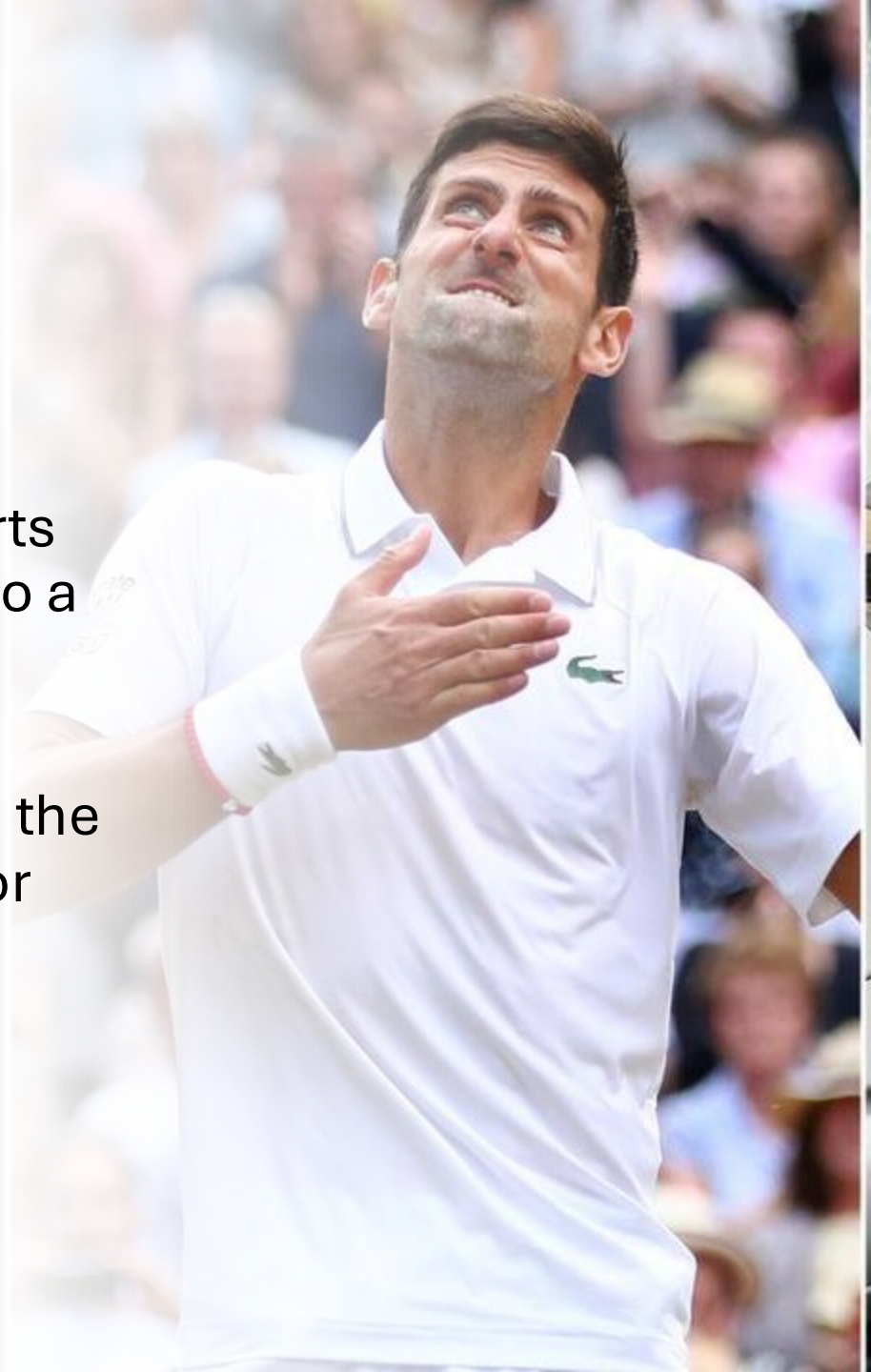
Analysis on English Premier League

By: Ayush Nelli(Group 18)

Why this dataset?

- I've always loved sports and have wanted to do a project like this.
- When I found this dataset, I knew it was the perfect opportunity for me.



