

# The simple history of 3 point shooting

Weiting Li

Summer II 2019

## Abstract

Wikipedia: "A three-point field goal (also 3-pointer, three or informally, trey) is a field goal in a basketball game made from beyond the three-point line, a designated arc surrounding the basket." This analysis is about using statistical methods: Plots, Tables..etc, to study the recent 3 point shooting trend in NBA. In order to finish this task, we want to trace back to where it begins and how it evolves.

## Introduction

Everybody who is watching NBA these years, should know that 3 pointers is like a storm in the league right now. It is so obtrusive that everyone knows how popular they are and how powerful they are on the court. I am also a big fan of NBA since my college life starts, and since then I have never missed any games of my favorite team: Houston Rockets, also the team that shoots historical record of 3 pointers attempts. Therefore, today, as a big fan I want to use statistical methods to discover the secret behind the 3 pointers.

## Description of the data-set I use

- Here is a introduction to the data below:
  - The origin of this data-set is on a website called Basketball reference:  
[https://docs.google.com/spreadsheets/d/1cpXLS\\_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1cpXLS_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing)  
([https://docs.google.com/spreadsheets/d/1cpXLS\\_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1cpXLS_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing))
  - This dataset consists of a few important variables that I am going to use: Player, POS(position), Team, 3PA(3 pointers attempts), 2PA(mid-range attempts), 3P%, 2P% etc.
  - The reason I choose this data-set is because this data-set has all seasons stats on record of NBA, and it starts from 1950 to the present, which means there are plentiful data for me to play with.

