# FINAL PROJECT REPORT

# ENGLISH PREMIER LEAGUE 2018-19

**Jas Patil, Dishay Shah**

**iSchool, University of Illinois at Urbana-Champaign**

**IS 525: Data Warehousing and BI**

**Prof. Michael Wonderlich**

**May 10th 2022**

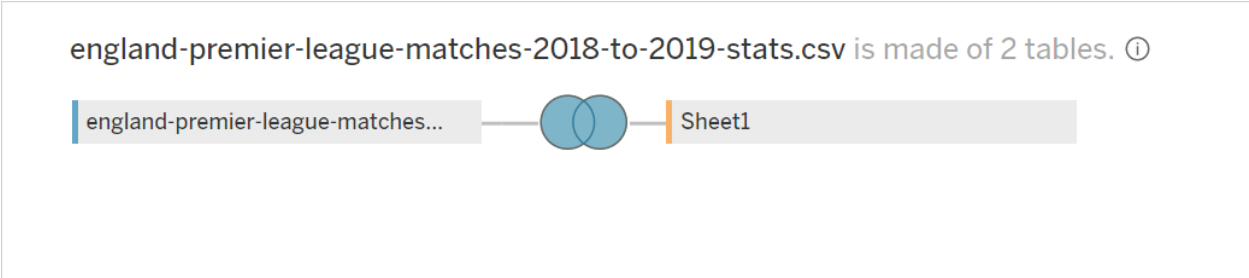
## 1. Introduction:

The most watched and played sport in the world has leagues in almost every country. We have taken the dataset for the most popular league of them all known as English Premier League where it is played in United Kingdom. As it is a year long event where 20 teams participate, 38 games are played by each team and so many other things to consider we need to find detail for the sporting enthusiasts.

The datasets that we have used for the project contains all the information to be analyzed for the year 2018-19. We selected this dataset because of our love and passion for the game and we needed to cross verify that the teams we support are better than the rest. It was hard to find accurate datasets as we had closely watched the entire season, so we knew the exact standings and some of the datasets out there had mistakes, so we thought of creating a dataset and join them using names so that the standings are accurate, and we get the right visualization and outcomes.

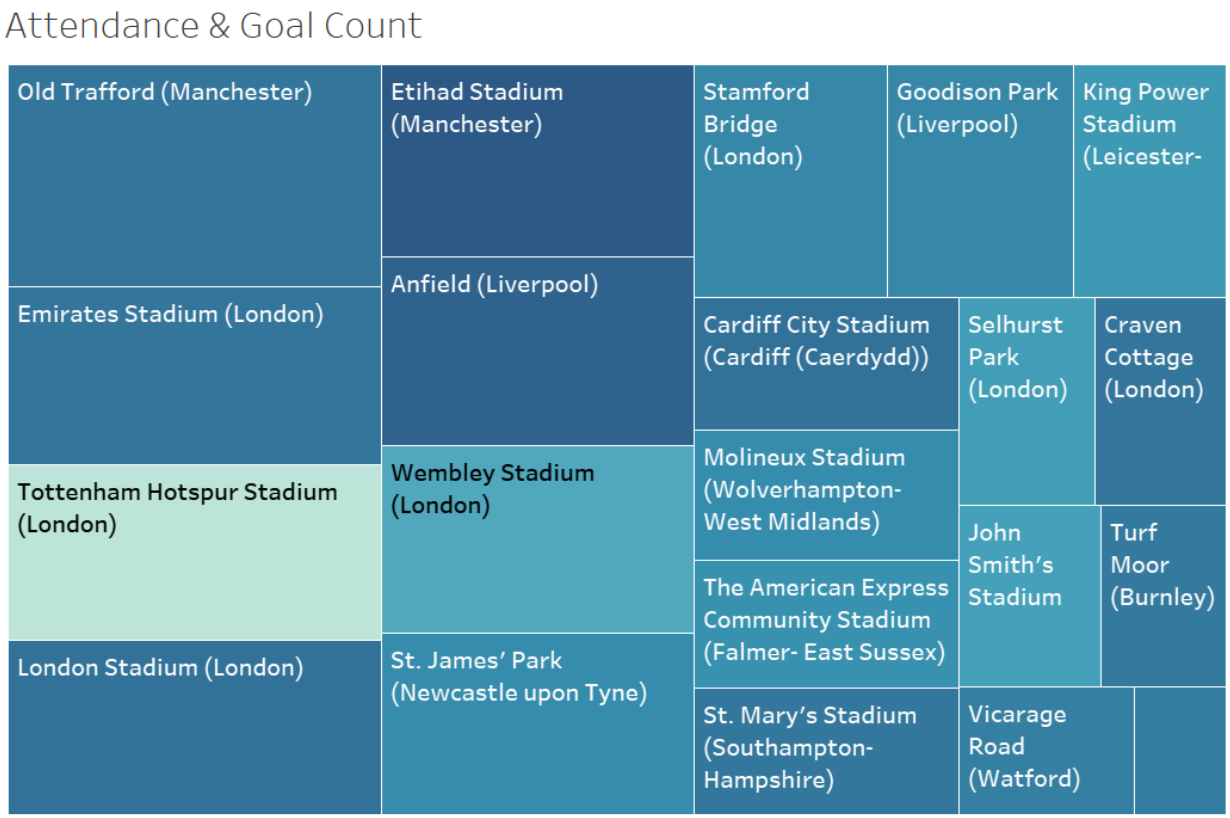
The main dataset has thousands of rows, and it was detailed at the very core and we got that dataset from Kaggle. Our work attempts to understand the patterns of teams and their performance for the particular season of 2018-19. Using the visualizations, we have analyzed the difference in both the trends, and based on the conclusions we will be able to design recommendations for every team and the areas where they can improve to increase their standings better for the next season.

