Stat 456 Final Project

Ji Hoon Woo

December 16, 2018

"Our Team Intentionally Lost the Game to Pick You"

Abstract

Baseball is the most well-known Sports for statistics because players act individually rather than performing in clusters. The movie "Moneyball" explicitly demonstrates how statistics can be implemented to make the team better performing. By the way, basketball is another interesting Sports requiring various statistical analysis and computation to find new outcome. In this project, 27 players of National Basketball Association 1st draft pick and each player's personal statistics are collected for experimentation. For those who do not have much knowledge about the NBA will assume that they may be all great players because each player was drafted 1st every season. I was also the person who believed their great performance statistics. However, some of them failed to prove their values due to the injury or the level of difficulty in NBA. Furthermore, only 60% of the players won the Rookie of the Year on their 1st season meaning that 40% of them were not the best rookies as fans or teams anticipated before the season begins. This paper then examines who were considered to be great rookies among other 1st draft pick players and compares any similarity or difference to find any interesting result. In detail, I divide them into 4 groups to investigate how each group is characterized than other groups. Aa a result, the advent of Shaquille O'Neal shocked NBA fans that his performance overwhelmed the other 1st draft pick players. He was an absolute outlier in 27 years of NBA history which will be examined in detail. Another interesting discovery from this analysis is the player's role of play. Their style of play is quite strictly assigned to certain tasks which will be explained in Main Results and Conclusions.

Introduction

There is annual event in National Basketball Association at the end of each season that fans and teams are looking forward to watching as much as the regular season. This is the NBA Draft. Not only the college undergraduates, but also international players are eligible to participate to be selected. Interestingly, here is the fun issue. There is a term known as "tanking" which refers to those teams from losing games on purpose to get a better draft choice (John, The NBA's New Lottery Will Merely Encourage More Teams To Tank Differently). Even though tanking rules are being changed to prevent teams to lose on purpose especially for the fans who pay the tickets to watch the game, there is still high chance for teams with more loses to obtain higher draft picks. My curiosity then begins to arise as I would like to know more about this. My fundamental question begins with "How much did the 1st round pick players reimbursed the team's previous year's sacrifice?" There must be some players who exceeded the teams' expectancy while some others are remembered as biggest bust. In order to begin my investigation, I had to choose variables that I would like to experiment and find interesting thoughts.

I chose 17 variables to implement my analysis from the official NBA Stats website. I searched individual player since Shaquille O'Neal in 1992. The reason I collected data from this year is because he is considered to be one of the best players in NBA history and wanted to observe if there is any player who outperformed than him in the 1st year of season. My 17 variables contain name, season each player played, position and personal statistics for only the rookie year. Points is obviously an important factor which leads to win the game, but there are also other variables we might consider significant. This paper is going to be full of analysis investigating how they are correlated, and which player has performed well from certain categories.

Goals

1. There is a certain role or aspect which team expects from the 1st draft pick player. What kind of characteristics are combined and correlated to define each different style of play?

   - The team assigns different playing role based on the player's ability and team's style of play. For instance, there are several terms describing each different style of play: Slasher, Swingman, Combo Guard etc. Each role of play requires certain type of skillsets to be excellent than other skills. In doing so, we are able to define both players' style of play and their necessary skillsets.

2. Fans and teams cannot be so definitive about which player outperformed the others because we have different measurement and expectancy based on the role. Then, how can we decide which player made better performance in rookie season than the others?

   - Players with fancy movements usually have more fans than the others but that does not mean that he is a better player. Then, we need to compare the actual statistics with those players having the similar style of play or with same position.

3. There must be some unknown facts we did not realize before the investigation. What are the interesting findings that have been brought up from this project?

- There are some assumptions that we believe to be true. For example, taller players are more likely to rebound the ball than smaller players. There must some other interesting facts which we did not expect, or certain player has unexpected statistics than similar type of players.

Main Results

	Season	Position	Height	Weight	Games	Minutes	Points	FGP	X3PP	FTP	REB	AST	TOV	STL	BLK	PF
Deandre Ayton	2018-2019	Center	216	113.0	28	30.9	15.8	0.598	0.000	0.758	10.2	2.5	1.9	0.5	0.9	2.9
Markelle Fultz	2018-2019	Guard	193	91.0	19	22.5	8.2	0.419	0.286	0.568	3.7	3.1	1.3	0.9	0.3	2.7
Ben Simmons	2017-2018	Forward	208	104.0	81	33.7	15.8	0.545	0.000	0.560	8.1	8.2	3.4	1.7	0.9	2.6
Karl-Anthony Towns	2015-2016	Center	213	112.0	82	32.0	18.3	0.542	0.341	0.811	10.5	2.0	2.2	0.7	1.7	3.0
Andrew Wiggins	2014-2015	Forward	203	88.0	82	36.2	16.9	0.437	0.310	0.760	4.6	2.1	2.2	1.0	0.6	2.3
Anthony Bennett	2013-2014	Forward	203	111.0	52	12.7	4.2	0.356	0.245	0.638	3.0	0.3	0.9	0.4	0.2	1.8
Anthony Davis	2012-2013	Center	208	114.8	64	28.8	13.5	0.516	0.000	0.751	8.2	1.0	1.4	1.2	1.8	2.5
Kyrie Irving	2011-2012	Guard	191	87.5	51	30.5	18.5	0.469	0.399	0.872	3.7	5.4	3.1	1.1	0.4	2.2
John Wall	2010-2011	Guard	193	94.8	69	37.8	16.4	0.409	0.296	0.766	4.6	8.3	3.8	1.8	0.5	2.5
Blake Griffin	2010-2011	Forward	208	113.0	82	37.9	22.5	0.506	0.292	0.642	12.1	3.8	2.7	0.8	0.5	3.1
Derrick Rose	2008-2009	Guard	191	91.0	81	37.0	16.8	0.475	0.222	0.788	3.9	6.3	2.5	0.8	0.2	1.5
Greg Oden	2008-2009	Center	216	125.0	61	21.5	8.9	0.564	0.000	0.637	7.0	0.5	1.4	0.4	1.1	3.9
Andrea Bargani	2006-2007	Forward	213	111.0	65	25.1	11.6	0.427	0.373	0.824	3.9	0.8	1.6	0.5	0.8	2.8
Andrew Bogut	2005-2006	Center	213	118.0	82	28.6	9.4	0.533	0.000	0.629	7.0	2.3	1.5	0.6	0.8	3.2
Dwight Howard	2004-2005	Center	211	120.0	82	32.7	12.0	0.520	0.000	0.671	10.0	0.9	2.0	0.9	1.7	2.8
LeBron James	2003-2004	Forward	203	113.0	79	39.6	20.9	0.417	0.290	0.754	5.5	5.9	3.5	1.6	0.7	1.9
Yao Ming	2002-2003	Center	229	141.0	82	29.0	13.5	0.498	0.500	0.811	8.2	1.7	2.1	0.4	1.8	2.8
Kwame Brown	2001-2002	Center	211	132.0	57	14.3	4.5	0.387	0.000	0.707	3.5	0.8	0.8	0.3	0.5	1.8
Kenyon Martin	2000-2001	Forward	206	106.0	68	33.4	12.0	0.445	0.091	0.630	7.4	1.9	2.0	1.1	1.7	4.1
Elton Brand	1999-2000	Forward	206	115.0	81	37.0	20.1	0.482	0.000	0.685	10.0	1.9	2.8	0.8	1.6	3.2
Michael Olowokandi	1998-1999	Center	213	122.0	45	28.5	8.9	0.431	0.000	0.483	7.9	0.6	1.9	0.6	1.2	3.0
Tim Duncan	1997-1998	Forward	211	113.0	82	39.1	21.1	0.549	0.000	0.662	11.9	2.7	3.4	0.7	2.5	3.1
Allen Iverson	1996-1997	Guard	183	75.0	76	40.0	23.5	0.416	0.341	0.702	4.1	7.5	4.4	2.1	0.3	3.1
Joe Smith	1995-1996	Forward	208	102.0	82	34.4	15.3	0.458	0.357	0.790	8.7	1.0	1.0	1.6	1.7	2.7
Glenn Robinson	1994-1995	Forward	201	109.0	80	37.0	21.9	0.451	0.321	0.796	6.4	2.5	1.4	0.3	3.9	2.7
Chris Webber	1993-1994	Forward	208	113.0	76	32.1	17.5	0.552	0.299	0.649	9.8	4.2	1.4	1.4	2.8	3.2
Shaquille O'Neal	1992-1993	Center	216	147.0	81	37.9	23.4	0.562	0.000	0.592	13.9	1.9	0.7	3.5	3.8	4.0

