

Balancing pressing intensity and defensive solidity against different passing styles

Berta Plandolit López

Introduction

Pressing strategies and defensive formations in football have evolved significantly over the last decades from rudimentary defensive tactics into sophisticated, coordinated systems. From early man-marking systems and zonal defending to highly structured pressing and defensive formations, modern football has seen a shift in how teams defend. Pressing aims to regain possession, disrupt the opponent's play and create scoring opportunities, however, if not executed efficiently it possesses several risks.

High pressure involves applying defensive intensity in the opponent's half, close to their goal. This approach allows teams to engage in active defense and take advantage of shorter distances to goal after winning possession. However, it comes with significant vulnerabilities, such as leaving space behind for long balls, exposing the team to dangerous counterattacks and making it harder to recover from individual mistakes. Additionally, high pressing can be physically demanding, as it requires covering larger areas of the field.

Making strategic decisions between defensive formation or high-pressure systems is crucial to a team's overall success, as it directly influences their ability to control the game, maintain defensive stability and exploit opportunities in attack. The right choice depends not only on the team's tactical philosophy and player capabilities but also on the specific characteristics of the opposition, such as their style of play, the technical ability of their players and the context of the match. Balancing these factors effectively can be the difference between dominating a game or leaving dangerous gaps for the opponent to exploit.

This research will focus on the trade-offs between aggressive pressing tactics and maintaining defensive solidity, particularly when playing against teams with skilled on-ball defenders and midfielders. The hypothesis is that teams with technically proficient defenders and midfielders who are comfortable in possession pose challenges to high-pressing systems and that in such cases, the pressing team may find its efforts neutralized, as the opponent can resist the pressure more easily.

To better understand these dynamics, this study employs a two-step approach, starting with clustering teams and then using a combination of exploratory analysis and modeling. First, we analyze different types of defenders and their abilities with the ball, breaking down their technical skills and decision-making under pressure. Following this, we

investigate the effect of defensive pressure on goal-scoring opportunities, aiming to determine whether pressing effectively disrupts play and, if so, to what extent. We will also examine whether the response to pressure varies depending on a team's defenders' passing abilities, while identifying key variables that significantly influence this effect.

Therefore, the central question this paper aims to explore is whether high-intensity pressing offers a strategic advantage against teams with skilled defenders and midfielders, particularly when considering the potential risks against the possible rewards of employing such tactics. With the growing influence of data-driven tactics in modern football, clubs increasingly rely on football analytics to fine-tune strategies, making this study relevant not for coaches and tactical staff.

Related work

The interaction between pressing intensity and defensive solidity in football has been the subject of much attention in recent years. This section reviews related work that contributes to the understanding of this balance, focusing on tactical behaviors, psychological factors and the implications of pressing strategies.

One of the key areas of investigation has been the tactical behaviors exhibited by teams employing different pressing strategies. For instance, studies have shown that teams utilizing high-pressing tactics often experience a trade-off between regaining possession and maintaining defensive organization. The research by Carr et al. highlights how teams playing a 4-4-2 formation with a high-press strategy can cover significant distances at high intensity, which can lead to fatigue and diminished performance over the course of a match (Carr et al., 2020). This finding underscores the importance of conditioning and tactical awareness in executing pressing strategies effectively. Moreover, the analysis of tactical actions in women's football teams participating in the UEFA Champions League has revealed that successful teams often adapt their pressing intensity based on the opposition's playing style, further emphasizing the need for flexibility in tactical execution ("Analysis of tactical actions of women's football teams participating in the UEFA Champions League", 2023).

Psychological factors also play a crucial role in the effectiveness of pressing strategies. Stanimirovic and Hanrahan's work on psychological predictors of job performance in professional sports suggests that the mental resilience and focus of players can significantly influence their ability to execute high-pressure tactics (Stanimirovic & Hanrahan, 2010). The psychological demands of maintaining intensity during pressing situations can lead to cognitive overload, affecting decision-making and execution under pressure. This aligns with findings from Blaser, who notes that shared knowledge and communication within teams can enhance collective performance, thereby mitigating the risks associated with high-intensity pressing (Blaser, 2019).

In addition to tactical and psychological considerations, the physical demands of pressing strategies are worth the attention. Research indicates that high-intensity pressing can affect players' overall performance and injury risk. For example, studies have shown that players engaged in high-pressing tactics exhibit elevated heart rates and perceived exertion levels, which can compromise their ability to maintain focus and execute technical skills effectively (Low et al., 2018; Low et al., 2021). This is particularly relevant in the context of amateur football, where players may not possess the same level of conditioning as their professional counterparts, leading to a higher incidence of injuries related to high-intensity activities (Szymiski et al., 2021). Therefore, implementing effective training and recovery protocols is essential for teams that prioritize pressing as a core component of their tactical approach.

The relationship between pressing intensity and defensive solidity is further complicated by the characteristics of the opposition. Teams with skilled on-ball defenders and midfielders can exploit gaps left by overly aggressive pressing, utilizing their abilities to break defensive lines and create scoring opportunities. Research by Goh et al. emphasizes that a significant proportion of goals scored in open play result from quick passing sequences, highlighting the importance of maintaining defensive structure while pressing (Goh et al., 2022). This necessitates a careful evaluation of pressing strategies based on the strengths and weaknesses of opponents, as well as the specific context of the match. The ability to adapt pressing intensity in response to the opponent's passing style can be a decisive factor in determining the outcome of a match.

The integration of technology and data analytics into football has also transformed the analysis of pressing strategies. Recent advancements in machine learning and big data analysis have enabled researchers to model the relationships between tactical actions and match outcomes more effectively. For instance, the study by Ilevoli et al. demonstrates how passing network indicators can be utilized to model football outcomes, providing valuable insights into the tactical dynamics of pressing and defensive organization (Ilevoli et al., 2021). Such analytical frameworks can assist coaches in making informed decisions regarding pressing strategies, ultimately enhancing team performance.

More specifically, clustering emerges as a significant analytical tool to categorize players and teams based on various performance metrics, tactical behaviors and styles. Common clustering methodologies employed in football analysis include k-means clustering, hierarchical clustering and self-organizing maps.

In their study, Behravan et al. introduce an advanced big data clustering approach utilizing a particle swarm optimization algorithm to identify player roles in football. This method demonstrates how clustering can facilitate meaningful comparisons between

players and provide a deeper analysis of team performance. Similarly, in the realm of team clustering, Diquigiovanni and Scarpa (2018) developed an innovative hierarchical clustering method to examine the playing styles of Italian Serie A teams during the 2015-2016 season. By representing each team's playing style as a network, they clustered teams based on the similarity of their community structures. This approach enabled the identification of prevalent tactics and their influence on match outcomes, offering a nuanced understanding of strategic variations among teams.

