IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Which Soccer Metrics Best Predict Winning? A Data-Driven Analysis Across Europe's Top Five Leagues

Bishrelt "Giga" Bat-Erdene

Newton North High School, Newton, MA, 02460, USA; gigabat0125@gmail.com

ABSTRACT: This study examines which team-level measures hold the highest promise for forecasting success within Europe's top five soccer leagues. By employing data during the 2023–24 and 2024–25 seasons, I tested outcomes (goal difference, goals, points), expected metrics (xGA, xG, xGD), style (possession, progression, G+A/90, xG+xAG/90), and a bespoke finishing efficacy measure. All analyses were carried out in Microsoft Excel with correlations, regression, cluster, and residual analysis. Points per game (Pts/MP) normalized success across leagues. Goal difference and expected goal difference (xGD) had the best correlations with Pts/MP, while chance creation metrics (xG+xAG/90 and G+A/90) were very predictive. Regression models found that defense was as important as attack: decreasing xGA was always associated with increased points. Finishing efficiency sets apart elite and mid-table clubs. Cluster analysis determined five stable play profiles: high-possession progressors, controlled buildup, vertical creators, deep-block survivalists, and direct counters, with comparable seasonal performance gaps. League-wise, the Premier League combined higher chance creation with results, and Serie A and La Liga were efficient with limited chances; the Bundesliga and Ligue 1 underperformed compared to chance creation. Taken together, elite soccer success is a balance of chance creation, defensive solidity, and clinical finishing. These insights provide actionable guidance for coaches and analysts while giving fans a data-driven lens to understand team performance.

KEYWORDS: Mathematics, Applied Statistics, Regression Analysis, Soccer Analytics, European Football.

Introduction

The growth of soccer analytics has revolutionized the way that clubs, analysts, and supporters gain insights into performance. Foretelling success with soccer is still multifaceted. Conventional metrics, including goals, shots, and possession, provide some understanding, but frequently they don't represent the whole picture of performance. The data and metrics of soccer analytics have grown exponentially over the past several years through the aid of public data and sophisticated metrics.[1] Sophisticated metrics now permit greater analysis of attacking and defensive effectiveness, but the argument rages over which metrics best portend success by leagues and seasons.[2-6]

Prior research frequently relied on observations of individual competitions or brief periods of play, and generalizability was therefore limited. Thus, while some have observed the Premier League to stress possession and chance creation, others have noted the Serie A and La Liga to exemplify defensive effectiveness and tactical compactness.[7,8] Scarce analyses have juxtaposed data across various leagues by using standardized information, and therefore, the question of which indicators of performance correspond best to success within varying competitive conditions remains.[6,9,10]

This research aims to fill that gap by studying team-level performance in Europe's top five leagues (Premier League, La Liga, Bundesliga, Serie A, Ligue 1) over two consecutive seasons (2023–24 and 2024–25). This study assesses both basic and advanced key performance indicators (KPIs) with systematic methods using Microsoft Excel. This procedure encompasses correlation, regression, clustering, and residual analyses to determine which metrics best predict points per match (Pts/MP), a normalized outcome of success. By contrasting across leagues and styles of play, this study aims to define a clear framework by which to define drivers of team success in contemporary soccer.

Methods

Data collection:

Downloaded team-level performance data for the 2023–24 and 2024–25 seasons of Europe's top five European soccer leagues (Premier League, La Liga, Bundesliga, Serie A, Ligue 1) from fbref.com. Raw tables (matches played, goals scored/conceded, expected goals [xG], possession, progressive passes, assists, etc.) were imported into Microsoft Excel and standardized across leagues. There was a total of 98 different teams over the two seasons included in the dataset. All metrics were summative at the full-season level, and both the 2023–24 and 2024–25 datasets represent complete seasons.[2]

Key performance indicators (KPIs):

I examined both basic and advanced KPIs, clustered into four sections:

- ***Outcome metrics:*** Goals For (GF), Goals Against (GA), Goal Difference (GD), Points, and Points per Match (Pts/MP).
- ***Expected metrics:*** Expected Goals (xG), Expected Goals Against (xGA), Expected Goal Difference (xGD), and Expected Goals + Expected Assists per 90 minutes (xG+xAG/90).

