

Football Data Visualization

Rohan Gurubhaiye, Vellore Institute of Technology, Chennai.

John Mathew, Vellore Institute of Technology, Chennai.

Abstract:

The aim of this project is to create a various visualization into which football match data can be provided as input and intuitive visualizations about all aspects of the game is generated as the output by following the process as data scrapping on python/R-Programming, and then performing visualisation on Tableau.

Keyword: Data Visualization, Football data visualization, Tableau

1.Introduction

In this project, we tried to visualize the various aspects of football match and player's statistics to get the better insight on the specific players, Squad. Using our visualization different teams can get insights like where are they lacking and what kind of strategies can work against which kind of teams. The Dataset is scrapped from the website: www.fbref.com using python and R-Programming. All the Data visualization where performed on Tableau using the data that was scrapped.

1.2 Software Requirements

- Python (Jupyter Notebook)
 - BeautifulSoup(Scrapping).
 - Pandas(Dataframe Editing)
- R-Studio
 - UnderstatR(Scrapping)
 - BS4
- Tableau(Visualization)

1.3 Challenges

The major challenge while working with this project was to extracting/scrapping the data from the respective websites.

The Feature selection for the visualization was quite difficult

2. Data Visualization and Analysis

The below are the visualizations that we performed on football match data.

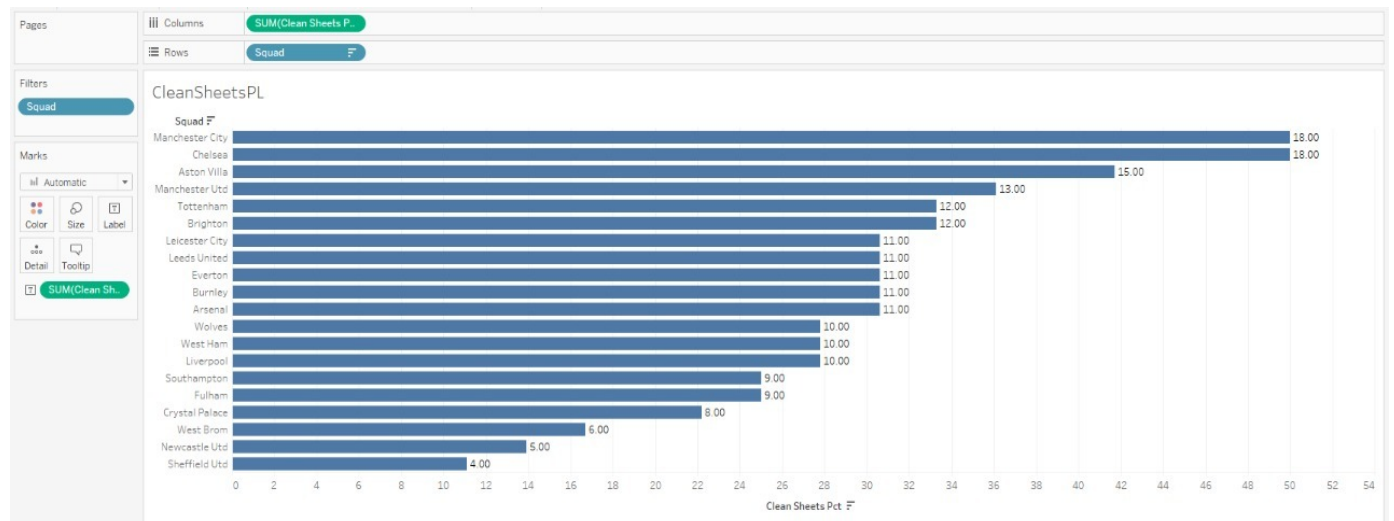


Fig 1: Clean Sheet.

In the above visualisation, we tried to show the ranking of various squad based on their Clean Sheet PL. After observing the plot, we can say that the Squad Manchester City Topped the Clean Sheets.

