Arsenal Season Stats, EPL '24-'25

Contents

Overview

A(r)senal Analysis is an analytical exploration of Arsenal's performance in the **2024-2025** season of the **English Premier League**. The team was chosen because of their ability to outperform league leaders and top contenders in the UEFA Champion's League (particularly their 3-0 win over Real Madrid) yet showing recurring inconsistencies in front of goal.

Throughout the season, fans and pundits alike heralded it as *"Arsenal's year."* With Manchester City's dominance waning, the path to the title seemed open. Yet by May, Liverpool sat atop the table while Arsenal once again finished second. The central question, then, is this: how can a squad strong enough to reach the Champions League semi-finals and challenge England's best still struggle to convert that dominance into goals?

Research Questions

- For a team that's beaten highly-rated teams such as Real Madrid and Manchester City, why is Arsenal's domestic performance inconsistent?
- How well do xG and xGA predict actual team performance?

Methodology

I scraped aggregated match data from FBRef and used Kaggle datasets for player-level statistics and the official fixture list. The datasets were stored in a Docker-based MariaDB instance in my homelab for reproducible querying. Using VSCode's SQL extensions, I cleaned and normalized data directly in SQL, then connected RStudio to the database for statistical exploration and visualization. With R, I performed statistical analyses to understand Arsenal's team performance and studied the correlation of metrics like xG and xGA alongside team performance to understand how well these metrics predict a team's performance.

*Disclaimer: Generative AI was used in this project to help guide the writing of code chunks and SQL queries. Generative AI was **not** used to write analyses or full code chunks.*

Data Sources

- EPL Player Match Data: https://www.kaggle.com/datasets/aesika/english-premier-league-player-stats-2425 (1 row per player, aggregated statistics for the 2024-2025 season)
- EPL Fixture Data: https://www.kaggle.com/datasets/secretglory/epl-fixtures-list-2024-2025 (Official match schedule as published)
- EPL Match Data: Scraped from https://fbref.com/en/comps/9/2024-2025/schedule/2024-2025-Premier-League-Scores-and-Fixtures (Aggregated match statistics, per team, for the entire season)
- EPL Season Table, 2024-2025: https://fbref.com/en/comps/9/2024-2025/2024-2025-Premier-League-Stats (League table from 2024 for context review)
- Arsenal Squad Data from 2023-2024: https://fbref.com/en/squads/18bb7c10/2024-2025/matchlogs/c9/schedule/Arsenal-Scores-and-Fixtures-Premier-League (Previous season summary stats for context review)
- EPL Season Table, 2023-2024: https://fbref.com/en/comps/9/2023-2024/2023-2024-Premier-League-Stats (League table from 2023 for context review)

Definition of Terms

- **GF**: Goals for
- **GA**: Goals against
- **GD**: Goal difference, (GF - GA)
- **xG**: Expected goals; the number of goals a team/player is expected to score in a match. **Higher is better**
- **xGA**: Expected goals against; the number of goals a team is expected to concede in a match. **Lower is better**
- **xGD**: Expected goal difference; The difference between xG and xGA. **Higher is better**

Performance in Perspective

Arsenal FC is a Premier League club with a storied history. Over the years, they've won 13 English top-flight titles, 14 FA Cups, and 17 FA Community Shields, among other honors. The club also produced *The Invincibles*, the legendary 2003–2004 squad that went an entire Premier League season unbeaten.

Unfortunately, that level of dominance was not to last. Financial constraints, largely attributed to the construction of the Emirates Stadium, forced a series of challenges that saw the club struggle for more than a decade before showing real signs of resurgence. Enter Mikel Arteta. The former Arsenal captain returned to the club mid-season in 2019 to take up the managerial role. Despite securing an FA Cup title later that season, it would take several years of rebuilding before consistent, sustainable improvement began to show.

Uphill Battles

The renewed optimism surrounding Arsenal at the start of the 2024–2025 season stemmed from their remarkable form in 2023–2024. There was room to grow, but this was the most competitive the team had been in a long while. Finishing 2nd, the team stood out for a couple of reasons, but particularly for their expected metrics at the end of the season. In this section, we use aggregated 2023-2024 season data to test how closely related expected metrics like xG and xGA are to actual team performance across the Premier League.