DOI: 10.36838/IJHSR84.30

• ***Style metrics:*** Possession (%), Progressive Passes (PrgP), Goals + Assists per 90 (G+A/90), xG per Possession, and Progressive Passes per Possession.
• ***Efficiency metric:*** To assess clubs' ability to turn expected opportunities into actual output, I developed a bespoke ***Finishing Efficiency*** metric:

$$Finishing\ Efficiency\ = \frac{G + A\ per\ 90}{xG + xAG\ per\ 90}$$

A full reference table of all KPIs used in this study is summarized in Table 1.

Table 1: Key performance indicators analyzed. These variables formed the foundation of all subsequent analyses.

KPI	Category	Description
GF	Outcome	Goals For (total)
GA	Outcome	Goals Against (total)
GD	Outcome	Goal Difference (GF – GA)
Pts	Outcome	Total Points in the season
Pts/MP	Outcome	Points per Match
xG	Expected	Expected Goals
xGA	Expected	Expected Goals Against
xGD	Expected	Expected Goal Difference
xG+xAG/90	Expected	Expected Goals plus Expected Assists per 90 minutes
Poss	Style	Possession Percentage
PrgP (Progression)	Style	Total Progressive Passes
G+A/90	Style	Goals plus Assists per 90 minutes
xG per Poss	Style	Expected Goals per unit of Possession %
Prog Passes per Poss	Style	Progressive Passes per unit of Possession %
Finishing Efficiency	Efficiency	Ratio of actual output (G+A per 90) to expected output (xG+xAG per 90)

Derived metrics (Excel formulas):
Key variables were calculated directly in Excel:
• ***Pts/MP:*** =Total Points / Matches Played
• ***Finishing Efficiency:*** (Goals + Assists per 90) / (xG + xAG per 90)
• ***Percentile ranking:*** =PERCENTRANK.INC(Range, Cell) to rank teams within distributions
• ***Quartile thresholds:*** =PERCENTILE.INC(Range, 0.25) and =PERCENTILE.INC(Range, 0.75) defined cutoffs for "Low," "Mid," and "High" buckets
• ***Efficiency bucket formula (example):*** =IF($B2<0.97,"Low",IF($B2<=1.02,"Mid","High"))
• ***Style labels:*** Possession, progression, and efficiency buckets were concatenated (e.g., =BH2&"-"&BI2) to generate interpretable team profiles such as *High-possession progressors* or *Direct counters*.

Statistical analyses:
I used MS Excel for the following multiple analyses to assess relationships between KPIs and team success (measured by Pts/MP):
• ***Correlation:*** =CORREL(Y-range, X-range) tested pairwise KPI associations.
• ***Multiple regression:*** Conducted using Excel's Regression tool within the *Data Analysis* add-in. Outputs included coefficients, standard errors, t-statistics, p-values, and R^2 values. Residuals were calculated by subtracting predicted Pts/MP from observed values.
• ***Clustering:*** Teams were grouped into five stylistic clusters (high-possession progressors, controlled buildup teams, vertical creators, deep-block survivalists, and direct counters) using percentile thresholds for possession, progression, and efficiency.
• ***Residual analysis:*** Compared observed vs. expected Pts/MP to identify over- and under-performers relative to KPI-based models.

Excel analysis tools:
• ***Pivot tables:*** Aggregated averages (e.g., Pts/MP, possession, efficiency) by league and cluster.
• ***Charts:*** Scatterplots (e.g., xG+xAG/90 vs Pts/MP, with bubble size = possession, color = efficiency) and residual plots. Team names were overlaid using *Format Data Labels → Value from Cells*.
• ***Conditional formatting:*** Applied to highlight efficiency levels and cluster differences in visual interpretation.

