

THANOS PAIDOULIAS

FOOTBALL PENALTY ANALYSIS

Contents

1	Analysis.....	1
1.1	Comparison of penalty conversion rates between men and women	1
1.1.1	Chi-square Method	1
1.1.2	Using Cramer's V	2
1.1.3	Bootstrap method	3
1.1.4	Final thoughts.....	4
1.2	Best five penalty takers in the dataset.....	5
1.2.1	Bootstrap approach.....	5
1.2.2	Bayesian approach	6
1.2.3	Custom metric approach.....	7
1.2.4	Penalty Taking Skill, combining Bayesian approach with Custom Metric.....	8
1.2.5	Conclusion	10
1.3	Strategic Selection: Ordering the First Five Penalty Shootout Takers.....	11
1.3.1	Literature review.....	11
1.3.2	Comprehensive Factors Influencing Penalty Taker Selection	12
1.3.3	Prioritizing Penalty Taker Selection: Strategic Factors.....	12
1.3.4	Strategic Selection of Shootout Penalty Takers: Analyzing Manchester United's 2020/2021 Season Data.....	13
1.3.5	Summary and limitations	15
2	References	16

List of Figures:

Figure 1 Conversion rates by gender	1
Figure 2 Count of penalties made and missed by gender.	2
Figure 3 Chi square test results table. P-value < 5%.....	2
Figure 4 Cramer's V summary table. Cramer's V close to 0 indicates not significant difference.	3
Figure 5 Bootstrap results. Very small mean difference and confidence interval contains also 0.....	4
Figure 6 Top 10 penalty takers using Bootstrap method.	6
Figure 7 Top 10 penalty takers using Bayesian method.	7
Figure 8 Top 10 penalty takers using custom metric for identifying penalty taking skill.	8
Figure 9 Penalty Taking Skill. A combination of Bayesian approach and Custom Metric. Top 5 players, full dataset.	9
Figure 10 Penalty Taking Skill. A combination of Bayesian approach and Custom Metric. Top 5 players, 99th percentile in terms of total attempts.....	9
Figure 11 Manchester United season 2020/2021, candidate penalty takers for shootouts.....	14
Figure 12 Manchester United Season 2020/2021 Penalty Shootout Strategy: Analytical Ranking of Top 5 Takers Based on Comprehensive Performance Metrics.....	14

1 Analysis

In this comprehensive analysis, we delve into the intricate world of soccer match data to uncover patterns and insights that shed light on suspected match irregularities. Through meticulous data manipulation and visualization, this report aims to not only identify matches with unusual outcomes but also to explore the broader trends and dynamics within the data. Employing a blend of advanced data processing techniques and interactive visual storytelling, we present a nuanced understanding of the factors that might influence match outcomes. This document serves as a detailed record of our analytical journey, encapsulating the methods, findings, and interpretations derived from our investigation.

1.1 Comparison of penalty conversion rates between men and women

This analysis aims to scrutinize the relationship between gender and penalty conversion rates in football (Figure 1), exploring whether there's a discernible difference in how male and female players perform during these critical moments of the game. Our investigation employs a detailed dataset encompassing penalties from a variety of competitions. We intend to use statistical methods such as the Chi-square test to identify any significant differences in conversion rates between genders, accompanied by effect size calculations like Cramer's V to assess the strength of any observed associations. Additionally, bootstrap resampling techniques will be applied to ensure the robustness of our findings, allowing us to present a comprehensive analysis of penalty conversion rates across genders.

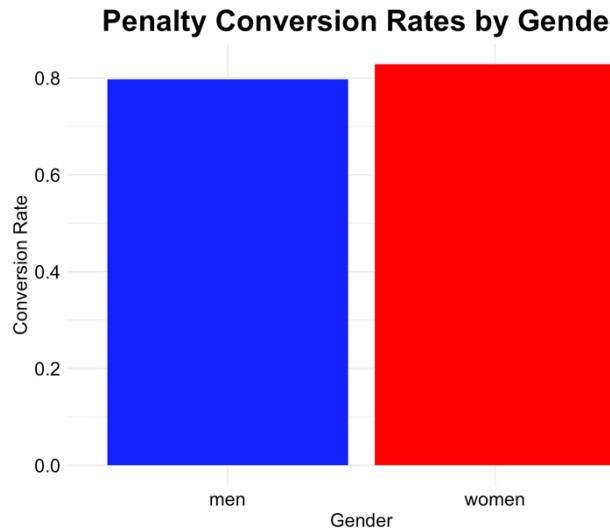


Figure 1 Conversion rates by gender.

1.1.1 Chi-square Method

Theory and Application of Chi-square Test

The Chi-square test, a cornerstone in the realm of statistical analysis, offers a robust method for evaluating the association between categorical variables. Fundamentally, it assesses whether observed frequencies in different categories deviate significantly from what would be expected under the assumption of independence. In the context of our analysis, this method is particularly suited to investigate whether gender (a categorical variable with two categories: men and women) influences the success of penalty kicks in football—a binary outcome (made or missed). By creating a contingency table (Figure 2) that cross-tabulates gender with penalty outcomes, the Chi-square test allows us to quantitatively determine if the observed distribution of successful and unsuccessful penalties aligns with or diverges from what would be expected if gender had no impact on penalty conversion rates.

Penalty conversion rates by gender			
Gender	Made	Missed	
Men	10546	2683	
Women	1544	320	

Figure 2 Count of penalties made and missed by gender.

Results of the Chi-square Test

Our application of the Chi-square test to the penalty kick data revealed a statistically significant difference in conversion rates between male and female players, with a p-value well below the conventional threshold of 0.05 (Figure 3). This result suggests that gender does have an effect on penalty kick outcomes in the dataset analyzed. The contingency table, which provided a breakdown of made and missed penalties by gender, served as the basis for this test, highlighting the utility of the Chi-square method in uncovering patterns within categorical data.

