The Art of the Three: Analyzing the Impact of Three Pointers on Team Success in the National Basketball Association

Nicholas Latchana
nl1776@nyu.edu

Jason Yang
jy2634@nyu.edu

Abstract: Basketball is a sport that, despite its rules rarely changing, has undergone several transformations in terms of the way it is played. As players become more skilled and are able to score further from the basket and beyond the three-point line with ease, we see that the game played today seems nearly completely different than it did several decades ago. With this evolution comes the simple question: is this transition to a three-point oriented game worthwhile? Our group aimed to look at if a pattern existed between team three-point success and overall success in standings. We decided to do this by analyzing two complete NBA seasons, nearly a decade apart, at both the team and player level. This task was accomplished through the utilization of Hadoop architecture, such as Hadoop HDFS, Hive, and Spark. As a whole, the entirety of the league's focus began to be shifted on shots of this variety, moving away from the traditional style that was previously associated with it, and it has become virtually impossible to mention basketball without associating it with the three-pointer. Through our analysis, we were able to determine that in both the 2008 and 2016 NBA seasons, teams that were able to shoot threes efficiently were rewarded with spots near the top of the standings. There were, however, outliers that existed that allowed us to prove a crucial point, which was that basketball is a multi-dimensional sport and success in one area does not promise overall success.

Keywords - basketball, three-point, evolution, multi-dimensional, Hadoop

I. Introduction

Currently in its 75th season, the National Basketball Association has risen in popularity over the last few decades and evolved to become one of the world's most popular leagues. While the league itself has expanded, we have observed an evolution of the style of basketball being played, as well. 20th century basketball was characterized by dominant centers, such as Wilt Chamberlain, Shaquille O'Neal, and Bill Russell, whose ability to score near the basket with relative ease made them valuable assets to their respective teams. In the 2000s, players like Kobe Bryant and Dwyane Wade dominated the game from the mid-range with their terrific jump shooting. Today, we see that players such as Stephen Curry and James Harden have generated incredible success for their teams with their utilization of the three-point shot. Players in today's NBA are more skilled and dangerous than those at any other time in history, considering their ability to excel at scoring from all of these aforementioned locations on the court. The diversification of scoring provides teams with countless possibilities that were virtually unseen in previous eras. As we saw that players had continued to expand their arsenals as time progressed, we began to ponder if this was a worthwhile venture. In other words, our group aimed to look at the impact this transition to three-point oriented basketball has had on the league and determine if this was a wise move in terms of team success.

As mentioned previously, the NBA already had a multitude of seasons for our analytic to be centered around, but the two seasons we chose were picked deliberately, in order to illustrate our point. First and foremost, we centered around the 2008-09 NBA season, a time in which the three-point shot was not nearly as big of an asset as it is today. It was not until the 2009 NBA Draft that the undisputed greatest shooter of all time, Stephen Curry, was drafted. Many fans and analysts alike accredit this evolution to Curry, and as a result, he has been dubbed as one of the most inspirational players of all time. Our second season was the 2015-16 NBA season, a season in which Curry himself set the NBA record for three-pointers made in a season at an astounding 402, shattering his own record of 272 that was set three years prior. Curry and his Golden State Warriors also broke the record for the best team record of all time, with a final record of 73-9. This was, statistically, the greatest team of all time in terms of three-point shooting ability and in overall success, but they ultimately failed to capture the NBA Championship, the league's most coveted trophy.

Basketball, like other sports, serves as a means of bringing individuals together. However, considering the manner in which today's game is so driven by analytics, we have noticed that there may be discrepancies in just how much certain fans are able to consume this knowledge. Prior to this study, my partner and I both considered ourselves to be fans of the NBA with differing levels of knowledge. However, neither of us had ever taken a deep dive into analytics and did not do much exploring of statistics outside of the games we watched. Despite this, it was virtually impossible for either of us to deny the fact that just from eyeballing, we could see that this current state of basketball had the highest three-point utilization of all time. As a result, we decided to see if the events unfolding in front of us would be supported by analytics, or if we were to see that the three-point shot was actually detrimental to the game we both loved. Our application was designed with the intention of being used by anyone with an interest in basketball as a means of breaking down this aforementioned "analytics barrier."

