

Data Visualisation Report: FIFA Soccer World Top 50 Players Data Visualisation

Jianan Xie



Fig. 1: The general data visualisation from Sofifa website for each FIFA soccer player, shown on the left two screenshots, and the final visualisation that has been implemented in the narrative visualisation.

Abstract— Showing multi-variables data for a soccer player normally are in text and number form. In this project, I compared many different visualisation strategies and decided to use interactive bubble char together with a radar and sunburst chart for my final narrative visualisation.

Index Terms—Interactive Bubble Chart, Multiple Views Visualisation, Multi-Variables Visualisation

1 INTRODUCTION

As soccer is the most popular sports in the world and FIFA 2018 World Cup is just around the corner. I want my narrative visualisation based on the dataset I got, to convey a comprehensive message about all general aspects of an individual soccer player, which may help people to increase their awareness and understanding of each soccer players, furthermore, I hope this narrative visualisation could help people to form their own model in order to let them have a better overview of their interested soccer player in the future.

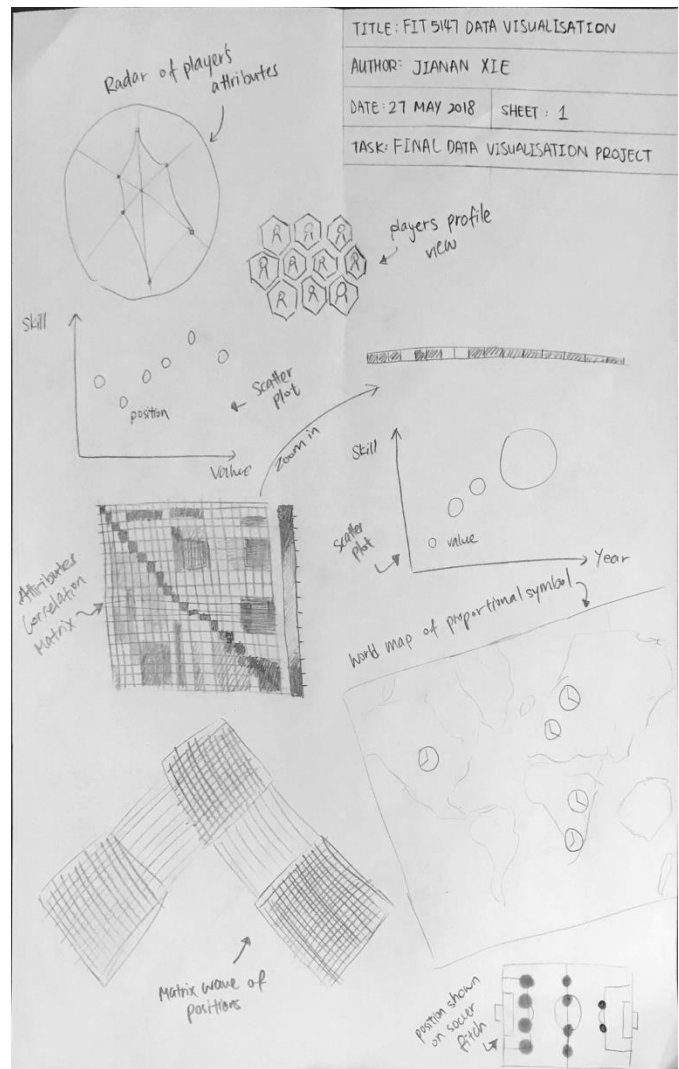
According to the feedback that I have collected from my data exploration, I realized it is a bit difficult to interpret soccer related techniques into a form that could be easier for people who do not have any soccer background to understand it, for example, the correlation matrix of soccer player's attributes, most of the people react that they do not understand what is the matrix try to tell after the first time I showed them without any explanation. I originally wanted my narrative visualisation works the benefits of any soccer-related professionals. However, I would like to see more people becoming passionate about the soccer sport after they viewed my project, therefore, I would like to design and implement my narrative visualisation for everyone.

