Analysis of Fantasy Football

John Wattenbarger

Jack Follis, Ph.D., Advisor

Department of Mathematics, Computer Science, and Cooperative Engineering

University of St. Thomas, Houston, Texas 77006

Purpose

This project was created to examine NFL statistics over the past 10 years and see how they relate to fantasy football. This project also seeks to discover how different position scoring has changed over time. These relationships and changes in NFL statistics were displayed using data visualization techniques.

Results

Scatterplots were made for each position to determine the relationship between yards, touchdowns, and fantasy points. Another Scatterplot was created to show the changes in fantasy football scoring over the past 10 years.

Process

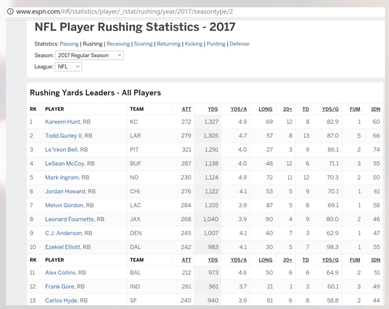
Data was gathered from ESPN.com using web scrapping (extracting data from web sites, and converting it into a useable format) with R (a programming language used for statistics). 10 years of NFL statistics were gathered and divided into 4 data tables based on player's NFL positions: quarterback, running back, wide receiver, and tight end.

```r
library('rvest')
library(ggplot2)
library(stringr)

# Take an HTML_Table from ESPN player stats, get the data,
getHtmlTable <- function(columns, url)
{
  webpage <- read_html(url)
  values <- html_table(webpage)
  table <- data.frame(matrix(unlist(values), ncol=columns, byrow=F))
  colnames(table) <- lapply(table[1, ], as.character)
  table <- unique(table)
  table <- table[-1,]

  return(table)
}

# makes the 1st column display the year
addYear <- function(table, year)
{
  table[,1] <- year
  colnames(table)[1] <- "Year"
  return(table)
}

# Creates and combines tables from the last 10 years
multiYearTable <- function(type)
{
  for(i in 1:10)
  {
    year <- 2007+i
    tempTable <- makeTable(type, year)
    tempTable <- addYear(tempTable, year)
    if (i == 1)
      tableCombined <- tempTable
    else
      tableCombined <- rbind(tableCombined, tempTable)
  }
}
```

Sample of R Code used:

rvest package was used to gather information from espn.com. ggplot2 package was used to create visuals

Web scraping with R

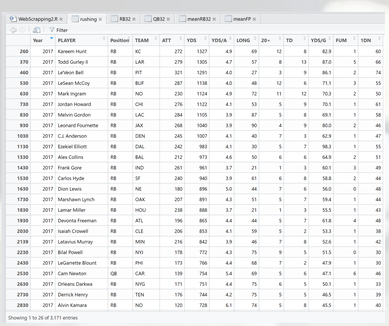


Webpage:

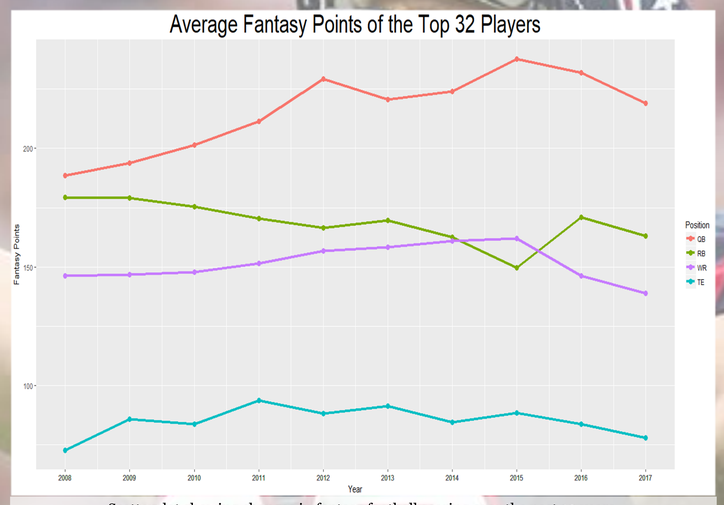
Passing, rushing, and receiving statistics were gathered from 2008-2017