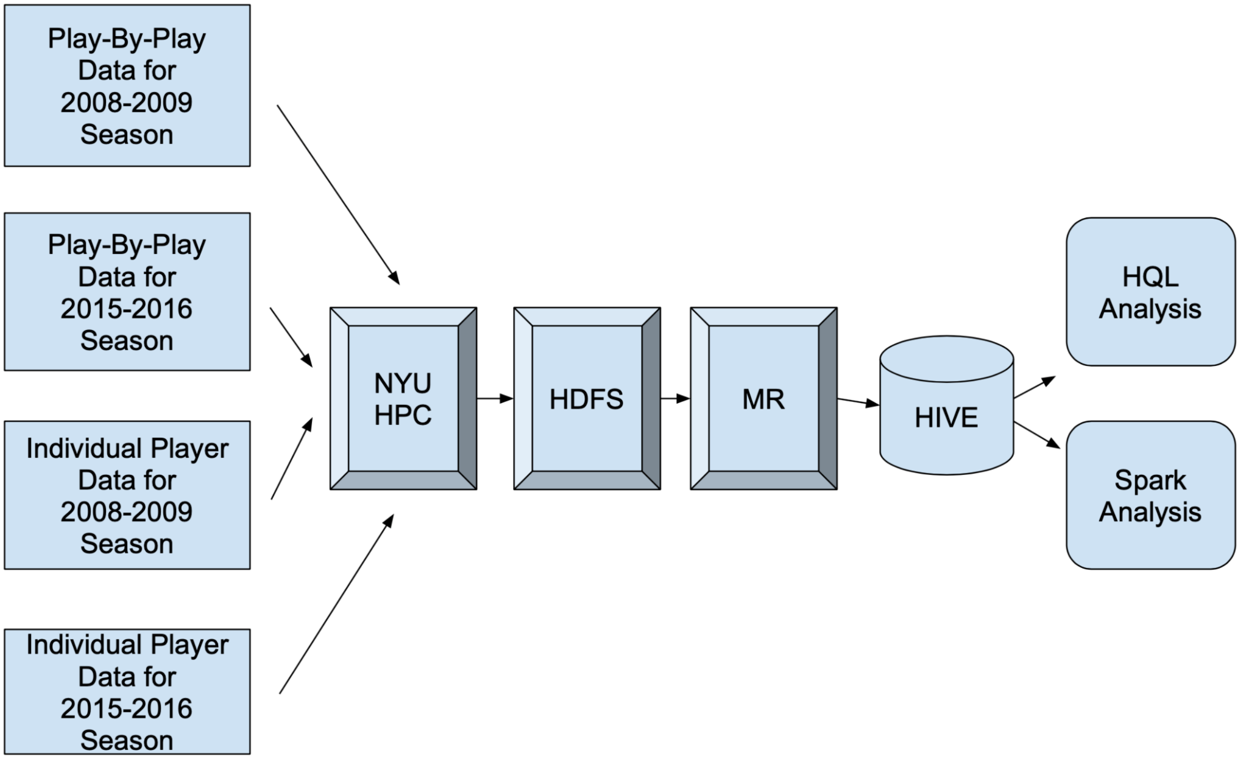

Figure 1: analytical technology flow

Data and Technology Flow Diagram: *Figure 1* depicts the flow of our data throughout the different technologies leveraged within our analytics application which sits atop the NYU Peel Cluster. The 4 raw datasets are first placed onto NYU HDFS followed by Hadoop MapReduce step which processes, cleans, and profiles our raw datasets and outputs clean data containing only the 3-point shot quantities we are interested in. The clean dataset output of this MapReduce stage is subsequently used in the Hive stage where an external table is created for each of the 4 cleaned datasets. A series of joins are executed to boil our 4 datasets down to 2 representing the 2008 and 2016 NBA season. This is followed by HQL queries and Spark Scala for analytics.

II. Motivation

The utility of our analytic is tremendous in the regard that it can serve a purpose to people at all levels of the game of basketball. My partner and I are both avid basketball fans, with him being a fan of the NBA and myself being a former player in my youth, and wanted to develop an analytic that benefited people on all ends of the spectrum, similarly to ourselves. Casual enjoyers of the game can use this analytic to analyze past seasons as a means of seeing just how revolutionary the three-point shot has been. Fans of specific teams and players have the ability to keep track of just them over the course of several years. Furthermore, our analytic is so in-depth that even personnel within the National Basketball Association can benefit by using it to track both the success of their own teams and players and the opposition's, as a means of seeing if both their offensive and defensive schemes require additional attention. Altogether, anyone with an interest in basketball can benefit from using this analytic as it provides a variety of statistics at varying levels of complexity.

