Basic Info

[NBA DATA VISUALIZATION](To be changed)

Yanran Ma, carriema
Qianchao Nie, nieqianchao
Shjie Jiang, jiangshijie

Background and Motivation

The National Basketball Association (NBA) is the pre-eminent men's professional basketball league in North America, and is widely considered to be the premier men's professional basketball league in the world. To evaluate player's and team's performance, many statistics are collected and available to analysis (e.g. http://stats.nba.com/). However, it's not very convenient for people to easily and intuitively get information behind the numerous statistical datasets.

This project is designed to give the people easy access to NBA statistical data available from stats.nba.com in a form of live visualization. For example, via the data visualization, people can easily know the information of players or teams, how the team's or player's performance change during the time, where is the most comfortable position that some plays shots their balls, etc. So people don't need to collect, clean or analyze the data, what's they need is just move or click their mouse, they can get an intuitive way to understand some NBA statics.

Project Objectives

 Some question we would like to answer through our data visualization :

Team level:
- The teams' information (e.g. geographical information via map visualization.)
- The wining rate of a team
- The ranking changes of the total teams during times
- Game history of one selected team with all others
- Change of teams' and players' performance during time
- Team performace comparision(Points, assists, rebounds, blocks)

Player level:
- Player's ability in different aspect (e.g. shoot, rebound, speed, Assistance, steal, etc.) in last season
- Comparison between two players
- Player's shot charts
- Player's shot frequency and field goal with different distance
- How is the performance of a player in a game? (by time/area)
- How valuebale is a player?

What's we would like to learn and accomplish:
- Collect data via crawler from NBA statistic website
- Clean and analysis the data
- Design interactive visualization which is easily understand by people and at the same time convey valuable information

Data and Processing

Data is come from NBA statistic website (http://stats.nba.com/), once we get the raw data, we will clean and process it by taking advantage of some python packages like Numpy, Pandas, etc.
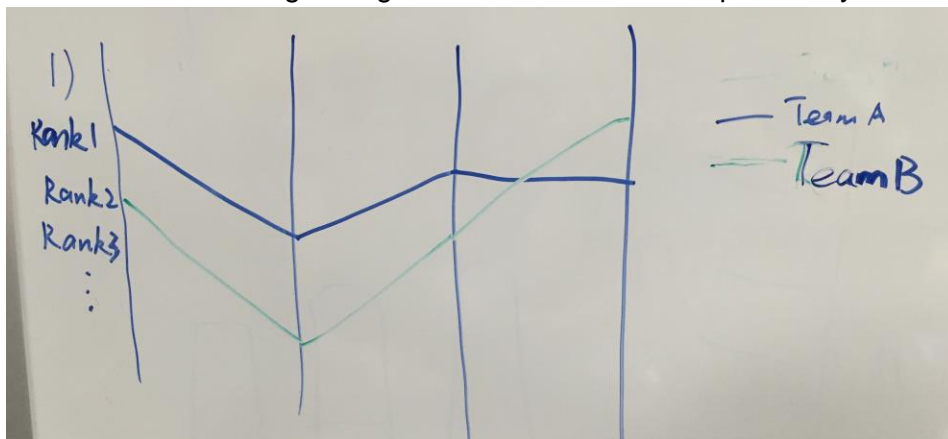API example:
http://stats.nba.com/stats/commonplayerinfo?LeagueID=00&PlayerID=202355&SeasonType=Regular+Season

Visualization Design

## Team Level Visualization:

We want to visualize the ranking changes for each team as time passed by.