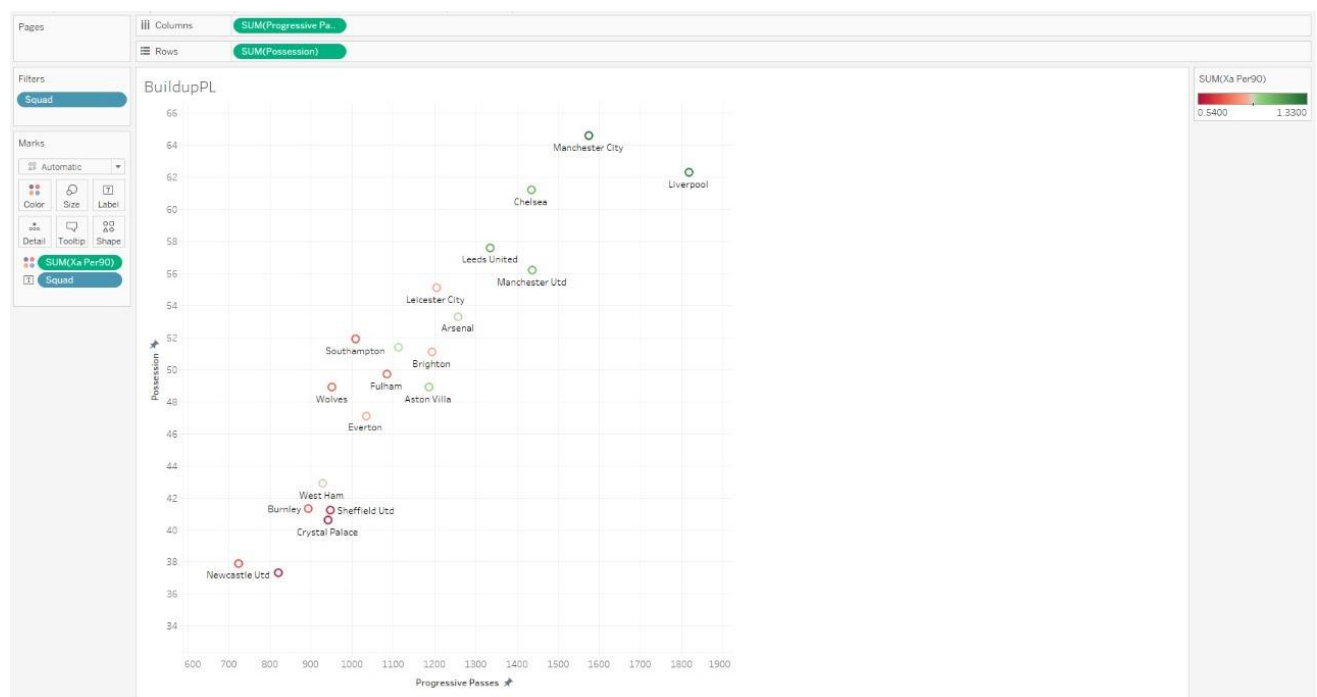


Fig 2: Possessions Versus Progressive Passes.

In the above visualisation, we tried to show the ranking of various squad based on their Possessions Versus Progressive Passes. We used Xa_Per_90 for colour mapping. After observing the plot, we can say that the Squad Manchester City, Liverpool, Chelsea are the Top 3 squads having high Possessions Versus Progressive Passes.



Fig 3: Touch Distribution: Attacking Third

In the above visualisation, we tried to show the ranking of various squad based on Touch Distribution: Attacking Third. We used Winning goals for colour mapping. After observing the plot, we can say that the Squad Manchester City toppled the Attacking Third.



Fig 4: Touch Distribution: Defensive Third

In the above visualisation, we tried to show the ranking of various squad based on Touch Distribution: Defensive Third. We used Winning goals for colour mapping. After observing the plot, we can say that the Squad Leeds United toppled the Defensive Third.



Fig 5: Touch Distribution: Mid Third

In the above visualisation, we tried to show the ranking of various squad based on Touch Distribution: Mid Third. We used Winning goals for colour mapping. After observing the plot, we can say that the Squad Leeds United toppled the Mid Third.

