Jeremiah Graff

IST 719: Tuesday Night 8pm

Final Project Poster Script

Data Source file from:

<https://www.fangraphs.com/leaders.aspx?pos=all&stats=bat&lg=all&qual=0&type=8&season=2019&month=0&season1=2019&ind=0&team=&rost=&age=&filter=&players=&startdate=&enddate=>

Unexecuted Code:

# Jeremiah Graff

# IST 719: 8pm Tuesday Night

# Final Poster Project R Script

# Dataset: 1,410 Observations of 24 variables (24\*4) \* (1,410/100)

# (96 \* 14.1) = 1353.6 >= 100: True

###########################################################################

## Part 1

# choosing baseball file that shows key batting statistics for qualified players

fname <- file.choose()

# reading the file in

baseball <- read.csv(file = fname, header = T, stringsAsFactors = FALSE)

# seeing the structure of the file...mostly ints and numbers

str(baseball)

colnames(baseball)[1] <- "player"

par(mfrow = c(2,2))

boxplot(baseball$WAR, horizontal = T, col = "lightgreen",

ylab = 'boxplot for player WAR', xlab = 'WAR', main = "Distribution of WAR")

plot(baseball$WAR, ylab = 'WAR', xlab = 'Player Index', main = "Distribution of WAR"

, pch= 17)

hist(baseball$WAR, col = "lightgreen", xlab = "WAR", ylab = "WAR Count",

main = "WAR Distribution Histogram", breaks = 18)

par(mfrow=c(1,1))

# aggregating the player data to see team WAR information

teams <- data.frame(aggregate(baseball$WAR,list(baseball$Team,baseball$League),sum))

# assigning new name to War cumulation column

colnames(teams) <- c('team','league','total\_war')

# removing the players who traded mid-season and didn't have a permanent team

teams <- teams[-1,]

str(teams)

# sorting the teams by total WAR

teams <- teams[order(teams$total\_war),]

# creating a color pallette to make AL teams red and NL teams blue

fill\_colors <- c()

for (i in 1:length(teams$league)) {

if (teams$league[i] == 'AL' ) {

fill\_colors <- c(fill\_colors, "Red")

} else {

fill\_colors <- c(fill\_colors, "Blue")

}

}

### team plot

barplot(teams$total\_war,names.arg = teams$team, horiz = T, las = 1,

xlim = c(-10,40),cex.names = 0.5, col = fill\_colors

, main = "Sorted WAR by Team"

,xlab = "Collective WAR"

,ylab = 'Team')

#looking to see top 10 players from 2019 in WAR terms

top10plyrs <- head(baseball,10)

# sorting the list of top players

top10plyrs <- top10plyrs[order(top10plyrs$WAR),]

# creating a color pallette to make AL players red and NL playerss blue

fill\_colors <- c()

for (i in 1:length(top10plyrs$League)) {

if (top10plyrs$League[i] == 'AL' ) {

fill\_colors <- c(fill\_colors, "Red")

} else {

fill\_colors <- c(fill\_colors, "Blue")

}

}

### player plot

barplot(top10plyrs$WAR,names.arg = top10plyrs$player,horiz = T,

las = 1, cex.names = 0.5, col = fill\_colors

, main = "Top 10 WAR Players"

,xlab = "Player's WAR"

,ylab = 'Player')

### multidimensional plot

league\_comp <- tapply(baseball$WAR,list(baseball$League,baseball$Division),sum)

league\_comp <- league\_comp[-1,-1]

barplot(league\_comp,beside = T

, col = c('tomato','lightblue')

, main = "WAR by League by Division"

, leg = T

, xlab = "Division"

, ylab = 'Collective WAR')