Clustering teams according to their defensive players

The primary objective of this section is to assess LaLiga teams by clustering them according to the on-ball abilities of their defenders. This step serves as a foundation for understanding how teams with different defensive skill sets respond to high-pressure situations.

We will use key metrics that reflect the defender's passing style in possession, such as their passing accuracy, range and performance under pressure. By grouping teams into distinct clusters, we can identify similarities and differences between these groups and explore whether teams with more skilled defenders behave differently when faced with defensive pressure.

The insights from this clustering analysis will help determine if certain defensive profiles are better suited to withstand pressing tactics, ultimately informing the study of how different teams handle pressure and create goal-scoring opportunities.

3.1 Data Collection and Preprocessing

The data used for this analysis is from the 2023-2024 LaLiga regular season. The dataset encompasses a total of 290 matches, including complete data for 10 teams -Rayo Vallecano, Real Madrid, Getafe, Valencia, Atlético de Madrid, Athletic Club, Cádiz, Mallorca, Barcelona and Sevilla. For the remaining 10 teams -Osasuna, Almería, Girona, Real Sociedad, Real Betis, Villarreal, Celta de Vigo, Granada, Deportivo Alavés and Las Palmas- data is only available for their matches against the selected 10 teams.

The data, provided by StatsBomb, includes both event data and 360 frames, giving a detailed view of players' actions and positioning during the matches.

Since the goal of this section is to analyze a team's performance on the defensive side of the field, the possessions have been filtered in order to include only those that start from the goalkeeper. This approach allows for a more controlled study of how teams build from the back. From these possessions, we extracted the events related to passing that took place in the defensive and midfield positions, specifically covering the spaces from 1 to 16, as shown in Figure 1.



Figure 1: Tactical Positions Guide. Source: StatsBomb Events Documentation, Appendix 1.

3.2 Key Metrics For Evaluation

The following metrics were calculated under the conditions outlined in the previous subsection. These metrics were selected to evaluate how accurate and skilled the defenders are with the ball and capture a comprehensive picture of the team's passing style. We focused specifically on passing ability, as it is a key component in responding to and breaking down high-pressure situations, making it a strong indicator of defensive skill.

To account for the varying number of matches available for each team in the dataset, all metrics have been averaged per game.

- **Total passes per game:** the average number of passes attempted by the defenders in each game. This metric provides an overall view of the team's passing frequency and involvement in ball circulation from the back.
- **Success Rate:** the proportion of the passes that successfully reached the intended target. This metric assesses the accuracy and reliability of the defenders.
- **Goalkeeper passes per game:** the average number of passes made by the goalkeeper. This metric helps to understand the team's playing style, whether the keeper is heavily involved in building the play.

- **Passes under pressure per game:** the number of passes attempted by defenders while under pressure from the opponent. This metric uses the StatsBomb feature *under_pressure*, which identifies events where the ball carrier is actively being pressured by an opponent.
- **Success rate under pressure:** the proportion of passes completed successfully under pressure. This metric reflects how accurate defenders pass under intense pressing from the opponent.
- **Average length and width of passes:** the average forward (length) and horizontal (width) distance of passes made by defenders. These metrics indicate the team's tendency to either play direct, forward passes or distribute the ball laterally across the field.

3.3 Clustering methodology

We employed K-means clustering as the chosen methodology. K-means is well-suited for this problem because it effectively partitions teams into distinct groups based on their defensive passing abilities. It is also a versatile method that provides flexibility in selecting the number of clusters, allowing the balance between meaningful insights from the data and a more arbitrary choice that considers the specific goals of the project.

The elbow method was used to select the optimal number of clusters, helping to balance interpretability with variance explanation. A range of cluster numbers were tested and measured the inertia reduction - which is the sum of squared distances between each data point and its cluster center. In Figure 2, we can observe that inertia drops significantly as we increase the number of clusters up to 4, after which the reduction becomes smaller. Given the goal of the project and the fact that the dataset contains 20 team, selecting 4 clusters seems to be an appropriate choice for this analysis.

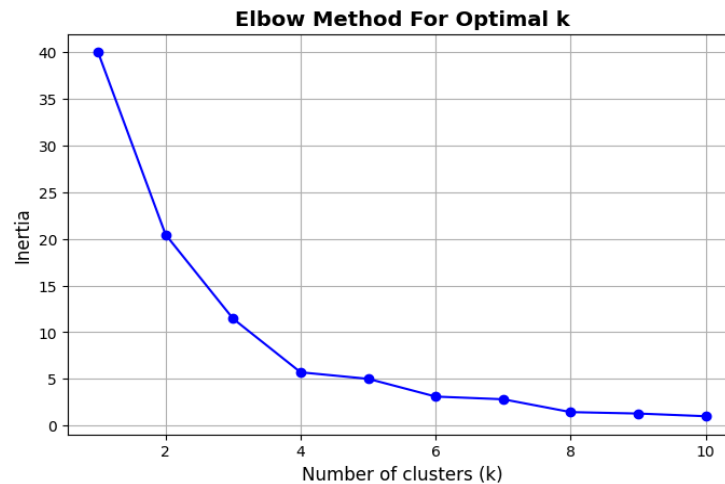


Figure 2: Inertia reduction curve for determining the optimal number of clusters using the elbow method.

The results of the clustering can be visualized in Figure 2, which shows a Principal Component Analysis (PCA) plot of the clusters. The PCA is used to reduce the dimensionality of the data and visualize the clusters in a 2D space, providing a clearer picture of how teams with similar defensive passing abilities group together. In this reduced dimensional space, we observe that the clusters are well-separated, with each cluster occupying approximately one quadrant.

Together with the PCA, on the right side of the figure, we include the final La Liga 2023-2024 table and Table 1 lists the teams in each cluster, color-coded according to the clusters shown in the PCA plot. Analyzing these clusters in relation to league performance, we observe that the green cluster includes 3 out of 4 top teams in the league. In contrast, the yellow cluster -the largest in terms of team count- contains more teams that finished in the lower parts of the league table. The remaining two clusters, purple and blue, have a more varied spread, with teams distributed across the ranking.