http://www.espn.com/nfl/statistics/player/_/stat/

Convert webpage to data table

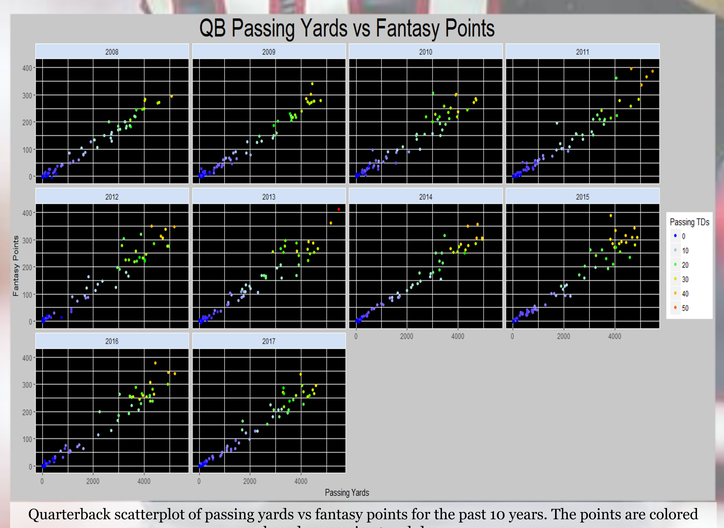


Data from the espn.com was gathered into a data table in R. Data was then cleaned up into a more useable format.

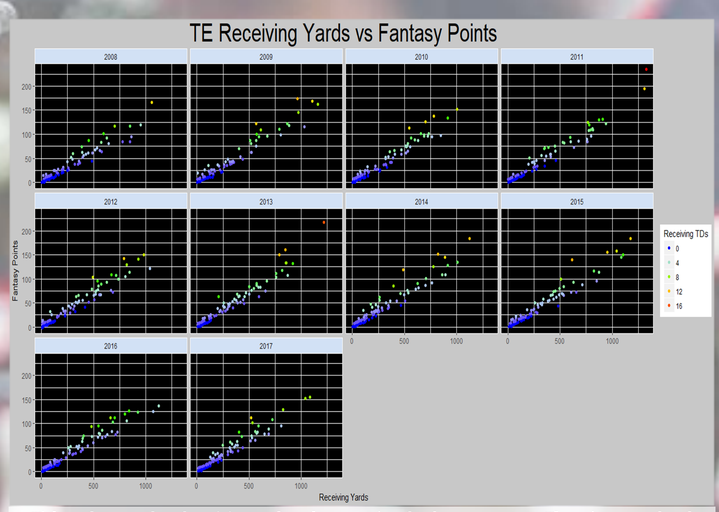


Scatterplot showing changes in fantasy football scoring over the past 10 years.

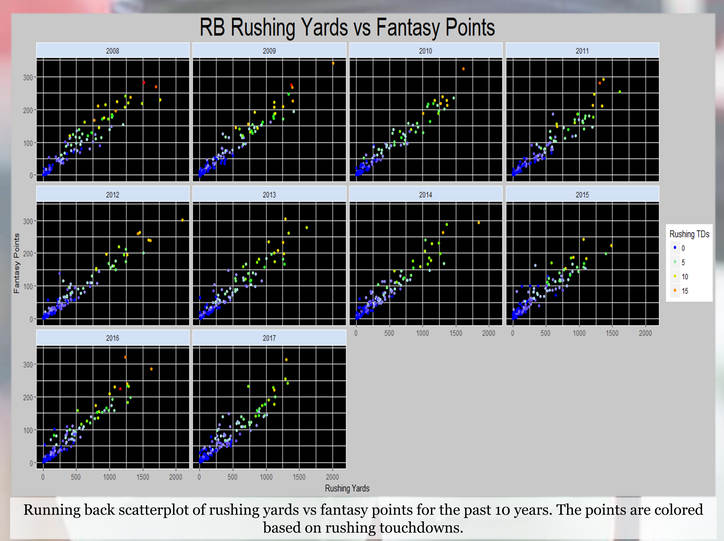
Wide receiver scatterplot of receiving yards vs fantasy points for the past 10 years. The points are colored based on receiving touchdowns.



Quarterback scatterplot of passing yards vs fantasy points for the past 10 years. The points are colored based on passing touchdowns.



Tight end scatterplot of receiving yards vs fantasy points for the past 10 years. The points are colored based on receiving touchdowns.



Running back scatterplot of rushing yards vs fantasy points for the past 10 years. The points are colored based on rushing touchdowns.

Conclusion

The NFL has become more of a passing league over the past decade, and this trend is evident in these graphs. Quarterback and wide receiver points generally increased, and running back points generally decreased from 2008 to 2016.The past two years (2016 and 2017) show that the NFL might be becoming more focused on rushing, as quarterback and wide receiver points were lower, while running back points were higher. These graphs also show that quarterback and running back points seem to have an inverse relationship.

Each position showed a strong relationship between yards and fantasy points. Running backs had the weakest relationship between yards and fantasy points, likely because running backs are reliant on both rushing and receiving. The tight end position had a few players greatly outperform their peers. In 2011 and 2013 there is a noticeable gap between the best tight ends and the others.