CSE 557A Final Project Proposal
"NBA Players Statistics and Performance Analysis in Decades"

Po-Hsu Chen, 448031
Teng Xu, 452663

- **Basic Info.**
  "NBA Players Statistics and Performance Analysis in Decades"
  - Po-Hsu Chen, 448031   pohsu.chen@wustl.edu
  - Teng Xu, 452663   teng.xu@wustl.edu

  Link: https://github.com/david148877/david148877.github.io

- **Background and Motivation.** Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

There is always a discussion about who is the best player in NBA and everyone will give their different answer. For me the answer is Kobe Bryant. However, I have to say that there must be someone do something better than Kobe in the court and I also believe that there must be someone did something better than Michael Jordan. As a result, I am curious who is the one that make something even greater than those super star such like MJ, Shaq and Kobe:)
As super fan of NBA and basketball, we decide to answer those question by our information visualization. As we both know, in many NBA website or related game, there is a database for players that you can check player's' performance. We think it is so cool if we can create our own database for those great player by our own style. I know that those existed database can show player's scaled stats and this is not what we want to do. What we will do is to show something new and also meaningful things to our audience. Having such passion to make our own database, we decide to choose this topic and show our creativity to you.

- **Project Objectives.** Provide the primary questions you are trying to answer with your visualization. What would you like to learn and accomplish? List the benefits.

Is this NBA player "strictly" better than that one?
Which NBA player did better on a spcefic field?
Which NBA player has more significant improvement on the stat you choose?
Which NBA player has more relationship with other players?
What is the key word for a NBA player? how about 10 words?
Making this project will give us the chance to get the answer of those question.
During the process of the project, we will directly see the comparison of two players and

we may find something surprise us.

During the process we will see the step of a improvement of a player, which may also inspire ourselves.

During the process we will dig more about a player, which will make those great player get known by us step by step.

Meanwhile, other NBA fans can visit our website to look the information that they may never seen before and also know the player they what know.

If there is a situation there two fans are arguing about who did best on stealing, we can just stop them and oper our website for them, the answer will come out.

An important thing of our project is, it is really fan for NBA fans because we will make it unique and creative.

- **Data.** From where and how are you collecting your data? If appropriate, provide a link to your data sources.

We are now exploring more options and possibilities. We plan to collect data from at least these websites:

https://www.teamrankings.com/nba/stat/win-pct-all-games
http://www.databasebasketball.com/leagues/leaguelist.htm
http://stats.nba.com/standings
http://stats.nba.com

We pretend to download and parse them through a crawler. Due to the structure of websites and the intention of banging robots, sometimes it doesn't work. In that case, we'll manually download those missing parts.

Meanwhile, we will try to dig the key work for each player and also the play's relationship with other players. Those may not be easily downloaded online. So we decide to find those by ourselves. Searching and learning will also be a joyful process for us.

- **Data Processing.** Do you expect to do substantial data cleanup? What quantities do you plan to derive from your data? How will data processing be implemented?

Yes, we expect a lot of data cleanup work to do. There are many json files provided by NBA official website. However, they are unlabeled and chaotic. There will also be some data we will find manually so that we will make it clean. We expect to get the stats of each player in the decade. We want to know what's the relationship between players and players. We also want to know what is the key word for each player and those will all be showed in our data. If it is necessary, we will implement the data processing in R to make our data more readable and clean.

