English Premier League Data Analysis Report

Reported by: Nguyen Minh Duc

Strategic Insights from 32 Seasons of EPL Match Data (1993/94-2024/25)

Table of Content

Executive Summary

This report presents actionable insights from an exploratory data analysis (EDA) of English Premier League (EPL) match data to address my three key business questions including home advantage, team playing styles, and high-value match identification. The analysis leverages popular technologies of Python to clean and preprocess raw data, SQL to query and explore new insights with the preprocessed data and Tableau to clearly visualize and articulate the central tendency of match statistics and performance metrics to uncover patterns to guide strategic decisions in stadium investments, player recruitment and marketing optimization.

Introduction

The EPL (English Premier League) is a global sports phenomenon, driving revenue through ticket sales, broadcasting, and sponsorships. This analysis leverages a dataset with 51 columns including match details (goals, shots, fouls,...) and derived metrics (e.g., shooting accuracy, performance indices) to address business-critical questions. The insights support clubs and media companies in showcasing analytical rigor and business acumen.

Data Overview

The dataset (***epl_cleaned_partial_explored_data.csv***) covers 12 150 matches in total across 32 seasons of English Premier League from the 1993/94 (dated 14/08/1993) to 2024/25 (dated 20/05/2025) season. The dataset is thoroughly preprocessed with a notable 0% of missing goals data to ensure a sturdy analysis. SQL (PostgreSQL) is used to query and compute valuable metrics including goal differences, shot accuracy, and excitement scores to enable deeper lenses of the match and team dynamics.

A-Z metric	A-Z value
Unique Seasons	32
Date Range	1993-08-14 to 2025-05-20
Total Matches	12150
Missing Goals Data	0

Figure 1: Overview of number of seasons, start-to-end dates, total matches and missing goal values

Key Business Insights

Home Advantage Analysis

A critical question in this analysis is the extent to which home advantage influences outcomes in the English Premier League (EPL), as understanding this dynamic informs strategic decisions around stadium investments and ticket pricing. The analysis reveals that home teams secure victories in approximately 45.8% of matches on average across the entire seasons combined from the start date of the league up to date which significantly outpaces away teams winning 28.8% of matches on average across the entire seasons combined from the start date of the league up to date. Otherwise, draw accounts for 25.4% of cases. The figures below show the tendency of winning percentage between the competition from the start up to date and the 10 most recent seasons do not change considerably and remain in favor of the home games. It is clearly demonstrated in Figure 1 that the average percentage of home team victories virtually doubles of the away teams which shows the real motivation playing at the home games.