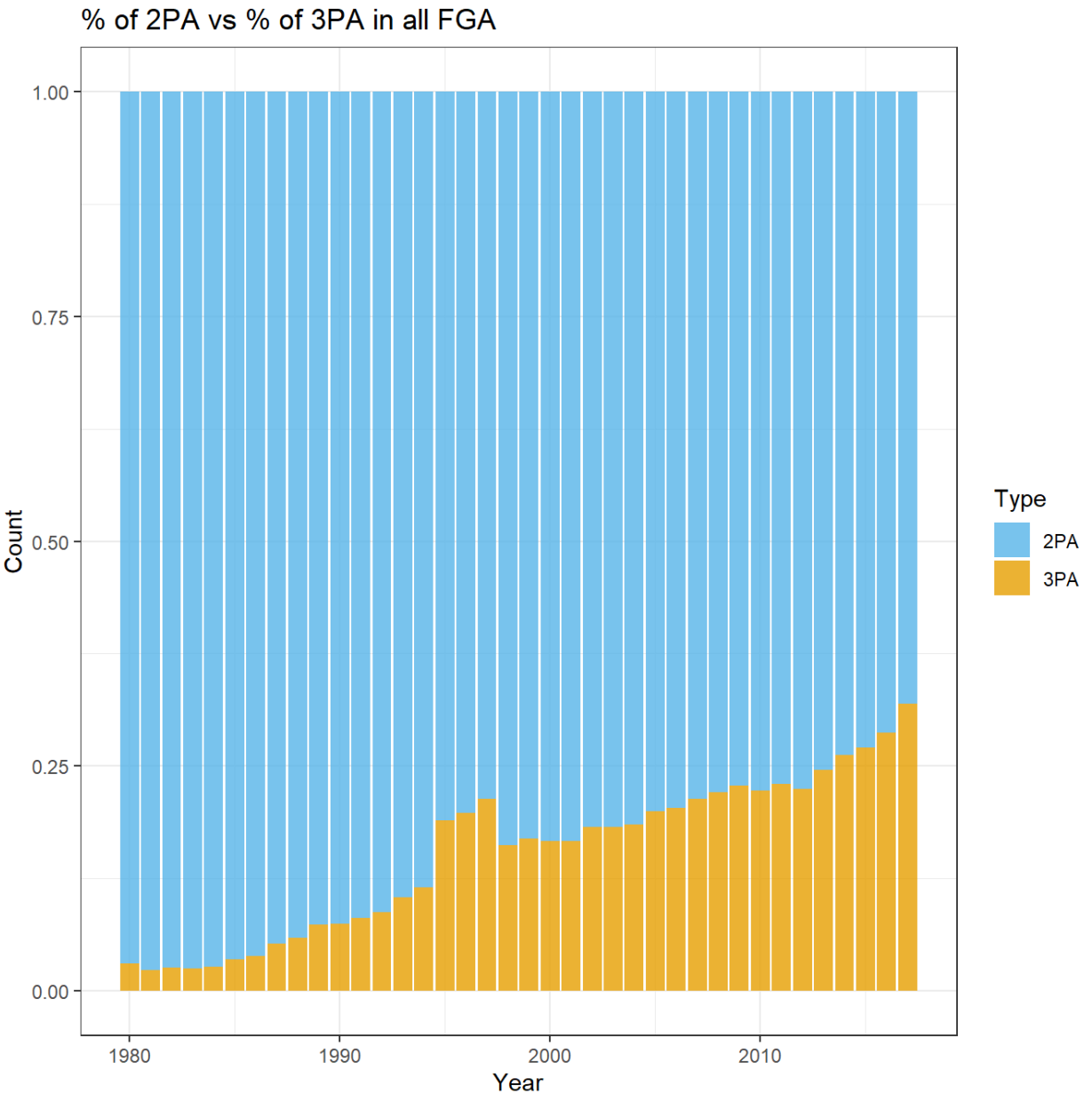
## Methods

- - To show how 3 pointers developed after it was firstly added to the rules of NBA league in 1979, I will be using statistical methods like plots. The reason to choose this method is because plots are more direct and easy to understand for all readers that may have zero knowledge background of this sport. Therefore, I will be using ggplot() mainly to show the results.
  - Also, I used table such ktable or pander functions to give results.
  - I used a package called wordcloud to present my result.

## Results and Discussion

### Example 1

Firstly, I create the graph below to show how percentage of 3PA in all attempts developed since 1980 compared to 2PA.



Discussion

According to the graph, it is very obvious that after 3 pointers been introduced to the league, players start from barely shooting a 3p to shooting 3 pointers very frequently every game.

Although 2 point shooting (mid-range, lay-up) still made up most of the shooting attempts, 3 pointers attempts percentage has increased dramatically and definitely is getting to be a important basketball skill in the league.

Example 2

Then, I want to determine how number of 3 pointers changed in every game compared to 2PA. Therefore, I create another graph which is about the mean of two type of shots per game from 1980 to 2017. And then also I want to see what the 3 pointers has become compared to what it was when first invented, along with how 2PA per game change after all these years too. So I create a table to show how much they have changed in 2017 compared to 1980.