Chi-square Test Summary				
	Statistic	Value	DF	P_Value
X-squared	X-squared	9.745	1	0.0018

Figure 3 Chi square test results table. P-value < 5%.

Limitations of the Chi-square Method

Despite its widespread use and applicability, the Chi-square test is not without its limitations. A primary concern is its sensitivity to sample size. In analyses involving very large datasets, even minuscule differences can appear statistically significant, potentially overstating the practical importance of the findings. Conversely, in studies with smaller samples, the Chi-square test might lack the power to detect genuine associations. Another limitation is its reliance on the assumption that each observation is independent and that all categories have sufficiently large expected frequencies—conditions that are not always met in practice. These constraints necessitate a careful interpretation of Chi-square results, complemented by additional analyses, such as effect size calculations, to gauge the practical significance of observed differences.

This analysis, while revealing significant differences in penalty conversion rates between genders, also underscores the importance of looking beyond mere statistical significance.

1.1.2 Using Cramer's V

Understanding Cramer's V

Cramer's V is a statistical measure that quantifies the strength of association between two categorical variables, offering insight into the relevance of the relationship observed in a Chi-square test. Unlike the Chi-square statistic, which merely indicates the presence of a significant association, Cramer's V assesses the magnitude of this association, providing values that range from 0 (no association) to 1 (perfect association). This measure is particularly useful in our context to understand not just if gender influences penalty conversion rates in football, but also how strong this influence is. In applying Cramer's V to our analysis, we aim to bridge the gap between statistical significance and practical significance, thereby offering a more nuanced understanding of the relationship between gender and penalty success.

Results from Cramer's V Analysis

The application of Cramer's V to our dataset yielded a value that, while indicating a statistically significant association between gender and penalty conversion rates as detected by the Chi-square test, suggests that the strength of this association is quite weak (Figure 4). This finding is critical as it highlights that, despite the statistical significance suggested by the Chi-square test, the practical difference in penalty conversion rates between genders may not be as impactful. Such results prompt

a more tempered interpretation of the initial Chi-square findings, advocating for a consideration of the magnitude of effect alongside its statistical significance.

Chi-square Test and Cramer's V Summary	
Metric	Value
X-squared	9.7450
Degrees of Freedom	1.0000
P-Value	0.0018
Cramer's V	0.0250

Figure 4 Cramer's V summary table. Cramer's V close to 0 indicates not significant difference.

Limitations and Considerations of Cramer's V

While Cramer's V provides valuable insights into the strength of association between categorical variables, it also comes with its own set of limitations and considerations. One key issue is that Cramer's V does not indicate the direction of the association, merely its strength. Therefore, it must be interpreted in conjunction with the contingency table or other descriptive statistics that can provide directional context. Additionally, Cramer's V is a measure of association, not causation; it cannot tell us why or how one variable influences another, only that they are related to some degree. Finally, like many statistical measures, Cramer's V assumes that the data are correctly collected and categorized, and it may not perform well with very small sample sizes or highly imbalanced data distributions.

1.1.3 Bootstrap method

Describing the Bootstrap Method and Its Positives

Bootstrap resampling is a powerful statistical technique that estimates the distribution of a statistic across many simulated samples. This method involves repeatedly drawing samples, with replacement, from a dataset and recalculating the statistic of interest each time. This method does not assume a specific distribution shape for the data (e.g., normal distribution) and can be applied to various types of data and statistical measures. By generating a distribution of the statistic from these resampled datasets, researchers can obtain measures of central tendency (like the mean) and variability (such as standard deviation), as well as construct confidence intervals. The bootstrap method is particularly advantageous in situations where the theoretical distribution of a statistic is complex or unknown, allowing for robust statistical inference based on empirical data distributions.

Results from the Bootstrap Analysis

In our analysis, bootstrap resampling was employed to assess the stability and confidence of the observed difference in penalty conversion rates between genders. By creating thousands of simulated samples and recalculating the difference in conversion rates for each, we were able to construct a distribution of the observed differences. The mean difference obtained from this distribution reinforces the findings from the Chi-square test and Cramer's V, offering a nuanced view of the effect size (Figure 5). Moreover, the confidence intervals generated through this method provide a range within which the true difference in conversion rates is likely to lie, offering valuable insights into the precision of our estimates and the likelihood of observing such a difference purely by chance. Zero value is inside the confidence interval; therefore we can conclude that there is no real difference between women and men penalty conversion.

Bootstrap Analysis Summary	
Metric	Value
Mean Difference	0.00025
Standard Deviation	0.00438
Confidence Interval Lower	-0.00926
Confidence Interval Upper	0.00935

Figure 5 Bootstrap results. Very small mean difference and confidence interval contains also 0.

Limitations and Considerations of Bootstrap Resampling

Despite its flexibility and utility, bootstrap resampling comes with its own set of limitations. One notable concern is its dependency on the original sample being representative of the population. If the initial sample is biased or too small, the bootstrap estimates may also be biased or imprecise. Furthermore, bootstrap methods can be computationally intensive, especially with large datasets and a high number of resampling iterations, potentially limiting their use in resource-constrained environments. Another consideration is the interpretation of bootstrap confidence intervals; different methods for constructing these intervals can yield slightly different results, necessitating careful consideration of the chosen method based on the data's characteristics.

1.1.4 Final thoughts

Our analysis aimed to discern whether there is a statistical difference in penalty conversion rates between men and women. Employing a combination of the Chi-square test, Cramer's V, and bootstrap resampling methods, we found a statistically significant difference between the genders, suggesting that gender does indeed influence penalty conversion outcomes in football. However, the effect size, as measured by Cramer's V, indicates that this difference, while statistically significant according to Chi-square test, is relatively small in practical terms. Bootstrap resampling reinforced the reliability of this finding, further affirming the presence of a difference but also echoing the subtlety of its impact. Therefore, the answer to the main question is although there is a difference in penalty conversion rates between men and women the practical significance of this difference is limited in the broader context of sports performance and decision-making.