Results

Elite teams consistently turned dominance into points:
Points per match (Pts/MP) was adopted as the overarching point of comparison across both seasons, as this adjusts success to account for leagues with varying overall counts of matches through the premier five leagues. The comparison of elite and relegation-level clubs through Pts/MP is depicted in Table 2. The values spanned between 0.32 (2024–25 by Southampton) and 2.65 (2023–24 by Leverkusen). The clubs at the top of each league table were those with the highest Pts/MP, with a direct relationship between this metric and league position.

Table 2: Elite vs. relegation-level teams by Pts/MP. This comparison highlights the large performance gap between elite and relegation clubs, confirming that Pts/MP clearly differentiates top and bottom performers across leagues.

Season	Highest Pts/MP (Elite club)	Pts/MP	Lowest Pts/MP (Relegation club)	Pts/MP
2023–24	Leverkusen (Bundesliga)	2.65	Granada (La Liga)	0.55
2024–25	Liverpool (Premier League)	2.21	Southampton (Premier League)	0.32

This comparison reveals the divide between elite and struggling clubs. Although playing styles varied, Pts/MP clearly distinguished between top performers from relegation-threatened clubs.

Chance creation was the strongest predictor of success:
Comparing various key performances (KPIs) with points per game (Pts/MP) revealed that chance creation and conversion were of higher significance than possession or efficiency by itself during both seasons, as revealed by Table 3.
• In 2023–24, the strongest associations with Pts/MP were goal difference (r = 0.97), expected goal difference (xGD, r = 0.91), and chance creation by G+A per 90 (r = 0.85) and by xG+xAG per 90 (r = 0.83).
• In 2024–25, the trend persisted: goal difference (r = 0.97) and xGD (r = 0.93) were the strongest predictors, with G+A per 90 (r = 0.87) and xG+xAG per 90 (r = 0.84) closely correlating with success.

• Possession and progression had moderate but significant associations (r = 0.73–0.83), and finishing efficiency was considerably weaker (r = 0.50).

These findings indicate that teams regularly creating good-quality chances and having healthy expected goal totals were significantly more likely to win, irrespective of possession control.

Table 3: Correlation of key metrics with Pts/MP. Correlation coefficients (r) for outcome, expected, and playstyle metrics across the 2023–24 and 2024–25 seasons. Chance creation (xG+xAG/90 and G+A/90) consistently showed the strongest link with success, while possession and finishing efficiency were weaker predictors.

KPI	Correlation_with_Pts/MP	
	2023-24 season	2024-25 season
GD	0.97	0.97
xGD	0.91	0.93
G+A/90	0.85	0.87
PrgP (Progression)	0.83	0.81
Poss	0.78	0.73
xG+xAG/90	0.83	0.84
Finishing_Efficiency	0.50	0.50

Efficiency separated the champions from the rest:

Regression models were applied to determine which of the KPIs predicted success, defined by Pts/MP. In both seasons, Goal Difference (GD) best correlated with Pts/MP (r = 0.97) by itself, but regression models were necessary to determine which of the advanced metrics continued to exist when considered together.

Table 4 shows the significant predictors of team success determined by regression models. Defensive strength was the clearest indicator of success in 2023–24. Defenses with lower expected goals against (xGA) were extremely significant ($p < 0.001$), so clubs that allowed fewer good chances were regularly among the best. When xGA was taken out of the model, ball progression (PrgP) came into significance (p = 0.034), and possession had a borderline significance (p = 0.075). Finishing efficiency was not significant in the model on either occasion.

In 2024–25, finishing efficiency came into play. The better the conversion of chances, the bigger a measurable advantage (p = 0.046 with xGA; p = 0.038 without xGA). Progression was significant again (p = 0.046 with xGA; p = 0.003 without xGA), and xGA was very predictive ($p < 0.001$). Creation of chance (xG) approached significance within the model without xGA (p = 0.068), but did not quite pass the threshold.

These results suggest that creating opportunities is important, but the very best clubs were separated by defensive solidity and effective finishing. Liverpool (2024–25) and Leverkusen (2023–24) both combined these qualities, whereas Manchester United and Sevilla underperformers were unable to convert solid KPI profiles into reliable points.

