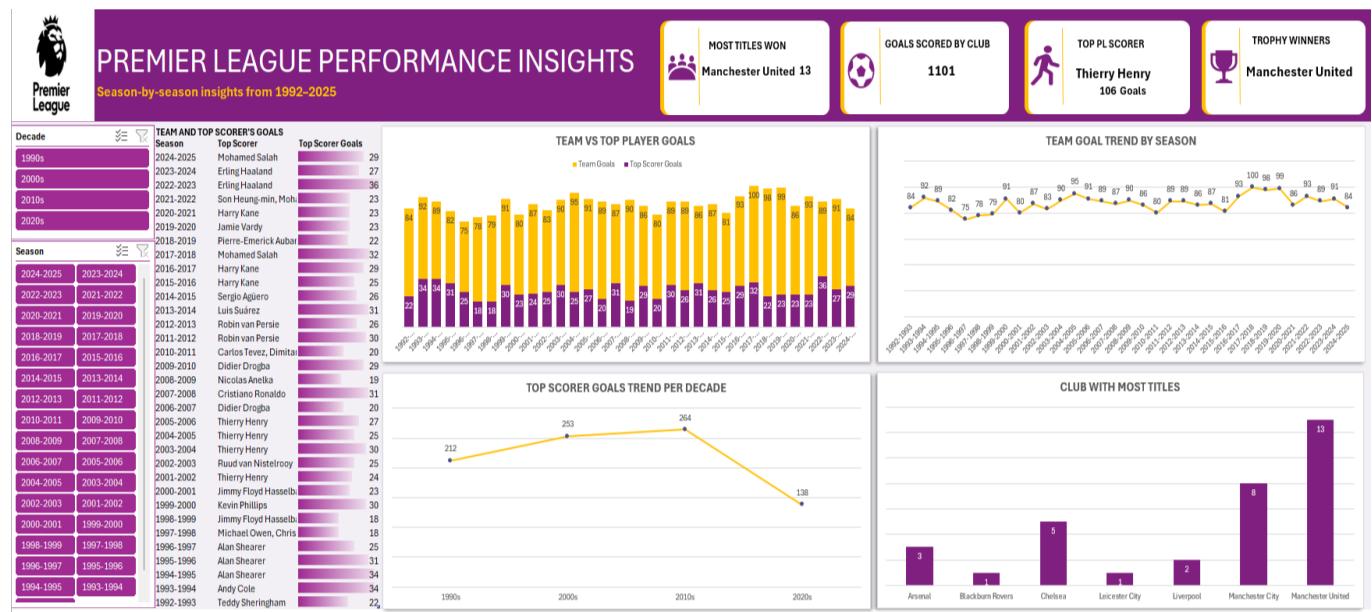


PREMIER LEAGUE DATA ANALYSIS (1992-2025)

ABSTRACT

This project explores 33 seasons of Premier League data (1992–2025) to uncover the relationship between team success and individual top scorers. Using Excel and Power Query, data was scraped from [FBref.com](#), cleaned, and visualized through interactive dashboards. The analysis compares team goal totals with top scorer contributions, tracks scoring trends by decade, examines title distribution among clubs and club goals per season.

Key findings reveal that team effort consistently outweighs individual brilliance in winning titles. The insights provide valuable context for understanding performance patterns and strategic team building in elite football.



INTRODUCTION

The English Premier League (EPL) stands as one of the most prestigious and competitive football leagues in the world. Over the past three decades, it has produced some of the most iconic football moments, legendary players, and dominant teams. With its global fanbase and billions in broadcast revenue, the performance of teams and players in the Premier League is constantly under scrutiny not only by fans, but also by analysts, scouts, and sports strategists seeking data-driven insights into what defines success at the highest level.

In this context, this project explores a fundamental yet often debated question:

Is a Premier League title more dependent on team performance or individual brilliance especially from the league's top goal scorers?