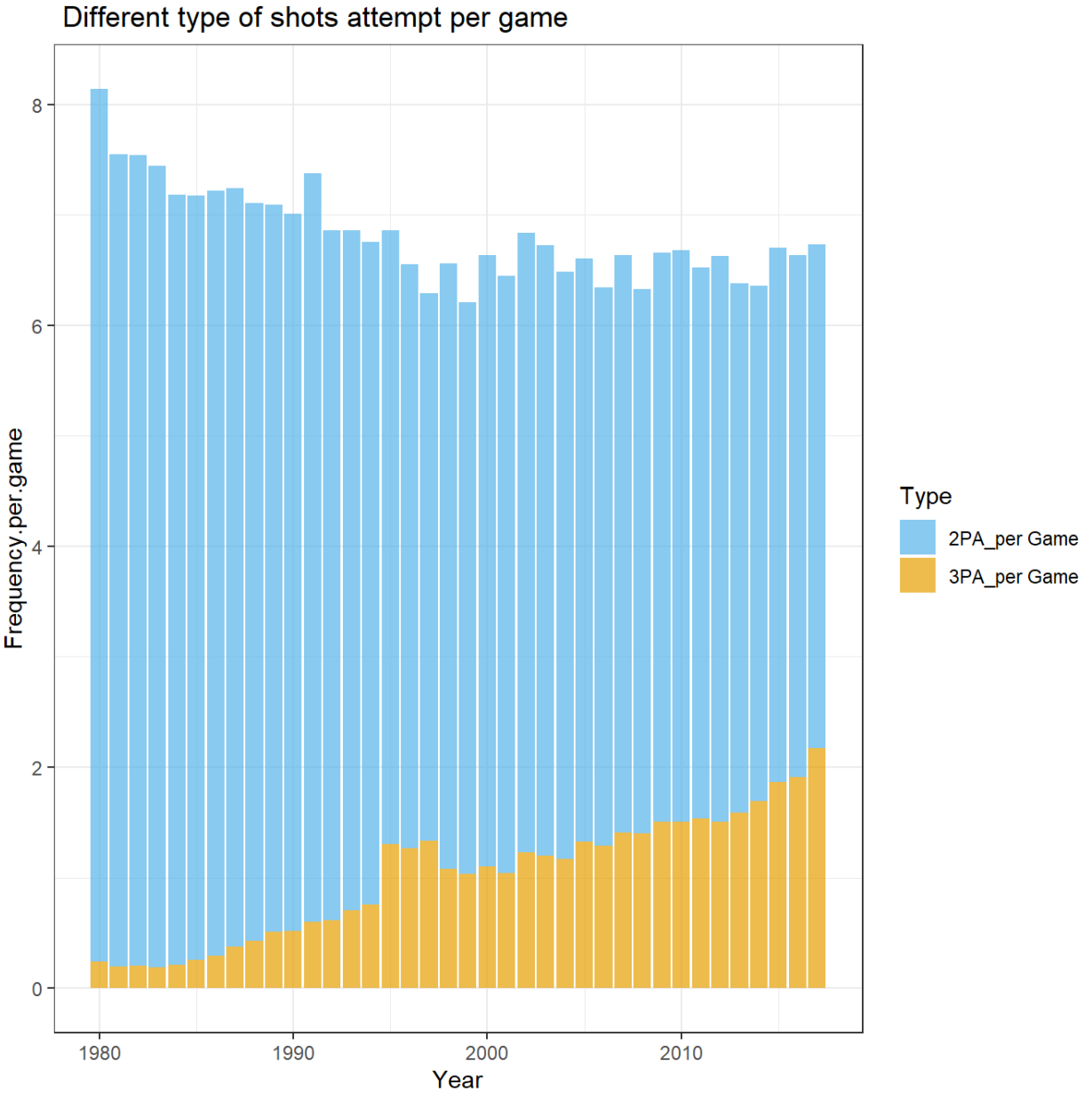


Table continues below

|      | diff.2PA | diff.3PA | max.3PAPG.1980 | max.3PAPG.2017 |
|------|----------|----------|----------------|----------------|
| 2017 | 0.5777   | 8.994    | 3.069          | 9.987          |

|      | max.3PPG.1980 | max.3PPG.2017 |
|------|---------------|---------------|
| 2017 | 1.154         | 4.101         |