The datasets had too much of unwanted data and more than necessary null columns, so we had to process it and clan it and then extract using MS excel. After getting the processed dataset, we identified relationships between the various entities of the dataset. Once we understood the data, the final task before gaining insights from the data was to visualize it. For creating elaborate visualizations and dashboards based on our data we used Tableau Desktop.



## 2. Visualizations:

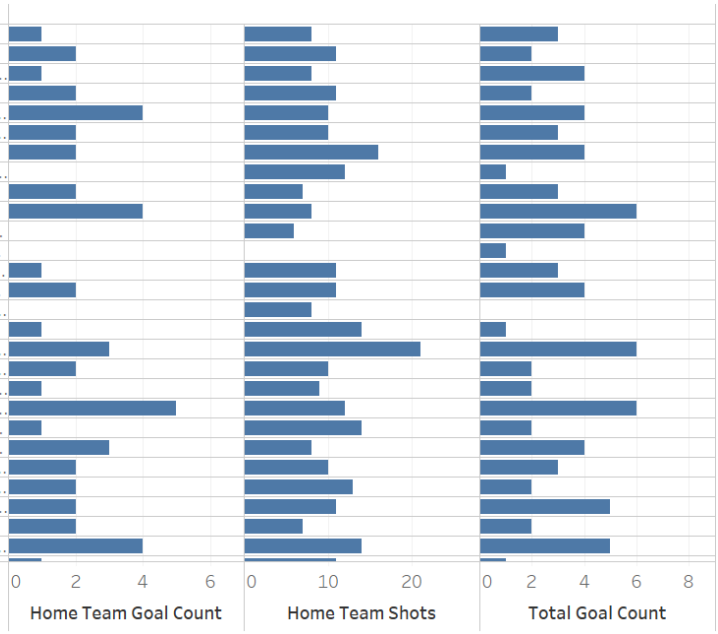
a. **Attendance and Goal count –** The first visualization that we did is just a basic heat map where we used Avg Attendance for the size of tree map, color of Total Goal count, Text for Stadium name and detail for team name. So basically it just a way to see the which stadium had the most attendance and the most goals in that stadium.




b. **Goals and Shots:** Here we have used column chart to explain how the home team performed vs all the rest of the teams and the game week when it happened. It tells in detail about how many goals that team scored and how many shots that team take.