III. Related Works

One article that my partner and I found to be incredibly useful when designing this analytic was the paper, *To Three or Not to Three,* by Tony Caporale and Trevor C. Collier. In their study, the two developed a model that would describe how coaches choose to allocate field goal attempts between two-point and three-point shots. They wrote, "The aspect of a basketball game in which coaches have the most control is their team's shot selection," and this statement is completely supported by my partner and I. After all, one potential use of our analytic was to determine if teams should continue to emphasize three point shots or if they should abandon this style of play entirely, shifting focus onto other areas. Caporale and Collier ultimately concluded that the most successful teams were the ones who were able to effectively allocate their shot selection, which does not always require emphasis on three-pointers, exclusively.

Similarly to the Caporale and Collier paper, our analytic greatly benefitted from our incorporation of ideas from *Three point shooting and efficient mixed strategies: A portfolio management approach.* In this research article from Mark Fichman and John O'Brien, it was determined that since the introduction of the three-point line in 1979, the probability of a successful three-point shot has exhibited an increasing trend for its mean and a decreasing trend for its variance. This was extremely significant to our project, as this makes it apparent that the league as whole is increasing in terms of their three-point making ability. It is this fact that makes it easier to justify why we see the numbers increasing as drastically as they have, but also allows us to question if this is worth transitioning away from a two-point oriented game, since its probability has stabilized throughout the league's history.

IV. Description of Datasets

In order to get the most complete statistics, our first two datasets were selected from Sports Statistics, a site that tracked and encapsulated play-by-play data for the entirety of two NBA seasons : 2008-09 and 2015-16. The 2008-09 dataset was contained in CSV format and consisted of 546,454 lines that detailed every single play that occurred in that season and as a result, its contents added up to a whopping 89.1MB. Similarly, the 2015-16 dataset was 568,334 lines and it also comprised all game events from that season, with a size of 95.1 MB. Since they were from the same source, the original schema for both featured over thirty columns and included a variety of information, such as player names, team abbreviations, and team IDs. However, in order to be used specifically for our purpose, we narrowed the schema down to include just five of these columns: **GAME_ID, TEAM_ABBREVIATION, HOMEDESCRIPTION, VISITORDESCRIPTION, SCORE**. First and foremost, we kept the Game ID in order for us to track individual games. Within these specific games, we kept the abbreviation of the team that was the one completing an action. Next, we decided to keep Home Description and Visitor Description, which were used to signify if the play was being done by the Home or Visitor team, and as a result, these two were never present together since only one team could have been credited per play. Finally, we kept the score, a major component that allowed us to ultimately keep track of wins and the success of teams.

To accompany the Play-by-Play dataset of the two seasons, we investigated Player specific attributes. The 2008 Player dataset contains player-specific statistics across the 2008-2009 NBA season. Likewise, the 2016 dataset contains season statistics for each active player in the 2015-2016 NBA season. Both datasets are sourced from Basketball Reference, a reputable source that tracks and aggregates NBA statistics. Both datasets are approximately 74KB with the 2008 Player dataset housing 583 observations and 2016 Player data housing 579 observations where each observation is an active player under the season of interest. Following the same schema, both datasets contain information on a Player's personal attributes such as age, the team they play for, offensive statistics on two-point and three-point shooting and defensive statistics such as total number of blocks and steals. It is important to know that each variable is a sum across the entire season which allows us to analyze which players performed best, especially in three-point shooting, our area of analysis. We utilized Hadoop MapReduce to reduce the original schema to only include the following variables: **Player, Team, 3-point makes, 3-point attempts, 3-point percentage,** and **total points.** This schema allows us to focus solely on 3-point related attributes necessary for our application.

