Andrew Ivanov

December 17th 2019

NBA Data Analysis Project

Executive Summary

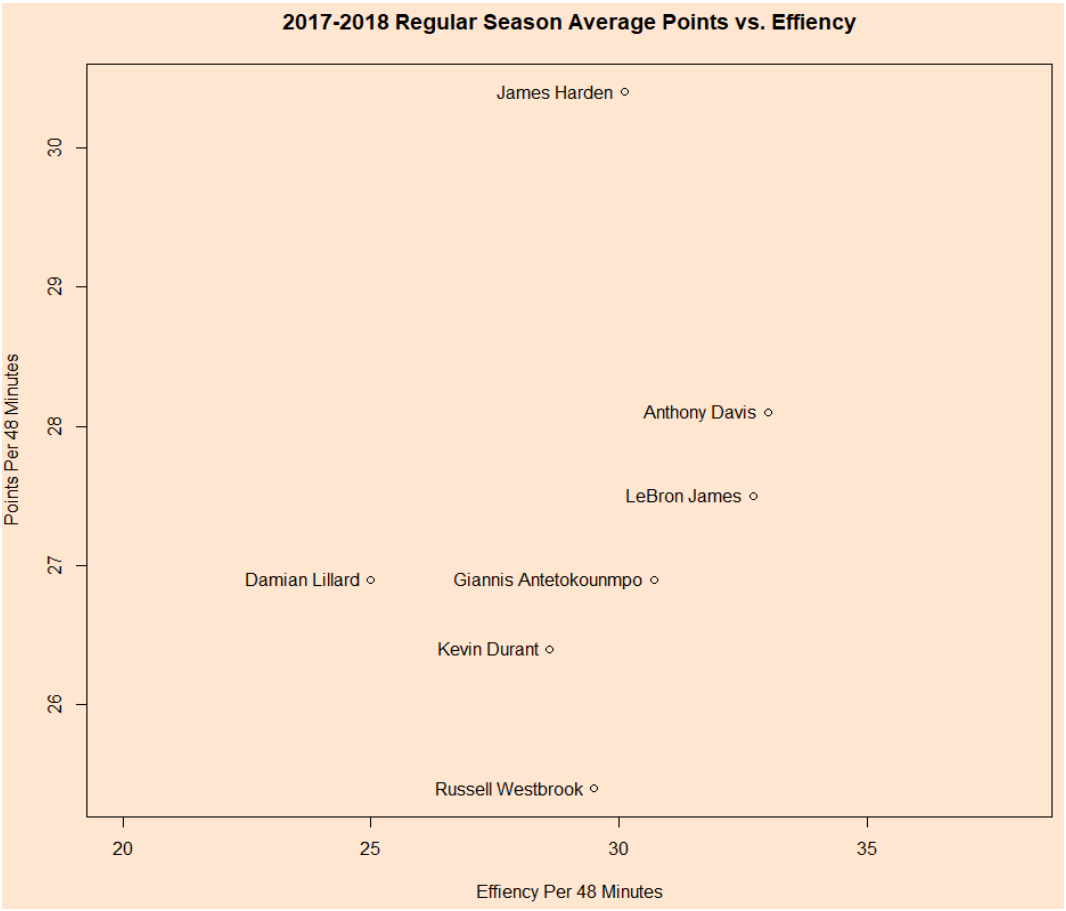
After analyzing the 2017-2018 and 2018-2019 I have a few interesting observations. A majority of players who perform well in regular season see a decline in efficiency and points per game in the playoffs. There was one player each season who saw an increase in efficiency and points per game in the postseason, and both players reached the NBA finals that year. In the 2017-2018 season, LeBron James averaged 34 points per game in the playoffs and 27.5 in the regular season. In the 2018-2019 season, Kawhi Leonard averaged 30.5 points per game in the playoffs and had an efficiency rating of 32. During the regular season he averaged 26.6 points per game and his efficiency was 26. I also looked at the highest scoring players of the last two NBA seasons and how they scored a majority of their points. In the 2017-2018 season, LeBron James led the NBA in scoring with a majority of his points coming from two-point field goals. In the 2018-2019 season, James Harden led the NBA in scoring. He led the NBA in three-point field goals made and free throws made for the second straight year.

Data gathering / wrangling / description

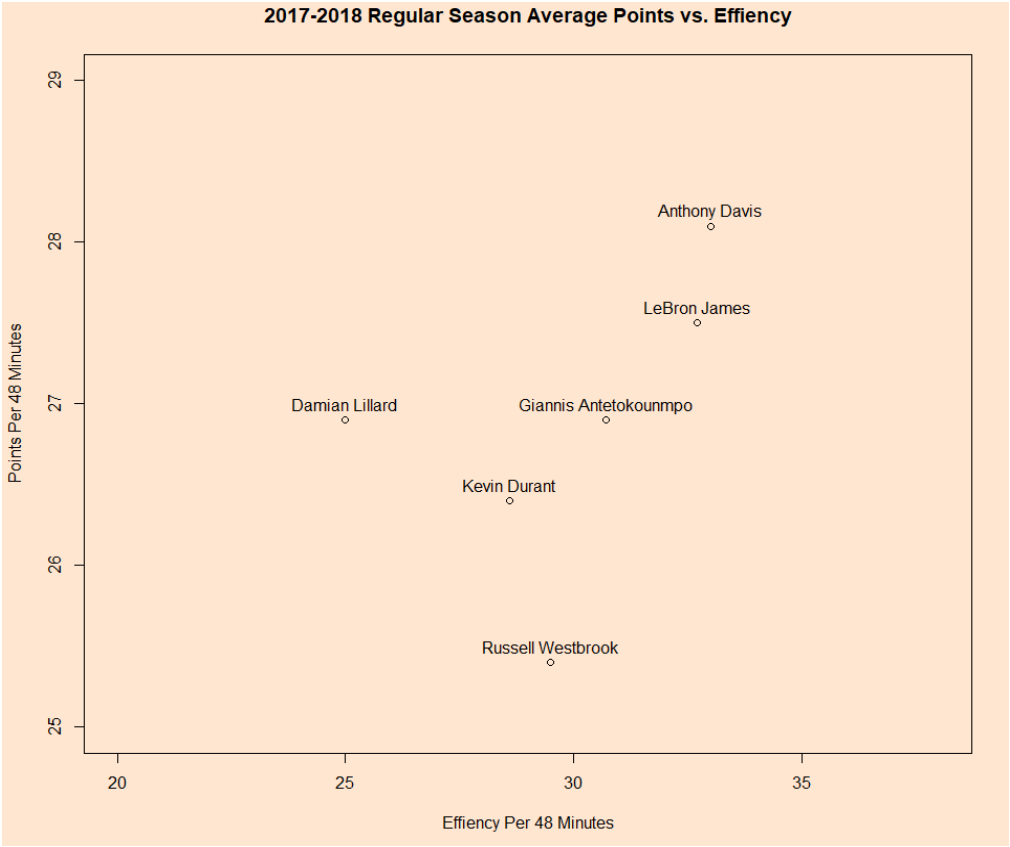
To use the nbastatR package in R, first I had to install the "remotes" package and run the install_github function to load nbastatR into RStudio. In nbastatR I was mostly using the metrics_leaders() function which provided me with statistics from the preseason, regular season, and playoffs for a given year. After running the metrics_leader function I had a matrix with 33 variables and rows equal to the number of players in the NBA in the season I specified. I was only interested in the highest scoring players in a particular season, so I used to dplyr filter function to select players that averaged more than 25 points per game. Afterwards, I used the select function to reduce the number of variables that I would be working with. Once I had the variables and players for the regular season data, I selected the name of the player and the playerID and saved them in the PlayersRegSeason variable. Since I was comparing regular season to playoff performance, I wanted to save the regular season players I was working with to make it easier to find them in the playoff dataset. Using a semi_join I join the PlayersRegSeason variable with the playoff dataset, so the dataset would only be the players that I had previously selected from the regular season. For the second part of the project, I had most of the data I needed to create the bar plots except for the total points from 2-point field goals and 3-point field goals. I already had the number of 2-point field goals and 3-point field goals made, so I created two new variables TwoPointFGS and ThreePointFGS. TwoPointFGS was equal to 2-point field goals made times two and ThreePointFGS was equal to 3-point field goals made times three.

