Final Report

Kecheng Liang

12/12/2018

# Introduction

Basketball is one of the most popular sports in the world. National Basketball Association(NBA) is the largest league for this sport. There are lots of interesting data in the game. In basketball, an assist is attributed to a player who passes the ball to a teammate in a way that leads to a score by field goal, meaning that he or she was “assisting” in the basket. Because an assist can be scored for the passer even if the player who receives the pass makes a basket after dribbling the ball. In some situtation it becomes hard to define whether it is a assist. We may think that player who play in the home game are more easily get the tenth assist when the player already have nine assists. Same thing may happen in rebound. I want to do the analysis whether it really happens. Also I will do other interesting graph to show the miracle NBA data.

# Data

The data is downloaded from website and the data from 2012 to 2018. It is well organized with 51 variables and I removed some useless variables.

Initial <- read.csv("2012-18\_playerBoxScore.csv")  
Initial$playAST <- as.numeric(Initial$playAST)  
Initial$playTRB <- as.numeric(as.character(Initial$playTRB))  
Initial <- Initial[,c(-1,-2,-3,-4,-5,-8,-11,-12,-13,-14,-15,-16,-17,-24,-48,-49,-50,-51)]  
away\_data <- filter(Initial,Initial$teamLoc=="Away")  
home\_data <- filter(Initial,Initial$teamLoc=="Home")  
nba\_total <- rbind(away\_data,home\_data)

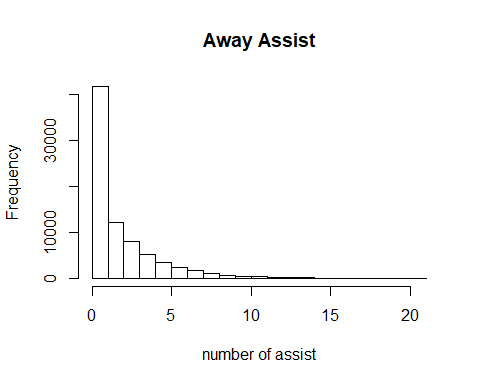
var <- data.frame(Variables = c("teamAbbr","teamConf","teamLoc","teamRslt","playDispNm","playStat","playMin","playPos","playHeight","playWeight","playPTS","playAST","playTO","playSTL","playBLK","playPF","playFGA","playFGM","playFG.","play2PA","play2PM","play2P.","play3PA","play3PM","play3P.","playFTA","playFTM","playFT.","playORB","playDRB","playTRB","opptAbbr","opptConf"),  
 Explanation = c("Abbreviation of team","Identifies conference of team","Identifies whether team is home or visitor","Identifies whether team has won or lost","Player display name","Identifies staring status of player","Player minutes on floor","Player position during game","Player height","Player weight","Points scored by player","Assists made by player","Turnovers made by player","Steals made by player","Blocks made by player","Personal fouls made by player","Field goal attempts made by player","Field goal shots made by player","Field goal percentage made by player","Two point attempts made by player","Two point shots made by player","Two point percentage made by player","Three point attempts made by player","Three point shots made by player","Three point percentage made by player","Free throw attempts made by player","Free throw shots made by player","Free throw percentage made by player","Offensive rebounds made by player","Defensive rebounds made by player","Total rebounds made by player","AAbbreviAtion of opponent","Identifies conference of opponent"))  
kable(var,caption = "Variables explanation")

Variables explanation

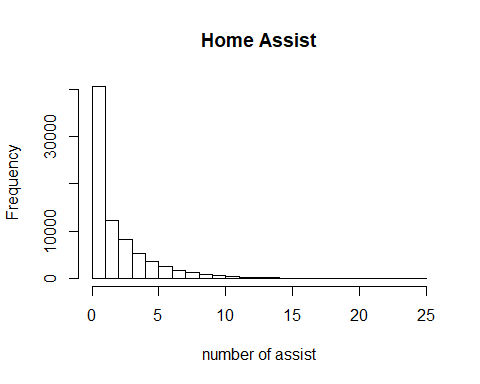
|  |  |
| --- | --- |
| Variables | Explanation |
| teamAbbr | Abbreviation of team |
| teamConf | Identifies conference of team |
| teamLoc | Identifies whether team is home or visitor |
| teamRslt | Identifies whether team has won or lost |
| playDispNm | Player display name |
| playStat | Identifies staring status of player |
| playMin | Player minutes on floor |
| playPos | Player position during game |
| playHeight | Player height |
| playWeight | Player weight |
| playPTS | Points scored by player |
| playAST | Assists made by player |
| playTO | Turnovers made by player |
| playSTL | Steals made by player |
| playBLK | Blocks made by player |
| playPF | Personal fouls made by player |
| playFGA | Field goal attempts made by player |
| playFGM | Field goal shots made by player |
| playFG. | Field goal percentage made by player |
| play2PA | Two point attempts made by player |
| play2PM | Two point shots made by player |
| play2P. | Two point percentage made by player |
| play3PA | Three point attempts made by player |
| play3PM | Three point shots made by player |
| play3P. | Three point percentage made by player |
| playFTA | Free throw attempts made by player |
| playFTM | Free throw shots made by player |
| playFT. | Free throw percentage made by player |
| playORB | Offensive rebounds made by player |
| playDRB | Defensive rebounds made by player |
| playTRB | Total rebounds made by player |
| opptAbbr | AAbbreviAtion of opponent |
| opptConf | Identifies conference of opponent |

# EDA

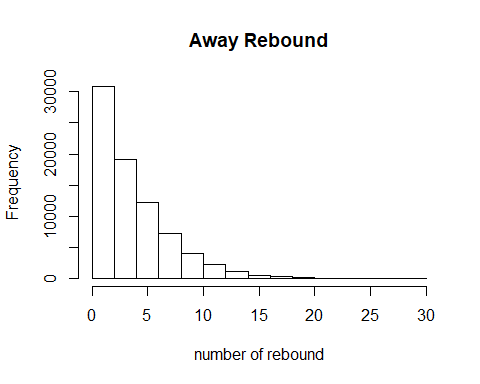
hist(x=away\_data$playAST,main = "Away Assist",xlab = "number of assist")



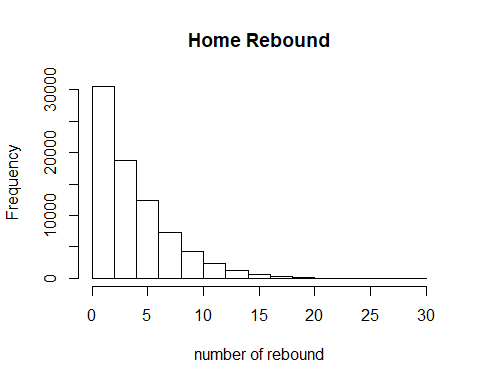
hist(x=home\_data$playAST,main = "Home Assist",xlab = "number of assist")



hist(x=away\_data$playTRB,main = "Away Rebound",xlab = "number of rebound")



hist(x=home\_data$playTRB,main = "Home Rebound",xlab = "number of rebound")



# Chi square test

##   
## Pearson's Chi-squared test  
##   
## data: total\_ast  
## X-squared = 95.549, df = 23, p-value = 8.177e-11

##   
## Pearson's Chi-squared test  
##   
## data: total\_trb  
## X-squared = 64.517, df = 29, p-value = 0.000164