Joe Hampton | Lane Sorell | Shane Reigodedios
CIS490 - Big Data
Project Proposal
6 November 2018

Problem Statement

Fantasy football is a group activity where men and women of all ages can vicariously enjoy their favorite sport by creating their own roster of players. The participants battle head to head, gaining points as their hand-picked players perform various actions during the seasonal games.

Objective

Help participants maximize their fantasy football points by providing a narrowed search of the available players, displaying only players who were involved in yardage gained + touchdowns. These picks will then be ranked by a rough estimate of the fantasy football points they have gathered since the start of the 2017 season (calculated as a function of various different categories (passing yards, touchdowns, etc.)) Therefore, allowing participants to draft a more effective team for 2018.

Background

As stated in the Problem Statement, fantasy football is an activity that accommodates all football-lovers. Fantasy football leagues can arise between groups of friends, acquainted colleagues/coworkers, or even amongst family members. Participants draft teams of real players that score points based on their in game performance. There are several different websites that support fantasy football games and each site has a different method for scoring. For our project, we decided to use ESPN's scoring system since it is the most popular platform for playing. They provide the following scoring background: [2]

PASSING	RUSHING
Standard scoring: · TD Pass = 4pts · Every 25 passing yards = 1pts · 2pt Passing Conversion = 2pts · Interceptions Thrown = -2pts **Custom options:** · Every <1, 5, 10, 20, 25, 50, 100> passing yards · Every <1, 5, 10> completions · Every <1, 5, 10> incomplete passes	**Standard scoring:** · TD Rush = 6pts · Every 10 rushing yards = 1pt · 2pt Rushing Conversion = 2pts **Custom options:** · Every <1, 5, 10, 20, 25, 50, 100 rushing yards · Every <1, 5, 10> rushing attem · TD Rush · 40+ yard TD rush bonus

Since most points are scored from yards gained and touchdowns we decided to formulate our scores based solely on these two statistics. This means that for passers, they would receive 4 points for every touchdown and 1 point for every 25 yards gained. For the runners and wide receivers, they would receive 6 points for every touchdown and 1 point for every 10 yards gained.

Related Work

With the relatively high popularity of fantasy football coupled with its ingrained competitiveness, it is not surprising that there are already many other (way more advanced) efforts to statistically funnel out the best scoring-players. One example is EPA, or Expected Points Added, a stat extrapolated from football analytics that, to some degree of accuracy, predicts the amount of fantasy points a certain player can add on a single play. [1]

Methodology

First and foremost, our team must understand the fundamentals of fantasy football, especially the scoring mechanic. Using the ESPN scoring standard scoring format [2], we will select a handful of high score-impacting actions that we can use to convert our dataset into an approximate fantasy football point value.

Our six different categories and their fantasy football point conversion rates were:

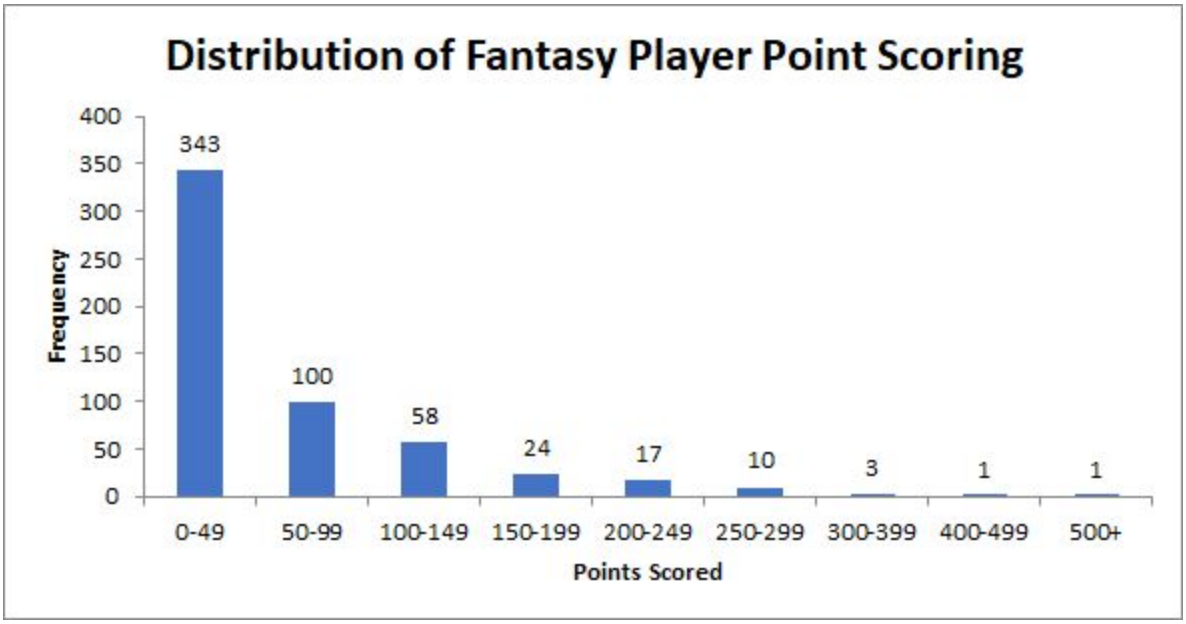
1. Rushing Touchdown (6 pts.)
2. Touchdown Reception (6 pts.)
3. Passing Touchdown (4 pts.)
4. Rushing Yards (1 pt per 10 yds.)
5. Receiving Yards (1 pt per 10 yds.)
6. Passing Yards (1 pt per 25 yds.)

Then, we will need to find a dataset. A trustworthy dataset for football statistics over a sizable amount of time (2 yrs+). Once found, we can transfer the data into MongoDB and apply MapReduce techniques to accumulate and then sort the data we are interested in.

Lastly, we simply need to display the highest scoring football players and their respective scores.

Results

Plays in database: 45,000+
Emits from mapper: 56,000+
Player number reduced to: 564



Name	Points
D.Prescott	518
D.Kizer	404
T.Gurley	323
R.Wilson	318
T.Brady	303
P.Rivers	288
M.Stafford	286
K.Cousins	281
B.Roethlisberger	279
A.Smith	274

Summary

Before we make any interpretations of the above results, a few things need to be stated. Firstly, not every NFL player in the league is included; like we mentioned earlier, we only looked at Quarterbacks, Receivers, and Running backs. In addition, within these three types of players, the points collected are still not exact. For example, QBs are meant to lose a point for every interception that they throw, however, this was not included into our implementation for simplicity's sake.

With these observations in mind, let's take a look at the results. It appears that D. Prescott (518 pts), D Kizer (404 pts), and T. Gurley (323 pts) are very high scoring players and would probably result in a more effective fantasy football team if they were drafted to a player's fantasy team.

For our future work, I would like to implement changes to address the aforementioned inaccuracies with the fantasy football score conversion. Ergo, it would incorporate other football player positions and formulate their score based on every standard ESPN scoring criteria [2].

References

[1] Foxsports These Are the Advanced Stats I Use to Win Fantasy Football Leagues <https://www.foxsports.com/fantasy/football/story/these-are-the-advanced-stats-i-use-to-win-fantasy-football-leagues-120414> December 6, 2018.
[2] Espn Scoring Formats <http://games.espn.com/ffl/resources/help/content?name=scoring-formats> December 14, 2018.
[3] Wikipedia <https://en.wikipedia.org/wiki/MongoDB> December 14, 2018.
[4] GitHub <https://github.com/ryurko/nflscrapR-data> December 14, 2018.
[5] Kaggle <https://www.kaggle.com/maxhorowitz/nflplaybyplay2009to2016> December 14, 2018.

Appendix

A1 Introduction
A1.1 Problem Statement
A1.2 Objective
B1 Background
B1.1 Figure (Scoring Format)
B1.2 Related Work
C1 Methodology
D1 Results
D1.1 Figure (Results)
E1 Summary
F1 References