TABLE 1.1

TABLE 1.1 is my dataset of NBA 1st draft pick players from 1992 in chronological order. Markelle Fultz, Ben Simmons, Blake Griffin and Greg Oden have the season record a year after their draft because they did not play their 1st season or played only few games due to the injury. The rest of players have their 1st season record because implementing their entire career statistics is making much harder to compare and analyze. There are players who are injury prone or experience early decline of abilities while players like LeBron James maintain his maximum ability all the time. In addition, players generally do not play for one team, but they transfer to the other team in order to receive much more money or to obtain NBA Championship Ring.

Term Abbreviation:

FGP	3PP	FTP	REB	AST	TOV	STL	BLK	PF
Field Goal %	3 Point %	Free Throw %	Rebound	Assist	Turnover	Steal	Block	Personal Foul Committed

TABLE 1.2

TABLE 1.2 represents original terms of abbreviation

	Height	Weight	Minutes	Points	FGP	X3PP	FTP	REB	AST	TOV	STL	BLK	PF
Height	1.0000	0.8609	-0.2415	-0.2400	0.5027	-0.2998	-0.1331	0.5702	-0.6278	-0.4658	-0.2411	0.3981	0.3765
Weight	0.8609	1.0000	-0.2708	-0.2059	0.3631	-0.3995	-0.2422	0.5336	-0.5722	-0.5225	-0.0597	0.4942	0.3276
Minutes	-0.2415	-0.2708	1.0000	0.8910	0.2746	0.1207	0.1601	0.4181	0.5288	0.6165	0.5036	0.3139	0.2091
Points	-0.2400	-0.2059	0.8910	1.0000	0.3091	0.2362	0.2742	0.4385	0.4938	0.5461	0.5086	0.3854	0.1845
FGP	0.5027	0.3631	0.2746	0.3091	1.0000	-0.3958	-0.0950	0.7613	-0.0510	-0.0509	0.1104	0.4429	0.5152
X3PP	-0.2998	-0.3995	0.1207	0.2362	-0.3958	1.0000	0.5940	-0.3735	0.2538	0.1775	-0.0169	-0.1486	-0.3011
FTP	-0.1331	-0.2422	0.1601	0.2742	-0.0950	0.5940	1.0000	-0.2632	0.0998	0.1492	-0.1601	-0.0496	-0.4091
REB	0.5702	0.5336	0.4181	0.4385	0.7613	-0.3735	-0.2632	1.0000	-0.2138	-0.0774	0.2614	0.6361	0.6115
AST	-0.6278	-0.5722	0.5288	0.4938	-0.0510	0.2538	0.0998	-0.2138	1.0000	0.7744	0.4372	-0.2885	-0.2437
TOV	-0.4658	-0.5225	0.6165	0.5461	-0.0509	0.1775	0.1492	-0.0774	0.7744	1.0000	0.2055	-0.3230	-0.1121
STL	-0.2411	-0.0597	0.5036	0.5086	0.1104	-0.0169	-0.1601	0.2614	0.4372	0.2055	1.0000	0.2479	0.2527
BLK	0.3981	0.4942	0.3139	0.3854	0.4429	-0.1486	-0.0496	0.6361	-0.2885	-0.3230	0.2479	1.0000	0.4992
PF	0.3765	0.3276	0.2091	0.1845	0.5152	-0.3011	-0.4091	0.6115	-0.2437	-0.1121	0.2527	0.4992	1.0000

Table 2

Then I began my investigation from simple analysis. This is a correlation matrix between 2 variables. There are some crystal-clear relationships which everyone takes for granted such as Height & Weight and Points & Minutes. Taller people tend to have more weights because they have more muscles and fat. Also, the longer the player play, the more points he is likely to score. Besides these facts, there are two other high correlations we need to observe carefully: REB (Rebound) & BLK (Block) and AST (Assist) & TOV (Turnover).