When interpreting differences in penalty conversion rates between genders, it's crucial to recognize that such disparities do not inherently imply causality. The observed variations might reflect underlying factors indirectly related to gender rather than the biological or physical capabilities tied to gender itself. Historical disparities in access to training facilities, differences in professional development opportunities, and the level of investment in women's football programs compared to men's are pertinent examples. Environmental and systemic differences could contribute to variations in penalty conversion rates between male and female players.

1.2 Best five penalty takers in the dataset

Evaluating a player's penalty-taking ability presents a nuanced challenge that intersects with the broader complexities of sports analytics. Traditional approaches, such as analysing simple conversion rates, often overlook the depth of context needed to accurately assess performance. Specifically, they may not adequately account for the statistical uncertainty inherent in smaller sample sizes or the number of attempts, potentially leading to misleading comparisons among players. This complexity is heightened when considering players with perfect records from a minimal number of attempts, which might not be truly indicative of superior skill compared to those with slightly lower success rates but over a more substantial number of attempts. Such scenarios underscore the necessity for a more sophisticated evaluation method that can encapsulate a player's proficiency in penalty kicks more comprehensively.

To deal with these challenges, our analysis employs a blend of statistical methodologies, primarily focusing on bootstrap techniques, Bayesian inference, and the construction of a custom metric. The bootstrap method allows us to empirically estimate the variability and confidence intervals of a player's success rate, offering a robust way to account for the uncertainty associated with limited data. Bayesian inference, on the other hand, provides a probabilistic framework that integrates prior knowledge about penalty kick success rates, enabling a dynamic update of our beliefs about a player's ability based on their performance data. Furthermore, we propose the development of a custom metric that combines success rate, attempt volume, and a measure of uncertainty into a single, weighted score. This metric aims to balance the various factors influencing penalty success, using predefined weights to reflect their relative importance systematically.

1.2.1 Bootstrap approach

In assessing a player's penalty-taking prowess, the bootstrap method stands out as a powerful statistical tool that enables the estimation of a player's success rate while accounting for the uncertainty inherent in the available data. This approach is particularly valuable when dealing with varying numbers of attempts across players, as it provides a way to generate an empirical distribution of success rates, which in turn helps calculate confidence intervals for these estimates. The introduction of pseudo-successes and pseudo-failures, derived from the overall success rate, is a strategic method used to stabilize the estimates from players with a limited number of attempts. By doing so, we leverage the collective performance to inform the estimation of individual success rates, imposing a form of regularization that pulls extreme values toward a more general expectation of performance.

Turning to the analysis results and the accompanying visualization, we implemented a filter to exclude players with fewer than four attempts, aligning with the lower 80th percentile of attempt distribution, to ensure a focus on a subset of data robust enough for meaningful analysis. However, the top 10 players highlighted in the plot (Figure 6) predominantly consist of individuals with relatively few attempts, all under 20, indicating that our approach using pseudo-counts might not fully resolve the issues arising from small sample sizes. Additionally, the sizeable confidence intervals observed for these top players suggest a significant degree of uncertainty regarding their ranking order. This uncertainty is visually represented in the plot, where the horizontal error bars indicate the range of the bootstrap confidence intervals for each player's adjusted success rate.

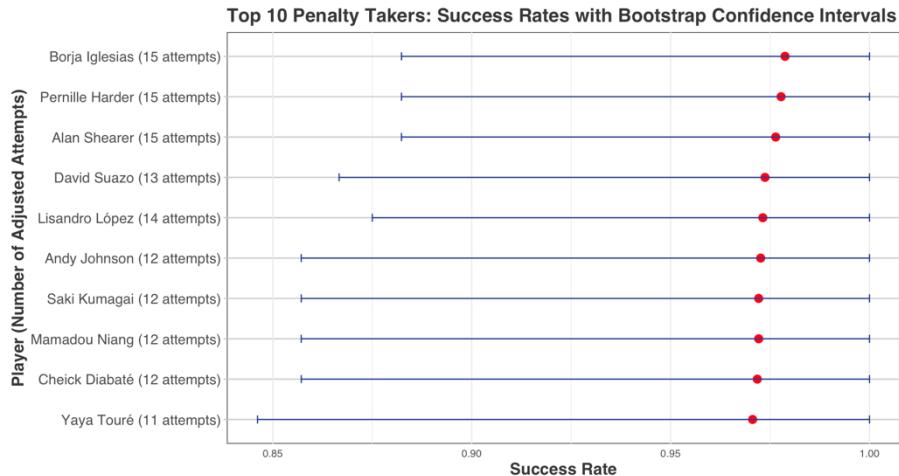


Figure 6 Top 10 penalty takers using Bootstrap method.

Such substantial confidence intervals challenge the reliability of the ranking, emphasizing the necessity of a more refined approach that can provide narrower intervals, particularly for those with fewer attempts, to improve the precision of the estimated success rates. Furthermore bootstrap is very dependable on the random seed which is used for the random bootstrapping. The analysis, therefore, underscores the need for additional methodologies or refinements to address the limitations observed when dealing with players with a sparse number of penalty attempts.

1.2.2 Bayesian approach

The Bayesian approach to analyzing penalty-taking ability provides a probabilistic evaluation that incorporates both the individual performance data and the collective insight drawn from the entire dataset. By adopting a Beta(1,1) prior, we ensure that our analysis starts from a position of minimal assumptions, allowing the actual data to shape the resulting success rate estimates for each player.

Our Bayesian model was executed via the MCMCbinomialbeta function in the MCMCpack package, leveraging the power of MCMC simulations to generate posterior distributions of success rates. These distributions yield not just a point estimate—the posterior mean—but also credible intervals that convey the uncertainty around these estimates.

The plot (Figure 7) illustrates the top performers, indicated by the posterior mean success rates, along with their corresponding credible intervals. While the intervals are informative, they are not necessarily narrower than those from the bootstrap approach. This similarity implies that despite the different methodologies, there is a consistent recognition of uncertainty tied to the number of attempts—particularly evident in players with fewer shots at goal.

