Game Changers: Examining Performance Metrics that Lead to FIFA World Cup Upsets

Simon Mathew Vincent

Abstract

Football, commonly known as soccer in some parts of the world, is the most popular sport globally, showing off an estimated 4 billion fanbase and a staggering economic impact exceeding $200 billion annually. The FIFA World Cup, organized every four years in a different host city, stands as a cultural event, generating over $7.5 billion in revenue in 2022 while uniting nations through the spirit of competition and national pride. The tournament's capacity for upsets captures audiences, allowing underdog teams to defy expectations and carve their names into football history. This study delves into match data from previous tournaments to uncover the tactical and statistical elements that contribute to these unexpected victories. By analyzing key factors such as efficiency in counterattacks, ball possession, passing accuracy, and opponent errors, we reveal distinct patterns that differentiate upsets apart from anticipated outcomes. Our findings illuminate the strategies that enable underdog teams to overturn the odds, highlighting the dynamic nature of these remarkable victories. This research offers valuable insights for analysts, coaches, and fans who aspire to deepen their understanding of the dynamics behind game-changing performances in international football.

Introduction

In the world of sports, an "upset" is defined as an unexpected outcome, typically when a less favoured team or player defeats a stronger opponent. A situation in which someone beats the player or the team that was expected to win (Cambridge Dictionary). Upset prediction is critical as it is a billion-dollar game. This concept is particularly common in international soccer, where the FIFA World Cup showcases teams from around the globe competing for the cup. Upsets fascinate fans and analysts, as they challenge assumptions of skill and strategy, often leading to memorable moments that define the tournament's history. From stunning last-minute goals to underdog performances, these events can completely shift the story of a competition.

Throughout FIFA's history, there have been numerous instances where lower-ranked teams have risen to the occasion, shocking the soccer community and altering the expected course of the tournament. Outstanding upsets, such as the United States defeating England in the 1950 World Cup or Senegal's triumph over France in 2002, serve as powerful reminders that in soccer, anything is possible (Thapliyal). These unexpected victories not only raise the status of the winning teams but also inspire aspiring players and create a sense of hope among fans that the game can bring surprise and excitement at any moment.

The purpose of this research is to better understand and study the underlying causes of upsets, from a football perspective such as the factors or metrics that affect an upset. I hypothesize that some important factors such as shots on goals, tackles, interceptions, and long balls increase the likelihood of upsets the most. This premise will be put to the test through a quantitative analysis of upsets in the 2022 Fifa World Cup. The insights gained from this study may be beneficial for football analytics by coaches and fans.

Literature Review

This literature review examines the significant variables influencing defensive performance, goal-scoring effectiveness, and passing organization in football.

Defensive performance

Defensive performance is when the team not in possession of the ball prevents the team in possession of the ball from scoring a goal. Four key defensive factors —defensive errors, tackle success rate, dual success rate, and clearances—were identified as critical to a team's defensive performance. Defensive errors emerged as the most influential factor, correlating with a 10% reduction in the probability of winning. Research indicates that these errors can lead to an imbalanced defence, particularly worsening toward the end of matches when players experience fatigue, consequently increasing the likelihood of a goal (Lepschy et al.).

Goal Scoring Effectiveness

The success of a football match is fundamentally determined by a team's ability to score goals, making it essential to consider various key performance indicators associated with scoring and defensive actions. Among these, shots on target serve as a crucial causal condition, as evidence shows that winning teams typically record a higher number of attempts on target. [citation] While some studies emphasize the importance of total shots about match outcomes [citations], the debate remains regarding the reason behind shooting, with some players taking shots simply to maintain rhythm rather than sole scoring motivation. In addition to shots on target, possession control remains a highly debated variable (Bloomfield et al.). This debate is about whether ball control is a strong or a weak factor linked to success. Influenced heavily by Barcelona's playstyle, possession is thought to correlate with success, yet other research suggests it is a weak predictor of match results (Yan et al.).

