



**COM-480 Process Book Report**

# **TAFRIKA**

**Spring 2024**

---

**Raed Ben Mustapha – Farah Briki – Khalil Achache**

# INTRODUCTION

The aim of this project is to offer a more detailed understanding of basketball analytics. Our visualizations are crafted for basketball enthusiasts eager to gain deeper insights into their favorite players and teams. They will also be valuable for coaches, helping them make informed decisions and devise plays that could lead to higher-percentage shots. Additionally, these insights can assist players in honing their game and shot selection by learning from the improvements and efficiency patterns of successful players.

As passionate NBA fans and data scientists, we were motivated to work on NBA visualizations because we are fascinated by the challenge of interpreting the vast array of statistics and data to uncover intriguing insights. Although there are many existing visualizations and analyses related to the NBA, we had some ideas for new perspectives that, to our knowledge, have not been thoroughly explored yet.

# DATA

For this project, two datasets were leveraged.

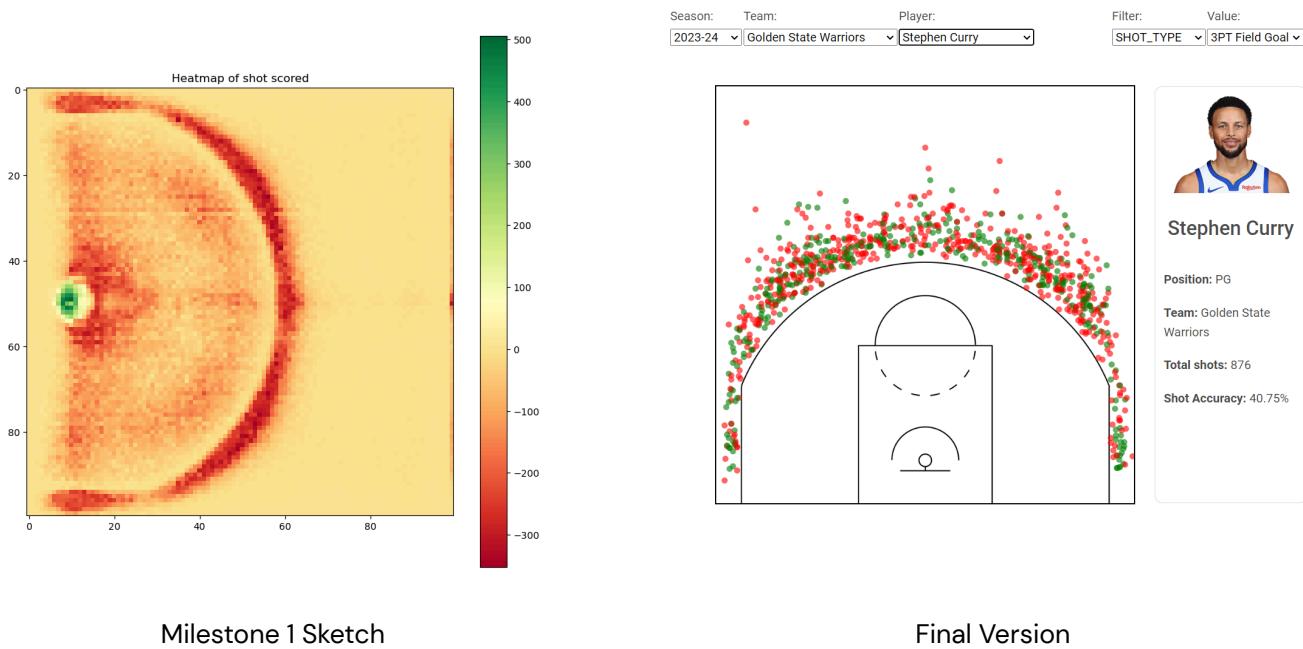
The primary dataset was sourced from a comprehensive compilation available on GitHub from [this link](#). The dataset provides a historical record of 4M NBA shots from 2004 to 2024, detailing player and team information, shot outcomes, and spatial data regarding shot locations.

The secondary dataset was collected from [Kaggle](#) and enriches the analysis with 128k shots with specific metrics from the 2014/2015 season, such as shot clock, dribbles and touch-time before the shot.