Problem Statement

While top scorers often receive the most attention, teams with strong collective performances are more likely to succeed. However, this relationship has not been clearly quantified across multiple seasons. Clubs, especially those building for long-term success, need to understand:

- **Are league titles typically won by teams with the top goal scorer?**
- **How much of a title-winning team's success depends on spreading goals across the squad?**
- **Do trends in team scoring and top scorer goals vary by era or tactical evolution?**
- **Which clubs have shown consistent dominance, and how is it reflected in their title count?**

Answering these questions requires historical data analysis, pattern recognition, and context-based interpretation.

Project Goals

This analysis focuses on the following objectives:

1. **Scrape and structure data** on Premier League winners and top scorers from the 1992–1993 to the 2024–2025 season using FBref.com.
2. **Clean and transform** the dataset in Excel using Power Query to make it suitable for trend analysis and dashboard visualization.
3. **Compare team goal outputs with top scorer goals** per season to assess their impact on title outcomes.
4. **Group seasons by decade** to identify broader historical trends in scoring and dominance.
5. **Visualize title distribution** among clubs to highlight long-term consistency or gaps in performance.

Through this project, the aim is to provide both football enthusiasts and stakeholders with clear, data-backed insights into what contributes to championship-winning campaigns helping bridge the gap between fan perception and analytical truth.

DATA DESCRIPTION

The dataset used in this project was manually scraped from [FBref.com](#) and covers **33 seasons** of the English Premier League, from **1992–1993 to 2024–2025**.

Data Cleaning & Transformation:

- Removed unnecessary columns that were not relevant to the analysis.
- Split combined columns (e.g., player name & goals, club name & goals) into separate fields for clarity.
- Added a new "Decade" column to group seasons by 1990s, 2000s, 2010s, and 2020s.
- Converted all data types (e.g., goals as integers, season as text) to ensure consistency.
- Appended and merged both tables where necessary to enable comparative analysis between team goals and top scorer goals.

This structured dataset served as the foundation for building insightful visualizations, calculating trends, and analyzing patterns in title-winning strategies over time.

The dataset contains the following variables:

Variable	Description	Data Type
Season	Premier League Season (e.g., “1995-96”)	Text
Winning Club	Name of the team that won the league title	Text
Team Goals	Total goals scored by the title-winning team	Numeric
Top Scorer	Name of player with the most goals that season	Text
Top Scorer’s club	The club which the top scorer played for	Text
Top Scorer Goals	Total goals scored by the top Scorer	Numeric
Decade	Grouping of the season into a decade (e.g., “2000s”)	Text

METHODOLOGY

This project followed a structured approach to collect, clean, transform, and analyze Premier League data using Microsoft Excel. The methodology is broken into the following phases:

1. Data Collection

- **Source:** Data was manually scraped from [FBref.com](#), a reliable public football statistics site.
- **Scope:** Premier League winners and top scorers from the **1992–1993** to **2024–2025** seasons.

2. Data Cleaning & Preparation (Excel Power Query)

- **Removed irrelevant columns** that did not contribute to the analysis.
- **Split combined values** (e.g., “Manchester City (99)” or “Salah (32)”) into separate columns for teams/players and their goal counts.
- **Standardized formats:**
 - Converted goals to integer data types.
 - Formatted seasons consistently as text (e.g., "1995–96").
- **Checked for duplicates and null values** — none were present post-cleaning.

3. Data Transformation

- Created a **new column to group seasons by decade** (1990s, 2000s, 2010s, 2020s).
- Appended and structured data into a format that allows season-by-season comparison between title-winning team goals and top scorer goals.

4. Data Analysis

- Analyzed:
 - Goal trends across seasons and decades.
 - Relationship between team success and individual top scorer performance.
 - Distribution of titles among clubs over 33 seasons.

6. Data Visualization

Created interactive **Excel charts** including:

- Line charts for **team goal trends over time**.
- Column charts for **title distribution across clubs**.
- Comparative visuals between **team goals vs. top scorer goals**.