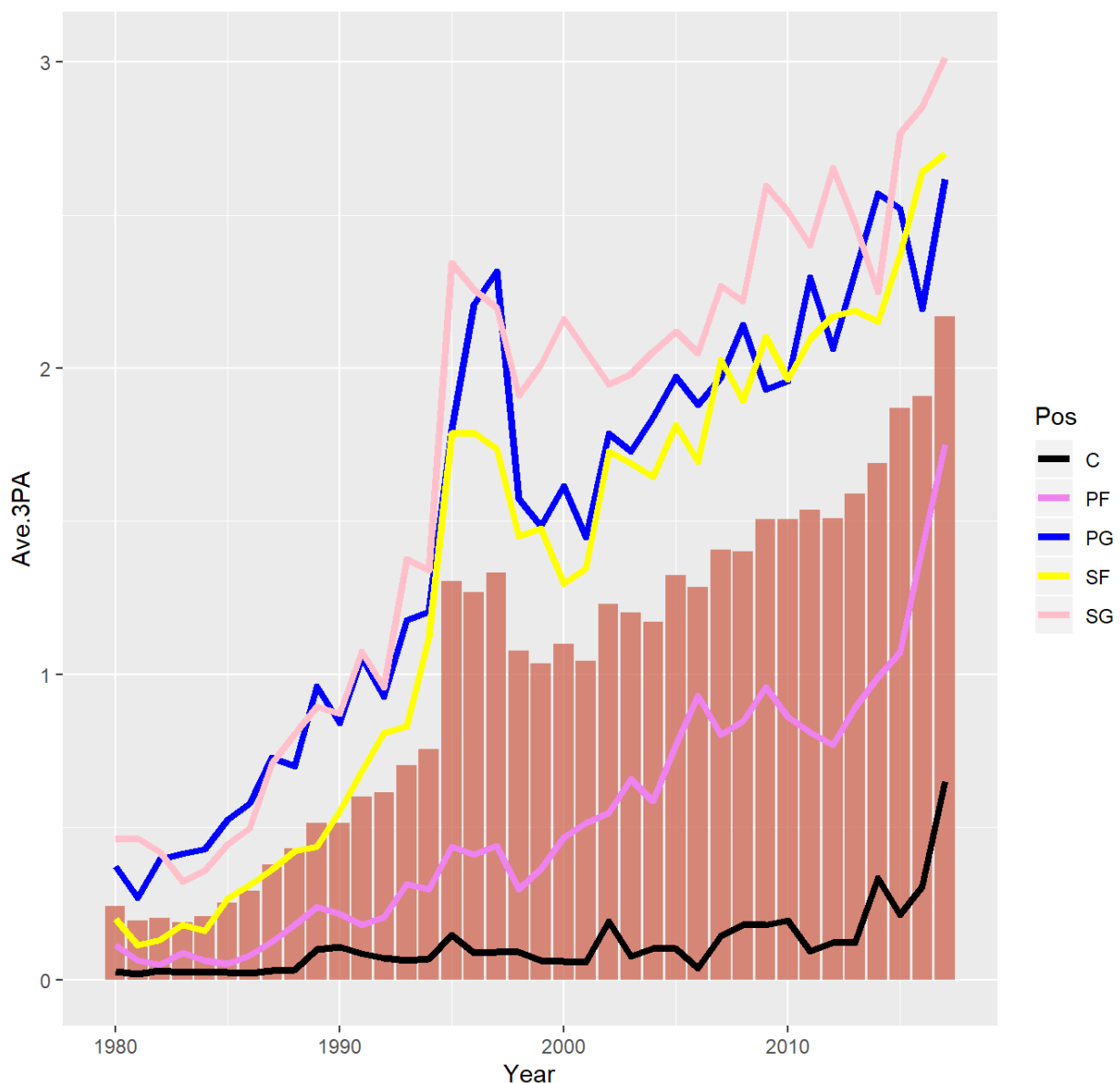
Discussion

According to the graph, it is easy to see that the observation complies with what we observed in the first example. 3PA per game increased dramatically for every player in every game. Interestingly, Players in 2017 shoots averagely 4.56 2Ps per game and 2.16 3Ps per game. Compared to NBA in 1980, the numbers are 7.89 2Ps per game and 0.24 3Ps per game. In other words, 2PA.per game in 2017 decreased by 42.3% compared to 1980 and 3PA. per game increased by 799%, which is almost 8 times the number in 1980. It is also incredible to see that the maximum of 3PA in 1980 was 3.069, but in 2017 it is 9.987, which means it increases about 3 times. And the max of 3P made every game was 1.15 but in 2017 it becomes 4.1(about 4 times).

## Example 3

For this example, I want to study about how 3PA and 2PA per game develops among different positions (PG, SG, SF, PF, C).

3PA per game in different positions



## Discussion

Unsurprisingly, The graph tells the fact that PG and SG are the first two positions adapted to 3 point shooting. It is not hard to understand this result because in the view of history, nba has always been a league promotes height and limb length. Especially in the 80s, which usually fans call it 'Ancient times', all of stars usually are centers. So in order to survive in the league, short guys usually have to learn much more skills than PF and C. But there is more to tell. Since 3P appears in NBA, it is not hard to see that everyone in the league is adapting to 3P shootig and learns how to shoot consistent 3Ps every game, because there is apparent growth of 3 point attempts per game in every positions. Also compared to players in center, number of players in PF obviously has a much higher increase after 2012. And this totally complies with my own experience that is NBA teams prefers PF that are capable of shooting 3Ps.

## Example 4

I am also interested in finding out who is the best 3 point shooter in NBA history.