# First Plot: Spatial Visualization

This interactive visualization allows for an in-depth exploration of NBA players' shooting performance, presenting data in a detailed and intuitive manner. It enables users to gain insights into shot patterns and efficiency by providing several analytical features.



## Features:

- Selection of Season, Team, and Player**
  - Users can begin their analysis by selecting the desired season from a dropdown menu, followed by the team to refine the player list.
- Comprehensive View of Shots**
  - The scatter map displays all shots taken by the chosen player during the season.
  - Each dot on the half-court diagram represents a shot, with green dots indicating successful shots and red dots indicating missed shots.
- Filtering by Categories**
  - The visualization includes options to filter shots based on specific categories, such as "Three-Point Field Goals," to display only three-point attempts. Additional filters allow users to focus on specific areas, such as shots from the left corner, right corner, or other regions of the court.
- Detailed Player Information**
  - A player card on the right provides key statistics for the selected player, including position, team, total number of shots taken, and shot accuracy for the selected shot zone.

## Insights Provided

- Shooting Patterns and Hotspots**
  - The scatter map helps identify the player's shooting hotspots and areas with lower accuracy, revealing where the player is most effective on the court.
- Performance Under Different Conditions**

- Filtering options allow analysis of how the player's performance varies across different shot types and court areas, uncovering tendencies such as proficiency in corner threes versus above-the-break threes.

- **Comparative Analysis**

- The visualization facilitates the comparison of shooting performance among different players, valuable for comparing teammates to know where to position players in offense or evaluating potential opponents to have a better defensive strategy.

- **Player Development**

- Coaches and players can develop targeted training drills to improve shooting from areas where they are less efficient. By focusing on these specific aspects, players can enhance their skills and add new strengths to their repertoire.

## Road to Final Version

**Initial Plan:** Initially, we aimed to create a groundbreaking visualization combining both offensive and defensive datasets. The idea was to allow users to place an offensive player in a specific spot on the court, position a defensive player in front of him, and then calculate the probability of a successful shot based on the defensive pressure. This innovative approach was designed to give deep insights into how specific defenders affect the shooting efficiency of different players.

**Challenges Encountered:** However, we faced several challenges that necessitated a shift in our approach:

1. **Limited Dataset:**

- The defensive dataset covered only one year, which provided insufficient data for robust analysis.
- The dataset included information about the closest defender and their distance from the shooter, but merging this with offensive data proved complex due to the limited number of data points.

2. **Complexity of Calculations:**

- Calculating a meaningful shot probability based on the defensive pressure required a comprehensive dataset and complex modeling, which was difficult with the available data.

**Revised Approach:** Given these challenges, we decided to focus on a more feasible and straightforward visualization that will still give us valuable insights:

- We shifted our efforts towards developing a made/missed scatter plot, which displays the locations and outcomes of shots taken by a player on a basketball court diagram.
- This plot provides valuable insights into shooting patterns and efficiency, offering a practical and informative tool for analysis.

## Steps to Final Plot:

1. **Data Preparation:**

- Cleaned the data. For instance, the X,Y location of the shots were not in the same range for the seasons between 2019–2021