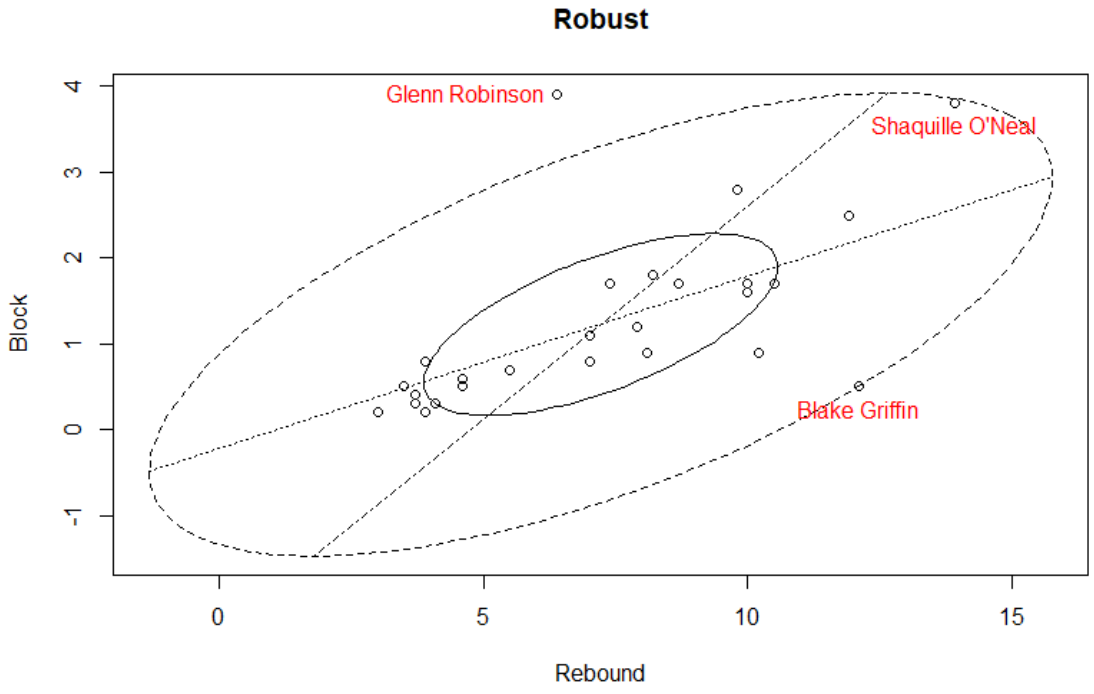

FIGURE 2.1

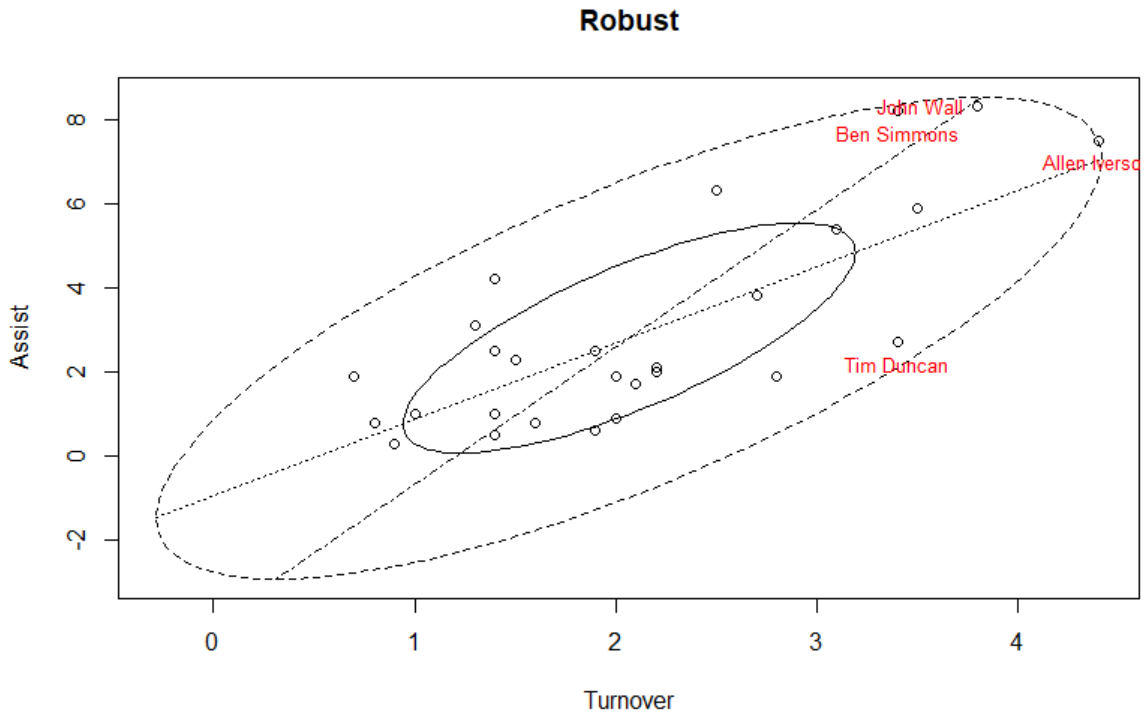

FIGURE 2.2

Both FIGURE 2.1 and FIGURE 2.2 explore the concept of relationship between two variables. In FIGURE 2.1, two important features of big-man aspect are chosen. The correlation between 2 variables is .6361 such that we may observe a linear positive trend. In other words, this can be interpreted as players with high number of rebounds would likely to block more. In FIGURE 2.2, there exists a higher correlation between Assist and Turnover with .7744. This is such as interesting discovery that players who attempt to possess the ball and pass it to other teammates are likely to occur more mistakes than any other style of players. There are some outliers from each plot that we should consider. In FIGURE 2.1, there are 3 outliers: Glenn Robinson, Shaquille O'Neal and Blake Griffin. Robinson was not a huge player, but he was a great blocker before his debut in NBA. In the contrary, Blake Griffin was good at rebound than blocking. O'Neal was both good and outperformed than the 1st draft pick players in his rookie season. In FIGURE 2.2, players indicated with red colors are outliers who used to handle the ball in that season except for Tim Duncan. It can be interpreted as Duncan used to make a lot of mistakes while possessing the ball.

