NBA_Final

Jiashun Cheng jchen66@wisc.edu, Shaodong Wang swang647@wisc.edu, Bret Harvestine bharvestine@wisc.edu, Huiyu Jiang hjiang94@wisc.edu, Yulin Liu yliu677@wisc.edu

5/5/2017

NBA Final: "Evolution of the Big Man"

Introduction:

This November, Marc Gasol, the center for the Memphis Grizzlies, hit the game-winning three-point shot against the LA Clippers to put them ahead 109 - 107. Fourteen games into the season, Gasol was 19 of 45 behind the three-point line, which already surpasses the 12-of-66 shooting from his first eight years in the league. In fact, before this year, the most made threes Gasol had in a season was three. So why such a severe increase?

In addition to Gasol, Karl-Anthony Towns, Joel Embiid, Kristaps Porzingis, and Myles Turner are all 7-ft tall and attempting at least one three per game, with Porzingis averaging 5.3 attempts per game on 39 percent shooting (Cato, 2017).

"Pick and roll is still the staple behind most offenses, but pick and pop has become just as important. Having a center who can clear the lane for speedy penetrating guards is ideal. Especially if that center can now be an outlet pass receiver for the penetrating guard, if the guard gets stopped by help defense," says sports analyst Warren Shaw.

Numerous NBA articles are mentioning the evolution of centers and power forwards and how their game play is consisting of more and more three point shot attempts, which prompted us to dig deeper into this phenomena. We want to provide statistical and visual evidence of whether or not we can find a relationship between the different centers and power-forwards in the league and their offensive ability or style of play.

We will observe a players' seasonal data in the center and power forward positions (2002-Present) to show that over the last 15 years there has been an evolution of the big men in the NBA where they are either versatile in offensive skill (wide range shooting and dribbling) or only getting easy baskets.

Data Description:

The data used in this analysis can be found at nba.com. We scraped shooting data containing every player in the league for each year, 1996-2016, and read it into a single data frame. We choose to only use data from 2002 and after for our analysis because 1996-2001 seasons have limited shot classifications, which complicates our clustering. An example of one row from our dataset is as follows with player Stephen Curry:

player_id	player_name	team_id	team_name	period	minutes_remaining
201939	Stephen Curry	1610612744	Golden State Warriors	1	10

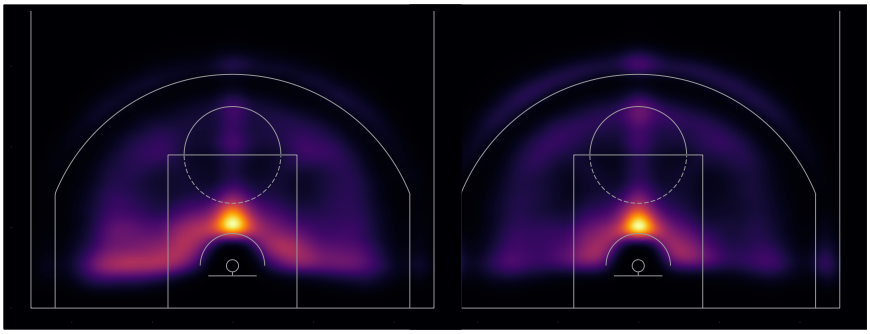
event_type	action_type	shot_type	shot_zone_basic	shot_zone_area	shot_zone_range
Made Shot	Cutting Finger Roll Layup Shot	2PT Field Goal	Restricted Area	Center(C)	Less Than 8 ft.

The type of shooting information collected contains a players' unique ID number, team ID, period (1,2,3,or 4), time_remaining, event_type (shot outcome), action_type (type of shot), shot_type(two or three points), shot_zone_basic, shot_zone_area, shot_zone_range, and the location of the shot.