2. **Interactive Features:**

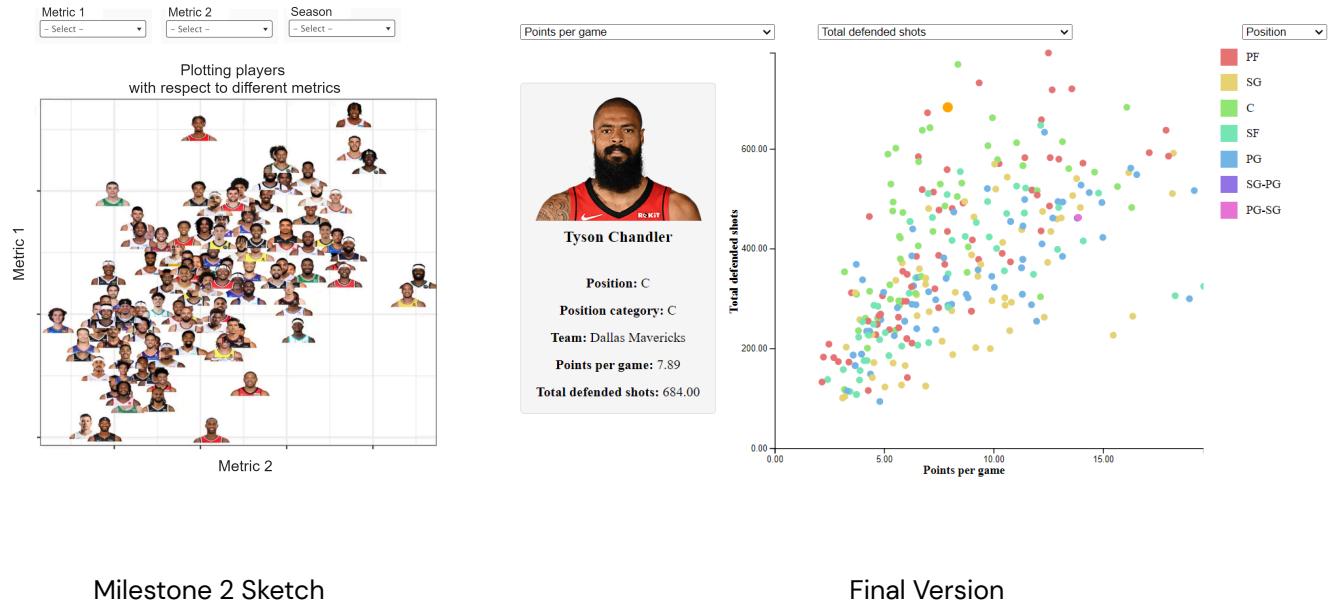
- Developed interactive dropdown menus for users to select the season, team, and player (making sure you only get the next dropdown only if you select the one before it).
- Implemented filters to allow users to focus on specific attributes.

3. **Visualization Development:**

- Created the scatter map using D3.js, plotting each shot as a dot on a half-court diagram, with green dots for successful shots and red dots for missed ones.

# Second Plot: Player Performance Comparison

This interactive scatter plot visualization enables exploration and comparison of NBA players' performance based on various offensive and defensive metrics during the 2014–2015 season. It provides a powerful tool for analyzing player efficiency and identifying standout performers.



## Features

- **Selection of Metrics for Comparison**
  - Users can begin by choosing two metrics from the dropdown menus to plot on the x-axis and y-axis. These metrics can be either offensive or defensive. For example, selecting "Points per Game" for the x-axis and "Three Pointers Made per Game" for the y-axis helps identify high-scoring players with effective three-point shooting.
- **Identifying Player Performance**
  - The scatter plot displays each player as a dot, positioned according to their performance in the selected metrics. Hovering over a dot reveals detailed information about the player, including their position, team and chosen metrics values for the player.
- **Categorization by Attributes**
  - A dropdown menu on the right allows categorizing players by team, position, or other attributes. Each dot is color-coded based on the selected category. Selecting "Team" color-codes the dots according to each player's team, facilitating comparisons within the same team or across different teams.
- **Filtering by Category**
  - To focus on specific groups of players, users can filter the scatter plot by selecting a category. This enables a more detailed analysis of player performance within a single team, aiding in identifying the best defenders, most efficient scorers, and optimal strategies for different players.

## Insights Provided

- **Offensive and Defensive Efficiency**
  - By selecting appropriate metrics, users can identify players who excel both offensively and defensively.
- **Player Roles and Strategies**

- Comparing metrics such as "Shot Accuracy" and "Dribbles Before Shot" differentiates between isolation players and catch-and-shoot players, providing insights into their scoring strategies and effectiveness.

- **Team and Position Analysis**

- Categorizing and filtering players by team or position allows detailed analysis of specific groups. This helps in understanding team dynamics, such as identifying the best defensive player for clutch situations or the most efficient scorer for the team.

- **Strategic Decisions**

- Coaches and analysts can use this visualization to make data-driven decisions about player roles and strategies. For example, determining whether a player should focus on catch-and-shoot opportunities, create their own shots through dribbling, or pass it to a better efficient teammate.