7. Documentation & Insight Generation

- Key findings were summarized into actionable insights.
- A comprehensive report was developed to communicate the results to stakeholders and peers via LinkedIn, Twitter, GitHub, and Medium.

ANALYSIS

This section explores key insights and trends extracted from 33 Premier League seasons (1992–2025), using Excel-based techniques such as Power Query, calculated columns, and custom visualizations.

1. Team Goal Trend Analysis (1992–2024)

Chart Title: *Team Goal Trend by Season*

Method:

- Cleaned scraped data using **Power Query**: split “Winner (Goals)” into “Team” and “Goals”.
- Created a **Pivot Table** to summarize team goals per season.
- Visualized data using a stacked line chart with markers with color distinction for clarity.

Insights:

- Peak of 100 goals (2017–18).
- 2016–2019 was a high-scoring period.
- Goal dips in 1997–98, 1998–99, 2011–12, etc.
- Slight decline after 2019–20.
- Average goals mostly above 85.

Interpretation:

- Benchmark 2016–2019 to identify peak strategies.
- Investigate drop in recent seasons—could suggest tactical shifts or squad depth issues.

2. Top Scorer vs. Title-Winning Team Comparison

Chart Title: *Team vs Top Player Goals*

Method:

- Used Power Query to split "Top Scorer (Goals)" into "Player" and "Goals".
- Created a **Pivot Table** to align top scorer goals with team goals per season.
- Visualized via a **stacked column chart** with clear labeling.

Insights:

- Rare overlap between top scorer and title-winning team.
- Suggests **individual brilliance doesn't always translate to trophies**.

Interpretation:

- Balanced scoring across a team is more effective.
- Clubs shouldn't overly rely on one forward for success.

3. Club Title Distribution

Chart Title: *Clubs with the Most Titles*

Method:

- Used a **Pivot Table** to count how many times each club appeared as a winner.
- Built a **horizontal bar chart** to showcase distribution.

Insights:

- Manchester United dominates with 13 titles.
- Manchester City, Chelsea and Arsenal trail with 8, 5 and 3 respectively.
- One-time winners like Blackburn and Leicester are outliers.

Interpretation:

- Long-term consistency is rare and valuable.
- Mid-table teams must focus on sustained development.

4. Top Scorer Goals Trend per Decade

Chart Title: *Top Scorer Goals Trend per Decade*

Method:

- Added a **calculated column** to classify seasons into decades (1990s, 2000s, etc.).
- Used a **Pivot Table** to sum top scorer goals by decade.
- Visualized with a **line chart** to highlight decade-wise trend.

Insights:

- Increase from 212 in 1990s to 264 in 2010s.
- Sharp drop to 138 in 2020s so far.

Interpretation:

- Goals now more evenly distributed across teams.
- Indicates tactical maturity, rotation, and improved defensive setups.

RESULTS

1. Goal Trends Over Time

- The **highest-scoring season** was in **2017–18**, with 100 goals scored by the title-winning team.
- From **2016–17 to 2019–20**, teams consistently scored **above 98 goals**, indicating a period of strong offensive performance.
- The **lowest-scoring seasons** include 1997–98 (75 goals) and 1998–99 (78 goals), showing how goal averages have fluctuated over time.
- A **slight decline** in recent seasons suggests possible tactical shifts, player transitions, or defensive improvements.