- **Visualization Design.** How will you display your data? Provide some general ideas that you have for the visualization design. Create **three alternative designs for your visualization**. Create **one final design that incorporates the best of your three designs**. Describe your designs and justify your choices of visual encodings. You use the Five Design Sheet Methodology.
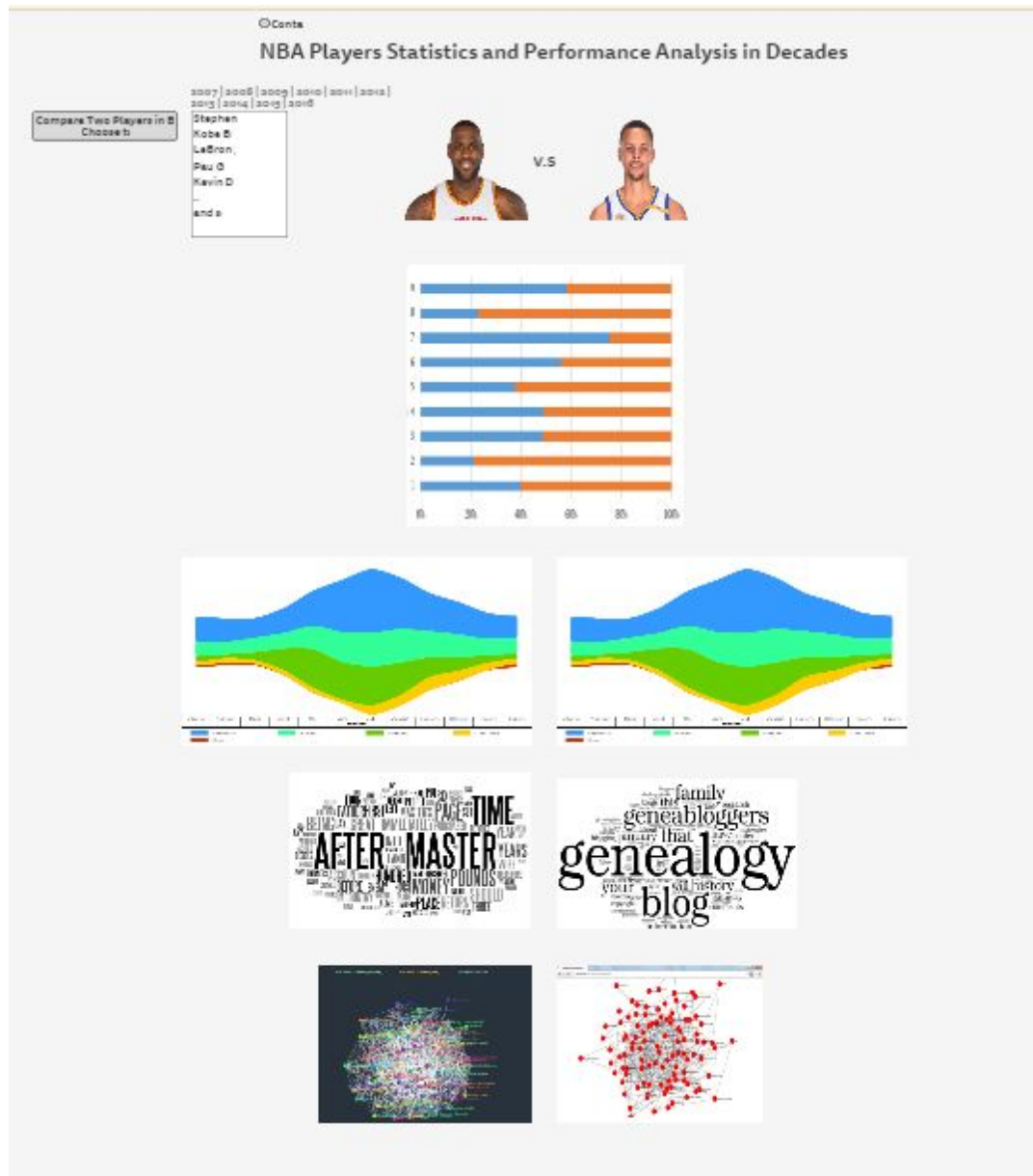
ideas:

Bar Graph in Contrast

ThemeRiver

Text Cloud

Relationship

Design 1:
Put all visualizations in the same canvas. Users scrolling down to see more visualizations.
Click the button "Compate Two players in Bar Graph"

Pros: User can directly see all the visualiztion in one time
Cons:To many information at same time may cause no focus