To start, we take a look at the defensive structure Arsenal had developed. For a team long associated with defensive lapses, Arsenal's 2023 season marked a significant turnaround, recording an xGA of **27.9**, the lowest in the league. This led to the team conceding only **29** goals throughout the season, with 2023 champions Manchester City conceding **34** in that same season.

Table 1: figure 1, Premier League Table, 2023

rk	team_name	gf	ga	gd	pts	xg	xga
1	Manchester City	96	34	62	91	80.5	35.6
2	Arsenal	91	29	62	89	76.1	27.9
3	Liverpool	86	41	45	82	87.8	45.7
4	Aston Villa	76	61	15	68	63.3	59.9
5	Tottenham	74	61	13	66	68.2	63.4
6	Chelsea	77	63	14	63	74.5	58.1
7	Newcastle United	85	62	23	60	76.0	61.4
8	Manchester United	57	58	-1	60	56.5	68.9
9	West Ham United	60	74	-14	52	52.3	71.1
10	Crystal Palace	57	58	-1	49	48.6	52.0
11	Brighton	55	62	-7	48	56.8	55.4
12	Bournemouth	54	67	-13	48	55.9	58.1
13	Fulham	55	61	-6	47	50.8	62.9
14	Wolves	50	65	-15	46	46.7	67.7
15	Everton	40	51	-11	40	54.0	55.2
16	Brentford	56	65	-9	39	58.2	56.0
17	Nottingham Forest	49	67	-18	32	49.9	53.3
18	Luton Town	52	85	-33	26	42.4	78.0
19	Burnley	41	78	-37	24	40.6	70.4
20	Sheffield United	35	104	-69	16	38.3	76.6

Liverpool, who followed with an xGA of **45.7**, conceded **41** goals which underscored how much Arsenal had closed the defensive gap with the very best. Arsenal's attack, meanwhile, showed potential despite the lack of a traditional number 9 player.

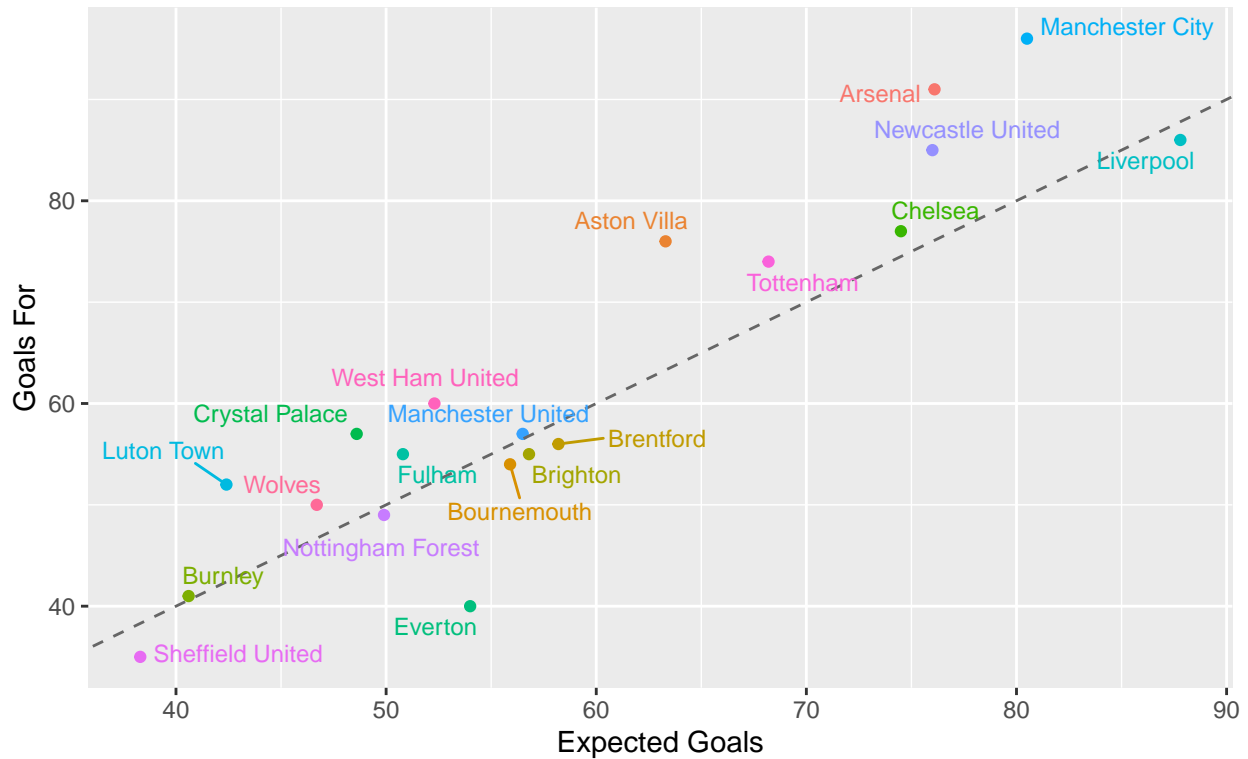## Goals For v.s. Expected Goals, EPL 2023