Data Analysis

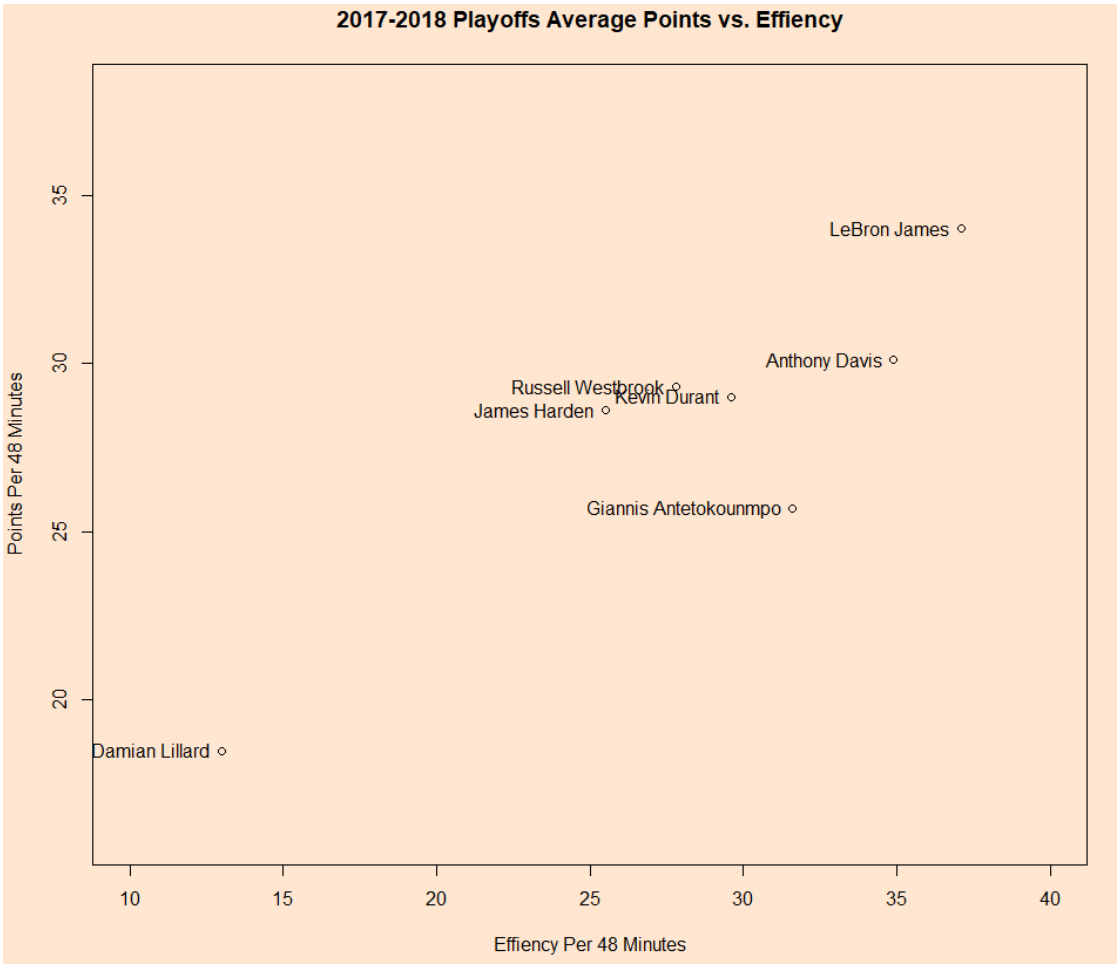
The first question I had that I wanted to explore in the dataset was whether the performance of players changes in the playoffs comparted to the regular season? Do players step up in the post season when it is win or go home? I plotted points score versus efficiency in a scatterplot to see which players were most efficient in the regular season and which players scored the most points in the regular season.



In the 2017-2018 regular season, James Harden led the NBA with an average of 30.4 points per game per 48 minutes. However, Harden was the 4th most efficient player behind Anthony Davis, LeBron James, and Giannis Antetokounmpo. It looked to me like the James Harden datapoint was an outlier, so I also created a scatterplot that excluded James Harden.