## Goals And Shots

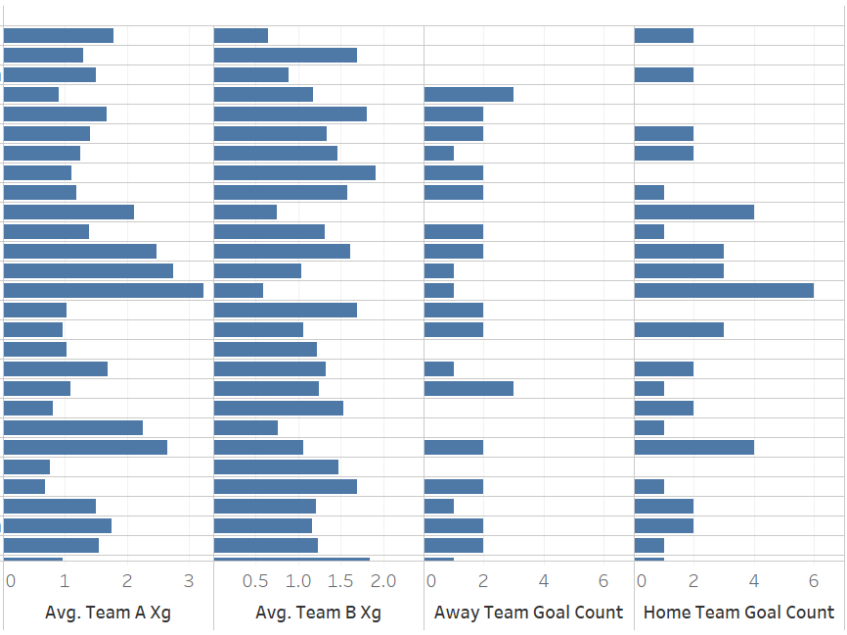
| Team Name | Away Team Name | Game .. | date GMT |
|---|---|---|---|
| AFC Bournemouth | Arsenal | 13 | Nov 25 2018 - 1:.. |
| | Brighton & Hove .. | 18 | Dec 22 2018 - 3:.. |
| | Burnley | 33 | Apr 06 2019 - 2:0.. |
| | Cardiff City | 1 | Aug 11 2018 - 2:.. |
| | Chelsea | 24 | Jan 30 2019 - 7:4.. |
| | Crystal Palace | 7 | Oct 01 2018 - 7:0.. |
| | Everton | 3 | Aug 25 2018 - 2:.. |
| | Fulham | 35 | Apr 20 2019 - 2:0.. |
| | Huddersfield To.. | 15 | Dec 04 2018 - 7:.. |
| | Leicester City | 5 | Sep 15 2018 - 2:.. |
| | Liverpool | 16 | Dec 08 2018 - 12.. |
| | Manchester City | 29 | Mar 02 2019 - 3:.. |
| | Manchester Unit.. | 11 | Nov 03 2018 - 12.. |
| | Newcastle United | 31 | Mar 16 2019 - 3:.. |
| | Southampton | 9 | Oct 20 2018 - 2:0.. |
| | Tottenham Hots.. | 37 | May 04 2019 - 1.. |
| | Watford | 21 | Jan 02 2019 - 8:0.. |
| | West Ham United | 23 | Jan 19 2019 - 3:0.. |
| | Wolverhampton .. | 27 | Feb 23 2019 - 3:0.. |
| Arsenal | AFC Bournemouth | 28 | Feb 27 2019 - 7:4.. |
| | Brighton & Hove .. | 37 | May 05 2019 - 3:.. |
| | Burnley | 18 | Dec 22 2018 - 12.. |
| | Cardiff City | 24 | Jan 29 2019 - 7:4.. |
| | Chelsea | 23 | Jan 19 2019 - 5:3.. |
| | Crystal Palace | 35 | Apr 21 2019 - 3:0.. |
| | Everton | 6 | Sep 23 2018 - 3:.. |
| | Fulham | 21 | Jan 01 2019 - 3:0.. |



