DS Final Project

Zachary Dougherty

November 29, 2018

```
season_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\nba-players-stats\\Seasons_Stats.csv"
,
  header = T,
  sep = ",")
#P1
players<- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\nba-players-stats\\Players.csv", sep = "," )
Player_data<- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\nba-players-stats\\player_data.csv",
  head = TRUE,
  sep = ",")
GSW_stats18<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW_statistics.csv",
                       header = TRUE,
                       sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW_statistics.csv'
```

```
row.names(GSW_stats18) <- c("team(18)", "team/G(18)", "L Rank(18)")
GSW_stats17<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2016-2017.csv",
                       header = TRUE,
                       sep = ",")
row.names(GSW_stats17) <- c("team(17)", "team/G(17)", "L Rank(17)")
GSW_stats16<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2015-2016.csv",
                       header = TRUE,
                       sep = ",")
row.names(GSW_stats16) <- c("team(16)", "team/G(16)", "L Rank(16)")
GSW_stats15<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2014-2015.csv",
                       header = TRUE,
                       sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW2014-2015.csv'
```

```
row.names(GSW_stats15) <- c("team(15)", "team/G(15)", "L Rank(15)")
GSW_stats14<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2013-2014.csv",
                       header = TRUE,
                       sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW2013-2014.csv'
```

```
row.names(GSW_stats14) <- c("team(14)", "team/G(14)", "L Rank(14)")
GSW_stats13<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2012-2013.csv",
                       header = TRUE,
```

```
                        sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW2012-2013.csv'
```

```
row.names(GSW_stats13) <- c("team(13)", "team/G(13)", "L Rank(13)")
GSW_stats12<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2011-2012.csv",
                        header = TRUE,
                        sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW2011-2012.csv'
```

```
row.names(GSW_stats12) <- c("team(12)", "team/G(12)", "L Rank(12)")
GSW_stats11<- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\GSW2010-2011.csv",
                        header = TRUE,
                        sep = ",")
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote
## = quote, : incomplete final line found by readTableHeader on 'C:\Users
## \bbygbbyg\Documents\DS_Final_Project\GSW2010-2011.csv'
```

```
#P2
twenty_ten_eleven_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2010_11_team_per_game_stats.csv",
  sep = ",")
twenty_eleven_twelve_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2011_12_team_per_game_stats.csv",
  sep = ",")
twenty_twelve_thirteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2012_13_team_stats_per_game.csv",
  sep = ",")
twenty_thirteen_fourteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2013_14_team_per_game_stats.csv",
  sep = ",")
twenty_fourteen_fifteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2014_15_team_stats_per_game.csv",
  sep = ",")
twenty_fifteen_sixteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2015_16_team_stats_per_game.csv",
  sep = ",")
twenty_sixteen_seventeen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2016_17_team_per_game_stats.csv",
  sep = ",")
twenty_seventeen_eighteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2017_18_team_stats_per_game.csv",
  sep = ",")
twenty_eighteen_nineteen_team_stats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\2018_19_season_stats_per_game.csv",
  sep = ",")
#P3
d_rose <- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\D_Rose_stats.csv",
                   sep = ",")
lbj <- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\LBJ_stats.csv",
```

```
                    sep = ",")
kd <- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\KD_stats.csv",
                    sep = ",")
sc <- read.csv("C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\Steph_Curry_stats.csv",
                    sep = ",")
russ_wb <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\Russell_Westbrook_stats.csv",
                    sep = ",")
james_h <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\James_Harden_stats.csv",
                    sep = ",")
#P4
nbaStats <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\nbastats.csv")
nbaRatings <- read.csv(
  "C:\\Users\\bbygbbyg\\Documents\\DS_Final_Project\\nbaratings.csv")

library(RColorBrewer)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
library(ggplot2)
library(stringr)
```

Introduction

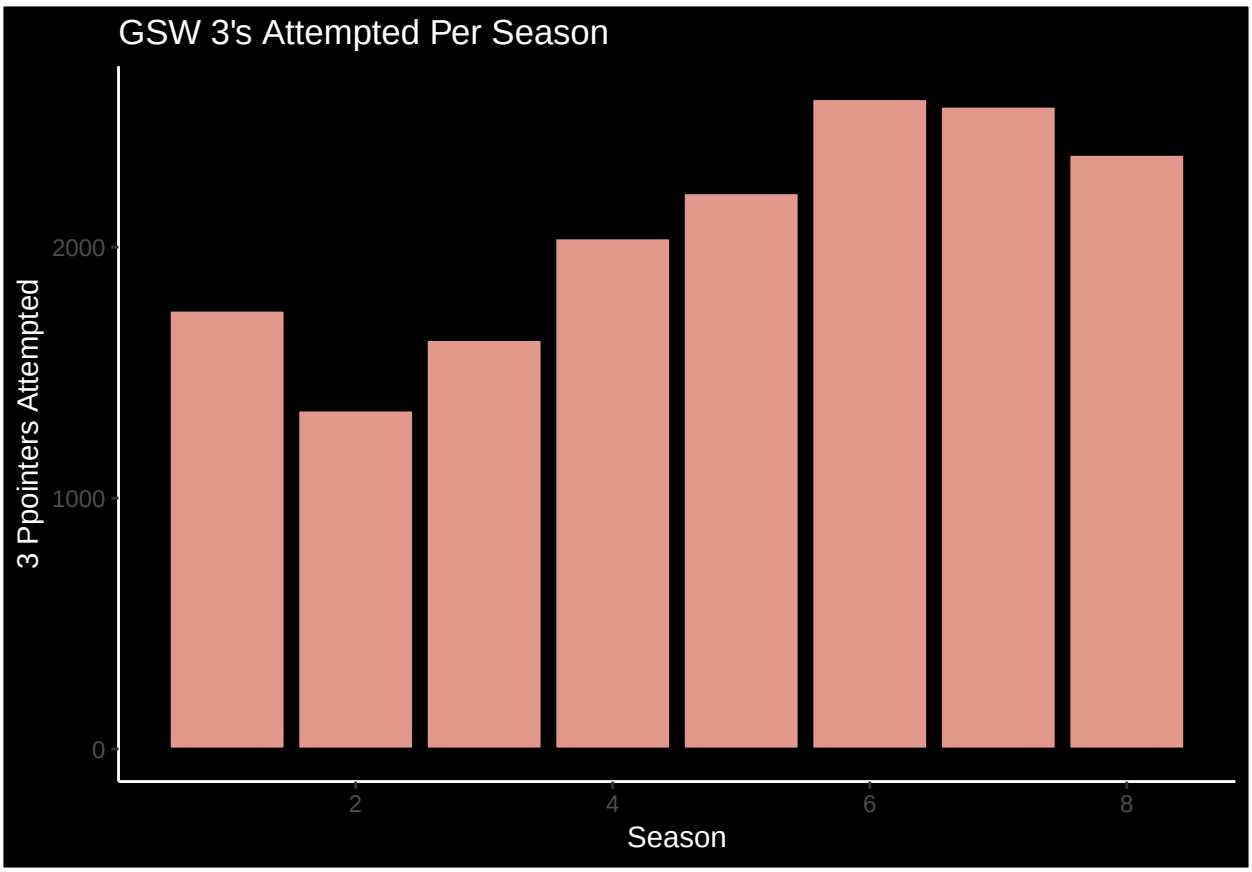
From the data collected from the GSW stats we were interested to see what factors separate the GSW from other teams, what makes them so much more efficient and better than others. To do so we broke it up the statistics so that we can notice a specific thing that GS is doing right that others can't do.

Part 1

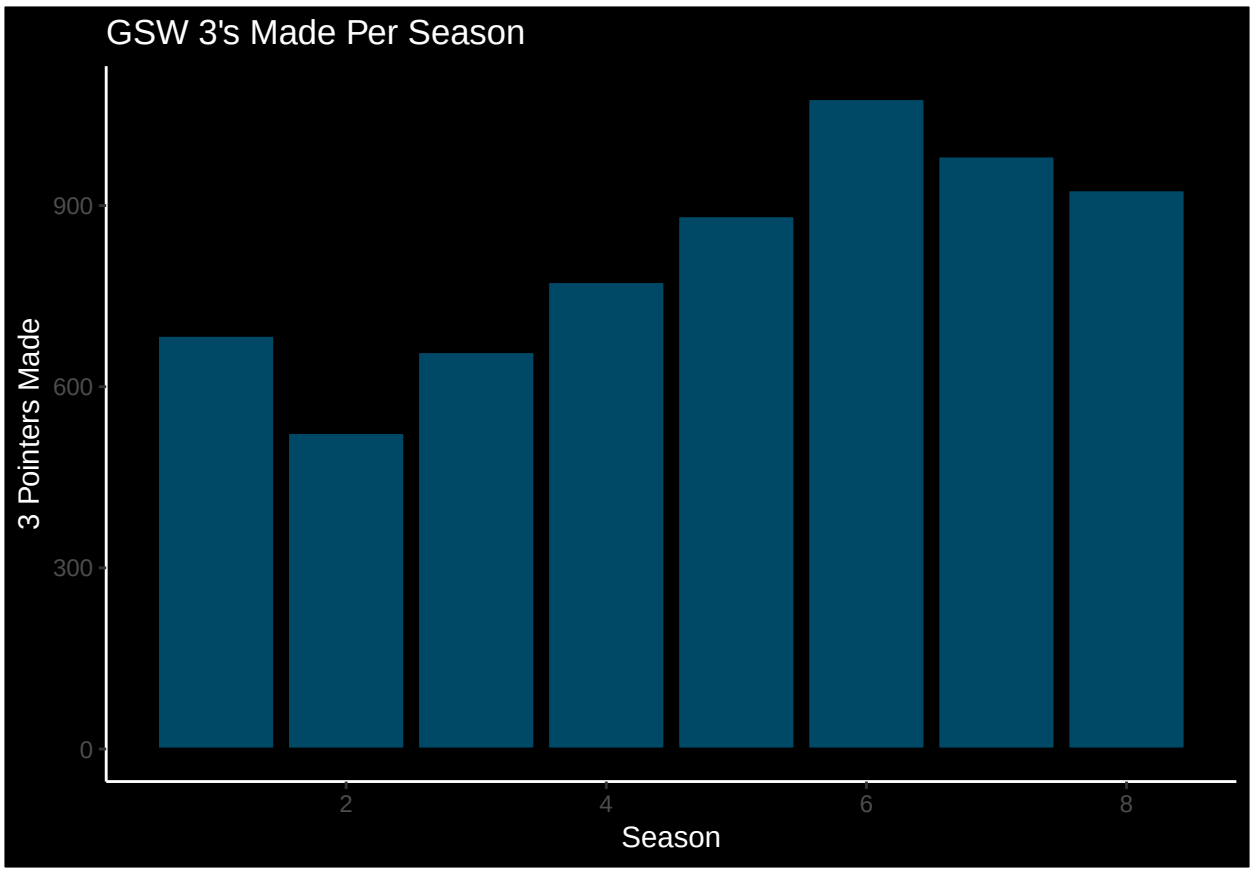
The average person knows that GS is known for shooting and making multiple 3 pointers. The starting five players on the warriors are all efficient when it comes to making field goals behind the arc. To our surprise, the data displays says that GS doesn't even take the most shots in the NBA. They are ranked 17 in the league for 3 pointers attempted in a season. They are ranked 17 but yet they have attempted 2,370 in the 2017-2018 season. GS isn't even ranked top 10 for 3 pointers attempted but yet they are number 1 for 3 point percentage. Their efficiency on the court for making 3 pointers has forced other teams to shoot more shots behind the arc to keep up, which often times results in their defeat. By putting more stress on other teams to catch up teams become desperate to make up the point difference. The team also lead the league with the highest average of points per game, 113.5. If you look back decades or even back in the 2000s, teams hardly