Table 4: Key predictors of team success (Pts/MP) from regression models. These models demonstrate that defense (low xGA) and progression consistently predicted points, while finishing efficiency distinguished champions in 2024–25.

Season	Significant predictors (p < 0.05)	Notes
2023–24	xGA (lower values = stronger defense), PrgP (Progression)	Defensive solidity was the strongest predictor; progression added some value, but finishing efficiency was not significant.
2024–25	xGA (lower values = stronger defense), PrgP (Progression), Finishing Efficiency	Defense remained crucial, but finishing efficiency also separated top clubs from others.

Full regression outputs with coefficients, t-statistics, and p-values are provided in the Excel files.

Finishing efficiency was the great divider:

To track how effectively clubs converted expected chances into actual output, I developed a bespoke Finishing Efficiency stat:

$$Finishing\ Efficiency = \frac{G + A\ per\ 90}{xG + xAG\ per\ 90}$$

A value greater than 1.0 signifies clinical conversion (more assists/goals than predicted), and a value below 1.0 denotes poor performance. This indicator was needed because of the way standard xG-based metrics only account for chance quality but not whether teams consistently finished those chances.[11]

Figure 1 shows the relationship between finishing efficiency and team success in both seasons. Looking across both seasons, the pattern was clear. In 2023–24, clubs like Atalanta, Roma, Real Madrid, and Leverkusen were among the most efficient finishers, regularly turning chances into points. On the other hand, Everton, Köln, Cádiz, and Lecce wasted too many opportunities and fell behind. The same trend showed up again in 2024–25. Nottingham Forest, Wolves, Bologna, and Holstein Kiel all finished at a high level, while Montpellier, Valladolid, Southampton, and Real Sociedad could not make their chances count.

In short, teams that were more clinical in front of goal usually ended up higher in the table, while those that struggled to convert chances often found themselves stuck in the bottom half.

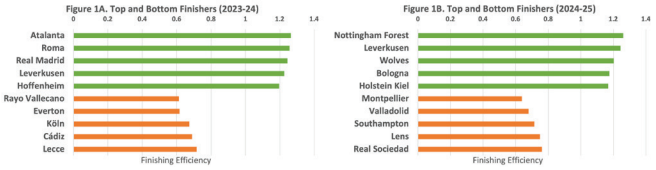

Figure 1: Finishing efficiency across Europe's top five leagues. (A) Top five and bottom five clubs by finishing efficiency in 2023–24. Teams above 1.0 consistently converted their chances, while those below 1.0 struggled to score. (B) Top five and bottom five clubs by finishing efficiency in 2024–25. Once again, strong finishers separated themselves from clubs that couldn't make the most of their opportunities.

Five clear playstyle clusters emerged across Europe:

Teams across both seasons were clustered into five groups using possession, progression, and

effectiveness, as shown in Table 5 and Figure 2:

1. ***High-possession progressors*** (e.g., Man City, Real Madrid, PSG, Bayern) – these clubs consistently paired heavy possession with elite finishing, averaging ~2.1 Pts/MP and ranking near the top of their leagues.
2. ***Controlled buildup teams*** (e.g., Napoli, Roma, Lille, Bologna) – patient and balanced possession produced solid but not elite outcomes, averaging ~1.7 Pts/MP.
3. ***Vertical creators*** (e.g., Atalanta, Monaco, Atlético) – direct forward play produced mid-table success (~1.8 Pts/MP).
4. ***Deep-block survivalists*** (e.g., Dortmund, Newcastle, Lyon, Aston Villa) – defensive setups averaged just ~1.2 Pts/MP, with many teams fighting to stay in mid-table.

DOI: 10.36838/IJHSR84.30

5. ***Direct counters*** (e.g., Everton, Union Berlin, Cagliari, Ipswich Town) – reactive low-possession play was the least effective, averaging ~0.9 Pts/MP and league ranks close to relegation.

Table 5: Cluster style summaries across two seasons. These results confirm five stable playstyle clusters with reproducible performance patterns across seasons, where high-possession progressors consistently dominated.

