

Shooting Data visualization and analysis

Peng Yuan - INFSCI 2415
Instruct by Prof. Lingfei Wu

Chart display:

Stephen Curry Postseason Shooting stats chart

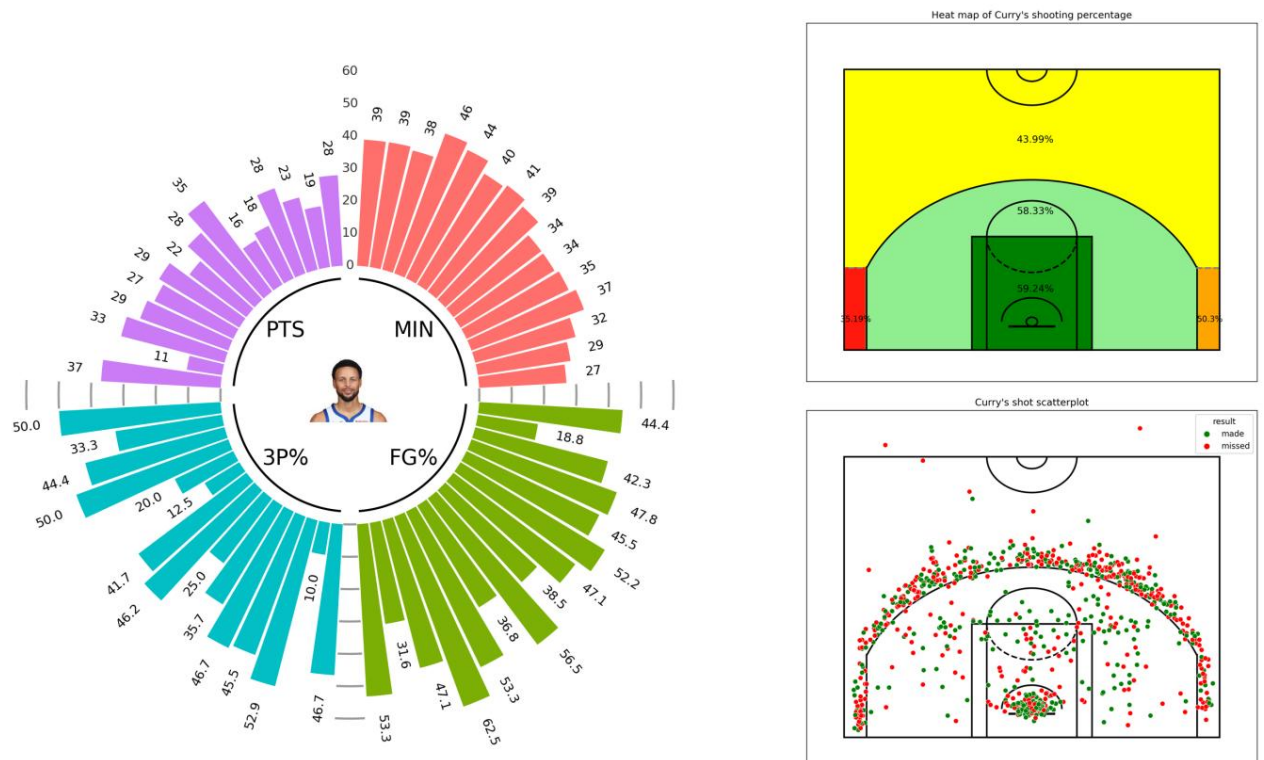


Figure Introduction

The main image on the left is a polarized bar graph, and the image presents Curry's 17-18 playoff season stats, including minutes played, points scored, three-point percentage and field goal percentage. In the center of the chart is the player's head-shot and classification name, and the top of the bar graph shows the value of each stat.

The two images on the right side are a heat map of shooting percentage by region, and a scatter plot of specific shots made. The heat map shows how well players are hitting shots within each region, and the scatter plot shows the player's preferred shooting spots.