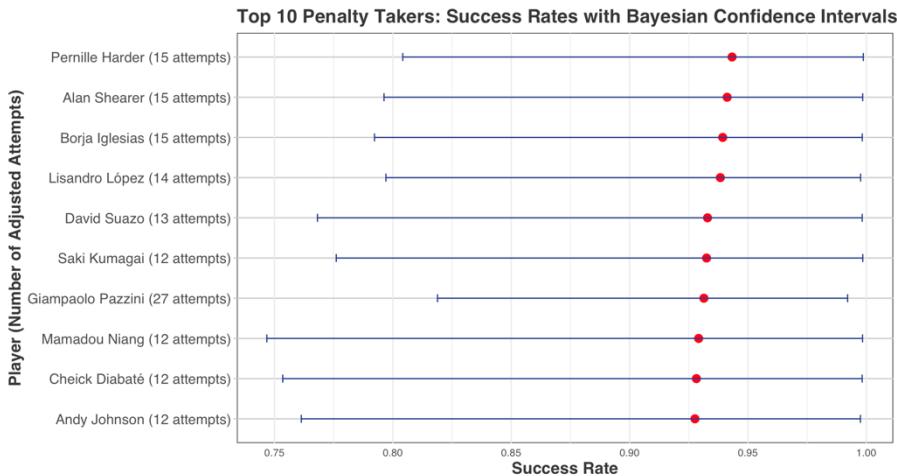


Figure 7 Top 10 penalty takers using Bayesian method.

Nevertheless, the presence of Pazzini, with a relatively higher number of attempts, in the top 10 is indicative of the model's capacity to identify consistent performers across varying amounts of data. It's important to note, though, that while the Bayesian method provides a solid foundation for estimating success rates, the order of top players should still be considered with caution, especially when the sample size is not robust.

In evaluating our model's performance, the Brier score comes into play as a measure of predictive accuracy. The low Brier score achieved here indicates a close alignment between the predicted probabilities and the actual outcomes, underscoring the model's effectiveness in capturing the true probability of a successful penalty kick. This successful prediction performance adds an extra layer of validation to the Bayesian approach, confirming its place as a valuable tool in the sports analytics arsenal.

1.2.3 Custom metric approach

In an endeavor to develop a comprehensive metric for evaluating penalty takers, we integrate several statistical features that reflect different aspects of a player's performance. The goal is to amalgamate these features into a single, cohesive metric that captures a player's penalty-taking ability more holistically than simple conversion rates. The features we consider are:

Conversion Rate (CR): This is the most straightforward measure, representing the proportion of penalties that a player successfully scores out of their total attempts. It is calculated by dividing the number of successful penalties by the total number of penalties taken.

Uncertainty Adjustment (UA): Recognizing that a conversion rate derived from a small number of attempts may not be as reliable as one derived from many attempts, we introduce an uncertainty adjustment. This adjustment is calculated using the standard error of the conversion rate, which considers both the number of attempts and the conversion rate itself. The Uncertainty Adjustment (UA) is then determined by taking the inverse of one plus the standard error. The closer the UA is to 1, the more reliable the conversion rate is deemed to be.

Normalized Penalties Scored (PS_norm): To put individual success in the context of the collective performance, the number of penalties a player scores is normalized by the highest number of penalties scored by any player in the dataset. This normalization ensures that the metric accounts for the range of performance while scaling individual successes against a benchmark of the best performance observed.

The composite score for each player is then calculated by taking the average of these three components, with equal weight given to each.

In analyzing the results (Figure 8), a striking distinction from earlier Bayesian and bootstrap approaches is the inclusion of players with a substantial number of attempts, such as Cristiano Ronaldo and Lionel Messi. This difference highlights a potential advantage of the custom metric: its ability to reflect a player's extended history and consistent performance over a larger dataset.

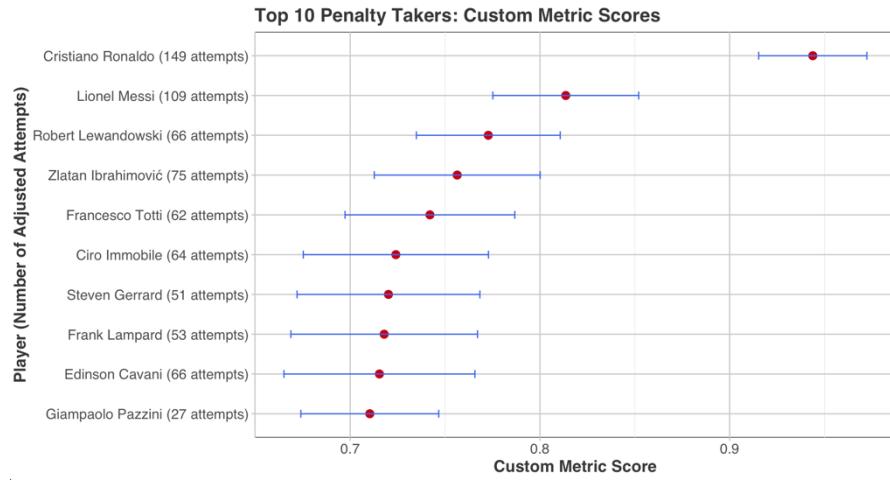


Figure 8 Top 10 penalty takers using custom metric for identifying penalty taking skill.

However, this approach is not without its challenges and shortcomings. One primary issue is the predefined, equal weighting of components. This may not accurately reflect the true influence each component has on a player's likelihood to successfully score a penalty. Ideally, weights should be derived from a modeling process, where numerous features are evaluated for their relationship with the target variable (success or miss). By identifying important features and estimating their coefficients through a regression or machine learning model, we could allocate weights that mirror the predictive importance of each feature.