Design 2:
After choosing two players, when users move their mouse on the button, it will show the corresponding image of the button (through hovering method).
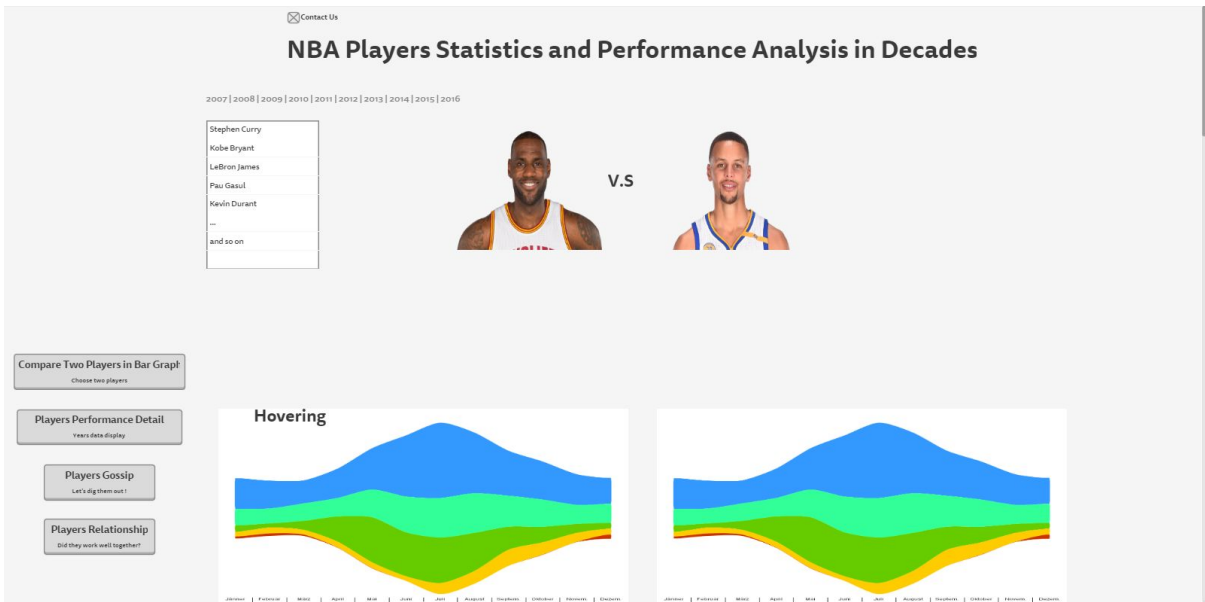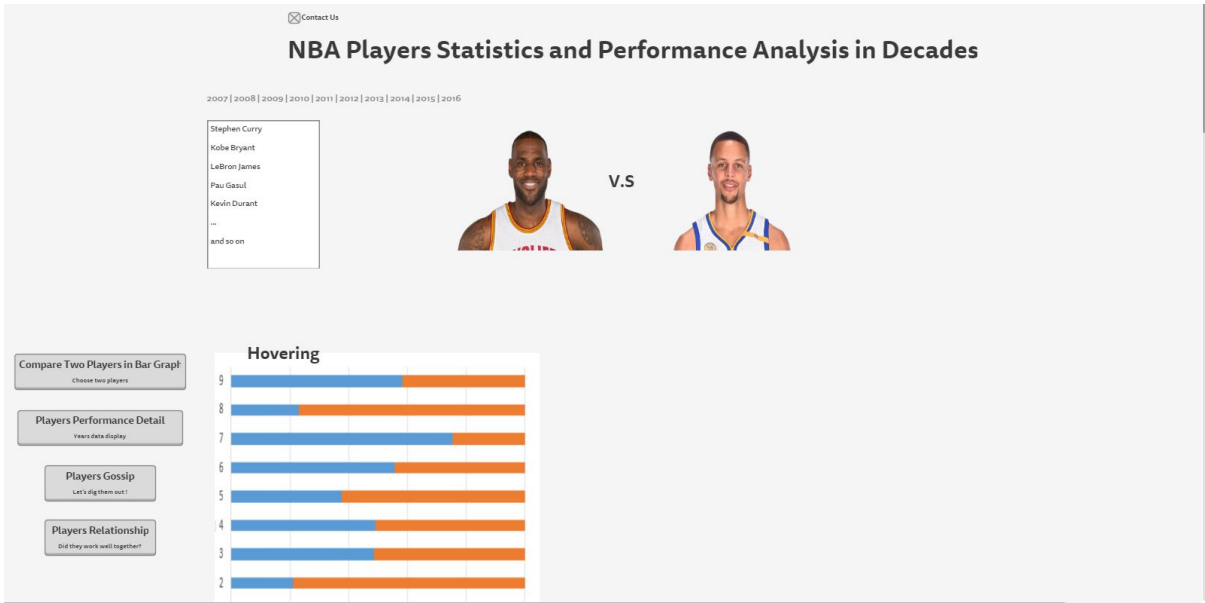
Design 3:
Show all visualization in the same canvas and users don't have to scroll down. It implies