ENGLISH PREMIER LEAGUE

20 teams

Each team plays 38 games whereas in NBA teams play 82

There are no playoffs

THE FINAL DAY

MW 37

Pos	Club	P	GD	Pts
1	Man City	37	60	88
2	Arsenal	37	61	86
3	Liverpool	37	43	79
4	Aston Villa	37	20	68
5	Spurs	37	10	63
6	Chelsea	37	13	60
7	Newcastle	37	21	57
8	Man Utd	37	-3	57
9	West Ham	37	-12	52
10	Brighton	37	-5	48
11	Bournemouth	37	-12	48
12	Crystal Palace	37	-6	46
13	Wolves	37	-13	46
14	Fulham	37	-8	44
15	Everton	37	-10	40
16	Brentford	37	-7	39
17	Nott'm Forest	37	-19	29
18	Luton	37	-31	26
19	Burnley	37	-36	24
20	Sheffield Utd	37	-66	16

*Everton deducted six points following a breach of the Premier League's Profitability and Sustainability Rules
*Nottingham Forest deducted four points following a breach of the Premier League's Profitability and Sustainability Rules

Bottom 3 teams are relegated

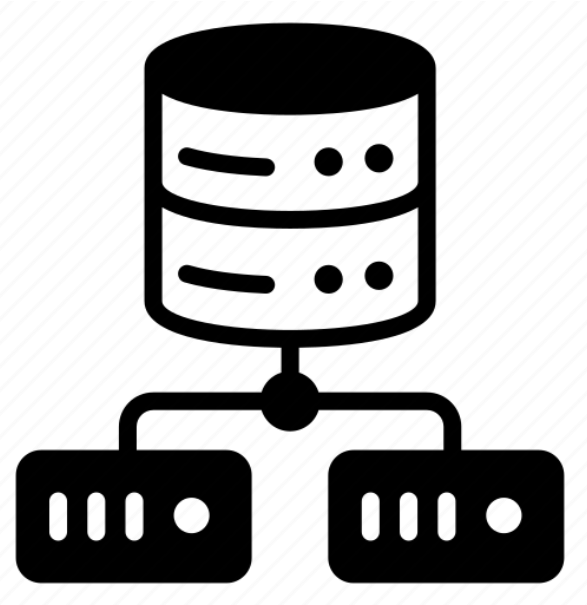
Players can’t be traded without their permission/consent