Another concern with the current method is that it doesn't account for the contextual factors that could affect penalty success, such as pressure situations, goalkeeper skill, or match conditions. Furthermore, the equal weighting scheme may not be the most effective way to synthesize the components into a single score. The fact that the results show significant differences compared to the earlier approaches underscores the necessity of a more dynamic weighting system that can adapt to the unique aspects of each player's data. Finally this metric favors players with high number of attempts and penalty takes.

1.2.4 Penalty Taking Skill, combining Bayesian approach with Custom Metric

In developing a nuanced metric for assessing penalty-taking skill, we've synthesized two distinct approaches into a single rating system, normalized on a 0-100 scale. This scale is particularly user-friendly and familiar to football enthusiasts, as it resonates with the rating systems often seen in popular football video games.

The new metric, named "Penalty Taking Skill," incorporates elements from both the Bayesian approach and the custom metric. The Bayesian component tends to favor players with high conversion rates, even if the number of attempts is relatively low. This aspect of the metric rewards accuracy and success rate above all. Conversely, the custom metric component recognizes and rewards players who have consistently taken and scored penalties over a higher number of attempts, valuing experience and a proven track record.

By averaging these two metrics, the Penalty Taking Skill rating aims to balance the precision of the Bayesian method with the broader perspective of the custom metric.

The first plot (Figure 9) shows the top 5 penalty takers from the full dataset. Notably, Cristiano Ronaldo tops this chart, demonstrating his exceptional penalty-taking prowess across a significant number of

attempts (149). This result is consistent with a general football consensus that Ronaldo is an elite penalty taker. The plot illustrates a mix of players, some with numerous attempts like Ronaldo and others with fewer attempts but still showing a high skill rating, reflecting the combined metric's dual nature.

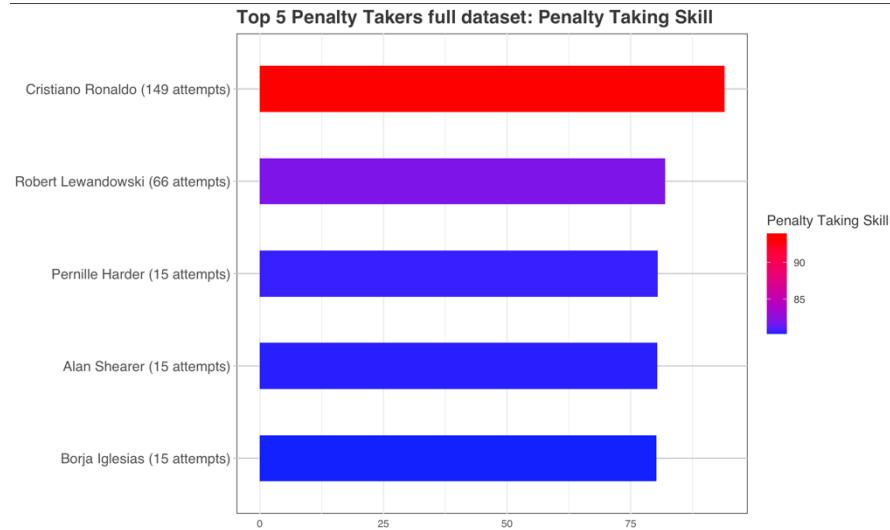


Figure 9 Penalty Taking Skill. A combination of Bayesian approach and Custom Metric. Top 5 players, full dataset.

The second plot (Figure 10) filters for players in the top 99th percentile by attempts, focusing on those with the most experience taking penalties. Once again, Cristiano Ronaldo features prominently, alongside other prolific players like Lionel Messi and Robert Lewandowski. This plot emphasizes the skill of players who not only score penalties consistently but also step up to take them in numerous high-stakes situations.

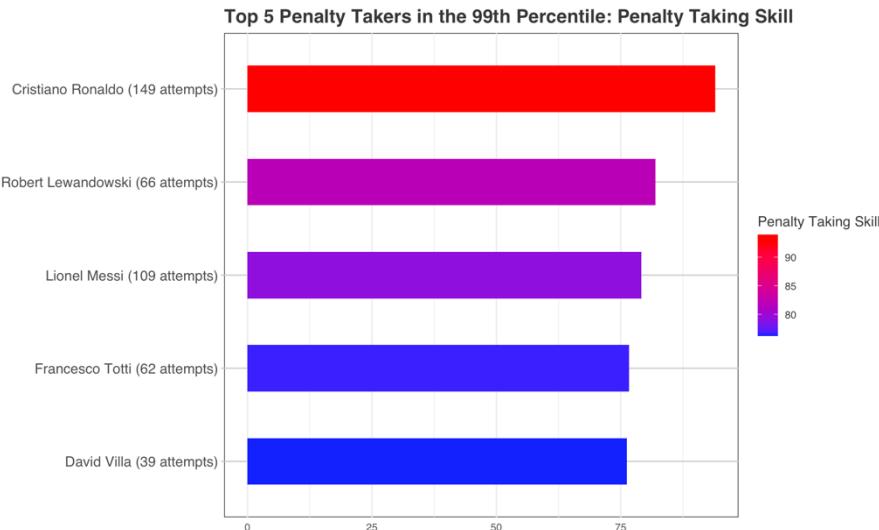


Figure 10 Penalty Taking Skill. A combination of Bayesian approach and Custom Metric. Top 5 players, 99th percentile in terms of total attempts.

However, the metric has its limitations. It incorporates the Bayesian approach, which can disproportionately favor players with perfect or near-perfect success rates from a smaller number of attempts. This could lead to an inflated skill rating for players who have not been tested as extensively as those with many attempts. Additionally, the custom metric component, while rewarding the quantity of penalties taken, may also privilege players who have had more opportunities, regardless of their success rate variability over time.

Moreover, the predetermined equal weighting of components in the custom metric does not account

for the possibility that some aspects of performance might be more predictive of future success than others. A more sophisticated model would weigh features based on their predictive power, which could be determined through statistical analysis or machine learning techniques.

Another drawback is that the conversion rate is a factor in both the Bayesian and custom metric calculations, which could result in this aspect of performance being doubly influential in the final rating.