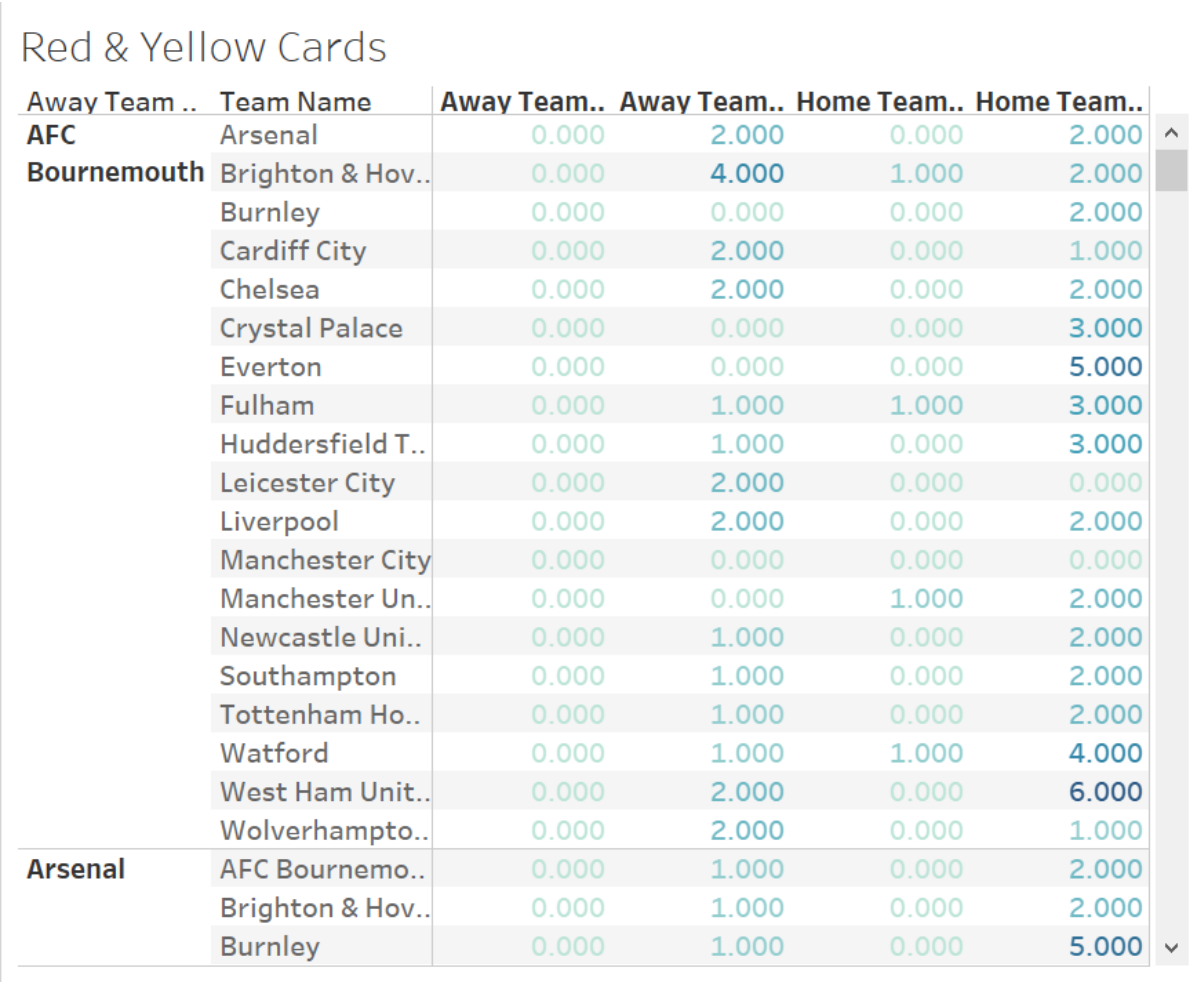

**c. Expected Goals –** This the most talked and anticipated attribute before any game. It talks about all games played in each week and expected goals for both home team and away team and actual goals in each game scored by both teams.

## Expected Goals

| Game .. | Away Team Name | Team Name |
|---|---|---|
| 1 | Brighton & Hove .. | Watford |
| | Burnley | Southampton |
| | Cardiff City | AFC Bournemouth |
| | Chelsea | Huddersfield To.. |
| | Crystal Palace | Fulham |
| | Everton | Wolverhampton .. |
| | Leicester City | Manchester Unit.. |
| | Manchester City | Arsenal |
| | Tottenham Hots.. | Newcastle United |
| | West Ham United | Liverpool |
| 2 | AFC Bournemouth | West Ham United |
| | Arsenal | Chelsea |
| | Fulham | Tottenham Hots.. |
| | Huddersfield To.. | Manchester City |
| | Liverpool | Crystal Palace |
| | Manchester Unit.. | Brighton & Hove .. |
| | Newcastle United | Cardiff City |
| | Southampton | Everton |
| | Watford | Burnley |
| | Wolverhampton .. | Leicester City |
| 3 | Brighton & Hove .. | Liverpool |
| | Burnley | Fulham |
| | Cardiff City | Huddersfield To.. |
| | Chelsea | Newcastle United |
| | Crystal Palace | Watford |
| | Everton | AFC Bournemouth |
| | Leicester City | Southampton |