Figure 2: Premier League 2023 GF/xG Scatterplot

In the chart above, teams above the line are those that over performed, or those that scored more goals than expected (GF - xG). We can see that Arsenal counted themselves among the top when it came to expected goals.

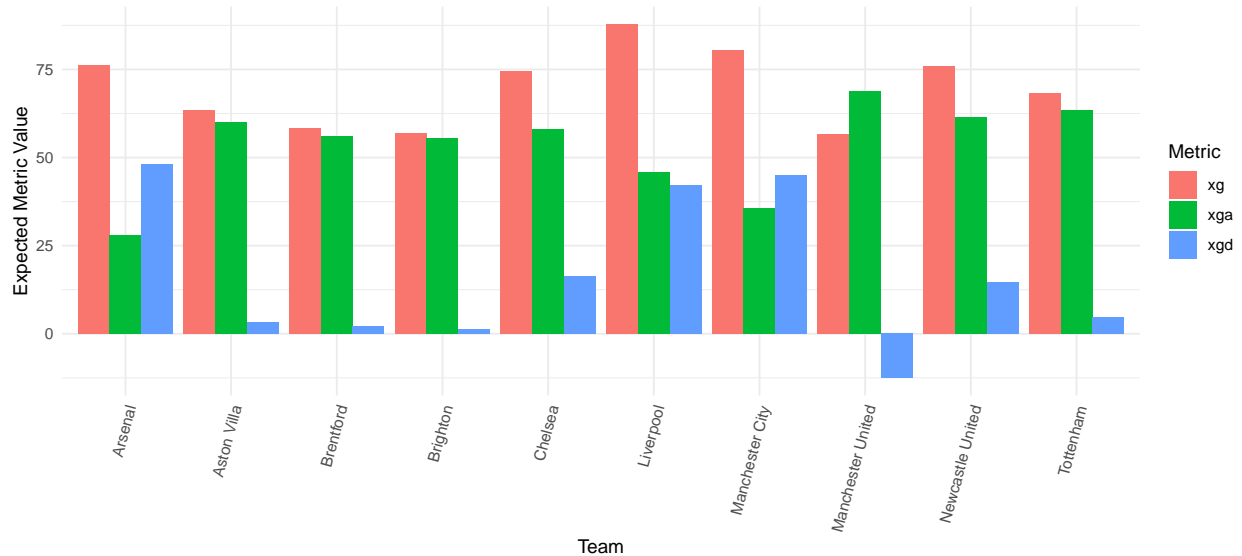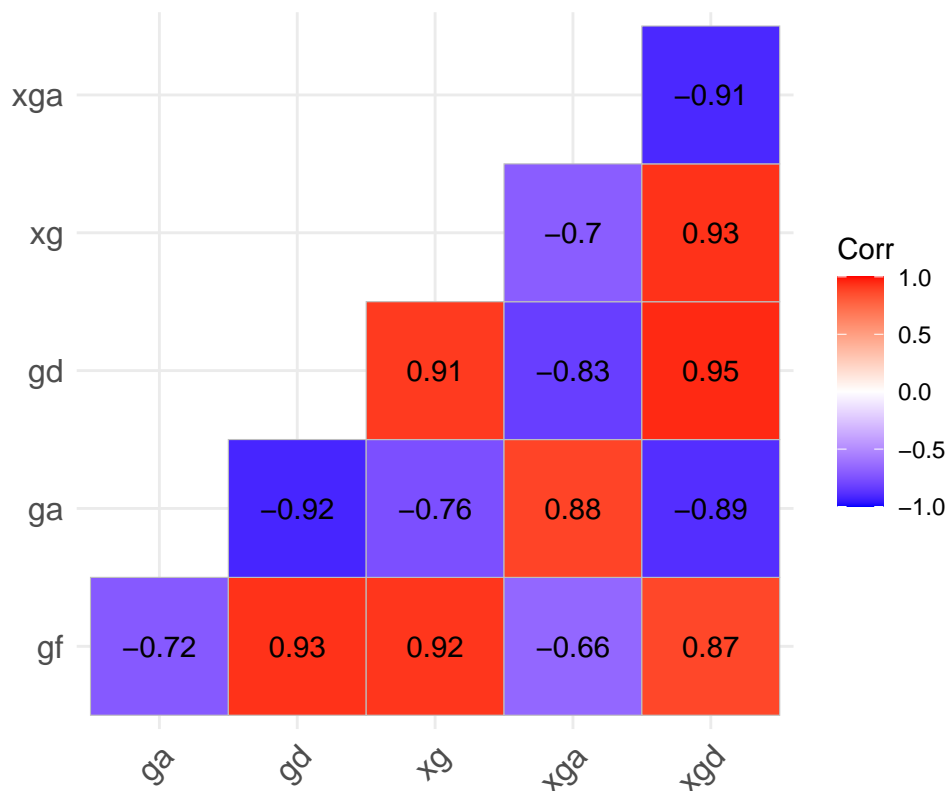## Top 10 Premier League Teams by xG (2023–24)