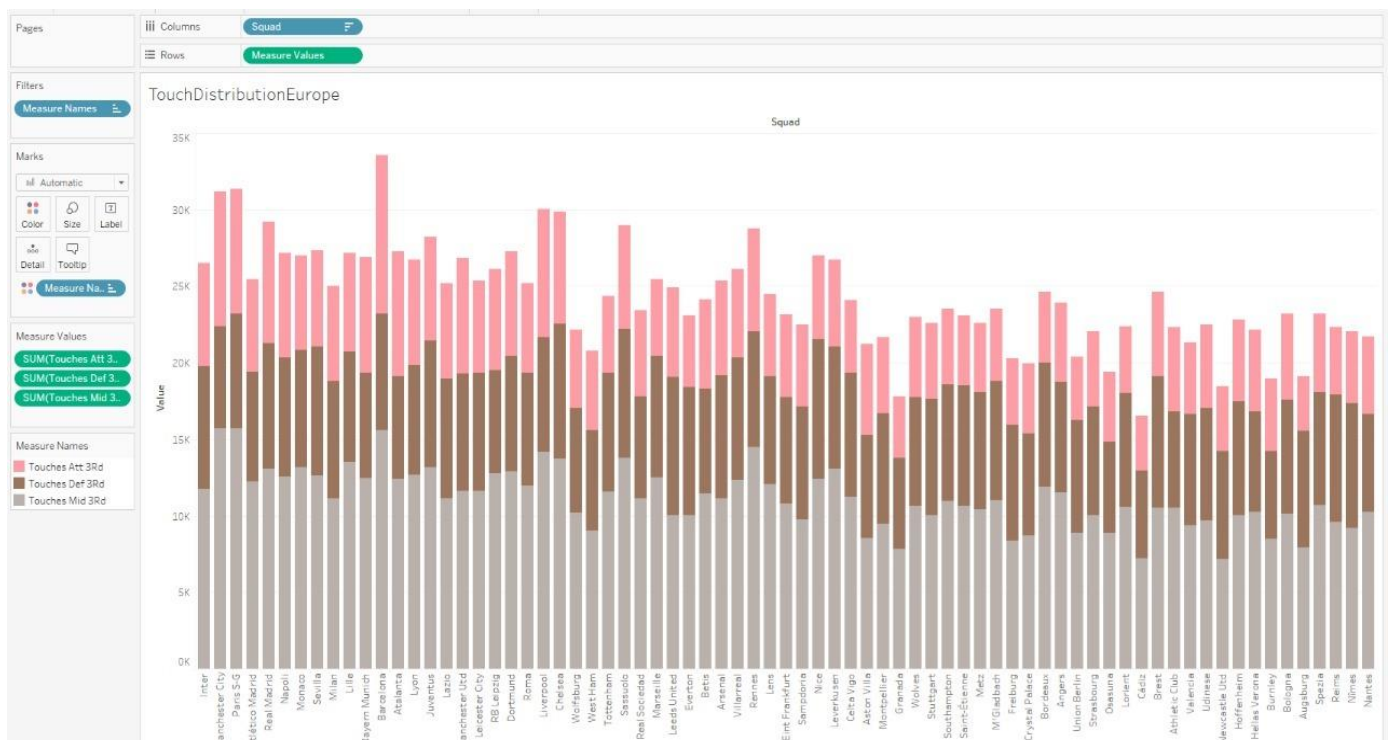


Fig 6: All Teams Touch Distribution Sorted by Number of Goals Scored.

In the above visualisation, we tried to All Teams Touch Distribution Sorted based on Number of Goals Scored.



Fig 7: Player xG per Situation-Freekicks

In the above visualisation, we tried to show the Player expected goal per Situation based on Free Kicks.