The greatest 3p maker in history

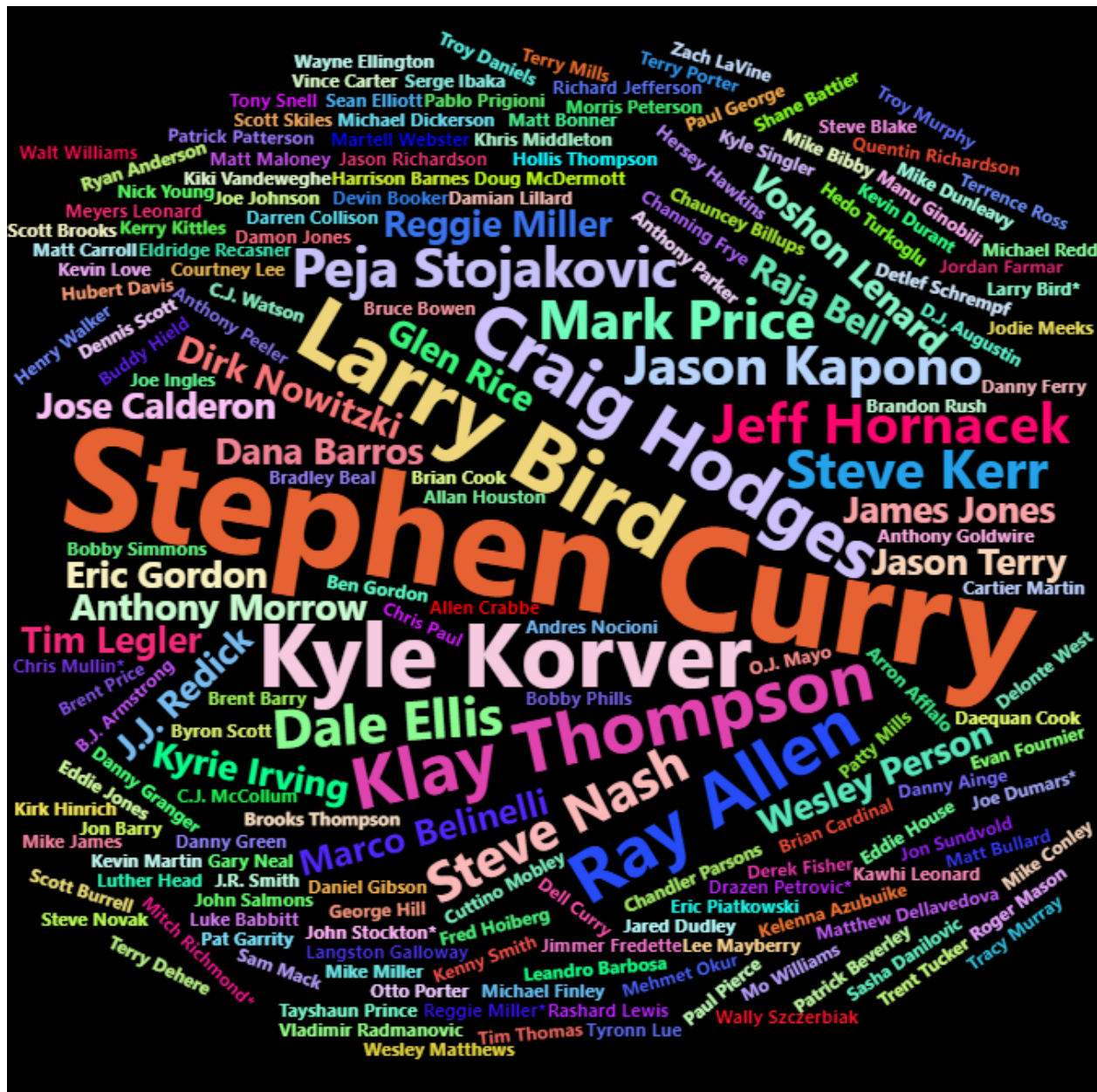
| Rank | Player           | Games | x3Pa | x3Pt. | x3P.made |
|------|------------------|-------|------|-------|----------|
| 1    | Ray Allen        | 1376  | 7962 | 0.399 | 3174     |
| 2    | Reggie Miller*   | 1389  | 6486 | 0.395 | 2560     |
| 3    | Kyle Korver      | 1173  | 5409 | 0.429 | 2322     |
| 4    | Vince Carter     | 1497  | 6119 | 0.375 | 2292     |
| 5    | Chauncey Billups | 1274  | 5810 | 0.386 | 2245     |
| 6    | Jason Terry      | 1359  | 5895 | 0.380 | 2242     |
| 7    | Joe Johnson      | 1379  | 5629 | 0.372 | 2096     |
| 8    | Peja Stojakovic  | 908   | 4932 | 0.401 | 1979     |
| 9    | J.R. Smith       | 950   | 5173 | 0.374 | 1937     |
| 10   | Jason Richardson | 1009  | 5148 | 0.374 | 1926     |

## Discussion

Based on stats such as Games played, number of 3P attempts, percentage of 3 pointers, and 3 point shots made in career, Ray allen is found to be the best 3 point shooter ever in NBA history. After him is Reggie Miller and Kyle Korver. This result does not surprise me because according to my experience, Ray Allen has been considered a great competitor of the best 3P shooters in history after he retired in 2014. He has made 3174 3ps out of 7962 attempts with a incredible accuracy of 0.399 (0.4).

