360 Degrees Analysis on the Most Beloved Football Club in English Premier League: Manchester United

ByeongJo Kong
System Design and Management,
Massachusetts Institute of Technology

Ta Hang Chen
System Design and Management,
Massachusetts Institute of Technology

ABSTRACT

Manchester United is one of the most popular soccer teams in the world. They have fans from all around the world, and fans are known for their passion for the team. The authors intend to create a visualization tool to show the significance of Manchester United in three perspectives, team performance, player statistics, and social media influence.

Keywords: English Premier League, Manchester United, Interactive Visualization, D3.js.

1 INTRODUCTION

Manchester United is considered the most popular soccer club in the world [1]. Its worldwide fan base reaches a 1.1 billion milestone in 2019. Several reasons make Manchester United one of the most successful soccer clubs in the world. Firstly, its team performance is among the top tier in the league. It has won 20 League titles, 12 FA Cups, five League Cups, and a record 21 FA Community Shields. Secondly, Manchester United has successfully created individual wonder players over several generations. For example, Best, Beckham, and Ronaldo. Fans support not only the club but these players too. Thirdly, Manchester United has a long and last relationship with non-European countries. It often tours to Asia countries and the United States, interacting with existing fans and attracting new ones.

All of these facts make us wonder whether we can create a visualization tool to show the significance of the team. By doing so, we will have a powerful tool to attract more people to join the Red Devil fan club! We decided to leverage the 6.859 final projects to create a visualization tool for users to explore three perspectives, team performance, player statistics, and social media influence. We aim to provide the users a website tool for users to understand Manchester United more through an interactive data visualization tool.

This work contains five charts: Transfer Fee Bubble Chart, Ranking Radar Chart, Winning Points Line Chart, Fans Bubble chart, and Interactive Map; From the visualization, we found that Manchester United has the top team performance, has a huge fan base in South East Asia countries, and spends the most money to keep the best players.

2 RELATED WORK

2.1 Manchester United History

Manchester United Football Club is a professional club based in Old Trafford, England, that competes in the Premier League, the top flight of English football. Nicknamed "the Red Devils," the club was founded in 1878, changed its name to Manchester United in 1902, and moved to its current stadium, Old Trafford, in 1910.

Manchester United has won more trophies than any other club in English football, with a record 20 League titles, 12 FA Cups, five League Cups, and a record 21 FA Community Shields.

On the commercial side, Manchester United was the highest-earning football club in the world for 2016–17, with annual revenue of €676.3 million, and the world's third most valuable football club in 2019, valued at £3.15 billion ($3.81 billion). Fans for Manchester United are known for their passion for the team and are very active on social media.

2.2 Bubble Chart

In our project, we used two types of bubble charts; one for visualizing the amount of money spent by the club for recruiting high-performing players and the other for visualizing the club's size of social media followers. Integrating the user interactivity into these charts required some levels of studies, particularly in dealing with the size, location, and movement of bubbles and creating a proper data pipeline for feeding the data in the correct formats.

2.3 Radar (Ring) Chart

This chart uses several important functions that regulate the overall behavior of the chart. Firstly, the ringchart(container) to construct a new instance of RingChart with default settings. The container can be an SVG or g element. size([width, height]) sets the dimensions to specified width and height and returns this chart. innerRadius(_) is used to set the chart's inner radius and palette(_) to set the color palette to a categorical scheme (for rankmap) or a continuous interpolator.

2.4 Interactive Map Chart with topoJSON

This work uses the topoJSON, which is an extension of GeoJSON and encodes the geography information. This thesis uses D3 together with topoJSON to visualize the geographic distribution of Facebook followers for Manchester United. Butler's work [2] clearly shows the ability to integrate topoJSON with D3. Also, this observable D3 demo is another work that shows how to integrate topoJSON with D3 date input [3]

3 METHODS

This section describes the techniques and algorithms used in work, including the front-end techniques and data transformation logic. Also, visualization encoding is also included in this chapter.

3.1 Bubble Chart: Player Transfer Fees

This chart shows the amount of money the club spent recruiting top players for the last three years. It offers two view options to the users, 'All' and 'By year.'

3.1.1 Data
The data was collected from the EPL players transfer records section in TransferMarket[4]. This online database holds the

decades of transfer fees records of the EPL players. For our project, we used the data from the year 2018 to 2020 (3 years).

3.1.2 Techniques and Design Decisions

Implementation of this chart is largely governed by the bubble chart creation function, which returns a function that will instantiate a new bubble chart given a DOM element to display it in and a dataset to visualize. One of the significant parts of this function is defining the sizes and locations of the bubbles. This operation regulates how the bubbles behave in terms of locations, movement, and size.

