Moneyball

Chris Huszar

August 13, 2018

# Fantasy Football Analysis

#Load Packages  
require('tidyr')

## Loading required package: tidyr

## Warning: package 'tidyr' was built under R version 3.4.2

require('stringr')

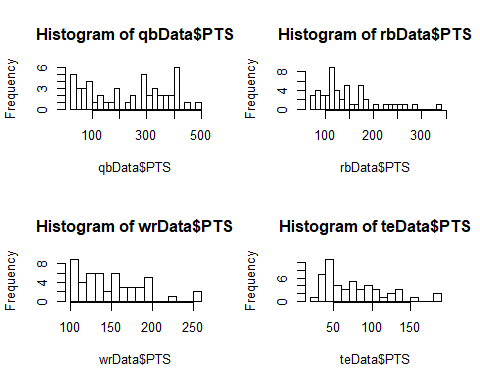
## Loading required package: stringr

## Warning: package 'stringr' was built under R version 3.4.3

require('ggplot2')

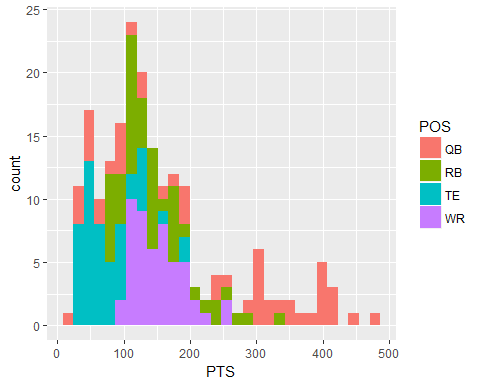
## Loading required package: ggplot2

#Read in Data  
datapath<-"C:/Users/cmhus/iCloudDrive/Analytics Projects/Football Data/"  
data<-read.csv(paste(datapath,"2018 Fantasy Football QB RB WR TE.csv",sep="/"))  
auctionData<-read.csv(paste(datapath,"2017 Draft Results.csv",sep="/"))  
  
#Scrub Data  
auctionData$auctionPrice<-as.numeric(gsub("[[:punct:]]", " ", auctionData$auctionPrice))  
  
splitNames<-data.frame(str\_split\_fixed(auctionData$playerTeamPosition, ",", 2))  
splitNamesTeam<-str\_trim(substr(splitNames$X2,1,4))  
splitNamesPos<-str\_trim(substr(splitNames$X2,5,7))  
  
auctionData$playerName<-splitNames$X1  
auctionData$team<-splitNamesTeam  
auctionData$pos<-splitNamesPos  
auctionData<-subset(auctionData[,c(4,5,6,1,3)],team!="" & pos!="K")  
  
#Set league info  
qbStart<-1  
rbStart<-2  
wrStart<-3  
teStart<-1  
  
teams<-8  
budget<-200  
  
#Subset Positions on 2017 points data  
qbData<-subset(data,POS=="QB")  
rbData<-subset(data,POS=="RB")  
wrData<-subset(data,POS=="WR")  
teData<-subset(data,POS=="TE")  
  
#Order by PTS  
qbData <- qbData[order(-qbData$PTS), ]  
rbData <- rbData[order(-rbData$PTS), ]  
wrData <- wrData[order(-wrData$PTS), ]  
teData <- teData[order(-teData$PTS), ]  
  
#Subset Positions on 2017 auction data  
qbAucData<-subset(auctionData,pos=="QB")  
rbAucData<-subset(auctionData,pos=="RB")  
wrAucData<-subset(auctionData,pos=="WR")  
teAucData<-subset(auctionData,pos=="TE")  
  
#Order by 2017 Auction prices  
qbAucData <- qbAucData[order(-qbAucData$auctionPrice), ]  
rbAucData <- rbAucData[order(-rbAucData$auctionPrice), ]  
wrAucData <- wrAucData[order(-wrAucData$auctionPrice), ]  
teAucData <- teAucData[order(-teAucData$auctionPrice), ]  
  
#look at histograms of each position pts distribution  
par(mfrow=c(2,2))  
hist(qbData$PTS, breaks = 20)  
hist(rbData$PTS, breaks = 20)  
hist(wrData$PTS, breaks = 20)  
hist(teData$PTS, breaks = 20)