## Example 5

Finally, in the last analysis, I want to study about who is the most dominant 3 point shooter, based on three data: 1. my dataset; 2. Survey online; 3. Three point contest winners.



## Dicussion

According to survey and history, the main stream of fans and nba consider Stephen Curry as the best 3p shooters that we ever met.

## Conclusions

According to my research, 3 point shooting has been more and more popular and significant in nowadays NBA. In 2017, It has increased 8 times of itself in 1980. And more importantly, 3 point shooting changed the NBA dramatically. Before it was invented, most of the basketballs stars are centers. But now, Stephen Curry, James Harden, Klay Thompson.. etc are all small guards that are super good at making three point FGs. It balanced the basketball world so that not only giants can play this sport. If you were a coach of a basketball team, I would highly recommend you to urge his/her players to practice more 3s daily. Because now it is a trend all over the world not just in NBA. If any one wants to play basketball well in the era we live in, this is absolutely the best skill everyone should learn. And also it will increase space on the court which promotes better offense. The golden era of 3 point shooting has arrived.

## Citations

1. Wordcloud2: <https://cran.r-project.org/web/packages/wordcloud2/vignettes/wordcloud.html#install-wordcloud2>  
(<https://cran.r-project.org/web/packages/wordcloud2/vignettes/wordcloud.html#install-wordcloud2>)
2. Survies:
  - 10 best 3-point shooters in NBA history, ranked|For The Win: <https://ftw.usatoday.com/2016/05/best-3-point-shooters> (<https://ftw.usatoday.com/2016/05/best-3-point-shooters>);
  - Ranking the 10 greatest three-point shooters in NBA history: <https://www.msn.com/en-us/sports/nba/ranking-the-10-greatest-three-point-shooters-in-nba-history/ss-AAxCwvJ>  
(<https://www.msn.com/en-us/sports/nba/ranking-the-10-greatest-three-point-shooters-in-nba-history/ss-AAxCwvJ>)
  - NBA History: The Greatest 3-Point Shooters Ever: <https://bleacherreport.com/articles/793178-nba-history-ray-allen-larry-bird-and-the-greatest-3-point-shooters-ever#slide0>  
(<https://bleacherreport.com/articles/793178-nba-history-ray-allen-larry-bird-and-the-greatest-3-point-shooters-ever#slide0>)
3. Wikipedia page I quote: [https://en.wikipedia.org/wiki/Three-point\\_field\\_goal](https://en.wikipedia.org/wiki/Three-point_field_goal)  
([https://en.wikipedia.org/wiki/Three-point\\_field\\_goal](https://en.wikipedia.org/wiki/Three-point_field_goal))
4. Where I collect my dataset: <https://www.basketball-reference.com/> (<https://www.basketball-reference.com/>)  
link to the data:  
[https://docs.google.com/spreadsheets/d/1cpXLS\\_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1cpXLS_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing)  
([https://docs.google.com/spreadsheets/d/1cpXLS\\_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1cpXLS_OYFeMvdtgnOF5HI6QIAjqZR4p98X1aOPyyhTg/edit?usp=sharing))