The first design is slope plot. Using slope plot can easily tell the ranking changes of each team. However, since there are too many teams in the league, the lines could be overlapping and hard to distinguish.
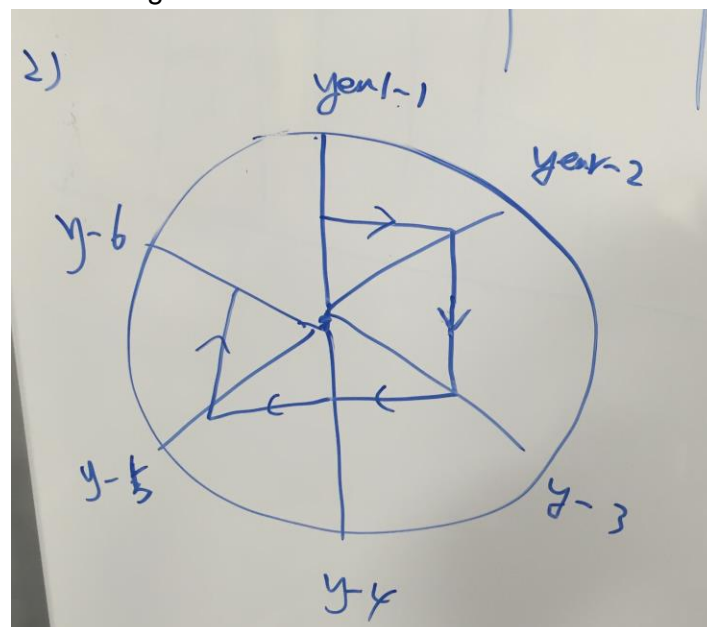
Second visualization is a combination of pie chart and slope chart. As there are almost 20 years historical data, we could save space using this method. However, it faces the same problem with the first one. The overlapping situation will be more serious since the space is limited.
