Professional Athletes: Skill Vs. Pay

Charles Fiscella

4/23/2020

## Summary

The problem that I am trying to address with my project is that some professional baseball athletes will get paid a crazy high salary simply because their name may be big and they may be the face of the team they play for. My goal is to analyze certain players from a Professional sport organization, the MLB or Major League Baseball, and see if their pay is based on how good they actually are with their stats and also if they start having a gradual decline with their stats and performance, will their poor performance have an impact on their pay. The main research question that I am trying to address is “Do athletes get paid more simply because of their name and because the fans love them or is it based on their performance?”. I got my data through different online resources that specialize in sports statistics in the MLB. I will be using R to go through and analyze the data, exploring any athletes who may have a large salary but have also possibly had a slight decline from years past and are still getting paid the same salary or may have a bigger salary now. With this problem I will also be trying to see if some rookie athletes who deserve getting paid more will actually get paid more or if the team simply can’t afford to pay them more even if they clearly deserve it. Even if you may not know a lot about sports or baseball, it is important to know how much money certain athletes may make while playing sports. Sometimes the amounts are a huge number and can sometimes shock people when they here the amount. Recently, some sports teams even made record breaking contracts with some players just to make it more appealing to want to join their organization. Reasons like this are why I want to explore the sports pay and stats area.

## Introduction

Professional sports, we all watch them. From soccer to baseball we all have that favorite team or player that we like to watch. However, some of these players are getting a crazy amount of money even if they don’t perform the greatest of their ability. Do players get paid based on their name and how they make a team look? Or do these players get paid based on their statistics and their past accomplishments? These questions will both be explored in my project. I wil be analyzing data from sites that specailize in baseball statistics including pay and roster and using this data to determine which of these questions is true.

## Data

MLB Player statistics was created by fangraphs at <https://www.fangraphs.com/leaders.aspx?pos=all&stats=bat&lg=all&qual=0&type=8&season=2019&month=0&season1=2019&ind=0&team=0&rost=0&age=0&filter=&players=0&startdate=2019-01-01&enddate=2019-12-31>

MLB Payroll information also created by fangraphs at <https://www.fangraphs.com/roster-resource/payroll/>

install.packages('csv')  
install.packages('readxl')  
MLBStats <- read.csv("FanGraphs Leaderboard.csv")

## Methodology

The first steps that I had to take in this project was cleaning FanGraphs Leaderboard csv file and combining it with the payroll data found on the website. The csv file had too many unncessary players stats that were listed in the file and it was very hard to read and understand for someone who may not know baseball stats that well. The goal is to make the data easy to read and understand, so I used excel to change the headers so that they may be easliy understood. I also added the payroll information for certain players to the file so that people can easily see how much each players salary exactly is. After moving the data into one data file, I brought the data into the R programs. I decided to install the “baseballDBR Package” as it is the perfect package for calculating BaseBall Statistics. Once I installed the package, I summarized all the statistics data in the csv file so that we see the average of each statistic. Then i added all the statistics together. This makes it easier to view that statistics since there are so many together.

hsb2 <- read.table("FanGraphs Leaderboard.csv", header=T, sep=",")  
attach(hsb2)  
  
hsb2[1:50, ]  
  
mean(hsb2$Games)  
mean(hsb2$Plate.Appearances)  
mean(hsb2$Home.Runs)  
mean(hsb2$Runs)  
mean(hsb2$Runs.Batted.In)  
mean(hsb2$Stolen.Bases)  
mean(hsb2$Walk..)  
mean(hsb2$Strikeout..)  
mean(hsb2$Batting.Average)  
mean(hsb2$On.Base.Percentage)  
mean(hsb2$Slugging.Percentage)  
  
sum(hsb2$Games)  
sum(hsb2$Plate.Appearances)  
sum(hsb2$Home.Runs)  
sum(hsb2$Runs)  
sum(hsb2$Runs.Batted.In)  
sum(hsb2$Stolen.Bases)  
sum(hsb2$Walk..)  
sum(hsb2$Strikeout..)  
sum(hsb2$Batting.Average)  
sum(hsb2$On.Base.Percentage)  
sum(hsb2$Slugging.Percentage)

plot(hsb2$Plate.Appearances,hsb2$Batting.Average, main="Games Played Vs. Batting Average ", xlab = "Plate Appearances", ylab = "Batting Average")

Once the data was plotted, you can see that the more plate appearances that a player has, the higher their batting aveage is. Of course you would expect this since if they don’t appear batting at the plate or in the lineup, they wouldn’t be able to accumulate a high batting average. Batting average is calculated by how many hits a player gets and how many times per game they get hits.

##Results

##Plot players Batting Average   
plot(hsb2$ï..Name,hsb2$Batting.Average, main = "Player Averages", xlab = "Name", ylab = "" )  
  
##Plot players Runs Batted In  
plot(hsb2$ï..Name,hsb2$Runs.Batted.In, main = "Runs Batted In", xlab = "Name", ylab = "" )  
  
##Plot Players On Base Percentage  
plot(hsb2$ï..Name,hsb2$On.Base.Percentage, main = "On Base Percentage", xlab = "Name", ylab = "" )  
  
##Plot Players Average Salary  
ggplot(data=hsb2, aes(x=ï..Name, y=Average.Salary)) + geom\_bar(stat="identity")

I then plotted out all the different factors that can conrtibute to players getting paid more or less, it allows you to see their stats and then compare them to the average salary that they make in a year.

As plotted out on multiple charts, you are able to see which player gets paid the most and which players don’t get paid as much

## Implications

In some cases, certain players that are very good statistics wise do not get paid what they deserve. Peter Alonso of the New York Mets broke the rookie home-run record and had an outstanding rookie season in 2019. However, he only gets paid the league minimum. Now in 2020 it is unknown what he will get paid since he had such amazing numbers the past year, it depeneds on what the indiviual organizations say about them. They may not be able to afford to pay players what they deserve, which is why some players choose to be traded the next year of eligibility.

## Conclusion

In my project, I have determined that there are some players who have a great year who the following year will not get paid what they deserved. There are also some players who don’t have as great as a year and still get paid a lot becauser the individual organization can afford to pay them. It comes down to if the individual team organization has money and can afford any big name or popular players, they are willing to hand over the money to get them, if the team doesn’t have as much money, they are more likely to just keep the player at their current salary and this may cause the player to want to be traded to another team in the future. In some cases if a player has a bad year and a team doesn’t like how they’re performing, they may buy that person out of their contract (pay the remaining amount of the players contract) in order to trade them to another team to make room on the roster for a younger, better player.

## References

FanGraphs. (2019). Major League Leaderboards " 2018 " Batters " Dashboard: FanGraphs Baseball. Retrieved May 03, 2020, from <https://www.fangraphs.com/leaders.aspx?pos=all>

FanGraphs. (2019). 2020 Payroll: RosterResource. Retrieved May 03, 2020, from <https://www.fangraphs.com/roster-resource/payroll/>