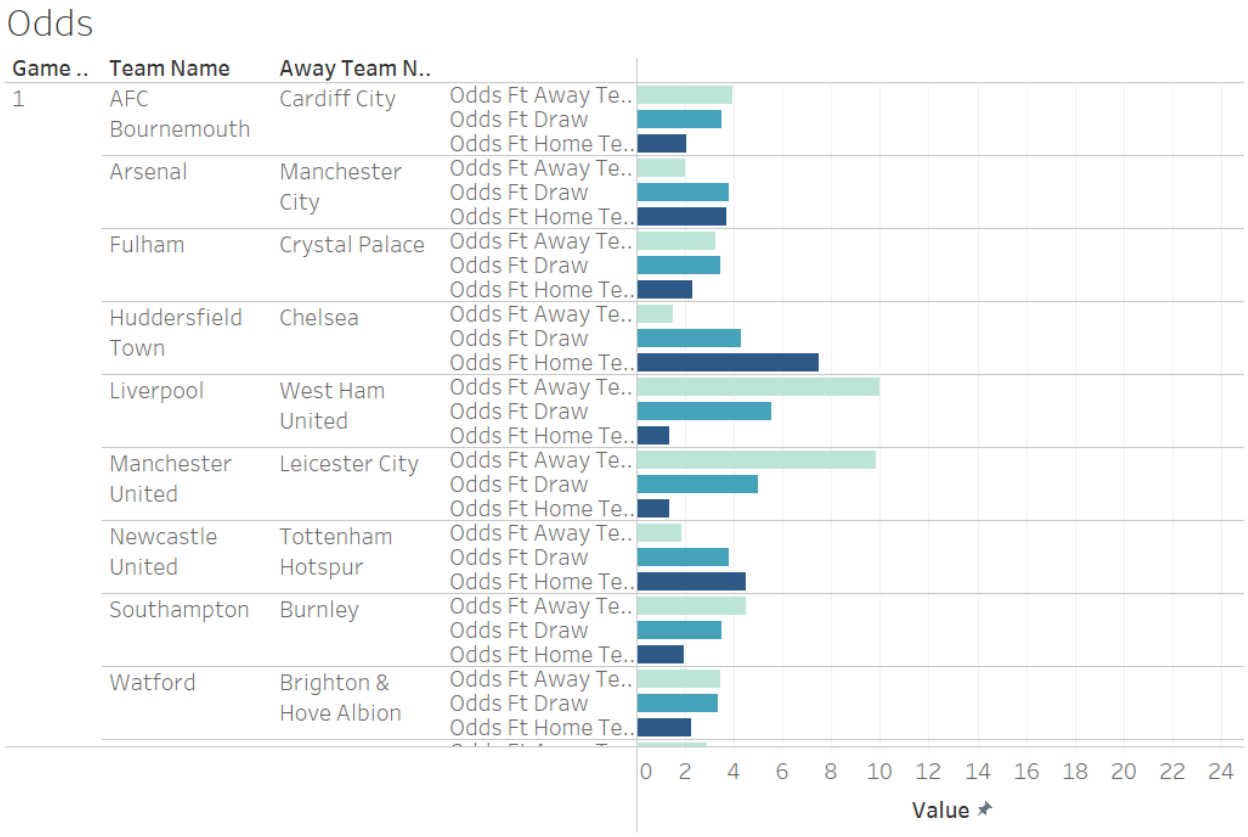
**d.** Red Cards and Yellow Cards - This visualization just shows the sportsmanship of each team by showing how many red or yellow cards did they get in that match.