2 DESIGN

First Design Sheet

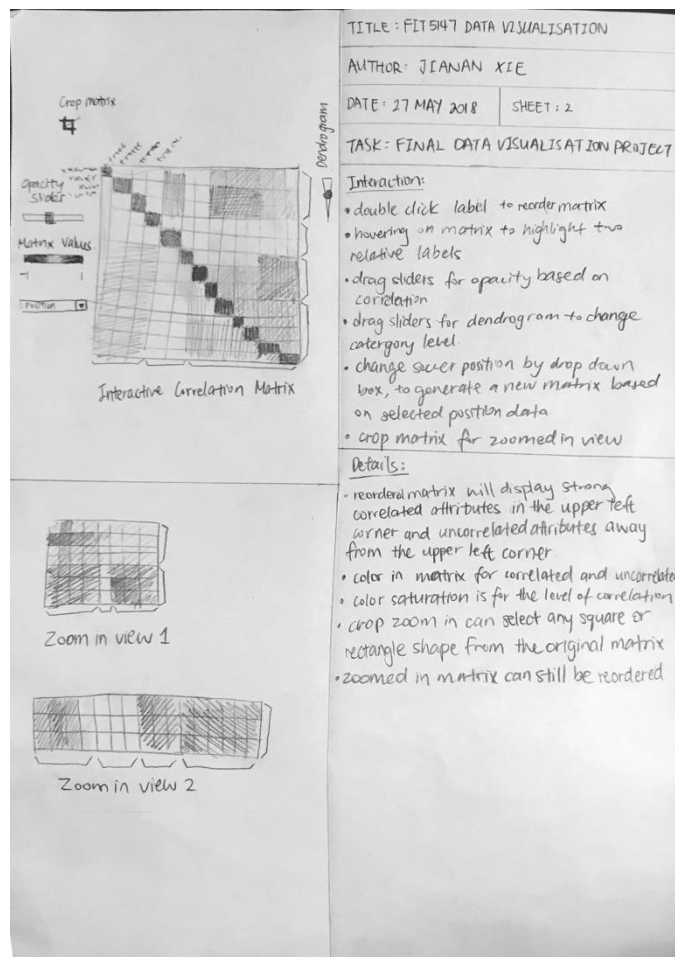
I have considered many possible visualisations to display the soccer players' data and attributes.

- Radar chart to display the six core abilities of a player.
- Joined hexagons view for display players profile. The similar players will be placed next to each other based on skills correlation.
- Scatter plots to display selected variables (eg. A player's value according to year and skills).
- Heat map correlation matrix for all attributes of players according to nationality and squad position.
- Matrix wave for correlation between different squad positions



- Proportional symbol for soccer players' geographical distribution with selected player's attributes (eg. which country has more fast players; which country produces more strikers).
- A soccer pitch view to bringing more visual feature for soccer squad's formation.

Second Design Sheet



An interactive heat map correlation matrix shows the correlations of all attributes from selected players' data. Filters can be applied to select players' data based on club, nationality or field position. The matrix can be reordered by click on an attribute.

This visualisation will be ideal for professional roles from soccer to investigate on soccer player study. This visualisation can significantly help a club or a team to set their player training plan and new player selection. However, as I have mentioned in the introduction, which is the correlation matrix is hard to understand without any explanation, and I would like to make my narrative visualisation more friendly to all people, especially for those who do not have any background knowledge.

The demo made by web application Clustergrammer with prepared TSV data from my dataset:
http://amp.pharm.mssm.edu/clustergrammer/viz_sim_mats/5afce40db119f04f5ef8bbc1/output2.tsv