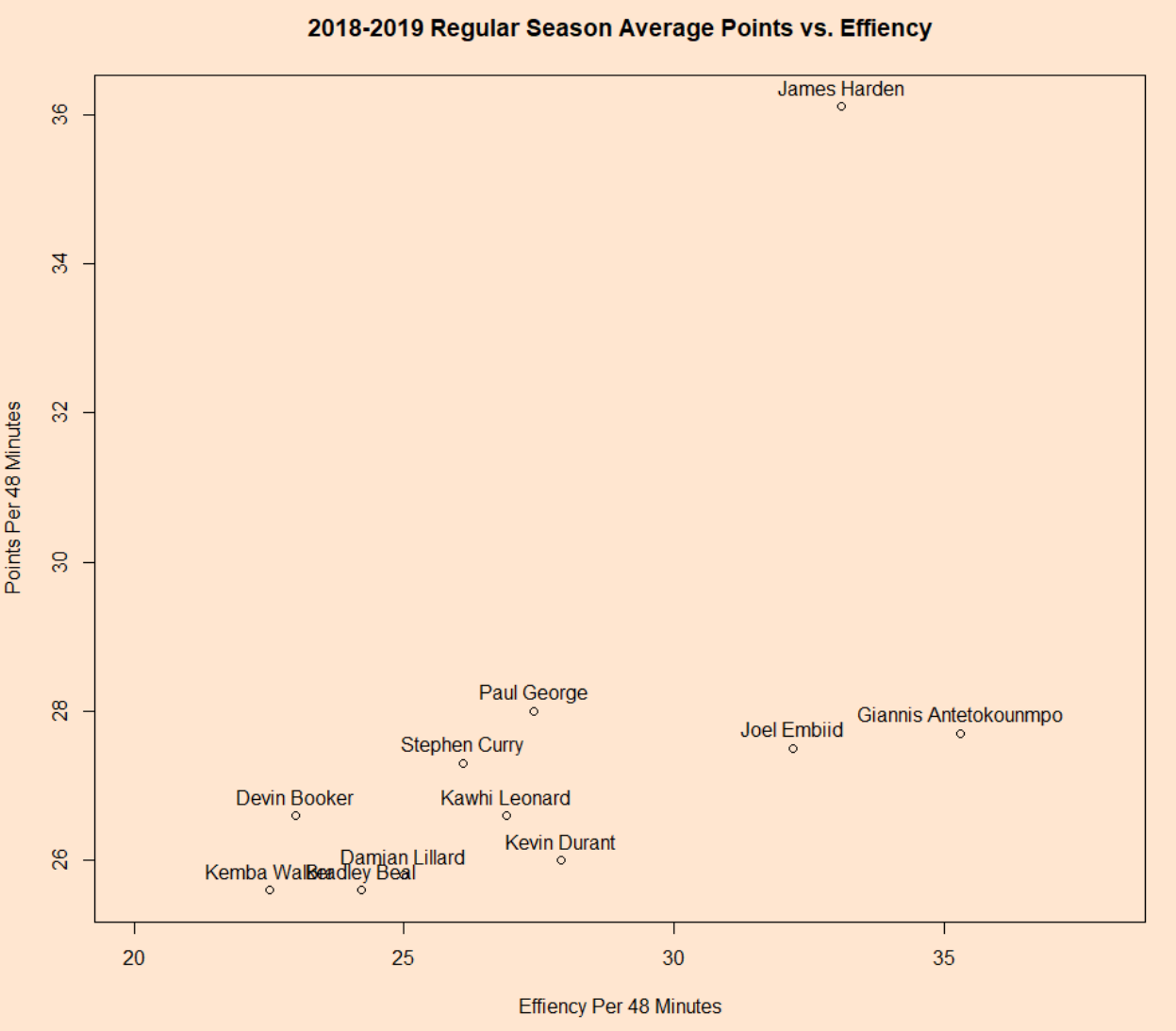


Afterwards I plotted the average points per game versus efficiency in the playoffs to see if the same players that were performing well in the regular season were also putting up similar numbers in the playoffs.

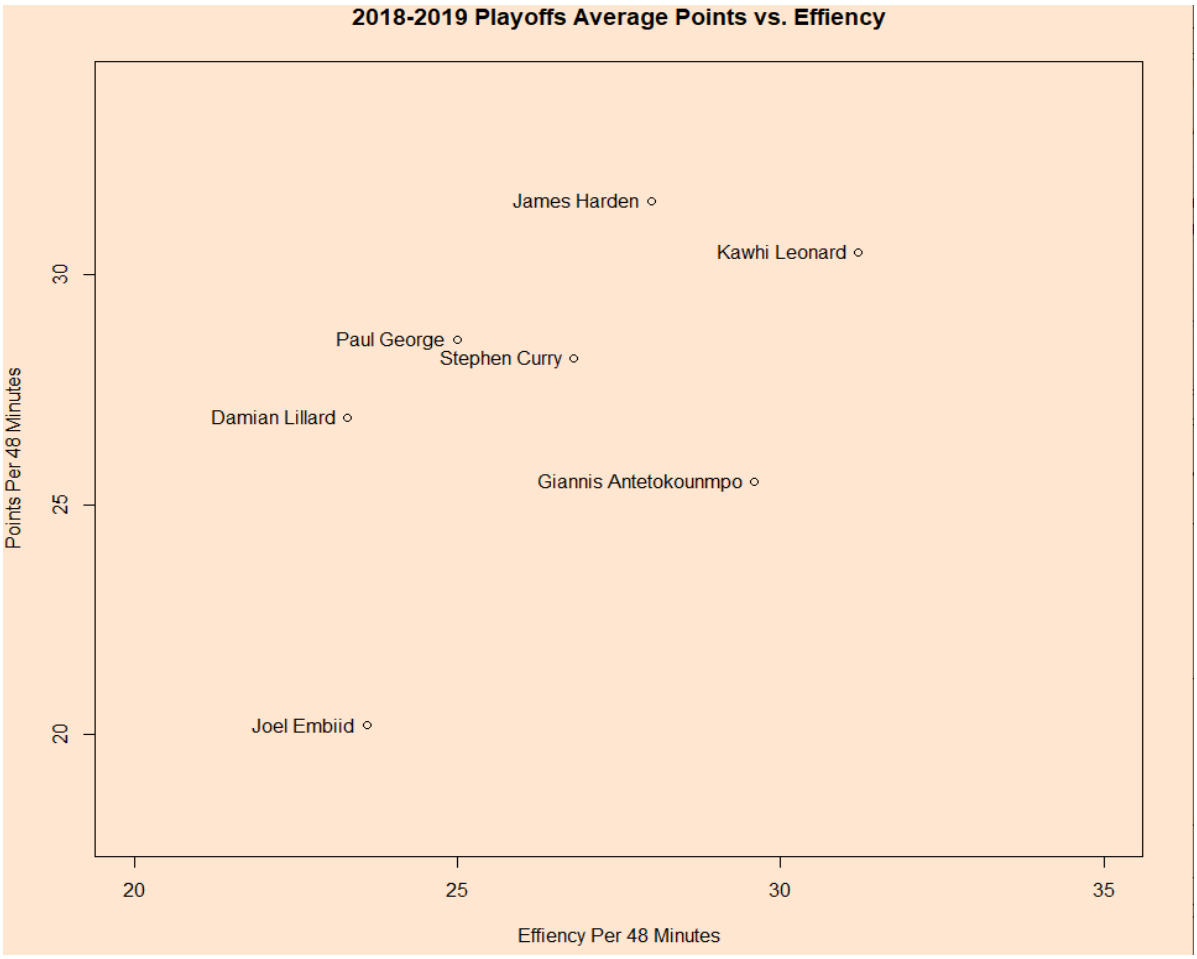


From the data we can see that Damian Lillard underperformed in the playoffs. He averaged less than 20 points per game in the postseason and his efficency per 48 minutes was less than 15. Compared to his numbers in the regular season where he averaged close to 27 points per game and had an efficency per 48 minutes of 25, he underperformed in both categories. When looking closer, the Portland Trailblazers lost 4-0 in the first round of the playoffs and Damian Lillard was the second leading scorer for the Trailblazers in all 4 games. We can also see that LeBron James stepped up in the playoffs and overperformed in the playoffs compared to his regular season numbers. James averaged 34 points per game, a 23.6% increase from his regular season average of 27.5. Other players in the dataset; James Harden, Russel Westbrok, and Giannis Atetokounmp also underperformed compared to their regular season statistics. However, Kevin Durant and Anthony Davis saw slight increases in their points per game and efficiency per 48 minutes.

