STAT230 Project Starting Point

Group D: Arthur, Donna, Kaitlyn, Michael

Date

Proposal

PURPOSE

DATA

POPULATION

RESPONSE VARIABLE

We chose Post Injury Play Time as our response variable. This is a quantitative variable, and is represented by the average number of downs per game after the player returns from the injury. The distribution is bi-modal; many observations are close to 0 downs - likely influenced by players who did not return from their injury, and many appear at 99 downs. Currently, we are unable to provide descriptive statistics for Post Injury Play Time, because we have not been able to mutate our data to become purely numerical due to the presence of units in the observations.

EXPLANATORY VARIABLES

Pre-season injury is a qualitative variable, categorized by Yes – the player sustained an injury before the season began, and No – the player did not sustain an injury before the season.

Player position is a qualitative variable, categorized by all the possible positions on a football team.

Season is a qualitative variable, categorized in (start of season/end of season) format: 2012/2013, 2013/2014, and 2014/2015.
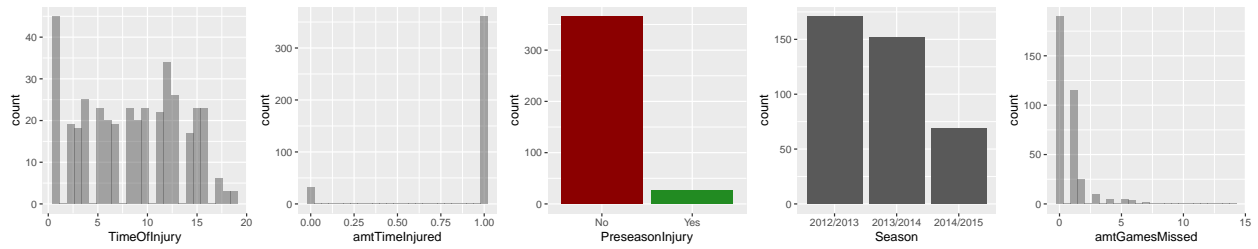
Number of games missed is a quantitative variable, measured by how many games in one season a player missed after sustaining the injury.

Average play time before the injury is a quantitative variable, measured in the average number of downs per game that occurred before the player sustained the injury.

Week of injury is a quantitative variable, measured in the number of weeks into the football season that the player sustained the injury.

Import Data & Wrangling

Exploratory Data Analysis (EDA)

Positions