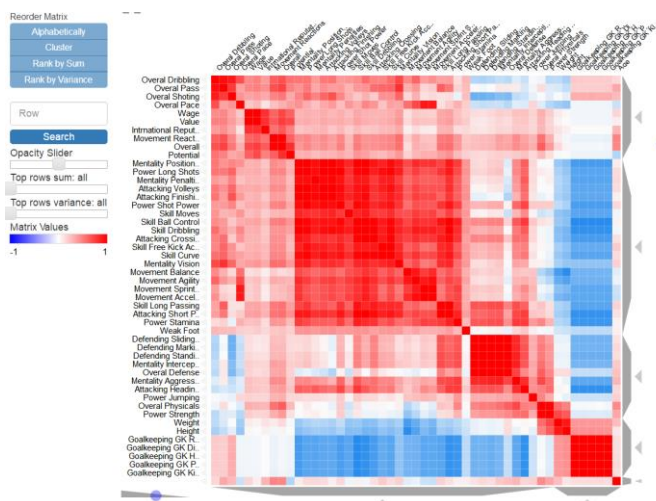
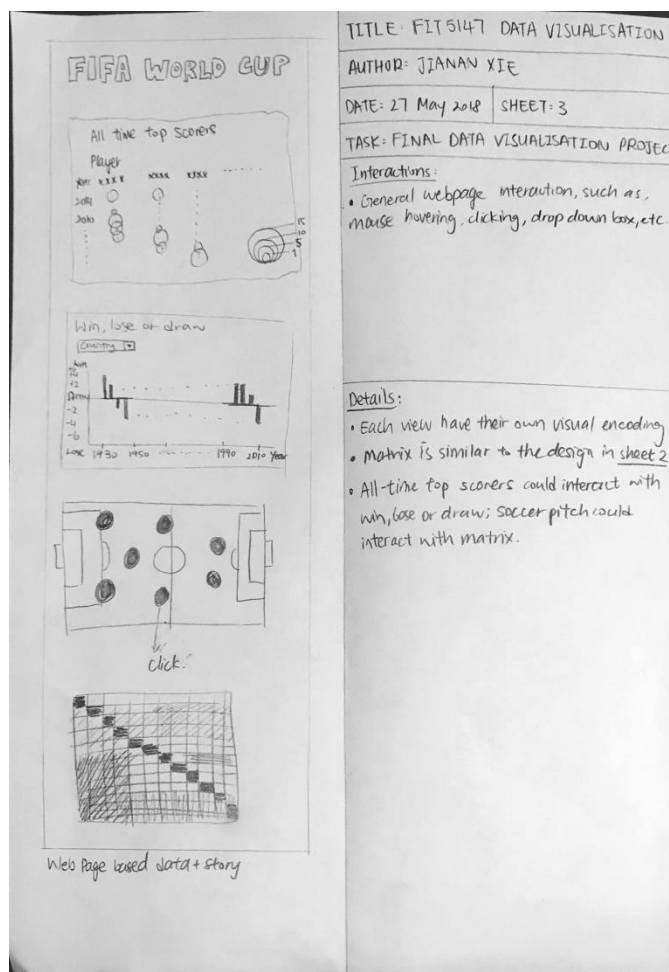


Fig. 2: Interactive correlation matrix by Clustergrammer

Third Design Sheet



The visualisation that I was thinking at this stage is a kind of data based storytelling web page, like an information web page with a storyline. I would make to make a web page about FIFA World Cup, from history games to the upcoming FIFA 2018 World Cup game. The page starts with a diagram that shows all-time top scorers of the games, for example, visualisation shown below.

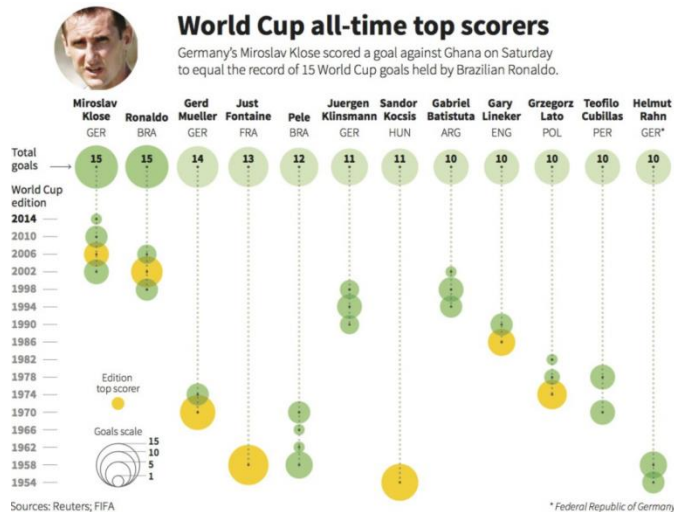


Fig. 3: Thomson Reuters: World Cup Top Scorers [1]

Followed by a visualisation of an all-time FIFA World Cup goal score bar chart by selected country, each positive value bar means the team has won the game with the number of goals scored; no bar means the game went out with a draw and negative bar means the team lost the game with the number of goals. The filter for this visualisation could be a drop-down box with all the countries that involved in World Cup games. Example visualisation showed below.