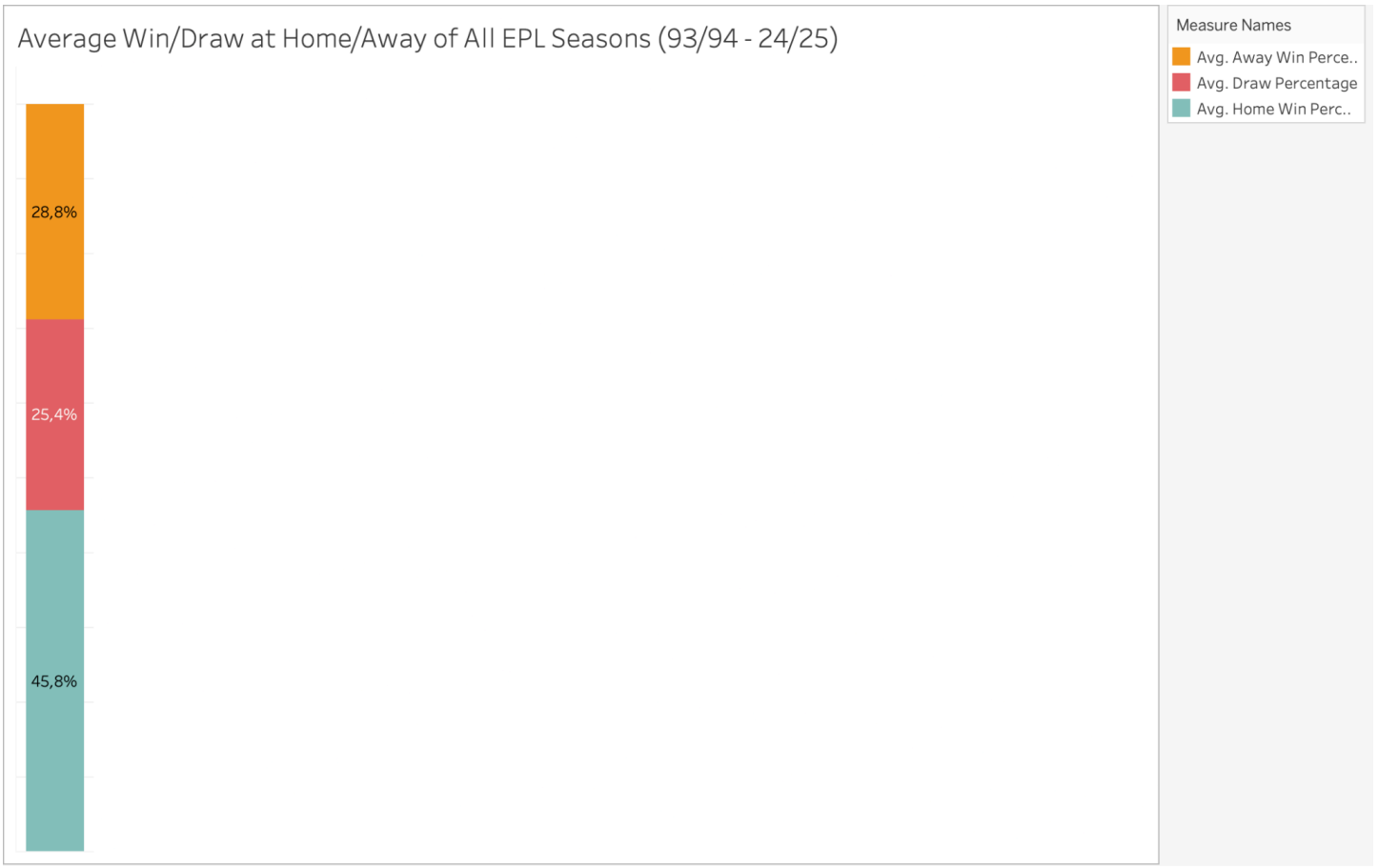

Figure 2: Home/away win, draw percentages throughout the EPL history

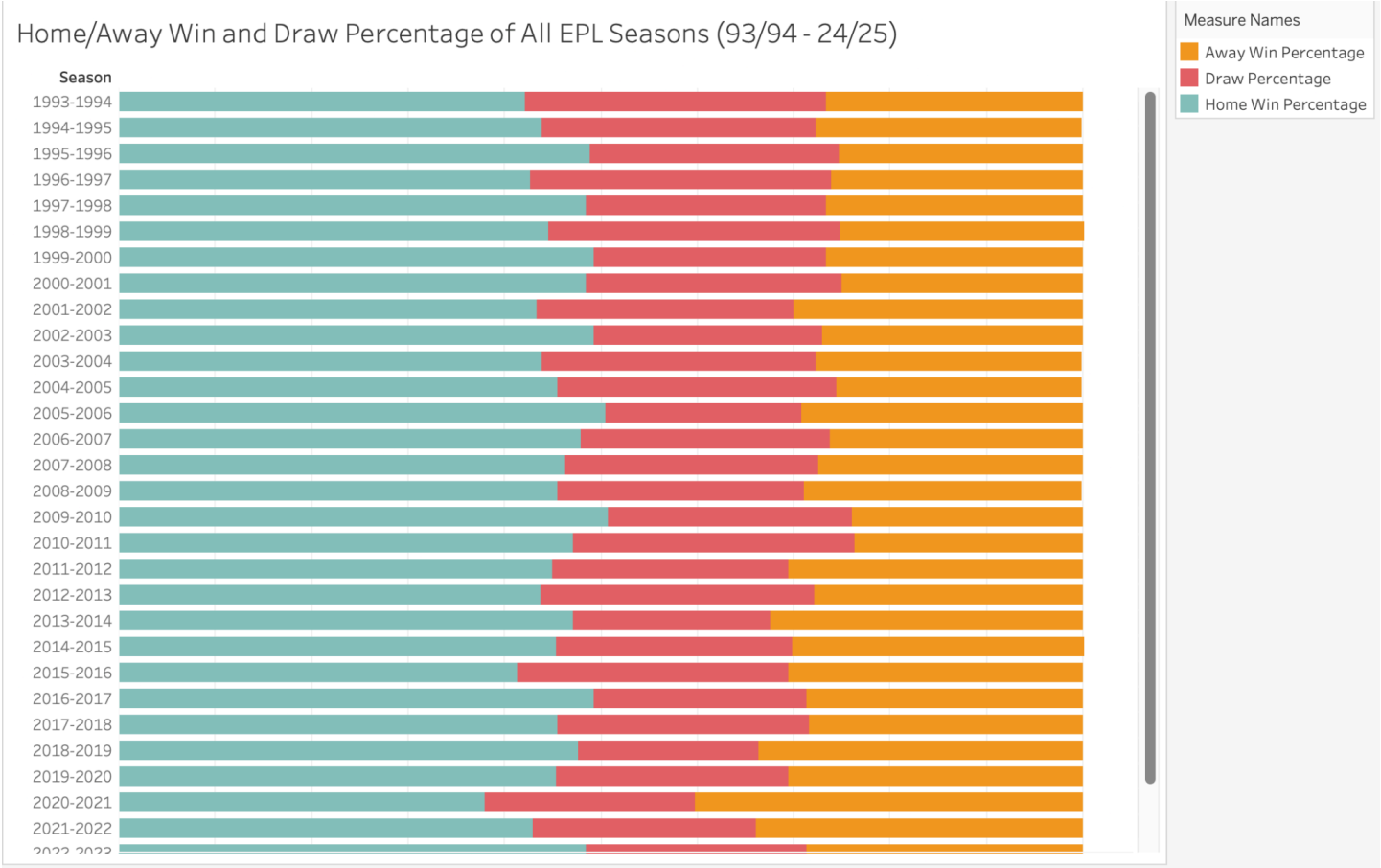

Figure 3: Home/away win, draw percentages by season

However, this home advantage is not evidenced to provide significant dominant scoring chances for home teams against the away teams across the history of this competition. In particular, home teams score an average of 1.533 goals per match compared to 1.165 goals for away teams which results in a notable 0.368 goal differential (approximately by 24%).