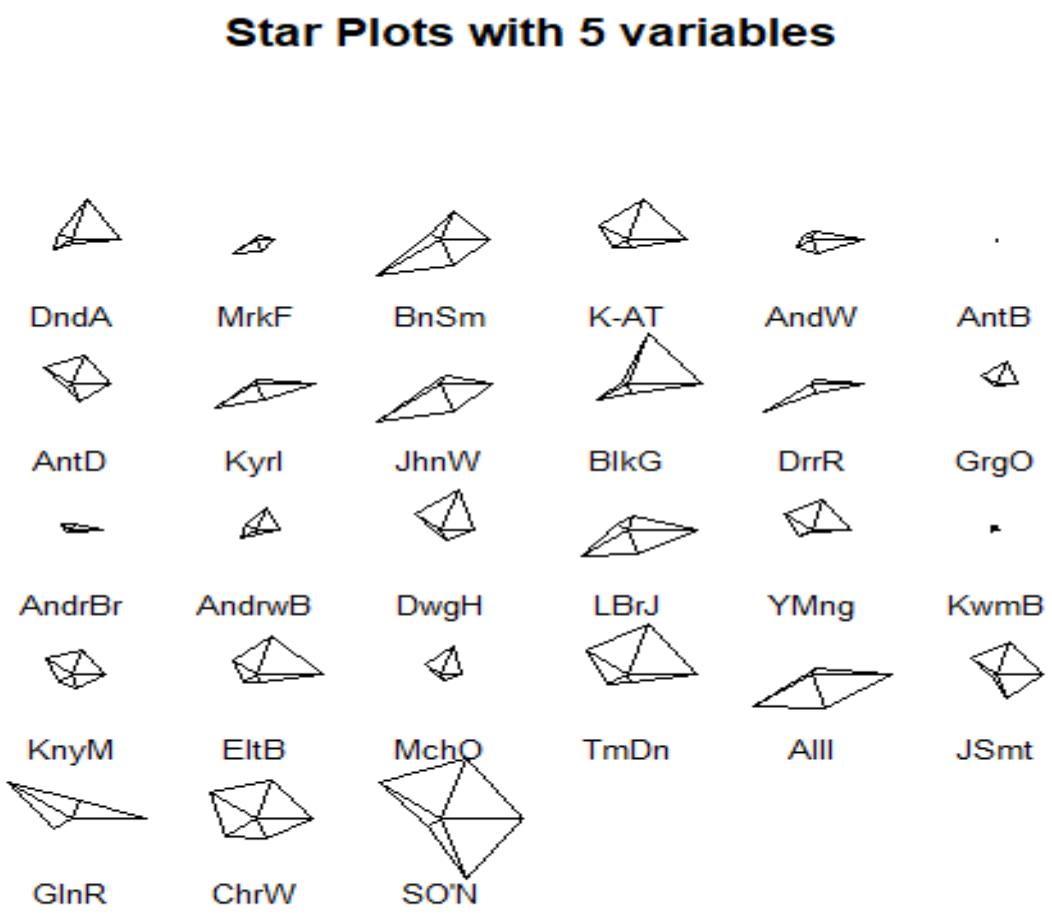

FIGURE 3

FIGURE 3 is basically displaying 5 different from top-right corner in clockwise order: Rebound, Points, Steal, Assist and Block. These variables are 5 representatives when it comes to personal statistics. Based on the FIGURE 2.1, NBA teams and fans expect high frequency of rebounds and blocks from the big-man while fast and small players tend to assist and steal the ball more often. Players with large shape indicates his versatility in many variables. In this plot, Shaquille O'Neal has the largest shape having great personal statistics from all variables except for Assist. In contrast, we may assume that Anthony Bennett and Kwame Brown might be the biggest bust draft picks that disappointed their teams and fans because they did not obtain any impressive statistics from 5 significant factors.

<Principal Component Analysis>

Eigenvalues

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8	Dim.9	Dim.10	Dim.11	Dim.12	Dim.13
Variance	4.523	3.667	1.473	0.940	0.641	0.449	0.411	0.311	0.257	0.127	0.086	0.077	0.038
% of var.	34.790	28.210	11.332	7.233	4.931	3.457	3.161	2.390	1.973	0.976	0.660	0.592	0.295
Cumulative % of var.	34.790	63.000	74.331	81.564	86.495	89.952	93.113	95.503	97.476	98.452	99.112	99.705	100.000

TABLE 3

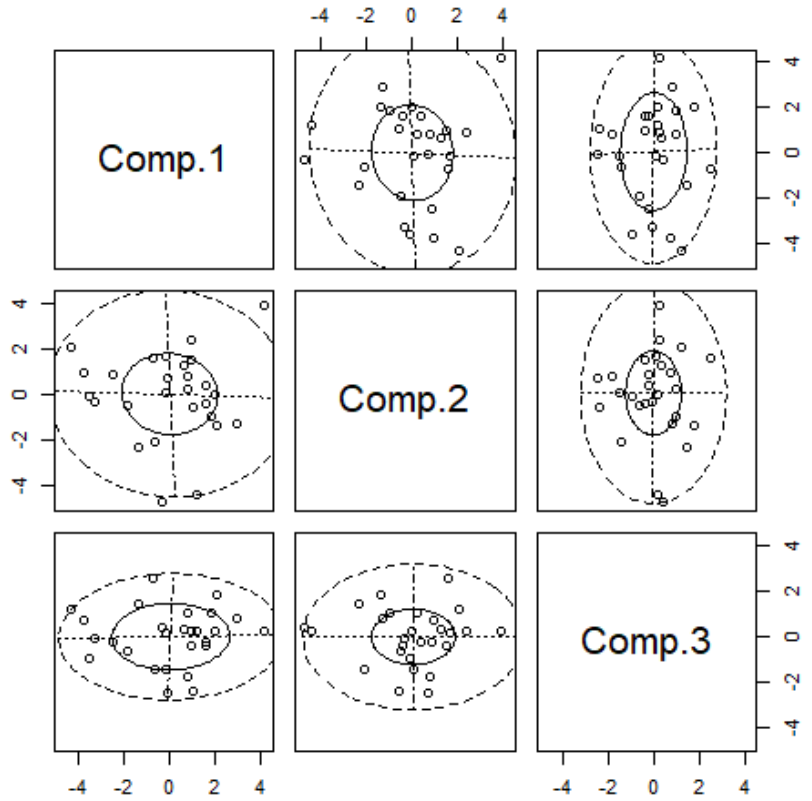

FIGURE 4.1

We briefly discovered the result that O'Neal was an invaluable player from the rookie season. Further investigation of player performance will be analyzed. Beforehand, we will compare and find any existing correlations among variables and transform them into a small number of uncorrelated variables. In TABLE 3, the first 3 principal components account for 74.3% of the variability in the data. In general, it is recommended to observe up to 95% of the cumulative variability but this makes our analysis difficult and hard to interpret. Thus, we will cover the first three components and experiment further analysis with different methods. The given weights for

the first three components as give below:

The 1st principal component

Z_1 = -.403(Height) -.403(Weight) + .104(Minutes) - .278(FGP) + .266(X3PP) +.192(FTP) -.333(REB) +.331(AST) +.292(TOV) -.286(BLK) - .286(PF)

The 2nd principal component

Z_2 = -.474(Minutes) - .474(Points) - .269(FGP) -.323(REB) -.263(AST) -.276(TOV) -.338(STL) - .242(BLK) -.225(PF)

The 3rd principal component

Z_3 = .205(Height) +.109(Weight) + .204(Points) + .504(X3PP) + .667(FTP) -.200(AST) -.142(TOV) -.213(STL) + .258(BLK) - .160(PF)

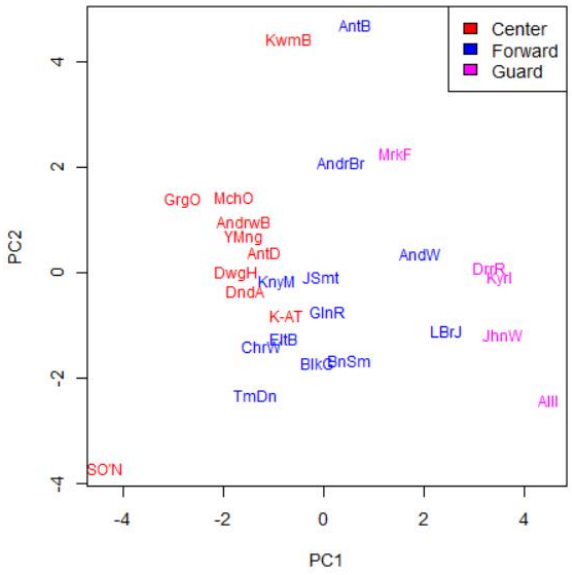

FIGURE 4.2: PC1 & PC2

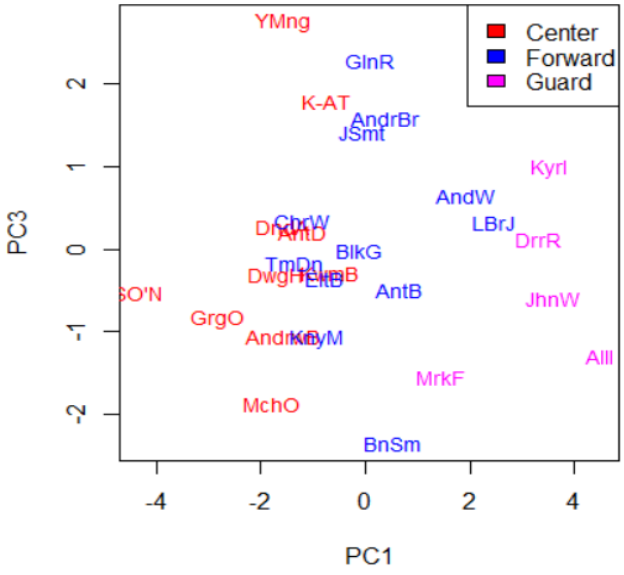

FIGURE 4.3: PC1 & PC3

Both are the PCA plots in first 3 principal components colored by positions. These plots infer some results that we already knew but demonstrate the actual relationship among similar variables. In FIGURE 4.2, Height and Weight have the largest weights in negative coefficients and there is a propensity that Centers (Taller players) are on the left side of the plot while Guards (Smaller players) are on the right. In case of PC2, players with less Minutes & Points have large positive influence such as Anthony Bennett and Kwame Brown. There is no certain pattern with position, but we can interpret it as Guard players scored high points and played longer.

Looking at the FIGURE 4.3, it is again hard to find a pattern of position with X3PP and FTP. However, it tells that Yao Ming & Glenn Robinson had great percentage of shooting success while Ben Simmons was awful. One thing to note is that Yao Ming rarely shot 3-point but had high percentage of Free Throw in rookie season.

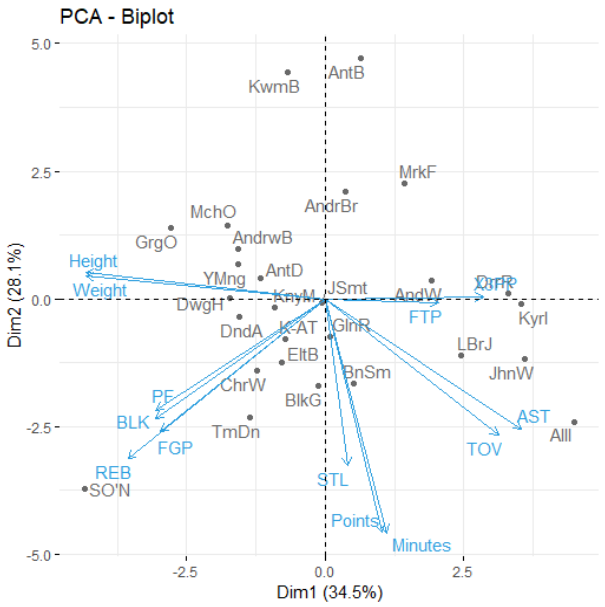

FIGURE 5.1: Biplot With O'Neal

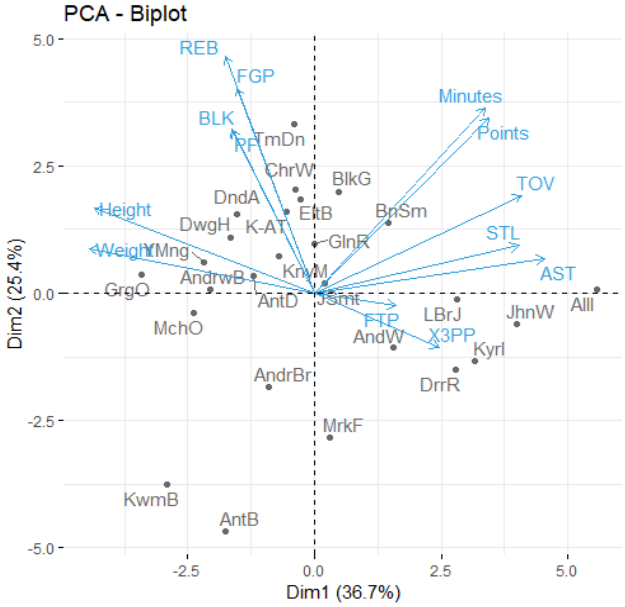

FIGURE 5.2: Biplot Without O'Neal

While I was doing the investigation with Principal Component Analysis, I was being curious of how much the outlier is affecting the change of coefficient weights. Then, I ran again the PCA without Shaquille O'Neal and observed the result of biplots which provide great relationship among variables. Surprisingly, the variance proportion from the first 2 components have been decreased from 62.6% to 62.1. This can be interpreted as the correlation between height and weight has been reduced while most of the variables tend to maintain same direction (correlation). That is, REB, FGP, BLK & PF have correlation which appear to be high among big-man players. STL, AST and TOV have correlation which are high among small players. Players that are highly related to each variable are marked close to the variable they belong.