Fig. 4: Thomson Reuters: Win, Lose or draw [2]

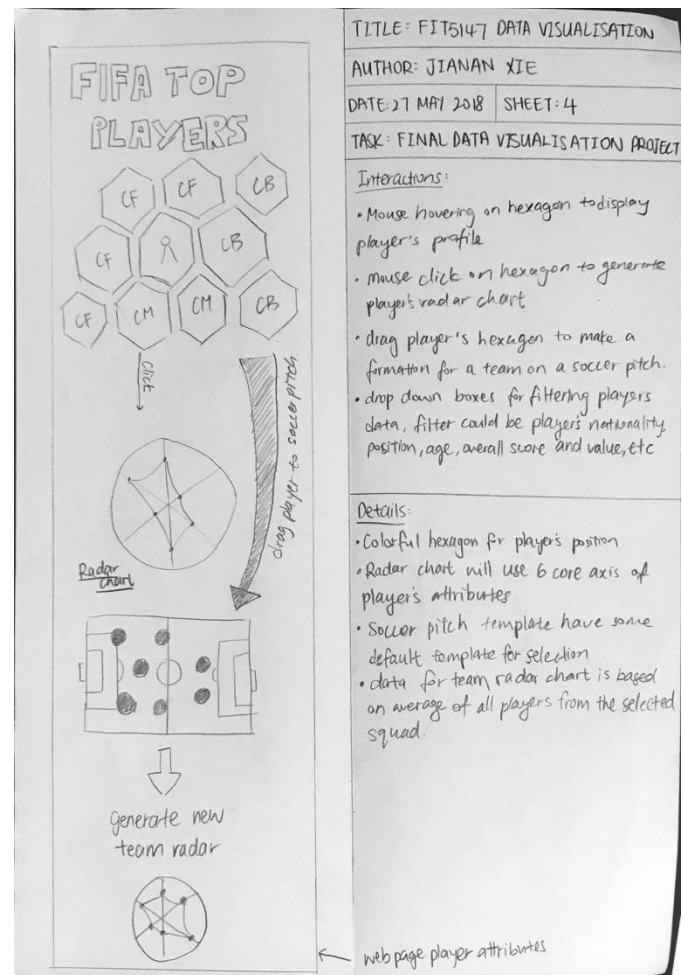
After the above two visualisations of FIFA World Cup history data, I would like to present a sort of soccer squad formation tactics view from a soccer pitch template formation, for the user to select each of field position in order to generate the related correlation matrix, data based on all the current FIFA players. An example is similar to design sheet 2.

The challenge for this design sheet is I need extra data to generate the first two visualisations, and it will also be difficult to make these three visualisations to interact with each other.

Fourth Design Sheet

A hexagon player's profile view would prefer a more visually striking feature for people to look on it. It could be colourful with lots of mouse interactions, such as click, hovering and dragging. Players with similar skills, traits and abilities will be placed close to each other. Mouse hovering on each hexagon to display detailed player's profile information, and mouse clicks on each hexagon to bring out the radar chart of six core abilities of the selected player.

Also, the pitch squad formation template has been designed more interactive with player's profile, which is allowed the user to drag players to make their own team, and a radar chart of the team will be generated after the squad been set.



The fourth sheet starts to change the design for not just soccer professionals, but to all people. Therefore, I have decided to remove the correlation matrix.

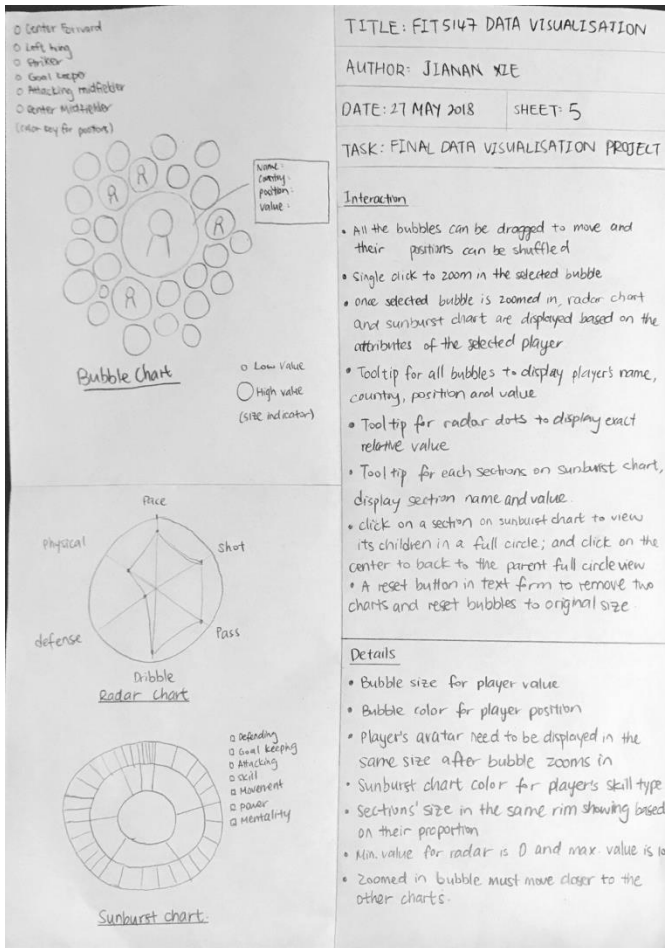
In order to increase the interactive feature of my design sheet 3, this design sheet visualisation is more capable of a web application. The weakness of this design sheet is the most of my data feature and attributes about a player is not been used, such as the details of all kind of skills. Therefore, I need to come out something which could help me to display multi-variables data.