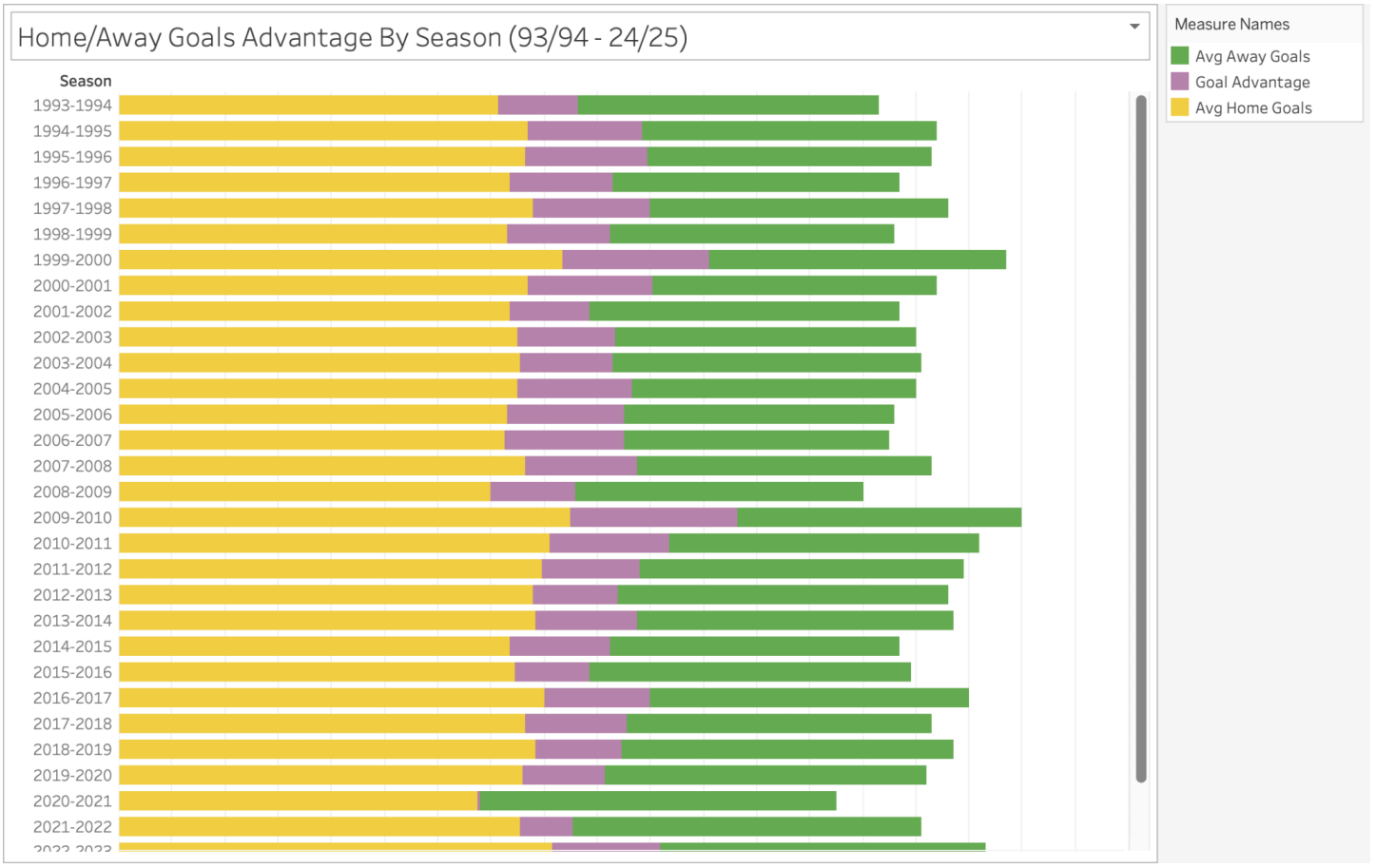

Figure 4: Home/away goals and their advantage percentages throughout the EPL history

Figure 5: Home/away goals and their advantage percentages by season

Based on the above analysis, it is clearly proven that there is no direct correlation home advantage of the winning rate of the home teams to the dominance in goals scored against away teams, but rather a very bonding connection comes from the warm-welcoming and genuine atmosphere and motivation by fans in the home games to the higher likelihood to win at the home games. Therefore, it is strongly advised that the football clubs should invest in the stadium with upgraded facilities and matchday experiences to enhance fan engagement that can amplify and strengthen the home advantage. Also from the business point of view, adopting dynamic ticket pricing strategies for high-demand home matches can optimize revenue by aligning prices with the increased likelihood of home victories, thereby enhancing the financial return on matchday operations.

Team Playing Style Effectiveness

Understanding which playing styles drive superior results in the EPL is essential for shaping player recruitment strategies and refining tactical approaches. The analysis indicates that teams employing a high-attack style, characterized by averaging more than 1.5 goals and 12 shots per game, achieve the highest seasonal points total, averaging 73.8 points. Teams with a strong defensive approach, conceding an average of 1.1 goals per game, follow closely with 59.7 points, while those with efficient attacking styles, marked by shot accuracy exceeding 40%, earn 42.6 points on average. Teams with a balanced approach only get 41.1 points on average followed by teams with efficient attacking play styles.