Cluster / Style	Avg Pts/MP		Avg League Rank		Notes
	2023–24	2024–25	2023–24	2024–25	
Controlled buildup teams	1.7	~1.7	5.3	~6.0	**Elite:** Bologna, Napoli; **Mid-table:** Roma, Lille, Real Sociedad
Deep-block survivalists	1.2	~1.2	11.4	~11.0	**Mid-table:** Dortmund, Villa, Newcastle; Relegation: Lecce, Granada, Monza
Direct counters	0.9	~0.9	14.8	~14.5	**Relegation:** Everton, Cagliari, Union Berlin, Ipswich
High-possession progressors	2.1	~2.1	3.8	~3.5	**Elite:** Man City, Arsenal, Liverpool, Real Madrid, Barcelona, Bayern, Leverkusen, Inter, Chelsea
Vertical creators	1.8	~1.8	5.0	~5.0	**Mid-table:** Atalanta, Monaco, Atlético, Lens

The same five clusters appeared in both seasons, and gaps in their performance were nearly identical. That stability emphasizes that the model registers enduring tactical profiles across Europe, and not short-term fluctuations.

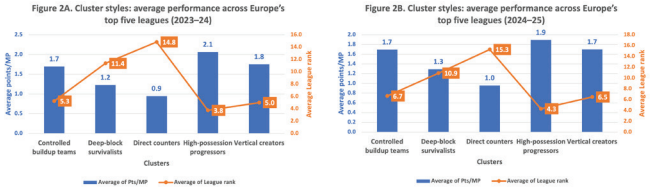

Figure 2: Cluster styles and average performance across Europe's top five leagues (A: 2023–24 and B: 2024–25). Bar and line charts show average Pts/MP (bars) and average league rank (line) for each of the five clusters. High-possession progressors consistently dominated, while direct counters struggled at the bottom.

Some clubs beat the numbers, others fell short:

Residual analysis compared actual performance (Pts/MP) with predicted values from regression models. This highlighted which clubs earned more or fewer points than expected based on their KPI profiles.

Overperforming and underperforming teams from residual analysis appear in Table 6. In the season 2023–24, Juventus, Inter, and Atlético Madrid were among the strongest overperformers, collecting points consistently higher than anticipated. Brest and Nice also stood out within Ligue 1 among the teams converting efficient play into better results. By contrast, Almería and Burnley emerged as clear underperformers, and even Bayern Munich collected points lower than anticipated with their superb metrics.

Napoli made the largest positive residual measurement in 2024–25, with their results exceeding the models by a significant amount. There were also bigger-than-expected outcomes for Freiburg and Fiorentina, while Roma and Rayo Vallecano contributed to additional surprises. Tottenham, on the other hand, came in way short of its expected values and hence was the biggest underperformer. There were also poor performances by Saint-Étienne and Rennes of Ligue 1, and by Holstein Kiel of the Bundesliga.

This analysis reveals how intangibles such as coaching, mentality, and the balance of the squad can propel or drop squads past or below statistical projections. Certain high-profile clubs like Bayern and Tottenham underperformed compared to expectations, maybe because of injuries or tactical switches. Interestingly enough, some clubs that overperformed in the residual analysis, like Inter and Juventus, also happen to have high-value rosters, implying that the market value of the roster could account in part for their outcomes and worth analyzing in future research.

Table 6: Top 5 overperformers and underperformers by residuals (2023–24 and 2024–25). Residuals = Actual Pts/MP – Predicted Pts/MP. Positive values show teams that outperformed their KPI profile, while negative values show underperformance.