Fifth Design Sheet

Interactive bubble chart has been used for the player's profile display, as the bubble chart is more flexible to re-position and to change its size, and the size of the bubble could be determined by the player's value, which is an extra visual encoding feature that the hexagons view. The dragging function would allow the user to re-position the view into what they want to look, such as same field position players can be put together; similar value players can be another group.

The radar chart is similar to the design sheet 4.

For the multi-variable data, I finally decided to use sunburst chart. Normally, the sunburst chart is for display hierarchy. The goodness of my data is all the detailed player attributes can be categorised into 7 groups; therefore, the sunburst chart could have the overall player rank score in the middle; the 7 groups in the first rim as children of the centre and as parents for the second rim; the detailed attributes in the second rim as children. The interaction would allow the user to click on a parent to change the parent to become the circle centre and its children to become the first rim with sections based on the data value proportions.



display the top 50 players, which is also a good set of player's distribution.

- d) For some of my features like the radar chart, it is required two layers of data in JSON. Therefore, when I load the data into my main visualisation, which is the node of the bubble chart. Instead of loading the data directly, I have changed the loading pattern to catch the data from JSON file. An example is below:

```
Weak_Foot: data.Weak_Foot,
Skill_Moves: data.Skill_Moves,
//data pattern for radar J.XIE
radarData: [
  {
    className: data.Name,
    axes: [
      {axis: "Pace", value: data.PAC},
      {axis: "Shot", value: data.SHO},
      {axis: "Pass", value: data.PAS},
      {axis: "Dribble", value: data.DRI},
      {axis: "Defense", value: data.DEF},
      {axis: "Physical", value: data.PHY}
    ]
  }
],
```

Fig. 5: Radar chart data pattern when it is loading to bubble node

In this case, the data for the radar chart will be loaded as in a pattern with two layers (the first layer is an array of className and axes, and the second layer is an array of axis and value). The pattern will help the radar JavaScript to perform better in the later coding.

- e) For the sunburst chart, I have read a couple of examples and tried each of them. I personally think they all contain two weaknesses.

- The label problem. Only static version of sunburst chart can properly display label. All the zoomable sunburst charts have some issue on display their label after spinning. The label positions are very hard to adjust as the lengths of the label are different and the area of the sector are different as well.
- The centre problem. Most of the zoomable version of sunburst chart is just using their centre for go back to its parent layer, but not use the centre to display the value and information for that category. Also, for static sunburst, the centre is used to display information, but it causes confusion when the mouse hovering on the different children. Think about if there is a large sunburst chart with many layers, the centre displays the name and value of a sector when mouse hovered on a very small sector, it would be hard for the user to notice where is displayed data from unless there is a deliberate design of changing opacity.

In order to solve this weakness, I have decided not to put a label on the chart to avoid causing nasty view. I added a legend on the side to use colour to indicate the sector name. Also, to display sector data details, I decided to use mouse hover tooltip.

- f) For the bubble player profile, some low-value players' avatar is small as the bubble size is based on the player's value. The avatar needs to change to a reasonable size after the bubble has been zoomed in. The example is shown below.

3 IMPLEMENTATION

D3.js has been used for the implementation of the project.

Data preparation:

I used the website CSVJSON (<https://www.csvjson.com/csv2json>) to convert my dataset into JSON file. All data manipulate activity commenced in the JavaScript code.