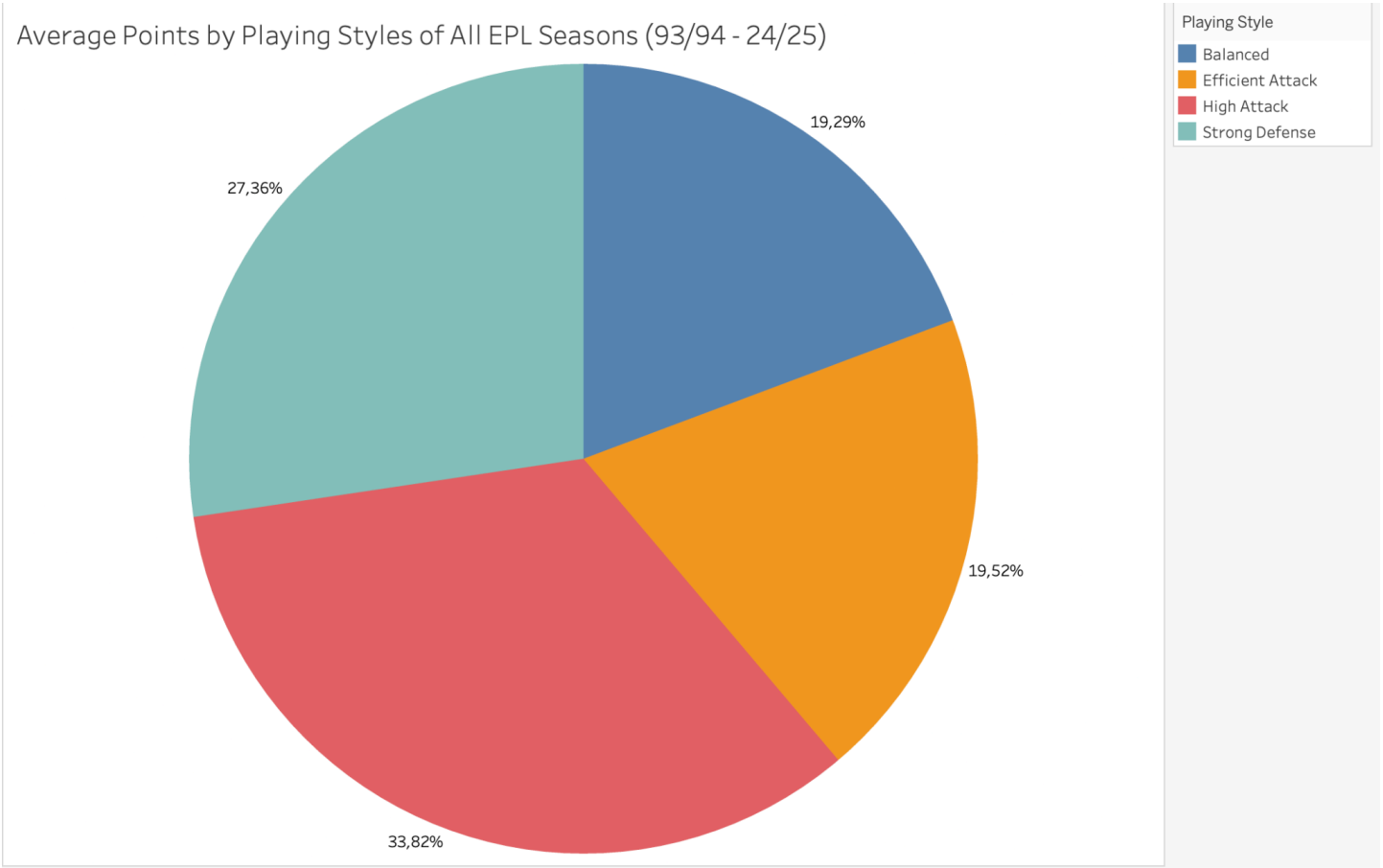

Figure 6: Average points by playing styles throughout the EPL history

It is surprising that teams employing the balanced and followed by efficient attack playing styles are among the most vulnerable in terms of the average goals conceded per game of 1.72 and 1.5 respectively. On the other hand, teams that

pursue high and efficient attacking football tend to create the most goals. It is also noticeable that teams with strong defensive foundations create more goals, which is more offensively impactful, than teams that pursue the playing style with the balance in attack and defense. It is also interesting to see from Figure 6 that teams adopting efficient attacking football tend to be prone to conceding more goals than teams adopting high attacking football. It is understandable that the more offensive the team is, the more dangerous threats the team will deliver. It creates a tension on the defensive duties on the opponent's side so that when the highly offensive team creates more goals, it puts a severe pressure on defenders to prevent the oppressive attack, making it difficult to recover mentally and physically from the amount of goals conceded.

From the observation, it is understandable that football teams nowaday gradually tend to adopt either a strong defensive playing style or highly attacking playing style to optimize the maximum positive output results, especially in a high-intensity football league like EPL.