scored over 100 points per game. When it did happen it was because the team was on fire making on their shot. The team has transformed the league by forcing others to keep up with their high scoring, resulting in teams taking more 3s than GS themselves. GS overall is a high scoring machine, they also lead the league in 2 point percentages while once again not being the top 10 team in 2 pointers attempted. They were ranked 20 in the league, even lower in rank compared to 3 pointers, and yet they make more than anyone else. The effectiveness of the team is seen through their ball movement and how they have set up their offense to always have options.

```
row.names(GSW_stats11) <- c("team(11)", "team/G(11)", "L Rank(11)")
GSW_stats<- rbind(GSW_stats11,GSW_stats12,GSW_stats13,GSW_stats14,
                  GSW_stats15,GSW_stats16,GSW_stats17,GSW_stats18)
season<- data.frame(seasons=1:8)
threePA11<- GSW_stats11$X3PA[1]
threePA12<- GSW_stats12$X3PA[1]
threePA13<- GSW_stats13$X3PA[1]
threePA14<- GSW_stats14$X3PA[1]
threePA15<- GSW_stats15$X3PA[1]
threePA16<- GSW_stats16$X3PA[1]
threePA17<- GSW_stats17$X3PA[1]
threePA18<- GSW_stats18$X3PA[1]
threePA<- rbind(threePA11,threePA12,threePA13,threePA14,threePA15,threePA16,threePA17,threePA18)
threePA_season<- cbind(season,threePA)

threePA_graph<- ggplot(threePA_season,
                       aes(x=seasons, y=threePA_season$threePA,fill=seasons)) +
  geom_bar(stat = "identity",fill="#e2988b",color="black") +
  xlab("Season")+
  ylab("3 Ppointers Attempted")+
  ggtitle("GSW 3's Attempted Per Season") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))
threePA_graph
```




```
threeP11<- GSW_stats11$X3P[1]
threeP12<- GSW_stats12$X3P[1]
threeP13<- GSW_stats13$X3P[1]
threeP14<- GSW_stats14$X3P[1]
threeP15<- GSW_stats15$X3P[1]
threeP16<- GSW_stats16$X3P[1]
threeP17<- GSW_stats17$X3P[1]
threeP18<- GSW_stats18$X3P[1]
threeP<- rbind(threeP11,threeP12,threeP13,threeP14,threeP15,threeP16,threeP17,threeP18)
threeP_season<- cbind(season, threeP)

threeP_graph<- ggplot(threeP_season, aes(x= threeP_season$seasons, y=threeP_season$threeP))+
  geom_bar(stat = "identity", color = "black",fill="#004966") +
  xlab("Season") +
  ylab("3 Pointers Made") +
  ggtitle("GSW 3's Made Per Season") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))
threeP_graph
```



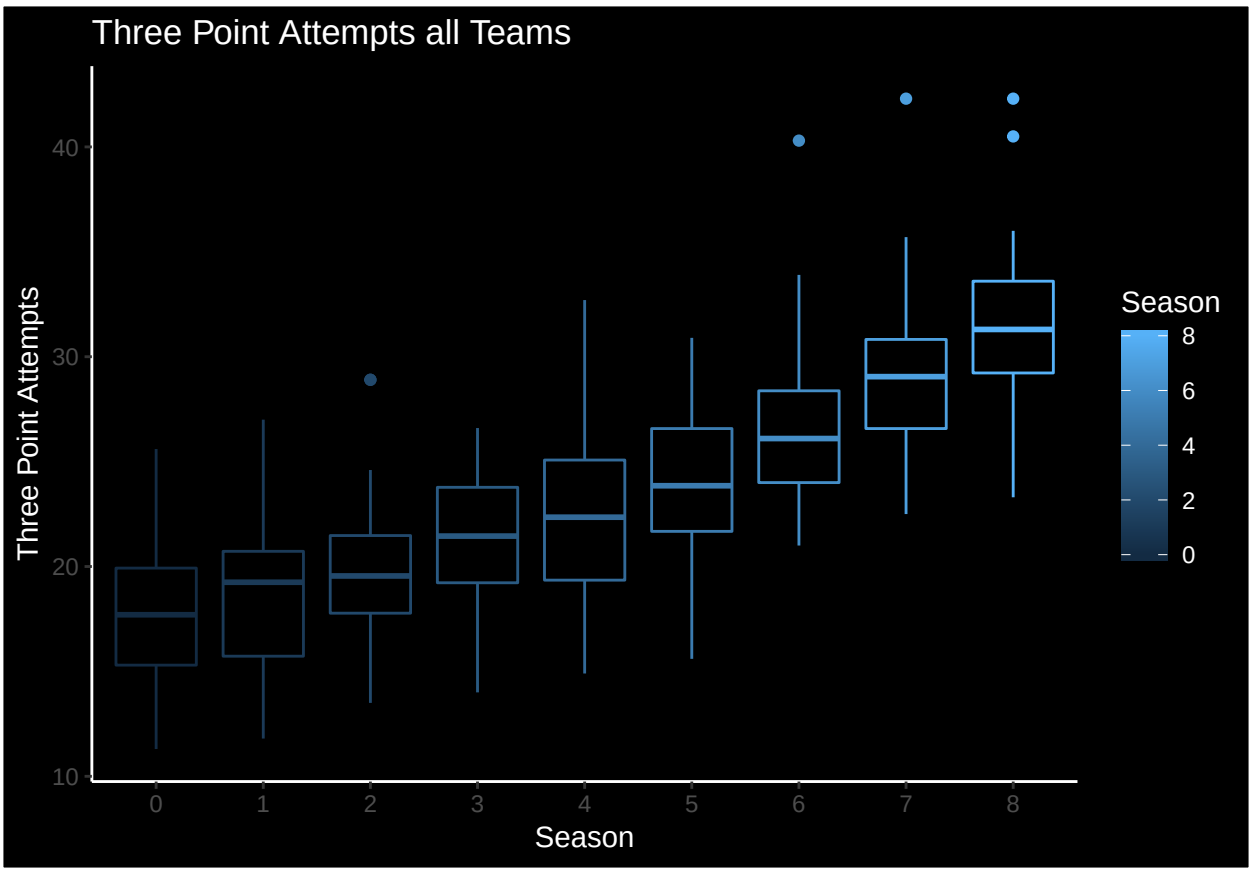
From what is displayed by the graphs you can see nearly see a correlation between the amount of 3s attempted and the percentage of 3s made to the success of the team. For the past 8 seasons you can see that the Warriors have adopted the three point shot to be their main attribute as a team. Out of the 8 seasons the warriors have lead the league in 3 point percentage four times, being top 5 in the league all 8 years. They have only lead 3 point attempts once in the past 8 seasons and that season they broke the record for the amount of wins in a season. You could say practice makes perfect because from 2010-2013 the team never broke 2000 3 pointers in a season. It was after the 2012-2013 that the team would have stay about 2000 for the next 5-6 years. Out of those 5-6 years the team won 3 championships and have made it to playoffs every single year since.

Part 2

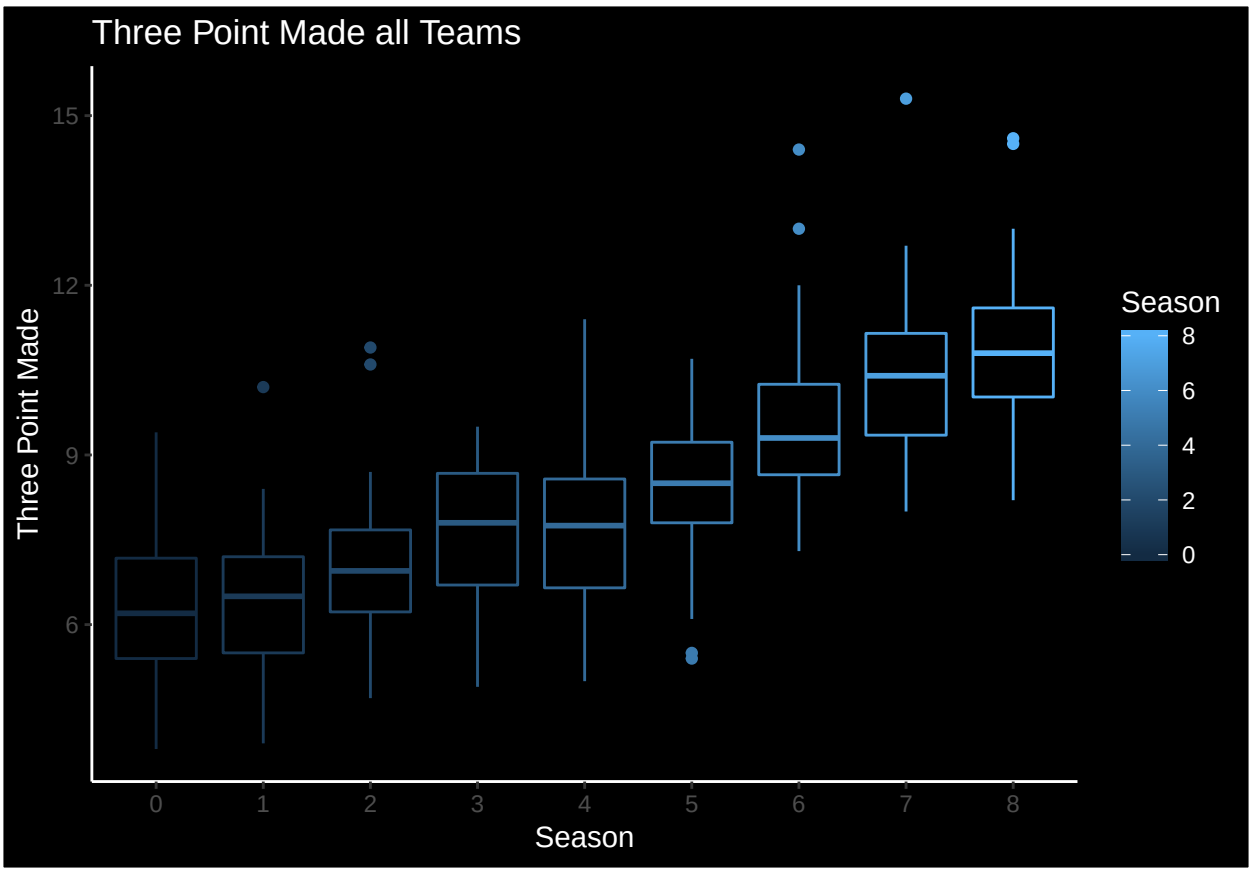
After analyzing Golden State's trends during the past nine seasons, we wanted to see how the rest of the league matched Golden State's trends. We decided to look at the 3 point attempts and makes and 2 point attempts and makes of all other teams besides Golden State for the past nine seasons.

```
#All stats for seasons 2010-11 to 2018-19
all <- rbind(twenty_ten_eleven_team_stats,
             twenty_eleven_twelve_team_stats,
             twenty_twelve_thirteen_team_stats,
             twenty_thirteen_fourteen_team_stats,
             twenty_fourteen_fifteen_team_stats,
             twenty_fifteen_sixteen_team_stats,
             twenty_sixteen_seventeen_team_stats,
             twenty_seventeen_eighteen_team_stats,
             twenty_eighteen_nineteen_team_stats)
```