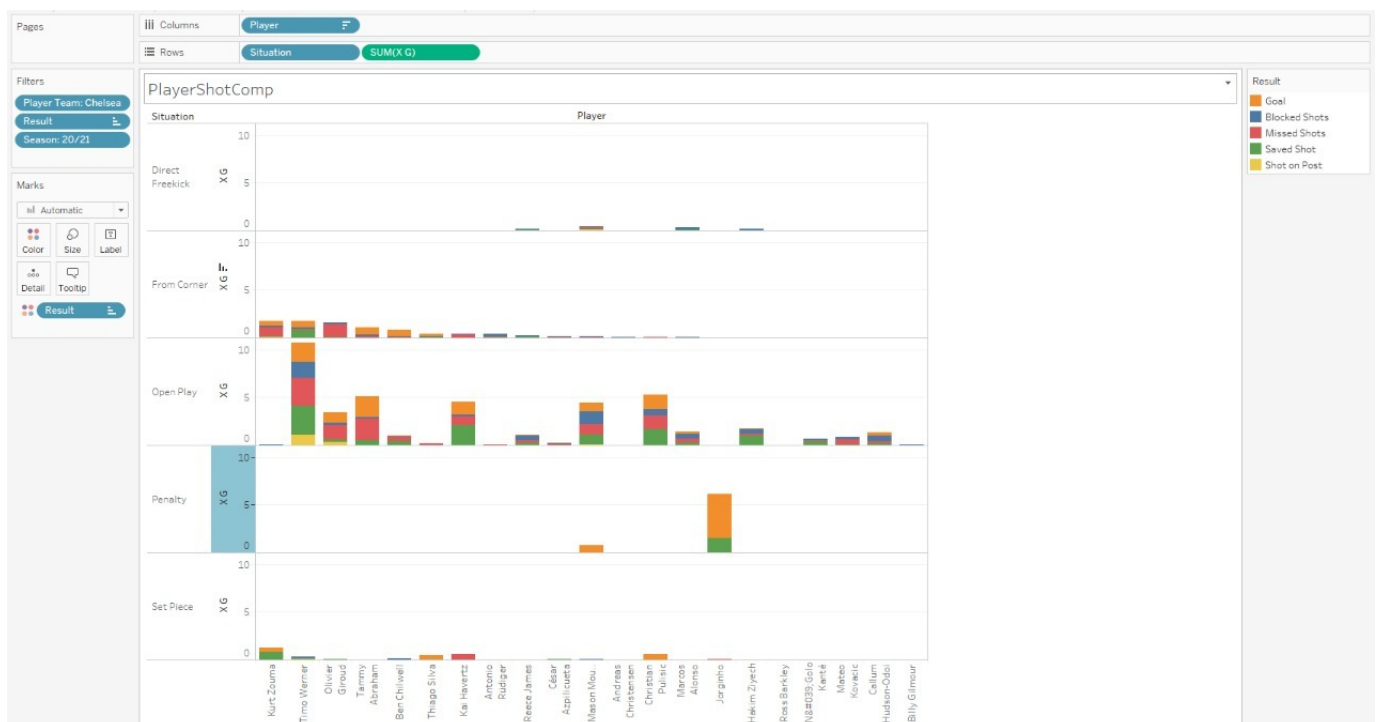


Fig 8: Player xG per Situation-Corner

In the above visualisation, we tried to show the Player expected goal per Situation based on Corner Shots.