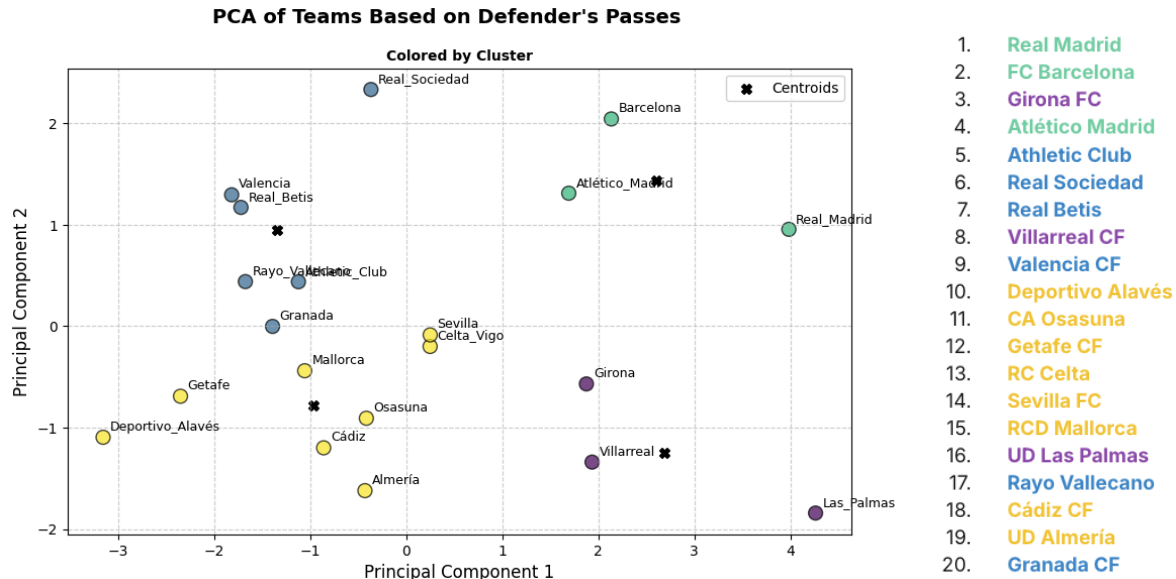


Figure 3: PCA of teams based on defenders' passing style. Each cluster is represented by a distinct color, with black crosses marking the centroids. On the right, La Liga 2023-2024 final standings, each team colored by cluster

Figure 4 illustrates the feature contributions to the principal components. This chart shows the impact of each metric on the two principal components, helping us understand which features most influence the clustering results.

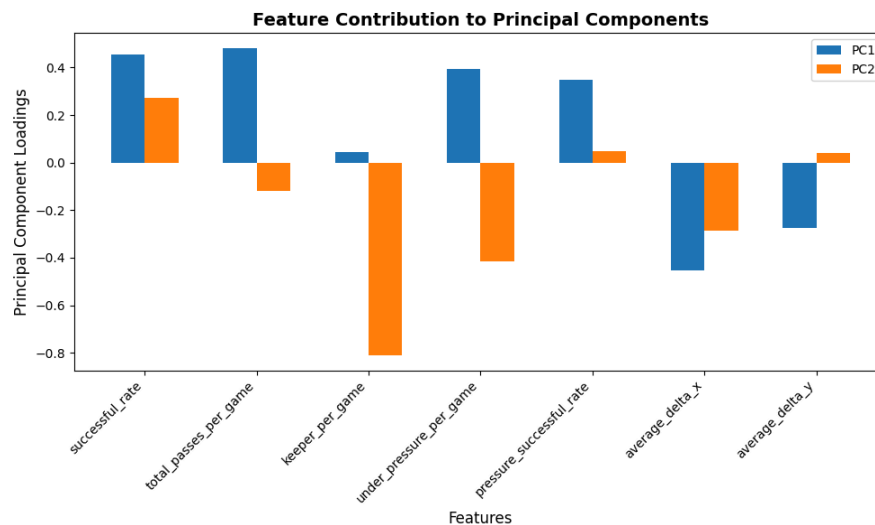


Figure 4: Feature Contributions to Principal Components.

The horizontal axis in the PCA plot (principal component 1) is influenced primarily by success rates and total passes per game and the number of passes attempted under pressure. Teams on the right side of the plot tend to have defenders who complete a higher volume of passes and do so with greater accuracy. The vertical axis (principal

component 2) is most strongly influenced by the number of passes the goalkeeper makes per game, meaning that teams located toward the top part of the plot involve their goalkeeper more frequently in play.

Table 1: Teams in each cluster.

Green	Barcelona, Real Madrid, Atlético de Madrid
Purple	Girona, Villarreal, Las Palmas
Yellow	Almería, Celta de Vigo, Cádiz, Deportivo Alavés, Getafe, Mallorca, Osasuna, Sevilla
Blue	Athletic Club, Granada, Rayo Vallecano, Real Betis, Real Sociedad

3.4 Results

While certain features contribute more heavily to each principal component, all calculated metrics play a role in defining the clusters. This means that the clustering does not depend solely on any single metric but rather on an interaction of all metrics, where each feature adds value to the separation of teams.

Figure 5 illustrates the relationship between success rate and total passes per game, where we can observe that the green and purple clusters contain the teams that attempt the highest number of passes per game in the defensive field, with also the highest accuracy. This indicates a preference for controlled ball distribution among these teams and precision to maintain possession. While in the other hand, teams with lower passing volumes or success rates, might adopt a more direct style of play or experience quicker turnovers in possession.

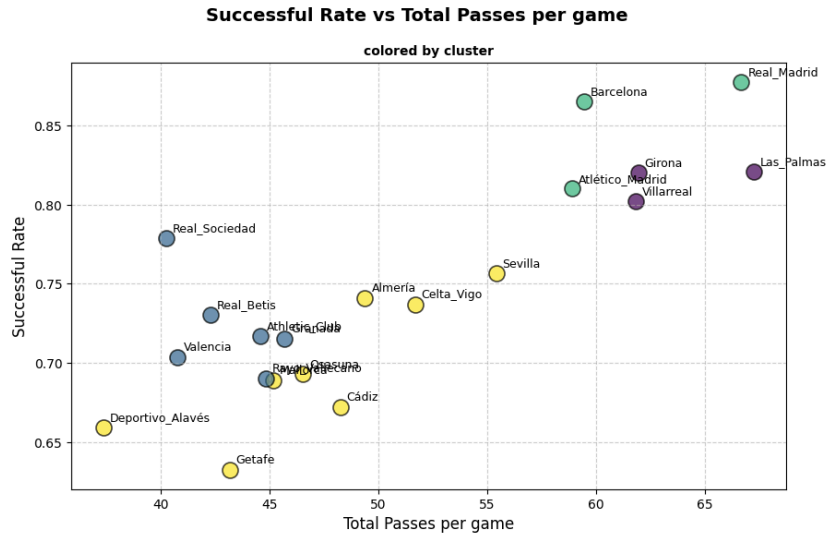


Figure 5: Distribution of team's Passing success rate vs Total passes per game, colored by cluster.