The next step was to examine the same data for the 2018-2019 season.



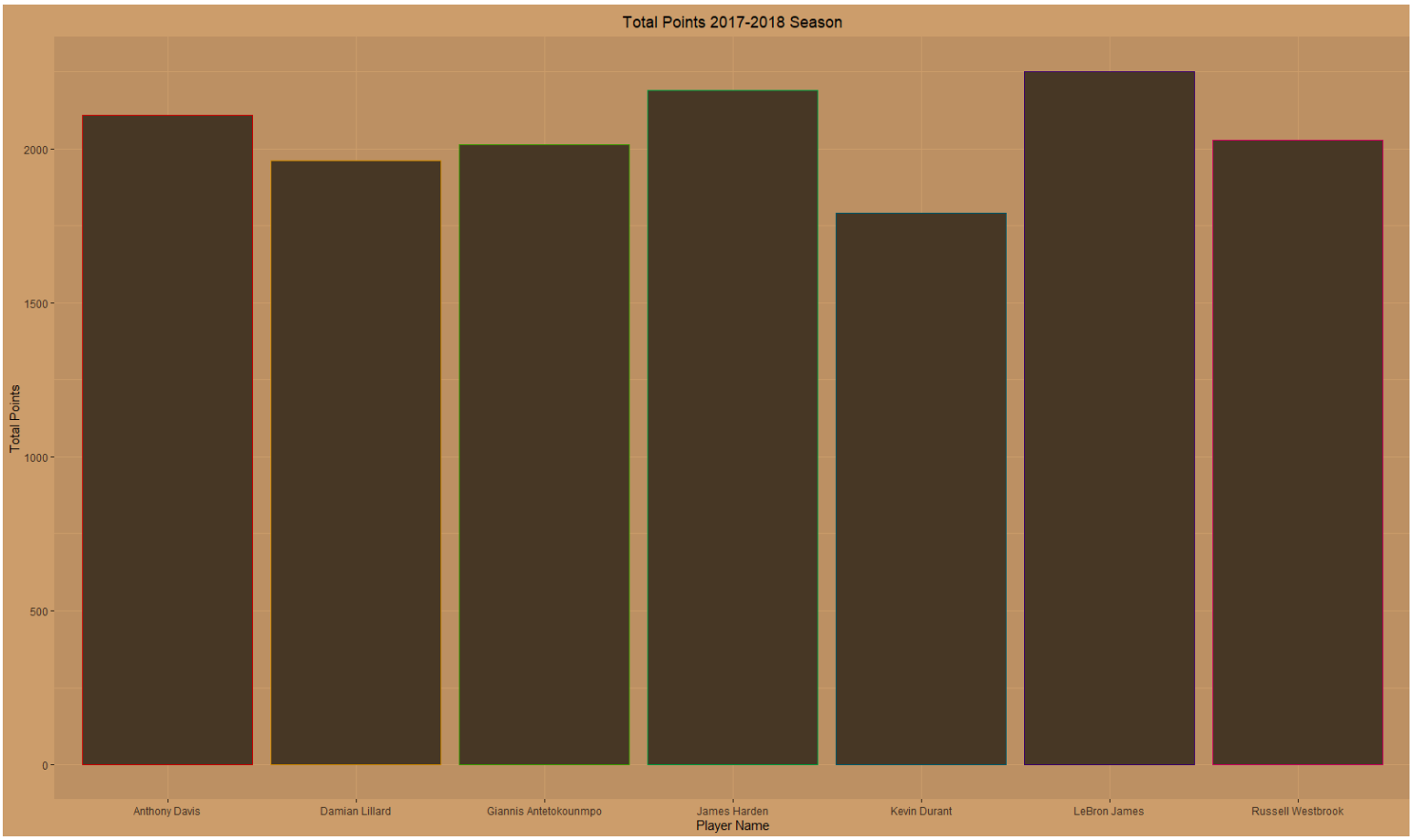
Once again James Harden was an outlier in the data. He scored 36.1 points per game in the 2018-2019 season, Paul George was behind Harden with 28 points per game.



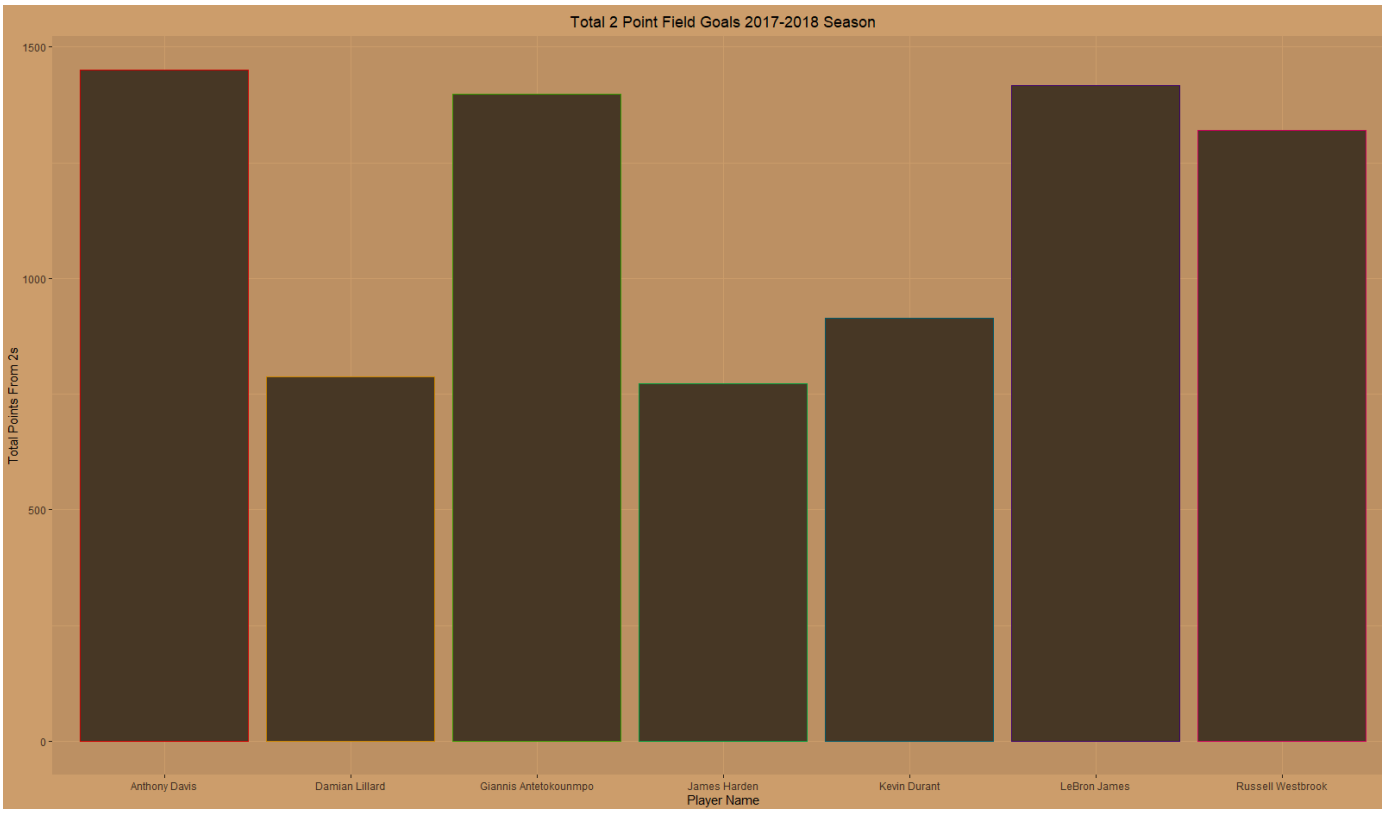
For the 2018-2019 playoffs, there are leading scorers from the regular season missing in the playoff dataset. This is because Bradley Beal, Kemba Walker, and Devin Booker's teams did not make the playoffs. Kevin Durant is the 4th player missing and that is because he was injured throughout the playoffs. Once again, the majority of the players in the dataset underperformed compared to the regular season. Joel Embiid struggled in the playoffs. In the regular season he averaged around 27 points per game and had an efficiency of 33. In the playoffs Embiid averaged 20 points per game and had an efficiency of 24. When we look at Philadelphia's playoff schedule, the 76ers played the Toronto Raptors in the 2nd round of the playoff. They were the best defensive team and managed to slow down Embiid. James Harden also struggled in the playoffs. His points per game dropped from 36 in the regular season to 31.6 in the playoffs. His efficiency also declined from 34 to 27. In the 2018-2019 playoffs, the only player to have a significant increase compared to his regular season statistics was Kawhi Leonard. In the regular season he averaged 26.6 points and had an efficiency of 26. In the playoffs he improved on both numbers. His average points per game was 30.5 and his efficiency was the highest of star players at 32.

