**ETL Project**

**Group:** Max, Tommy, Erin

**Goal:** We have decided to build a data inventory of NFL data. The inspiration for this project is that there are many resources for data but each resource has limits – limited free datasets, limited years, and data that is unrelated – and we would like to have a data source in the future that has a much greater flexibility.

With this in mind, we plan to do the following:

1. Focus on data from 2009 – 2013 since this time period is the most applicable for the datasets
2. We will take player data from NFLScrape and append to this dataset the NFLSavant player information
3. We will take game data from the NFLScrape and append NFLSavant weather data

**Future:** If we have time we will look at aggregating play-by-play game data and combine that with the NFLScrape game data and NFLSavant weather data (see #3).

Sources:

* <https://github.com/ryurko/nflscrapR-data>
  + Player info
  + Play by Play data
* <http://nflsavant.com/about.php>
  + CSV of All NFL Weather From 1960 to 2013
  + Every NFL Player in CSV format (Early years to 2013) – shows additional player info

Optional Sources:

* <https://www.kaggle.com/maxhorowitz/nflplaybyplay2009to2016>
* <https://sportsdata.io/developers/api-documentation/nfl> (not likely to be used)