

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

- What factors contribute to Arsenal's inconsistent domestic performance, despite victories over top-tier teams like Real Madrid and Manchester City?
- To what extent do expected goals (xG) and expected goals against (xGA) reliably predict actual team outcomes?
- How does Arsenal's goal-scoring distribution among players impact their overall offensive effectiveness?

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: Scrapped 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.

Expecting Success

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: Premier League Table Summary, 2023-2024 Season

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