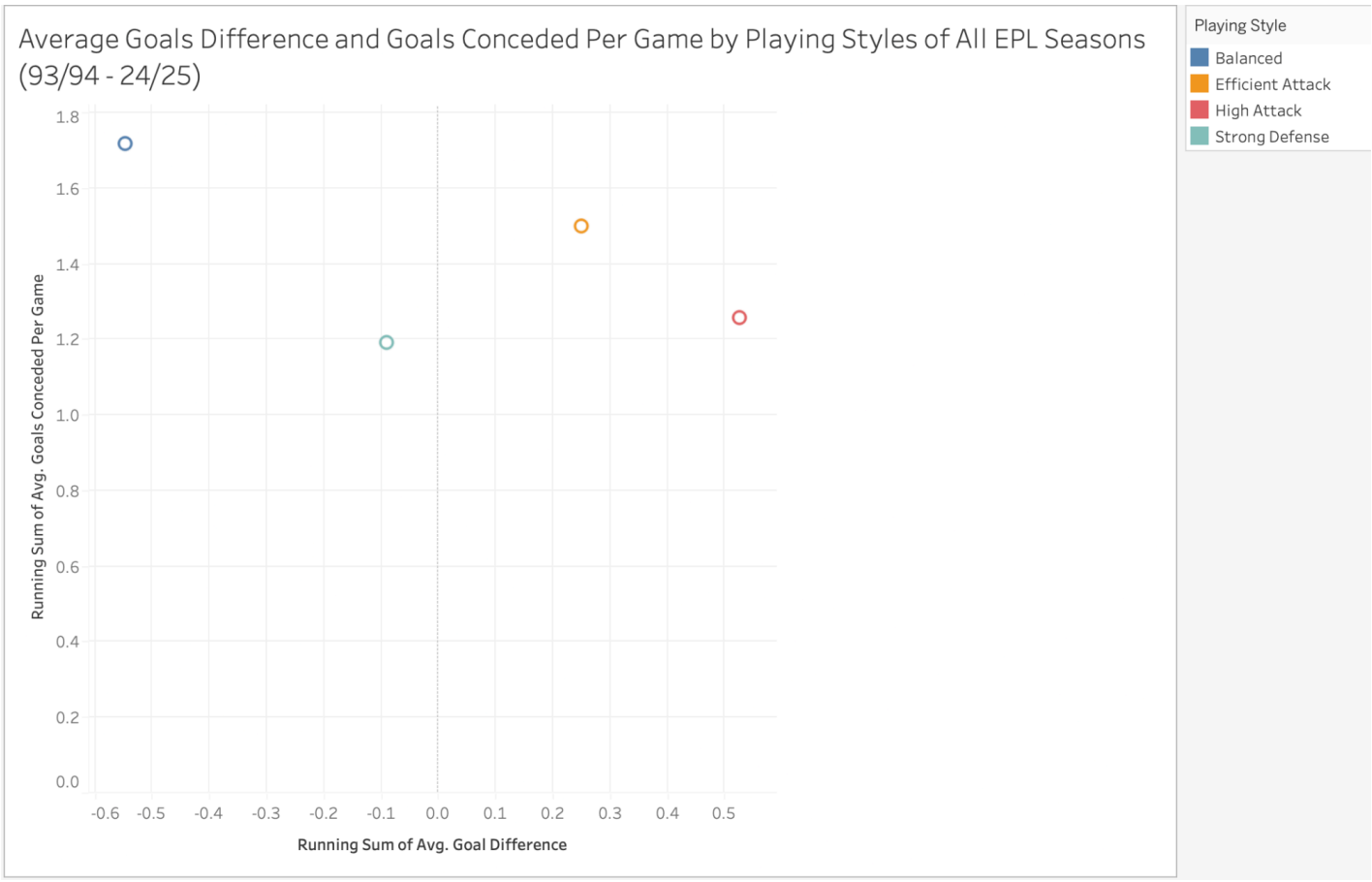

Figure 7: Average goals difference and goals conceded per game by playing styles throughout the EPL history

The analysis proves even more accurately, especially in this season (2024/25) that teams pursuing either highly attacking or strong defensive playing styles tend to achieve their own targets in the start of the season, however things do not go well with teams adopting efficient attacking and especially balanced profiles. The table below shows the final results at the end of the season:

	Club	MP	W	D	L	GF	GA	GD	Pts	Last 5
1	Liverpool	38	25	9	4	86	41	45	**84**	W L D L D
2	Arsenal	38	20	14	4	69	34	35	**74**	D L D W W
3	Man City	38	21	8	9	72	44	28	**71**	W W D W W
4	Chelsea	38	20	9	9	64	43	21	**69**	W W L W W
5	Newcastle	38	20	6	12	68	47	21	**66**	W D W L L
6	Aston Villa	38	19	9	10	58	51	7	**66**	L W W W L
7	Nottm Forest	38	19	8	11	58	46	12	**65**	L D D W L
8	Brighton	38	16	13	9	66	59	7	**61**	W D W W W
9	Bournemouth	38	15	11	12	58	46	12	**56**	D W L L W
10	Brentford	38	16	8	14	66	57	9	**56**	W W W L D
11	Fulham	38	15	9	14	54	54	0	**54**	W L L W L
12	Crystal Palace	38	13	14	11	51	51	0	**53**	D D W W D
13	Everton	38	11	15	12	42	44	-2	**48**	L D W W W
14	West Ham	38	11	10	17	46	62	-16	**43**	L D W L W
15	Man United	38	11	9	18	44	54	-10	**42**	D L L L W
16	Wolves	38	12	6	20	54	69	-15	**42**	W L L L D
17	Tottenham	38	11	5	22	64	65	-1	**38**	L D L L L
18	Leicester City	38	6	7	25	33	80	-47	**25**	L W D W L
19	Ipswich Town	38	4	10	24	36	82	-46	**22**	L D L L L
20	Southampton	38	2	6	30	26	86	-60	**12**	L L D L L

Figure 8: The complete EPL standing table this season

It is important to note that teams with highly attacking playing profiles tend to be either big teams like Liverpool, Manchester City, Arsenal, Chelsea chasing for winning premiership titles or mid-table teams fighting for a top 5 position to be qualified for the Champions League tournament for the upcoming season, like Brighton, Nottingham Forest, Bournemouth. There is only one team with a strong defensive profile this season, Everton, successfully escaping from the relegation zone to a lower-level league. It is also true that all of the teams with balanced playing styles struggle to achieve any particular expected targets at the start of the season. Teams with balanced profiles, notably, Leicester City, Ipswich Town and Southampton will be relegated at the end of the season while the only efficient attacking team, Brentford, hardly achieve any target this season at the 10th position on the league table.