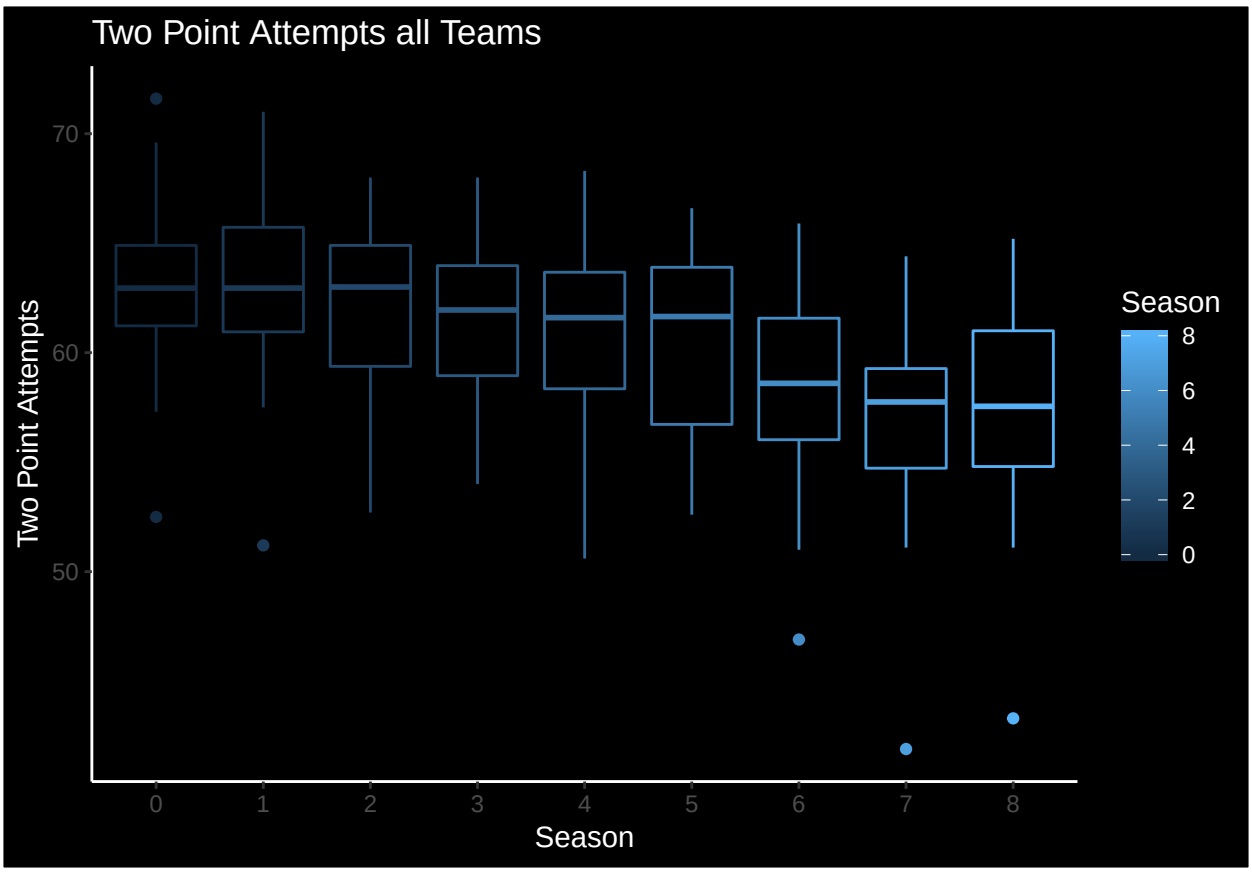
```r
#Want to check rest of the league without Golden State
all <- filter(all,Team!="Golden State Warriors"&Team!="Golden State Warriors*")
#Some plots
three_att_box <- ggplot(all,aes(x=as.factor(Season),y=X3PA,color=Season)) +
  geom_boxplot(fill="black") +
  xlab("Season") +
  ylab("Three Point Attempts") +
  ggtitle("Three Point Attempts all Teams") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

three_made_box <- ggplot(all,aes(x=as.factor(Season),y=X3P,color=Season)) +
  geom_boxplot(fill="black") +
  xlab("Season") +
  ylab("Three Point Made") +
  ggtitle("Three Point Made all Teams") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

two_att_box <- ggplot(all,aes(x=as.factor(Season),y=X2PA,color=Season)) +
  geom_boxplot(fill="black") +
  xlab("Season") +
  ylab("Two Point Attempts") +
  ggtitle("Two Point Attempts all Teams") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

two_made_box <- ggplot(all,aes(x=as.factor(Season),y=X2P,color=Season)) +
  geom_boxplot(fill="black") +
  xlab("Season") +
  ylab("Two Point Made") +
  ggtitle("Two Point Made all Teams") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

three_att_box
```



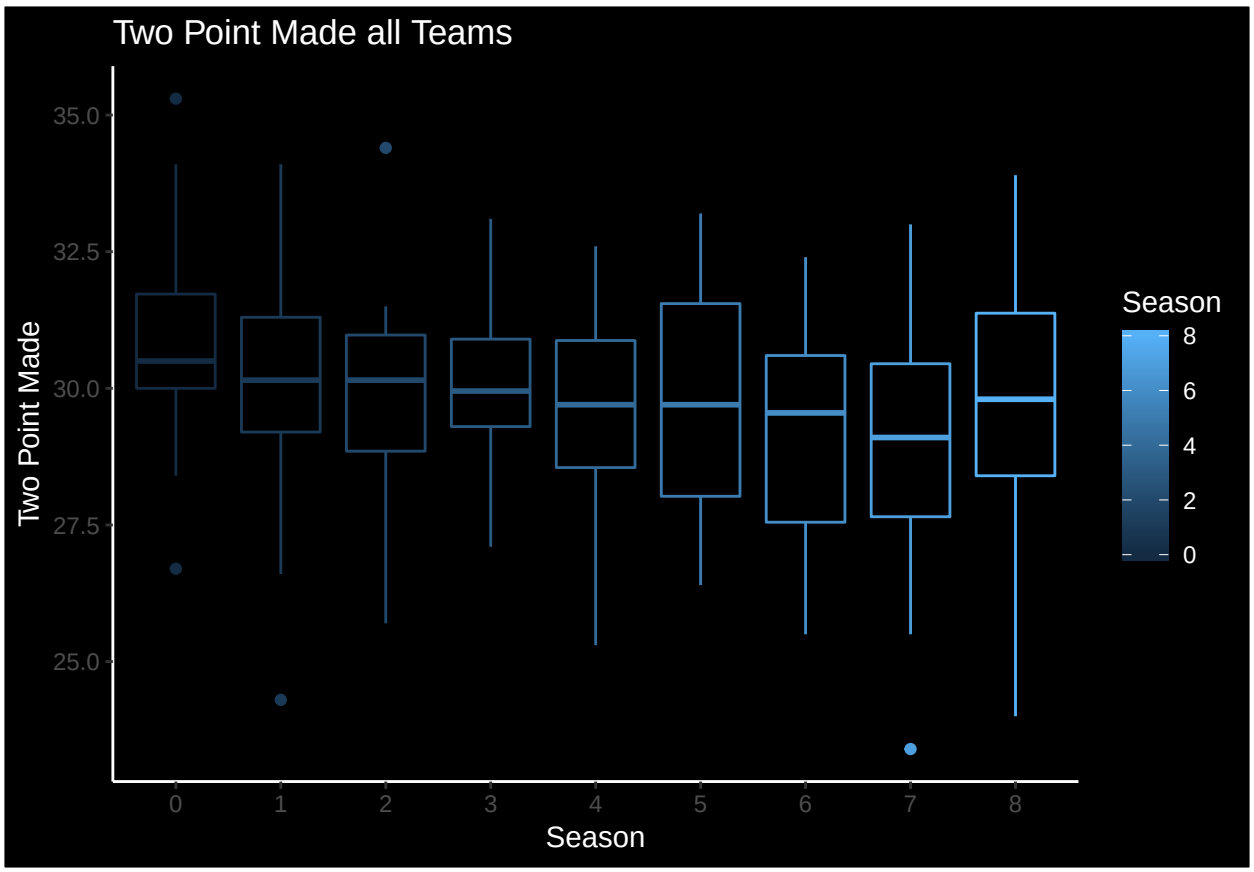

```
three_made_box
```


Three Point Made all Teams
Three Point Made
15
12
9
6
0
1
2
3
4
5
6
7
8
Season
Season
8
6
4
2
0

```
two_att_box
```


Two Point Attempts all Teams
Two Point Attempts
70
60
50
0
1
2
3
4
5
6
7
8
Season
Season
8
6
4
2
0

```
two_made_box
```



From these boxplots, we can see that the average 3 point attempts and makes has increased steadily even before Golden State's exploitation of the shot. However, after the 2014-15 season, we can see multiple outliers above the mean. For instance, during the 2016-17 season, one team was attempting nearly 40 3 point shots per game, about 15 more than the average. The next two season have outliers nearing the 45 attempts mark. Clearly, certain teams have tried to emphasize the 3 point shot in their gameplan, and the high number of 3 point shots made speaks to the efficacy of this shift. The number of 2 point attempts has slightly decreased over the past nine seasons, with many low outliers, some being almost 20 attempts below averages. These plots may show that a general trend in the league is occurring, with a few teams drastically altering their game plan in an attempt to emulate the success of the Golden State Warriors.

