
MILESTONE 3: PROCESS BOOK

Ruiqi Yu(340546), Junzhe Tang(340747)

Contents

1	Brief Introduction	2
1.1	Motivation	2
1.2	Brief Work	2
2	Visualization Implementation	2
2.1	Geographical data visualization: Map	2
2.2	Winning percentage over the years: Line Chart	3
2.3	3-pointers analysis: Line and Bar Chart	3
2.4	Teams' Style Comparison: 5 Star Chart	3
2.5	Player's Information	4
2.6	Player's Shoot Distribution: Box Chart	5
3	Challenge in the project	5
3.1	Map	6
3.2	Data Resource	6
4	Peer Assessment	6
4.1	Junzhe Tang	6
4.2	Ruiqi Yu	6
5	Discussion	6

1 Brief Introduction

1.1 Motivation

In the early stage of our project, we found that basketball information is not well visualized in terms of data. NBA is the most famous basketball league all over the world, but the NBA official website and a famous reference website Basketball-Reference seldom offer visualization modules for NBA fans. So the goal of our project is to realize some interesting data visualization in the sport area.

We think that the data visualization of sport events can contribute the application of data, which can have a great impact on modern basketball events. Currently, moneyball theory is a hot theory in many sport events and lots of teams' managers from different sport events are using it. [1] In our eyes, moneyball theory makes teams' managers and even fans trust and focus more on the data, and data visualization is one important part of using data.

1.2 Brief Work

Our project is mainly posted on nbadata.com. In our project, we have done: 1) Teams home geographical data visualization; 2) Teams' Win Percentage change over the years; 3) 3-pointers analysis visualization; 4) Teams' Style Comparison from 2 different ways.

2 Visualization Implementation

In this part, we will introduce our implemented visualization and the paths to achieve them.

2.1 Geographical data visualization: Map

NBA has 30 teams and they have their own stadiums in different cities. NBA teams are divided into western league and eastern league, which is connected to the playoff schedule. Therefore, knowing the location of teams are important, which is a kind of geographical information. Map is a good tool to realize the visualization of geographical information. We use the map to achieve the data visualization of the location of teams' home. As shown in Figure 1b, we can know the name of the state and which teams are in the state by moving the cursor on the map.

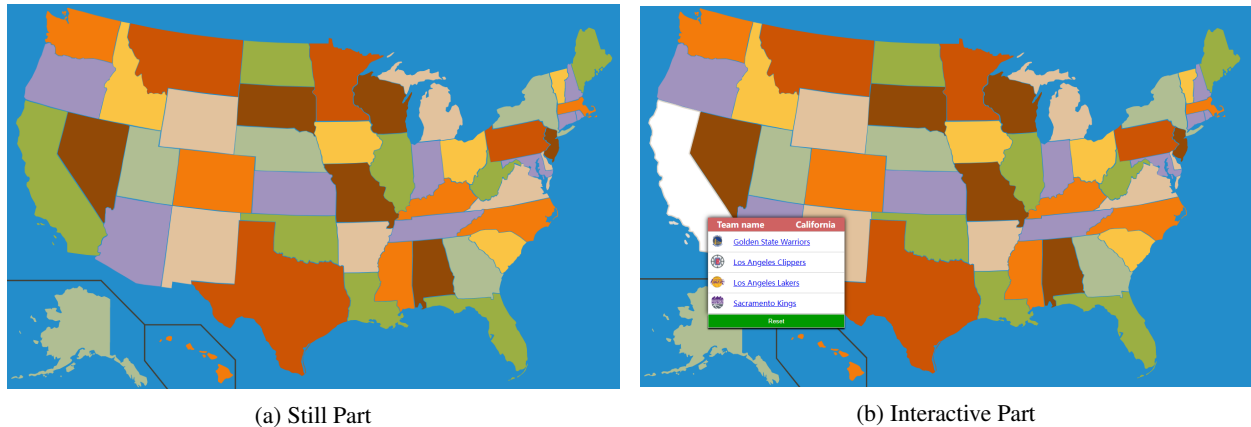


Figure 1: Map Functioning

To achieve this visualization, the implementation steps are:

1. **Storage:** Store the name of states as an object → Store each state team name as an object, or as an array if a state has more than 2 teams ('State Name': ['Team Name 1', 'Team Name 2']) → Store each team's logo as an object.
2. **Map:** The map is made of multiple vector faces, with specific data available online for each state in the United States. We construct multiple face vectors in equal scale and visualize them, give them different colors, and finally add responses.