However, there is only one outlier from this analysis and that is Tottenham. Although Tottenham is among top teams with a highly attacking profile, they not only struggle to compete either for the title winning or top 5 position but also they nearly drop their position to the relegation zone.

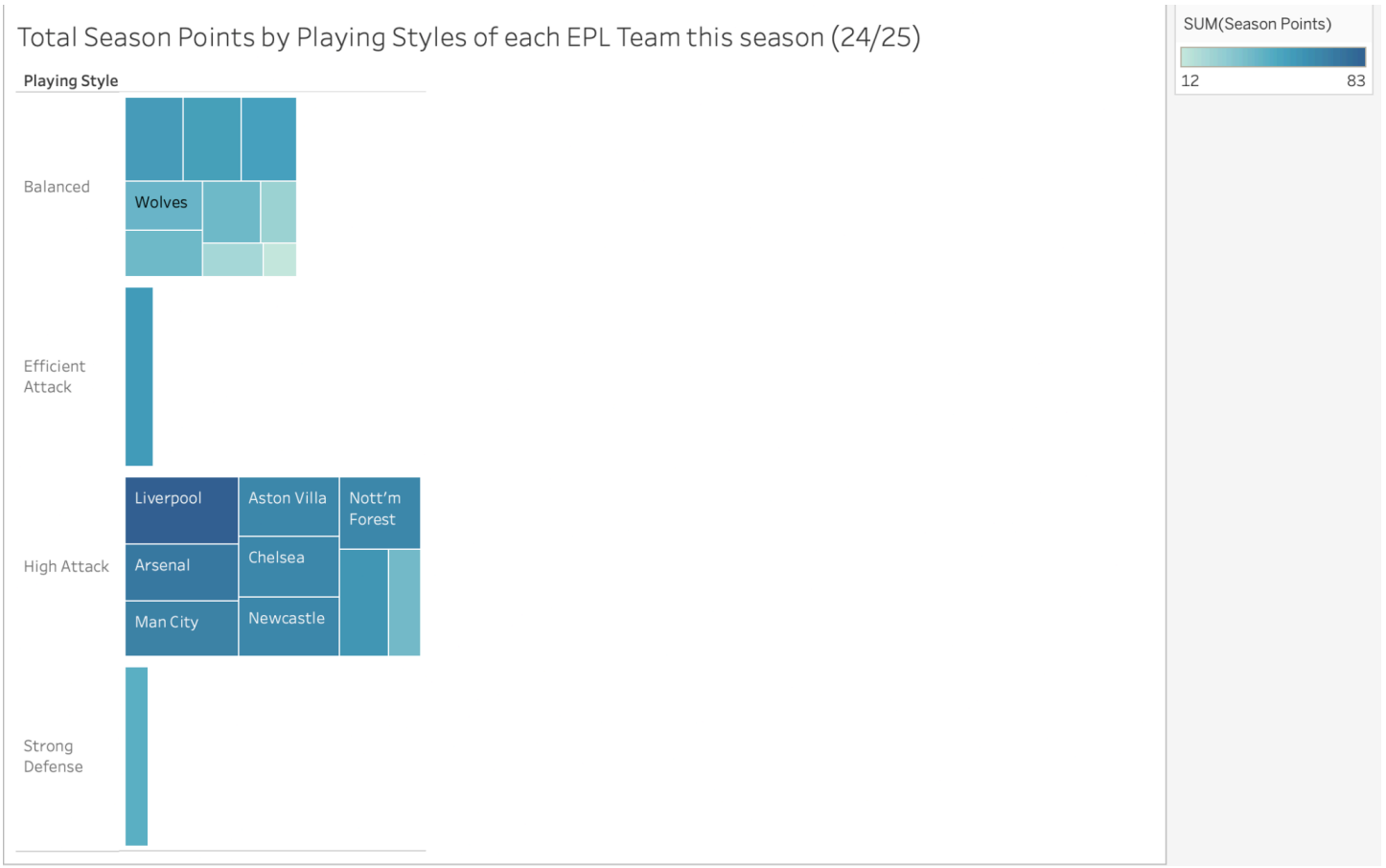

Figure 9: Total season points based on playing styles of each EPL team this season

These insights suggest that clubs should prioritize recruiting versatile forwards who excel in highly offensive attacking football to reinforce attacking flexibilities while also considering steady defenders to strengthen teams favoring a defensive approach. Tactically speaking, coaches should aim to incorporate aggressive attacking play with strong defensive resilience to achieve consistent performance across seasons in order to enhance overall team competitiveness.

High-Value Match Identification

Determining what constitutes a commercially valuable match is pivotal for optimizing television scheduling, ticket pricing and marketing efforts in the EPL.