Passing Organisation

Furthermore, possession is often associated with increased passing frequency or crosses. Crosses also help directly in scoring goals. (Pulling et al.). Crosses also play a pivotal role in creating goal-scoring opportunities, as they can directly lead to goals. Previous studies indicate a strong link between the effectiveness of crosses and the success of winning teams, justifying that they are an important indicator in performance analysis (Pulling et al.). Moreover, the defensive line breaks a newer FIFA indicator which is critical when considering how teams pressure their opponents. A defensive line break is when the attacking team plays the ball beyond the farthest player in that line. These breaks are especially significant when they occur near the attacking third, where they can lead to scoring opportunities. (Yan et al.).

Methods

Data Collection

For my research, I utilized FBref, a comprehensive platform to track statistics for football teams and players globally (FBref). My primary focus was on the match statistics from the 2022 FIFA World Cup. I collected data for 32 teams from the 2022 FIFA World Cup. From this platform, I collected data on upsets matches and compiled them. It is important to acknowledge the limitations present in this methodology. This data collection primarily relied on statistics from the 2022 World Cup and there were 64 matches played between 32 teams out of which 30 matches were upsets. Hence, it may not provide a comprehensive and clear view of historical trends or patterns in football upsets. Additionally, data availability for earlier tournaments and matches was not utilised, which could affect the usefulness of the findings.

Data Analysis

To determine whether a match is an upset, I analysed the Expected Goals (xG's) metrics, specifically, if the team with the lower xG's emerged victorious, it is classified as an upset. Following this process, I gathered a range of data to identify common factors among underdog teams that contributed to their success in these unexpected victories (Figure 1). This can create a correlation between these metrics and success rates.

Match	Team	Goals	xG	Possession	Passing Accuracy	Shots on Target	Saves	Fouls	Corners
Senegal vs Netherlands	Senegal	0	0.9	46	78	27	1	13	6
Senegal vs Netherlands	Netherlands	2	0.7	54	78	30	4	13	7
Argentina vs Saudi Arabia	Aregntina	1	2.2	69	83	36	0	7	9
Argentina vs Saudi Arabia	Saudi Arabia	2	0.1	31	66	67	5	21	2
Germany vs Japan	Germany	1	3.1	74	87	32	2	6	6
Germany vs Japan	Japan	2	1.5	26	68	33	8	14	6
Belgium vs Canada	Belgium	1	0.8	53	83	33	2	11	4
Belgium vs Canada	Canada	0	2.4	47	81	10	2	14	4
Qatar vs Senegal	Qatar	1	1	45	79	30	2	7	6
Qatar vs Senegal	Senegal	3	0.8	55	84	38	2	12	6
Tunisia vs Australia	Tunisia	0	0.9	59	74	29	1	16	5
Tunisia vs Australia	Australia	1	0.6	41	66	22	4	15	2
Poland vs Saudi Arabia	Poland	2	1.6	37	74	33	4	18	4
Poland vs Saudi Arabia	Saudi Arabia	0	1.7	63	83	27	1	15	5
Japan vs Costa Rica	Japan	0	0.9	57	85	23	0	22	5
Japan vs Costa Rica	Costa Rica	1	0.1	43	78	25	3	9	0
Korea Republic vs Ghana	Korea Republic	2	1.8	63	82	27	0	13	12
Korea Republic vs Ghana	Ghana	3	1.6	37	77	43	4	9	5
Tunisia vs France	Tunisia	1	0.5	35	74	60	3	14	7

Figure 1. Data Sample of Upset Matches

In my research, I'm utilizing Karl Pearson's correlation coefficient, which is a statistical measure that quantifies the strength and direction of the linear relationship between two continuous variables. It is denoted by the variable 'r' and the value of "r" ranges from -1 to 1. A strong positive correlation (values close to 1) indicates that as one variable increases, the other variable also tends to increase consistently. For example, a strong positive correlation might be observed between shots on target and goals scored. On the other hand, a weak or negative correlation (values close to -1) suggests that as one variable increases, the other tends to

decrease, or that there is little to no linear relationship. For instance, a weak negative correlation could be seen between possession and goal conceded. By employing Pearson's correlation in my research, I aim to identify and analyse the patterns and relationships present in my data effectively.

Results

Average Normalised Difference in Metrics

The results show which factors are crucial for underdog teams to win a match. The key metrics with a positive average normalised difference were shots on target, fouls, saves, tackles, interceptions, clearances, and goal kicks. The first table compares different soccer stats between winning and losing teams, showing which factors contribute most to an upset (Figure 2). Negative bars below zero are stats where the underdog team had less control than opposition.

