**Hypothesis:**

Artificial turf is more impact absorbent and it releases the player’s cleats slower than the natural turf which requires the player to apply more force and effort to start a movement. This might lead to overuse of muscles and therefore cause injuries.

We will study in detail all the movements, plays and the player data to identify any commonality that could lead us to probe this hypothesis.

**NFL Data Points**

* Quarter when the injury happened
* Time played by the player on the injury game
* Time played on artificial turf before injury
* Number of game on the season
* Games played in artificial turf
* Previous game before injury game (Artificial/natural)
* Player position and team (Special, Defense or Offense)
* Player hand (Left/Right)
* Training turf (Artificial/Natural)
* Number of seasons played by the player
* Previous injuries by the same player
* Player movement direction and position
* Player weight
* Weather and temp
* Humidity

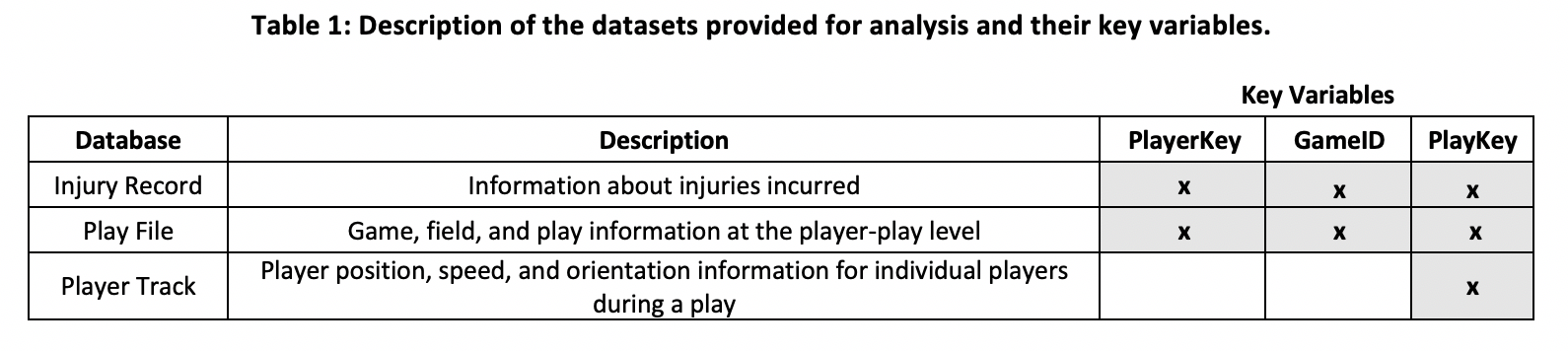
Players Biometrics

EDA exploratory data analysis

## Data Files

There are three files provided in the dataset, as described below:

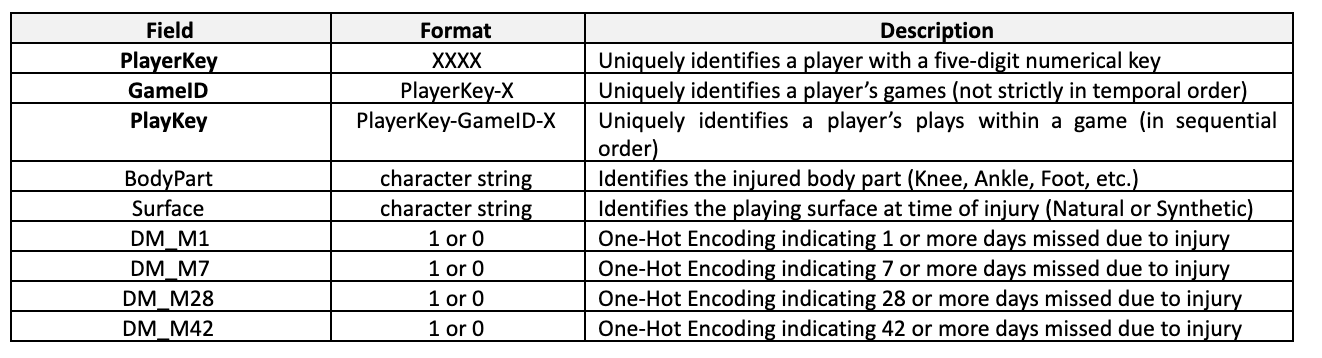
* Injury Record: The injury record file in .csv format contains information on 105 lower-limb injuries that occurred during regular season games over the two seasons. Injuries can be linked to specific records in a player history using the PlayerKey, GameID, and PlayKey fields.
* Play List: – The play list file contains the details for the 267,005 player-plays that make up the dataset. Each play is indexed by PlayerKey, GameID, and PlayKey fields. Details about the game and play include the player’s assigned roster position, stadium type, field type, weather, play type, position for the play, and position group.
* Player Track Data: player level data that describes the location, orientation, speed, and direction of each player during a play recorded at 10 Hz (i.e. 10 observations recorded per second).



## Field and Key Definitions

The following provides a description of each field contained within the datasets and their corresponding formats and descriptions. Key Variables are designated in bold.