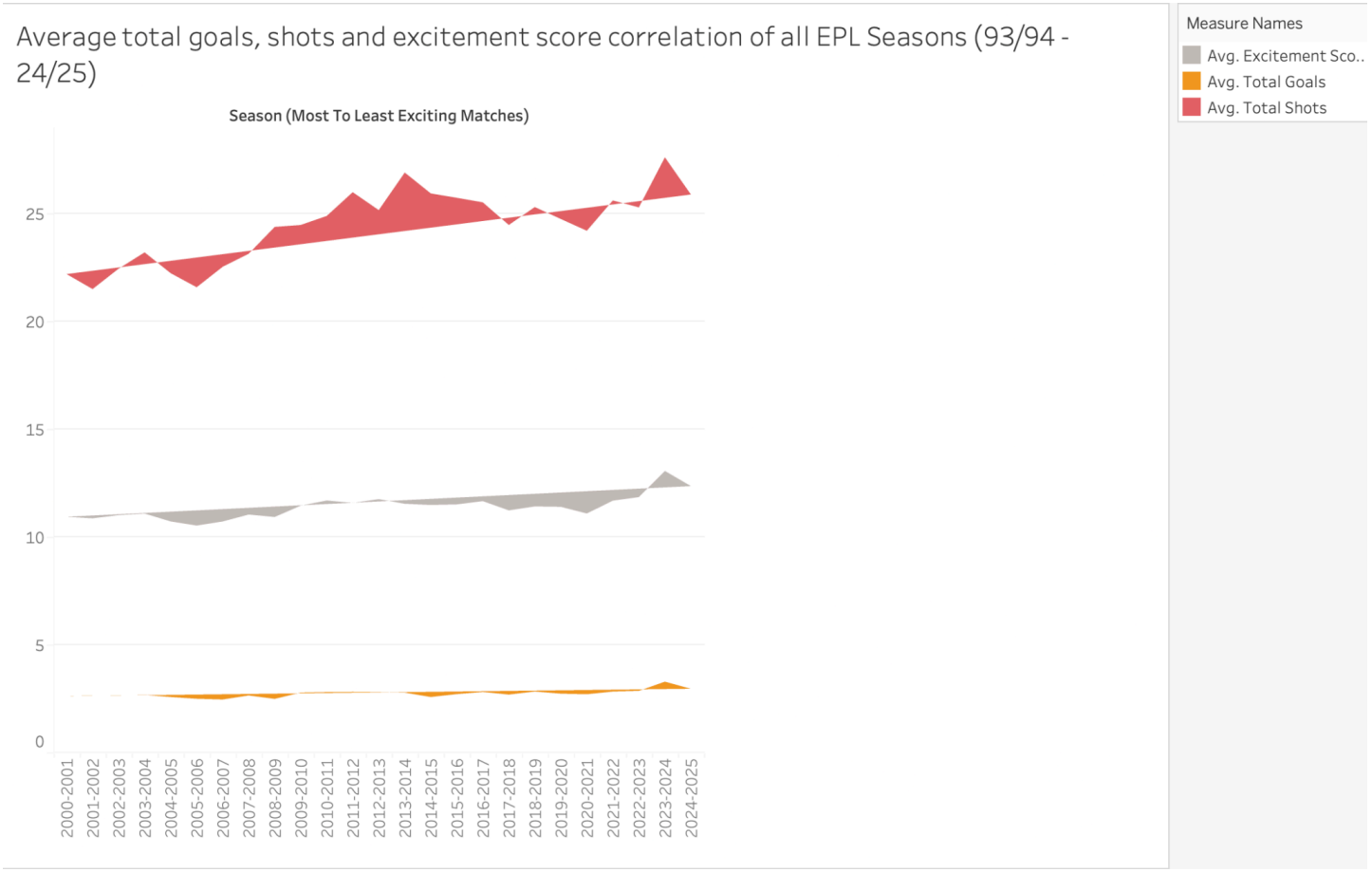

Figure 10: Correlation of the average total goals, shots and excitement throughout EPL history

As can be seen from the above figure, the more shots taken do not automatically convert effectively into more goals scored. The highlight was in the season 2013/14, there was an average of 26.88 total shots taken and only 2.77 total goals on average. That is approximately 10.3% of goal conversion rate which is too low for the effective goal scoring outputs. Also, it can be seen a mutual connection between the amount of goals scored and the excitement scores by the fans. It is proven statistically that when the team scores more goals, the excitement scores will be higher which leads to more match satisfaction. Here is the formula used to calculate the excitement score:

$$\textbf{Excitement} = (\textbf{Goals} \times 2) + (\textbf{Shots} \times 0.1) + (\textbf{Close\_Game\_Bonus}) + (\textbf{Cards} \times 0.5)$$

It can be seen that there is a strong need for clinical and technical offensive players who can maximize the goal conversion rate in order to drive engagement from fans. The below figure even further finds the right time to take an action in the market in order to maximize the business profits.