3. **Cursor triggers an event:** Add a cursor movement event to each state element. (for loops are added for each state. When the cursor moves to one state, the function obtains the cursor's x and y coordinates and also get the label.) → Get the attribute label of the element and take the corresponding value by the original first step of storing the object. → Obtain the number of teams, team name and team log for each state, output the values to the page and add them to the pre-designed element template

Toronto Raptors has been placed in New York State wrongly, and it will be discussed later in chapter 3.1.

2.2 Winning percentage over the years: Line Chart

Winning percentage is an important factor to know the performance of a team. For most basketball websites, they offer the winning percentage of teams in a specific year. We hope to offer a chart of teams' winning percentage along years in Figure 2, from which people can know the ups and downs of teams' performances. By moving the cursor on the map, visitors can get the specific information. As shown in Figure 2b, the chart offers the selection function. If people want to focus on one or a few teams, you can choose them in the selection function zone at the right side of the chart.

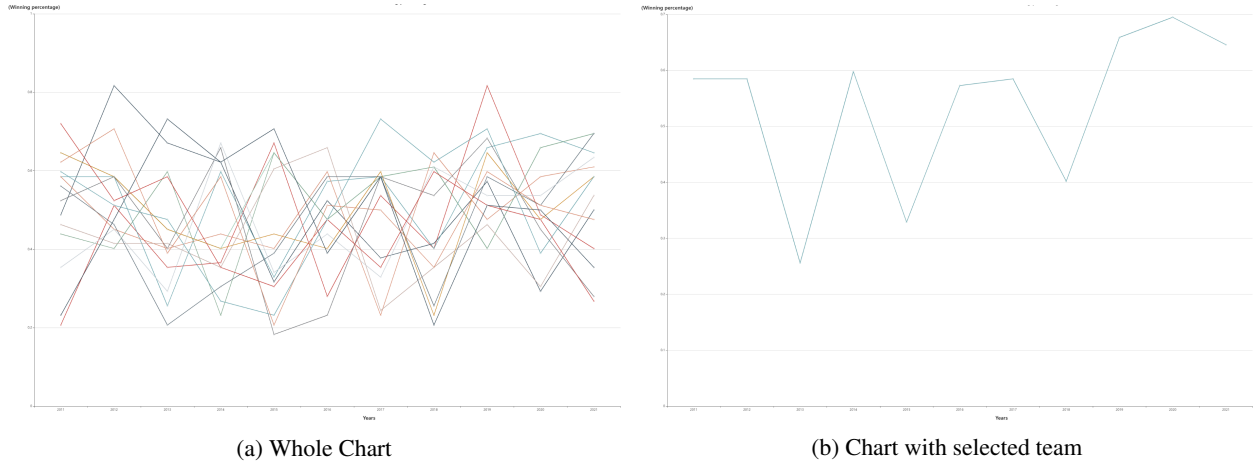


Figure 2: Winning Percentage Over the Years Visualization

The achievement steps are similar for this and the next section, so we only discuss it here:

1. **Use dom to get elements**
2. **Configure the graph data by adding to the option object**(let option = {})
3. **Set each configuration in Option:** Whether the chart is a line chart or a bar chart can be verified here.
4. **Generate a canvas for the data configured**

2.3 3-pointers analysis: Line and Bar Chart

One application of the aforementioned moneyball theory on basketball event is to raise the tactics importance of 3-pointers. For the traditional basketball theory, players should get as close as possible to the basket, which will raise the hit rate, and thus 3-pointer rate is not high. But a new theory was born that 3 pointers is one of the most efficient score methods since the expected value of 3-pointers is higher in a low hit rate within acceptable range. Due to the new theory, more players have more attempts of 3-pointers. Therefore, we want to do some visualization on 3-pointers related information to know the change of 3-pointers strategic position of different teams over the years.

2.4 Teams' Style Comparison: 5 Star Chart

Different NBA teams have their own styles and tactics due to the difference of abilities of players and coaches, which will lead to different match's data. The visualization of some specific information can show the differences between teams. Points(PTS), blocks(BLK), rebounds(REB), steals(STL), and assists(AST) are five basic teams data, and we use them as our compared targets. We use the following equation to normalize 5 kinds of data.