Figure 1 is a screenshot of the chart. Each bubble represents the amount of money spent by the club to recruit a particular player. Its size indicates the amount of the money. In terms of color-encoding, the red circles represent Manchester United, blue circles top 5 clubs, and the rest green circles are the rest of high-performing and low-performing clubs depending on the density of the color. As shown below, there are two view options; the 'All' option allows users to see the transfer fee records all mixed in the same area, 'By year' option allows users to see separated bubbles by years.
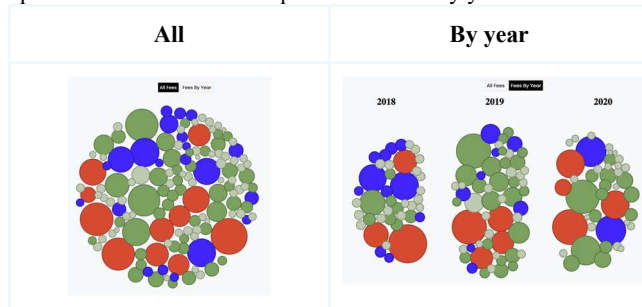

Figure 1. Players transfer fees spent by EPL clubs

3.2 Radar Chart: Team Ranking

This radar chart was used to visualize the 20 years of EPL clubs' ranking history. It offers various view options, allowing users to have more control in optimizing the visualization for their needs.

3.2.1 Data

The data was collected from the Premier League History Dashboard hosted under the official EPL website [5]. The website offers various information related to the performance of EPL clubs. We used the ranking information hosted under the Seasons review section.

3.2.2 Techniques and Design Decision

Implementing this chart required the use of various functions that controls the creating and shaping of the elements in the chart, e.g. ringchart(container), size([width, height]), innerRadius(_), palette(_). We wanted the users to control ASC/DESC orders of the bubbles, deciding whether the inner circle is the higher rank or vice versa. The users can also control the size of the inner radius, shapes of nodes, legend placement, and themes (light vs. dark) based on their preferences. Each pile of circles represents one year. As the default setting, the outer circle is the 1st place, whereas the inner circle is the last place in the ranking.
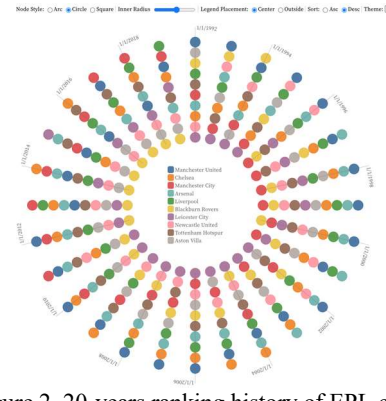

Figure 2. 20-years ranking history of EPL clubs

3.3 Multi-Line Chart: Winning Points

This multi-line chart shows Manchester United's team performance compared to other teams. The metric we use is Points. A team is awarded 3 points after winning a match or one point if the game ends as a draw or tie.

3.3.1 Data

Date used in this chart is collected from SportsMo[6] and the year range is from 2000 to 2020. The additional data transformation is to calculate the average number for each team. The final chart presents a comparison for the top 10 teams. The definition of winning point is that a team is awarded 3 points after winning a match or one point if the game ends as a draw or tie.

3.3.2 Techniques and Design Decision

This chart uses the D3 Observable multi-line techniques to present the data. The data was first collected into two arrays, data array and year array. The final chart is to plot the data based on year and Points. During the Implementation, we discussed several issues trying to get the best experience for the users.

We decided not to provide a zooming function because the data level is only at the year level. Zooming functions will not get more detailed information, such as by date. Also, we decided not to include all the 43 teams because it will complicate the screen. The final decision is to include the top 10 teams in the visualization. Lastly, to further simplify the meaning, the unselected lines are marked in a light grey color to focus on the desired comparison.

In this work, we select line charts as encoding because line encoding can convey the trend information, meaning we want to convey to the users. Lastly, this work selects winning points as the visualized metric because it considers both win and lose scenarios and can represent an overall performance for the team.
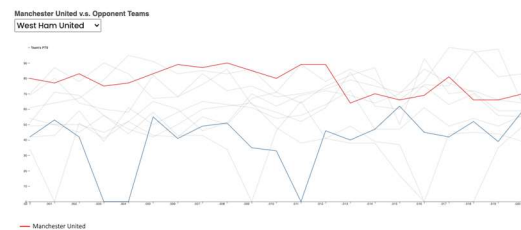

Figure 3. Points of Manchester Utd. and other competing clubs

3.4   Bubble Chart: Social Media Followers

We again used the bubble chart to visualize the number of social media followers each EPL club has. In the visualization we offer two options; one for the follower numbers by clubs, and the other one for by first team squad.

3.4.1   Data

The data was collected from PMPR Excellence. It is a website that provides comprehensive analytic data for the EPL clubs. Both the number of followers by clubs and the first team squad were obtained through the PMPR website [7].