1.2.5 Conclusion

The quest to identify the top five penalty takers through various statistical techniques—bootstrap, Bayesian, custom metrics, and a combined approach—has provided a comprehensive picture of penalty-taking prowess. Each method offers unique insights, with the bootstrap highlighting the uncertainty in penalty success rates, the Bayesian emphasizing conversion rates, and the custom metric valuing the sheer number of penalties taken. By integrating these methods, the combined metric attempts to balance these factors, offering a more rounded evaluation of a player's penalty-taking ability.

Despite the robustness of the combined approach, it is crucial to note that no single metric perfectly captures the multifaceted skill of penalty taking. The dual influence of conversion rates, the potential bias towards players with fewer attempts, and the arbitrary nature of the weights assigned in the custom metric all pose challenges that must be acknowledged. Yet, these analyses move us closer to understanding the qualities that distinguish the best penalty takers.

1.3 Strategic Selection: Ordering the First Five Penalty Shootout Takers

Addressing the challenge of determining the optimal sequence of penalty takers for a team preparing for shootouts is a task that straddles both theoretical and practical realms. Theoretically, the approach to solving this problem involves statistical analysis, historical data examination, and psychological factors consideration. Practically, it requires application of these findings to real-world scenarios, considering individual player conditions and team dynamics that may not be fully captured in the data.

Penalty shootouts stand as the ultimate test of nerve and skill in football, a moment where games, and sometimes the fate of entire tournaments, hinge on the narrowest margins. For teams and coaches, the arrangement of penalty takers can be as strategic as it is psychological, with the order potentially influencing the outcome of the shootout. The process of determining this order is multifaceted, blending quantitative analysis with a nuanced understanding of player psychology and team chemistry.

1.3.1 Literature review

The research paper "Penalty shootouts are tough, but the alternating order is fair" by Silvan Vollmer, David Schoch, and Ulrik Brandes presents an empirical analysis of penalty shootouts in association football, focusing on whether the conversion rates during regulation or extra time differ from those during shootouts, and whether the alternating shooting order offers an advantage to the team kicking first. The study analyzes approximately 50,000 penalties from European men's football competitions, about one-third of which are from more than 1,500 penalty shootouts.

Key findings include:

- Shootout conversion rates are lower than during regular game play, attributed more to the worsened performance of shooters rather than improved performance of goalkeepers.
- Statistically, there is no advantage for either team in the alternating shooting order, suggesting fairness in the current system.
- Detailed analyses complement these findings, exploring factors like player selection, fatigue, and first-mover advantage.

Other studies that explore the football penalties are the following:

- Almeida, Volosovitch, and Duarte (2016) investigated penalty kick outcomes in UEFA club competitions from 2010-2015, identifying situational, individual, and performance factors that affect penalty success.
- Apesteguia and Palacios-Huerta (2010) focused on psychological pressure in competitive environments, using a randomized natural experiment to demonstrate how pressure affects performance, particularly in penalty kick scenarios.
- Arrondel, Duhautois, and Laslier (2019) explored the decision-making process under psychological pressure, specifically examining the shooter's anxiety during penalty kicks.
- Bar-Eli and Azar (2009) conducted an empirical analysis of shooting strategies and goalkeepers' preferences during penalty kicks in soccer, offering insights into the tactical aspects of executing and defending penalty kicks.
- Brams and Ismail (2018) discussed the fairness of sports rules, with a focus on making penalty shootout rules fairer, suggesting modifications to current practices to enhance equity.
- Brinkschulte, Furley, and Memmert (2020) addressed the common assumption that English football players perform poorly in penalty kick situations, using statistical analysis to challenge this narrative.

By examining these studies, we can gather that penalty kick outcomes are influenced by a complex interplay of psychological, tactical, and situational factors. The insights from this literature review underline the importance of understanding these dynamics to improve penalty kick strategies and performance in competitive football matches.

1.3.2 Comprehensive Factors Influencing Penalty Taker Selection

In structuring the approach to select the first five penalty takers for a shootout, we delve into the multidimensional aspects that can influence penalty kick outcomes. These considerations integrate empirical findings with common football wisdom and strategic insights:

- **Psychological Factors:** Anxiety and pressure can significantly impact a player's performance. Studies have shown that players may experience different levels of pressure based on their position in the lineup. The initial penalty takers may feel the burden of setting a strong precedent, while subsequent takers may face the escalating tension of sustaining the lead or catching up.
- **Player's Historical Performance:** Past performance data is a crucial indicator of a player's potential success. Players with a high conversion rate during regular and high-pressure game situations are likely to replicate this success during shootouts.
- **Fatigue and Physical Condition:** The physical state of players after a full match can affect their performance. Fatigued players may not perform at their peak level, hence their order in the shootout should consider their endurance levels and any signs of injury.
- **Goalkeeper's Tendencies and Weaknesses:** The analysis of opposing goalkeepers can influence the selection and order of penalty takers. Understanding a goalkeeper's diving patterns, reaction times, and weaknesses can guide the strategy to exploit these factors.
- **Tactical Considerations:** The order of penalty takers may be arranged to apply psychological pressure on the opposition. A strong initial taker can set a confident tone, while reserving other high-performing takers for critical moments can sustain pressure on the opposing team.
- **Player Confidence and Composure:** The mental state of players just before the shootout is a significant consideration. Those who exhibit confidence and composure under pressure may be better suited to take on the early or decisive penalties in the shootout sequence.

By integrating these factors, coaches can formulate a strategic order for penalty takers that maximizes the team's chances of success. While empirical studies provide a statistical foundation for these decisions, the real-time assessment of players' conditions and the match context are equally imperative to finalize the lineup.