Figure 3: Premier League 2023–2024 Season xG Summary

The combination of a strong defence with a budding attacking mindset would result in an xGD of **48.2**. Manchester City, on the other hand, finished the season with an xGD of **45**, a little lower than Arsenal's.

4

This naturally drives us towards the question: *do xG and xGA* **actually** *correlate to the actual performance of a team?*.

Correlation Matrix: Premier League 2023–2024

In this heatmap, we can see that for the 2023-2024 season the expected metrics accurately described a team's performance. The relationship is particularly strong with *xG* and *Goals For*, showcasing an *r* of **0.92**. While not as strong as that of xG-GF, **xGA and GA** also show a tight relationship with an **r** of **0.88**. The result then is a tightly related xGD-GD r of **0.95**.

Arsenal FC would then end that campaign with a mean of **2.39** goals per match. Combined with an average **0.76** goals conceded per match, the team appeared poised for another strong run in the following season.

Who Needs A Striker?

Which brings us to the actual season in question. As an overview, let's take a look at the Premier League table at the end of the 2024-2025 season.

Table 2: Figure 4: Premier League Table, 2024

rk	team_name	gf	ga	gd	pts	xg	xga
1	Liverpool	86	41	45	84	82.2	38.6
2	Arsenal	69	34	35	74	59.9	34.4
3	Manchester City	72	44	28	71	68.1	47.7
4	Chelsea	64	43	21	69	67.8	47.3
5	Newcastle United	68	47	21	66	63.8	45.5
6	Aston Villa	58	51	7	66	56.1	50.1
7	Nottingham Forest	58	46	12	65	45.5	48.9

rk	team_name	gf	ga	gd	pts	xg	xga
8	Brighton	66	59	7	61	58.7	54.6
9	Bournemouth	58	46	12	56	64.0	48.5
10	Brentford	66	57	9	56	59.0	55.4
11	Fulham	54	54	0	54	49.0	47.2
12	Crystal Palace	51	51	0	53	60.4	49.1
13	Everton	42	44	-2	48	41.8	46.2
14	West Ham	46	62	-16	43	47.0	59.7
15	Manchester United	44	54	-10	42	52.6	53.8
16	Wolves	54	69	-15	42	43.7	58.1
17	Tottenham	64	65	-1	38	58.8	63.3
18	Leicester City	33	80	-47	25	32.6	71.9
19	Ipswich Town	36	82	-46	22	34.4	72.7
20	Southampton	26	86	-60	12	32.7	84.8