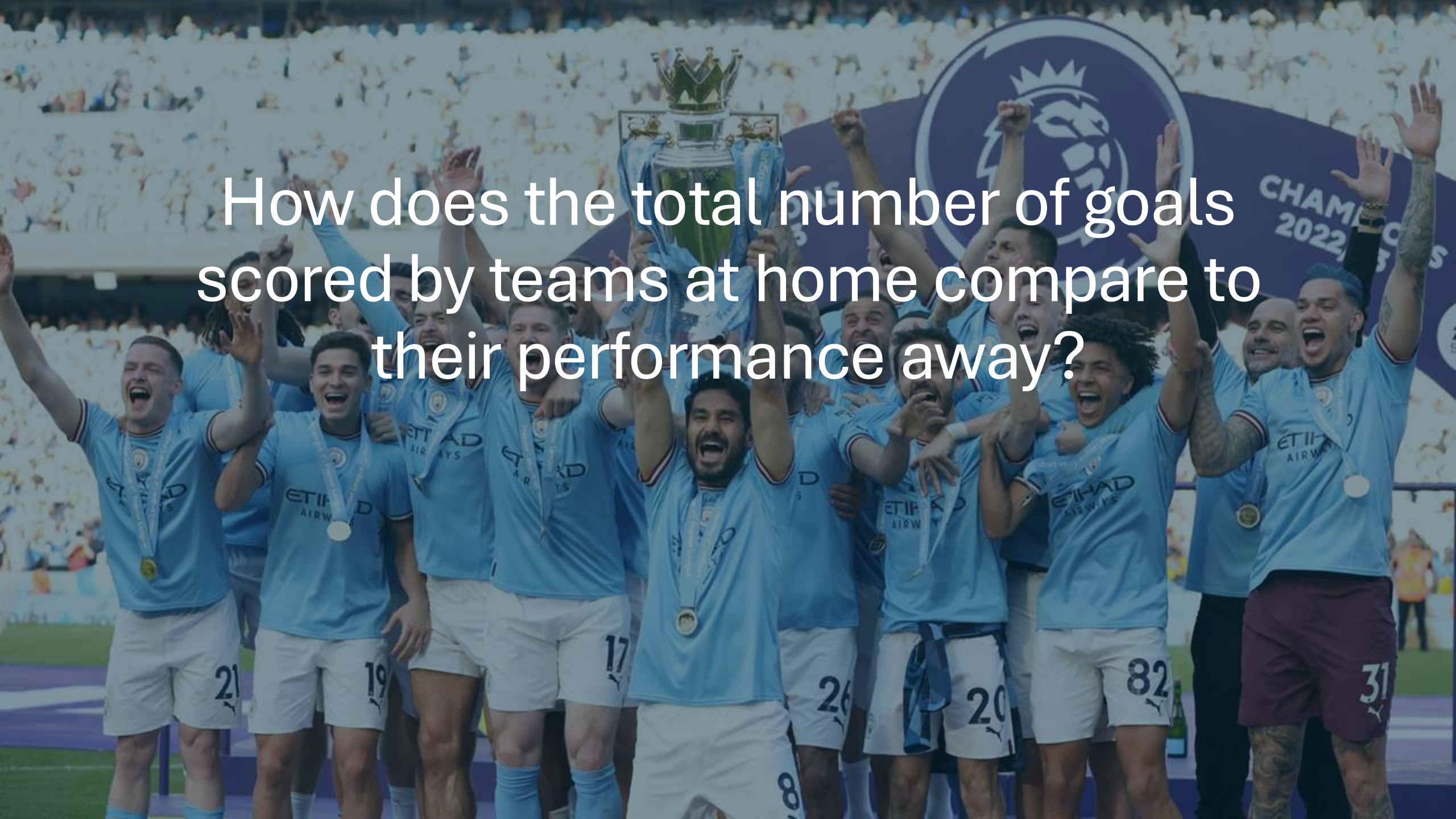
Dataset

- Seasons covered 93-94 to the current
- Each season's data is stored in a separate csv file

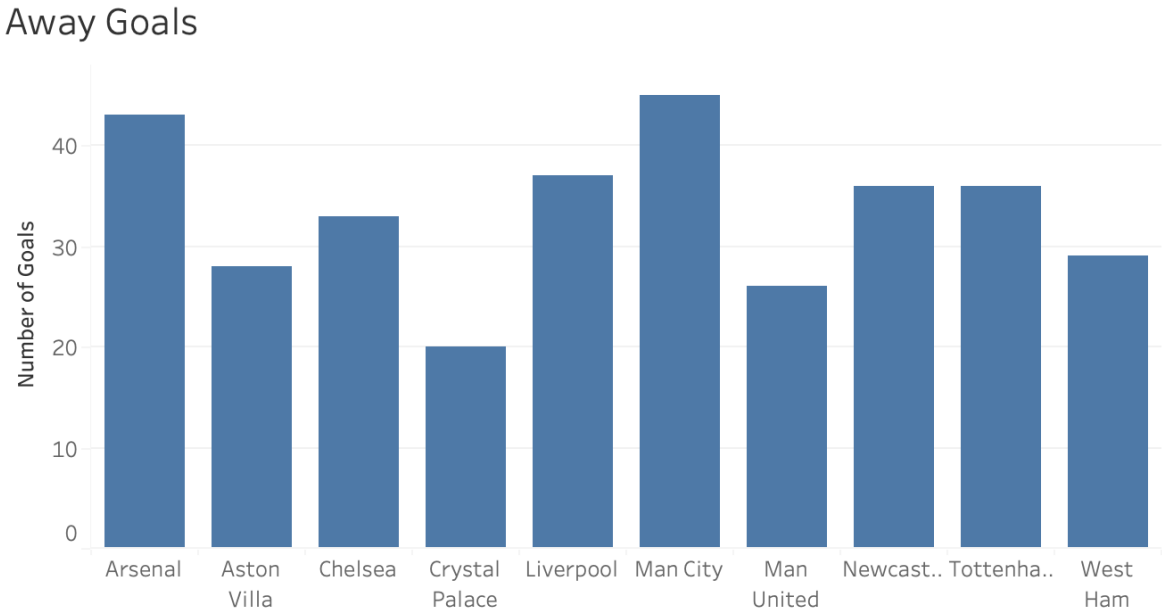
Match-level Statistics:

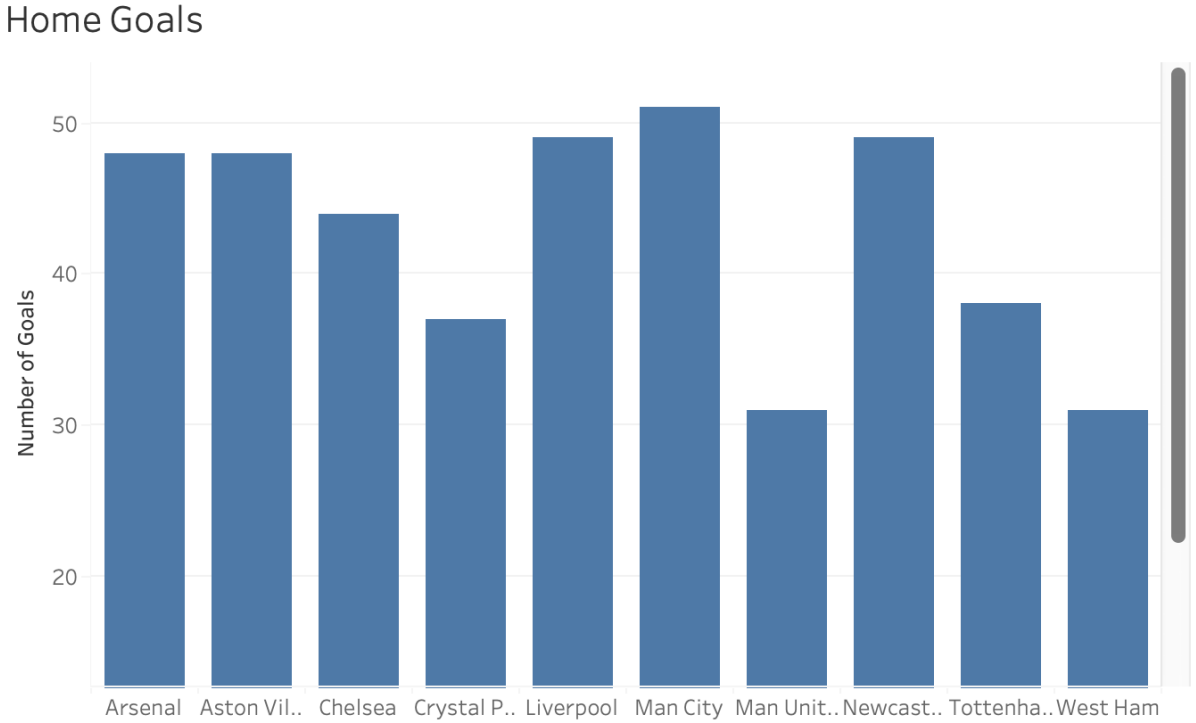
- Full-time & Half-time goals 🥅
- Attempted shots, Shots on target ⚽
- Yellow and Red Cards 🟨🟥
- Fouls, Corners 🚩




How does the total number of goals scored by teams at home compare to their performance away?