$$x^* = \frac{x}{max} \quad (1)$$

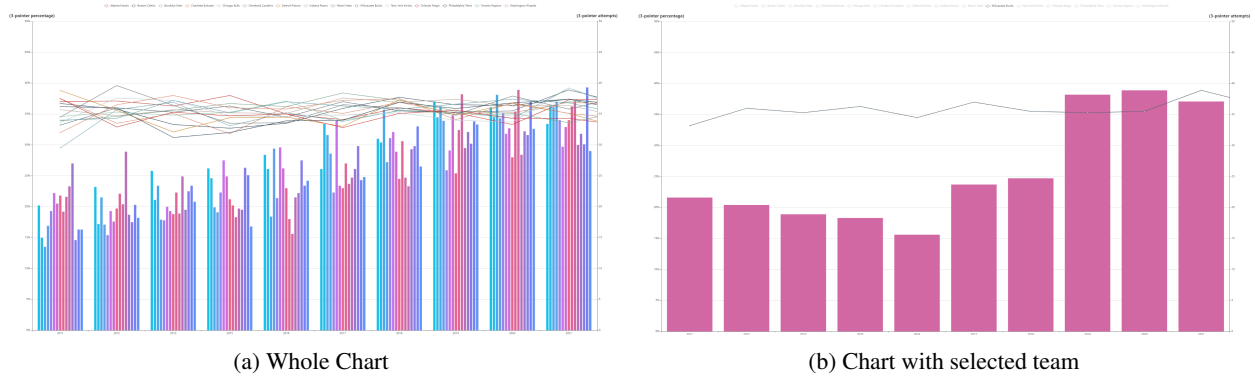


Figure 3: Winning Percentage Over the Years Visualization

With the chart in Figure 4, visitors can compare two teams in terms of 5 factors. Usually, higher x^* means better performance in the area and owing the larger area means that the team has more comprehensive competitive strength. The 5 star chart can achieve comparison between two teams in an intuitive way.

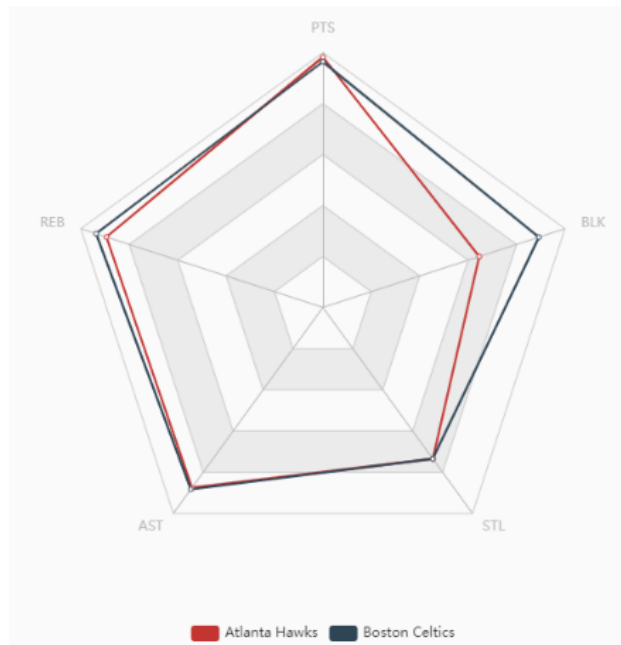


Figure 4: 5 Star Chart

The implementation of the chart is as followed:

1. **Design the chart style.**
2. **Use for loops to add events for t_{qd} in class.**
3. **Store teams' names and logo.**
4. **Add a red team selection functioning area.**

2.5 Player's Information

We also offer the visualizations of the player. In the chart, field goal percent(FG%), 3-pointer field goal(3P%), free throw(FT%), minutes played(MIN), and other aforementioned factors are included in this chart.

To implement the color bar in the percent-related data, the steps are:

1. for loops get the class name present in the div: xrbfb.
2. Set a background color as gradient at first. The absolute positioning element width and height are equal to 100%. During the loop of xrbfb, loop the number of sub-elements.

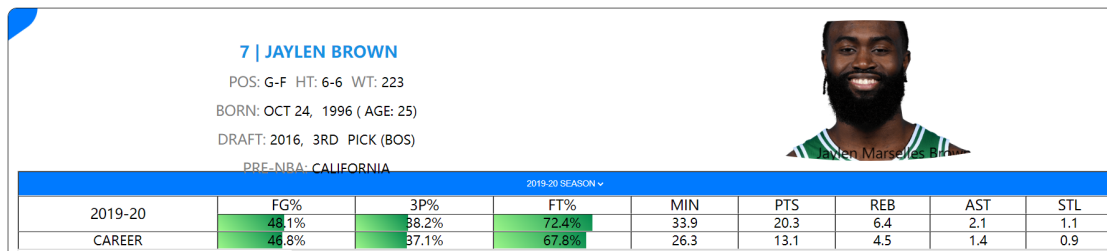


Figure 5: Player Information

2.6 Player's Shoot Distribution: Box Chart

The shoot attempt and percentage of players are also an important factor to study a player. We are inspired by the heat map. If we can get the specific shooting data of the player on the court, then we can visualize the player based on his shooting position, hit rate and number of shots. In the chart, visitors can know which area has a more attempt and percent from the area.