Again, second place despite having the league's best defensive structure. Arsenal's defense conceded only **34** goals throughout the season, beating the champions Liverpool who conceded **41** goals. This only makes it apparent that their difficulties in getting the ball to the back of the net still outweighed their defensive capabilities.
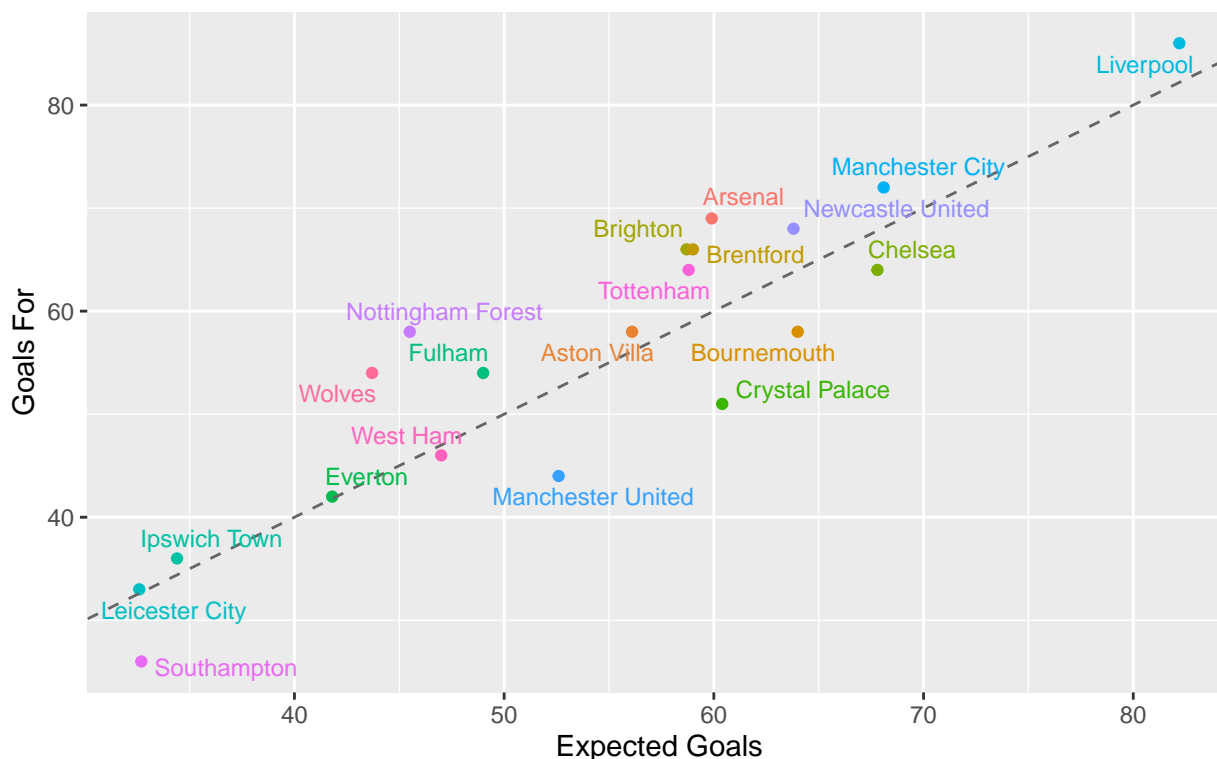

Figure 5: Premier League 2024 GF/xG Scatterplot

But who is responsible for that task? At this point in time, Arsenal had its number 9 in the form of Kai Havertz, a player who *isn't necessarily a traditional striker*. Combined with their aggressive midfield, we can see where Arsenal pulled most of its goals from.

Table 3: Figure 5: Cumulative goal share, 2024

player_label	goals	goal_share	cum_share
Kai Havertz (FWD)	9	0.1363636	0.1363636
Gabriel Martinelli (MID)	8	0.1212121	0.2575758
Leandro Trossard (MID)	8	0.1212121	0.3787879
Mikel Merino (MID)	7	0.1060606	0.4848485
Bukayo Saka (MID)	6	0.0909091	0.5757576

On this table we have the top 5 goal shares from that season. We can see that most of the team's goals come from the midfield players with a total of **29** goals between them, and Kai Havertz providing a further **9** goals on top of that.

In fact, in terms of cumulative shares, Kai Havertz, Gabriel Martinelli, Leandro Trossard, and Mikel Merino all claim a combined **48.48%** of Arsenal's total goals from 2024-2025. But while these players threatened the goal when they were on the pitch, injuries would limit their availability. Let's take a look at the players with the most minutes:

Table 4: Figure 6: Top 5 Arsenal Player Attendance

player_name	position	appearances	minutes	goals	assists
David Raya	GKP	38	3420	0	0
William Saliba	DEF	35	3041	2	0
Declan Rice	MID	35	2833	4	7
Thomas Partey	MID	35	2799	4	0
Leandro Trossard	MID	38	2550	8	7

From this list, only Leandro Trossard provided a big conribution to Arsenal's total goals. While it could be argued that Declan Rice provided **7** assists aside from **4** goals, his role in the team is less "striker" and more "controller". The rest of the attacking or high scoring players would end up injured, further limiting Arsenal's goal-scoring edge.