Season	Top 5 Overperformers (highest residuals)	Residual	Top 5 Underperformers (lowest residuals)	Residual
2023–24	Juventus (Serie A)	+0.52	Almería (La Liga)	–0.45
	Nice (Ligue 1)	+0.45	Burnley (Premier League)	–0.40
	Inter (Serie A)	+0.41	Bayern Munich (Bundesliga)	–0.38
	Atlético Madrid (La Liga)	+0.37	Salernitana (Serie A)	–0.38
	Brest (Ligue 1)	+0.35	Darmstadt 98 (Bundesliga)	–0.37
2024–25	Napoli (Serie A)	+0.61	Tottenham (Premier League)	–0.65
	Freiburg (Bundesliga)	+0.32	Holstein Kiel (Bundesliga)	–0.44
	Fiorentina (Serie A)	+0.31	Saint-Étienne (Ligue 1)	–0.41
	Rayo Vallecano (La Liga)	+0.30	Rennes (Ligue 1)	–0.39
	Roma (Serie A)	+0.30	Ipswich Town (Premier League)	–0.36

League style profiles:

League comparisons across the board revealed separate stylistic variations between Europe's top five leagues. Table 7 and Figure 3A-B summarize how finishing efficiency, expected goals plus expected assists (xG+xAG), possession, and results (Pts/MP) aligned.

In 2023–24, the Premier League led in attacking production (2.83 xG+xAG/90) and possession (51.4%), and recorded the highest Pts/MP (1.52) too. La Liga and Serie A recorded lower xG+xAG (2.29 and 2.19) with similar Pts/MP maintained (1.48 and 1.46), but with higher efficiency in turning fewer chances into results. Ligue 1 trailed with both low finishing efficiency (0.94) and points (1.40). Bundesliga recorded plenty of opportunities (2.66 xG+xAG) but yielded only a modest 1.42 Pts/MP, indicating underperformance relative to output.

In 2024–25, the Premier League again led by both chance creation (2.65 xG+xAG) and outcomes (1.53 points/match). Serie A and La Liga tied at 1.48 points/MP despite Serie A generating fewer chances (2.17 vs. 2.31 xG+xAG), again suggesting greater efficiency. Ligue 1 edged up (1.45 points/MP) but still lagged, while the Bundesliga was again the least efficient with only an average of 1.41 points/MP despite lofty xG+xAG.

Overall, these patterns convey two consistent themes: (1) the Premier League combined volume (chances and possession) with results, while (2) Serie A and La Liga showed efficiency in converting fewer chances into similar points. In comparison, the Bundesliga and Ligue 1 underperformed relative to their chance creation.

Table 7: League averages of style metrics and results. These results show consistent inter-league patterns, with the Premier League combining the most chance creation and results, while Serie A and La Liga achieved efficiency from fewer chances.

Season	League	Finishing efficiency	xG+xAG/90	Poss (%)	Pts/MP
2023–24	Bundesliga	1.04	2.66	50.3	1.42
	La Liga	0.99	2.29	50.9	1.48
	Ligue 1	0.94	2.36	50.1	1.40
	Premier League	1.01	2.83	51.4	1.52
	Serie A	1.01	2.19	50.7	1.46
2024–25	Bundesliga	1.05	2.53	50.3	1.41
	La Liga	0.97	2.31	50.8	1.48
	Ligue 1	0.97	2.59	50.3	1.45
	Premier League	1.01	2.65	50.9	1.53
	Serie A	1.03	2.17	50.9	1.48

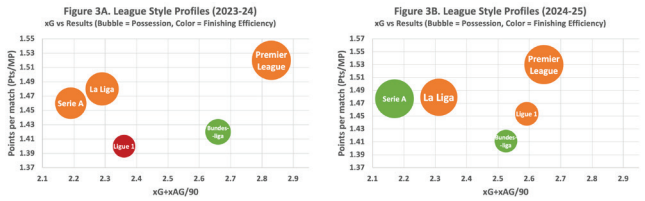

Figure 3: League style profiles (A: 2023–24, B: 2024–25). Bubble charts showing average league styles. X-axis: xG+xAG per 90; y-axis: Pts/MP. Bubble size = possession; color = finishing efficiency. The Premier League consistently combined the highest chance creation with strong results. Serie A and La Liga achieved similar points with fewer chances, while the Bundesliga generated high xG but underperformed. Ligue 1 trailed in both seasons.

Discussion

This research shows that chance creation, defensive solidity, and finishing efficiency best explain success across the continent's top five leagues. Goal difference, predictably, correlated nearly perfectly with Pts/MP, but more detailed analyses revealed consistent patterns that extend well beyond outcomes.

