# NBA players' performances on court: A Tableau visualization project

Jingheng Huan
Duke Kunshan University
Kunshan, China
jingheng.huan@dukekunshan.edu.cn

Yihang Fu
Duke Kunshan University
Kunshan, China
yihang.fu@dukekunshan.edu.cn

Xueting Li
Duke Kunshan University
Kunshan, China
xueting.li@dukekunshan.edu.cn

## I. Introduction

The NBA is considered the premier stage for both basketball players and fans. In recent times, the style of play in the NBA has evolved from its traditional roots to a more contemporary approach, with a strong emphasis on enhancing players' shooting abilities. Consequently, basketball fans have ignited a hot debate over identifying the NBA's top shooter. In our research, we tried to examine online discussions on Weibo and determined the most frequently-mentioned players within these conversations. Additionally, we utilized an API to query on-court performance data from the NBA website, which we then used to create visual representations of players' performance across various attributes not limited in shooting performances.

## II. Methodology

### A. NBA forum word cloud

We obtained Weibo discussion data by using a Python scraper specifically designed for keyword searches [1]. We configured it to fetch weibos containing the keywords *NBA* and *shooters* in 2023. Following the data collection, we cleaned the raw data and translated them from Chinese to English. Then, we counted the frequency of mentions of different NBA shooters. At last, we used Tableau to visualize this data in a word cloud format, enabling easy identification of the most talked-about NBA shooters on Weibo in 2023.

### B. NBA player Shotchart and other shot data visualization

After we get the most frequent mentioned NBA shooters, we tried to visualize their shooting skills in a more direct way. First of all, we used NBA API *shotchartdetail* function to query nearly 20000 pieces of player's shooting data [2]. The table I presents the data description of the outputs.

After the data query, we used Tableau to create a shot chart. The X and Y in the data descriptor represents the coordinates of a certain shot at court. After importing the shot coordinates, we inserted a map, which is a basketball half-court sketch to carry these shot spots. One thing need to be careful is that the parameters of the map (for example the range) should be well-adjusted to match the shot coordinates. We finally added more build-in filters, such as shot zone, shot type, etc for more interaction options. Besides the shot chart, we also provided other visualizations showing the average shot distance of each game he played.

TABLE I
Data Descriptor for NBA shot chart data

| Field | Data Type | Description |
|---|---|---|
| player_id | string | The unique identifier for the player who attempted the shot. |
| player_name | string | The name of the player who attempted the shot. |
| team_id | string | The unique identifier for the team of the player who attempted the shot. |
| team_name | string | The name of the team of the player who attempted the shot. |
| game_id | string | The unique identifier for the NBA game in which the shot was attempted. |
| game_date | date | The date on which the game was played. |
| period | integer | The period in which the shot was attempted (e.g., 1, 2, 3, 4, OT). |
| minutes_remaining | integer | The number of minutes remaining in the period when the shot was attempted. |
| seconds_remaining | integer | The number of seconds remaining in the period when the shot was attempted. |
| event_type | string | The type of event (e.g., "Made Shot," "Missed Shot," "Free Throw"). |
| action_type | string | The type of action or play associated with the shot (e.g., "Jump Shot," "Layup"). |
| shot_type | string | The type of shot (e.g., "2PT Field Goal," "3PT Field Goal"). |
| shot_zone_basic | string | The basic description of the shot zone (e.g., "In The Paint (Non-RA)" for shots in the paint). |
| shot_zone_area | string | The area of the court where the shot was attempted (e.g., "Center(C)" for shots in the center). |
| shot_zone_range | string | The range of the shot (e.g., "Less Than 8 ft" for shots near the basket). |
| shot_distance | integer | The distance from the basket where the shot was attempted (in feet). |
| X_loc | float | The X-coordinate of the shot location on the court (in basketball court units). |
| Y_loc | float | The Y-coordinate of the shot location on the court (in basketball court units). |
| made | boolean | Whether the shot was made (true) or missed (false). |
| game_event_id | integer | A unique identifier for the specific event within the game. |

### C. NBA player Radar Chart

We also created a Radar Chart aiming to evaluate players' performances in a comprehensive perspective. We used the *playercareerstats* function from the NBA API to get players' overall data in the 2022-2023 NBA season. The table II is the data descriptor of our queried dataset. However, making a Radar Chart requires us to compare players at a certain contributes horizontally. Therefore, we did an easy data scaling using the method shown in Algorithm 1. We divided each value at a certain column by the largest value of this column.