D3 template selection:

- Bubble Chart [3]
- Force-Directed Graph [4]
- Zoomable Sunburst on d3.js v4 [5]
- D3.js - Radar Chart or Spider Chart [6]

Implement decisions:

- As I am very new to D3, I have decided to use others D3 template to add features and modify the display.
- I originally wanted to put the radar chart inside of the zoomed in a bubble, but after lots of attempts, I still cannot get the radar showing inside of a bubble, so I decided to put the radar aside of the bubble chart.
- In the beginning, I want to display all players' profile in bubbles for those who are going to play in the FIFA World Cup 2018. There are approximate 600 more players. I have tested my project, which is capable to run more than 800 bubbles, but it would make the visualisation very messy and hard to find any key point. So eventually, I decide just to



Fig. 6: Small bubble causing a small avatar in the middle as the avatar loaded based on the size of the circle

- g) For bubble zoom out, I originally use click on any non-zoomed in bubble area to zoom it out, but this will cause the zoomed in a bubble to zoom out when clicking on the sunburst chart. So I changed to click on the zoomed in a bubble to zoom it out.
- h) Instead of move the zoomed in bubble to the middle of that SVG, I chose to move the bubble to the right side to make it closer to the other views.
- i) The max value or radar has been set to 100 and min value is 0, and there are 5 layers of the circle to display the data value, between layers are 20 indifferent.
- j) Players' avatars have been downloaded in IMG folder to increase the visualisation performance and offline using. I could use image link for the avatars but it will take a while for visualisation to initialisation.
- k) All JavaScript including D3 has been downloaded into the JS folder for offline using.

4 USER GUIDE

Interaction with a bubble chart

Interaction 1 - Click on a bubble to zoom in a bubble and generate radar and sunburst chart with selected player's data.

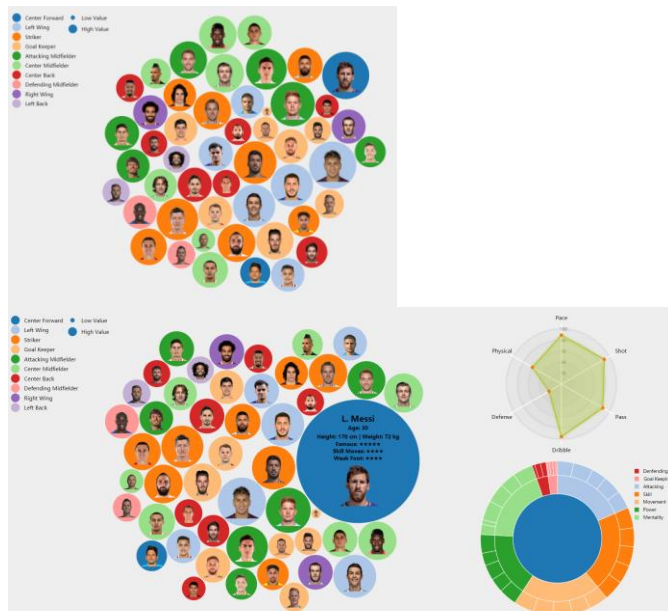


Fig. 7: Click on a bubble after initialising the visualisation

Interaction 2 – Click on the zoom in a bubble to zoom it out, and radar and sunburst chart will disappear.

Interaction 3 – Dragging bubble to re-position the view.

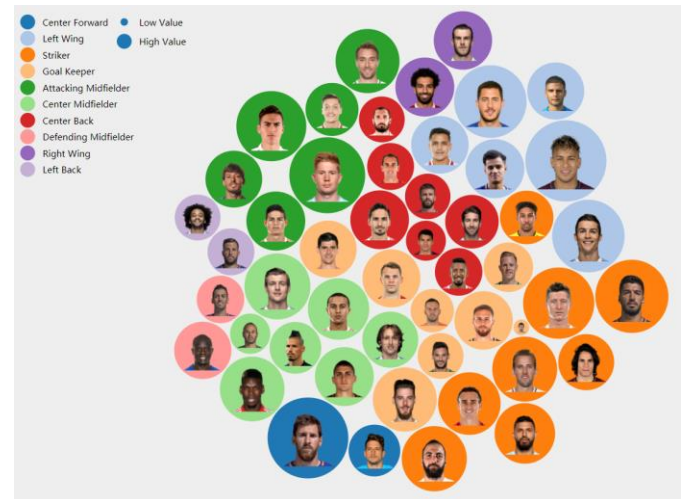


Fig. 8: Players at same field position has been repositioned next to each other to compare their value easier.

Interaction 4 – Mouse hovering on a bubble to display player's name, nationality, field position and value.