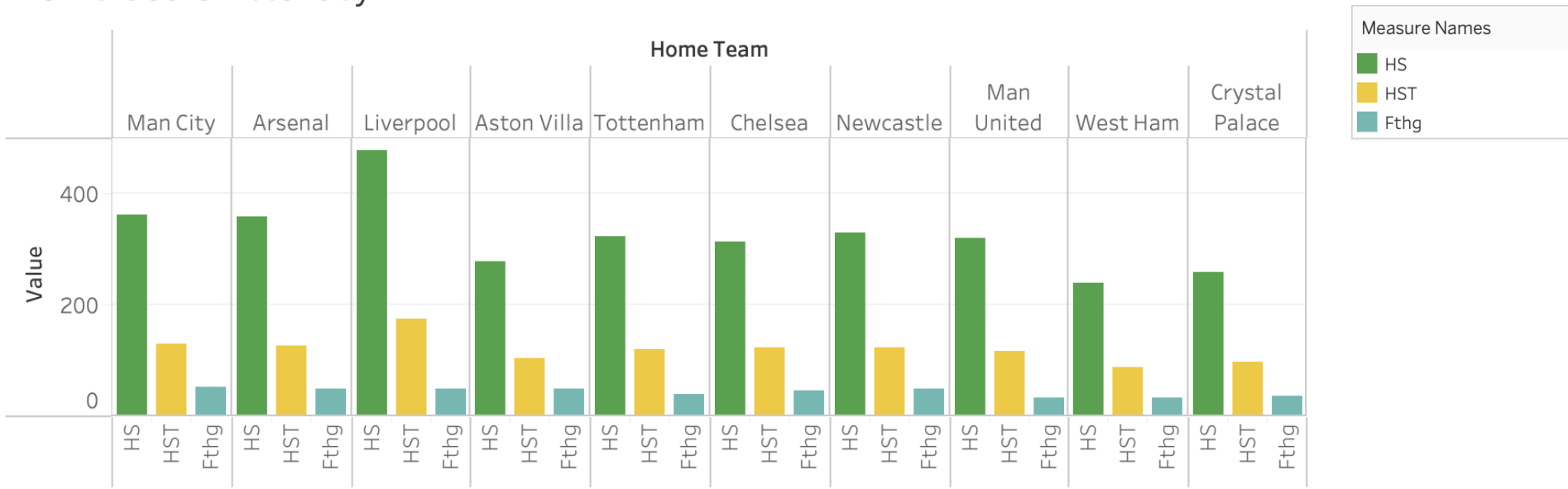
23-24 Season


Away Goals
Number of Goals
40
30
20
10
0
Arsenal
Aston Villa
Chelsea
Crystal Palace
Liverpool
Man City
Man United
Newcast..
Tottenha..
West Ham

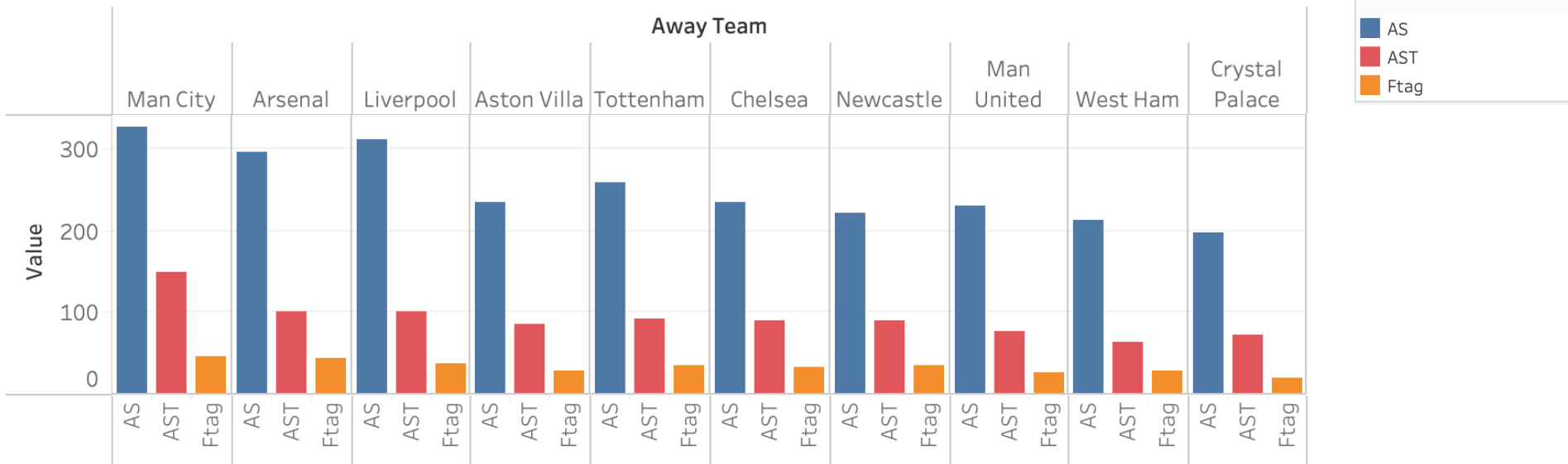

Home Goals
Number of Goals
50
40
30
20
Arsenal
Aston Vil..
Chelsea
Crystal P..
Liverpool
Man City
Man Unit..
Newcast..
Tottenha..
West Ham

Do teams in higher standings show better shooting performance?