V. Analytic Stages

A: *Preprocessing and Extraction*

A1**:** As mentioned previously, the two play-by-play datasets were rather intensive, something that was not necessary for our project since we were only interested in the three-point aspect of basketball. After our initial copy of these datasets onto the NYU Peel Cluster, it was important that we removed this unnecessary schema in order to not only make the data easier to work with, but also more relevant. Therefore, we decided to use the Mapper portion of this MapReduce job as a means of extracting only data pertinent to our project, which ended up only being the game ID, the team name, the play that was occurring (whether it be by the home or visiting team), and the score. This was done by using the game ID as the key, since the aforementioned four schemas were all being tracked on a game-by-game basis and the dataset as a whole was of every game within the season.

In the Reducer stage of our MapReduce job, our first task was to neglect all plays that did not include a three-point shot, as these plays did not serve any importance in our study. During our first time, we did this by filtering out all plays that did not include the phrase "3PT Jump Shot," which we believed to be sufficient in removing these actions. However, this led to our results containing a far smaller number of both three-point makes and misses. We corrected this by simply filtering out all plays that did not include "3PT," as we noticed that some three-point shots were classified in other ways, such as "3PT Pullup Jump Shot," and as a result, were not being counted in our final statistics.

Despite us associating four different schemas with the Game ID, another problem that was encountered was when less than four pieces of information was written on certain lines. We deduced that the length of this values array was dependent on the type of play that was occurring, and as a result, would not be uniform throughout the dataset. When the type of play occurring was a miss by the home team, this array would be of length 2, since the only pieces of information being recorded were **TEAM_ABBREVIATION** and **HOMEDESCRIPTION.** Similarly, an array of length 3 indicated that the play was a miss by the visiting team as the only pieces of information being recorded were **TEAM_ABBREVIATION and VISITORDESCRIPTION.** Although we saw that two pieces of information are recorded for both of these events, the reason that the latter is of length 3 is because the null **HOMEDESCRIPTION** is in the middle of the two recorded schemas and is counted towards the array length. Finally, an array of length 4 represented a make by either the home or visitor team and is the only time when a value is recorded for **SCORE.** When this happened, our Reducer updated the score in order to have the most complete version at the end of the game. We then separated the score to determine how many points both the home and visitor team had, in order to determine who won the game, something that would ultimately be used when we developed our standings. At this point, following the conclusion of the game, the Reducer would output two lines, each containing the following statistics for their respective teams: the game ID, the team abbreviation, the number of made threes that game, the number of missed threes, and an int that would print 1 if they were the winners, or 0 if they had lost.

***A2*:** The 2008 and 2016 Player specific datasets were first placed onto the NYU Peel cluster within the HDFS system. From there, we preprocessed and cleaned the data using Hadoop

MapReduce which followed the following steps. Our Mapper processed each line within our dataset, removed social media handles and extraneous information originally concatenated to each player's name, and also extracted the player's **Team, 3-point makes, 3-point attempts, 3-point percentage,** and **total points.** The Player's team was written to the Reducer as the Key and the extracted fields of **Name, 3-point makes, 3-point attempts, 3-point percentage,** and **total points** were written to the Reducer as values. Within the reducer, we seek to only retain for each of the 30 NBA teams, the player who attempted the most 3-Point attempts. With the Team being the key, we looped over every player under the same team and wrote out the player with the most attempts. The Reducer outputted a cleaned dataset containing 30 rows, each representing the team's player who attempted the most 3-point shots with the following variables: **Name, Team, 3-point makes, 3-point attempts, 3-point percentage,** and **total points.** Both the 2008 and 2016 Player datasets underwent the same said procedure.

B: ***Hive Tables and Data Aggregation***

B1: Following the completion of these MapReduce jobs, the next step was to use the output to create two tables in Hive, one for each of the two seasons that our study was centered around. Each of these tables had five columns: **gameID, team, madeThrees, missedThrees, attempts,** and **win.** Below is an example of the Hive tables created for both the 2008 and 2016 seasons, titled 2008Stats and 2016Stats.