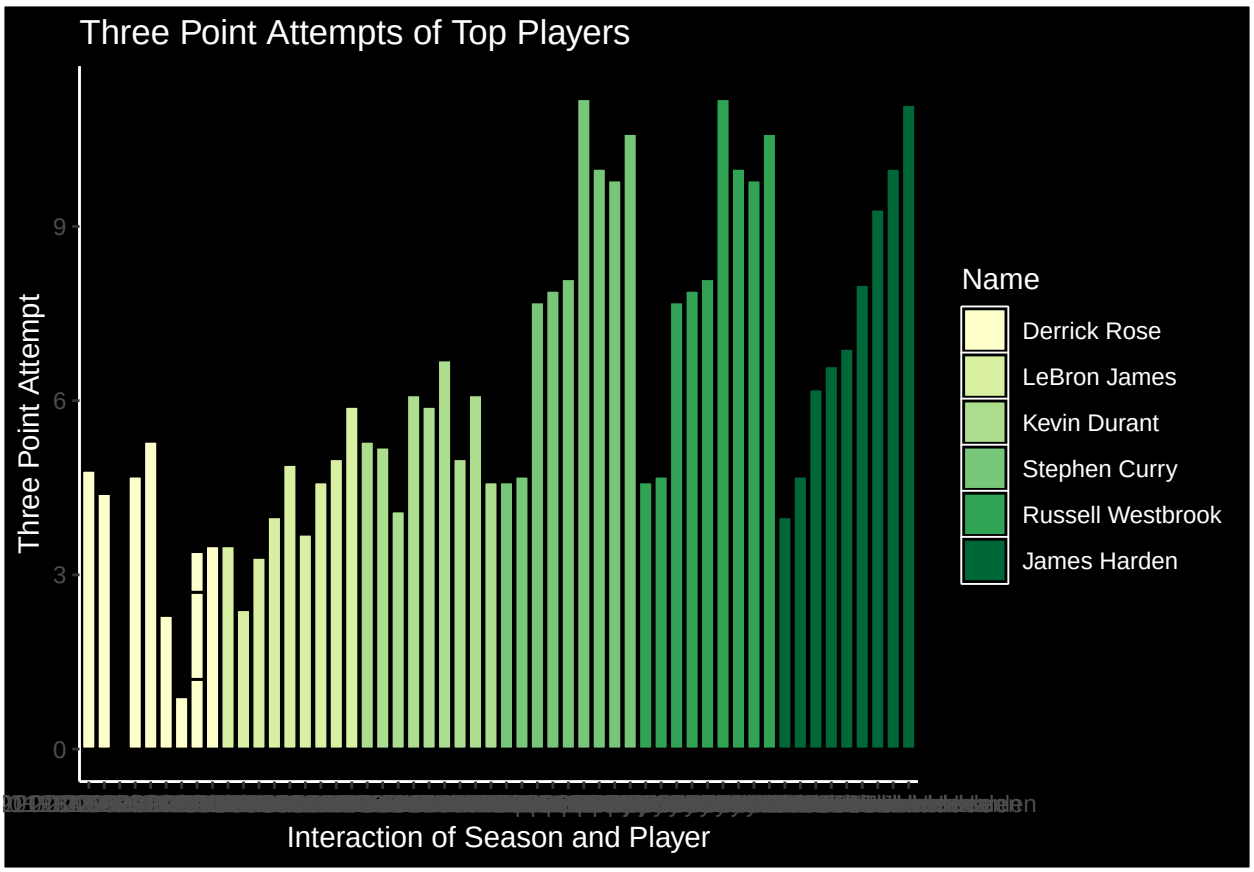
Part 3

Though team trends are important to observe, teams are influenced by the players on them, and so it is necessary for us to analyze how individual players have changed their game as well. We chose all the players that had either won MVP, were the scoring champion or led the league in win shares in the past nine seasons. Win shares is a complex statistic that essentially rates how many wins on a team can be attributed to a particular player on that team, it is a measure of how important that player was to earning wins for the team. There were only six players that shared all these different awards over the past seasons, LeBron James, James Harden, Derrick Rose, Kevin Durant, Stephen Curry, and Russell Westbrook.

```
#Using league MVP's, Pt champions, and WS champions

#Adding name columns
d_rose <- cbind(d_rose,"Name" = rep("Derrick Rose"))
lbj <- cbind(lbj,"Name" = rep("LeBron James"))
kd <- cbind(kd,"Name" = rep("Kevin Durant"))
```

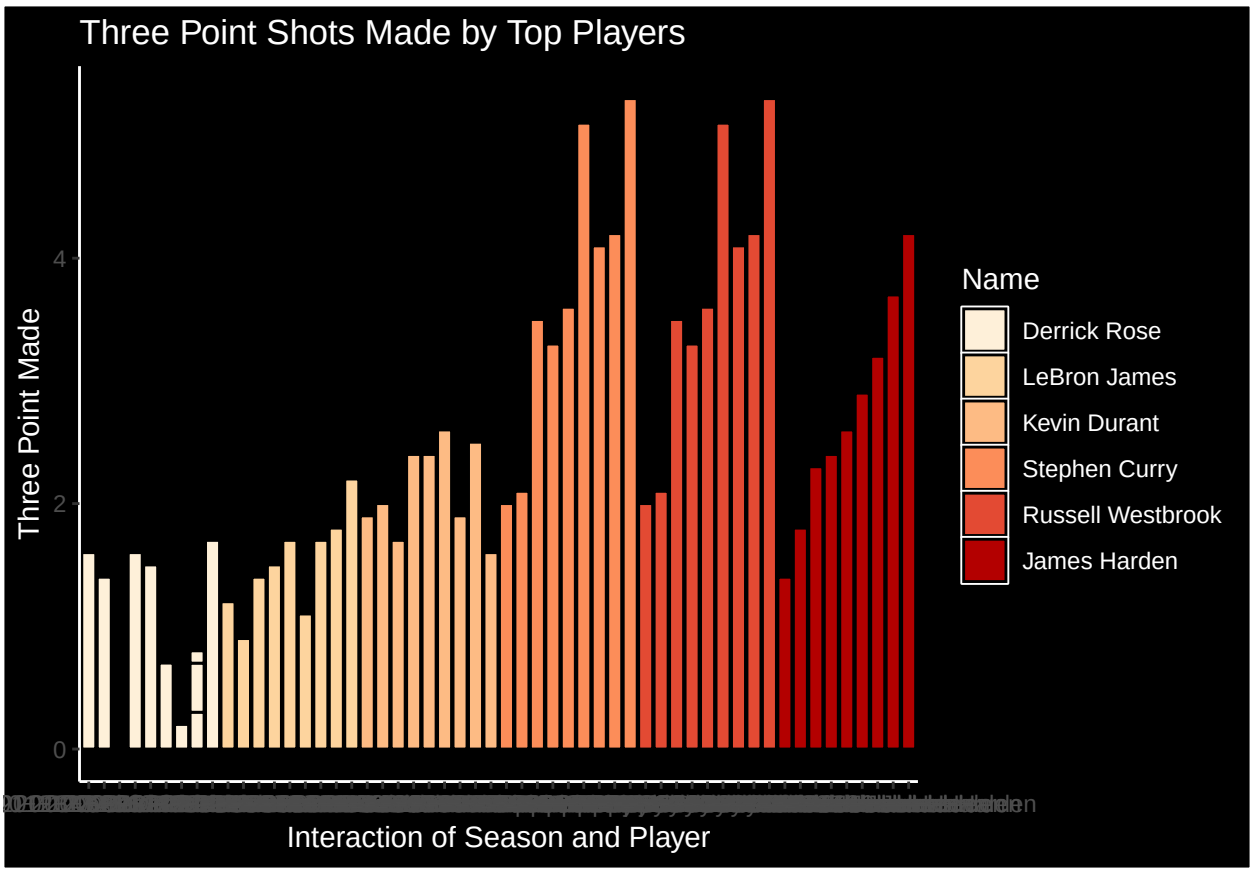
```r
sc <- cbind(sc,"Name" = rep("Stephen Curry"))
russ_wb <- cbind(russ_wb,"Name" = rep("Russell Westbrook"))
james_h <- cbind(james_h,"Name" = rep("James Harden"))
#Combining all top players statistics
top_players <- rbind(d_rose,lbj,kd,sc,russ_wb,james_h)

#Plots
top_players_three_att <- ggplot(top_players,
                                aes(x=interaction(Season,Name),
                                    y=X3PA,fill=Name)) +
  geom_bar(stat = "identity",color = "black") +
  xlab("Interaction of Season and Player") +
  ylab("Three Point Attempt") +
  scale_fill_brewer(palette = "YlGn") +
  ggtitle("Three Point Attempts of Top Players") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

top_players_three_made <- ggplot(top_players,
                                aes(x=interaction(Season,Name),
                                    y=X3P,fill=Name)) +
  geom_bar(stat = "identity",color = "black") +
  xlab("Interaction of Season and Player") +
  ylab("Three Point Made") +
  scale_fill_brewer(palette = "OrRd") +
  ggtitle("Three Point Shots Made by Top Players") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

top_players_three_att
```

```
## Warning: Removed 1 rows containing missing values (position_stack).
```




```
top_players_three_made
```

```
## Warning: Removed 1 rows containing missing values (position_stack).
```



It is important to keep in mind exactly what we are trying to look for in these plots. We want to see if players who are performing at the highest possible level are utilizing the three point shot more during games. We can see that three players truly stand out from the rest, those being Stephen Curry, Russell Westbrook, and James Harden. These three players have all attempted and made more 3 point field goals than their peers on the plot, and all three of these players have been either MVP, scoring champion, or leader in win shares during the past three seasons. On the opposite extreme, we see a player like Derrick Rose who, after getting injured and losing some of his athleticism and playing ability, fell from utilizing the 3 point shot. But, during his past two seasons with Minnesota, he has gotten to a point where this season, he is making the most 3 point shots per game in his entire career, and is considered to be having an NBA comeback. LeBron James has also increased his 3 point attempts and makes per game, and is considered the best player in the NBA today.