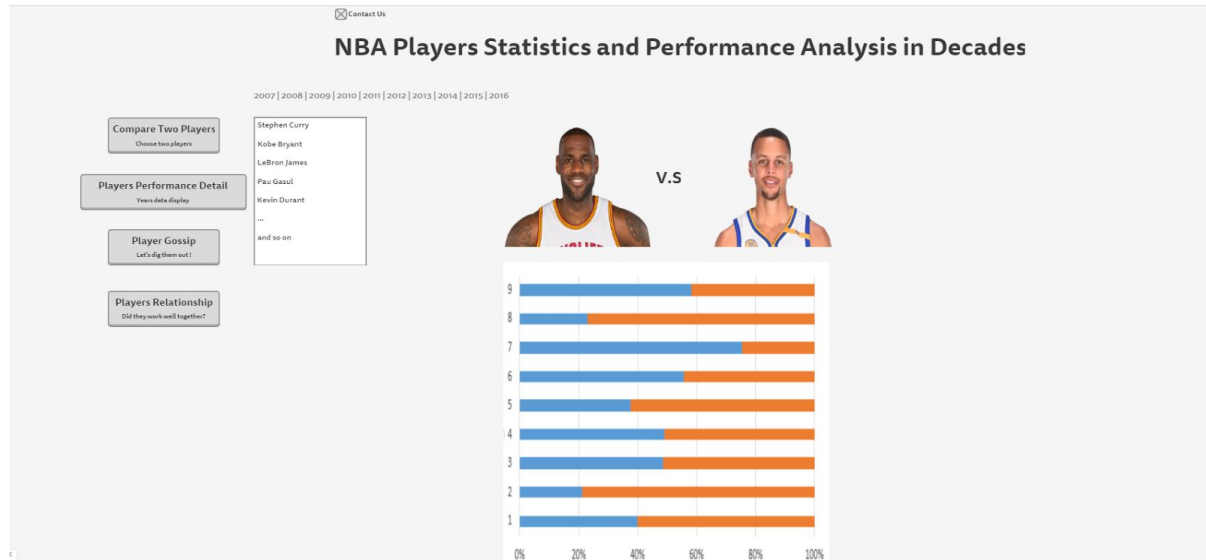**that the left side different to the right side by players.**
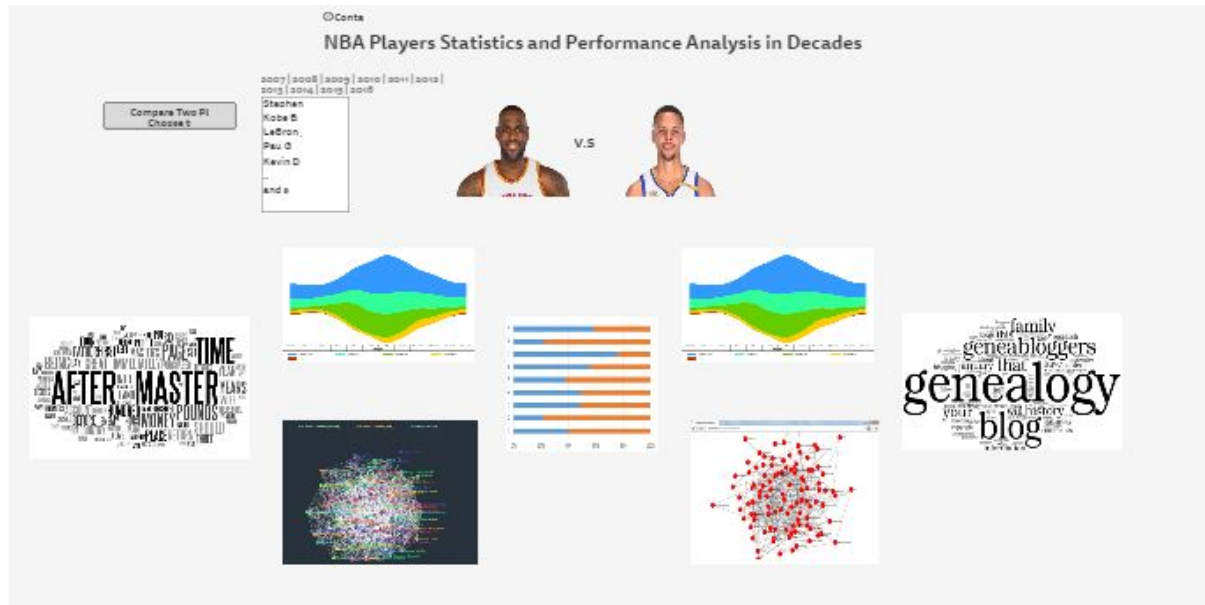
Pros:Users can access all visualization in the same time.
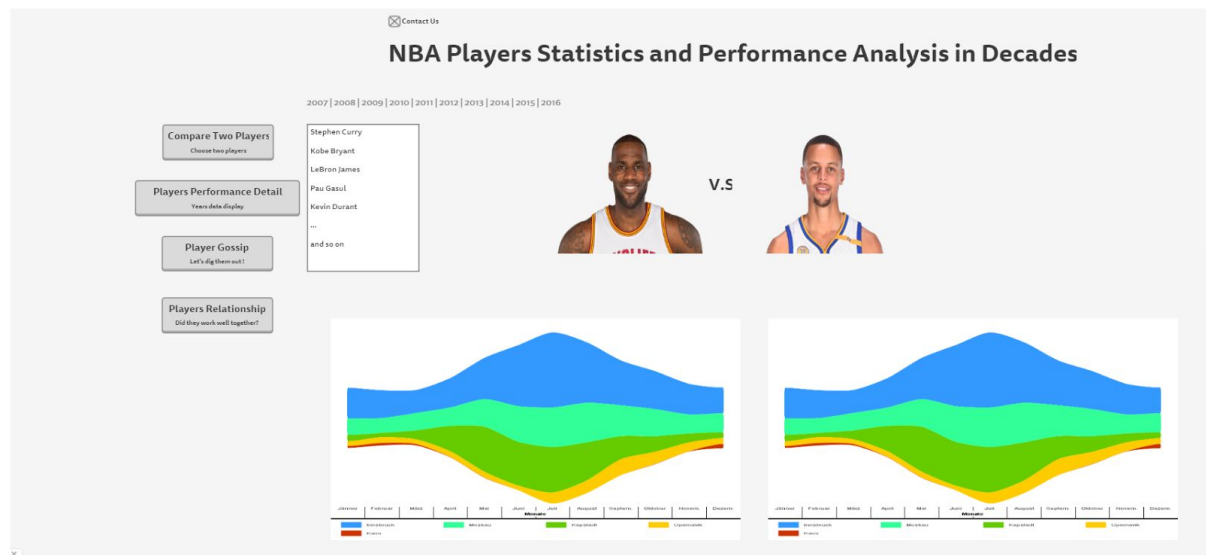**Cons: It's a little confused to users.**