In Figure 6, we can observe how clusters are also clearly divided by the average number of passes made by the goalkeeper per game. Teams in the yellow and purple rely on their goalkeeper more frequently, with over 13 times per game on average. This tendency suggests a style that either integrates the goalkeeper as an active participant in ball distribution or involves receding to the goalkeeper to reset play under pressure. In contrast, teams in the blue and green clusters utilize their goalkeeper less frequently in this role, indicating alternative approaches where defenders or midfielders handle distribution more directly, possibly to maintain a quicker tempo or reduce risk.

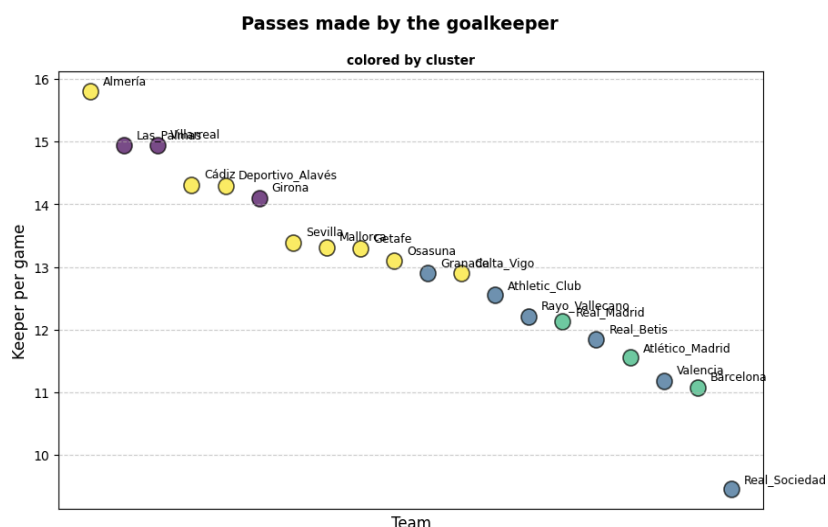


Figure 6: Distribution of team's average number of passes made by goalkeeper per game.

In Figure 7, we explore how teams manage passes under pressure by examining the relationship between the number of passes attempted under pressure per game and the

successful rate of those passes. While the distinctions between clusters are less pronounced in this plot, the green cluster stands out for having the highest passing accuracy under pressure, with Real Madrid showing the highest success rate among all teams. Additionally, Las Palmas in the purple cluster is notable on the far right side of the plot. This team attempts the most passes under pressure per game with a quite good accuracy.

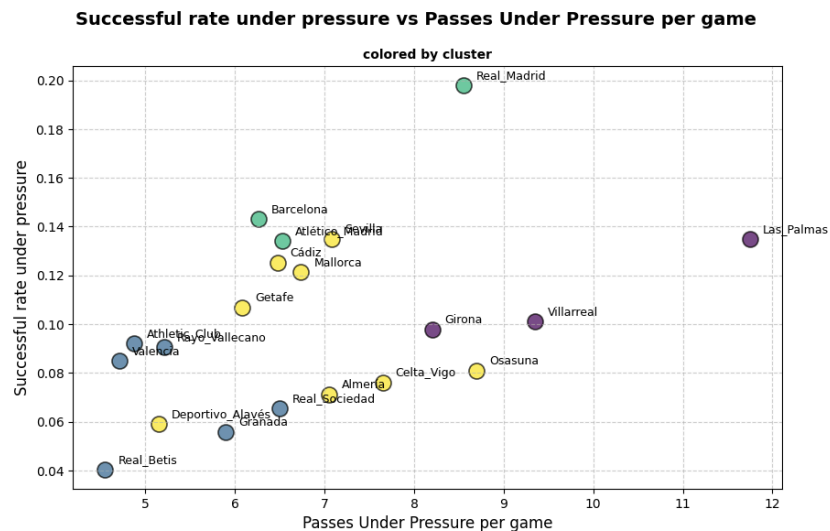


Figure 7: Distribution of teams' success rate under pressure vs average number of passes made under pressure per game.

In Figure 8, we can analyze teams based on their average pass length and direction by examining the Average Delta X (horizontal pass distance) and Average Delta Y (vertical pass distance). The yellow cluster, which includes teams like Getafe and Deportivo Alavés, show the highest values for both Average Delta X and Delta Y, meaning that these teams tend to make longer passes both horizontally and vertically. On the other hand, the green cluster appears towards the left side of the plot, indicating more lateral passing.

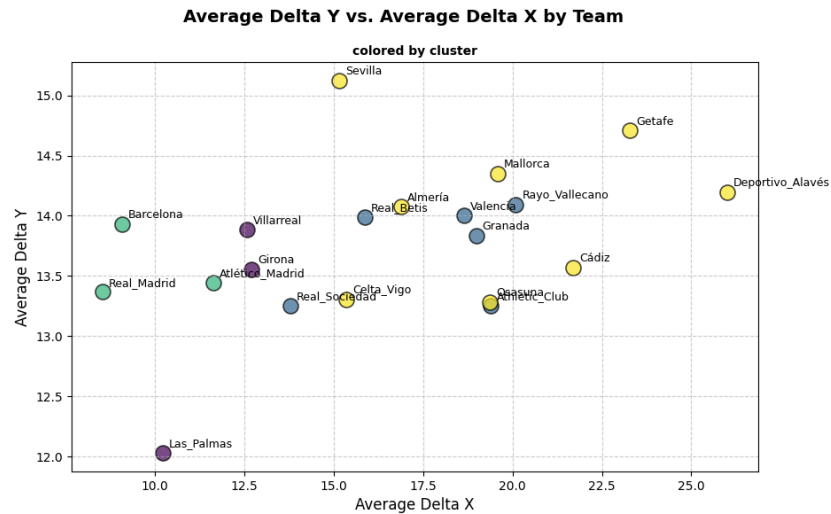


Figure 8: Distribution of team's average passing length and width.

Overall, this clustering analysis underscores the strategic diversity among La Liga teams in terms of defensive passing styles. It sets the stage for further investigation into how these clusters interact with pressing tactics and for understanding of team strengths and potential vulnerabilities.

Study of the Effect of Pressure on Goal Opportunities

This section explores the relationship between defensive pressure and the creation of goal opportunities in LaLiga matches during the 2023.2024 season. Pressing in football is widely known as a method for disrupting the opponent's play and win the ball back or delay their offensive actions. However, it's efficiency and the trade-off between significantly affecting the opponent's gameplay and the potential risks exposed, requires further exploration.

The objective of this section is to quantify the impact of defensive pressure on the likelihood of creating goal-scoring opportunities. Additionally, we'll investigate whether this effect varies across clusters of teams based on their defensive profiles, as identified in the previous section. By doing so, we aim to discover if any certain cluster has an easier time breaking pressure or whether their gameplay remains unaffected when closely defended.