After cleaning the data, five dimensions were defined to describe the player: scoring (PTS), shooting (FG3_PCT), rebound (REB), defense (Defense), and assist (AST). Since

comparison needed, radar chart was our best option. By creating four calculated fields (PATH, DEGREE, X, Y), we could assign the according index values to each dimension.

TABLE II
DATA DESCRIPTOR FOR NBA PLAYER STATS AT SEASON 2022-2023

| Field | Data Type | Description |
|---|---|---|
| PLAYER_ID | string | The unique identifier for the player. |
| LEAGUE_ID | string | The unique identifier for the basketball league (e.g., NBA). |
| Team_ID | string | The unique identifier for the team. |
| GP | integer | Games played in the player's career. |
| GS | integer | Games started in the player's career. |
| MIN | float | Total minutes played in the player's career. |
| FGM | float | Total field goals made in the player's career. |
| FGA | float | Total field goals attempted in the player's career. |
| FG_PCT | float | Field goal percentage in the player's career. |
| FG3M | float | Total three-point field goals made in the player's career. |
| FG3A | float | Total three-point field goals attempted in the player's career. |
| FG3_PCT | float | Three-point field goal percentage in the player's career. |
| FTM | float | Total free throws made in the player's career. |
| FTA | float | Total free throws attempted in the player's career. |
| FT_PCT | float | Free throw percentage in the player's career. |
| OREB | float | Total offensive rebounds in the player's career. |
| DREB | float | Total defensive rebounds in the player's career. |
| REB | float | Total rebounds in the player's career. |
| AST | float | Total assists in the player's career. |
| STL | float | Total steals in the player's career. |
| BLK | float | Total blocks in the player's career. |
| TOV | float | Total turnovers in the player's career. |
| PF | float | Total personal fouls in the player's career. |
| PTS | float | Total points scored in the player's career. |

**Algorithm 1:** Normalize Columns in a DataFrame

**Data:** DataFrame $df$, List of columns to normalize $columns\_to\_normalize$

**foreach** $column$ ***in*** $columns\_to\_normalize$ **do**

$max\_value \leftarrow$ maximum value in $df[column]$ ;

// Find the maximum value

$df[column] \leftarrow df[column]/max\_value$ ;

// Normalize the column

**Result:** Normalized DataFrame $df$

## III. RESULTS

In this section, we will introduce our visualization sketches and their corresponding effects on our research target.

### A. Word Cloud

This word cloud shown in 1 plot effectively distills complex textual data into an easily digestible format. By presenting the frequency of mentions as varying text sizes, the word cloud immediately directs attention to the most popular NBA shooters, allowing for quick and intuitive interpretations. This aligns perfectly with the study's objective of understanding the public sentiment surrounding NBA shooters on Weibo.