What's important is what we don't see—strikers who *aren't* named Kai Havertz. With Havertz being the sole striker, the team is reliant on the German to provide that kind of attacking behavior that midfielders and wingers don't come naturally with.

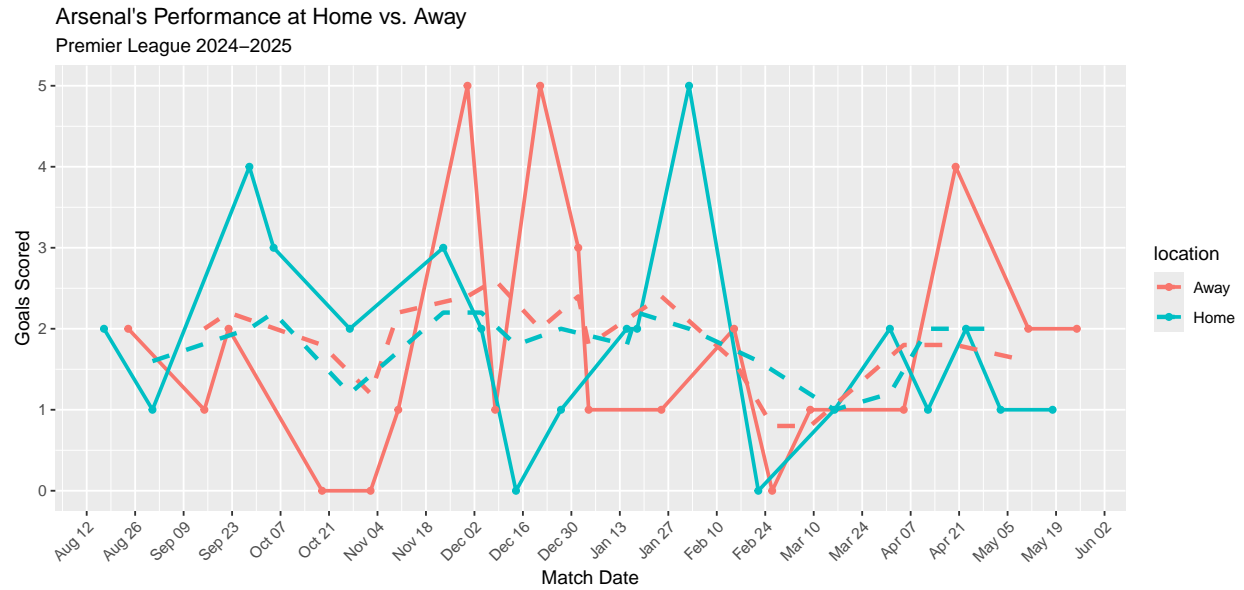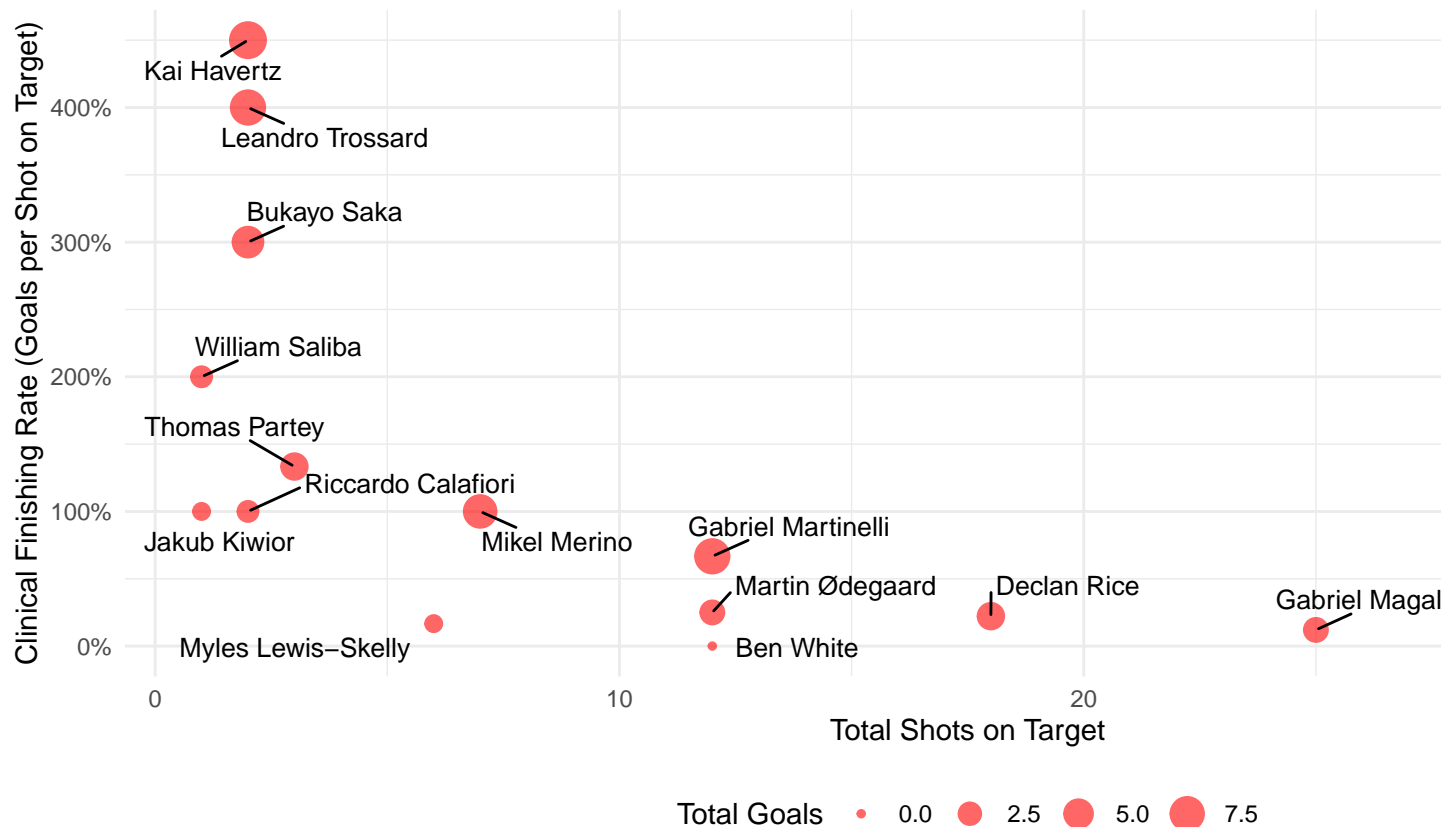Now, let's take into account Arsenal's goal scoring trends over the season.