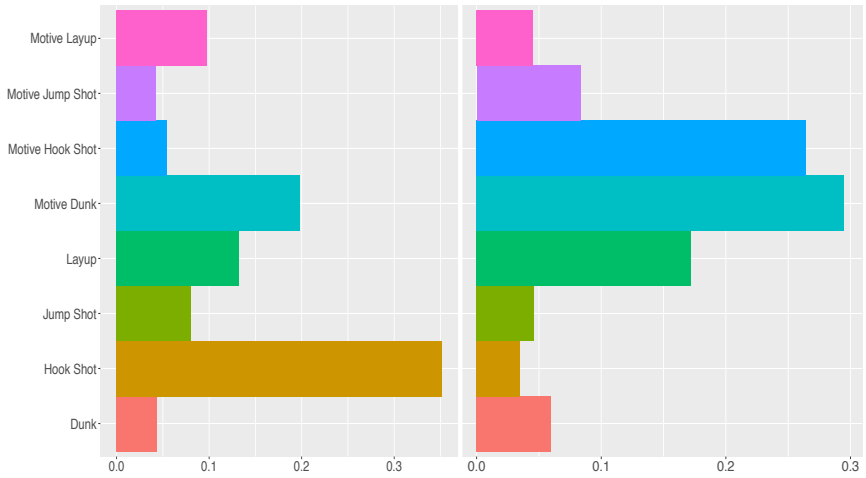
Data Cleaning and Manipulation:

Once the data is cleanly loaded, we use applications of the dplyr package to filter and select players of the center and power forward posistion. We then manipulate our dataset with group_by() for each player and each year in order to calculate the total number of shot attempts by classification of the shot. There are a total of 67 classifications used to describe a shot in this dataset, so we simpify them into only 8 categories: dunk, hook shot, jump shot, layup, motive dunk, motive hook shot, motive jump shot, and motive layup. Note: motive indicates the player was dribbling before the shot attempt. This transformed data frame contains 1,657 rows. To clean our data, we remove players that take less than 300 shots per year to ensure we are looking at players with consistent minutes played and shot attempts.

To get an initial glance at our data, we construct heat map comparisons between active players and hall of fame players to visually see where shots are being taken on the court by centers and power forwards. We can see a distinct difference in the heatmap for the Active bigs (featured right), which shows darker patterns beyond the three-point line and lighter patterns near the basket and around the baseline vs. the hall of fame players' featured on the left. Therefore, active players now are taking more long range shots according to these to plots.



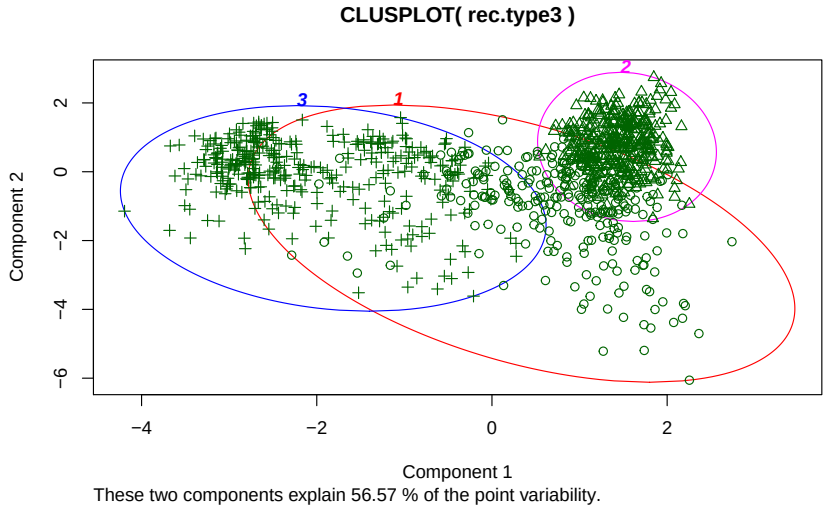
Next, we look to compare the distribution of shot types for active big players and hall of fame players in the center positions under our 8 categories.



As expected after observing the heat maps, the two distributions have an obvious change in shooting patterns. The NBA's active bigs are increasing the rate of self-made dribble plays. Specifically, the increase in Motive Hook Shots, Motive Dunk, and Motive Jump Shot are three categories that show there has been a development in the shooting tendacies for centers and power forwards.