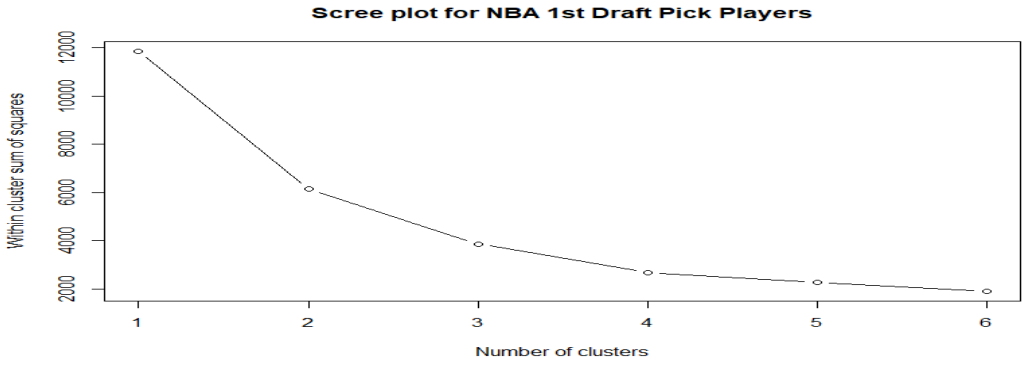

FIGURE 6.1

Then I thought I need another way to cluster my players other than the position. Position is just a mere notation referring to their formal style of play and there must be some other way to cluster players based on their personal statistics. I decided to use K-means clustering because this method is widely used for the most time and simple. Based on the scree plot, the optimum number of clusters are 4 because most of sum of squares are covered by 4.

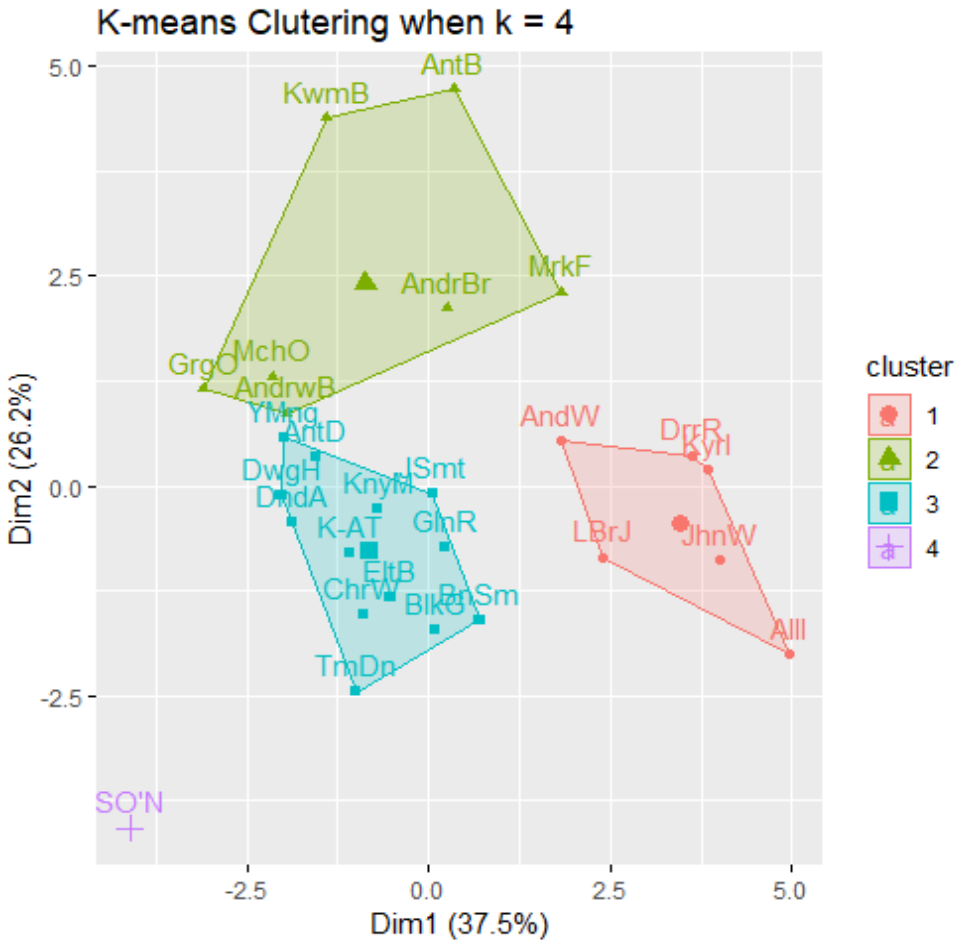

FIGURE 6.2

FIGURE 6.2 is the output of K-means clustering with 10 random sets. From this output, we can once again identify the outstanding performance of Shaquille O'Neal who only comprises of one cluster by himself. There is no contradiction that no other players made striking impression than him. Cluster 1 is a group of ball handlers focusing who possess the ball often and find chance to pass the ball to others otherwise he scores by himself. In Cluster 3, they are mostly tall and proficient in big-man style of play. On the other hand, Cluster 4 has many bad statistics regards to PC2 coefficients. In other words, there are considered to be the bust picks of the season.