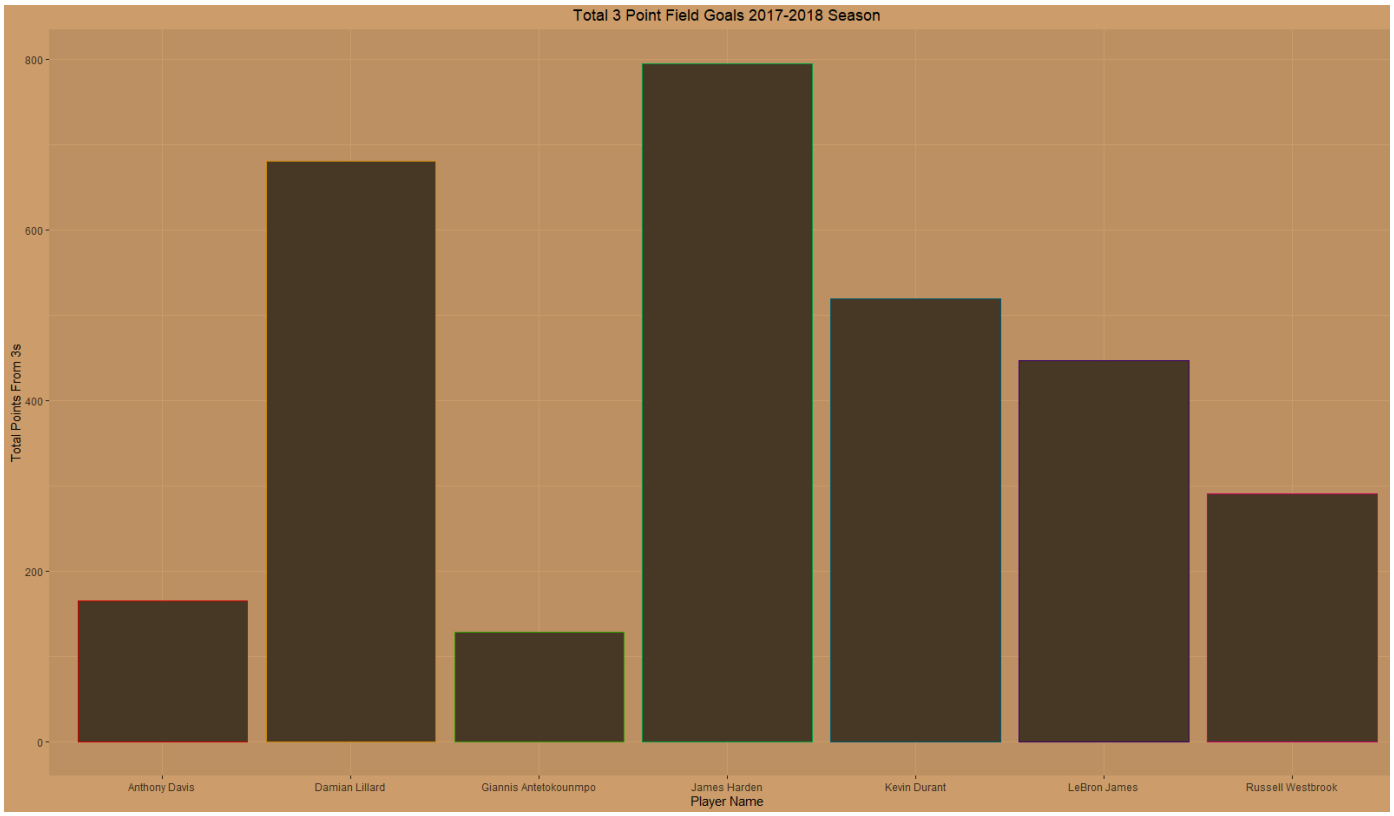
James Harden led the NBA in scoring during the 2017-2018 and 2018-2019 seasons. He had the highest points per 48 minutes in 2017-2018 and led the NBA in total points in 2018-2019. How did he score most of his points? Did his style of offense differ from other players?