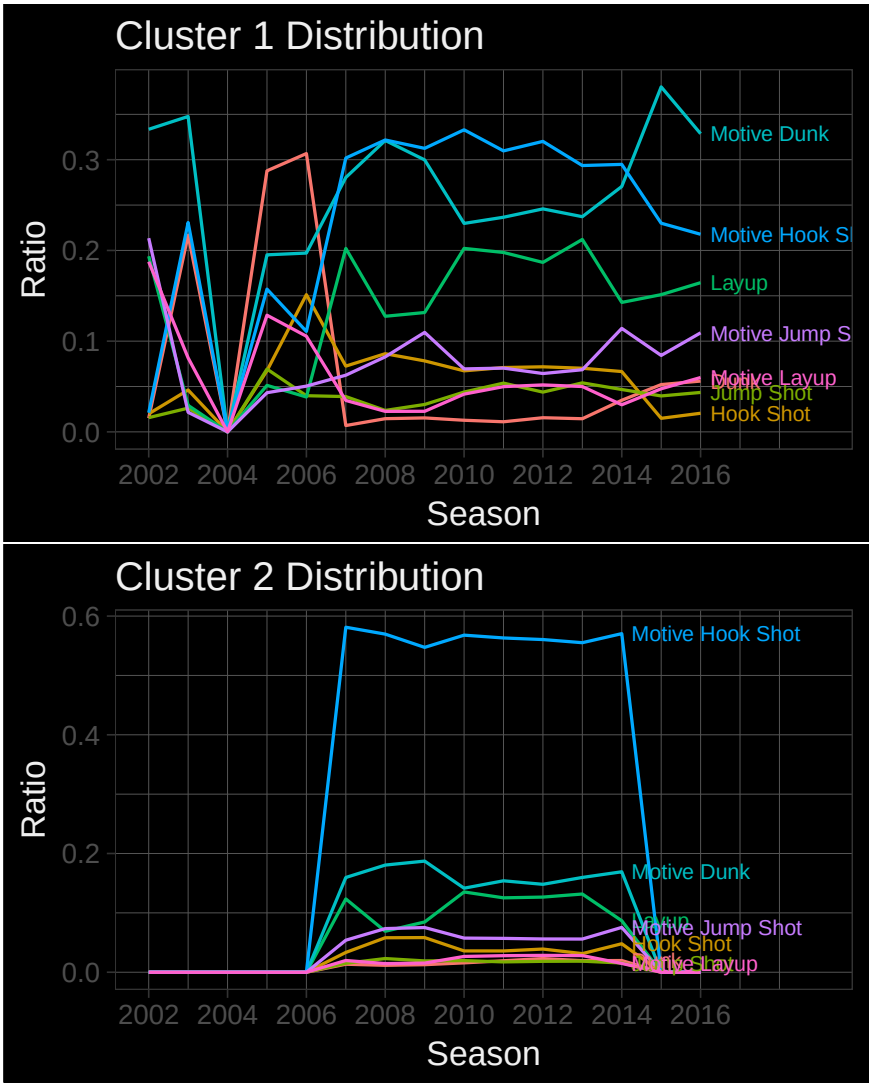
Discovering the Evolution:

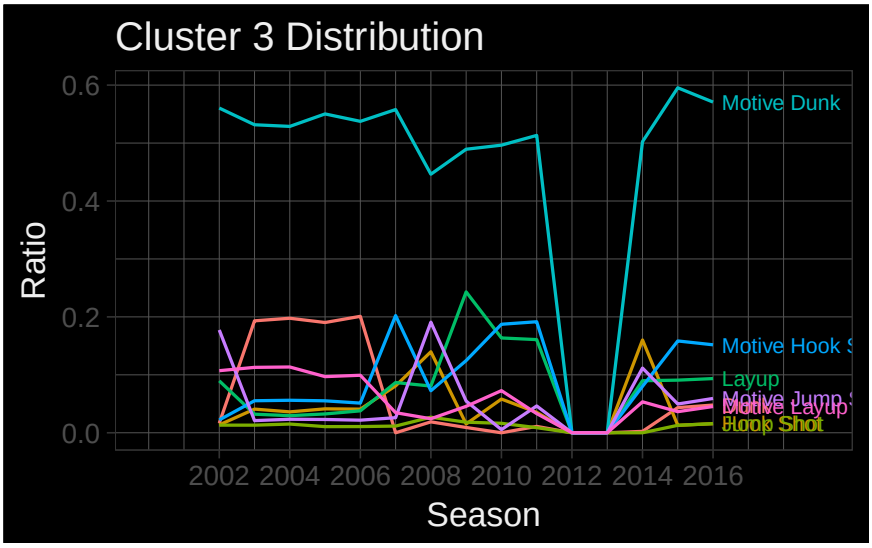
We run a one time K-Means clustering of every player in the dataset from 2002-2016 in order to reveal underlying players' characteristics and how the number of similar players changes. We begin with one cluster and repeat for up to 6 clusters, finding that three clusters represented the data best.