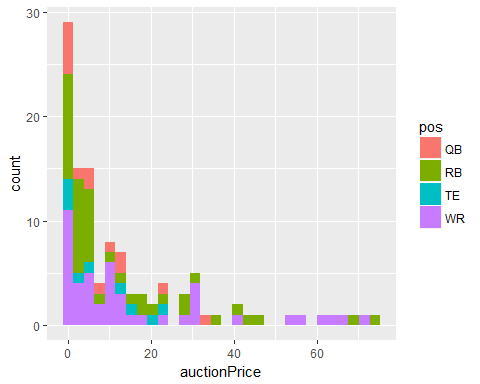
#View stacked historgam to get level of pts and number of players for each position and pts level  
par(mfrow=c(1,2))  
ggplot(data, aes(x=PTS, fill=POS)) +  
 geom\_histogram()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

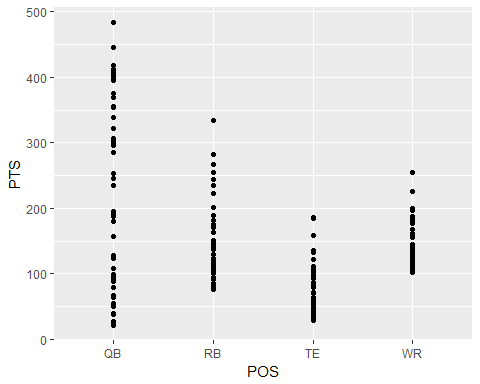


ggplot(auctionData, aes(x=auctionPrice, fill=pos)) +  
 geom\_histogram()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



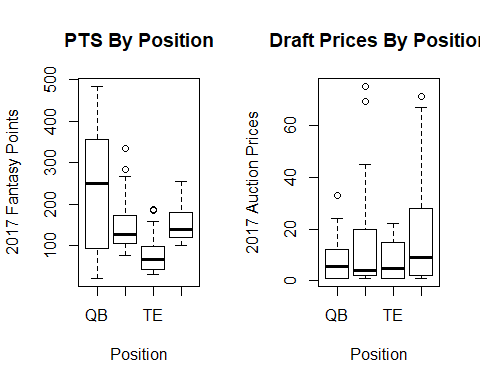
par(mfrow=c(1,2))  
ggplot(data, aes(y=PTS,x=POS)) +  
 geom\_point()



ggplot(auctionData, aes(y=auctionPrice,x=pos)) +  
 geom\_point()

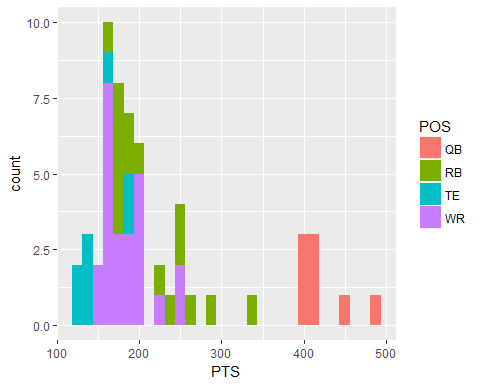


par(mfrow=c(1,2))  
boxplot(PTS~POS,data=data, main="PTS By Position",   
 xlab="Position", ylab="2017 Fantasy Points")  
  
boxplot(auctionPrice~pos,data=auctionData, main="Draft Prices By Position",   
 xlab="Position", ylab="2017 Auction Prices")



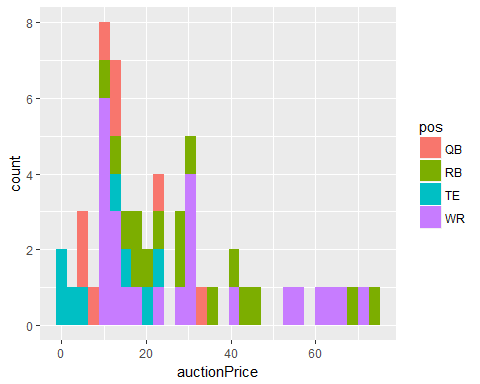
#Look at just the starting spots by position  
qbDataTop<-head(qbData,teams\*qbStart)  
rbDataTop<-head(rbData,teams\*rbStart)  
wrDataTop<-head(wrData,teams\*wrStart)  
teDataTop<-head(teData,teams\*teStart)  
  
topStarters<-rbind(qbDataTop,rbDataTop,wrDataTop,teDataTop)  
ggplot(topStarters, aes(x=PTS, fill=POS)) +  
 geom\_histogram()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