First, goal difference and expected goal difference (xGD) were the best predictors of success, each correlated extremely highly with Pts/MP. Moreover, metrics of chance creation like xG+xAG per 90 and G+A per 90 correlated strongly with performance, demonstrating that those teams that regularly created good quality chances best of all were set up to win. This confirms earlier research that associated success with passing sequence and shot creation to match outcomes.[2,8] This also agrees with earlier studies emphasizing expected goals as a measure of attacking strength, and these results confirm that combining xG with assists (xAG) provides added predictive value.[5,6]

Second, regression analyses revealed that defense was just as determinative. Lower expected goals against (xGA) throughout again predicted higher Pts/MP, affirming that elite clubs distinguish themselves not just by scoring but by restraining the quality of chances against. Liverpool (2024–25) and Leverkusen (2023–24) proved this balance, whereas struggling clubs like Sevilla and Manchester United failed to turn solid attacking metrics into results because of defensive vulnerability.

Third, the bespoke finishing efficiency was another important differentiator. Not the best individual predictor, but consistently accountable for why clubs overperformed or underperformed relative to their expectations. Real Madrid, Arsenal, and Girona overperformed through clinical finishing, and Burnley and Granada could not turn chance creation into points. Efficiency, while fluctuations were evident, appears important to distinguish between champions and mid-table clubs.[11]

Tactical clustering confirmed these results. High-possession progressors were the overwhelming winners of both seasons, repeatedly averaging above 2 Pts/MP. Direct counters fared poorly with an average of below 1 point per game. The persistence of these five stylistic clusters across two seasons reveals that the model defines stable rather than short-term tactics, as predicted by previous research demonstrating that possession is valuable depending on the situation, like whether a team is winning or losing by a goal, or facing a stronger or weaker opponent.[7,10,13]

Comparisons across leagues showed a significant style difference. The Premier League achieved the largest attacking volume and had the strongest outcomes. At the same time, Serie A and La Liga were efficient by taking fewer chances and turning out with similar results. The Bundesliga and Ligue 1, however, took many chances but failed to live up to the expected performances, and this hints at structural or tactics-based inefficiencies. Such inter-league variations therefore stress that the relationship between metrics and success depends on a wider set of competitive conditions.[2,9] The Premier League's historical dependence on long-ball methods may account for modest finishing efficiency despite high chance creation. The difference in aggregate roster talent probably impacts how successful some styles of play prove within leagues. The reduced success of deep-block survivalist teams, once characteristic of the style long associated with José Mourinho's clubs, indicates this approach grows less successful in the contemporary game. Further, leagues like the Bundesliga and Ligue 1 face talent concentration issues, since top strikers tend to move to powerhouses like Bayern or PSG and impact finishing efficiency and league competitiveness.

This study has some limitations. It is based on 98 team-season observations across two years and does not include player-level factors, injuries, or other game context variables. These factors could affect how broadly the results apply, and future studies could explore them with more detailed data. Future work could also include Champions League matches and player-level metrics to see how individual performance and additional competitions influence Pts/MP. Potential future research could also look at how team payroll or market value corresponds to Pts/MP to gauge the impact of player quality on success. Another area for potential research is whether teams whose goal distributions are more balanced also tend to be more consistently successful across a given season. Moreover, set-piece statistics like free-kick frequency and conversion rate could also be looked at to assess their impact on points outcomes overall. Taken together, all this would create a richer understanding of how economic, tactical, and player-level variables impact team performance across the top leagues in Europe.

DOI: 10.36838/IJHSR84.30

These findings indicate that elite soccer success can best be described by a mix of chance creation, defensive solidity, and clinical finishing. These metrics can help coaches, analysts, and recruitment departments set training priorities, evaluate team strengths and weaknesses, and inform tactical or player acquisition decisions. The results here serve both as an academic framework and as practical insights for those applying soccer analytics. They also demonstrate that robust analyses can be performed using widely available tools such as Microsoft Excel when used thoughtfully and with clear reasoning.