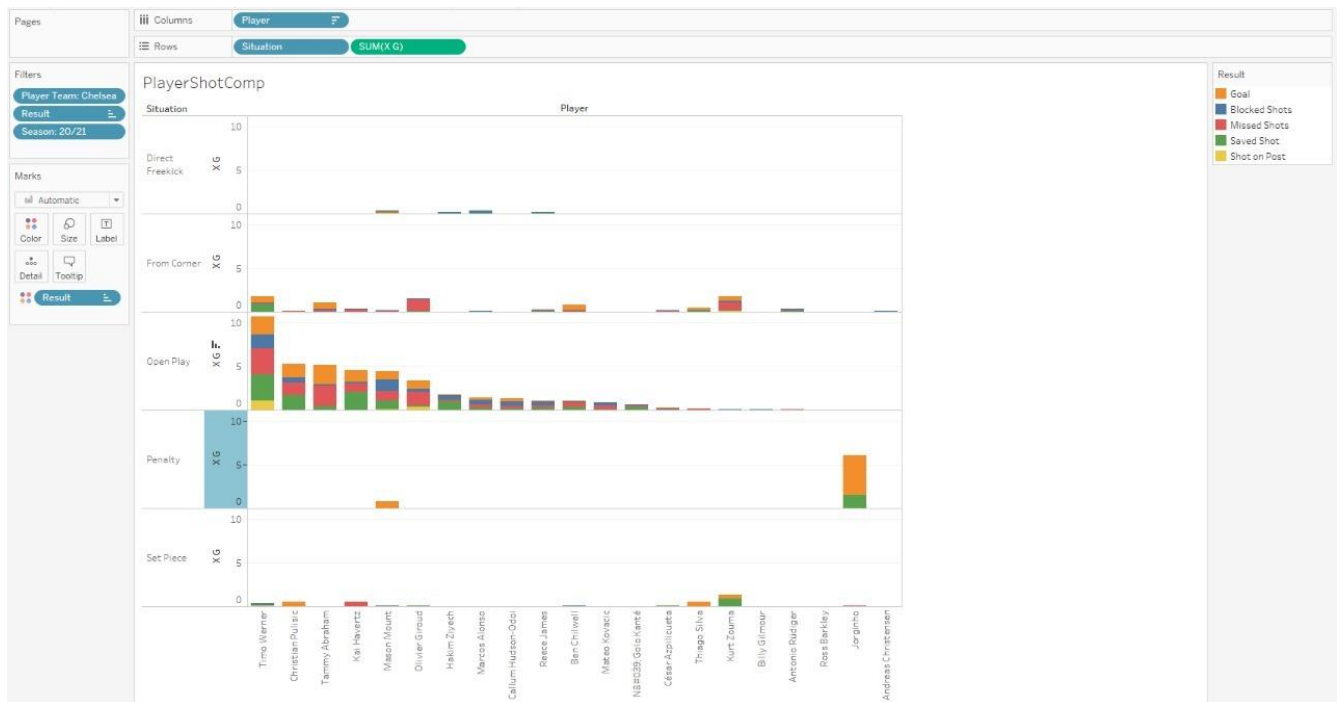


Fig 9: Player xG per Situation-Open Play

In the above visualisation, we tried to show the Player expected goal per Situation based on Open Plays.