### B. Shot Chart

The figure 2 shows the specific shot performances at the court. The blue dots are the shots made by the players and the yellow signs are the miss shots of the players. Furthermore,

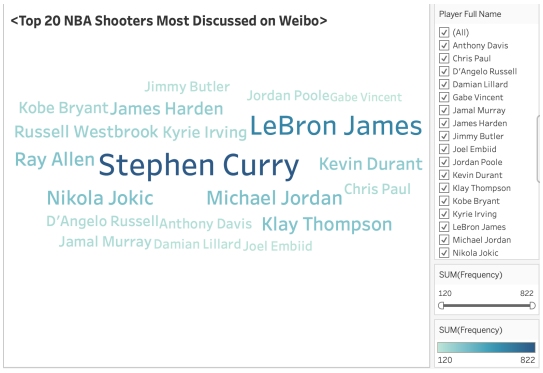


Fig. 1. The Word Cloud Plot of Top 20 NBA Shooters Most Discussed on Weibo in 2023.

we can select a specific player and see more information of his shooting performance. As shown in Figure 3, we selected the shots made farther than 24 feet (outside of three-point lines). Also, we can select other filters for example showing his performances in guest games.

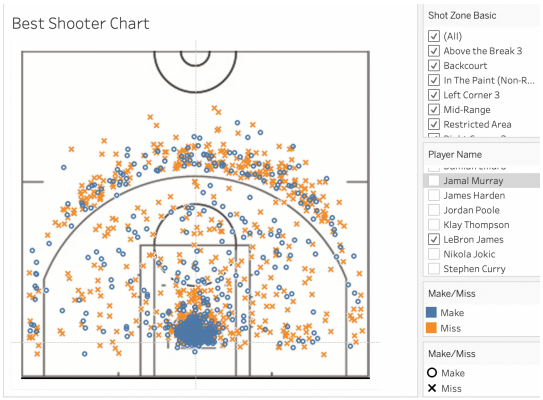


Fig. 2. The shot charts of outstanding shooters in the league.

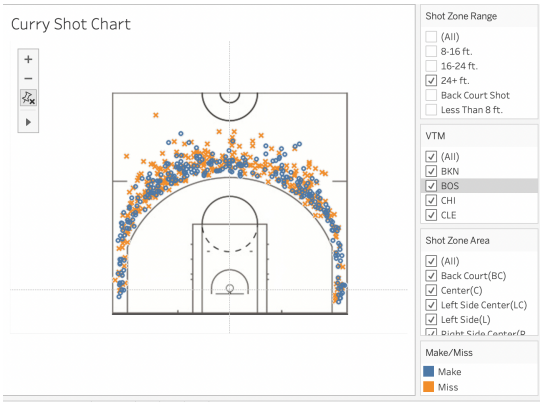


Fig. 3. Shots made 24 feet or farther by Stephen Curry.

We also plotted another box plot to show his shot distance in each game in a more direct way. Each box represents his shot distance information on a specific day. A pie chart is

presented in figure 5 to show one player's shooting distance preference.

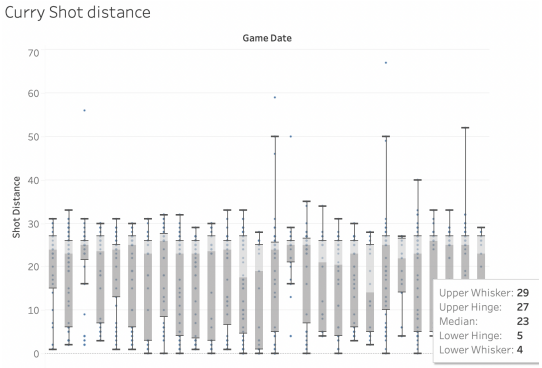


Fig. 4. Box plot showing Stephen Curry's shot distance information.

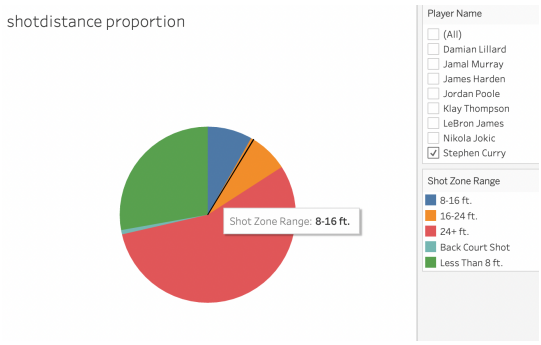


Fig. 5. Pie Chart shows the shot distance of a certain player.