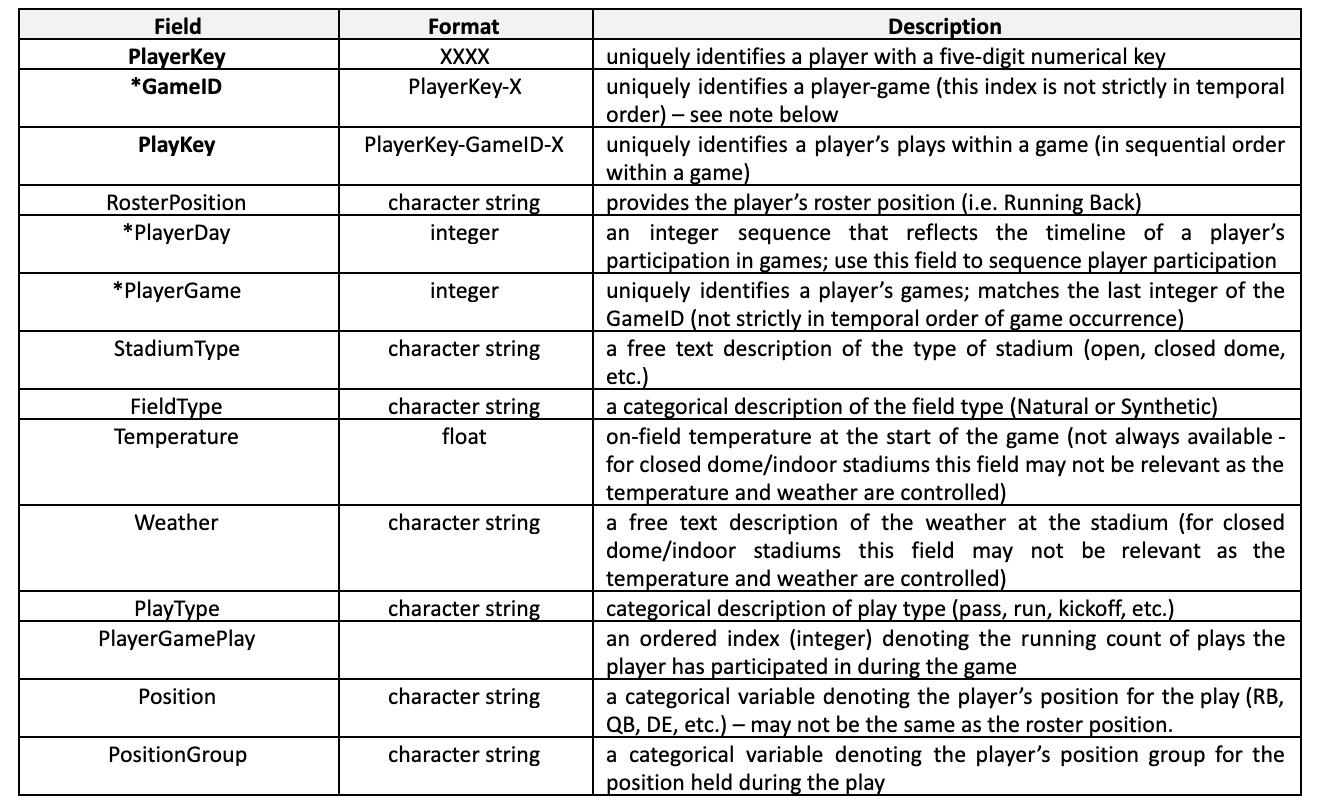
### Injury Data File



Note that there is not a PlayKey available for every injury. This indicates that the game in which the injury occurred is known, but the specific play in which the injury occurred was not noted at the time of injury.

### Play List

The play file contains information about each player-play in the dataset, to include the player’s assigned roster position, stadium type, field type, weather, play type, position for the play, and position group



\*Important Note: The GameID field is a unique identifier of player games but does not strictly reflect the order in which the games were played. The PlayerDay is an integer sequence that provides an accurate timeline for player game participation. In order to generate an accurate timeline of an individual player’s game participation, the PlayerDay variable should be used. The interval between days in the PlayerDay field for an individual player accurately reflects the interval in days between that player’s participation in games. Every player has a PlayerDay = 1 (note that this date is not the same for all players). Some players may have negative values for PlayerDay, which simply indicates participation in a game that occurred before their individually assigned PlayerDay = 1.

### Player Track Data

The player track file in .csv format includes player position, direction, and orientation data for each player during the entire course of the play collected using the Next Gen Stats (NGS) system. This data is indexed by PlayKey (which includes information about the player and game), with the time variable providing a temporal index within an individual play.

When processing the player track data, it is recommended to calculate velocity using the x, y position data and use those calculated velocities for any analysis (although we have provided the speed variable reported by the NGS system). The origin for the x and y coordinates is defined as the corner of the home endzone and home sideline (see Figure 1). The angles defined by orientation and direction are referenced from the y-axis of the coordinate system.

