Randal Morris

Dr. Mok

IS412

9/25/15

Project Description and Activities

**The Problem**

Using a single fantasy sites projections for a league is ineffective. People can use multiple sites to help, but it’s still ineffective.

**The Solution**

Using R to pull data from multiple sites and create a more accurate projection. The scripts I am developing will take data from multiple fantasy football sources (ESPN, CBS, Fantasy Pros, etc.) and allow users to generate either a season projection (total points for a season) or weekly projections. The technically knowledge of those who use the scripts will be minimized to having to edit a league settings file. After getting the weekly projection scripts working, I intend to use the R Package Shiny to create the same for use as a web app. This will require less technical knowledge and increase convenience.

My initial focus is weekly projections because the season has started and it has a dual focus towards weekly pay leagues such as Draft Kings or Fan Duel. Many of those who play in fantasy football leagues pay as well. By using the descriptive area of analytics this will allow users to make better decisions on who to start/sit, pay league choices, and drafting for leagues.

**How The System Will Work**

The user has to modify the League Settings file for their league to allow for accurate projections. This is just editing numbers and will have comments to help those with less technical knowledge. For users who are using the weekly projections need to also adjust the week as well.

**Weekly Projections Technical Workings**

There will be a main script that will allow a user to run all the scripts from one. There will be 2 scripts for Running Backs, Wide Receivers, and Tight Ends so leagues with points per reception (points for each catch) is included. Each position will do these processes (RB, WR, and TE will do 2 repetitions). Once they complete, then the script will perform the Value Gap Analysis. This is to show any players who are under or over rated based on other team projections. For example, If QB#1 is projected 300 passing yards, but all the receivers (RB, WR, TE) only have a total reception yards of 200 then we must determine if the QB is over rated (which can reduce his points for the week) or if the receivers are under rated (which can increase their points).

The first process in the script pulls data from each source. This process reads the projection data from sources and the rank data from Fantasy Pros Expert Consensus ranks. Then create a data table for each source.

The second process is cleaning and standardizing the data tables to removes rows that are not players, trims the number of players to Top 40 for RB, WR, TE. Rest of the positions will be trimmed to the Top 32. Unwanted columns that are not important are removed. Columns are renamed to be uniformed. All columns with numbers are converted to numeric. Since some sites combine player names, team, and positon in the same column so we must separate them. Then we reorder each data table for some more uniformity.

The third step is to merge each data table into a master projection data table.

The forth step is to subset the projections data table into a new data table that contains only the name and each source projection.

The fifth step is the calculations steps. There are 3 main ways to view projections across multiple sources, Average, Robust Average, and Hodges-Lehmann. This process also adds in the Std Deviation, and ceiling/floor.

The sixth process sorts and orders our projections. Then displays any graphs and saves the data.

**Season Projections Technical Workings**

The season projections will have a similar process as the weekly projections. There is a main script that runs all the scripts. This process is work in progress at best since the season has started, a weekly projection has a higher priority at this moment. Data and CSV files will be save for every projection source and calculations at various times in the process. Graphics will be created for density of each projection source (points related to positon), and various calculations.

The first process in the script pulls data from each source. This process reads the projection data from sources and the rank data from Fantasy Pros Expert Consensus ranks. Then create a data table for each source.

The second process is cleaning and standardizing the data tables to removes rows that are not players, Unwanted columns that are not important are removed. Columns are renamed to be uniformed. All columns with numbers are converted to numeric. Since some sites combine player names, team, and positon in the same column so we must separate them. Then we reorder each data table for some more uniformity.

The third process merges the data in to a master projection data table.

The forth process begins the calculations. The calculations done are projections based on league settings, Wisdom of the Crowd, Risk, Value over replacement, Avg cost, roster optimization, risk optimization, and simulations to increase projected points and lower risk.