We also wanted to examine if the roles of certain positions changed with league wide trends. We wanted to focus on the center, as the center is typically a position played inside the key and doesn't traditionally come out past the three point line looking for a shot. We wanted to see whether big men are also joining this trend and compare that to the 3 point attempts and makes of point guards, a position which traditionally does take three point shots. We decided to only include players who played at least 50 games a season and 25 minutes per game, otherwise players that have little minutes played could skew the data.

```
season_stats_post_2010 <- subset(season_stats,Year>2009)

#Positions
#Have to play at least 50 G
pgs <- filter(season_stats_post_2010,Pos=="PG"&Tm!="TOT"&G>50&MP>25)

PG2010 <- summarise_all(filter(pgs,Year==2010),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
```

```
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2011 <- summarise_all(filter(pgs,Year==2011),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2012 <- summarise_all(filter(pgs,Year==2012),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2013 <- summarise_all(filter(pgs,Year==2013),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2014 <- summarise_all(filter(pgs,Year==2014),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2015 <- summarise_all(filter(pgs,Year==2015),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2016 <- summarise_all(filter(pgs,Year==2016),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
PG2017 <- summarise_all(filter(pgs,Year==2017),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
pgs_avgs <- rbind(PG2010,PG2011,PG2012,PG2013,PG2014,PG2015,PG2016,PG2017)
pgs_avgs$Pos <- rep("PG")

centers <- filter(season_stats_post_2010,Pos=="C"&Tm!="TOT"&G>50&MP>25)

C2010 <- summarise_all(filter(centers,Year==2010),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2011 <- summarise_all(filter(centers,Year==2011),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2012 <- summarise_all(filter(centers,Year==2012),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA
```

```
## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA
```

```
## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2013 <- summarise_all(filter(centers,Year==2013),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA

## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA

## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2014 <- summarise_all(filter(centers,Year==2014),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA

## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA

## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2015 <- summarise_all(filter(centers,Year==2015),funs(mean))
```

```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA

## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA

## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2016 <- summarise_all(filter(centers,Year==2016),funs(mean))
```

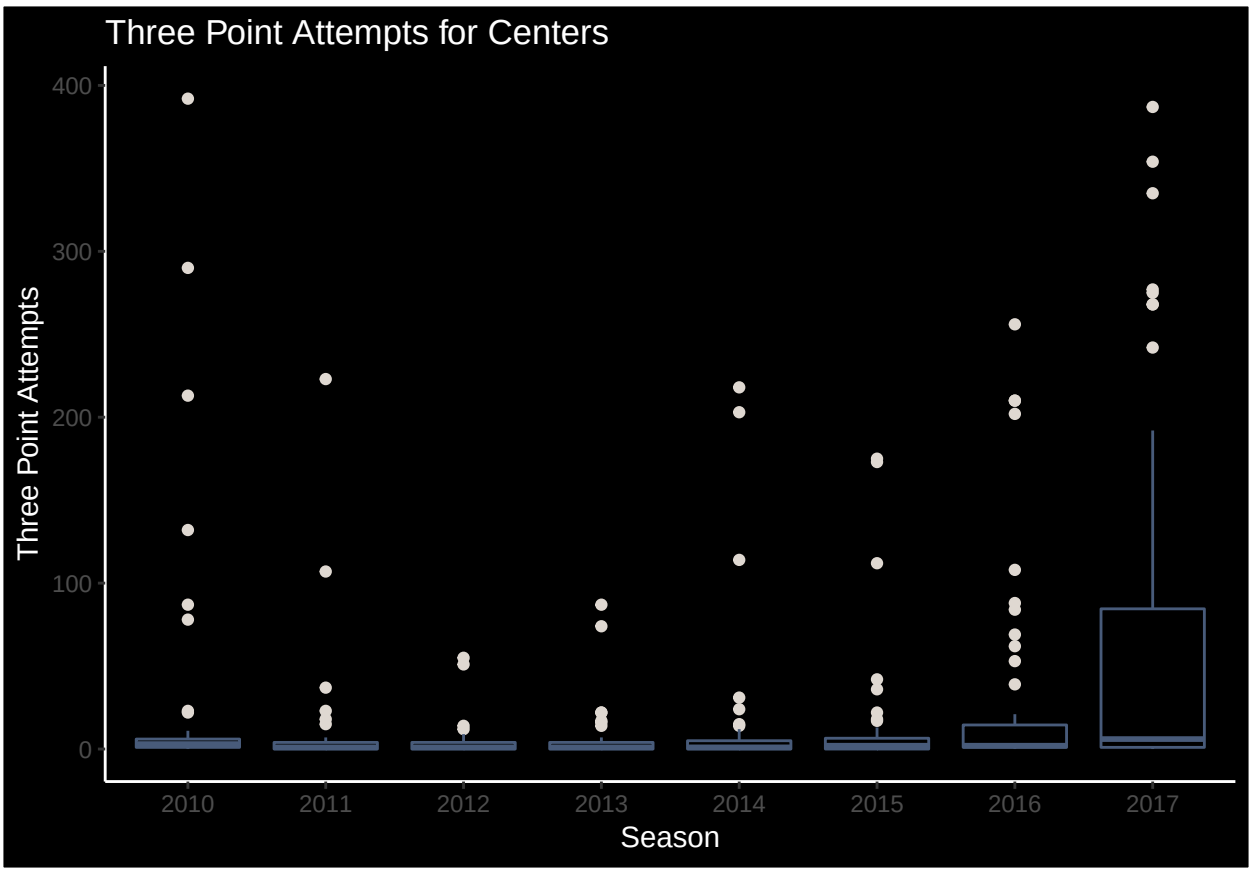
```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA

## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA

## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
C2017 <- summarise_all(filter(centers,Year==2017),funs(mean))
```

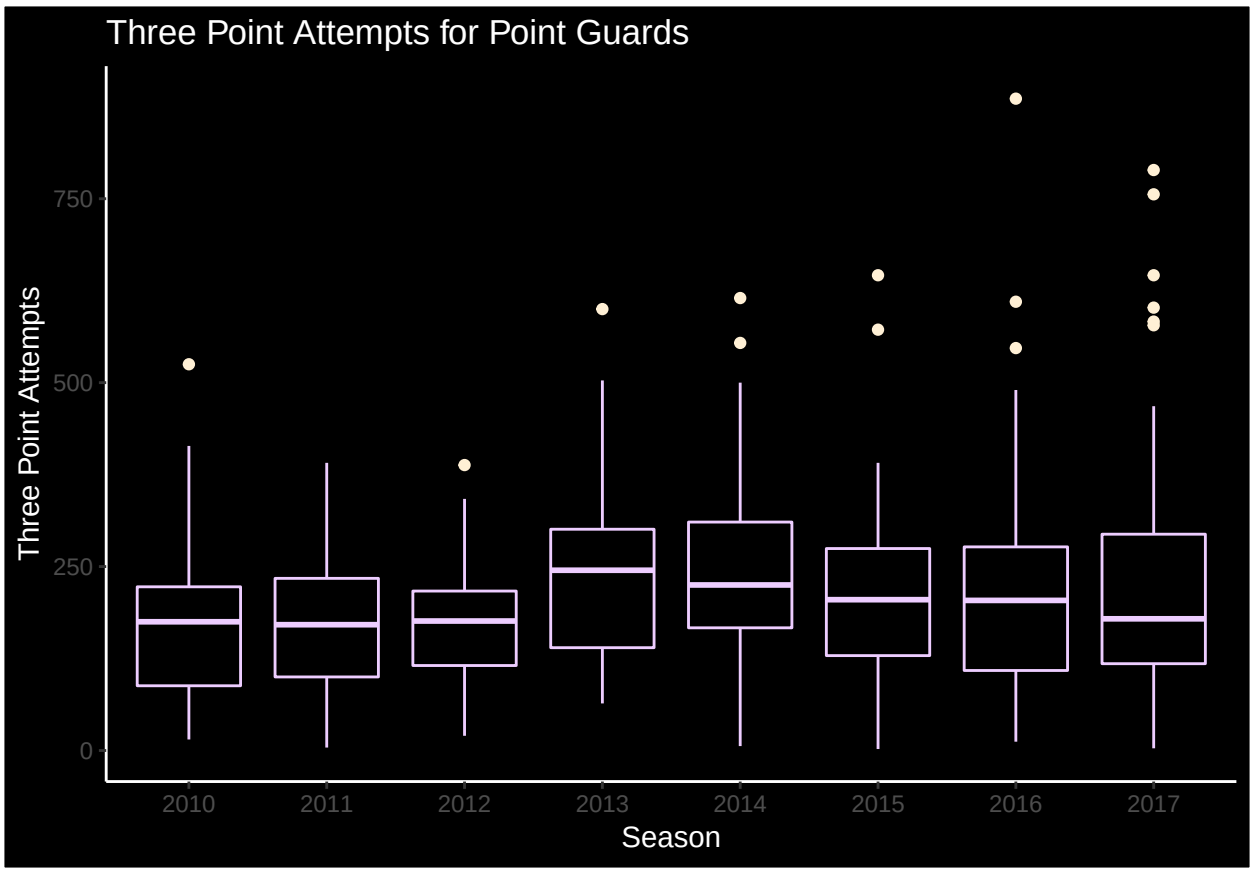
```
## Warning in mean.default(Player): argument is not numeric or logical:
## returning NA

## Warning in mean.default(Pos): argument is not numeric or logical: returning
## NA

## Warning in mean.default(Tm): argument is not numeric or logical: returning
## NA
```

```
center_avgs <- rbind(C2010,C2011,C2012,C2013,C2014,C2015,C2016,C2017)
center_avgs$Pos <- rep("C")

avg_pos <- rbind(center_avgs,pgs_avgs)


#Plots
centers_3PA <- ggplot(centers,aes(x=as.factor(Year),y=X3PA)) +
  geom_boxplot(color="#485b7a",fill="black",outlier.color = "#dfd8d1") +
```