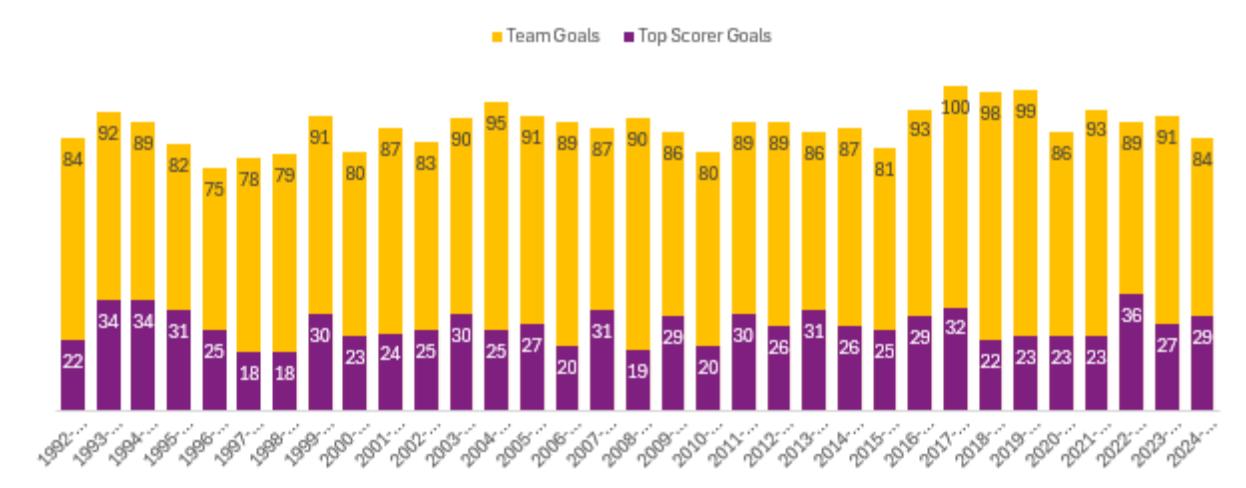
TEAM GOAL TREND BY SEASON



2. Top Scorer vs. League Winner

- Most **top scorers did not belong to the league-winning teams**, highlighting a weak correlation between having the top scorer and winning the title.
- This reveals the importance of **balanced team play** and distributing goal contributions rather than relying on one player.