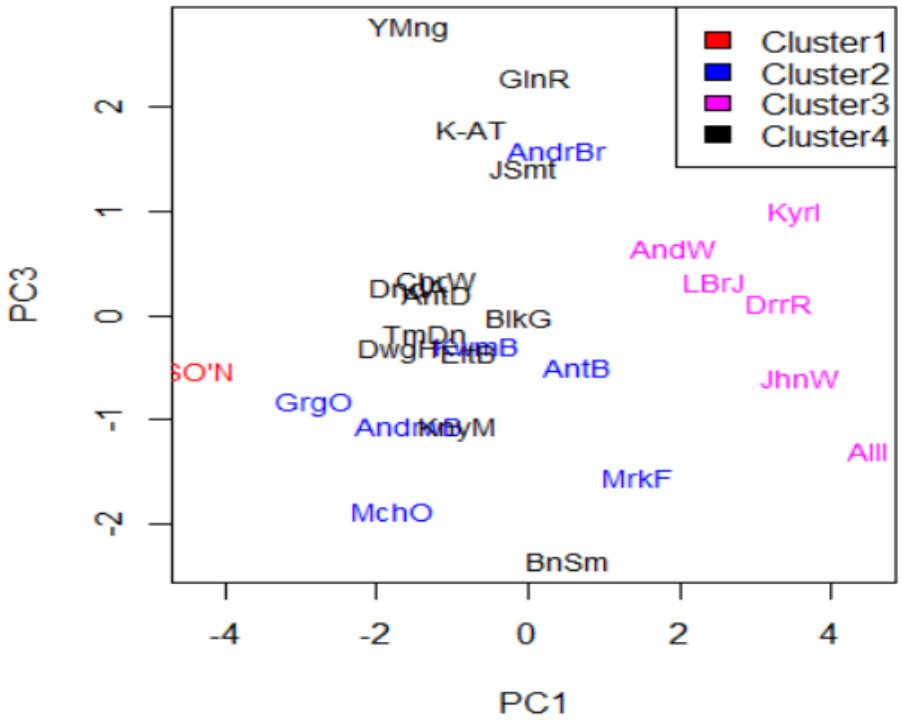

FIGURE 6.3

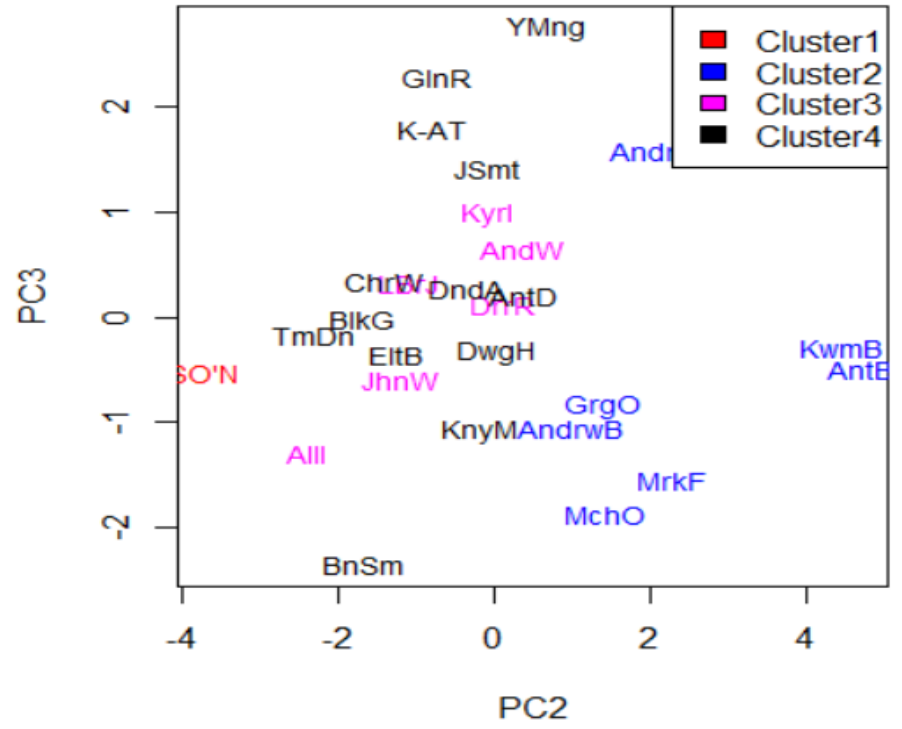

FIGURE 6.4

I added both FIGURE 6.3 & FIGURE 6.4 for additional explanation. In my analysis, there were not as useful as the FIGURE 6.2 because these Clusters are now spread apart and making much harder to interpret. In other words, PC 1 & PC 2 are providing much of the story.

Conclusions

There were a lot of limitations in this investigation because we are only measuring the personal statistics without considering team statistics. In other words, personal statistics cannot guarantee team's win because basketball is a teamplay sports. We also need to consider defensive rate, teamwork and other potential factors to determine how worth the player was. Furthermore, if the player was on the team that has many great players, they have high probability of win, and this correlates to better personal statistics.

In addition, each player had different length of minutes played. There are players who get exhausted easily while others play better as they play longer. The players who are injury prone or cannot maintain their healthy condition during the season will have worse statistics than the others.

The final conclusion from the analysis is that the 1st season of Shaquille O'Neal was superb than anyone else. He was named the 1993 NBA Rookie of the Year and became the first rookie to be voted an All-Star starter since Michael Jordan in 1985 (NBA.com Staff, 1993 NBA All-Star recap). Players from Cluster 1 in FIGURE 6.2 were also outstanding that all the players except for John Wall won the NBA Rookie of the Year on that season. Also, they possessed the ball long amount of time by looking at their number of assists and turnovers. Therefore, we can say that 1st draft pick players having great ability with ball-handling skills are likely to win the Rookie of the Year. Cluster 3 only contains about half of the award winners. What is more astonishing is that players from Cluster 2 failed to win the Rookie of the Year! They are all remembered as the bust 1st draft pick players giving a lot of disappointment for both fans and their teams.