## Road to Final Version

**Initial Focus:** For this visualization, we concentrated on the 2014–2015 season data, which offered a wealth of insightful columns, including defensive metrics, dribbles, and touch time before the shot. Initially, we considered using this data for our first plot to combine offensive and defensive insights. However, due to the desire to cover multiple years in the first plot and the complexity of using the defensive spatial data, we decided to fully leverage the detailed metrics available in the 2014–2015 dataset for this comparative analysis.

### Rationale:

- **Rich Dataset:**

- Key columns included data on dribbles before the shot, touch time, and defensive pressure, which are critical for understanding player efficiency and strategies.

- **Focused Analysis:**

- By focusing on a single season, we could delve deeper into the nuances of player performance, using the full range of available metrics to position players based on their offensive and defensive contributions.
- This approach allowed for a more detailed and accurate comparative analysis, highlighting standout performers within the context of that season.

### Implementation Steps:

1. **Data Integration and Cleaning:**

- Cleaned the 2014–2015 dataset and merged it with our all\_shots dataset for that season, we faced some problems matching every shot from both datasets. which we fixed and ended up with most of the shots.
- Calculated various insightful metrics into a unified dataset, preparing it for visualization.

2. **Metric Selection and Plotting:**

- Developed dropdown menus for users to select two metrics to plot on the x-axis and y-axis. Metrics included offensive stats like points per game and defensive stats like opponent shot accuracy when the player is the primary defender
- Created a scatter plot using D3.js, plotting each player as a dot based on their performance in the selected metrics.

3. **Interactive Features:**

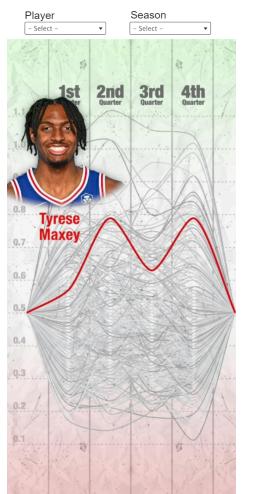
- Implemented hover functionality to display detailed player information when hovering over a dot, including position, team, and key statistics.
- Added categorization options, allowing users to color-code players by team, position, or other attributes.

4. **Filtering and Detailed Analysis:**

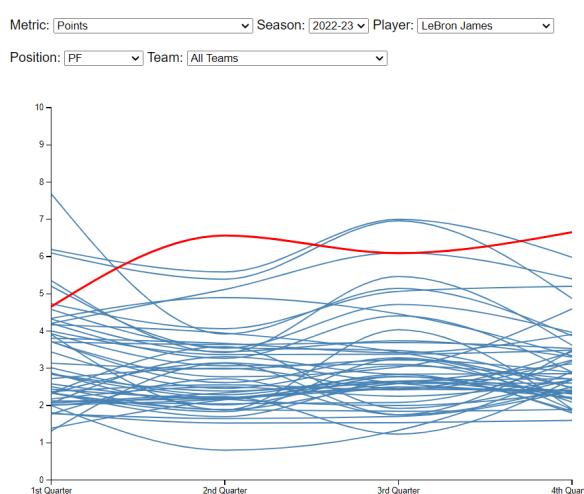
- Developed filtering capabilities to enable users to focus on specific groups of players, such as a particular team or position.
- Enhanced the plot with detailed player cards that appear when a dot is hov, providing in-depth performance metrics for selected players.

# Third Plot: Quarterly Performance

This interactive line graph visualization provides a detailed analysis of NBA players' performance across each quarter of the game for the selected season. It highlights how players perform in different quarters, emphasizing those who excel in crucial moments.



Milestone 2 Sketch



Final Version

## Features

- **Selecting Metrics for Analysis**
  - Users begin by choosing the performance metric they want to analyze from the dropdown menu. Options may include points per quarter, field goal percentage (FG%), three-pointers made, and other relevant statistics. For example, selecting "Points" allows users to see how players score in each quarter.
- **Choosing Season, Team, and Player**
  - Dropdown menus are used to select the season, team, and player for analysis. For instance, the visualization can display data for Kyrie Irving from the 2022-23 season.
- **Exploring the Performance Lines**
  - Each line on the graph represents a player's performance across the four quarters. The x-axis denotes the quarters (1st, 2nd, 3rd, and 4th), while the y-axis shows the value of the selected metric. Hovering over or clicking on a line highlights the player's performance details, and a player card appears on the right side of the graph, showing key statistics.
- **Player Card Details**
  - The player card provides detailed information about the selected player's performance for the chosen season.
- **Filtering by Team and Position**
  - The graph can be filtered by selecting specific teams or positions from the dropdown menus. This helps narrow down the analysis to particular groups of players. For example, choosing "Dallas Mavericks" displays lines only for players on that team, making it easier to compare their quarter-by-quarter performance.