2008stats.gameid	2008stats.team	2008stats.madethrees	2008stats.missedthrees	2008stats.attempts	2008stats.win
20801201	LAL	3	12	15	1
20801201	MEM	8	11	19	0
20801202	IND	8	14	22	0
20801202	CLE	8	13	21	1
20801203	WAS	3	11	14	1
20801203	TOR	4	13	17	0
20801204	DET	4	10	14	1
20801204	CHI	5	11	16	0
20801205	NJN	9	15	24	1
20801205	CHA	0	6	6	0

Figure 2: 2008Stats

2016stats.gameid	2016stats.team	2016stats.madethrees	2016stats.missedthrees	2016stats.attempts	2016stats.win
21501201	SAS	3	13	16	0
21501201	GSW	10	17	27	1
21501202	NYK	7	12	19	1
21501202	TOR	7	13	20	1
21501203	CLE	11	18	29	1
21501203	ATL	8	22	30	0
21501204	ORL	7	16	23	1
21501204	MIL	3	12	15	0
21501205	BOS	8	16	24	0
21501205	CHA	14	18	32	1

Figure 3: 2016Stats

B2: Following the MapReduce stage which cleaned and profiled our 2008 and 2016 Player datasets, we load this data into 2 external Hive tables with both having the following column schema: **team, playername, threep, threepa, threepp,** and **points.** A snapshot of the said Hive table is depicted below.

player2008.team	player2008.playername	player2008.threep	player2008.threepa	player2008.threepp	player2008.points
ATL	Mike Bibby	167	428	0.39	1176
BOS	Ray Allen*	199	486	0.409	1434
CHA	D.J. Augustin	108	246	0.439	850
CHI	Ben Gordon	173	422	0.41	1699
CLE	Mo Williams	183	420	0.436	1443

Figure 4: player2008

player2016.team	player2016.playername	player2016.threep	player2016.threepa	player2016.threepp	player2016.points
ATL	Kyle Korver	158	397	0.398	739
BOS	Isaiah Thomas	167	465	0.359	1823
BRK	Bojan Bogdanović	129	338	0.382	887
CHI	Nikola Mirotić	135	346	0.39	777
CHO	Kemba Walker	182	490	0.371	1689

Figure 5: player2016

B3: Following the creation of an external table for each of our 4 datasets as described and depicted in ***B1*** and ***B2***, we proceeded to join the 2008 Play-by-Play dataset with the 2008 Player dataset using a Left Outer Join on **team** name. This resulted in a table containing aggregated Team statistics with the respective player statistics as well. This Join action was repeated on the 2016 NBA season datasets. Snapshots of the two joined datasets are shown below.

nba2008.gameid	nba2008.team	nba2008.madethrees	nba2008.missedthrees	nba2008.attempts	nba2008.win	nba2008.playername	nba2008.threep	nba2008.threepa	nba2008.threepp	nba2008.points
20800001	BOS	3	11	14	0	Ray Allen*	199	486	0.409	1434
20800001	CLE	3	12	15	1	Mo Williams	183	420	0.436	1443
20800002	CHI	5	7	12	1	Ben Gordon	173	422	0.41	1699
20800002	MIL	6	8	14	1	Richard Jefferson	116	292	0.397	1607
20800003	LAL	7	3	10	1	Kobe Bryant*	118	336	0.351	2201

Figure 6: nba2008

nba2016.gameid	nba2016.team	nba2016.madethrees	nba2016.missedthrees	nba2016.attempts	nba2016.win	nba2016.playername	nba2016.threep	nba2016.threepa	nba2016.threepp	nba2016.points
21500001	ATL	8	19	27	0	Kyle Korver	158	397	0.398	739
21500001	DET	12	17	29	1	Kentavious Caldwell-Pope	114	369	0.309	1105
21500002	CHI	7	12	19	1	Nikola Mirotić	135	346	0.39	777
21500002	CLE	9	20	29	0	J.R. Smith	204	510	0.4	955
21500003	GSW	9	21	30	1	Stephen Curry	402	886	0.454	2375

Figure 7: nba2016

C. HQL Analysis and Graphs

Step B produced two cleaned and aggregated datasets, nba2008 and nba2016, that each store both team and player specific statistics for their respective seasons. Using these external tables, we conducted a multitude of Hive HQL queries in the next stage of our application.

