CIS 490 Final Presentation Fantasy Football

Joe Hampton | Lane Sorell | Shane Reigodedios

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.

➔ **How can we apply big data techniques?**

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.

➔ **Okay, where do we start?**

Fantasy Football Scoring:

Studying the scoring format found on espn.com, we were able to identify 6 of (what we thought to be) the most score-impacting actions [1].

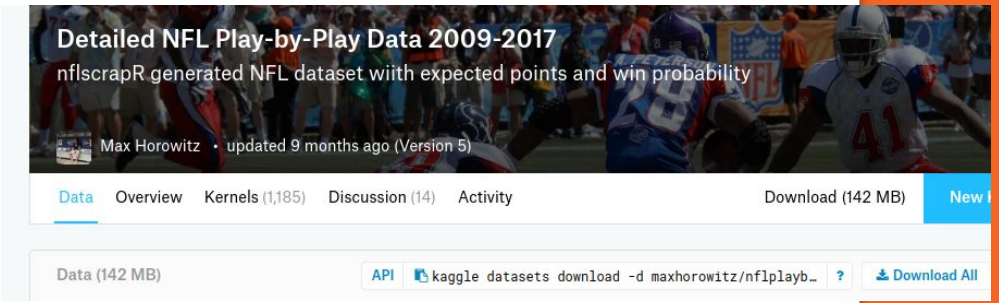
1. TD Reception (6pt)
2. TD Rush (6pt)
3. TD Pass (4pt)
4. Passing Yards (1pt per 25)
5. Rushing Yards (1pt per 10)
6. Receiving Yards (1 pt per 10)

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

➔ **Now we need data...**

Data:

After some searching, we found a Play by Play extensive dataset with data from plays all the way back to 2009. We decided only the most recent complete season offered by the dataset (2016) would be relevant to our objective [2].



➔ **Great! Now where do we put it...**

MongoDB:

With a data source finally found, we were able to import it as a CSV file into MongoDB, a cross-platform document-oriented database. Program [3].

(The free-tier option for MongoDB couldn't accommodate the size of our initial CSV file, so we had to run some Excel scripts to remove some data that we knew we had no use for)

➔ **Finally, we can make some real progress..**

WRENCH!:

Upon further inspection, the data was NOT consistent. Fields like Touchdown had unintelligible data entries, like -3 OR 17 instead of the expected values of. 1 & 0 (a touchdown was scored or it was not). Additionally, there was zero explanation of how any data fields were filled out.

(No worries though, we found a better data source of the 2017 season [4])

➔ **Something Something MapReduce?**

MapReduce - Mapper:

```
var mapFunction1 = function() {
  var value = {
    touchdowns: 0,
    yards: parseFloat(this.yards_gained),
    IsQB: false
  };
  if(this.play_type == "run"){
    value.touchdowns = parseFloat(this.rush_touchdown);
    emit(this.rusher_player_name, value);
  }
  if(this.play_type == "pass"){
    value.touchdowns = parseFloat(this.pass_touchdown);
    emit(this.receiver_player_name,value);
    value.IsQB = true;
    emit(this.passer_player_name,value);
  }
};
```

➔ MapReduception?

MapReduce - Reducer(+finalizer):

```
var reduceFunction1 = function(name, vals) {
  reducedVal = { touchdowns: 0, yards: 0, IsQB: vals[0].IsQB};
  for (var idx = 0; idx < vals.length; idx++) {
    reducedVal.touchdowns += vals[idx].touchdowns;
    reducedVal.yards += vals[idx].yards;
    }
  return reducedVal;
};
```

```
var finalizeFunction1 = function(name, vals) {
  vals.points = 0;
  if(vals.IsQB == false){ vals.points += Math.floor(vals.yards / 10.0);
    vals.points += vals.touchdowns * 6;
  }else{ vals.points += Math.floor(vals.yards / 25.0);
    vals.points += vals.touchdowns * 4;}
    return vals;
};
```

➔ Now for the part you've all been waiting for...

Results:


```
D.Prescott,"{"touchdowns":28,"yards":3503,"IsQB":false,"points":518}"
D.Kizer,"{"touchdowns":16,"yards":3089,"IsQB":false,"points":404}"
T.Gurley,"{"touchdowns":19,"yards":2092,"IsQB":false,"points":323}"
R.Wilson,"{"touchdowns":37,"yards":4261,"IsQB":true,"points":318}"
T.Brady,"{"touchdowns":32,"yards":4393,"IsQB":true,"points":303}"
P.Rivers,"{"touchdowns":28,"yards":4408,"IsQB":true,"points":288}"
M.Stafford,"{"touchdowns":29,"yards":4270,"IsQB":true,"points":286}"
K.Cousins,"{"touchdowns":31,"yards":3940,"IsQB":true,"points":281}"
B.Roethlisberger,"{"touchdowns":28,"yards":4190,"IsQB":true,"points":279}"
A.Smith,"{"touchdowns":27,"yards":4169,"IsQB":true,"points":274}"
C.Wentz,"{"touchdowns":33,"yards":3444,"IsQB":true,"points":269}"
D.Brees,"{"touchdowns":25,"yards":4225,"IsQB":true,"points":269}"
C.Newton,"{"touchdowns":28,"yards":3837,"IsQB":true,"points":265}"
J.Goff,"{"touchdowns":29,"yards":3702,"IsQB":true,"points":264}"
L.Bell,"{"touchdowns":11,"yards":1954,"IsQB":false,"points":261}"
B.Bortles,"{"touchdowns":23,"yards":3903,"IsQB":true,"points":248}"
K.Hunt,"{"touchdowns":11,"yards":1782,"IsQB":false,"points":244}"
```


➔ **That's a Wrap, Thanks for listening!**

References

[1] http://games.espn.com/ffl/resources/help/content?name=scoring-formats

[2] https://www.kaggle.com/maxhorowitz/nflplaybyplay2009to2016

[3] https://en.wikipedia.org/wiki/MongoDB

[4] https://github.com/ryurko/nflscrapR-data