## Appendix

### Code for Example 1

```
nba <- read.csv("Seasons_Stats.csv") # read the file
test.1 <- select(filter(nba, nba$Year > 1979), Year, Pos, Player, G, X2PA, X3PA, FGA) # First, filter it.
year <- c(1980:2017) # a list of years in this case.
test.1 <- na.omit(test.1) # delete all NAs
#str(test.1)

Sum.2 <- tapply(test.1$X2PA, test.1$Year, sum) # sum of 2p attempts every year.
Sum.3 <- tapply(test.1$X3PA, test.1$Year, sum) # sum of 3p attempts every year.
Sum <- tapply(test.1$FGA, test.1$Year, sum) # sum of all goal attempts every year.

joint1 <- rbind( # a data.frame to combine these variables
  data.frame(x = year, Count = Sum.2/Sum, Type = "2PA"),
  data.frame(x = year, Count = Sum.3/Sum, Type = "3PA")
)

ggplot(joint1) + geom_bar(aes(x = joint1$x, y = Count, fill = Type), alpha = 0.8, stat = "identity") + labs
(x = "Year", title = "% of 2PA vs % of 3PA in all FGA") + theme_bw() + scale_fill_manual(values=c( "#56B
4E9", "#E69F00")) # ggplot of 2PA% versus 3PA% with years as x-axis.
```

### Code for Example 2

```

test.2 <- select(filter(nba, nba$Year > 1979),Year, G, X2PA, X3PA, FGA, X3P)
# firstly, I create a variable called "X2PA.G" for every player, 2 pointer attempt every game.
test.2 <- mutate(test.2, X2PA.G = test.2$X2PA/test.2$G)
#Also, one variable called "X3PA.G" for 3PA too.
test.2 <- mutate(test.2, X3PA.G = test.2$X3PA/test.2$G)
x1 <- tapply(test.2$X2PA.G, test.2$Year, mean) #I take the mean of the X2PA.G for each year.

x2 <- tapply(test.2$X3PA.G, test.2$Year, mean) # for X3PA.G too.

joint3 <- rbind( # a data.frame to combine these variables
  data.frame(x = year, Frequency.per.game = x1, Type = "2PA_per Game"),
  data.frame(x = year, Frequency.per.game = x2, Type = "3PA_per Game")
)

ggplot(joint3) + geom_bar(aes(x = joint3$x, y = Frequency.per.game, fill = Type), alpha = 0.7, stat = "identity") + labs(x = "Year", title = "Different type of shots attempt per game") + theme_bw() + scale_fill_manual(values=c( "#56B4E9", "#E69F00"))

# ggplot of X2P.G vs X3P.G with year as x-axis.

diff.2PA <- x1[38]/x1[1] # to see how much X2PA.G change after years
diff.3PA <- x2[38]/x2[1] # to see how much X3PA.G change after years
t2017 <- subset(test.2, test.2$Year == 2017)
#head(t2017)
t1980 <- subset(test.2, test.2$Year == 1980)
#head(t1980)

max.3PAPG.1980 <- max(t1980$X3PA/t1980$G) # max of 3PA in 1980
max.3PAPG.2017 <- max(t2017$X3PA/t2017$G) # max of 3PA in 2017

max.3PPG.1980 <- max(t1980$X3P/t1980$G)
max.3PPG.2017 <- max(t2017$X3P/t2017$G)

pander(cbind(diff.2PA, diff.3PA, max.3PAPG.1980, max.3PAPG.2017, max.3PPG.1980, max.3PPG.2017)) # table to show the result

```

## Code for Example 3



```

attach(test.1)

PG <- subset(test.1, test.1$Pos == "PG") #subset for PG
SG <- subset(test.1, test.1$Pos == "SG") #subset for SG
SF <- subset(test.1, test.1$Pos == "SF") #subset for SF
PF <- subset(test.1, test.1$Pos == "PF") #subset for PF
C <- subset(test.1, test.1$Pos == "C") #subset for C

Ave.3PA <- test.1$X3PA/test.1$G # average 3PA per game
PG.3PA <- PG$X3PA/PG$G # average 3PA per game for PG

SG.3PA <- SG$X3PA/SG$G # etc...
SF.3PA <- SF$X3PA/SF$G
PF.3PA <- PF$X3PA/PF$G
C.3PA <- C$X3PA/C$G

#Data frames I use
Average <- aggregate(data = test.1, Ave.3PA ~ Year, FUN = "mean", na.rm = TRUE)
PG.3PA.PG <- aggregate(data = PG, PG.3PA ~ Year, FUN = "mean", na.rm = TRUE)
SG.3PA.PG <- aggregate(data = SG, SG.3PA ~ Year, FUN = "mean", na.rm = TRUE)
SF.3PA.PG <- aggregate(data = SF, SF.3PA ~ Year, FUN = "mean", na.rm = TRUE)
PF.3PA.PG <- aggregate(data = PF, PF.3PA ~ Year, FUN = "mean", na.rm = TRUE)
C.3PA.PG <- aggregate(data = C, C.3PA ~ Year, FUN = "mean", na.rm = TRUE)

# ggplot

ggplot()+ geom_bar(data = Average, aes(x = Year, y = Ave.3PA), alpha = 0.7, fill = 'coral3', stat = "identity") +
  geom_line(aes(x= Average$Year, y = PG.3PA.PG$PG.3PA, colour = "PG"), size = 1.5) +
  geom_line(aes(x= Average$Year, y = SG.3PA.PG$SG.3PA, colour = "SG"), size = 1.5) +
  geom_line(aes(x= Average$Year, y = SF.3PA.PG$SF.3PA, colour = "SF"), size = 1.5) +
  geom_line(aes(x= Average$Year, y = PF.3PA.PG$PF.3PA, colour = "PF"), size = 1.5) +
  geom_line(aes(x= Average$Year, y = C.3PA.PG$C.3PA, colour = "C"), size = 1.5) +
  scale_colour_manual(name="Pos", values=c( C="black", PG="blue", SG="pink", SF="yellow", PF="violet")) +
  labs(title = "3PA per game in different positions")