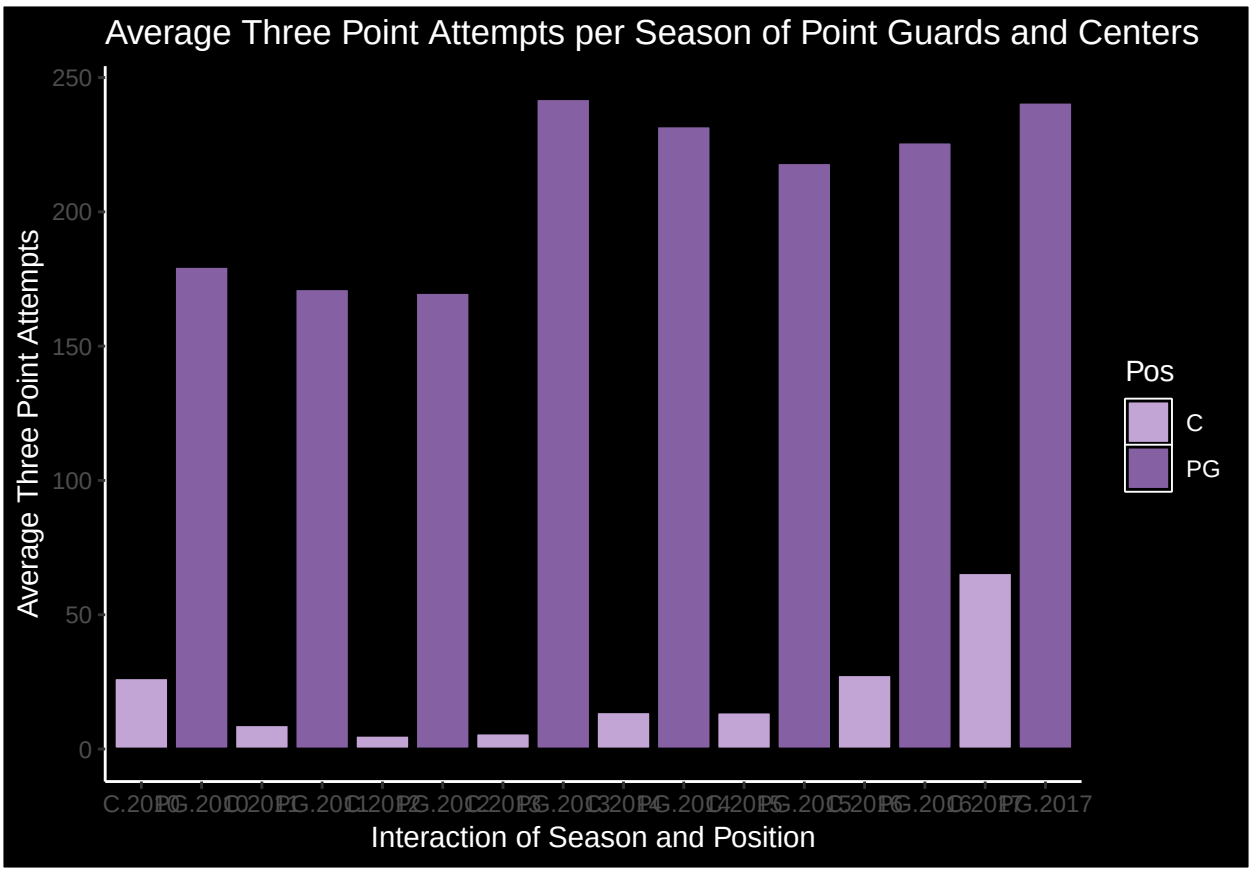
```
  ylab("Three Point Attempts") +
  xlab("Season") +
  ggtitle("Three Point Attempts for Centers") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

pgs_3PA <- ggplot(pgs,aes(x=as.factor(Year),y=X3PA)) +
  geom_boxplot(color="#edcdff",fill="black",outlier.color = "#fff0d5") +
  ylab("Three Point Attempts") +
  xlab("Season") +
  ggtitle("Three Point Attempts for Point Guards") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

pos_mean_three_att <- ggplot(avg_pos,aes(x=interaction(Pos,Year),y=X3PA,fill=Pos)) +
  geom_bar(stat = "identity",color="black") +
  xlab("Interaction of Season and Position") +
  ylab("Average Three Point Attempts") +
  scale_fill_manual(values = c("#c3a5d5","#8560a2")) +
  ggtitle("Average Three Point Attempts per Season of Point Guards and Centers") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "white"),
        plot.background = element_rect(fill = "black"),
        legend.background = element_rect(fill = "black"),
        legend.text = element_text(color = "white"),
        title = element_text(color = "white"))

centers_3PA
```


Three Point Attempts for Centers
Three Point Attempts
400
300
200
100
0
2010
2011
2012
2013
2014
2015
2016
2017
Season

```
pgs_3PA
```




```
pos_mean_three_att
```



These plots offer some interesting conclusions about positions in the nba. The bar plot would make it seem that centers are taking more 3 point attempts than they have in the past, but not at that high of a rate, as point guards are still taking well over 4 times the amount that centers are, even this most recent season. However, the box plots show a similar trend among centers as in team trends. Though the mean amount of 3 point attempts is still very low for centers, there are many outliers that show individual players taking steps toward a more three point focused strategy. There are many centers that are attempting more 3 point shots than the average for point guards over many seasons. The amount of centers doing this has increased over the years as has the amount of 3 point shots those players are taking. Thought point guards are still making 3 point shots and attempting them at a high rate, centers could begin to challenge the abilites of smaller players in the coming seasons.

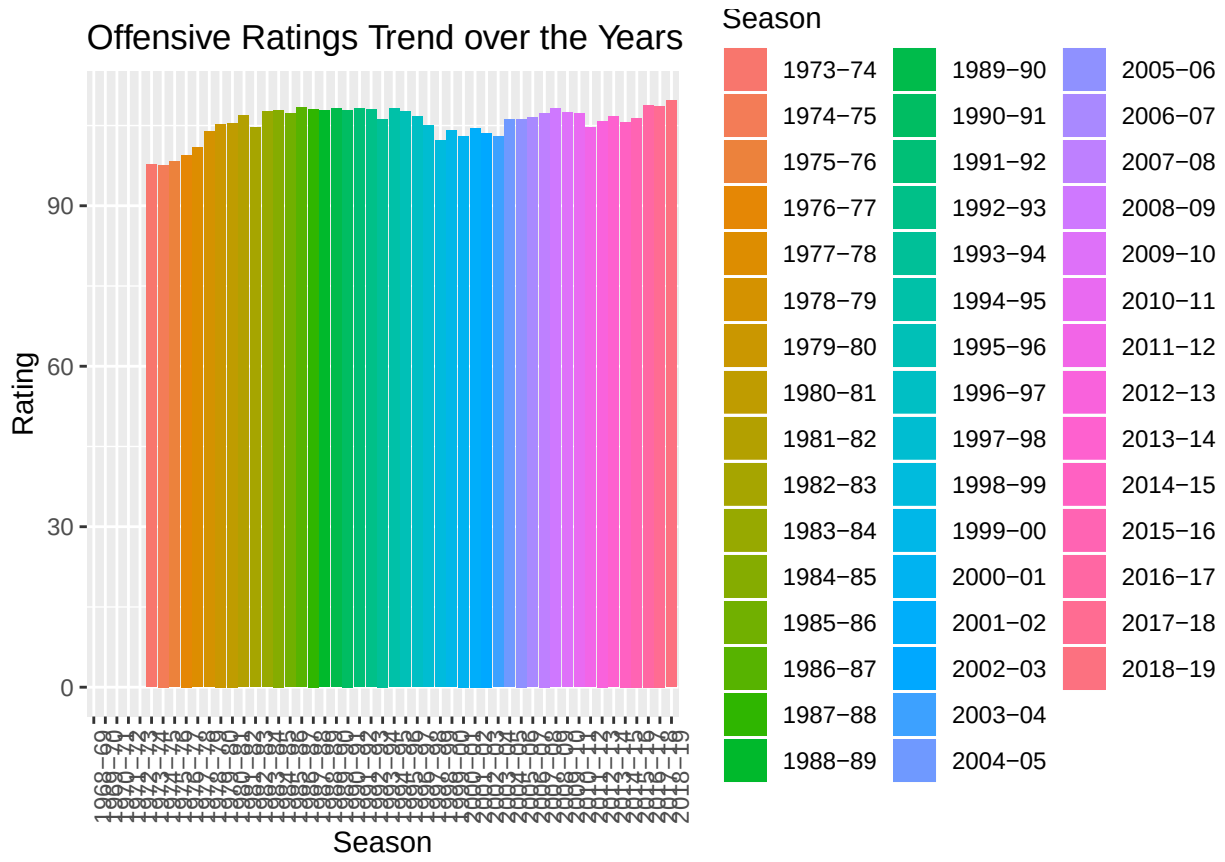
Part 4

The 3 point shot started in the NBA in 1979, so its impact can be seen and measured based on what happened before and after 1979. The Golden State warriors became an elite 3-point shooting team in 2014 and won three NBA championships since 2014. We will examine the impact of the 3-point shot and the impact of the rise of the Golden State Warriors as an elite 3-point shooting team.