Conclusion

Analysis of Europe's premier five leagues over two seasons reveals that success is not based on possession or attacking output alone. The best predictors of success were expected goal difference (xGD), chance creation (xG+xAG/90 and G+A/90), defensive solidity (xGA), and finishing efficiency. These metrics help explain why elite clubs outperformed others, and why some teams either exceeded or fell short of their expected statistical outcomes. By using a simple, reproducible Excel-based framework, this study demonstrates that league-wide and tactics-based comparisons can be made. Although high-possession progressors inevitably came out on top, defensive and efficient sides in Serie A and La Liga also recorded excellent results. The findings provide evidence that success in modern football entails not only creating opportunities but also defending solidly and finishing clinically.

Acknowledgments

I would like to thank my advisor and AP Statistics teacher, Mr. Adam Peloquin, for his valuable guidance on this project. I am also grateful to Newton North High School for its academic support and encouragement, and to my family for their constant support.

References

1. Anderson, C.; Sally, D. *The Numbers Game: Why Everything You Know about Soccer Is Wrong*; Penguin Books: New York, 2013.
2. FBref.com. Team Statistics for Europe's Top Five Leagues, 2023–24 and 2024–25 Seasons. Retrieved August 2025. https://fbref.com/en/.
3. FBref.com. Expected Goals (xG) Explained. Retrieved August 2025. https://fbref.com/en/expected-goals-model-explained/.
4. MLSSoccer.com. Soccer Analytics 101: Expected Goals. Published April 16, 2020. https://www.mlssoccer.com/news/soccer-analytics-101 (accessed September 2025).
5. Mackenzie, R.; Cushion, C. Performance Analysis in Football: A Critical Review and Implications for Future Research. J. Sports Sci. 2013, 31 (6), 639–676. https://doi.org/10.1080/02640414.2012.746720.
6. Sarmento, H.; Marcelino, R.; Anguera, M. T.; Campaniço, J.; Matos, N.; Leitão, J. C. Match Analysis in Football: A Systematic Review. *J. Sports Sci.* 2014, *32* (20), 1831–1843. https://doi.org/10.1080/02640414.2014.898852.
7. Collet, C. The Possession Game? A Comparative Analysis of Ball Retention and Team Success in European and International Football, 2007–2010. *J. Sports Sci.* 2013, *31* (2), 123–136. https://doi.org/10.1080/02640414.2012.727455.
8. Rathke, A. An Examination of Expected Goals and Shot Efficiency in Soccer. *J. Hum. Sport Exerc*. 2017, *12* (2), 514–529. https://doi.org/10.14198/jhse.2017.122.15.
9. Liu, H.; Hopkins, W.; Gómez, M. Á. Modelling Relationships between Match Events and Match Outcome in Elite Football. *Eur. J. Sport Sci.* 2016, *16* (5), 516–525. https://doi.org/10.1080/17461391.2015.1042527.
10. Pollard, R.; Reep, C. Measuring the Effectiveness of Playing Strategies at Soccer. *Statistician* 1997, *46* (4), 541–550. https://www.jstor.org/stable/2988603.
11. Davis, J.; Robberechts, P. Biases in Expected Goals Models Confound Finishing Ability. *arXiv* 2024, arXiv:2401.09940. https://arxiv.org/abs/2401.09940.
12. Hughes, M.; Franks, I. Analysis of Passing Sequences, Shots and Goals in Soccer. *J. Sports Sci.* 2005, *23* (5), 509–514. https://doi.org/10.1080/02640410410001716779.
13. Lago-Peñas, C.; Dellal, A. Ball Possession Strategies in Elite Soccer According to the Evolution of the Match-Score: The Influence of Situational Variables. *J. Hum. Kinet*. 2010, *25* (1), 93–100. https://doi.org/10.2478/v10078-010-0036-z.

Author

Bishrelt "Giga" Bat-Erdene is a senior at Newton North High School. He is the varsity soccer captain and the founder of a student-led tutoring program. Recognized by NSTEM and NSHSS for academic excellence, he plans to pursue finance and data-driven research in college.