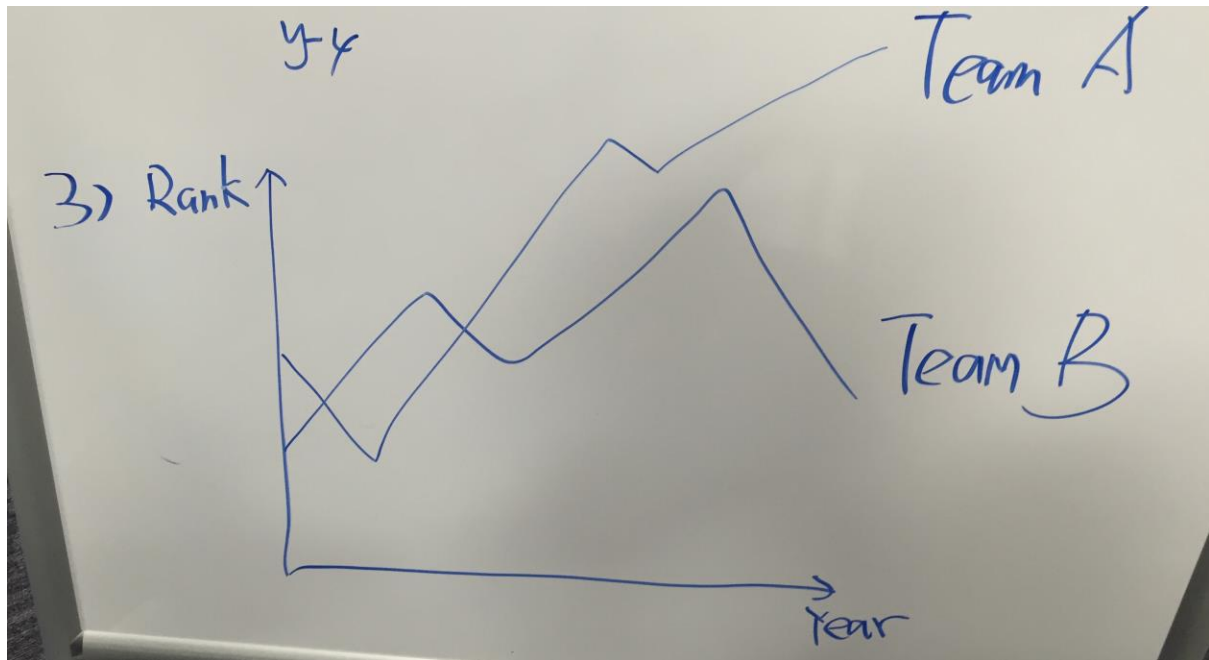
The third design is a simple line chart. It can shows the trends of the rankings. However, as the x-axis year is categorical, line chart would be confused.
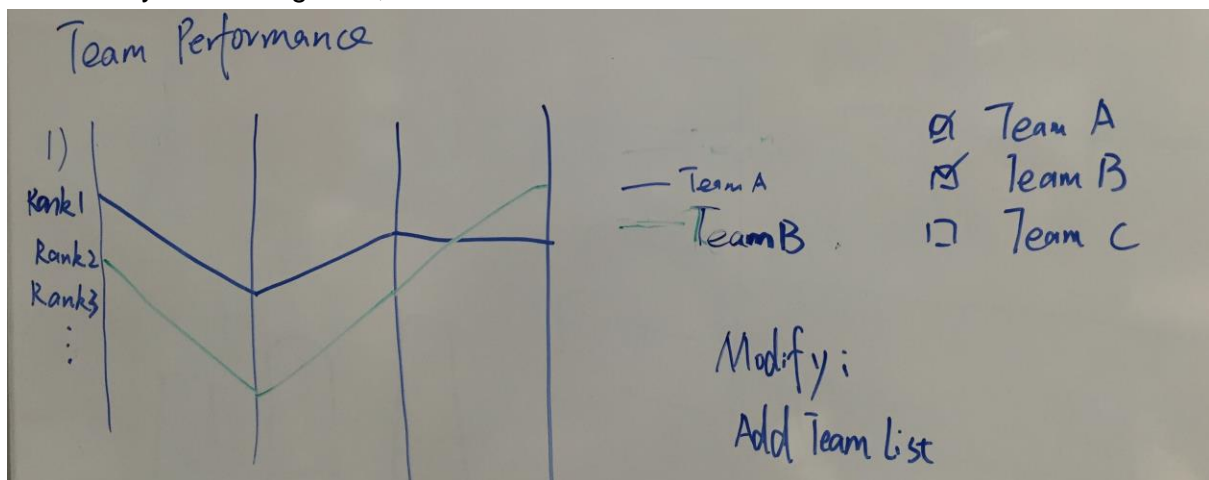
The biggest problem is how to clearly show the ranking trends of each team. So we added a team list. Only the team selected will shows on the graph. In this way, the number of lines will reduced and user can clearly see how the ranking changes of teams they interested.
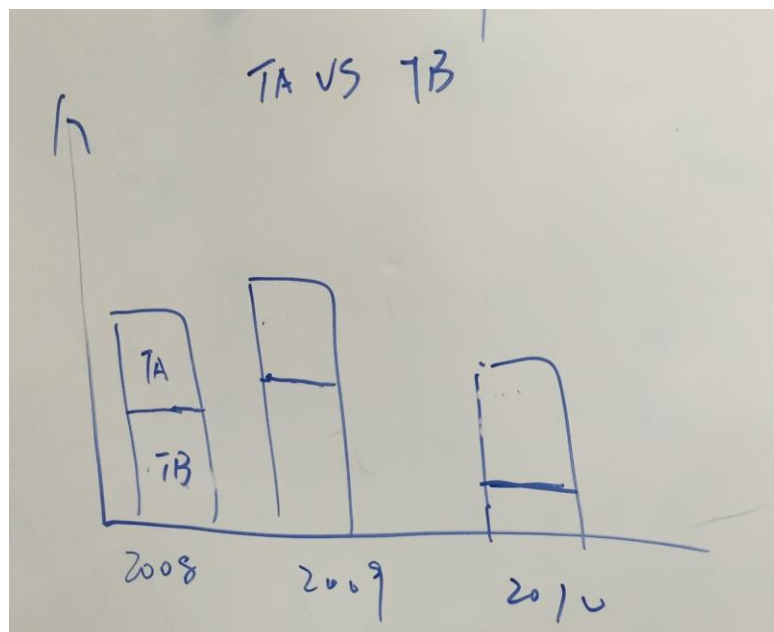
Game history:

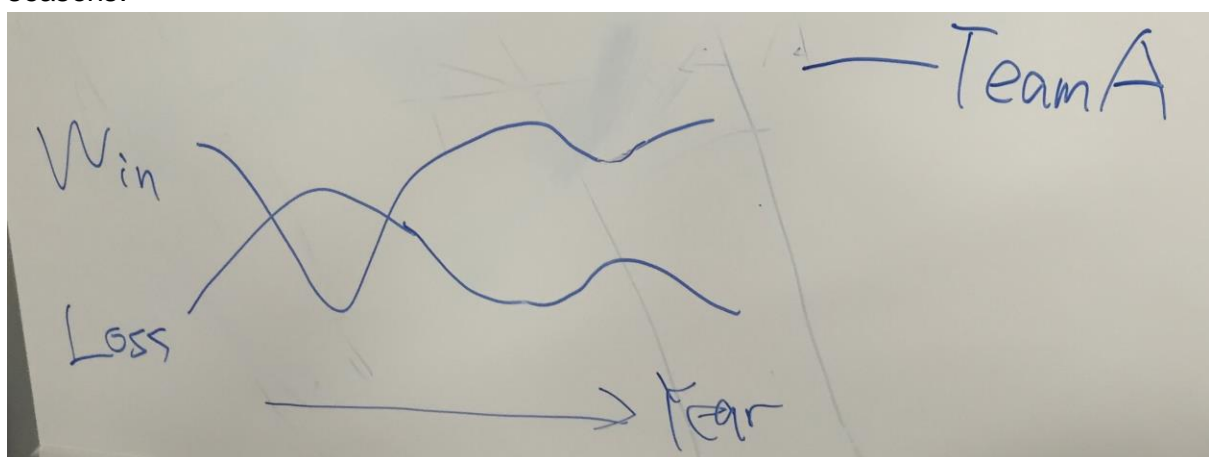In this part, we would like to show the performance of a team (wins and losses).
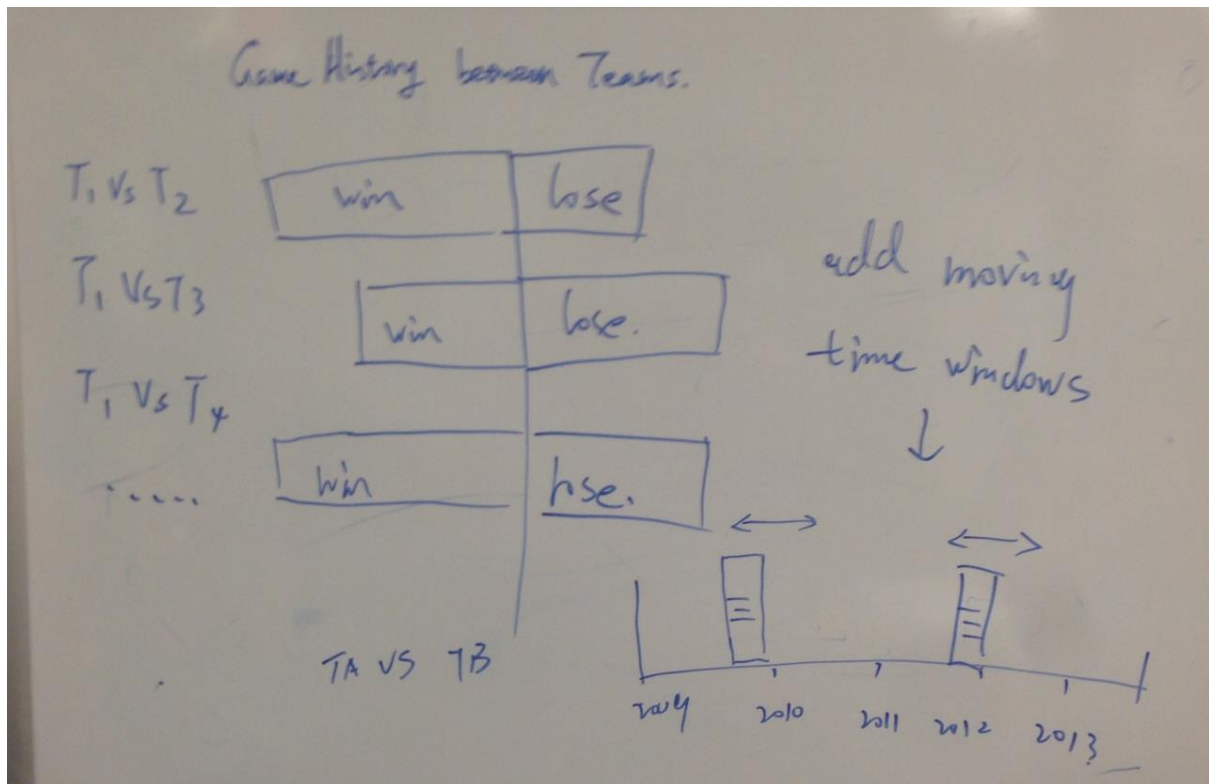
For design 1, one could select one team and get the game history of this team with all other teams in given season.