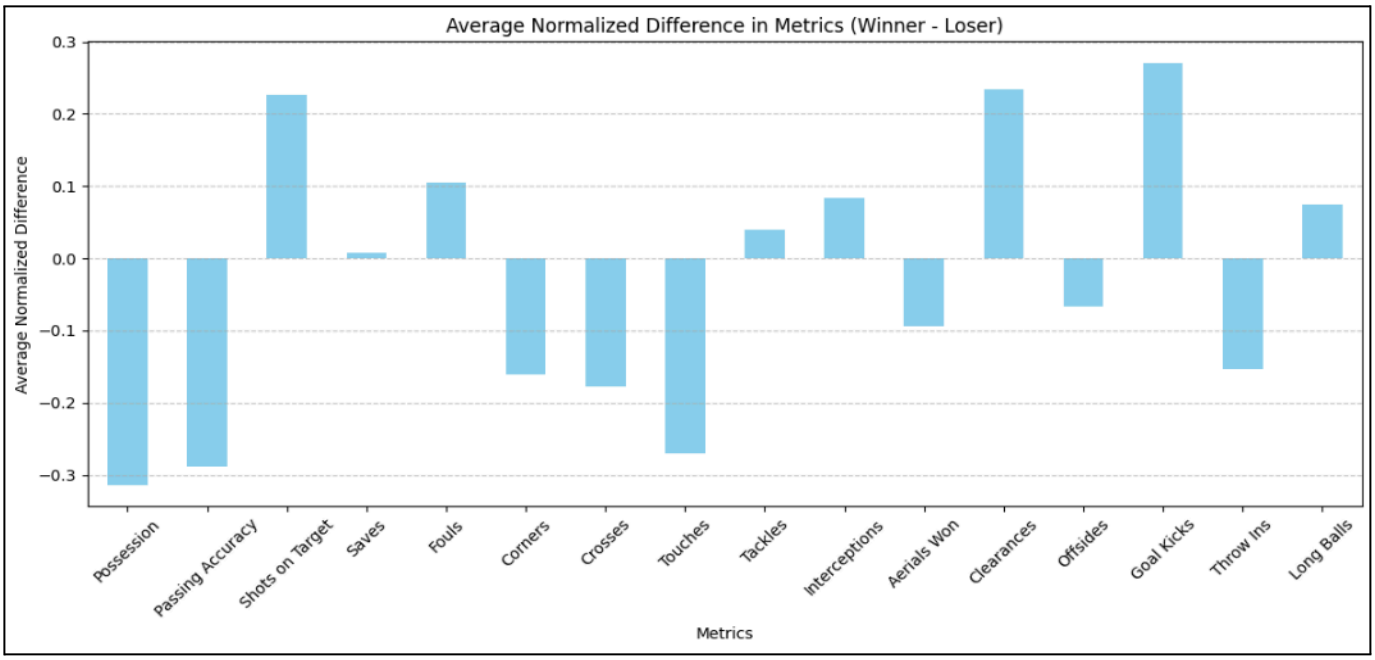

Figure 2. Average Normalized Difference in Metrics

As shown in Figure 2, possession, passing accuracy, and touches are all lower for winners, meaning upsets often happen when the underdog lets the other team control the ball but still finds a way to win. Positive bars above zero are stats where the underdog team had more control. Notably, shots on target, clearances, and goal kicks are higher, showing that underdogs tend to defend a lot, take their chances efficiently, and clear the ball more often. It shows that underdogs often win despite having less possession, lower passing accuracy, and fewer touches, meaning they let the stronger team control the game but remain defensively solid. Instead of dominating play, they focus on efficiency, having more shots on target, clearances, and goal kicks, indicating a counter attacking approach. More long balls and aerial duels won suggest a direct style rather than complex passing. Essentially, upsets happen when

the weaker team defends well, stays disciplined, and takes their limited chances more effectively.