Figure 6: Arsenal Home vs. Away Goals, 2024–2025

With the dashed-line showing the rolling averages, we can see that Arsenal's scoring form still suffers from inconsistencies. This only supports the argument that the team just didn't have the depth they needed to cover for absences caused by injuries.

But is it the players' fault? After all, personal challenges could cause players to underperform. This graph shows how efficient each player is in terms of goal-scoring by comparing their **goals per shot on target** against their **total shots on target**. The size of each point determines how many goals that player actually scored.

Arsenal's Clinical Finishing, 2024–2025

Comparing finishing rate against volume of shots on target

With the Y-axis showing us the player's *clinical finishing rate* and the X-axis showing *shots on target*, it's clear that Arsenal's attacking formation can't hit the target as much as they wanted to. Take Martin Odegaard, for example; the player is capable of getting into a position where he can shoot a ball on-target, but the man faces difficulties getting past the keeper. On the other hand, we have players like Kai Havertz who only have a few shots on-target but have a high finishing rate (higher than 100%!). This is an interesting point with the data that shows how the definition of a "Shot on Target" doesn't actually define a goal. For example, a player might shoot and "miss", but have the ball deflect into the goal via own-goal.

What we can then derive from the chart is that, while still a capable team, Arsenal's problem remains the same; they **need** a number 9, a traditional striker that converts chances into actual goals.

The team would end the season with a total xG of **59.9**, a far cry from last season's **76.1**. Despite scoring more than their xG, that's still a **-21.29%** change in their expected goal performance, a tell-tale sign that their chance creation is **not** what it was.

So, to anser the question of, "Who needs a striker?" the answer is definitely Arsenal.

Filling in the Gaps

[section discusses predictive stats/conclusions]