The three graphs were spliced together to better compare and analyze the data. For example, the scatter-plot and heat-map are used to analyze Curry's preference for offensive tools and how he hits them. And the bar chart on the left side reflects

Curry's overall performance data in the playoffs, echoing the graph obtained from the specific data visualization on the right side.

Concept and analysis statement

In the main chart on the left, the red bar represents minutes played, the purple bar represents points scored by the player, the blue bar represents three-point percentage and the green bar represents the player's shooting percentage. The values shown for three-point percentage and field goal percentage are percentage shares.

The following results were found by analyzing the data in the main chart. In a full 48-minute NBA game, Curry's playing time is almost in the 30-45 minute range, which represents how important Curry is to the team. Looking at the scoring numbers, we see that Curry has scored more than ten points in every game. Even in the case of the playoffs there is a single game of 35 points, which is the performance of the superstar. Even more surprising is that Curry's three-point shooting percentage is above 40% in most cases, except for a few lower games, which is a very scary statistic. With the high level of defensive intensity in the playoffs, his three-point percentage is almost the same as his total shot mileage. This performance is without a doubt unprecedented.

The basketball court on the right consists of lines, boxes and arcs. The basketball court is constructed by adjusting the position of the lines and arcs in the constructed planar coordinate system. In the heat map of shooting percentage, the color representation of the shooting percentage sorted from low to high is: red, yellow, orange, light green, dark green. In the scatter plot, each point represents a shot taken by a player, with red points representing missed shots and green points representing hit shots.

The scatterplot shows that Curry prefers to shoot threes from 45 degrees on the right side when at the top of the arc, but more and more prefers to shoot threes from the left bottom corner when in the bottom corner. This may have to do with the Warriors' playbook, where the left side is usually the weak side of the defense, so the left corner is usually a better shooting spot for bottom corner threes. Most two-point shots are scored at the rim, and Curry is far less selective about mid-range shots as an offense than he is about layups and three-point shots. A comparative analysis of the heat map of hits reveals that while there are more shots from the left bottom corner, they are hit at a much lower rate than other areas. The top of the arc area as a whole is hitting over 40 percent of its three-point shots and almost 60 percent of its two-point attempts. All of these hit rates are a direct reflection of Curry's scoring efficiency and preference for offensive tools.

Data and method

There are two main datasets used, one is publicly available data from the web that contains all the players' 17-18 shooting seasons, and the other set of data comes from Kaggle's website, which contains Curry's specific stats for each of his games during the 2009-2023 seasons.

The primary data source for the bar chart on the left is the detailed game-by-game data set for Curry brought to you by Kaggle. After sifting through Curry's playoff data from the 17-18 seasons, the four types of data shown above were visualized. In the composition, '0' was inserted in the middle of the four kinds of categorical data to realize the insertion of the scale bar, in addition, by adjusting the angle and rotation direction of the bar chart, the ring display of the bar chart was realized. In order to make the graph display more aesthetically pleasing, the hit rate data is displayed as a percentage, matching the numerical scales of scoring and playing time.

The primary data source used for the two plots on the right is the publicly available 17-18 season player shooting Dataset. After filtering Curry's shooting data, a specific visualization of shooting hits and coordinates is presented. After partitioning the basketball court through the constructed right-angle coordinate system, a scatter plot is constructed based on the coordinates of the shooting position of each point in the data. After that, the area is divided according to the coordinates, and the shooting percentage situation in each area is calculated. Among them, the way to determine whether the three-point area or two-point area shooting is to calculate the distance from the shooting position to the center of the circle formed by the three-point line. After calculating the hitting rate, different colors are used to fill in the different regions.

Github Code and Dataset

The specific code and dataset have been uploaded to Github. The correspondence between the dataset and the code has also been explained in the README file.

Here is the link to the work:

<https://github.com/York-Peng/INFSCI-2415-INFORMATION-VISUALIZATION/tree/main>