1.3.3 Prioritizing Penalty Taker Selection: Strategic Factors

When deciding on the optimal sequence of players for penalty shootouts, a hierarchy of considerations must be meticulously structured. These considerations draw on empirical research, tactical acumen, and ingrained football intuition. Here is the order of factors deemed pivotal for the strategic selection of penalty takers:

1. **Optimal Player Ordering:** The sequence of penalty takers is paramount, with the goal to progress from the most optimal to less optimal choices, acknowledging that many shootouts may not extend to the fifth kicker. The initial taker plays a vital role in setting a positive psychological state and boosting team morale.
2. **Closing Strategy:** The fifth kicker, who may potentially face the most consequential moment, should be a blend of skill, experience, and psychological solidity, ready to shoulder the climax of the shootout pressure.
3. **Quantitative Assessment:** A composite metric, akin to the one developed previously (1.2.4), which amalgamates conversion rates with the volume of experience (Penalty Taking Skill), serves as a quantitative backbone for selection.
4. **Psychological Resilience:** Gauging a player's confidence and composure throughout the match provides insights into their current mental fortitude. Players who stay calm and perform well under pressure should be chosen first.
5. **Shootout Experience:** Prior exposure to shootout conditions is invaluable. Those with a history of composed performance in similar high-stress situations warrant a higher place in

- the order.
6. **Experienced Players:** Players with more years in the game usually handle stress better. Their know-how can help calm the intense pressure of a penalty shootout.
 7. **Physical Readiness:** Despite a penalty requiring a singular effort, awareness of any player's physical limitations due to fatigue or injury is essential, as it may subtly influence performance.
 8. **Positional Proficiency:** Traditionally, forwards and midfielders, adept in scoring scenarios, are preferred over defenders and goalkeepers, though exceptions based on individual prowess should be acknowledged.
 9. **Other Tactical Nuances:** Additional nuances, such as recent penalty track records, goalkeeper insights, and left-footed versus right-footed dynamics, might also inform the selection process.

This hierarchy is not prescriptive but rather suggestive, offering a framework that balances statistical evidence, player psychology, and situational tact. It is an amalgam of data-driven rationale and the more qualitative elements that define the sport's unpredictable nature.

1.3.4 Strategic Selection of Shootout Penalty Takers: Analyzing Manchester United's 2020/2021 Season Data

For our practical analysis, we chose Manchester United for the 2020/2021 season as a case study due to a substantial number of penalties taken: 28 across the season by 12 distinct players. This large sample of penalties and players provides a robust foundation for combining theoretical insights with practical observations.

The methodology for selecting our penalty takers is grounded in a combination of empirical data and well-established football insights. I generated additional data points for each player to evaluate their performance under pressure—a crucial factor in a penalty shootout scenario. I defined "under pressure" situations as those during penalty shootouts or moments in the game where the player's team was not leading (i.e., drawing or losing).

The extra features that I created are the following:

- **Shootout Attempt:** The number of penalties each player took during shootouts.
- **Shootout Success:** The number of successful penalties during shootouts.
- **Shootout Success Percentage:** The player's success rate during shootouts.
- **Under Pressure Attempt:** The total number of penalties taken when the player's team was either in a shootout or not leading in the game.
- **Under Pressure Success:** The number of successful penalties under the defined pressure conditions.
- **Under Pressure Success Ratio:** The success rate of penalties under pressure.

Please note that the additional metrics included in our analysis were derived by considering the complete career statistics of each player. I merged these new columns with the pre-existing dataset, which already included each player's overall penalty attempts, successes, failures, and a calculated metric of penalty taking skill (1.2.4). The combined dataset (Figure 11) forms the basis for our strategic selection and ordering of Manchester United's top penalty takers.

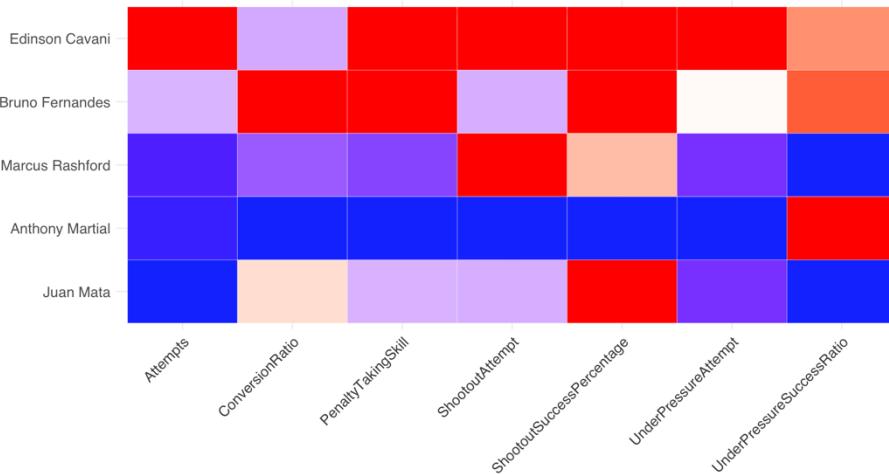
Comprehensive Penalty Performance Analysis of Manchester United Players (2020/2021 Season)														
player_id	player	team_id	team	season_name	Attempts	Successes	Failures	PenaltyTakingSkill	ShootoutAttempt	ShootoutSuccess	ShootoutSuccessPercentage	UnderPressureAttempt	UnderPressureSuccess	UnderPressureSuccessRatio
4350	Edinson Cavani	123	United	2020/2021	66	52	14	70.811230	3	3	1.000000	39	32	0.8205128
805	Bruno Fernandes	123	United	2020/2021	30	26	4	70.846925	1	1	1.000000	23	19	0.8260870
4224	Marcus Rashford	123	United	2020/2021	13	10	3	55.787328	3	2	0.6666667	9	7	0.7777778
1863	Anthony Martial	123	United	2020/2021	12	9	3	53.767222	0	0	0.000000	6	5	0.8333333
1971	Juan Mata	123	Manchester	2020/2021	11	9	2	59.593778	1	1	1.000000	9	7	0.7777778
4009	Alex Telles	123	Manchester	2020/2021	3	2	1	39.634031	1	1	1.000000	2	1	0.500000
4306	Fred	123	Manchester	2020/2021	2	1	1	26.481606	2	1	0.500000	2	1	0.500000
1543	Axel Tuanzebe	123	Manchester	2020/2021	1	1	0	59.183366	1	1	1.000000	1	1	1.000000
1987	Victor Lindelöf	123	Manchester	2020/2021	1	1	0	58.914959	1	1	1.000000	1	1	1.000000
1295	Luke Shaw	123	Manchester	2020/2021	1	1	0	58.911913	1	1	1.000000	1	1	1.000000
1536	Daniel James	123	Manchester	2020/2021	1	1	0	58.301692	1	1	1.000000	1	1	1.000000
1651	David de Gea	123	Manchester	2020/2021	1	0	1	9.669463	1	0	0.000000	1	0	0.000000