## Insights Provided

- **Quarterly Performance Insights**
  - This visualization helps identify how players perform in each quarter, revealing patterns such as strong starts, mid-game consistency, or clutch performances in the fourth quarter.

- **Clutch Performance Identification**
  - By focusing on the fourth quarter, users can spot players who significantly improve their performance in crucial moments. This is valuable for recognizing clutch players who can be relied upon in tight game situations.
- **In-Depth Player Analysis**
  - Clicking on a player's line provides a detailed view of their quarter-by-quarter performance, offering insights into their scoring trends, shooting efficiency, and other key metrics throughout the game.
- **Team Strategy Development**
  - Coaches and analysts can use this tool to develop game strategies based on player performance trends. For example, identifying players who excel in the final quarter can inform decisions on who to rely on during critical periods of the game.

## Road to Final Version

**Initial Concept:** Our initial concept for this project was to create an innovative visualization tool that would allow users to analyze NBA players' performance in a detailed and intuitive manner. We envisioned a platform where users could explore how players perform in each quarter, identifying patterns and recognizing clutch performers. This idea led us to develop the interactive line graph visualization.

**Challenges Encountered:** During the development process, we faced several challenges that required us to adjust our approach:

1. **Data Integration and cleaning:**
  - We used the full all\_shots dataset to have multiple seasons and most importantly recent ones. In a later step we discovered that players changing teams mid-season was causing weird lines in our visualization so we had to do some preprocessing for our players before exporting data needed for the visualization.
2. **Metric Selection and Customization:**
  - Deciding on the most relevant metrics for analysis required careful consideration. We wanted to offer users a variety of options, including points per quarter, field goal percentage (FG%), and three-pointers made, among others even though it would have been great to have more detailed metrics such as defensive metrics.
3. **User Interface and Experience:**
  - We aimed to create a seamless experience where users could quickly filter and explore performance data. Even though data of all players looks kinda messy we decided to keep it to keep track of a player performance compared to the whole league, and by filtering on teams you can see only players of that team with a more appealing graph.
  - Choosing the right line width was complicated because we have a lot of players on the graph and making the line bigger makes the graph incomprehensible meanwhile making it smaller makes clicking on a player more complicated so we had to find the best balance between the two.

## Development Steps:

1. **Data Collection and Preparation:**
  - Data was cleaned and formatted for seamless integration into our visualization tool.
2. **Building the Interactive Interface:**
  - We developed dropdown menus for selecting the season, team, and player. This allowed users to customize their analysis and focus on specific players or teams.
  - The interface was designed to be user-friendly, enabling easy navigation and interaction with the data.
3. **Creating the Line Graph Visualization:**
  - Using D3.js, we built the line graph to display player performance across the four quarters. Each line represents a player's performance, with the x-axis showing the quarters and the y-axis displaying the selected metric.
  - We implemented interactive features such as clicking on lines to highlight player performance details and display player cards with key statistics.

# Peer Review

Apart from the central choices for the projects that were taken between all members of the group, we provide below a table detailing the individual work done by each member of the group :

Ali Raed Ben Mustapha	Farah Briki	Khalil Achache
<b>Dataset:</b> <ul style="list-style-type: none"><li>• data for the 1st plot</li><li>• data for the 2nd plot</li></ul>	<b>Dataset:</b> <ul style="list-style-type: none"><li>• data preprocessing</li><li>• data for the 3rd plot</li></ul>	<b>Website Setup:</b> <ul style="list-style-type: none"><li>• Found the template</li><li>• Set up the initial website</li></ul>
<b>Second Plot</b>	<b>Third Plot</b>	<b>First Plot</b>
<b>Process Book</b>	<b>Process Book</b>	<b>Process Book</b>
<b>Screencast</b>		