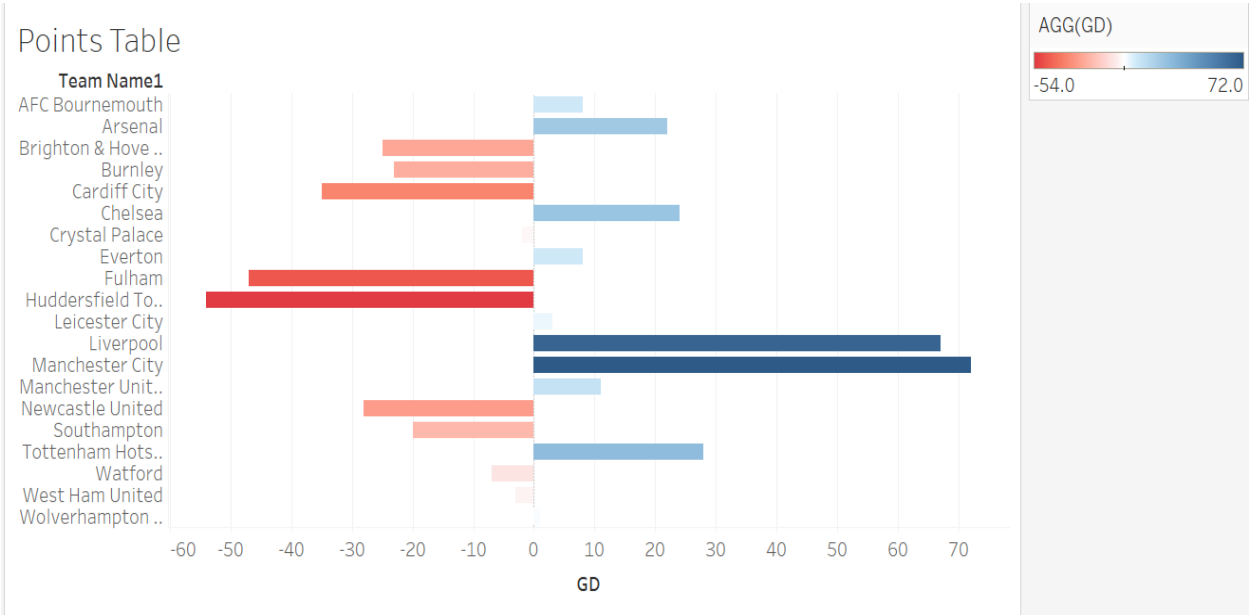
Red & Yellow Cards

| Away Team .. | Team Name | Away Team.. | Away Team.. | Home Team.. | Home Team.. |
|---|---|---|---|---|---|
| AFC Bournemouth | Arsenal | 0.000 | 2.000 | 0.000 | 2.000 |
| | Brighton & Hov.. | 0.000 | 4.000 | 1.000 | 2.000 |
| | Burnley | 0.000 | 0.000 | 0.000 | 2.000 |
| | Cardiff City | 0.000 | 2.000 | 0.000 | 1.000 |
| | Chelsea | 0.000 | 2.000 | 0.000 | 2.000 |
| | Crystal Palace | 0.000 | 0.000 | 0.000 | 3.000 |
| | Everton | 0.000 | 0.000 | 0.000 | 5.000 |
| | Fulham | 0.000 | 1.000 | 1.000 | 3.000 |
| | Huddersfield T.. | 0.000 | 1.000 | 0.000 | 3.000 |
| | Leicester City | 0.000 | 2.000 | 0.000 | 0.000 |
| | Liverpool | 0.000 | 2.000 | 0.000 | 2.000 |
| | Manchester City | 0.000 | 0.000 | 0.000 | 0.000 |
| | Manchester Un.. | 0.000 | 0.000 | 1.000 | 2.000 |
| | Newcastle Uni.. | 0.000 | 1.000 | 0.000 | 2.000 |
| | Southampton | 0.000 | 1.000 | 0.000 | 2.000 |
| | Tottenham Ho.. | 0.000 | 1.000 | 0.000 | 2.000 |
| | Watford | 0.000 | 1.000 | 1.000 | 4.000 |
| | West Ham Unit.. | 0.000 | 2.000 | 0.000 | 6.000 |
| | Wolverhampto.. | 0.000 | 2.000 | 0.000 | 1.000 |
| Arsenal | AFC Bournemo.. | 0.000 | 1.000 | 0.000 | 2.000 |
| | Brighton & Hov.. | 0.000 | 1.000 | 0.000 | 2.000 |
| | Burnley | 0.000 | 1.000 | 0.000 | 5.000 |