### C. Radar Chart

As shown in the figure 6, we can see Damian Lillard has an outstanding performance in scoring; Stephen Curry has an outstanding performance in shooting; Nikola Jokic has outstanding performances in rebound, defense, and assist. We can simply conclude that the proficiency of certain players lies in their attacking abilities, while defense may not be their forte.

## IV. Discussion

The word cloud visualizes the frequency of mentions for various NBA shooters on Weibo in 2023. The word cloud offers immediate insight and is easy to understand, saving users time in data interpretation. However, it lacks depth and nuance, not providing reasons behind the frequency of mentions or the sentiment of those mentions. The data is limited to Weibo and the year 2023, which may not be fully representative. To overcome these limitations, future iterations could incorporate sentiment analysis and expand the scope to include multiple platforms or years.

From the Shot Chart we can easily prove the changing trend of the playing style in NBA that the shot attempts in the paints and at the 3-point lines is much more than the attempts in the mid-range. We also find that Guards are more productive

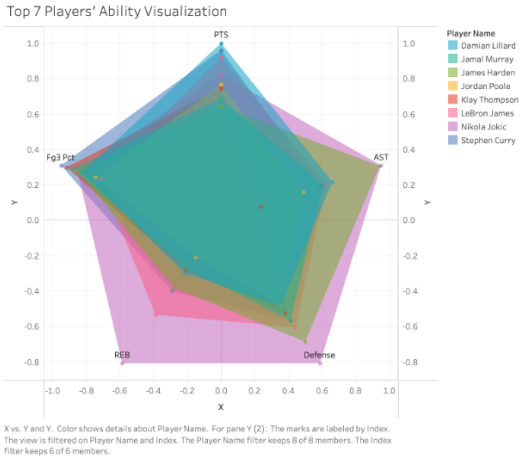


Fig. 6. Radar Chart that shows the player attributes comparison.

in 3-points shots and Stephen Curry is the most productive and efficient shooter. Some players have different preference zones for 3-point shooting. For example, Jamal Murray is a great shooter, however, he shots at the left side more than the right side. Also, Jordan Poole, the youngest player in the list, prefers to shoot at wings instead of the corners. Interestingly, Jokic, the MVP player in NBA only shoots at the three-point arc. Nevertheless, although we can easily find the characteristics of players' shooting prefrences through the shot chart, it cannot provide information in a quantitatively efficient way. In the future, we will look for more useful techniques to make visualizations on exploring NBA player's data.

The Radar Chart is an interesting way to compare players in a comprehensive insight. However, the attributes provided are not precise enough to represent the player's abilities in a certain aspect. For example, player's defense ability cannot only consider his stealing and block abilities. More advanced defensive data should be mined to fully show one's performances in defense. In the next step of our research, we will focus on mining more advanced data and do more comprehensive evaluation on player's performances.

## V. Conclusion

This study employed word cloud, shot charts, and radar charts to analyze NBA players' performance and popularity. The visualizations revealed key insights into shooting preferences and multi-faceted skills, generating actionable data for fans, analysts, and players. While offering a snapshot of current trends, this research also identifies the need for more advanced metrics and broader data analysis for a comprehensive understanding of player performance.

## Acknowledgment

We express our sincere gratitude to Professor Tong for her hard work. Also, we have many thanks to Forbes, who is one of the founders of the NBA API, for his patient assistance when we encountered problems while using his API.

## REFERENCES


[1] "dataabc/weibo-search: 获取微博搜索结果信息，搜索即可以是微博关键词搜索，也可以是微博话题搜索," GitHub, Apr. 19, 2023. https://github.com/dataabc/weibo-search (accessed Sep. 15, 2023).

[2] "swar/nba_api: An API Client package to access the APIs for NBA.com," GitHub, Sep. 15, 2023. https://github.com/swar/nba_api (accessed Sep. 16, 2023).