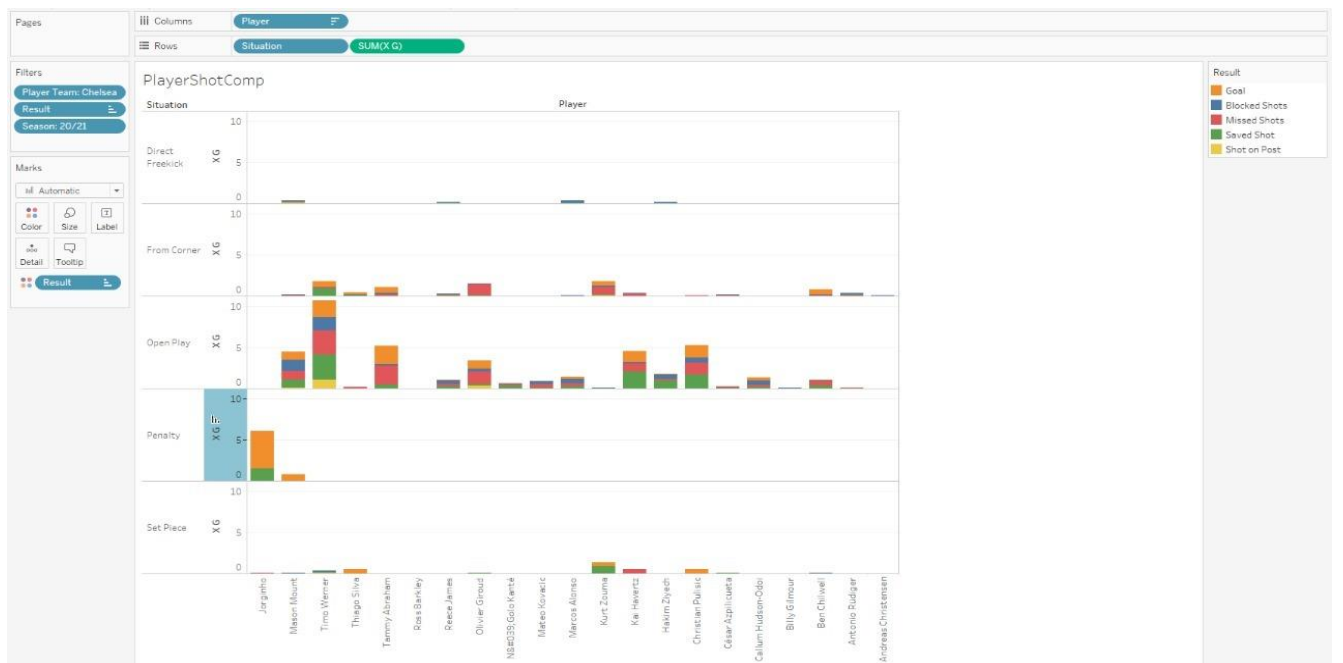


Fig 10: Player xG per Penalties

In the above visualisation, we tried to show the Player expected goal per Situation based on Penalties.



Fig 11: Player xG per Situation-Set Piece

In the above visualisation, we tried to show the Player expected goal per Situation based on Set Piece.

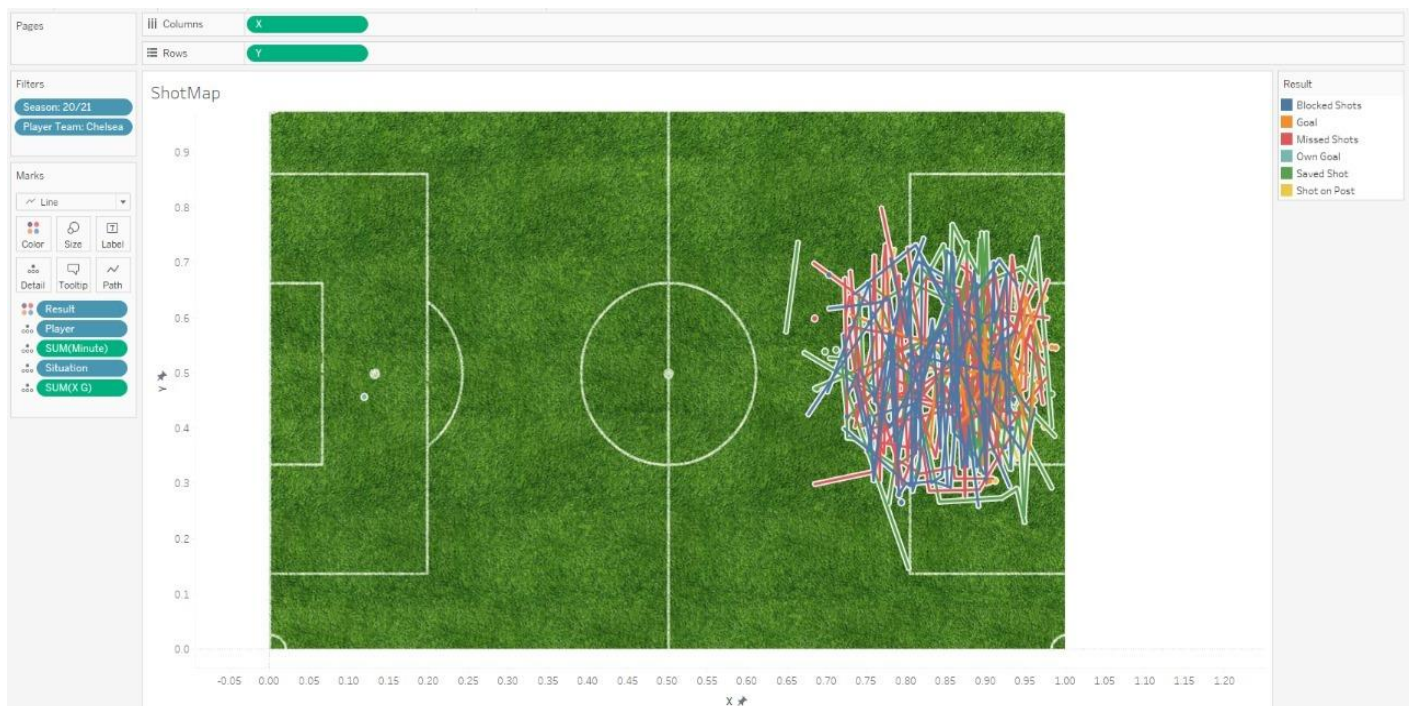


Fig 12: Full Trajectory of Various Shots

In the above visualisation, we tried to show trajectory of the various shots of team Chelsea.