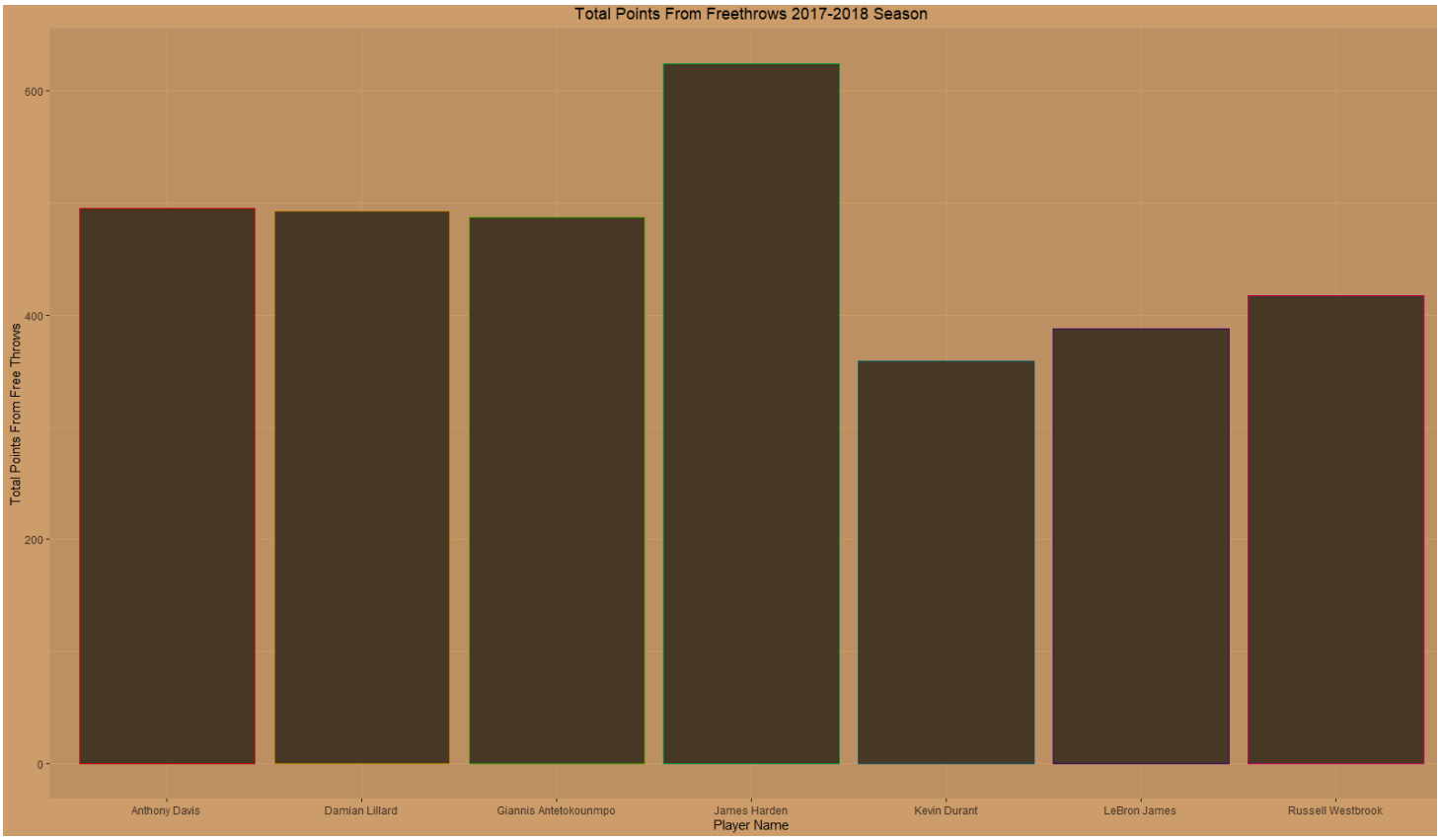
2017



In 2017, James Harden led the NBA in points per 48 minutes, but LeBron James led the NBA in total points. However, the way Harden and James scored their points was different. James focused more on close two-point field goals, while Harden scored with three-point shots and drawing contact to shoot free throws.

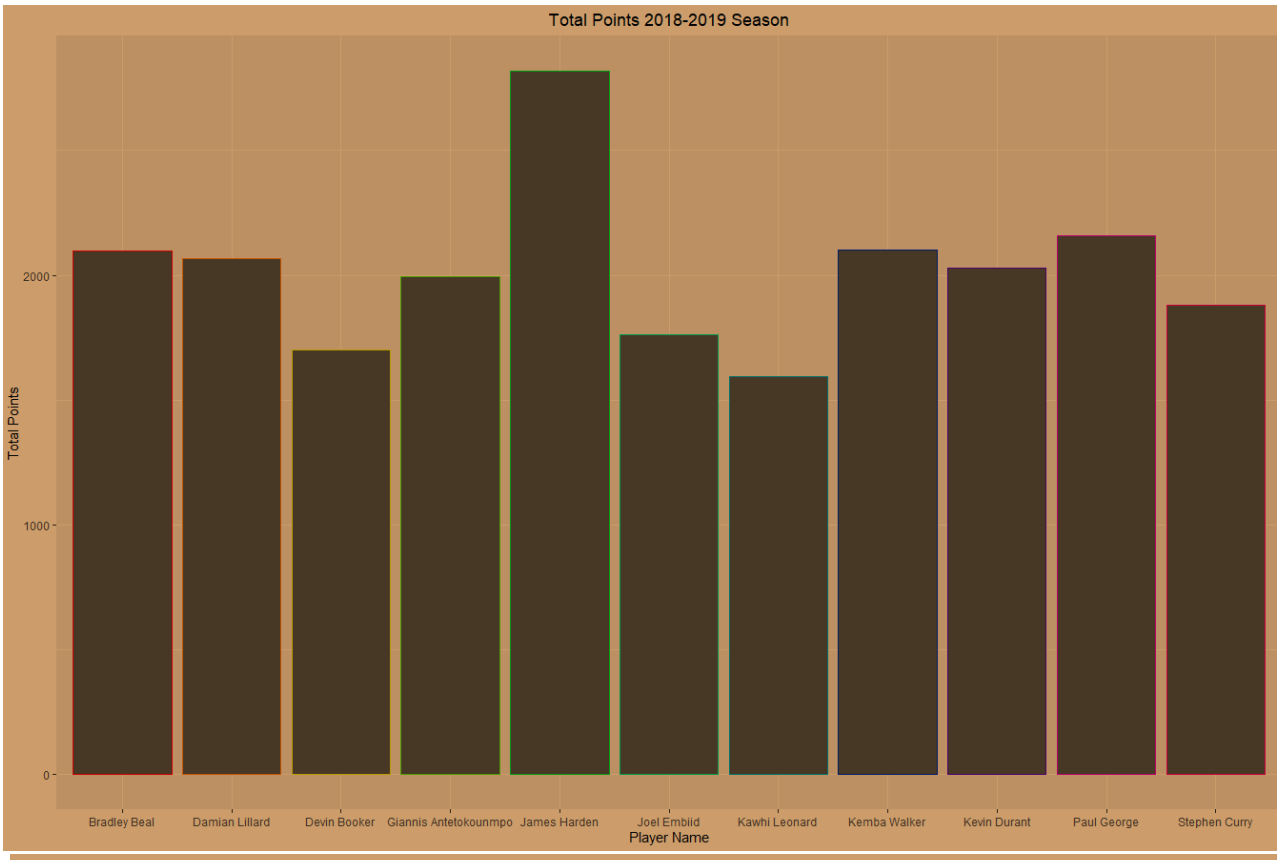

Total 2 Point Field Goals 2017-2018 Season
Total Points From 2s
1500
1000
500
0
Anthony Davis
Damian Lillard
Giannis Antetokounmpo
James Harden
Kevin Durant
LeBron James
Russell Westbrook
Player Name