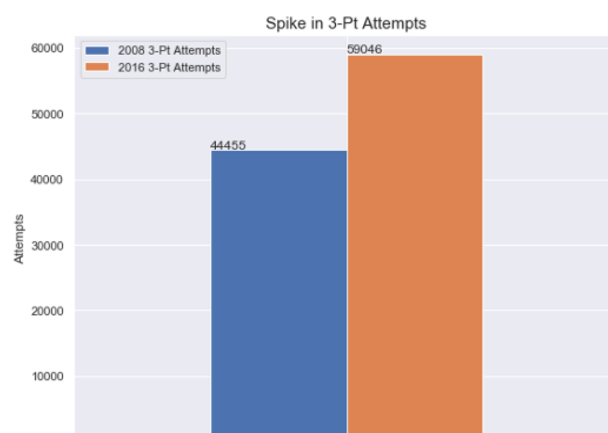

Figure 8: difference in 3pt attempts

C1: We aggregated the **attempts** column for both *nba2008* (*Figure 6*) and *nba2016* (*Figure 7*) to show the difference in 3-point attempts between the 2008 and 2016 NBA season. Depicted by *Figure 8*, there were roughly 14,591 more 3-point attempts in the latter season. With further queries, we found that this statistic deduces down to an average of 12 more attempts per game, 4 more made threes on average and an average of 12 more points. This shows that attempts have drastically increased and are being incorporated into playbooks with more significance. As attempts have substantially increased, we expected more made threes per game, which is the case as illustrated in *Figure 9*. The queries and results established in this section of analytics confirmed our initial hypothesis, which was that the three-point shot is being utilized more, but also serves as a slingshot into more complex deep driven queries specified in ***C2.***

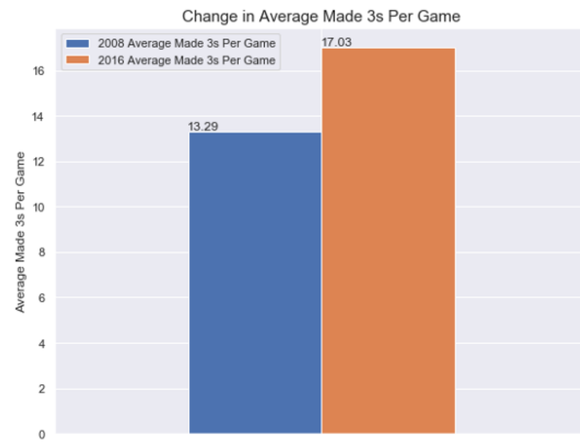

Figure 9: difference in average 3pt makes per game

C2**:** With a much higher volume of three point attempts in 2016 as compared to 2008, we sought to analyze whether players were shooting at a higher percentage (i.e. are they becoming more accurate). At this stage in the analysis, it was important to present to our readers that shooting 33% from 3-point equates to shooting 50% from 2-point shots; any player whose percentages are above these averages were considered to be very efficient and elite shooters. With this in mind, we found that the league 3-Point Percentage (3PP) in 2008 is 38.6%, where the 2016 average is 36.7%. This was an unexpected result as we expected players to be better at 3-point shooting, rather than becoming worse. We established in ***C1*** that the volume of 3-point shots increased substantially, showing that the 3-point shot did indeed change how players are shooting. It is unrealistic to expect that the percentage of made threes would increase at an even rate with the volume, so despite the league average decreasing, the idea that it remained relatively close reemphasized just how skilled the league as a whole was getting. With domain knowledge, we can say that in 2008, when 3-point shooting was not emphasized, only elite shooters and players at the position of Guard or Forward were shooting 3s. The concept of a "stretch big," or a player at PF or C shooting threes, was relatively uncommon and was not considered a vital part of their games at this time. However, In 2016, players of all positions including Power Forward and Center were attempting 3s, diversifying this percentage. Nevertheless, we see that 3PP is relatively the same but the area of focus is that 3-point attempts have increased drastically; we seek to better understand this finding in subsequent analytic stages.

C3**:** Our previous queries found that the volume of 3-point shots increased drastically with 3PP remaining relatively the same. We attributed this to the fact that in 2008, only elite shooters were shooting 3 pointers at such high rates and under this same cloud, we sought to examine the traits and habits of top shooters in relation to their 3-point shots. Our queries found that in both 2008 and 2016, the top 5 shooters sorted by 3PP shot over 40%, an extremely efficient and accurate percentage. This was expected as we were analyzing solely the best shooters in the league, but differences between the two generations of shooters still became apparent when analyzing their 3-point statistics. Namely, top shooters in 2016 attempted almost double the 3-point attempts than that of 2008 top shooters. This was inline with our findings in *C1*. Furthermore, we discovered that points from 3-point shots accounted for about 37% of a player's total points in 2008. This meant that a third of their total season points came about from 3-point shots, with the other two-thirds scored from 2-point shots and free throws. In 2016, this proportion jumped to an astonishing 53%. What this meant is that the best shooters in 2016 were not just shooting more 3s, but shooting 3s had become their specialty and often, their sole purpose. This signified that the best shooters in today's game can focus their gamestyle towards playing outside the arc and utilizing the 3-point shot to a high usage rate. This is far different from the game style of 2008 where only elite shooters shot 3s and basketball was mostly played within the arc, so we can see that this has transitioned into a more diversified game style where players can focus on different aspects of the game. Illustrations of our findings can be found on *Figure 10* and *Figure 11*.

