**Predicting NFL Play Calls for Optimum Success**

Williams Anosike

Problem Definition

To be successful in the NFL a team needs to be competent at two main aspects: Calling the most beneficial plays at a certain situation and executing those plays. Teams that combine these two aspects to a considerably high degree often are the most successful and are usually in the post-season every year. So what do I wish to accomplish?

I want to use data science techniques to establish a system for coaches to make the best play calls. This will put them in the best situation to win and create a successful football club.

Dataset

Publicly available NFL data sources has been a major obstacle in the creation of modern, reproducible research in football analytics. Maksim Horowitz, Ron Yurko, and Sam Ventura built and released nflscrapR an R package which uses an API maintained by the NFL to scrape, clean, parse, and output clean datasets at the individual play, player, game, and season levels. This dataset has full game play-by -play from the 2009 NFL season to the 2018 NFL season. There are 255 columns and 449371 rows. The dataset is quite large, so to minimize the size of our problem, I will limit it to the Washington Redskins.

The dataset contains information such as:

* Home/Away team
* Type of play
* Player personnel on the field for any play
* Time of play
* Result of play

I will use all available resources to create a machine learning model that will help the Redskins make the best decisions for their 2020 NFL Season Schedule.