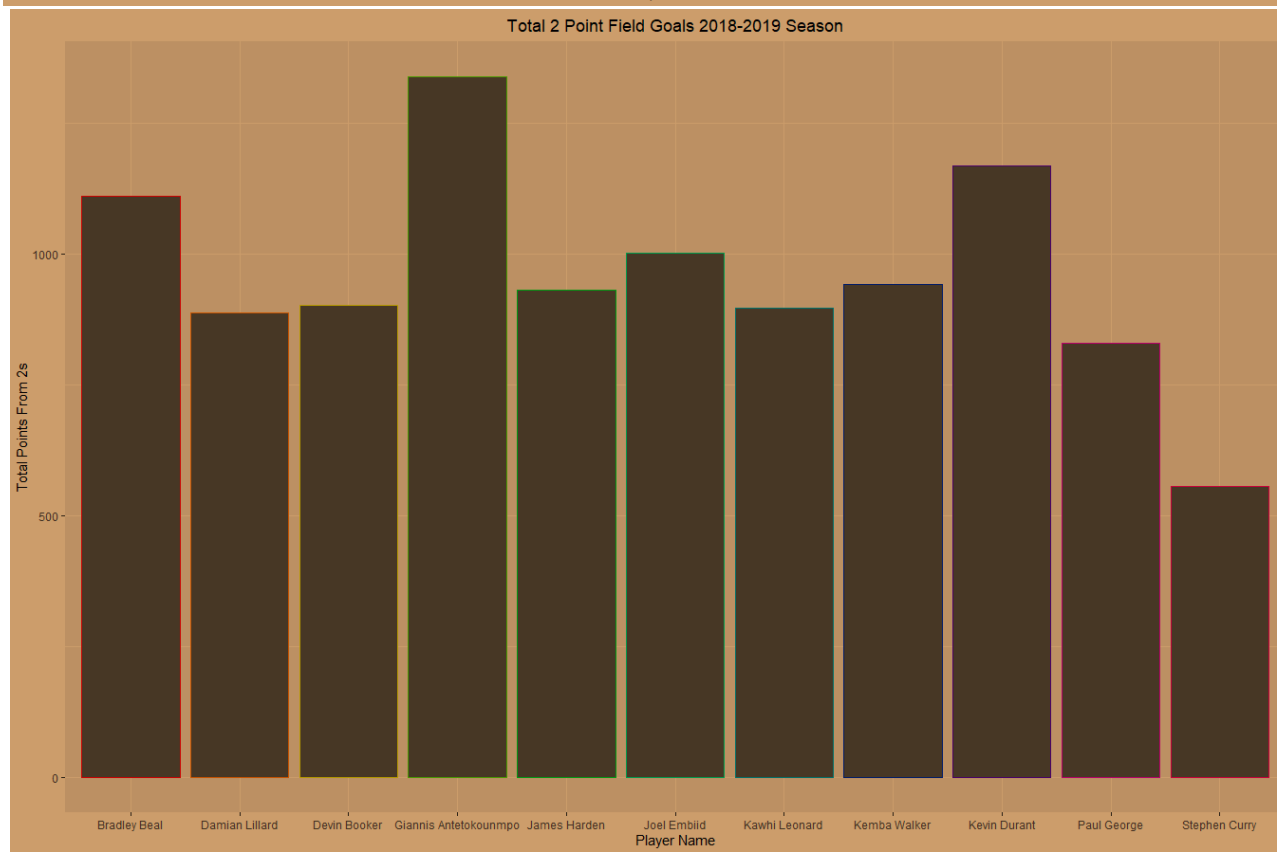

Total 3 Point Field Goals 2017-2018 Season
Total Points From 3s
800
600
400
200
0
Anthony Davis
Damian Lillard
Giannis Antetokounmpo
James Harden
Kevin Durant
LeBron James
Russell Westbrook
Player Name