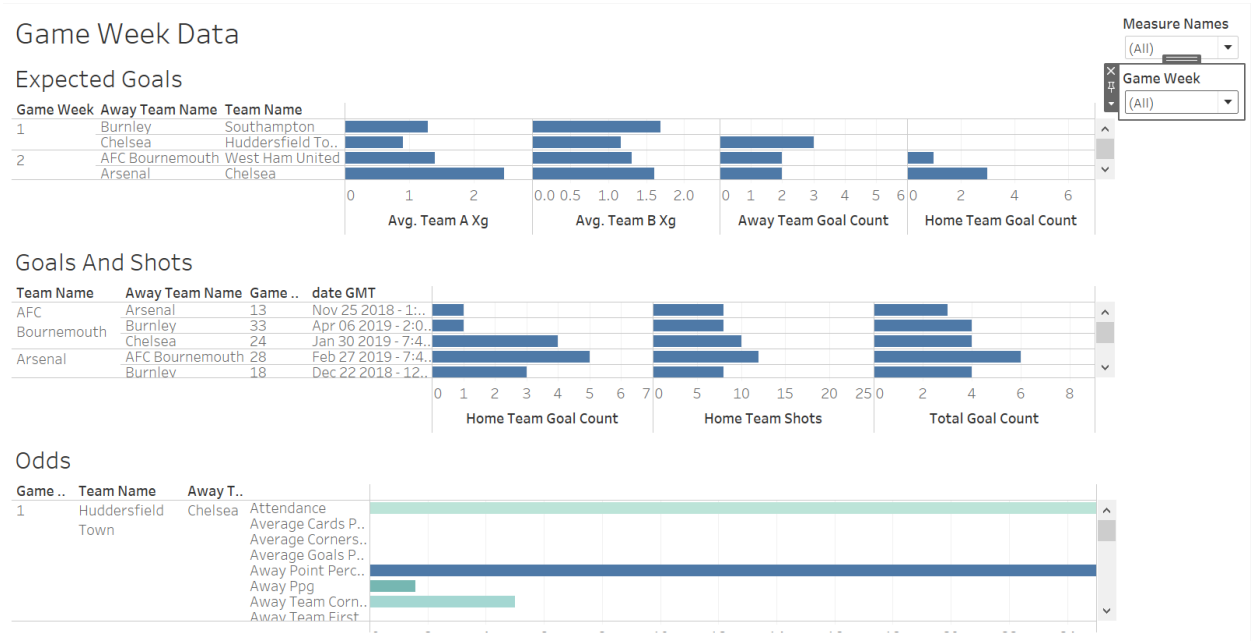
**e.** **Odds –** It is legal in United Kingdom to bid money on matches and so we made a visualization to show odds of home team winning vs away team winning vs draw. And we used measure values to show the color so all the matches have same colour bar and they can be read easily.