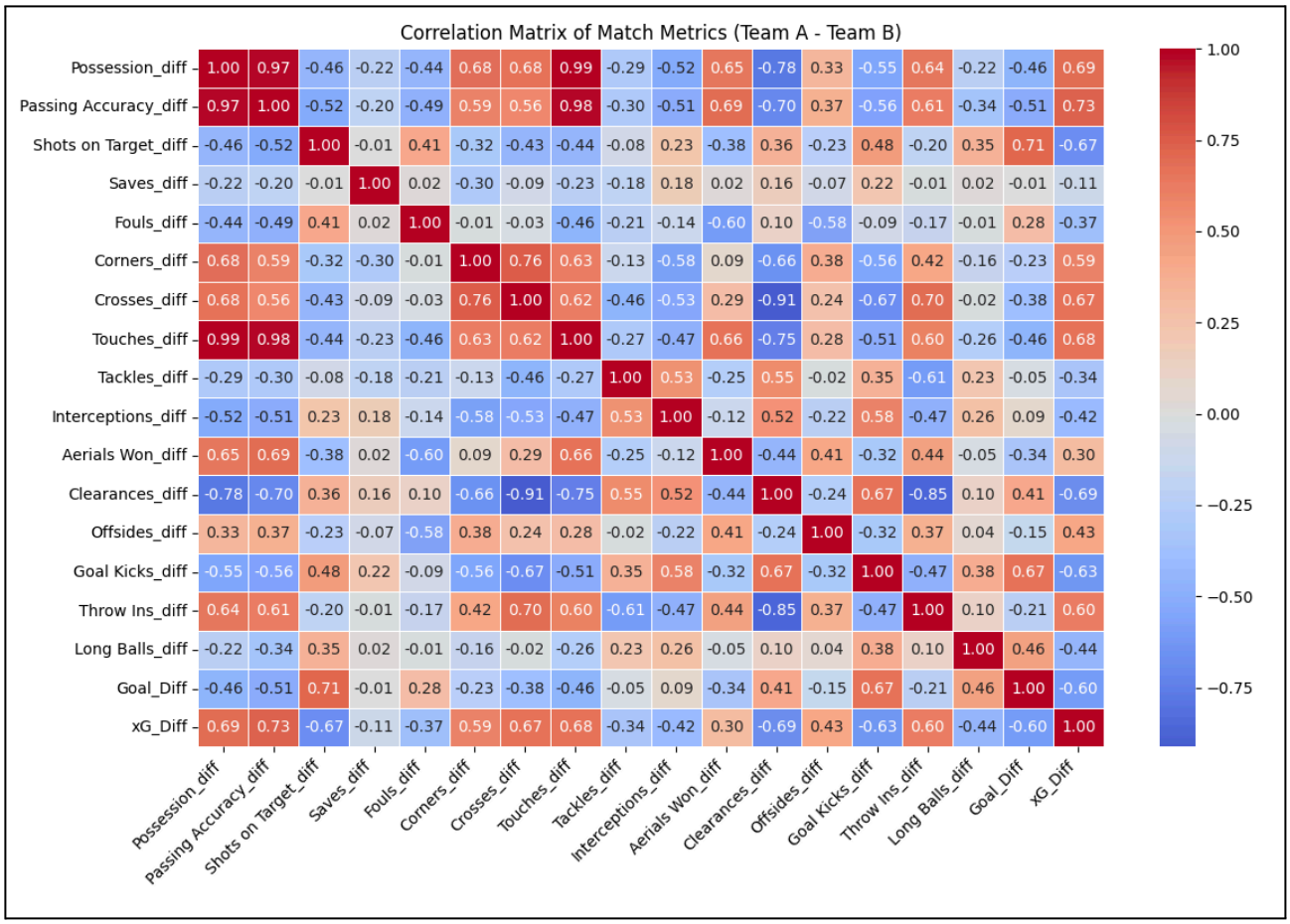

Figure 3. Correlation Matrix of Match Metrics

The third figure is a correlation matrix that shows how different soccer match statistics relate to each other (Figure 3). The numbers and colours indicate whether two stats move together (positive correlation, shown in red) or move in opposite directions (negative correlation, shown in blue). For example, possession and passing accuracy have a strong positive correlation, meaning that teams with more possession usually pass more accurately. Similarly, goal difference (Goal_Diff) and shots on target are positively related, suggesting that teams that create more accurate shots tend to score more and win games. On the other hand, clearances and possession have a strong negative correlation, meaning that teams that clear the ball a lot usually don't have much possession, which is common for underdog teams. The matrix also shows that expected goals (xG_Diff) are strongly linked to shots on target and goal difference, confirming that good goal-scoring chances often lead to winning.The line of 1.00s going from the top left to the bottom right is the main diagonal, which shows that each variable always perfectly correlates with itself. Overall, this figure helps identify which stats matter most in a game and how different styles of play are connected. Teams that dominate possession tend to rely on passing, while defensive teams focus on clearances, long balls, and aerial duels.

Discussion

The study has several limitations that should be addressed in future research. It did not analyse historical data from previous years and used only thirty upsets from the FIFA World Cup 2022, which could enhance the understanding of trends over time. Additionally, psychological factors such as player motivation and mental health, as well as team cohesion, were not measured, potentially overlooking critical influences on performance. Demographic variables like player age and team size were also omitted, limiting the depth of the analysis. As a recommendation for future investigations, researchers should incorporate longitudinal approaches, include assessments of psychological and interpersonal factors, and examine key demographic variables to develop a more comprehensive understanding of the dynamics at play.

Future research should delve deeper into the unpredictable nature of human psychology to enhance our understanding of how and why upsets occur across different contexts, including personal, social, and professional spheres. Understanding the underlying psychological mechanisms can provide valuable insight into the factors that lead to unexpected outcomes. Additionally, the use of emerging tools like artificial intelligence presents a promising approach for improving