3.4.2   Techniques and Design Decision

The design of this chart was inspired by the work of Nicolas Lambert and Kaho Cheung [8], who developed comprehensive instructions on implementing this bubble chart. Thanks to the instruction, implementing the chart was relatively easy. The data wrangling and transforming took the most of my time and efforts, as the format of obtained data did not fit properly into the chart's pipeline. In terms of design decisions, the size of the bubble represents the size of followers. We colored Manchester United in red to help the audience spot the team instantly. The users can also toggle between two view options as described above.
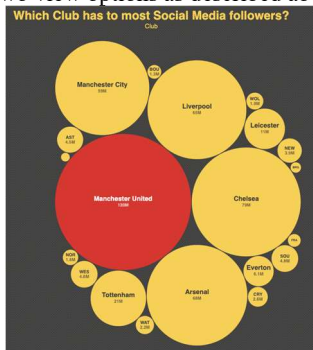

Figure 4. Number of social media followers of EPL clubs

3.5   Interactive Map: Social Media

This chart includes geographic distribution information for Manchester Club Facebook followers.

3.5.1   Data

The data used in this chart are collected from the Trackanalytic website [9]. The data contains the total number of Facebook followers for Manchester United's official Facebook site and its geographic distribution. The data is clean and is used directly in this project.

3.5.2   Techniques and Design Decision

In order to use the topoJSON, several data transformation steps are applied. For example - a filter for country code to avoid "undefined" countries. Once the data is prepared, the program maps the three digits country code to topoJSON for the visualization. Also, the tooltips function is enabled for this map.   Tooltip function is enabled because users may want to understand the detailed information for a specific area. We also decided to use color encoding to represent the population of Facebook followers.
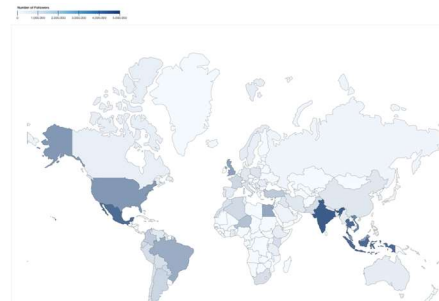

Figure 5. Facebook Follower Geographic Distribution

4   Discussion

On the player statistic side, without a doubt, Manchester United spends the most money in acquiring the players. This finding makes sense because Manchester United is one of the richest clubs in the world because they have a huge fan base.

From the Ranking Radar Chart and Winning Points Line Chart, it is clear that Manchester Club is consistently among the top-performing teams. Their performance contributes to their popularity. One thing worth noting is that the performance is not stable in recent years from 2018. It is critical for the club manager to have a strategy to stabilize the team performance.

On the social media side, it is rather surprising that Asia has strong performance, especially in SEA countries such as India, Indonesia, and Thailand. This may be because the English Empire brought soccer to these countries at the beginning of 20 century. This finding suggests that the club can host more Asia Tour to these countries to attract more fans!

5   Future Work

From a technical perspective, the tooltip function for the bubble chart - player transfer fees can be improved. Currently, it only shows text-based information about the player. However, it would be better to include a photo of the player and his profile. Having the player's photo can allow users to associate the data with the player quickly.   For the interactive map chart, it would be nice to implement navigation buttons to zoom in and out to a specific region quickly.

From a data perspective, it would be interesting if we can have more data to analyze the relationship between the number of transfer fees spent by the club correlates with the team performance. It is commonly seen that football clubs spend a great deal of money to recruit star players, but the player's performance does not meet the expectation. Having this analysis can have some interesting insights.

Also, in terms of the social media data, Twitter data can show different insights from Facebook data. It would be interesting if we can compare Facebook with Twitter data. This comparison can generate insights for the club on how to attract more fans on social media.

References

[1]   B1- R. Fay, "Manchester United named as world's most popular football club," *Manchester Evening News*, Aug. 17, 2019. https://www.manchestereveningnews.co.uk/sport/football/football-news/manchester-united-fans-news-latest-16771943 (accessed May 16, 2021).

[2]   P. Butler, "Mapping Temporal Datasets with D3," *Cartographic Perspectives*, no. 81, pp. 44–48, 2015.

[3]   topojson/topojson-simplify. TopoJSON, 2021. Accessed: May 19, 2021. [Online]. Available: https://github.com/topojson/topojson-simplify

[4]   "Premier League - Transfer records." https://www.transfermarkt.com/premier-league/transferrekorde/wettbewerb/GB1 (accessed May 20, 2021).

[5]   "Premier League History, Records, Origins & Historical Stats." https://www.premierleague.com/history/dashboard (accessed May 20, 2021). EPL table, PMPR Excellence, https://www.prexcellence.com/epl-tables

[6]   "Premier League 1999-00 Stats Centre Overview," *Sports Mole*. https://www.sportsmole.co.uk/football/premier-league/1999-00/ (accessed May 19, 2021).

[7]   "English Premier League Social Influencer Tables," *PMPR Excellence*. https://www.prexcellence.com/epl-tables (accessed May 20, 2021).

[8]   COVID-19 Bubble Chart with D3 Render, Oberservablehq, https://observablehq.com/@unkleho/covid-19-bubble-chart-with-d3-render

[9]   "Trackalytics | Social Media Statistics / Analytics & Website Statistics / Analytics Tools." https://www.trackalytics.com/ (accessed May 19, 2021).