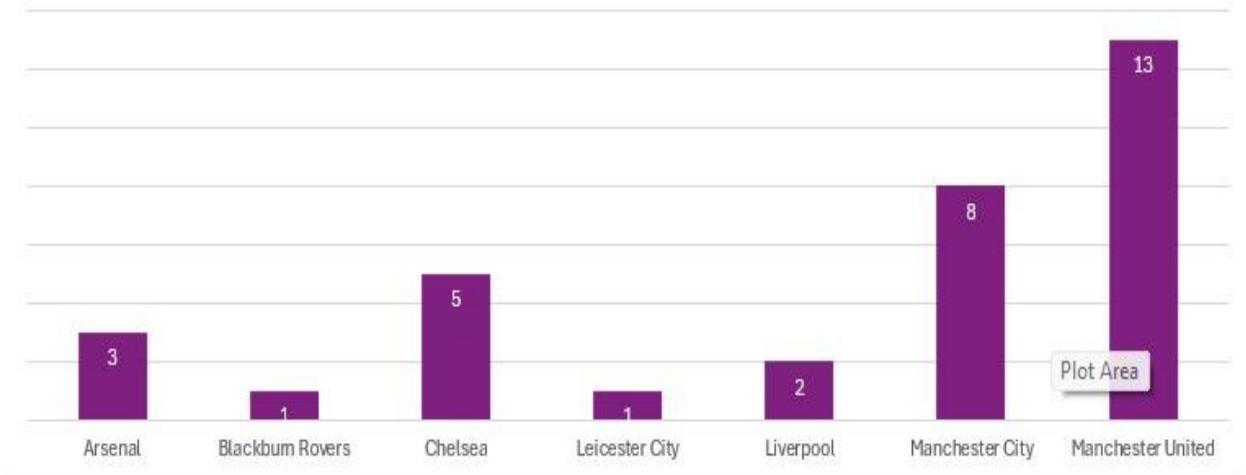
TEAM VS TOP PLAYER GOALS



3. Title Dominance

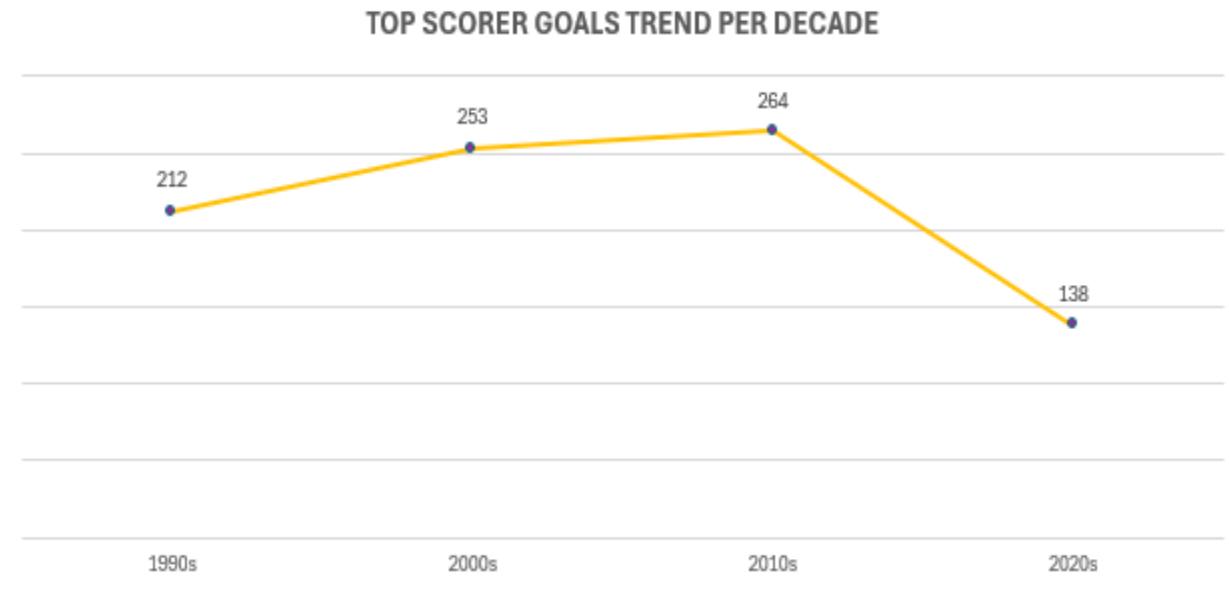
- **Manchester United** leads with 13 titles, followed by Chelsea (5) and Arsenal (3), showcasing long-term dominance by a few clubs.
- Only **5 clubs** have won the Premier League since 1992, revealing a **high concentration of success** at the top.
- Other title-winning clubs (like Leicester and Blackburn) are **rare outliers**.

CLUB WITH MOST TITLES



4. Top Scorer Goals Trend per Decade

- Top scorer totals increased from **212 (1990s)** to **264 (2010s)**, reflecting a rise in individual scoring.
- A significant drop to **138 goals** in the 2020s (as of 2024) may indicate a growing trend of **goal distribution across teams**, reduced season lengths, or evolving tactics.



Discovered Patterns

- **Goal scoring peaked in the mid-to-late 2010s**, with both teams and individual players achieving higher numbers.
- **Team success is more about consistency and depth**, not just individual brilliance.
- **Title wins are heavily concentrated** among a few elite clubs, pointing to structural advantages in recruitment, finances, and management.

DISCUSSION

This analysis of Premier League data from the 1992–1993 to the 2024–2025 season reveals several notable trends in goal scoring, team success, and title distribution that offer meaningful insights into the evolution of the league's competitive dynamics.