Fig. 9: Tooltip to display a player's name, nationality, field position and value.

Interaction with radar chart - Mouse hovering on an axis dot to display the value.



Fig. 10: Tooltip to display axis value for Shot

Interaction with sunburst chart

Interaction 1 – Click on a sector to zoom in.

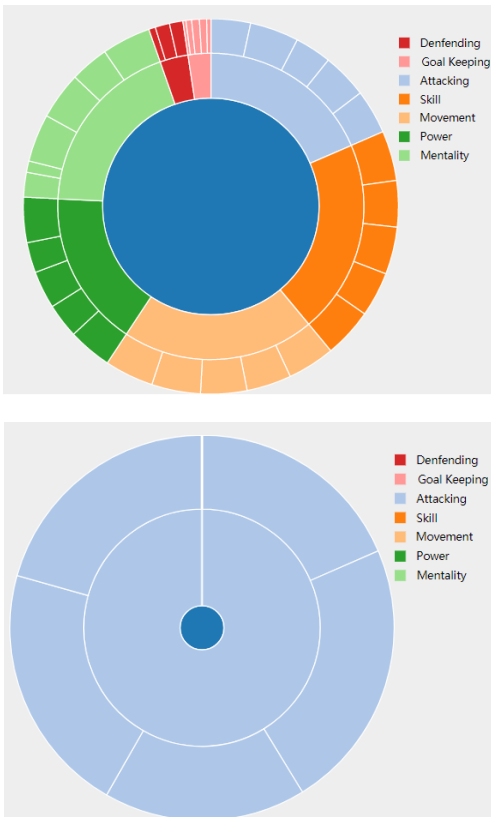


Fig. 11: Click on Attacking sector to zoom in

Interaction 2 – Click on the centre to zoom out.

Interaction 3 – Mouse hovering on a sector to display the label and value.

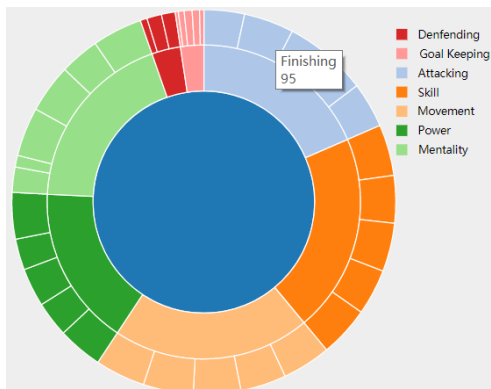


Fig. 12: Tooltip to display sector label and value

5 CONCLUSION

Achieved

The narrative visualisation has fulfilled my design and resolved the challenge to display the multi-variable data. The visualisation can provide comprehensive information about all attributes of a player. The dragging function of the bubble chart gives the user more flexible to re-group the bubbles. The sunburst also can show the proportion of their skills distribution.

Reflection

I have learnt a lot of things about data visualisation throughout this project and this unit.

- I cannot say I am confident to take any D3 visualisation project, but I should be competent to create any basic data visualisation by D3.
- For D3 programming, manipulating with data is the most difficult part, such as display label on sunburst chart is a very challenge task, the ideal approach for this could be just displayed a label on centre and first and second rims where the sectors are fairly big enough for the label.
- The other part about D3 programming is to learn how to join code and sub functions together because most of main functions or interactions are made up by a small piece of code.
- When getting stuck on coding, try everything and every approach of programming solutions first. If it is still cannot be resolved, like my display radar inside zoomed in a bubble, then change the design for this part.
- Design data visualisation from a different type of data. As I have said I am a kind of regret to choose my soccer player dataset because the multi-variable is very hard to interpret into an effective and fancy visualisation. It is always challenging as there are so many things need to be included in a single diagram, and the diagram must making sense for the user and reader.
- The design is very important. It is very easy to get lost in the beginning stage of the project, which I viewed heaps of different visualisations, but very hard to come out my own idea. I think the most important mindset at this stage is to clear your mind and think what you really want to tell your audience. The visualisation may not need to be so complex and so gigantic, but it needs to be precisely interpreting your dataset to your audience with the topic that you would like to carry out. For my case, I originally wanted my visualisation to have some nice looking players profile features for general audience combined with some visualisation for professional soccer people, but it was really hard for me to think the interaction between each visualisation, like my design sheet 3.