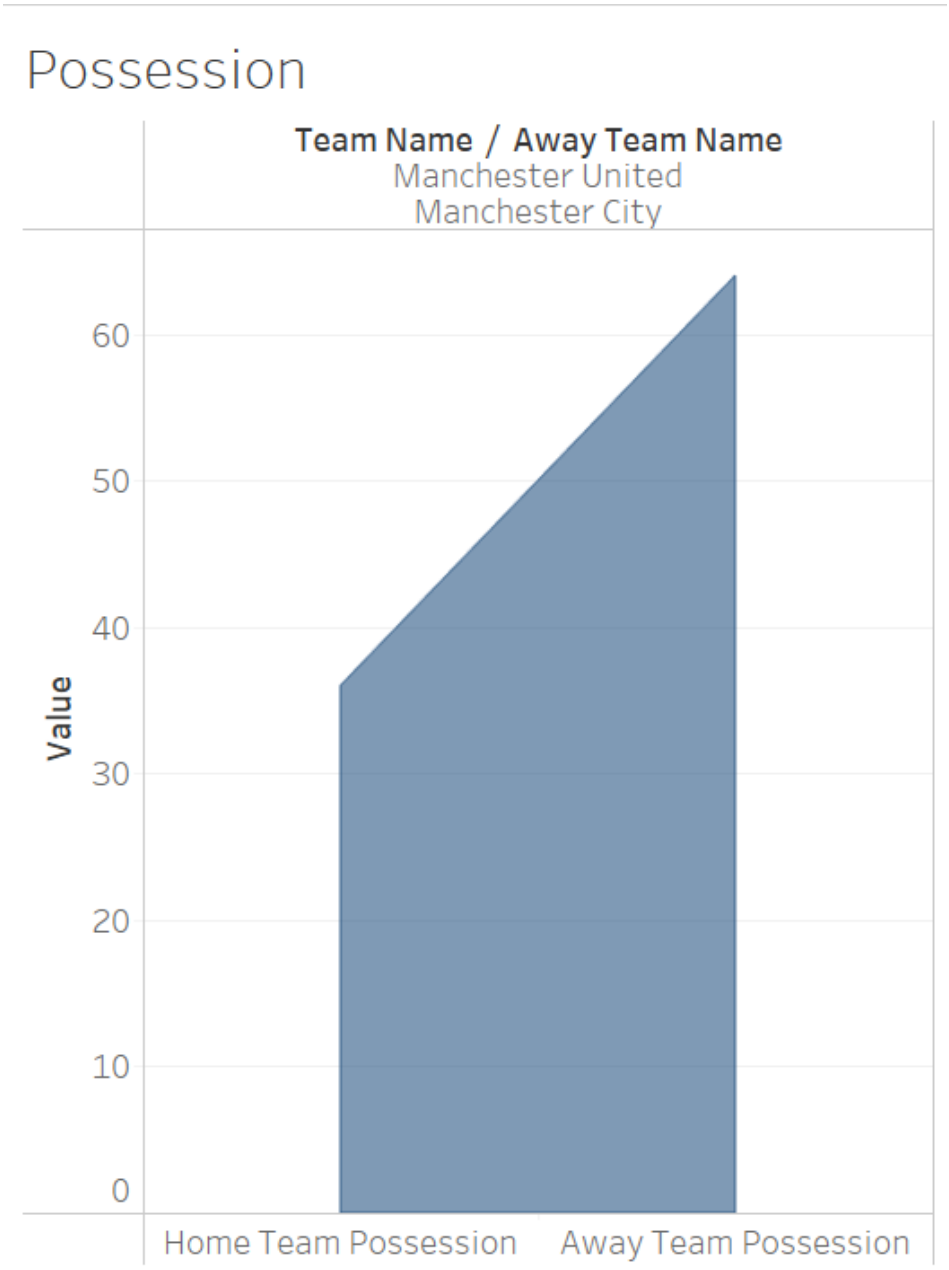

**f. Points Table –** This table shows the points table of goal differences of each team and the blue colour signifies that the team has positive difference and the red colour signifies the negative difference.



g. **Game Week Data –** This dashboard talks 3 measures between any game week selected. We can see the goals and shots each team made and scored, expected goals before the game and odds of which team were more to win the game. We can also see the time when the match took place. We have also made a duplicate of these goals and shots measure so that it does not change the other dashboard.



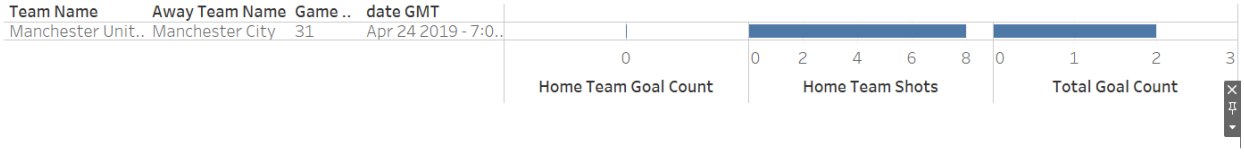

h. **Possession –** This is the most important attribute of the game as it talks about which team had the possession in that game i.e which team had the ball and controlled the game during the entire duration of the game.




i. **Dashboard 2 –** This dashboard talks about goal count of the entire game between any two teams, the stadium where the match took place, the number of cards which the home and away team got in the entire game and possession.

## Game Data Analysis

### Goals And Shots

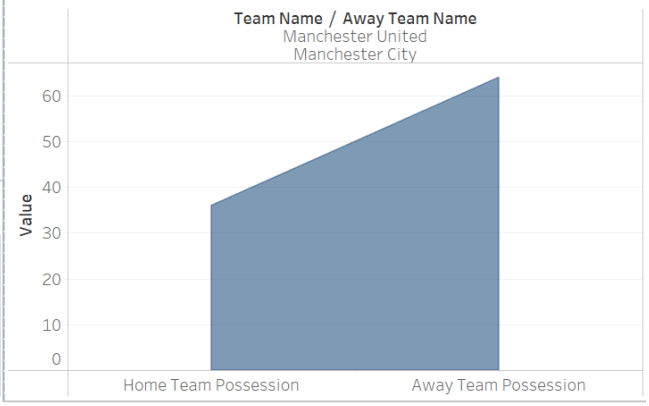



### Attendance & Goal Count


### Red & Yellow Cards

| Away .. | Team .. | Away Tea.. | Away Tea.. | Home Tea.. | Home Tea.. |
|---|---|---|---|---|---|
| Manchester City | Manchester United | 0.000 | 2.000 | 0.000 | 2.000 |

### Possession




**References – uploading the csv used for the project.**