#Now the same with auction values  
qbAucDataTop<-head(qbAucData,teams\*qbStart)  
rbAucDataTop<-head(rbAucData,teams\*rbStart)  
wrAucDataTop<-head(wrAucData,teams\*wrStart)  
teAucDataTop<-head(teAucData,teams\*teStart)  
  
topAucStarters<-rbind(qbAucDataTop,rbAucDataTop,wrAucDataTop,teAucDataTop)  
ggplot(topAucStarters, aes(x=auctionPrice, fill=pos)) +  
 geom\_histogram()

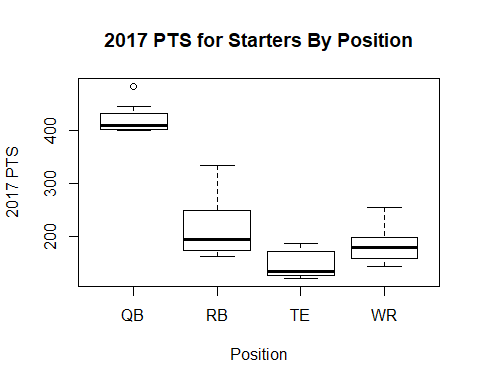
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



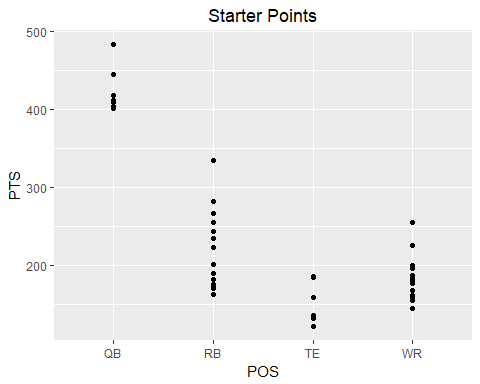
#Check the spread in PTS of the starters by position  
qbspread<-head(qbDataTop$PTS,1)-tail(qbDataTop$PTS,1)  
rbspread<-head(rbDataTop$PTS,1)-tail(rbDataTop$PTS,1)  
wrspread<-head(wrDataTop$PTS,1)-tail(wrDataTop$PTS,1)  
tespread<-head(teDataTop$PTS,1)-tail(teDataTop$PTS,1)  
  
starterspread<-data.frame(cbind(c('QB','RB','WR','TE'),rbind(qbspread,rbspread,wrspread,tespread)))  
rownames(starterspread)<-NULL  
colnames(starterspread)<-c('Position','Pts Spread')  
  
starterspread

## Position Pts Spread  
## 1 QB 82.5  
## 2 RB 171  
## 3 WR 110.5  
## 4 TE 65

par(mfrow=c(1,1))  
boxplot(PTS~POS,data=topStarters, main="2017 PTS for Starters By Position",   
 xlab="Position", ylab="2017 PTS")



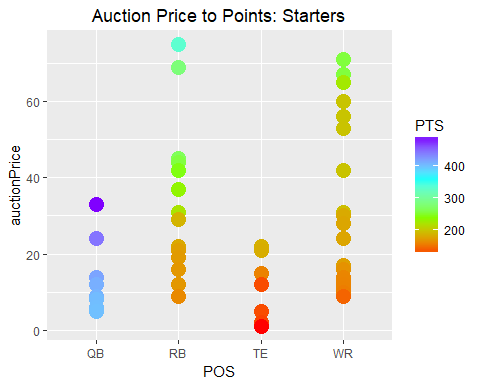
#Another look  
ggplot(topStarters, aes(y=PTS,x=POS)) +  
 geom\_point() +  
 ggtitle("Starter Points") +  
 theme(plot.title = element\_text(hjust = 0.5))