For the data visualisation, I am interested and passionate about many of things and features, but they are getting really hard for me to implement them. I have learnt the D3 course from Udacity online course during week 9 and week 10, and I still found the knowledge that I have learnt is not enough for me to get my assignment 3 D3 programming done quickly. I spoke to one of my friend who is an experienced web developer about my concern, and he told me the problem might be my foundation is not solid enough. Therefore, I changed my approach to digest the D3 knowledge by study more JavaScript and logic. It was very helpful to me in order to complete my final visualisation.

As this is my first semester in my Master of Data Science, there are lots of new techniques and terms that I have encountered but I have not gone deep to study them, such as data wrangling which I will study in my next semester, if I know the skill of that, it definitely would help me to optimise my dataset in order to increase the quality of my data exploration and visualisation.

Hindsight

Apparently, there are lots of things that I can do to make this visualisation better.

- If I have enough time, I would like to add filters for all players to display the user selected group of players on the interactive bubble chart, but it needs to be carefully designed. I choose 50 bubbles because it is the best number of visualisation in this case, however, my dataset contains more than 15000 players' data, and it could be very difficult to identify a player if the bubble is too small. The filter could be various, such as age, country, club, potential ranking and international reputation etc., basically, all attributes can be filters for player selection.

- b) I also want to add a year for my visualisation as my dataset has each player's data attributes from a different year, then I could have another visualisation underneath my main bubble chart to display some statistic data visualisation, like the scatterplot in my design sheet 1, example showed below.

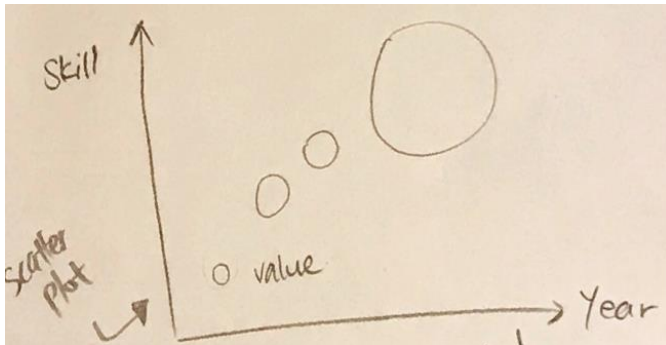


Fig. 13: Scatterplot of a player's value as a bubble, the player has grown his skills as the time goes by to increase his value

- c) I want to do is to re-position bubbles into a soccer pitch formation. The challenge could be making the re-position into formation compatible with the dragging bubble function, as the force-directed function has been used, one bubble dragging movement will affect all other bubbles. The circle packing chart can be used to group players at same field position together to put on the related soccer pitch position, the example showed below.

Circle Packing

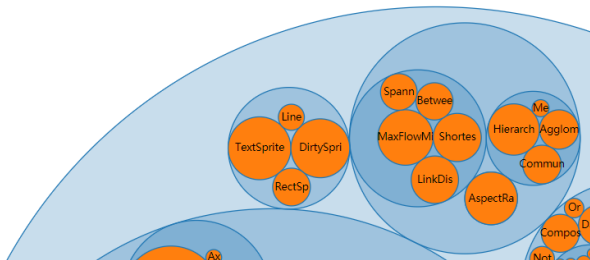


Fig. 14: Circle Packing could help to group players bubble into a soccer squad formation with a soccer pitch template

6 Reference

- [1] T. Reuters. World Cup Top Scorers. Retrieved from: https://d17h27t6h515a5.cloudfront.net/topher/2016/September/57e9a59d_11-worldcuptopscorers/11-worldcuptopscorers.png [URL]
- [2] T. Reuters. Win, Lose or draw. Retrieved from: https://d17h27t6h515a5.cloudfront.net/topher/2016/September/57e9a683_11-winloseordraw/11-winloseordraw.png [URL]
- [3] M. Bostock. Bubble Chart. Retrieved from: <https://bl.ocks.org/mbostock/4063269> [URL]
- [4] M. Bostock. Force-Directed Graph. Retrieved from: <https://bl.ocks.org/mbostock/4062045> [URL]
- [5] E. Trott. Zoomable Sunburst on d3.js v4. Retrieved from: <https://bl.ocks.org/maybelinot/5552606564ef37b5de7e47ed2b7dc099> [URL]
- [6] N. Bremer. D3.js - Radar Chart or Spider Chart - Adjusted from radar-chart-d3. Retrieved from: <http://bl.ocks.org/nbremer/6506614> [URL]