Appendix

Packages Installation

```
if (!require("MVA")){
    install.packages("MVA")
    require("MVA")
    }
if (!require("ca")){
    install.packages("ca")
    require("ca")
    }
if (!require("FactoMineR")){
    install.packages("FactoMineR")
    require("FactoMineR")
    }
if (!require("ggrepel")){
    install.packages("ggrepel")
    require("ggrepel")
    }
if (!require("factoextra")){
    install.packages("factoextra")
    require("factoextra")
    }
if (!require("cluster")){
    install.packages("cluster")
    require("cluster")
    }
```

FIGURE 2.1 & 2.2

```
b <- players[,c("REB","BLK")]
lab2 <- c("Shaquille O'Neal","Glenn Robinson","Blake Griffin")
outliers2 <- match(lab2,players$Player)
bvbox(b,main="Robust",xlab="Rebound",ylab="Block",method="robust")
text(b$REB[outliers2], b$BLK[outliers2], labels = lab2,col = c("red"), cex = 1, pos = c(1,2,1,1))

# Assist & Turnover
# To find out whether ball-handlers tend to make more mistakes
c <- players[,c("TOV","AST")]
lab3 <- c("Ben Simmons","John Wall","Allen Iverson","Tim Duncan")
outliers3 <- match(lab3, players$Player)
bvbox(c,main="Robust",xlab="Turnover",ylab="Assist",method="robust")
text(c$TOV[outliers3],c$AST[outliers3], labels = lab3,col = c("red"), cex = .8, pos = c(1,2,1,1))
```

FIGURE 3

```
# Star plots using Points, Rebound, Block, Assist and Steal as variables
stars(scale(NBA[,c("Points","REB","BLK","AST","STL")]), nrow=5, ncol=6, main = "Star Plots with 5 variables")
```

FIGURE 4.1

```
# Pairwise plots of PCA
r <- range(pca$scores[,1:3])
pairs(pca$scores[,1:3], xlim=r, ylim=r,
      panel = function(x,y,...){
          bvbox(cbind(x,y), add = TRUE)
      })
```

FIGURE 4.2 & 4.3

```
# Plot in 1st two principal components based on their positions
z1 <- pca$scores[,1]
z2 <- pca$scores[,2]
NBA$Position <- as.numeric(players$Position)
names <- rownames(NBA)
plot(z1,z2,xlab="PC1",ylab="PC2",type="n")
Center <- NBA$Position == "1"
text(z1[Center],z2[Center],labels = names[Center],col = "red", cex = .9)
Forward <- NBA$Position == "2"
text(z1[Forward],z2[Forward],labels= names[Forward],col="blue", cex = .9)
Guard <- NBA$Position == "3"
text(z1[Guard],z2[Guard],labels= names[Guard],col="magenta", cex = .9)
colors <- c("red","blue","magenta")
leg.txt <- c("Center","Forward","Guard")
legend("topright",legend=leg.txt,fill=colors)
```

```
z1 <- pca$scores[,1]
z3 <- pca$scores[,3]
plot(z1,z3,xlab="PC1",ylab="PC3",type="n")
Center <- NBA$Position == "1"
text(z1[Center],z3[Center],labels = names[Center],col = "red", cex = .9)
Forward <- NBA$Position == "2"
text(z1[Forward],z3[Forward],labels= names[Forward],col="blue" , cex = .9)
Guard <- NBA$Position == "3"
text(z1[Guard],z3[Guard],labels= names[Guard],col="magenta", cex = .9)
colors <- c("red","blue","magenta")
leg.txt <- c("Center","Forward","Guard")
legend("topright",legend=leg.txt,fill=colors)
```

FIGURE 5.1 & 5.2

```
# Biplot with comparisons
fviz_pca_biplot(pca, repel = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
fviz_pca_biplot(pca.revised, repel = TRUE,
                col.var = "#2E9FDF", # Variables color
                col.ind = "#696969"  # Individuals color
                )
```

FIGURE 6.1

```
# K-means Clustering
n <- dim(NBA)[1]; k <- 7
wss <- rep(0,k)
xm <- apply(NBA,2,mean)
for(i in 1:n){
wss[1] <- wss[1]+sum((NBA[i,]-xm)^2)
}
for(i in 2:k){
model <- kmeans(NBA,i)
wss[i] <- sum(model$withinss)
}
plot(1:k,wss,type="b",xlab="Number of clusters",
ylab="Within cluster sum of squares",main="Scree plot for NBA 1st Draft Pick Players")
```

FIGURE 6.2

```
means<- 4
km4 <- kmeans(scale(NBA),means, nstart = 10) # No consideration of their position, but classifing them by their
personal stats
c.names <- as.character(players$Player) # clustering without position
for(i in 1:means){print(paste("Cluster",i))
print(c.names[km4$cluster == i])
}
fviz_cluster(km4,data=scale(NBA)) + ggtitle("K-means Clutering when k = 4")
```

FIGURE 6.3 & 6.4

```
z1 <- pca$scores[,1]
z3 <- pca$scores[,3]
plot(z1,z3,xlab="PC1",ylab="PC3",type="n")
cluster1 <- km4$cluster == "1"
text(z1[cluster1], z3[cluster1],labels = names[cluster1],col = "red", cex = .9)
cluster2 <- km4$cluster == "2"
text(z1[cluster2], z3[cluster2],labels= names[cluster2],col="blue", cex = .9)
cluster3 <- km4$cluster == "3"
text(z1[cluster3], z3[cluster3],labels= names[cluster3],col="magenta", cex = .9)
cluster4 <- km4$cluster == "4"
text(z1[cluster4], z3[cluster4], labels = names[cluster4], col = "black", cex = .9)
colors <- c("red","blue","magenta","black")
leg.txt <- c("Cluster1","Cluster2","Cluster3","Cluster4")
legend("topright",legend=leg.txt,fill=colors)

z2 <- pca$scores[,2]
z3 <- pca$scores[,3]
plot(z2,z3,xlab="PC2",ylab="PC3",type="n")
cluster1 <- km4$cluster == "1"
text(z2[cluster1], z3[cluster1],labels = names[cluster1],col = "red", cex = .9)
cluster2 <- km4$cluster == "2"
text(z2[cluster2], z3[cluster2],labels= names[cluster2],col="blue", cex = .9)
cluster3 <- km4$cluster == "3"
text(z2[cluster3], z3[cluster3],labels= names[cluster3],col="magenta", cex = .9)
cluster4 <- km4$cluster == "4"
text(z2[cluster4], z3[cluster4], labels = names[cluster4], col = "black", cex = .9)
colors <- c("red","blue","magenta","black")
leg.txt <- c("Cluster1","Cluster2","Cluster3","Cluster4")
legend("topright",legend=leg.txt,fill=colors)
```

References

John, Allen St. “The NBA's New Lottery Will Merely Encourage More Teams To Tank Differently.” *Forbes*, Forbes Magazine, 1 Oct. 2017, www.forbes.com/sites/allenstjohn/2017/09/30/the-nbas-new-lottery-will-merely-encourage-more-teams-to-tank-differently/#3318567a2a1c.

NBA.com Staff. “1993 NBA All-Star Recap.” *NBA HISTORY*, NBA.com, 24 Aug. 2017, www.nba.com/history/all-star/1993.