To conduct this study, we employ a two-step approach:

1. **Exploratory analysis:** we begin by visually assessing goal opportunities and associated across possessions, stratified by whether or not high defensive pressure was applied.

2. **Modeling pressure's impact:** we then use statistical techniques to isolate the effect of pressure while controlling for confounding variables such as game location, team rankings and match context. Additionally, we examine if this effect differs across clusters.

By combining these two, this analysis serves as a foundational step for further studies, that can offer valuable insights for coaches and analysts looking to refine pressing strategies.

3.1 Data collection and preprocessing

The data used for this study is the same as in the last section, comprising the 2023-2024 La Liga season. However, the scope of possessions considered here has been refined to better fit the objectives of analyzing the effect of defensive pressure.

In this case, the possessions selected were those that begin in the defensive field while excluding counterattacks, to make sure we only focus on structured buildup plays rather than transitional moment, when other unmeasured latent factors might confound the effect. Additionally, possessions that happened after a red card have also been excluded, to avoid scenarios where the dynamics of the match are heavily skewed due to an imbalance in players on the field.

3.2 Metrics

The metrics used for both the exploratory analysis and the modeling are categorized into three groups: outcomes, pressure measurements and control variables. Each metric was carefully chosen and tested to provide insights into how defensive pressure impacts goal-scoring opportunities while accounting for confounding factors.

The **outcomes** represent the performance indicators used to evaluate the effectiveness of possessions under defensive pressure:

- **Goal Opportunities:** defined as the times an attacker enters the opponent's penalty area, as highlighted in Figure 9. This was selected to reflect how well a team manages to break the pressure, moving the ball into dangerous attacking positions.
- **Expected Goals (xG):** measures the quality of scoring chances after a shot is taken. This metric is calculated with the descriptive variable `statsbomb_xg` from the event type `Shot` provided by StatsBomb. (Source: [StatsBomb xG](#)).

- **On-Ball Value (OBV):** captures the overall effectiveness of the possession by adding up the probabilities of positive outcomes by the end of the possession. (Source: [StatsBomb OBV](#))

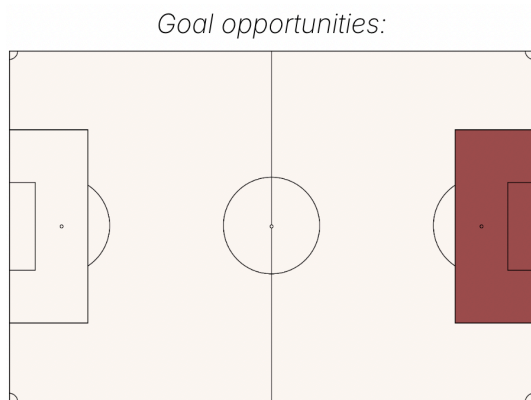


Figure 9: Goal opportunities representation.

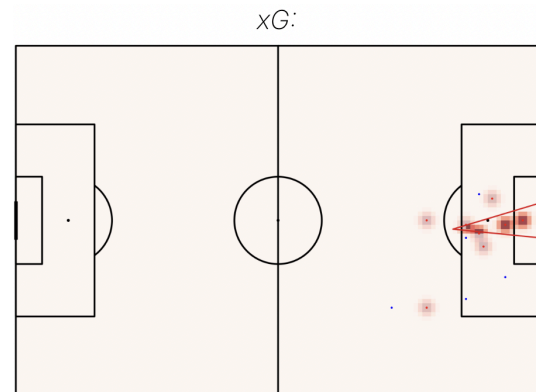


Figure 10: Expected goals representation. Source: <https://statsbomb.com/articles/soccer/upgrading-expected-goals/>

The following **pressure measurements** quantify the level of defensive intensity applied by the opposing team.

- **Distance to nearest defender:** the distance between the player in possession of the ball and the closest defender in 360-frame. Lower distances indicate higher pressure levels.
- **Number of defenders on the goal side of actor:** counts the number of opposing defenders in the defensive midfield of the attacking team.
- **Sum of distance to nearest defenders:** calculates the total distance from the every teammate to their closest defenders in one 360 frame. This metric reflects how collectively the defensive team pressures the ball carrier and nearby players.
- **Overlap Convex Hull:** Represents the proportion of the overlapping area between the defending team's convex hull and the attacking team's convex hull. A higher overlap suggests tighter spatial pressure applied by the defensive team, see Figure 11.

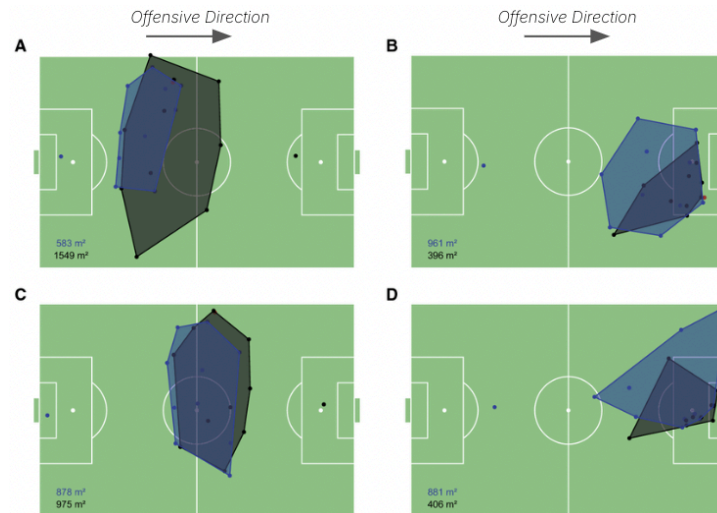


Figure 11: Representation of Overlap Convex Hull.

Control variables were incorporated to address potential confounding effects, ensuring that the influence of defensive pressure was properly isolated:

- **Difference in Rankings:** The difference between the rankings of the two teams at the end of the season. A positive value indicates that the team under analysis is ranked higher than its opponent.
- **Game Conditions:** Including factors such as whether the team is playing at home or the number of spectators present during the match.
- **Clusters:** represents the defensive profiles of teams as identified in the clustering analysis. This variables allows for an exploration of whether different defensive profiles respond differently to pressure.

3.3 Exploratory Analysis on a Possession-Level

The exploratory analysis provides an initial assessment of how defensive pressure affects the measured outcomes (Goal Opportunities, Expected Goals and On-Ball Value). By categorizing possessions into two groups, those under pressure and those not under pressure, the outcomes are compared to uncover potential differences and patterns in performance.

Possessions were labeled as "under pressure" based on the following predefined thresholds for pressure metrics, which were determined based on the range of metric values and informed by quantile analysis:

- Distance to Nearest Defender: Less than 3 units.

- Overlap Convex Hull: Greater than 50%.