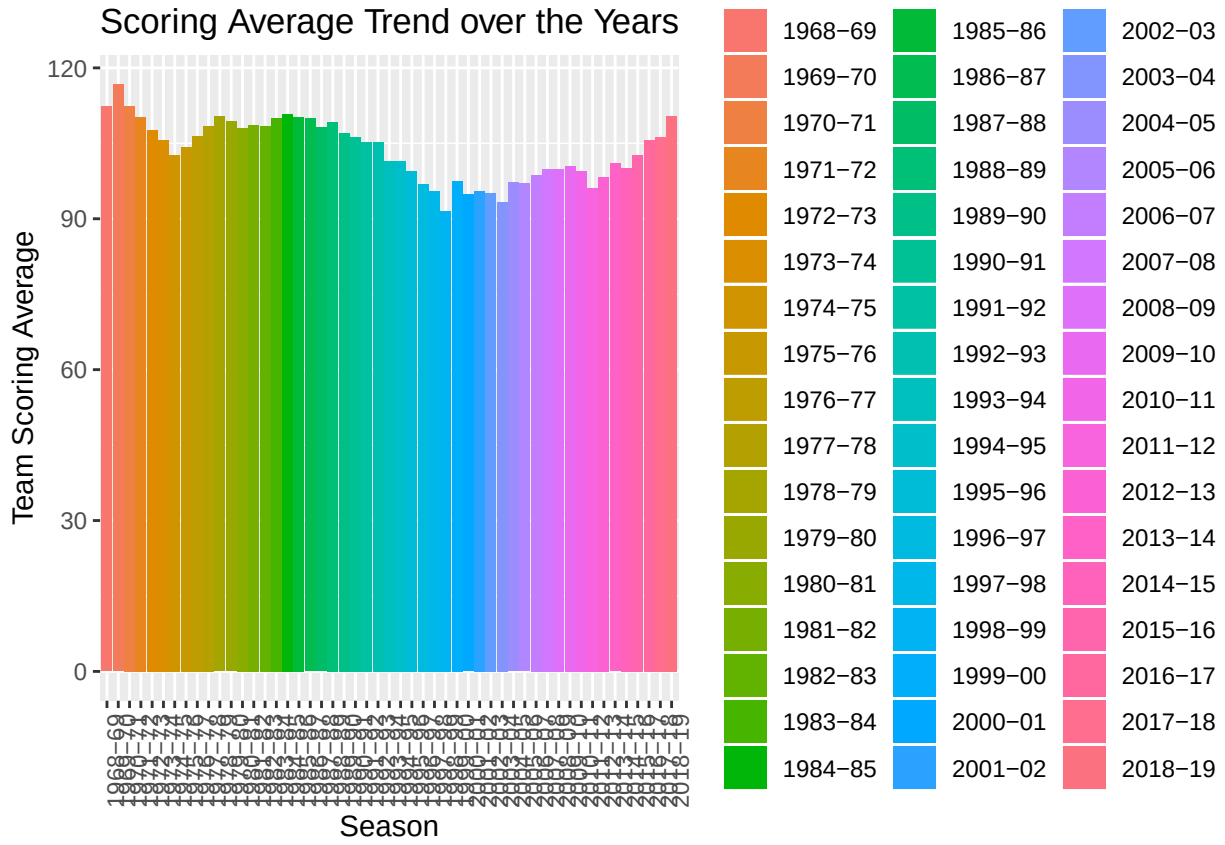
```
nbaStatsTibble <- as_tibble(nbaStats)
nbaRatingsTibble <- as_tibble(nbaRatings)
# offensive efficiency by years
offensiveRatings <- nbaRatingsTibble %>% group_by(Season) %>% summarize(Rating=mean(ORtg))
# impact the 3-point line had on scoring trends since
# the Golden State Warriors
# became an elite 3-point shooting team
```

```
# and won NBA champships.
scoringAvg <- nbaStatsTibble %>% group_by(Season) %>% summarize(Avg=mean(PTS / G))
# how the 3-point shot changed shot
# selection over the years
threePointPct <- nbaStatsTibble %>% group_by(Season) %>%
  summarize(threePT=mean((X3PA / FGA) * 100))
# graph of the offensive ratings, scoring average
# and two vs three point shooting averages over the
# years to visualize the data. Keep in mind the
# 3-point line started in 1979 and the Golden
# State Warriors elite 3-point shooting team
# happened in 2014
# offensive ratings plot
p <- ggplot(offensiveRatings, aes(y=Rating,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("Offensive Ratings Trend over the Years") +
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1))
```

```
## Warning: Removed 5 rows containing missing values (position_stack).
```

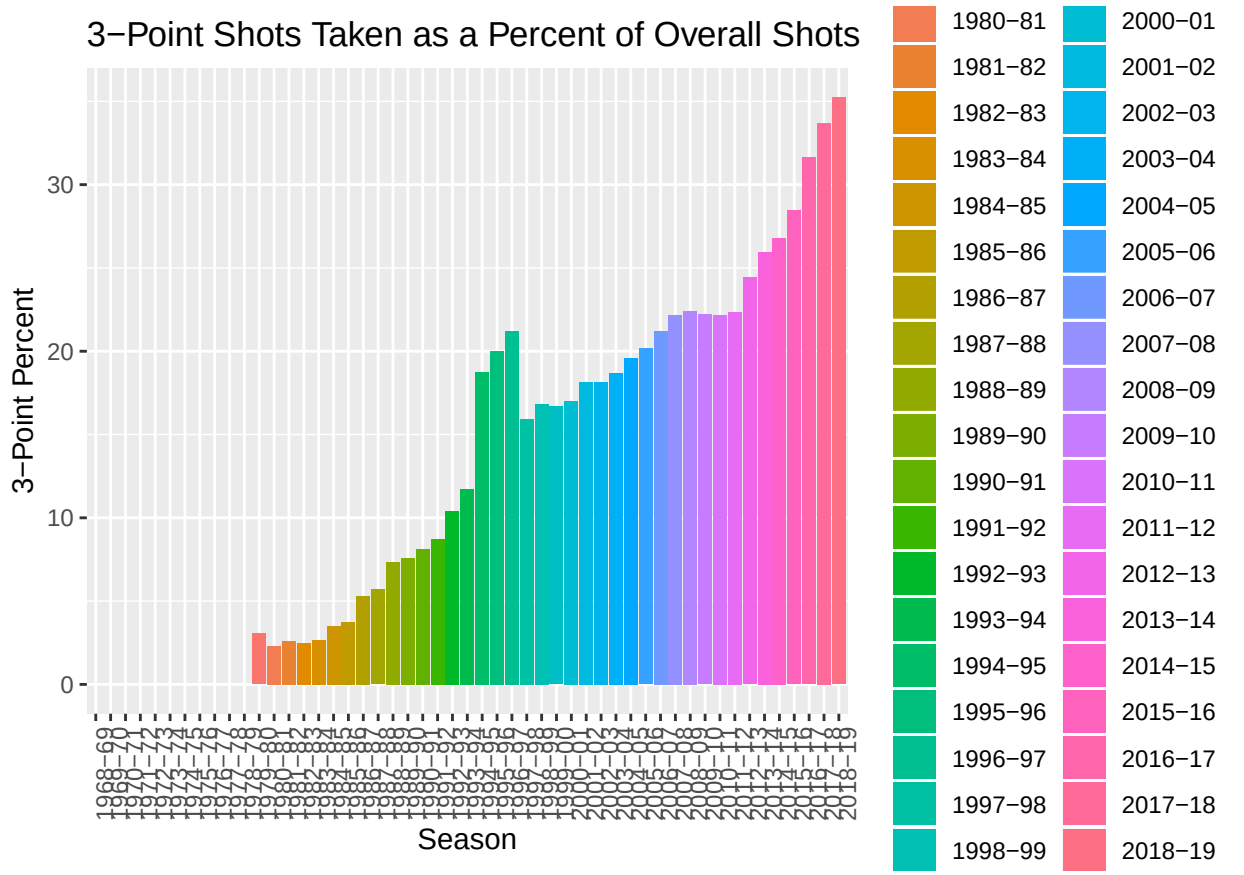



```
# scoring average
p <- ggplot(scoringAvg, aes(y=Avg,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("Scoring Average Trend over the Years") +
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1)) +
  ylab("Team Scoring Average")
```




```
# percentage growth of 3-point shots
p <- ggplot(threePointPct, aes(y=threePT,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("3-Point Shots Taken as a Percent of Overall Shots ") +
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1)) +
  ylab("3-Point Percent")
```

```
## Warning: Removed 11 rows containing missing values (position_stack).
```



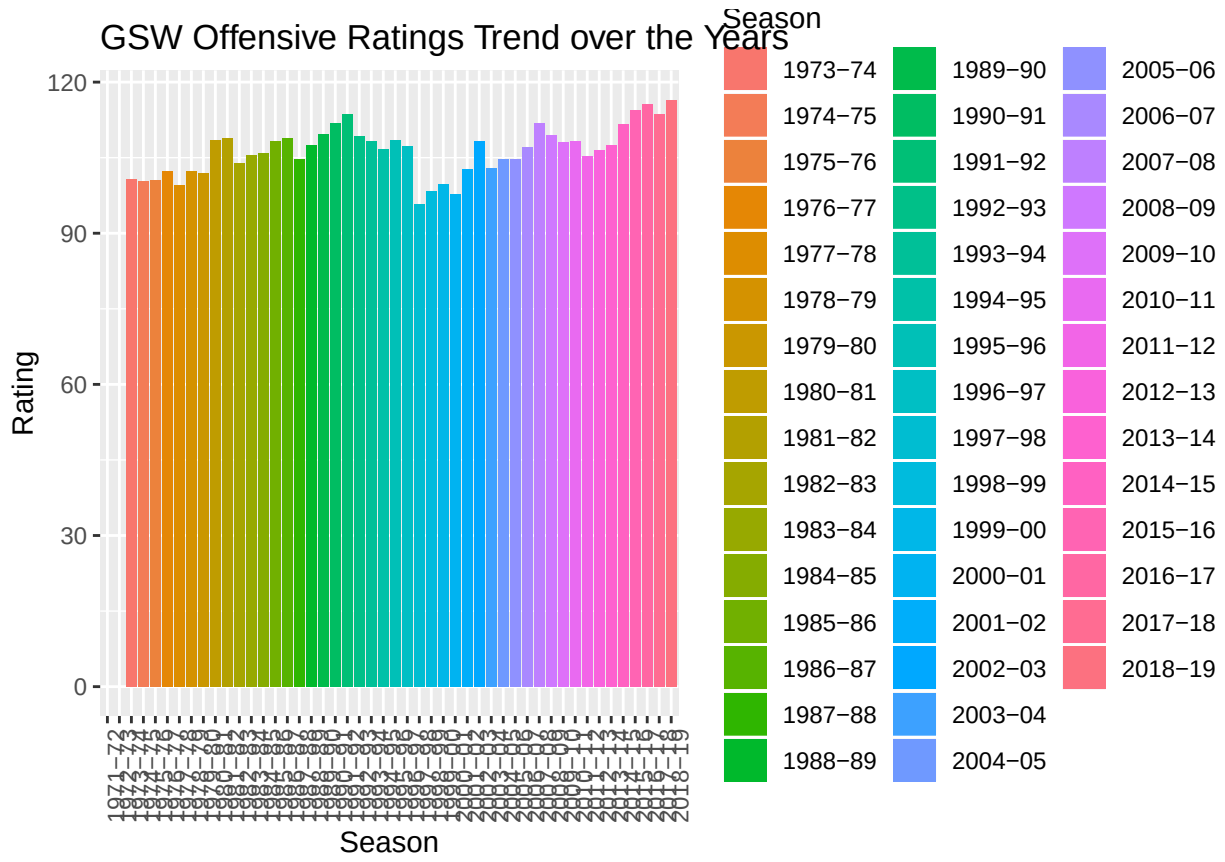
When the 3-point shot was introduced in 1979, we see an immediate jump in offensive ratings between 1981 - 1983. After 1983 the ratings fluctate a bit but remain higher compared to the time of only 2-point shots peaking 1991-92. There is a significant drop in 1996-97 to the lowset level ever then a slow rise. Since the Warriors became an elite 3-point team in 2014 winning three NBA titles the offensive ratings has been on the rise reaching the highest levels ever in the NBA. Team scoring averages follow a similar trend, but the highest average was back in 1991-92 when offensive ratings were at an all time high. In 1991-92 3-point shooting was at an all time high corresponding to the historic scoring and offensive ratings 3-point shots taken percentage continued to rise with dips in the early 2000's and early 2010's. We see a noticeable and meaningful jump in 2014 reaching an all time high in 2016-17 as the Warriors became an elite 3 point shooting team winnig three championships. We can see teams attempting to emulate their success driving all league stats higher. The 3-point line has made a large impact on the NBA driving all offensive stats higher overall and teams like the Warriors have had a lot of success winning multiple championships because of the 3 point shot.