player2008.team	player2008.playername	player2008.threep	player2008.threepa	player2008.threepp	player2008.points	proportion
CHA	D.J. Augustin	108	246	0.439	850	0.3811764705882353
CLE	Mo Williams	183	420	0.436	1443	0.3804573804573805
SAS	Roger Mason	166	394	0.421	969	0.5139318885448917
SAC	Kevin Martin	115	277	0.415	1254	0.2751196172248804
CHI	Ben Gordon	173	422	0.41	1699	0.305473808122425

Figure 10: proportion for he top 5 shooters in 2008

player2016.team	player2016.playername	player2016.threep	player2016.threepa	player2016.threepp	player2016.points	proportion
LAC	J.J. Redick	200	421	0.475	1226	0.4893964110929853
GSW	Stephen Curry	402	886	0.454	2375	0.5077894736842106
ORL	Evan Fournier	156	390	0.4	1213	0.38582028029678483
CLE	J.R. Smith	204	510	0.4	955	0.6408376963350786
ATL	Kyle Korver	158	397	0.398	739	0.6414073071718539

Figure 11: proportion for he top 5 shooters in 2016

D. Scala Analytics and Graphs

D1: Does shooting and making more threes allow a team to be more successful? We sought to answer this question by tracking the progression of 2008's best and worst teams, comparing their 2008 league standings and three-point statistics to theirs in 2016. The Los Angeles Lakers were the NBA champions and thus, the best in the league in 2008 with the following 3-point statistics: (3PM: 542, 3PA: 1505, 3PP: 0.36, Wins: 58). However, they ultimately dropped to being the second worst team in 2016 with the following statistics: (3PM: 626, 3PA: 1977, 3PP: 0.317, Wins: 22). While their volume increased, we noticed that their 3PP decreased, letting us know that they attempted to keep up with this transforming league without having a team equipped for it. Examining the worst team in 2008, Oklahoma City Thunder. their 2008 statistics are as follows (3PM: 321, 3PA: 933, 3PP: 0.344, Wins: 25). They jumped from being the worst in the league this year to being the fifth best in the league in 2016 with the following statistics: (3PM: 678, 3PA: 1941, 3PP: 0.349, Wins: 56). We notice that the Thunder's change in attempts were much greater than that of the Lakers while keeping the same 3PP, an indication that they also transformed with the league but took the additional steps to secure a roster of players capable of making threes at an impressive level, unlike the Lakers. *Figure 12* shows a comparison of these three-point makes and attempts of these teams during each season, and it becomes evident that their successes in three-point shooting and wins are linked. While 3-point shooting surely played an important factor in each team's success, it was certainly not the sole conclusive factor. There are simply too many variables and confounders that attribute to a team's success, such as playmaking, team chemistry, injuries, etc. that make it impossible to simply attribute success or failure to a single area like 3-point shots. What we can conclude is that 3-point shots have undeniably changed the way players and coaches play the game, and although it is not the only deciding factor in team and player success, it has become one of the most important.