```

## Code for Example 4

```

# filter the dataset
new.nba <- select(filter(nba, Year >= 1980 & X3P. > 0), Player, G, X3PA, X3P, X3P.)

# get the best 3p maker
new.nba = new.nba %>% group_by(Player) %>% summarize(
  Games = sum(G),
  x3P.made = sum(X3P),
  x3Pa = sum(X3PA),
  x3Pt. = round(x3P.made/x3Pa, 3)
) %>% filter(x3P.made >= 200, x3Pt. >= 0.37) %>%
  arrange(desc(x3P.made), desc(x3Pt.)) %>%
  mutate(Rank = min_rank(desc(x3P.made)))

#re-order the variables
Greatest.3pm = select(new.nba, Rank, Player, Games, x3Pa, x3Pt., x3P.made)

# table
kable(head(Greatest.3pm, 10), escape = FALSE, caption = " The greatest 3p maker in history") %>% kable_styling("striped", full_width = T) %>%
  column_spec(6, color = "white", background = "green") %>% column_spec(1, bold = T, color = "red", background = "black")

```

## Code for Example 5

```
require(devtools)
library(wordcloud2) # wordcloud2

#All time rank of 3P shooter based on 3p made.

Best3p.<- nba %>% group_by(Player) %>%

  summarise(x3P.m = sum(X3P),
            x3Pa = sum(X3PA),
            x3P. = x3P.m / x3Pa)%>% filter(x3P. > 0.37 , x3Pa >= 2000) %>% arrange(desc(x3P.)) %>% head(10)

#According to survey online

Best3pS_fans <- c("Jason Terry", "Klay Thompson", "Dale Ellis", rep("Larry Bird",2), rep("Kyle Korver",3),
"Steve Nash", rep("Reggie Miller",2), "Steve Kerr", rep("Stephen Curry",5), rep("Ray Allen",2))

# Three point contests
Player <- c("Larry Bird", "Larry Bird", "Larry Bird", "Dale Ellis", "Craig Hodges", "Craig Hodges", "Craig Hodges", "Mark Price", "Mark Price", "Glen Rice", "Tim Legler", "Steve Kerr", "Jeff Hornacek", "Jeff Hornacek", "Ray Allen", "Peja Stojakovic", "Peja Stojakovic", "Voshon Lenard", "Quentin Richardson", "Dirk Nowitzki", "Jason Kapon", "Jason Kapon", "Daequan Cook", "Paul Pierce", "James Jones", "Kevin Love", "Kyrie Irving", "Marco Belinelli", "Stephen Curry", "Klay Thompson", "Eric Gordon", "Devin Booker")
X3P.Contest <- data.frame( Player)

#data frame
Best3pMakers <- data.frame(Player = c(as.character(X3P.Contest$Player), as.character(Greatest.3pm$Player), as.character(Best3p.$Player), as.character(Best3pS_fans))) %>% count(Player) %>% arrange(desc(n))

#make the wordCloud
Best3pMakers %>% wordcloud2(size=0.5, color='random-light', backgroundColor="black")
```