prediction accuracy. By maximizing AI's ability to analyse large datasets and identify complex patterns in the sport of soccer, researchers can develop more complex and predictive models. These models could include variables such as emotional state, social influences, and cognitive biases, which play a crucial role in decision-making processes. Ultimately, this broad approach could lead to more effective strategies for anticipating and managing upsets, making valuable contributions to fields such as psychology, behavioral economics, and risk management.

Conclusion

In conclusion, the results of this study provide clear evidence that validates the original hypothesis that factors like shots on goal, tackles, interceptions and long balls increase the likelihood of upsets. The key takeaways highlight the significant impact of shots on goal, tackles, interceptions and long balls, supporting the initial predictions made of the research. However, it is important to acknowledge the limitations of the study, including only a small dataset analysed, psychological factors and mental health. These limitations suggest that while the findings may be helpful, further research is necessary to fully understand the implications. Future work should focus on a vast dataset, including different factors such as mental health and using AI which would help to deepen our understanding of this topic and address the questions that remain.

References

Bloomfield, J., et al. “Effects of score-line on team strategies in FA premier league soccer.” 16 September 2020, https://eprints.qut.edu.au/204511/. Accessed 17 February 2025.

Cambridge Dictionary. *Upset*. Definition. Learner's Dictionary ed., Cambridge, 2025, dictionary.cambridge.org. Accessed 11 February 2025. Online.

FBref. *2022 World Cup Scores & Fixtures*. 2022. *FBref*, https://fbref.com/en/comps/1/schedule/World-Cup-Scores-and-Fixtures. Accessed 24 February 2025.

Lepschy, Hannes, et al. “Success Factors in the FIFA 2018 World Cup in Russia and FIFA 2014 World Cup in Brazil.” *Frontiers in Psychology*, vol. 12, 2021, www.frontiersin.org. Accessed 11 February 2025.

Pulling, C., et al. “Analysis of crossing at the 2014 FIFA World Cup.” vol. 18, no. 4, 2017. *Analysis of crossing at the 2014 FIFA World Cup*, https://www.tandfonline.com/doi/abs/10.1080/24748668.2018.1509255. Accessed 18 February 2025.

Thapliyal, Mukesh. “FIFA World Cup: 10 of the Biggest Upsets of All time.” *Times of India*, 24 November 2022, timesofindia.indiatimes.com. Accessed 11 February 2025.

Yan, W., et al. “How to win in the FIFA World Cup Qatar 2022? A study of technical and tactical indicators based on fuzzy-set qualitative comparative analysis.” *Frontiers in Psychology*, 2023, pmc.ncbi.nlm.nih.gov. Accessed 11 February 2025.