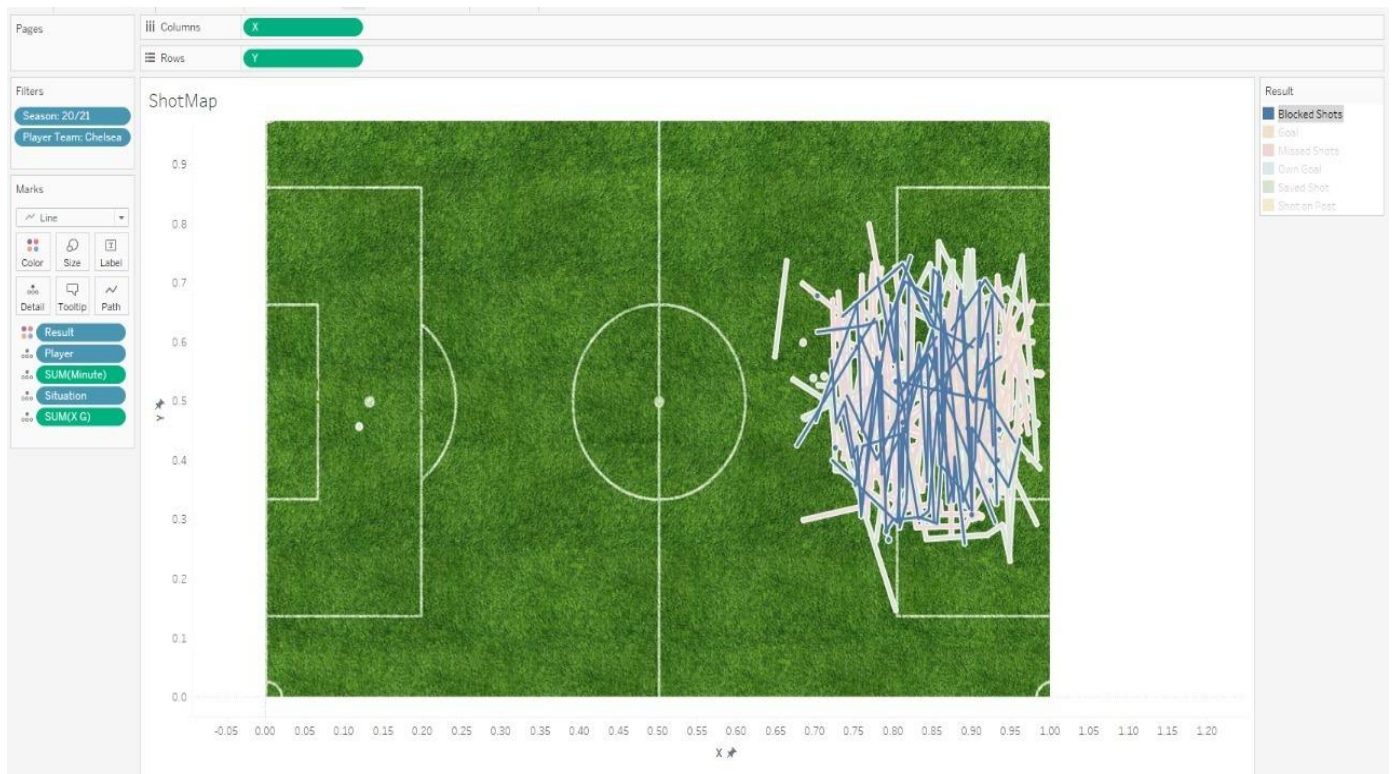


Fig 13: Blocked Shots

In the above visualisation, we tried to show trajectory of the blocked shots of team Chelsea.



Fig 14: Goal Shots

In the above visualisation, we tried to show trajectory of the shots that were goals of team Chelsea.



Fig 15: Missed Shots

In the above visualisation, we tried to show trajectory of the shots that were missed by team Chelsea.