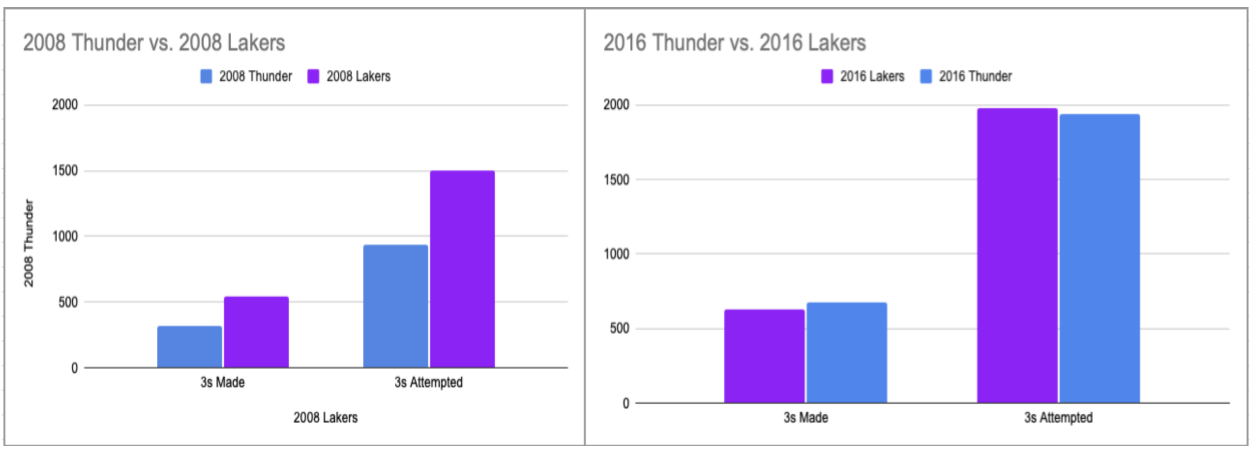

Figure 12: made and attempted threes for the Thunder and Lakers in both seasons

VI. Conclusion

Since the introduction of the 3-point line in 1979, the 3-point shot has and will continue to influence every aspect of basketball as shown in our analytic. Made three-point baskets can be demoralizing to teams attempting to come back into a game, while missed three-point shots provide the opposition with the ability to take leads and hold them. Through our analysis of both team and player data from different decades, we found that the 3-point shot is utilized much more substantially in the 2016 season than it had been in 2008. Namely, the number of attempts and average made threes per game rose substantially, but most importantly, we saw a transition of how the 3-point shot was being utilized. In 2008, 3-point shot points accounted for 30% the top shooters total points, which indicated that other shots were equally as important to players of the time. Players could not simply rely on their ability to make threes as a basis for being in the NBA, and it was necessary for their games to be multifaceted. Fast forward to our analysis of 2016, 3-point shot points accounted for roughly 50% of top shooter's points. This indicates that in today's game, players can specialize in the 3-point shot, changing the way coaches utilize each player on the team. We see athletes like Duncan Robinson, a player notorious for nothing but his ability to hit three-point shots, being awarded a 5 year, $90 million dollar deal this offseason, whereas it is likely that he would not have even been on a roster just a few years prior. While our analysis showed that teams who shot more threes with an above average percentage experienced more success over time, it is imperative to note that the 3-point shot was not, anc will not ever be, conclusive in determining a team's success. Success in basketball is a multivariate formula, meaning that simply shooting and making more 3-point shots does not guarantee team success, something that was observed in the 2016 season as mentioned in Section I. The analytics derived from this application allowed us to quantify just how the 3-point shot has changed how basketball is played compared to 2008 and 2016, allowing us to generate insights and thesis into the impact of the 3-point shot. The game of basketball will continue to evolve over time and the emphasis on the 3-point shot is a prime example of a recent transition.

VII. Acknowledgements

We would like to extend our gratitude to Professor Ann Malavet for her guidance and resourcefulness, as her insights and feedback were invaluable to the completion of this application. We would also like to thank the NYU High Performance Computing Team for allowing us to access cluster resources and their assistance with issue troubleshooting. Our data used were generously provided by Sports Statistics and Basketball Reference, and related research provided by the two papers specified in Section III aided in the project progression. Any and all errors presented in this application are our own.

VIII. References

- https://content.iospress.com/articles/journal-of-sports-analytics/jsa154
- https://link.springer.com/content/pdf/10.1007/s12122-014-9193-5.pdf
- https://sports-statistics.com
- https://www.basketball-reference.com
- https://www.basketball-reference.com/leaders/fg3_season.html
- https://www.espn.com/nba/story/_/page/nbarank22932314/nbarank-game-changers-25-most-influential-basketball-players-ever
- https://bleacherreport.com/articles/10009480-duncan-robinson-re-signs-with-heat-on-5-year-90m-contract-largest-ever-for-udfa
- https://bleacherreport.com/articles/94423-is-the-oklahoma-city-thunder-the-worst-nba-team-ever
- https://www.nba.com/lakers/history/backtobackchampionshtml