Visualizations and tests were created to understand whether possessions under pressure consistently result in fewer goal-scoring opportunities and lower overall effectiveness compared to those without pressure.

In the first stage, we compared **all the possessions** in the dataset split between whether they had been under pressure or not. As shown in Table 2, the outcomes were consistently higher for possessions without pressure. Statistical testing confirmed that the differences between the two groups were statistically significant across all metrics, highlighting the clear impact of defensive pressure on possession outcomes.

Table 2: Comparison of possession outcomes under pressure and without pressure.

Metric	Pressure applied	No pressure	Significant difference
Goal Opportunity rate	8.88%	14.55%	yes
xG per Possession x100	0.51	0.81	yes
OBV per Possession x100	12.17	23.2	yes

For xG and OBV, the Shapiro-Wilk test indicated that these metrics are not normally distributed under either condition. Therefore, the Mann-Whitney U test, a non-parametric alternative, was used to compare distributions. The resulting p-values were approximately zero, strongly supporting the conclusion that the differences are statistically significant.

For Goal Opportunity rate, which represents categorical data (success vs. no success), a Chi-square test was applied. The test yielded a p-value below 0.05, indicating that the observed differences in Goal Opportunity rates between pressured and unpressured possessions are unlikely to have occurred by chance.

Expanding on this analysis, Figure 12 illustrates the variation in the outcomes under pressure across the four different clusters. We can observe that pressure consistently reduces performance across clusters, but the degree differs among them.

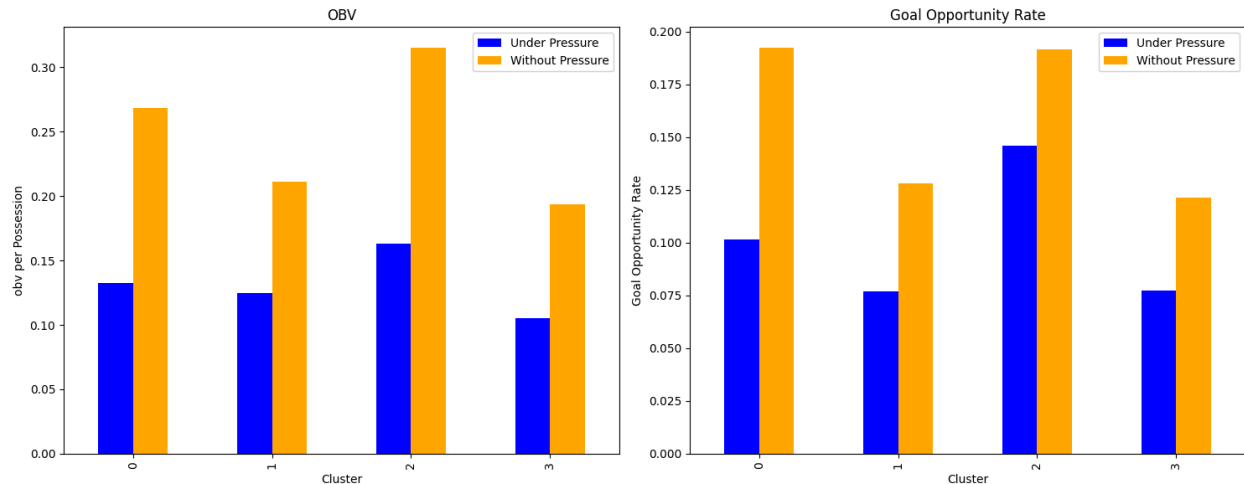


Figure 12: Impact of defensive pressure on performance metrics (OBV and Goal Opportunity rate) across clusters.

With these results, we cannot identify any consistent conclusions or patterns across clusters, highlighting that the impact of pressure on performance to be team-specific. Therefore, Figure 13 provides a comparison of the goal opportunity rates under pressure and without pressure for each team.

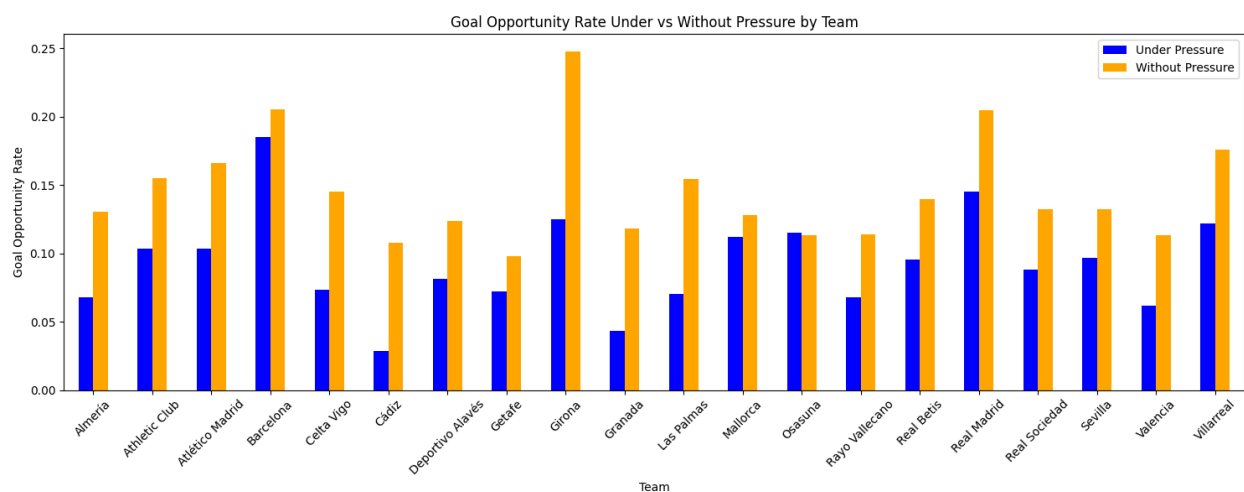


Figure 13: Comparison of goal opportunity rates under pressure and without pressure for each team.

Here, we can draw some interesting observations. Barcelona maintains a relatively stable performance regardless of whether they are under pressure or not, showcasing their adaptability and resilience in pressured scenarios. Girona shows the largest drop in goal opportunity rates under pressure, suggesting they may struggle to cope with defensive intensity. However, without pressure, they record the highest goal opportunities in the

league. And, interestingly, Osasuna stands out as the only team that performs slightly better under pressure.

Across the three analyses, possessions under pressure consistently showed diminished performance compared to those without pressure. These findings reaffirm the effectiveness of applying defensive pressure to disrupt opponent play and limit goal-scoring potential. When analyzing clusters, no definitive patterns emerged to suggest that any defensive profile is inherently more resistant to pressure. Instead, the team-specific results revealed nuanced behaviors.