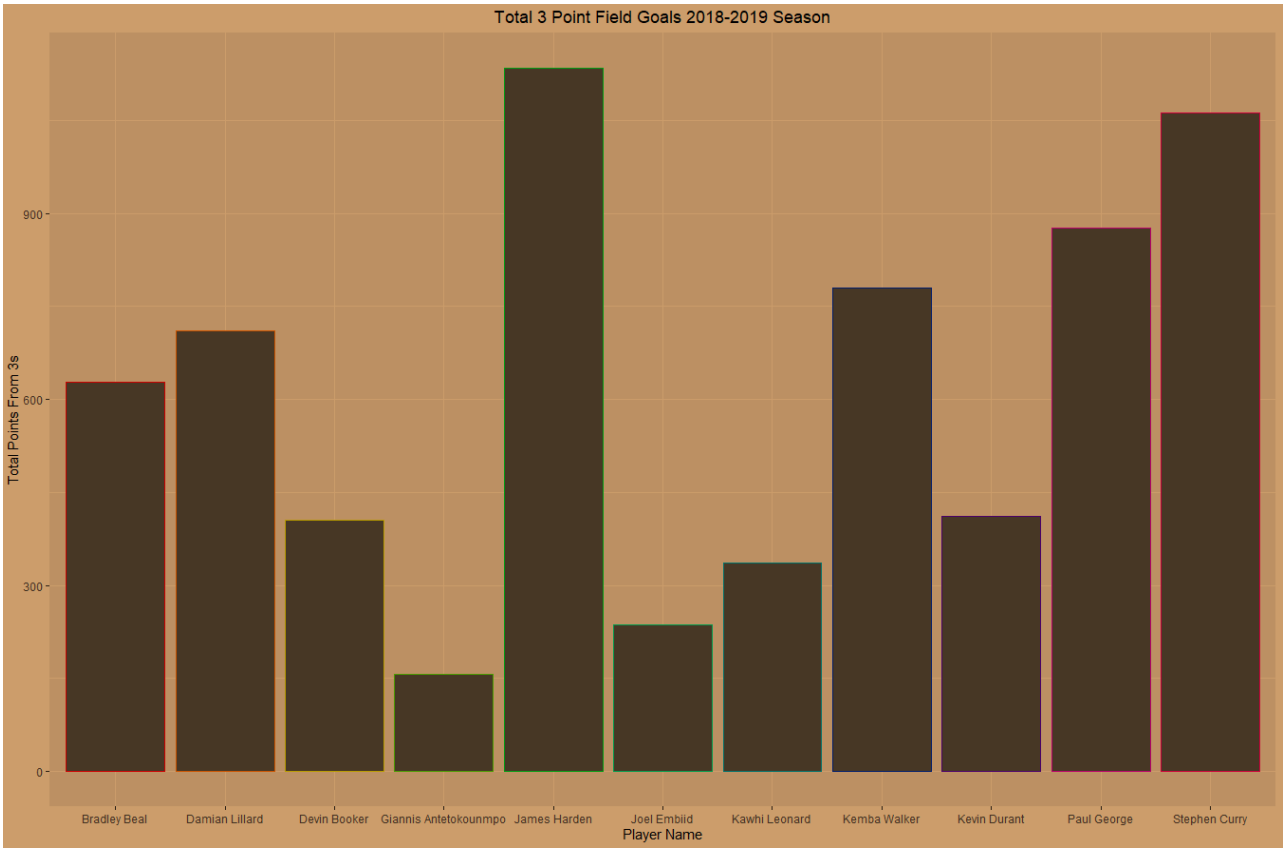


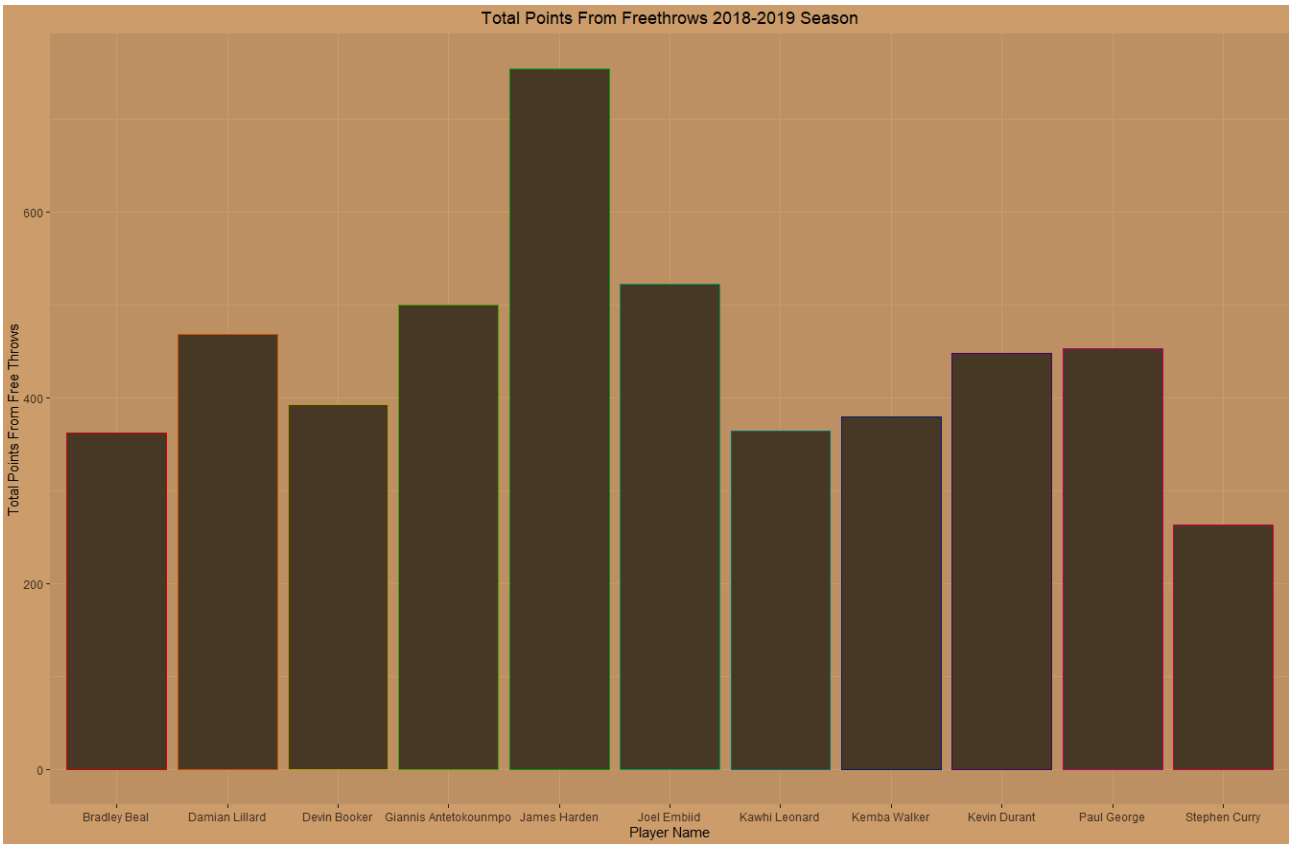
Out of the top 7 highest scoring stars in the 2017-2018 season, James Harden had the least two-point field goals made. He had the most three-point field goals made and has the most free throws made. A reasonable explanation for low number of two-point field goals made could be that Harden is getting fouled on his drives and layups which is leading to free throws. LeBron James was the leading scorer in the NBA in 2018. He scored most of his points from 2-point field goals. He was 2nd in 2-point fields, 4th in 3-point field goals, and 6th in free throws made. Anthony Davis was the 3rd leading scorer in 2018. He also scored a majority of his points from 2-point field goals. He was 1st in 2-point field goals, 2nd in free throws made, and 6th in 3-point field goals.

2018


Total Points 2018-2019 Season
Total Points
2000
1000
0
Bradley Beal
Damian Lillard
Devin Booker
Giannis Antetokounmpo
James Harden
Joel Embiid
Kawhi Leonard
Kemba Walker
Kevin Durant
Paul George
Stephen Curry
Player Name


Total 2 Point Field Goals 2018-2019 Season
Total Points From 2s
1000
500
0
Bradley Beal
Damian Lillard
Devin Booker
Giannis Antetokounmpo
James Harden
Joel Embiid
Kawhi Leonard
Kemba Walker
Kevin Durant
Paul George
Stephen Curry
Player Name


Total 3 Point Field Goals 2018-2019 Season
Total Points From 3s
900
600
300
0
Bradley Beal
Damian Lillard
Devin Booker
Giannis Antetokounmpo
James Harden
Joel Embiid
Kawhi Leonard
Kemba Walker
Kevin Durant
Paul George
Stephen Curry
Player Name


Total Points From Freethrows 2018-2019 Season
Total Points From Free Throws
600
400
200
0
Bradley Beal
Damian Lillard
Devin Booker
Giannis Antetokounmpo
James Harden
Joel Embiid
Kawhi Leonard
Kemba Walker
Kevin Durant
Paul George
Stephen Curry
Player Name

James Harden led the NBA in total scoring during the 2018-2019 NBA season. He had over 500 more points than the second leading scorer Paul George. Harden struggled with two-point field goals like in the 2017-2018, but he led the NBA in three-pointers made and free throws made. He made over 200 more free throws than the second-best player Joel Embiid.

Conclusions

I found it interesting that NBA stars tend to underperform in the playoffs. Fans assume that in do or die and elimination scenarios players would step up, but according to data from the last two NBA seasons this is not the case. The past two NBA seasons almost every player that performed well in the regular season struggled to put up similar numbers in the postseason. This could be because they are playing better competition than in the regular season and because teams play a best of 7 series where the opposing team can strategize to limit an opposing star player. Additionally, NBA analysts have been discussing how the game has changed and how the NBA has become a shooters league. The total points scored bar plots and James Harden support their arguments. Harden is leading the NBA in scoring with high level and efficient shooting. He has led the NBA in three-pointers made the last two seasons and has been top two in the NBA in scoring both seasons as well.