Figure 11 Manchester United season 2020/2021, candidate penalty takers for shootouts.

In forming our penalty shootout strategy, we meticulously selected the top five Manchester United players by scrutinizing their overall penalty performance, experience in taking penalties and experience under pressure, with a particular focus on attackers and attacking midfielders due to their higher propensity for taking penalties (Figure 12).

Manchester United Penalty Shootout Strategy: Top 5 Takers Ranked

Color Scale: Blue (Low) — White (Medium) — Red (High)



Note: The color scale represents normalized performance metrics.

Figure 12 Manchester United Season 2020/2021 Penalty Shootout Strategy: Analytical Ranking of Top 5 Takers Based on Comprehensive Performance Metrics.

- Edinson Cavani and Bruno Fernandes were chosen as the first two takers, reflecting our strategy to start strong. Both players exhibit exceptional Penalty Taking Skills, a trait represented by the red intensity in this metric on our heatmap. This initial duo is critical;
- Marcus Rashford occupies the third slot, strategically placed at the midpoint of the sequence to sustain the initial momentum. His heatmap indicators present a blend of low to high performance, suggesting a reliable penalty taker capable of handling the mounting pressure as the shootout progresses.
- We placed Anthony Martial fourth. While he has shown a varied performance, his placement is designed to leverage his strong points just before the final kick.
- For the decisive fifth penalty, we've reserved Juan Mata, notwithstanding that his performance in several metrics surpasses Martial's. His red zones in Shootout Success Percentage reveal his competency in high-stakes situations. The final penalty often bears the brunt of the shootout's pressure, and Mata's experience and composure make him uniquely suitable for this crucial moment.

This deliberate ordering is not only rooted in their skills as reflected in the heatmap but also their

positions. Attackers and attacking midfielders typically have more experience and success with penalties, underlining the rationale for our selection. Their larger volume of penalties taken, as seen in the Attempts metric, reinforces their expertise and justifies their inclusion over teammates with fewer penalties to their name.

In conclusion, our choice of takers is a blend of empirical data analysis and an appreciation for the mental and situational demands of a penalty shootout. The heatmap serves as a guide, but the final order is also heavily influenced by understanding the nuanced dynamics of pressure and player specialties.

1.3.5 Summary and limitations

Critical Factors for Shootouts: In summary, literature and practical observations concur that several critical factors influence the outcome of penalty shootouts. Key among them is the psychological resilience of the players, their historical performance under pressure, and their experience in similar high-stake situations. The physical state of the players, their position roles (attackers and midfielders often fare better), and the tactical order of the lineup, also play significant roles in determining the success rate of penalties during a shootout.

Approach to Player Selection: Our approach to selecting Manchester United's top 5 penalty takers for the 2020/2021 season involved an analytical review of historical penalty data, emphasizing attempts, successes, and a Penalty Taking Skill metric reflective of their overall prowess. We also factored in the players' performance during high-pressure situations, both in shootouts and in-game scenarios where the team was not leading. This analysis led to a strategic ranking that aligns with both empirical data and widely accepted football insights, aiming to maximize the team's potential for success.

Limitations of the Analysis: While our method offers a robust framework for choosing penalty takers, it's important to acknowledge its limitations. We do not have in-game physiological data, such as heart rate measurements, which could offer deeper insights into a player's pressure handling capabilities. Furthermore, the absence of comprehensive data on goalkeepers' performance and their historical interaction with our selected penalty takers leaves a gap in our analysis. Our evaluation of penalty taking skills and the final selection for Manchester United does not account for such data, which could influence penalty outcomes. We also lack individual player performance data from training sessions, which could impact the selection. Despite these limitations, our analysis presents a systematic approach to forming a penalty shootout strategy, which should be combined with real-time observations and coaching staff insights for the most effective outcome.

2 References

- Almeida, Carlos H., Anna Volossovitch, and Ricardo Duarte. "Penalty kick outcomes in UEFA club competitions (2010-2015): the roles of situational, individual and performance factors." *International Journal of Performance Analysis in Sport* 16.2 (2016): 508-522.
- Apesteguia, Jose, and Ignacio Palacios-Huerta. "Psychological pressure in competitive environments: Evidence from a randomized natural experiment." *American Economic Review* 100.5 (2010): 2548-2564.
- Arrondel, Luc, Richard Duhautois, and Jean-François Laslier. "Decision under psychological pressure: The shooter's anxiety at the penalty kick." *Journal of Economic Psychology* 70 (2019): 22-35.
- Bar-Eli, Michael, and Ofer H. Azar. "Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences." *Soccer & Society* 10.2 (2009): 183-191.
- Brams, Steven J., and Mehmet S. Ismail. "Making the rules of sports fairer." *Siam Review* 60.1 (2018): 181-202.
- Brinkschulte, Michel, Philip Furley, and Daniel Memmert. "English football players are not as bad at kicking penalties as commonly assumed." *Scientific Reports* 10.1 (2020): 7027.
- Vollmer, Silvan, David Schoch, and Ulrik Brandes. "Penalty shootouts are tough, but the alternating order is fair." *arXiv preprint arXiv:2310.04797* (2023).