3.4 Modeling pressure's impact on performance

To quantify the relationship between defensive pressure and team performance, the data was aggregated at a period level per game, separating it into first and second halves. To isolate the effect of pressure while accounting for potential confounding factors, we employed two statistical modeling approaches:

- Linear Regression Model: used to estimate the average effect of pressure on outcomes while controlling for variables like team rankings, game context and clustering groups.
- Mixed Linear Model: this model incorporates random effects to account for cluster-level variability. By doing so, we aim to understand if the relationship between pressure and performance differs significantly across clusters.

3.4.1 Linear Regression Model

In this first model a linear regression was employed, with On-Ball Value (OBV) as the dependent variable. Since it's a period-level analysis, the OBV was aggregated as an average across all possessions within a specific game period. The primary independent variable of interest was defensive pressure, measured through the four pressure metrics outlined in Section 3.2: Distance to Nearest Defender, Number of Defenders on the Goal Side, Sum of Distances to Defenders and Overlap Convex Hull. The model also incorporated clusters (representing defensive profiles of teams), the difference in ranking and interaction terms to capture nuanced effects.

The regression model was defined as follows:

$$Performance\ Metric = \beta_0 + \beta_1\ Pressure + \beta_2\ Cluster + \beta_3\ (Pressure \times Cluster) + \epsilon$$

This formulation allows the model to estimate the overall effect of pressure on performance outcomes while considering potential differences between clusters and other relevant variables.

The results of the linear regression model revealed a general trend: as defensive pressure increases, the expected performance tends to decrease. These findings support the previous hypothesis that high pressure disrupts offensive play, making it harder for teams to get more goal opportunities, or, in terms of OBV, having on average higher probability of scoring by the end of possessions. However, most individual pressure metrics and their interaction terms with clusters were not statistically significant, meaning that these variables do not have a clear or consistent effect on OBV in isolation.

One of the more notable findings was the statistically significant impact of the difference in ranking between teams. Higher-ranked teams tend to generate slightly higher OBV, highlighting the importance of opponent strength as a critical factor in performance outcomes. Teams with greater technical and tactical capabilities are better equipped to mitigate the disruptive effects of defensive pressure. Despite these insights, the model's R-squared value of 0.124 indicates that only a small proportion of the variance in OBV is explained by the variables included in this model. This highlights the need to account for additional factors that may influence performance, such as player-specific attributes, tactical adjustments or match-specific dynamics.

3.4.2 Mixed Linear Model

In this second model, a mixed linear model was employed, also focusing on OBV as the dependent variable. Mixed models are useful when data have hierarchical or grouped structures, such as repeated measures data, where some observations are related or nested within groups. Mixed models account for both the overall trend (fixed effects) and the variability between or within these groups (random effects). Therefore this model extends the linear regression framework by including random effects to account for potential variability between clusters of teams.

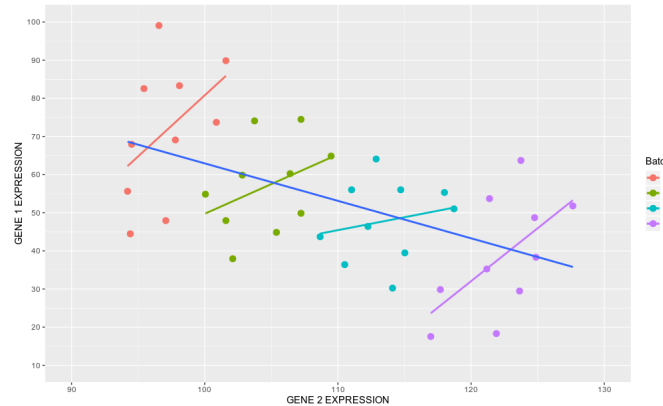


Figure 14: Illustration of Mixed Linear Model (MLM) vs. Linear Regression. The linear regression is represented by the overall trend blue line, while the MLM captures variability across four distinct batches (colored lines).

Source: [towardsscience.com, How Linear Mixed Models Work](https://towardsscience.com/How-Linear-Mixed-Models-Work/).

The independent variables in this model include the four pressure metrics outlined in Section 3.2 as well as the difference in team rankings. By incorporating clusters as random effects, the model will capture whether the defensive profiles influence how they respond to pressure.

The results of the mixed linear model showed the same general trend as the previous linear model: when defensive pressure is higher, the probability of scoring of the offensive team tends to decrease. Additionally, as observed in the linear regression model, the difference in ranking between teams was again a statistically significant predictor.

Interestingly, the random effects for clusters were not statistically significant. This suggests that the defensive capabilities of teams, as captured by the clusters, do not significantly alter how pressure impacts their OBV. In other words, teams grouped by their defensive profiles do not exhibit different responses to defensive pressure in terms of expected goals, at least in this scenario. The lack of significant random effects suggests that other factors, possibly related to individual player attributes, tactical nuances, or match-specific conditions, may play a more prominent role in shaping how teams respond to defensive pressure.

3.5 Results

This section presents the key findings from the analysis, which was conducted through a multi-step approach combining clustering, exploratory data analysis and statistical modeling.

- Defensive pressure reduces success per possession: across all metrics analyzed, including expected goals (xG) and goal opportunity rate and On-Ball Value (OBV), the results consistently indicate a negative effect of defensive pressure. When teams face higher pressure, their offensive performance, as measured by these metrics, tends to decrease. This reinforces the hypothesis that defensive pressure in deed disrupts offensive play, making it more difficult for teams to create and convert opportunities.
- No significant differences between clusters: despite grouping teams based on their defensive capabilities into clusters, no substantial differences were observed in how these clusters respond to pressure. This suggests that either responses to defensive pressure cannot be generalized across the defined clusters responses to defensive pressure may be highly team-specific and cannot be generalized across the defined clusters, or at least, not in the clusters studied in this research. Each team may possess unique tactical or technical adaptations to cope with pressure.
- Team-specific tends: as observed in the visual analysis, FC Barcelona displayed remarkable consistency, maintaining stable performance levels regardless of whether they were under pressure or not. In contrast, Girona was the most affected team, showing the largest drop in goal opportunity rates under pressure.
- Opponent strength matters: the difference in rankings between opposing teams repeatedly showed statistical significance. Higher-ranked teams perform better and generate more opportunities when facing lower-ranked teams. This indicates that stronger teams are better at attacking and breaking the pressure.
- Insufficient controls and missing variables: the lack of significant factors throughout the analysis revealed that the initial assumptions and selected variables were likely incomplete or insufficient for capturing the full picture. Leaving out variables such as player-level attributes or possession durations might have confounded the effects of pressure, leaving key aspects of team performance unexplained.