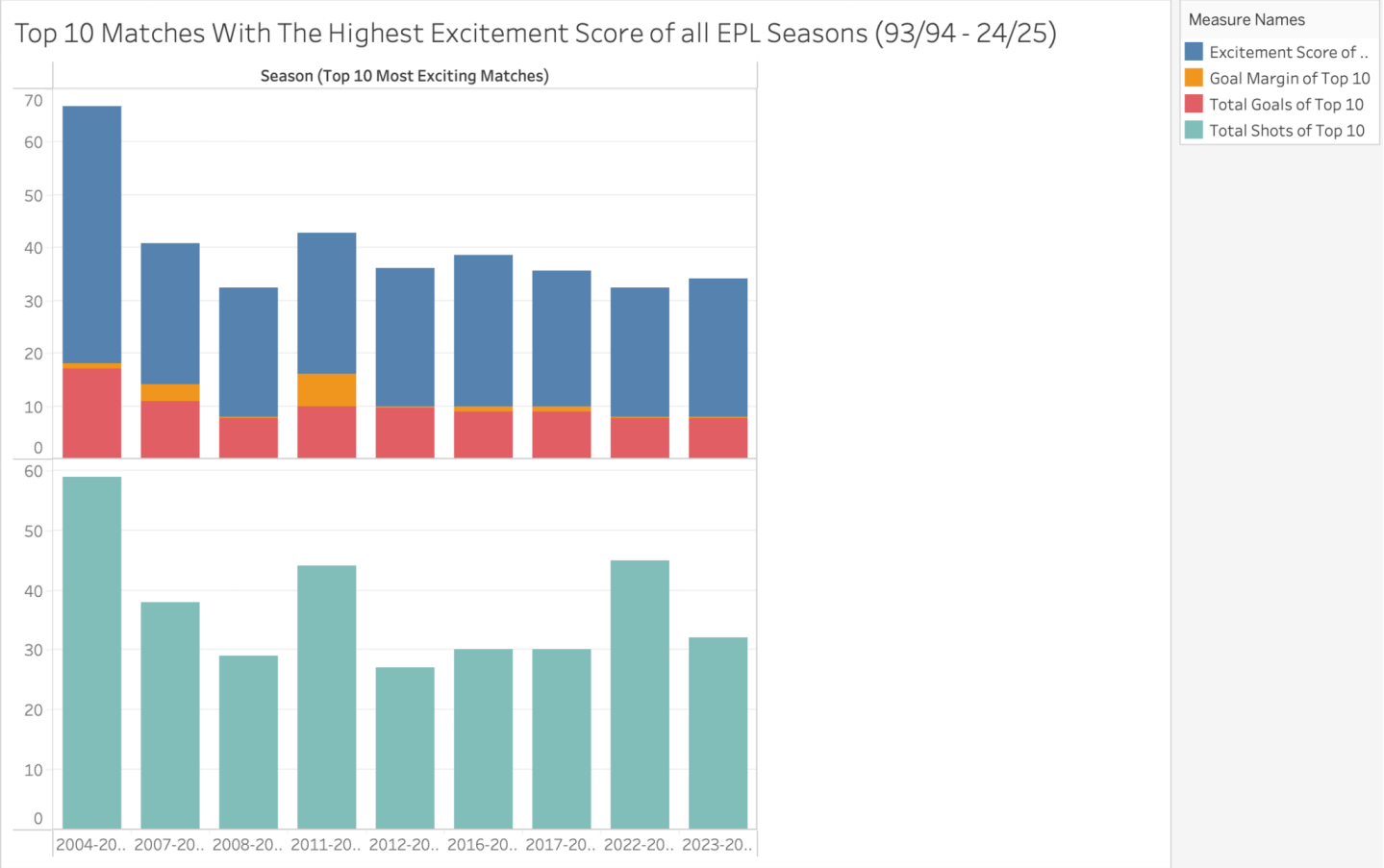

Figure 11: Top 10 matches with the highest excitement score in the EPL history

From this figure, out of all generic metrics, for instance total goals or total shots that have been analyzed before, it is foreseeable that the top 10 most exciting matches typically involve modern-day football which started and peaked from the 2004/25 season. However, while the rest of matches shown above are still in the top 10, the excitement scores started to considerably fall down from the 2007/08 season then keep fluctuating until now and haven't recovered and reached a higher climax from the peak in the 2004/05 season. It leads to a question: Is it the right time to invest in players with the right profiles to reach a new height in football excitement as well as a long-term business prospect.

Given the observed decline in excitement scores since the 2004/05 peak, as shown in Figure 10, clubs and stakeholders should consider this an opportunity to invest in talent acquisition strategies that target players capable of delivering the clinical finishing needed to restore and surpass historical excitement levels. By aligning recruitment with data-driven insights into match excitement dynamics, clubs can position themselves to create a new era of compelling football which drives both fan satisfaction and financial returns.

Conclusion

This analysis of 32 EPL seasons reveals three critical insights with clear strategic implications. Home advantage remains powerful, with home teams winning 45.8% versus 28.8% for away teams which is motivated more by atmosphere rather than scoring dominance. Clubs should prioritize stadium upgrades and fan experience investments over marginal player signings while implementing dynamic ticket pricing strategies for high-demand home matches to maximize revenue potential.

The data exposes a tactical truth: extremes consistently outperform balance. High-attack teams average 73.8 points while balanced teams manage just 41.1 points. This season confirmed the trend as three balanced teams face relegation while attack-focused Brighton and Nottingham Forest secured European positions. Teams should abandon balanced approaches to moderate levels of attack and defense. Instead, it should be encouraged to incorporate either high-attack systems with versatile forwards and strong defensive foundations built around clinical counter-attacking capabilities to maximize the competencies.

Most significantly, match excitement has declined since 2004/05 despite increased shot volume, revealing a clinical finishing crisis with conversion rates hovering around 10%. Clubs should target recruitment toward technically superior finishers who can convert chances at higher rates, creating both competitive advantage and enhanced commercial value through more engaging matches that drive fan satisfaction and broadcasting appeal.

That being said, the future of EPL belongs to data-driven clubs that embrace tactical polarization, maximize atmospheric advantages and solve the efficiency puzzle that transforms shots into goals and excitement into sustainable revenue streams.