Fig 16: Saved Shots

In the above visualisation, we tried to show trajectory of the shots that were Saved by the team opposite to Chelsea.

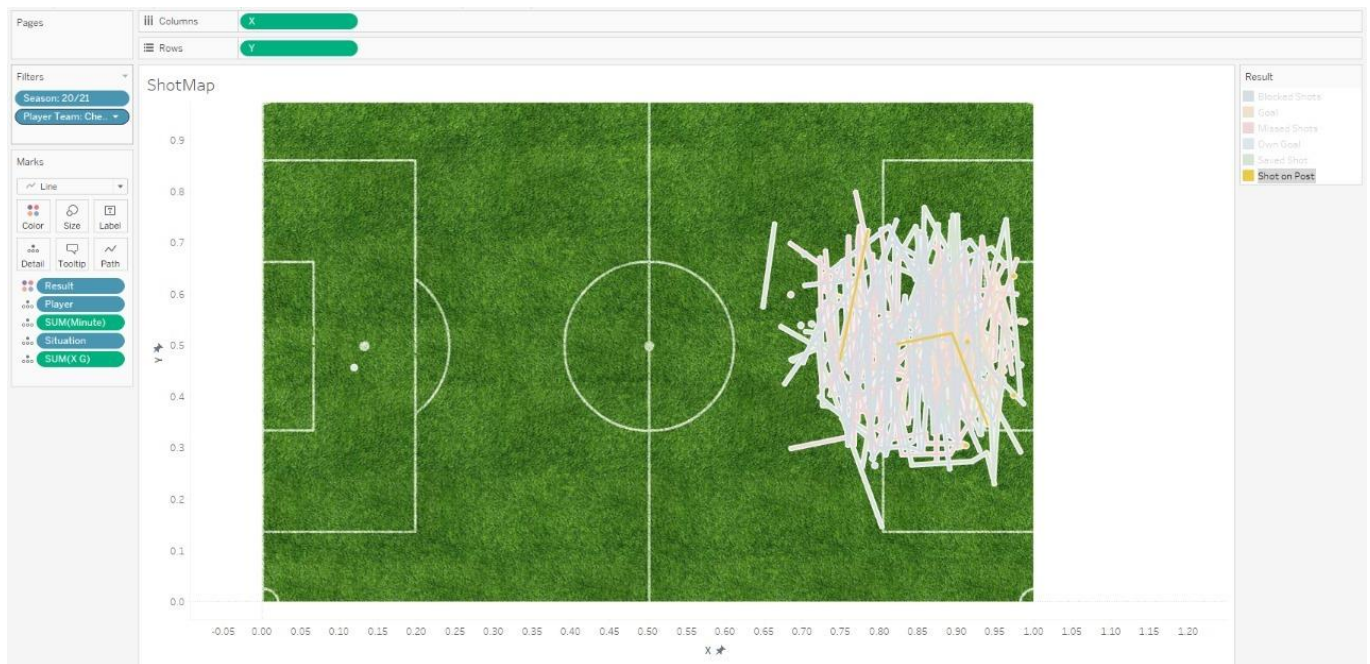


Fig 17: Shots on Post

In the above visualisation, we tried to show trajectory of the shots that were on post by team Chelsea.

3.Conclusion

After Exploring the football dataset, we get various insights on the statistics of different players, performance of teams based on various expects like types of shots (missed, goal, saved, blocked, post), Player expected goals as per situation (penalties, set-piece, corner, open play, free kicks), touch distribution. All this insight can help the team to proceed further to work where are the lacking, deciding strategies to tackle the opponent teams or to study the opponent teams or for further improvements.

5.References:

<https://realpython.com/beautiful-soup-web-scraper-python/>

<https://github.com/ewenme/understatr>

<https://dhvanil.com/football>

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., and Carpendale, S. Empirical Studies in Information Visualization: Seven Scenarios. IEEE transactions on visualization and computer graphics 18, 9 (2011), 1520–1536.

Lau, A., and Vande Moere, A. Towards a Model of Information Aesthetic Visualization. Conference on Information Visualisation (IV'07), IEEE (2007), 87–92.