Home Goals Accuracy

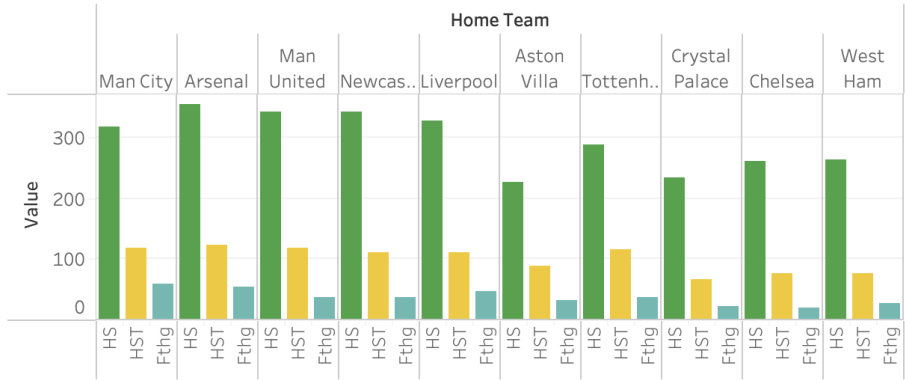


Away Goals Accuracy

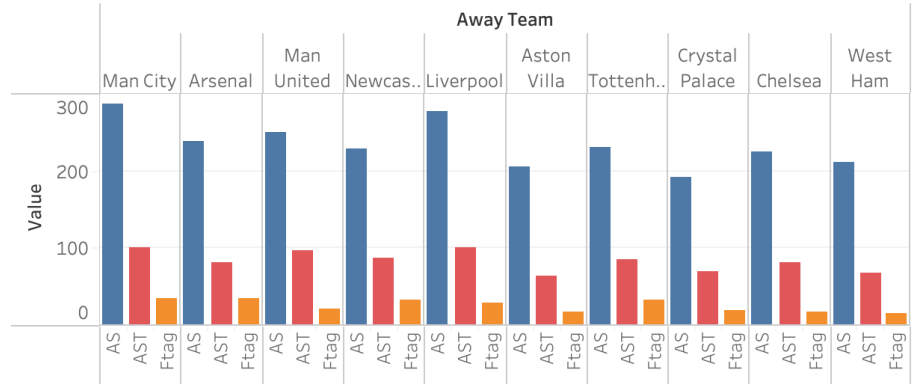


Team	Shot Accuracy(%)	Shot Efficiency(%)	Rank
Man City	40.289855	13.913043	4
Arsenal	34.550989	13.850837	5
Liverpool	35.063291	10.886076	13
Aston Villa	36.964981	14.785992	2
Tottenham	36.769759	12.714777	7
Chelsea	39.233577	14.051095	3
Newcastle	38.656987	15.426497	1
Man United	35.090909	10.363636	15
West Ham	33.333333	13.333333	6
Crystal Palace	37.061404	12.500000	8