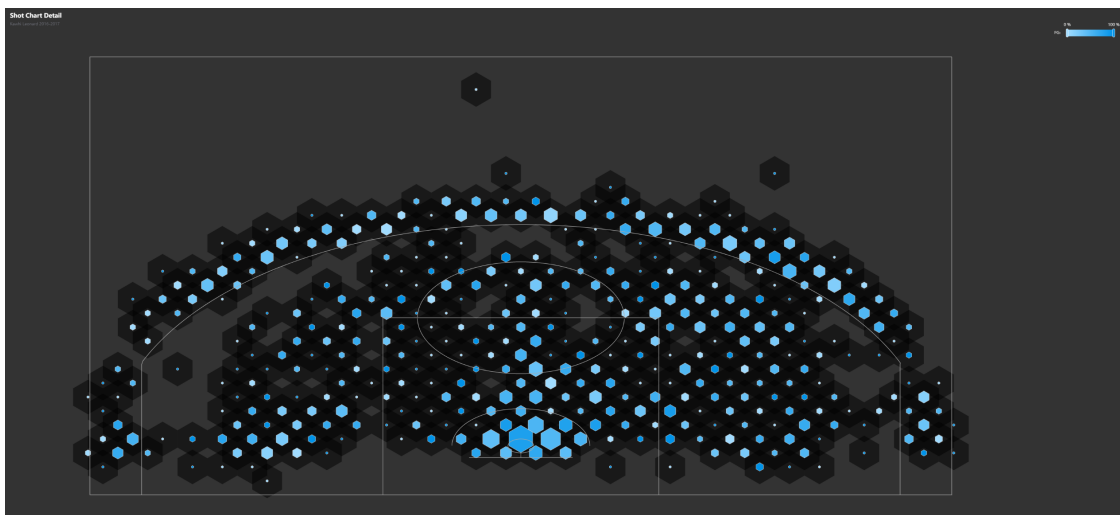


Figure 6: Player Shoot Box Chart

The implementation main steps are as followed:

1. **Court Division:** The basketball court is set as a two-dimensional plane, and the plane is divided equally into hexagons of the same size.
2. **Generate the distribute:** The size of the colored part of the hexagon is determined by the number of shot points falling into the hexagon, and the color of the hexagon is determined by the hit rate. The hexagon with no shooting point data falling into it is black

3 Challenge in the project

Before mentioning some technical challenges, we want to mention that one of our teammates just dropped the course and the lack of one teammate is a big challenge. So if our work is not so good, please forgive us.

3.1 Map

As mentioned before, Toronto Raptors is placed in the wrong state. Toronto Raptors is one special team because its home is located in Canada instead of United States. We can find the map of United States, but we cannot find the map of United States with the state of Toronto. To resolve the tricky thing, we place Toronto Raptors in New York State since it is very close to New York State.

3.2 Data Resource

The initial dataset mentioned in milestone 1 can offer basic data. For some charts we hope to implement, the data is hard or cannot be drawn from the dataset. Besides, we cannot find related dataset. Therefore, we need to search for some of them and build our own dataset, but little experience of using web makes us build it in a raw method and it increases the workload.

4 Peer Assessment

4.1 Junzhe Tang

Junzhe is our website manager. He has lots of experiences of building websites and achieve data visualizations. So he mainly work on the building and maintaining the website. If we have some tricky ideas on visualization, he is the key people to finalize the achievement.

4.2 Ruiqi Yu

Ruiqi has taken on the job of looking for, preprocessing all the datasets throughout our project. Since he is interested in basketball event and often surfs sport websites, most of the visualization ideas are offered by him. He also wrote and organized the process book.

5 Discussion

Our project work is kind of different from the initial plan for these reason:

- 1) **The leave of one teammate forces us to adjust our work structure.** For some charts, we only **offer visualizations for part of the data**, as one year instead of ten years as planned, since we think incomplete information's chart can display our implementation and hope to avoid repeating work and thus decrease the workload;
- 2) We think the visualization implementation methods of some initial goals are similar. We have changed our target and we implement different and more visualizations.

References

- [1] *Moneyball*. 2022. URL: <https://en.wikipedia.org/wiki/Moneyball> (page 2).