For design 2, one could select two teams and get all the game history of them in past seasons.

For design 3, the aggregated game stats of one team is given..

To combine the dimensions of all three design (by team, by season, and by total), we add a time slide in design 1: One could move either left and right bar of the slide to select one or multi season, then the game history of selected team with all others will be showed as well as the total stats of the team in the time span.

Player performance:

We hope to visualize the ability (score, block, rebound, steal, assist) of each user.

The first picture is an aster chart, it displays pie slices as lengths extending outward to the edge. Each different pie slice refer to player's one aspect ability (score, assistance, steal, rebound and block). The radius of these slice represented strength of ability. For example, the longer is one radius, the stronger is this aspect ability. The drawback of this design is hard to compare performances between two different players. Because overlapping of slices will lost some information.
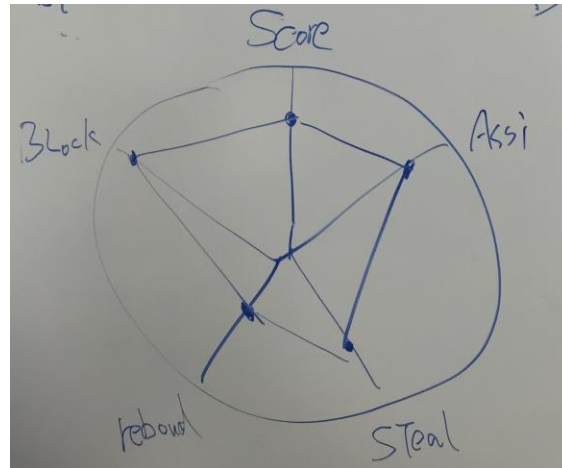
The second design is a radar chart. The edge of this chart is a circle, which stipulate the border and represent the maximum magnitude of the variable (aspect of ability). Inside the circle, there are five spokes that refer to different aspect ability. The data length of a spoke is proportional to the magnitude of the variable for the data point relative to the maximum magnitude of the variable across all data points.
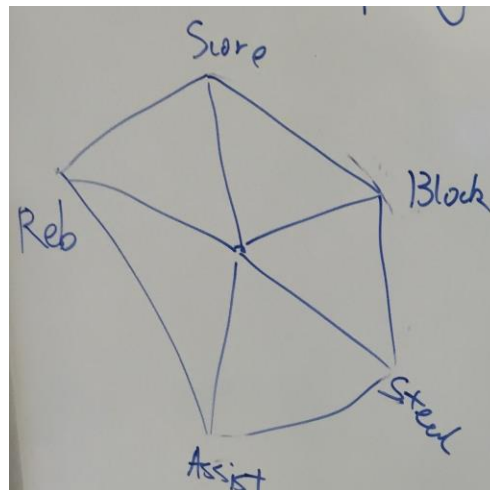

The third design is also a radar chart. The difference compare to design 2 is that there isn't outer circle. The benefit of this design is more simple and easy to understand. In addition, it's very convenient to compare performance between different players.
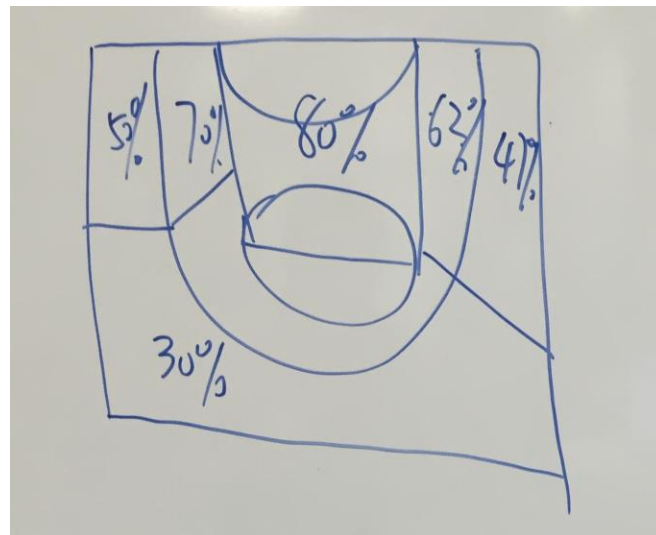
Shooting performance:

We would like to have each player's shooting chart, like the different shooting position preference and ability, like some player have stronger shooting ability in one specific position.
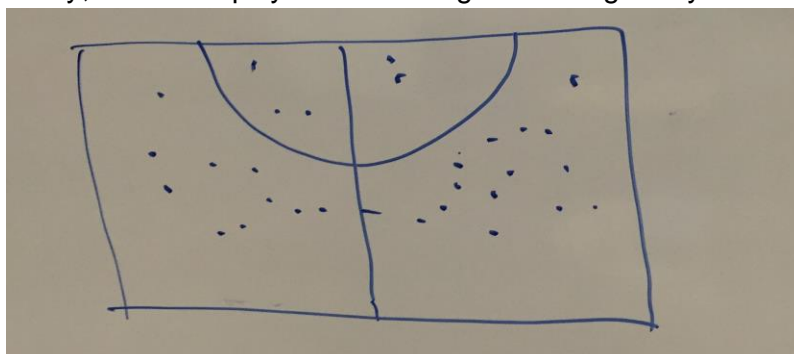
The first design is to plot all shooting positions of one player, so from the picture we can figure out the shooting preference from the density. The denser is one position, the more preference the player would like to shot at this position. The disadvange of this design is we can just get the position preference, but we can't get the performance at one specific position.
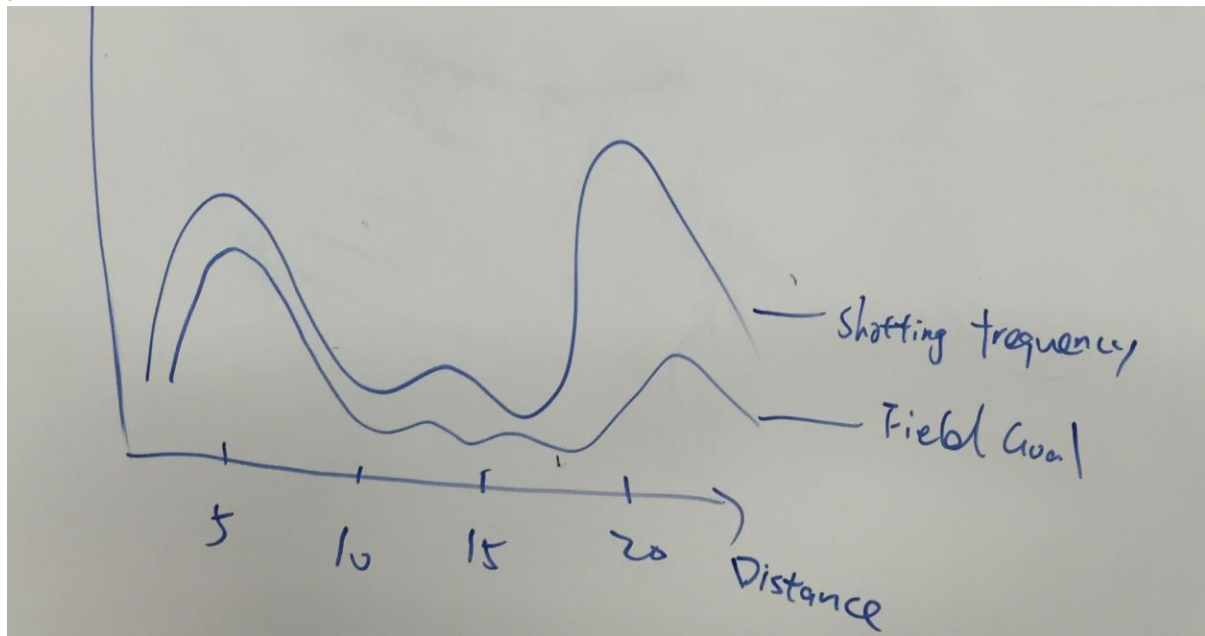
The second design is to use two axes, the x axis refer to the distance between the shooting position and basketry，the y axis is to represent some values, either the shooting frequency and field goal. The disadvange for this design is we just know the distance but we don't know exact position. For the positions that have the same distance, player's performance might not be same.

The third design is to give shooting performance to different position field. In this design, we don't consider each separate position point, instead we combine near positions as one position field. In this field, we calculate the player's performance and show it on the picture. We will choose this design as our final design, because it solve design one and design two's drawbacks and easy to understand.

Must-Have Features

Time window: Time span selection should be included in the form of slides or calendar, so users could select time periods they are interested.
Check box: To select a team or player from the all others in NBA, we need a chck box or a list in other form.

Optional Features

Interaction map: The home page is a US map and the pins the location of each team. When you click the pin of a team, you can see the team information and players' information of that team.
Animation transform: Use animation transformation during the change of graphs.

Project Schedule

11.1 - 11.7	Prepare and clean data
11.8 - 11.12	Create team level visualization.
11.13 - 11.15	Create game history visualization.
11.16 - 11.19	Create player performance visualization.
11.19 - 11.23	Create shooting performance visualization.
11.24 - 12.13	Add more animations and interactions. Designed the organization of website.