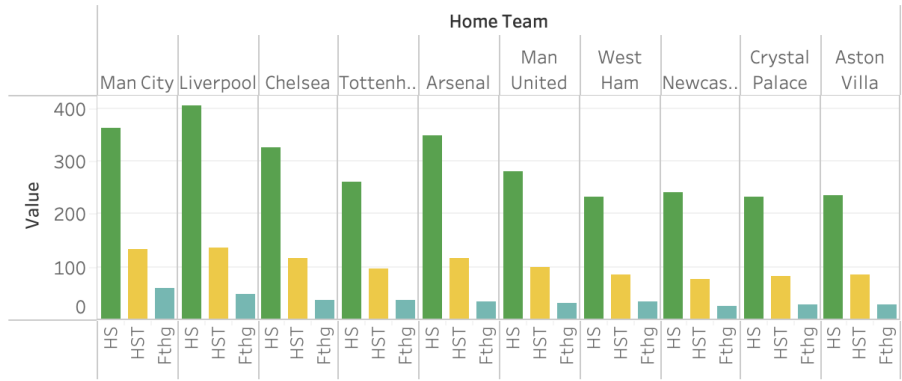
Home 22-23



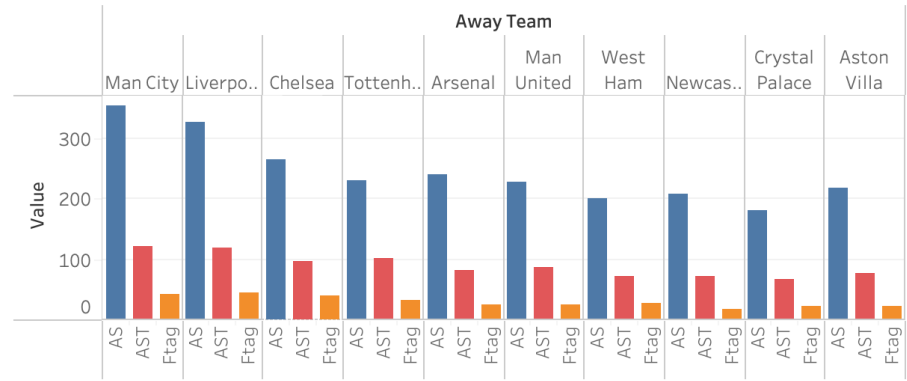
Away 22-23



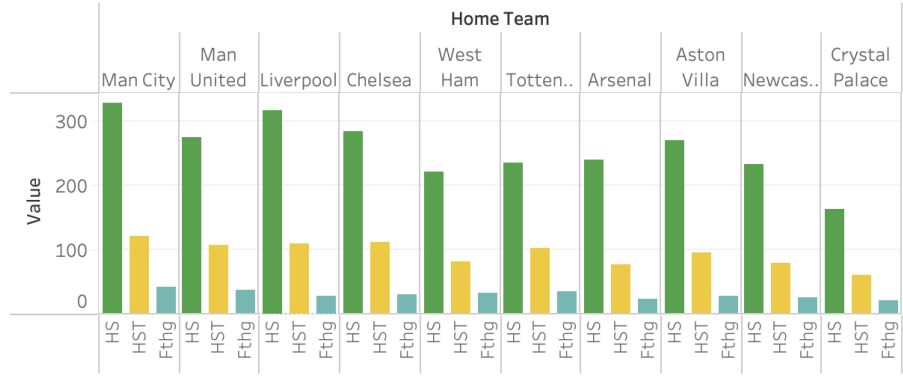
Home 21-22



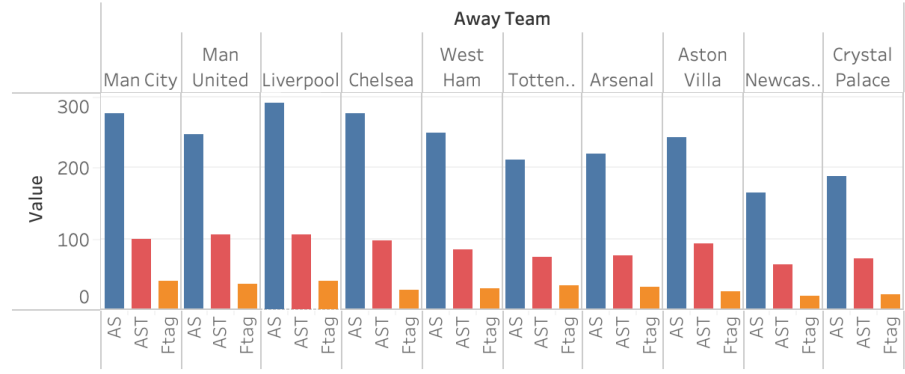
Away 21-22



Home 20-21

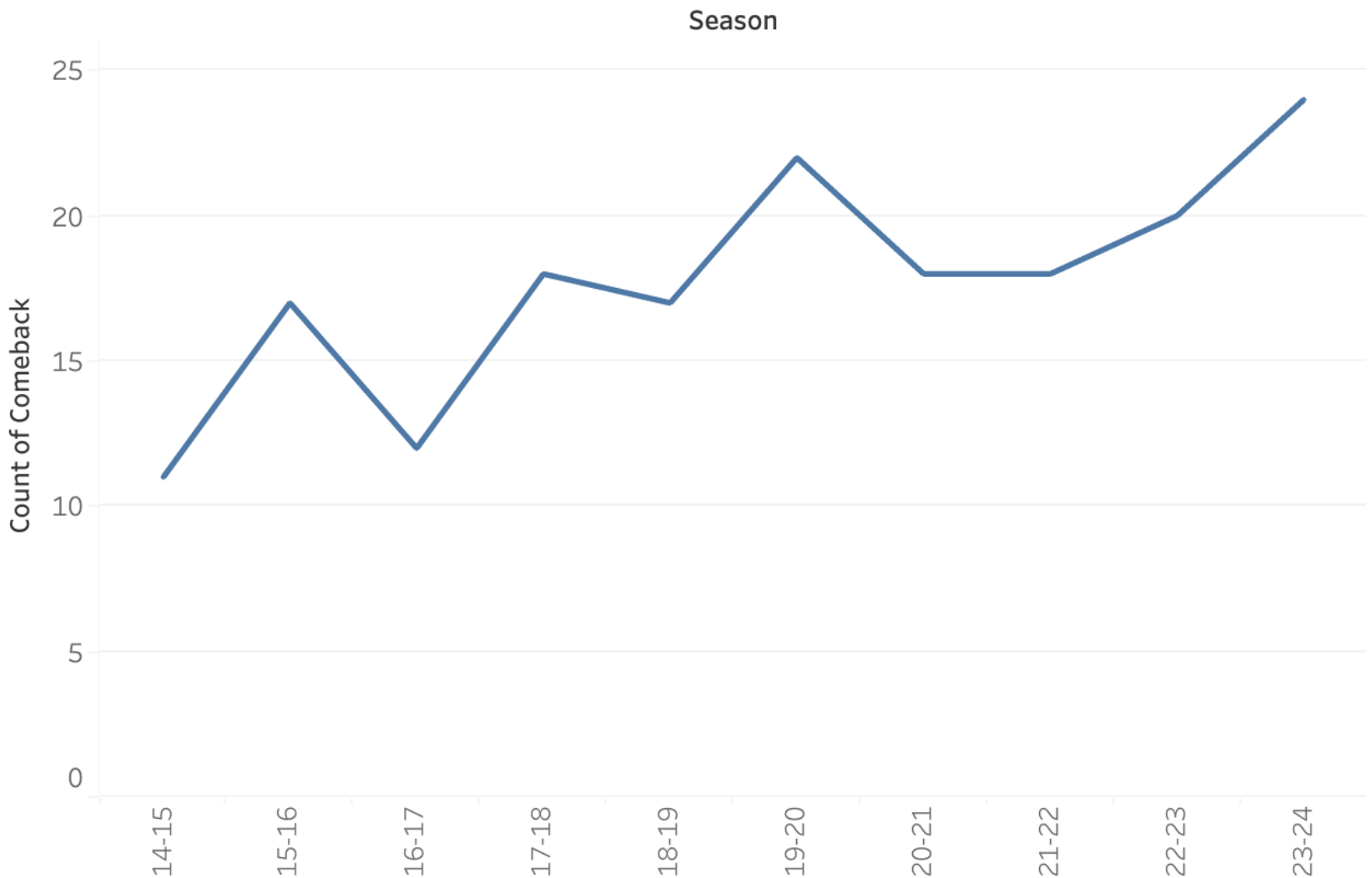


Away 20-21



How has the frequency of teams losing at halftime but coming back to win evolved over the last 10 seasons?

Comeback Trends



Regression

```
Please enter the match details:
Enter home team name:  Chelsea
Enter away team name:  Arsenal
The match will end in a draw with a score of 1 - 1
```

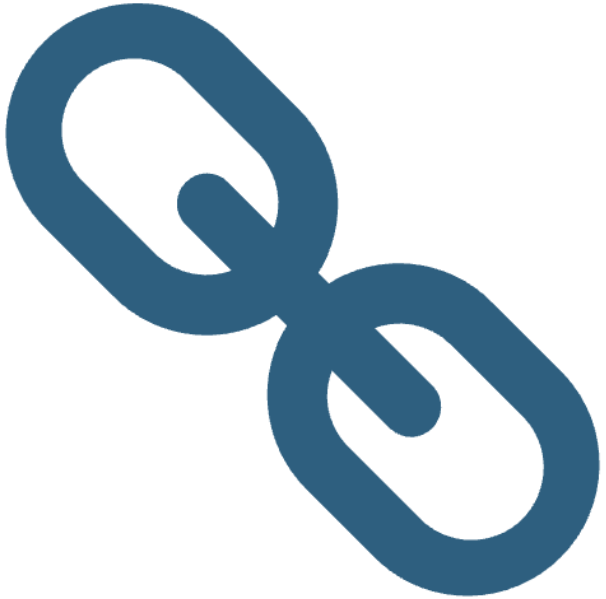
Evaluation Metrics

```
Home Team Evaluation:
RMSE: 0.8391132537575478
R²: 0.602427362592673 7

Away Team Evaluation:
RMSE: 0.7774088201739228
R²: 0.596688156240757 1
```

Cross Validation

```
Home Team Cross-Validation (R²):
Mean R²: 0.5816639965826267
Standard Deviation of R²: 0.014922501916495577

Away Team Cross-Validation (R²):
Mean R²: 0.5624417913003865
Standard Deviation of R²: 0.027739919359555814
```

References/Links



- Dataset:

https://datahub.io/core/english-premier-league

https://github.com/datasets/football-datasets/tree/main/datasets/premier-league

- Data Story:

https://ayushnelli03.github.io/DataStory/