figures between the defensive and linebacker you will see injuries for line back position are greater.

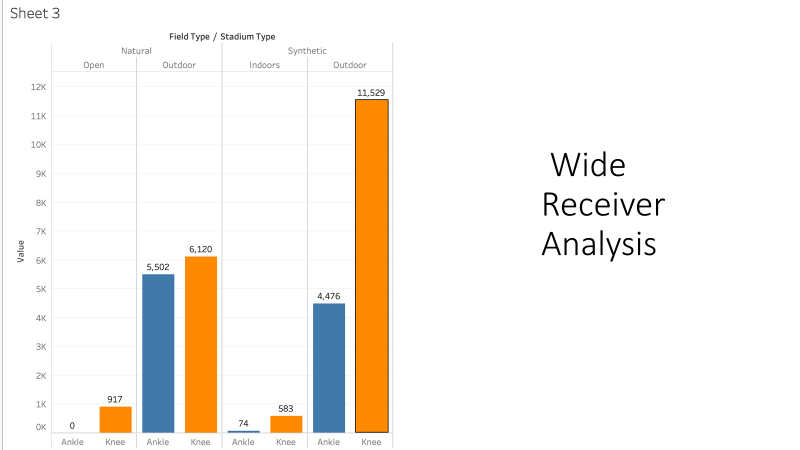


Figure 7. Wide receiver analysis

Here we are using field type ,stadium type and the amount of injuries to identify the number of injuries in indoors and outdoors in figure 7. The number of knee injuries on synthetic floor in outside stadium are more when compared to natural field type. Knee injuries are higher in all three types of positions types.

**Conclusion**

This project was taken to analyze to see the number of injuries in particular type of position, field type , stadium type because we wanted to identify most least preferred ground types .The reason being the investors and sponsors would spend so much of money on players well being and the profit would be maximized only when the important players play matches without missing. We came to conclusion to advise investors , stakeholders, sponsors, players and fans is playing on synthetic field will result in less number of injuries .

**References**

[1]https://www.kaggle.com/zinovadr/trained-nfl-data-ready-for-analysis

[2]https://en.wikipedia.org/wiki/National\_Football\_League

[3] <https://github.com/esrinivas08/NFL-Data-Analysis>