The Golden State Warriors became an elite NBA team around the same time they became the most prolific NBA 3 point shooting team winning 3 NBA championships since 2014.

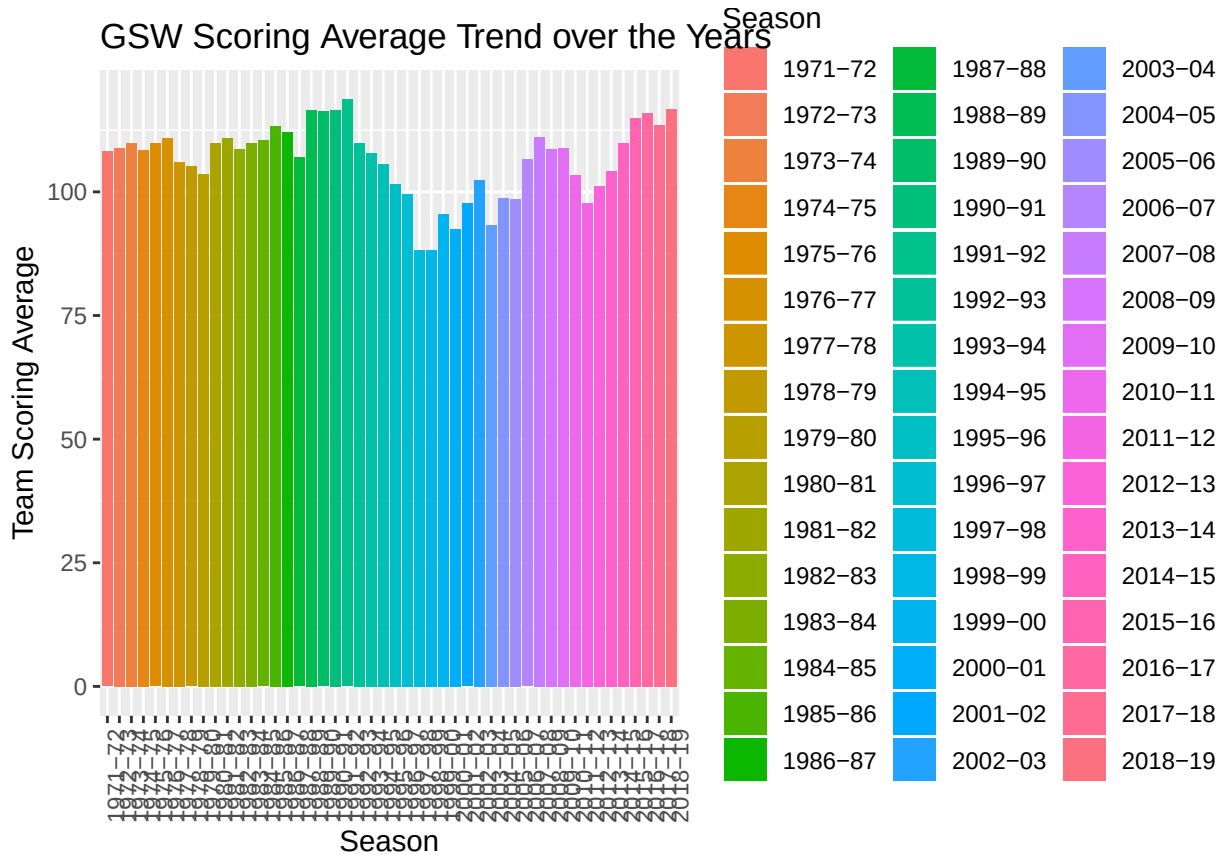
```
GSWStatsTibble <- nbaStatsTibble %>% filter(Tm == "GSW" | Tm == "GSW*")
GSWRatingsTibble <- nbaRatingsTibble %>% filter(Tm == "GSW" | Tm == "GSW*")

offensiveRatings <- GSWRatingsTibble %>% group_by(Season) %>% summarize(Rating=mean(ORtg))
scoringAvg <- GSWStatsTibble %>% group_by(Season) %>% summarize(Avg=mean(PTS / G))
threePointPct <- GSWStatsTibble %>% group_by(Season) %>%
  summarize(threePT=mean((X3PA / FGA) * 100))

# offensive ratings plot
p <- ggplot(offensiveRatings, aes(y=Rating,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("GSW Offensive Ratings Trend over the Years") +
```

```
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1))
```

```
## Warning: Removed 2 rows containing missing values (position_stack).
```

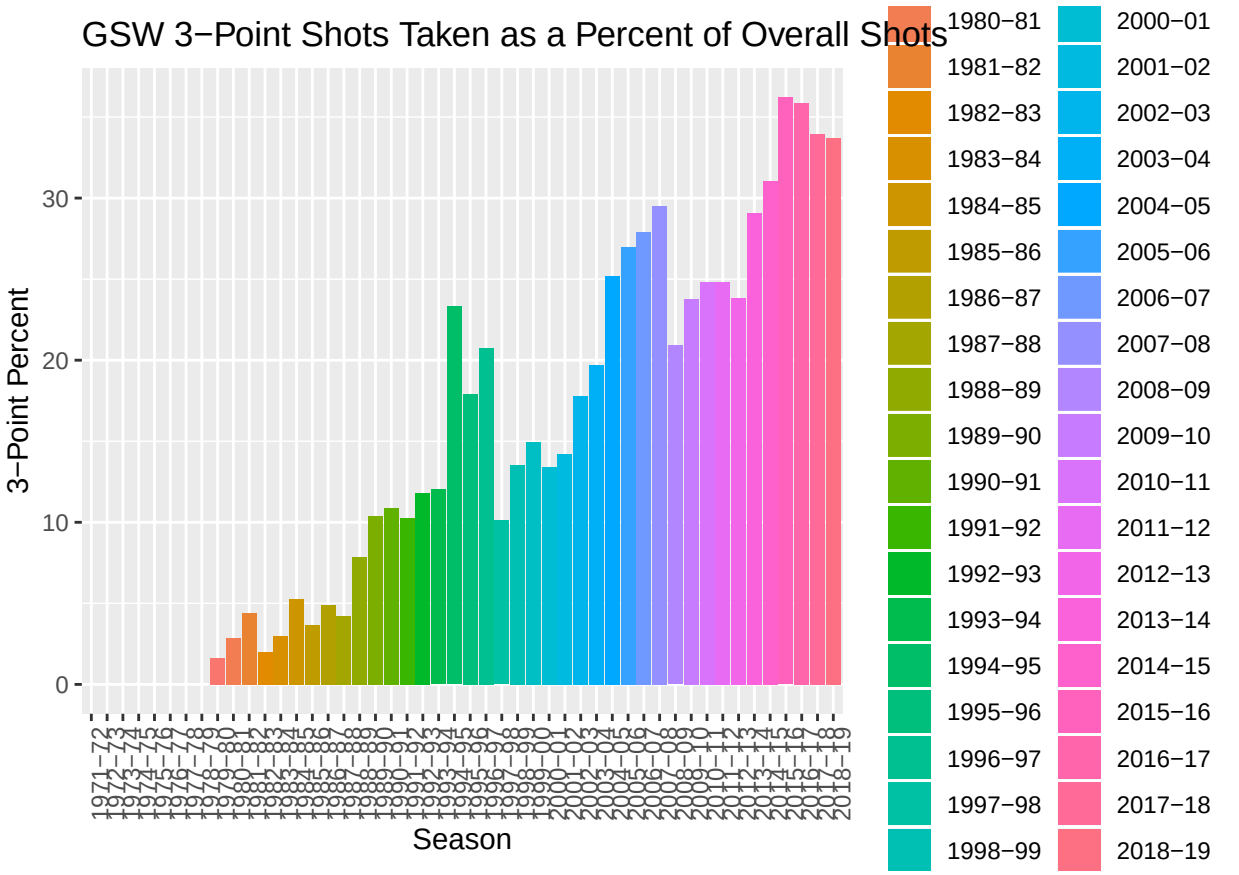



```
# scoring average
p <- ggplot(scoringAvg, aes(y=Avg,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("GSW Scoring Average Trend over the Years") +
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1)) +
  ylab("Team Scoring Average")
```




```
# percentage growth of 3-point shots
p <- ggplot(threePointPct, aes(y=threePT,x=Season, fill=Season))
p + geom_bar(stat = "identity") +
  ggtitle("GSW 3-Point Shots Taken as a Percent of Overall Shots ") +
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1)) +
  ylab("3-Point Percent")
```

```
## Warning: Removed 8 rows containing missing values (position_stack).
```



Looking at the Warriors stats we can see a steady and meaningful rise in there offensive ratings, and scoring average as there number of 3-point shots continued on a dramatic rise and a lot of championship success. Major impact on the rise of the Warriors.

Conclusion

From all the data and plots, we can observe a real trend in the modern nba. So many different stats point to the increased use of the 3 point shot in all teams and positions. We can see how centers have begun shooting 3 pointers at a much higher rate. This trend that GSW has set has even changed the way certain positions the game. Upward trends of 3 point percent of shot attempts show a favoring of the shot. Other teams have begun the climb towards GSW's dominance and it is undeniable that this modern dynasty is redefining the NBA. More research into specific data, like more individual players, could help us determine even more effects of this trend. Either way, GSW has created a unique, powerful and influential dynasty and it's never been a better time to be a fan.