Ultimate Design:
Display specific visualization when users click a certain button.

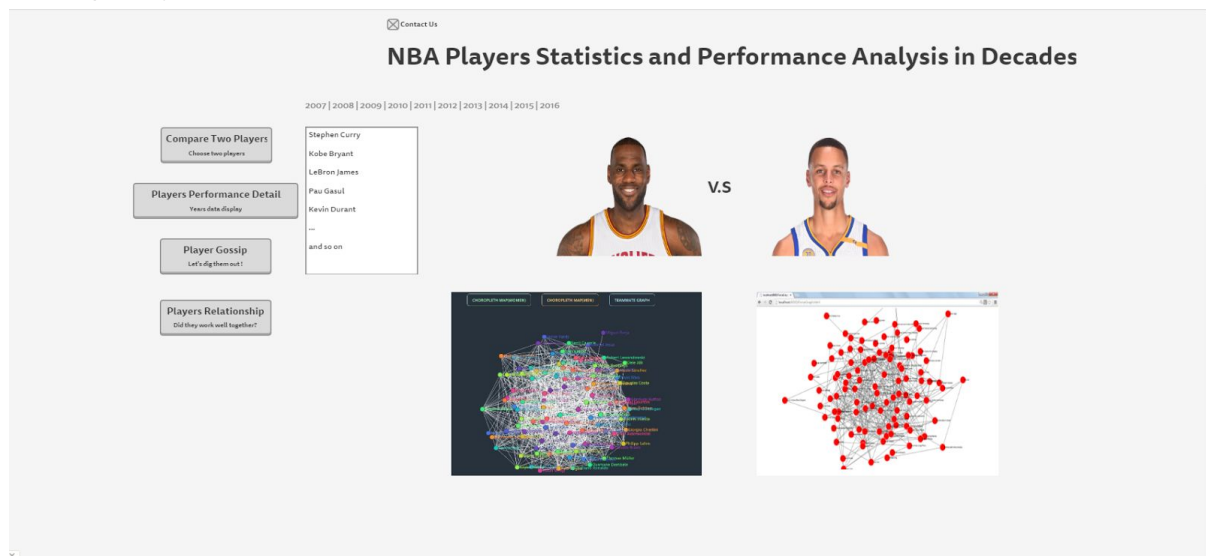When click "Copare Two Players in Bar Graph" button:

Clicking "Players Performance Detail" button:

Clicking "Players Gossips" button:

Clicking "Players Relationships"

Pros:User can choose what they are interested in and won't be bothered with other element.
Cons:User cannot have a direct view about how our visualization like and has no evidence for them to choose.

- **Must-Have Features.** List the features without which you would consider your project to be a failure.
    - A interface that user can choose the two player they want to compare. This interface will list all the player in specific year and user can also choose the year they want to look at.
    - A contrast chart that directly show the comparison between two players. User can directly see how one player do in one field compared with how other player do by the length of bar in the chart. Meanwhile, user can also see how this player compare with other in a overall view.
    - A stream river that show how this player improve in specific stats. We will bring a updated version of stream river. This time we will allow user to choose which stats they want to know and joint stats will also be available. The stream river will change according to how user choose.
    - A relationship map that shows the relationship between one player and others. The relationship map will based on a force-based map. Player will be represented by ball and relationship will be represented by line. User can drag the ball and all the line will follow the dragging to move. User can show which player has a wide relationship in this legend.
    - A text cloud that shows the key word for each player. User can see what is the hottest word related to this player and also some other keyword for him. User can know how to describe this player in the most accurate and cutting-edge way. User can also check is there any common word between two player.

- **Optional Features.** List the features which you consider to be nice to have, but not critical.
    - A share functionality will be cool. We are thinking to have a button that if users click it, they all be able to share our web on Facebook directly. But first of all, maybe we should try to create a button which can copy the link of our web.
    - Letting users able to make comments on each player maybe is a great idea. People can present their opinions and discuss on our web.
    - Maybe we can integrate upcoming games and build a functionality to bet on unknown results. Friends can log in and make a bet with others, and the rewards can be chores.

- **Project Schedule.** Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.