Limitations

This study has provided valuable insights into the relationship between defensive pressure and team performance; however, several limitations must be acknowledged:

- Cluster definition: the clustering process involved using the elbow method to determine the optimal number of clusters (ranging from 1 to 20). While this method is widely accepted, it inherently involves a degree of subjectivity in identifying the

point at which adding more clusters yields diminishing returns. This subjective element may have influenced the granularity and interpretability of the clustering results.

- Pressure quantification: several metrics were explored to quantify defensive pressure and while they capture key aspects of pressure, the thresholds used to define "under pressure" were somewhat arbitrary. Specifically, thresholds were determined based on quantiles, which, although practical, may not fully reflect the nuances of real-game scenarios. This approach could have introduced bias into the classification of possessions as pressured or unpressured.
- Incompleteness of data: the dataset used in this study was limited by camera vision and event-tracking technology, which may have influenced the quality and accuracy of pressure metrics. Additionally, complete data was only available for 10 teams, while the remaining 10 teams were only included for their matches against the selected 10. This limited scope reduces the generalizability of the findings across all La Liga teams.
- Omission of possession duration: the study did not account for the length of possessions (time), a factor that likely plays a significant role in determining a team's ability to respond to pressure, which may have left an important aspect underexplored.

These factors should be considered and carefully addressed in future research.

Conclusions and Future Work

This study offers an examination of the influence of defensive pressure on team performance, with a particular focus on teams featuring skilled on-ball defenders and midfielders. The overall goal of this research was to equip technical staff with data-driven insights to adjust strategies effectively against different types of teams. By clustering teams based on their defensive passing styles, we aimed to profile teams into different groups and understanding their unique characteristics. The second part of the study involved a deeper analysis of how high pressure affects the offensive team's scoring potential and whether different clusters exhibit varying responses to such pressure.

Through the clustering process and the analysis of critical metrics such as xG, goal opportunity rates and On-Ball Value (OBV), several conclusions were drawn. These findings have also informed potential directions for future research, ensuring that the insights gained can serve as a foundation for refining tactical approaches and addressing existing gaps in the analysis.

The analysis consistently demonstrated that while the impact of defensive pressure on team performance is evident, it does not differ by cluster. Across all subsections, the results showed how defensive pressure reduces team performance metrics. However, as further reinforced in Section 3.3, responses to pressure appear to be team-specific, with no significant differences observed across the defined clusters. This suggests that the current clustering approach, based on mainly defenders passing style, might not fully capture the ways teams handle pressure. Future studies could refine clustering methods by incorporating additional metrics or leveraging individual player-level knowledge to create more precise groupings.

Moreover, we have also consistently seen the importance of the opponent and their final ranking position. Teams ranked higher than their opponent perform better and generated more opportunities when under pressure. This underscores the importance of considering opponent strength in any analysis of team performance and pressure. Stronger teams likely have better tactical systems and player capabilities to cope with defensive intensity, which further impacts their ability to create scoring opportunities.

The models used -Linear Regression and Mixed Linear Model- while informative, have limitations in terms of the variables considered and the variability captured, most of the variables analyzed were not statistically significant. Incorporating factors -such as individual player attributes, possession durations and team formations- in future models could significantly improve the ability to fully explain and provide a more complete understanding of the dynamics at play.

To uncover broader trends, future work could expand the analysis to include data from multiple seasons or leagues. This would help identify consistent patterns across different contexts and reduce potential biases from focusing on a single season. Alternatively, narrowing the focus to a single team with richer, player-level data could provide deeper insights into the specific mechanisms through which teams handle defensive pressure and adapt their tactics.

In conclusion, while this study demonstrates the clear impact of defensive pressure on performance, it also highlights the complexity of team-specific responses and the need for richer data and advanced modeling techniques to uncover deeper insights into pressing strategies in football. Additionally, the findings suggest that the current clustering approach may not fully capture the nuanced ways teams handle pressure,

indicating that either teams cannot be effectively grouped into clusters based solely on defensive metrics or that a different type of clustering—potentially incorporating additional or alternative variables—would be required. These findings lay the groundwork for future research that can better inform tactical decision-making and performance analysis.

Acknowledgments

Marc Garnica Caparrós from the German Sport University Cologne for his invaluable support.

Data was provided by StatsBomb.

References

Karpa, Ihor & Ripak, Ihor & Lobasyuk, Vitaliy & Shanta, Ivan. (2023). Analysis of tactical actions of women's football teams participating in the UEFA Champions League. *Sports games*. 4(30). 14-23. <https://doi.org/10.15391/si.2023-4.02>

Behravan, I., Zahiri, S. H., Razavi, S. M., & Trasarti, R. (2019). Finding roles of players in football using automatic particle swarm optimization-clustering algorithm. *Big Data*, 7(1), 35–56. <https://doi.org/10.1089/big.2018.0069>

Blaser, M. (2019). Shared knowledge and verbal communication in football: changes in team cognition through collective training. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00077>

Carr, R., Mullen, R., & Williams, M. (2020). Differences in high intensity running when playing a 4-4-2 formation with a high press strategy: a case study from the english championship. *International Sports Studies*, 42(2), 21-31. <https://doi.org/10.30819/iss.42-2.03>

Diquigiovanni, J., & Scarpa, B. (2018). Analysis of association football playing styles: An innovative method to cluster networks. *Statistical Modelling*, 19(1), 28–54. <https://doi.org/10.1177/1471082X18808628>

Goh, A., Drinkwater, E., Harms, C., Scanlan, M., Newton, R., & Ma'ayah, F. (2022). Characteristics of goals scored in open play at the 2017 and 2018 australian national cerebral palsy football championship. *International Journal of Sports Science & Coaching*, 18(3), 858-866. <https://doi.org/10.1177/17479541221095941>

levoli, R., Gardini, A., & Palazzo, L. (2021). The role of passing network indicators in modeling football outcomes: an application using bayesian hierarchical models. *Asta*

Advances in Statistical Analysis, 107(1-2), 153-175.
<https://doi.org/10.1007/s10182-021-00411-x>

Stanimirovic, R. and Hanrahan, S. (2010). Psychological predictors of job performance and career success in professional sport. *Sport Science Review*, 19(1-2), 211-239.
<https://doi.org/10.2478/v10237-011-0013-z>

Szymiski, Dominik & Opitz, Sabine & Pfeifer, Christian & Rupp, Markus & Angele, Peter & Alt, Volker & Krutsch, Werner & Krutsch, Volker. (2021). High injury rates and weak injury prevention strategies in football referees at all levels of play. *Scandinavian Journal of Medicine & Science in Sports*. 32. 10.1111/sms.14083.

YILDIRIM, S. (2023). Investigation of the relationship between the tactical skills used by football players and their maximum performance levels. *Turkish Journal of Sport and Exercise*, 25(3), 472-481. <https://doi.org/10.15314/tsed.1363117>