1. Team Success vs. Individual Excellence

One of the most striking findings is the weak correlation between the league's top scorers and title-winning teams. In most seasons, the player who scored the most goals did not belong to the club that won the league. This trend suggests that **team success is less dependent on individual brilliance and more on overall squad balance, consistent performance, and tactical depth**. Clubs with effective attacking distribution across multiple players, supported by strong defensive frameworks, appear to be more successful than those that rely heavily on a single prolific scorer.

2. Trends in Team Goal Scoring

The data indicates a **significant increase in team goals** during the period from 2016–17 to 2019-20, with goal tallies reaching a peak of 100 in the 2017–18 season. This period can be characterized as an offensive peak, potentially driven by tactical innovations, world-class player acquisitions, and managerial strategies focused on high pressing and attacking football. However, the slight decline in team goals observed in recent seasons suggests a possible tactical shift. This may reflect increased defensive rigidity among clubs, changes in game tempo, or even broader factors such as fixture congestion and squad rotation.

3. Decadal Trend in Top Scorer Goals

When grouped by decade, the total number of goals scored by top scorers shows an upward trend from the 1990s (212 goals) to the 2010s (264 goals), followed by a steep decline in the 2020s (138 goals). This decline may point to **a growing trend toward goal distribution**, where multiple players contribute to scoring, as well as improved defensive setups across teams. Additionally, structural changes like compressed schedules (e.g., due to COVID-19), squad rotation policies, and increased parity in competition may be influencing individual scoring trends.

4. Title Distribution and Competitive Balance

Another important aspect of this study is the **high concentration of league titles among a few elite clubs**, with Manchester United, Manchester City, Chelsea, and Arsenal dominating the winners' list. This pattern reflects long-standing disparities in financial power, recruitment networks, and organizational stability. The emergence of outlier champions such as Blackburn Rovers and Leicester City illustrates that success is possible outside the traditional powerhouses, though it remains rare. These findings highlight the importance of long-term strategic planning, robust academy systems, and adaptive management for sustained success.

Implications

The findings of this analysis are valuable for club management, analysts, and fans seeking to understand performance patterns and strategic indicators. Clubs may use this information to:

- Benchmark team and player performance across decades.
- Evaluate the efficacy of recruitment and tactical strategies.
- Identify factors contributing to peak goal-scoring performance.
- Rethink resource allocation for long-term competitiveness.

RECOMMENDATIONS

Based on the trends and insights uncovered in this analysis of Premier League data from the 1992–1993 to 2024–2025 seasons, the following recommendations are proposed for stakeholders, particularly club analysts, coaches, and decision-makers:

1. Adopt a Balanced Team Strategy Over Individual Reliance

Given the weak correlation between top scorers and league titles, clubs should **prioritize team-wide attacking contributions** rather than over-relying on a single player. Investment in a well-rounded squad and system-oriented play is more likely to yield consistent success.

2. Investigate and Emulate Peak Performance Periods

Clubs should study the tactical and structural approaches of teams that peaked in the **2016–2019** period. Understanding factors such as pressing intensity, managerial style, and player dynamics during high-scoring seasons can inform strategic planning and recruitment.

3. Enhance Defensive and Tactical Adaptability

With recent declines in goal tallies, it is important to ensure tactical flexibility. Clubs should **develop systems that can adapt** to evolving league dynamics, including stronger defensive setups and more evenly distributed goal-scoring responsibilities.

4. Benchmark Performance Using Decadal Trends

Clubs and analysts can benefit from viewing performance through **decade-based comparisons**. This provides a broader context for evaluating progress, setting targets, and identifying underperformance patterns.

5. Promote Sustainable Long-Term Planning

The concentration of titles among a few clubs emphasizes the value of **organizational consistency, talent development, and financial planning**. Emerging or underperforming clubs should invest in infrastructure, analytics, and youth development to build long-term competitive capabilities.

6. Leverage Data-Driven Decision Making

Clubs should **incorporate historical and performance data into decision-making frameworks** for scouting, tactical adjustments, and performance reviews. Using tools like Excel dashboards or BI platforms can support clearer strategic insights.