We then make three variables to store the player names of each cluster. The ratio of the number of very shot

action in each group per year is calculated using source("clean data.R"), which also returns a data frame with the number of players in each cluster over the 15 seasons. Now, we construct line plots showing the change in ratio over the years for each cluster.



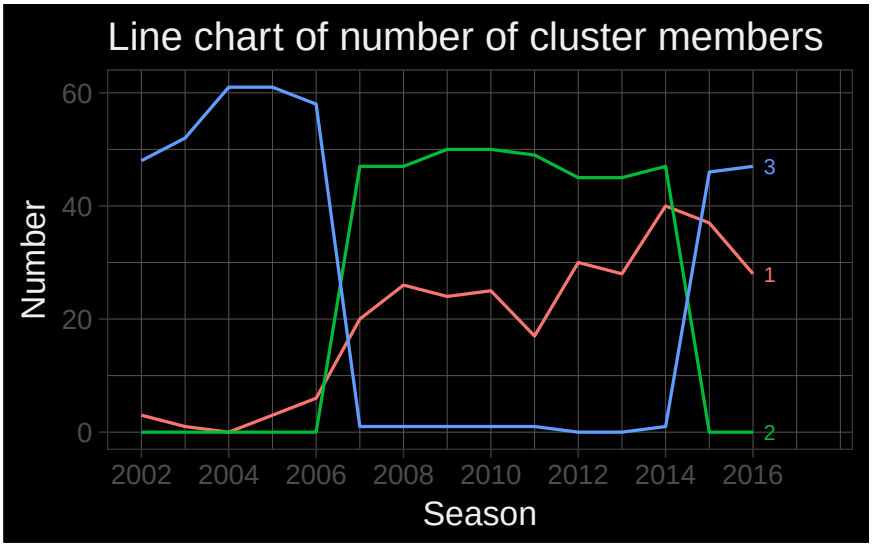


Observe that Cluster 1 depicts the proportion of shot attempts that are Motive Dunk, Motive Hook Shot, Layup, and Motive Hook Shot all show a relatively positive slope beginnning around the 2004 season. Conversely, Motive layup, Jump Shot, Hook Shot all have low ratios meaning these types of shots are less common. This cluster appears to be representative of a versatile big man.

Cluster 2 indicates a clear dominance in type of shot with Motive Hook Shot being the most attempted shot while the other categories are very uncommon. This cluster is representative of post play around the restricted area (the painted rectangle) or mid-range shooting.

Cluster 3 similarly has one shot type that is more prevalent than any other, which is Motive Dunk. Notice that the ratio of shot attempts by shot type decreases in every category beginning around the 2008 season to 0.0 until the 2012 season where it has a steep increase in Motive Dunk. Cluster 3 is representative of big players who are mainly shooting easy baskets under the rim or dunking; these players take little to no shots.

There is an evident distinction in the type of shots between the three clusters. Now, lets observe how the number of players in each cluster changes for every year. Again, we construct a line plot to display this information.



From this plot, we are able to support our hypothesis that there has been an evolution in the style of play for centers and power forwards. We can visually see that over the last 15 years the number of differnt types of bigs in the league changes periodically. The most evidently case of evolution in player characteristics has

happened over the past 3 years where we see that cluster 2 is almost becoming non-existent. Thus, meaning Centers and Power Forwards are no longer being utilized for post play and mid range shots. Rather, bigs are now either active in the offense and shooting from a wide range on the court or they are only on the court for their size and to get the easy cleanup around the rim.

Appendix:

1) Cato, T. (2016, November 26). Marc Gasol and the NBA's biggest centers are shooting threes now. Retrieved May 5, 2017, from SBNation website: http://www.sbnation.com/2016/11/22/13673818/ marc-gasol-3-pointers-stats-highlights-nba-trend

2) Mandell, N. (2017, January 17). The NBA's big men are shooting more and more 3s. Retrieved May 5, 2017, from For the Win website: http://ftw.usatoday.com/2017/01/ brook-lopez-myles-turner-antawan-jamison-nba-3-pointers

3) Shaw, W. (2017, March 6). The evolution of the NBA center has created a new type of center called the stretch 5. Retrieved May 5, 2017, from Basketball Gods website: http://www.basketballgods.net/ evolution-nba-center-created-new-type-center-called-stretch-5-new1/