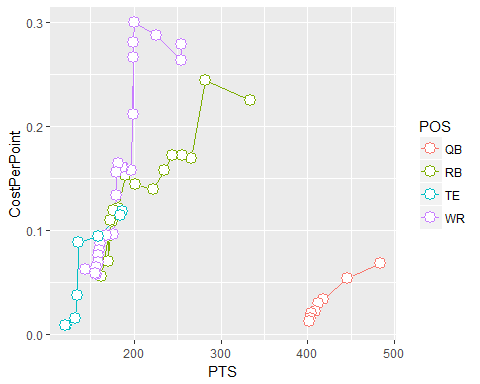
startersPtsValue<-cbind(topStarters[,c('POS','PTS')],'auctionPrice'=topAucStarters[,'auctionPrice'])  
  
startersPtsValue$CostPerPoint<-startersPtsValue$auctionPrice/startersPtsValue$PTS  
startersPtsValue

## POS PTS auctionPrice CostPerPoint  
## 1 QB 483.8 33 0.068210004  
## 2 QB 445.3 24 0.053896250  
## 3 QB 418.3 14 0.033468802  
## 4 QB 411.8 12 0.029140359  
## 5 QB 409.0 9 0.022004890  
## 6 QB 403.8 8 0.019811788  
## 7 QB 402.8 6 0.014895730  
## 8 QB 401.3 5 0.012459507  
## 51 RB 334.0 75 0.224550898  
## 52 RB 282.5 69 0.244247788  
## 53 RB 266.5 45 0.168855535  
## 54 RB 255.5 44 0.172211350  
## 55 RB 244.0 42 0.172131148  
## 56 RB 235.0 37 0.157446809  
## 57 RB 222.5 31 0.139325843  
## 58 RB 201.0 29 0.144278607  
## 59 RB 189.5 29 0.153034301  
## 60 RB 182.0 22 0.120879121  
## 61 RB 176.0 21 0.119318182  
## 62 RB 175.5 19 0.108262108  
## 63 RB 173.0 19 0.109826590  
## 64 RB 171.0 16 0.093567251  
## 65 RB 171.0 12 0.070175439  
## 66 RB 163.0 9 0.055214724  
## 101 WR 255.0 71 0.278431373  
## 102 WR 254.5 67 0.263261297  
## 103 WR 226.0 65 0.287610619  
## 104 WR 200.0 60 0.300000000  
## 105 WR 199.8 56 0.280280280  
## 106 WR 199.0 53 0.266331658  
## 107 WR 199.0 42 0.211055276  
## 108 WR 196.5 31 0.157760814  
## 109 WR 187.5 30 0.160000000  
## 110 WR 186.5 30 0.160857909  
## 111 WR 182.5 30 0.164383562  
## 112 WR 180.0 28 0.155555556  
## 113 WR 179.5 24 0.133704735  
## 114 WR 176.5 17 0.096317280  
## 115 WR 168.0 16 0.095238095  
## 116 WR 161.0 14 0.086956522  
## 117 WR 160.5 13 0.080996885  
## 118 WR 159.5 12 0.075235110  
## 119 WR 159.5 11 0.068965517  
## 120 WR 159.0 11 0.069182390  
## 121 WR 156.5 10 0.063897764  
## 122 WR 156.5 9 0.057507987  
## 123 WR 155.0 9 0.058064516  
## 124 WR 144.5 9 0.062283737  
## 151 TE 186.5 22 0.117962466  
## 152 TE 184.5 21 0.113821138  
## 153 TE 159.0 15 0.094339623  
## 154 TE 135.5 12 0.088560886  
## 155 TE 135.0 5 0.037037037  
## 156 TE 132.0 2 0.015151515  
## 157 TE 122.0 1 0.008196721  
## 158 TE 121.5 1 0.008230453

ggplot(startersPtsValue, aes(POS, auctionPrice ,colour =PTS)) +   
 geom\_point(size=5) +  
 scale\_colour\_gradientn(colours=rainbow(4)) +  
 ggtitle("Auction Price to Points: Starters") +  
 theme(plot.title = element\_text(hjust = 0.5))



ggplot(data=startersPtsValue, aes(x=PTS, y=CostPerPoint, group = POS, colour = POS)) +  
 geom\_line() +  
 geom\_point( size=4, shape=21, fill="white")



ggplot(data=startersPtsValue, aes(x=CostPerPoint, y=PTS, group = POS, colour = POS)) +  
 geom\_line() +  
 geom\_point( size=4, shape=21, fill="white")

