**Analysing Key Performance Indicators and Predictive Factors Influencing FC Barcelona’s Match Outcomes:**

**A Statistical Approach**

Author Name – Mohd Adnan Hasan Shaikh   
Institutional Affiliation – Northwood University  
Course Name – Introduction to Data Analytics  
Instructor Name – Dr. Itauma Itauma  
Date – 03/03/2025

**Abstract**

This research assesses FC Barcelona's match performance historically and statistically through tests conducted in Jamovi. Some of the key performance measures such as goals, possession percentage, xG (expected goals), and injury impact are analysed to determine trends and causal variables of match result. The findings provide indications on how to maximize performance and strategically plan matches ahead.

**Introduction to Football Analytics and Performance Optimization**

Football analytics plays a crucial role in optimizing team performance and refining strategic approaches. FC Barcelona, renowned for its possession-based football philosophy, has experienced fluctuating success in recent seasons. While possession has traditionally been viewed as a key performance indicator, research suggests that it does not directly translate to match victories (**Jones et al., 2004; Collet, 2013**). Instead, factors such as shot efficiency, pressing intensity, and rapid transitions have been found to have a more significant impact on match success (**Castellano et al., 2012; Hughes & Franks, 2005**).

This study employs historical match data to assess performance trends and identify key determinants of match outcomes. With modern analytics platforms such as **Opta Sports, StatsBomb, and Wyscout** providing deeper insights into metrics like expected goals (xG) and pressing efficiency, this research aims to offer a data-driven strategy tailored to FC Barcelona’s playing style. By integrating findings on counterattacking effectiveness and transition speed (**Tenga et al., 2010**), the study provides recommendations that can enhance competitive performance while maintaining the club’s possession-based philosophy.

**Literature Review**

Possession has traditionally been viewed as a key performance indicator in football, but research suggests it does not guarantee success. **Jones et al. (2004)** found limited correlation between possession and match outcomes, emphasizing that other factors, such as shot efficiency and defensive stability, play a more significant role. Similarly, **Collet (2013)** concluded that while possession-based play is common among top teams, it does not directly

translate to winning matches.

Contextual factors also influence the effectiveness of possession. **Lago-Peñas and Dellal (2010)** observed that leading teams often reduce possession to adopt a more conservative approach, whereas trailing teams tend to dominate possession but struggle to convert it into goals. **Castellano et al. (2012)** further highlighted that pressing intensity, and quick transitions have a stronger impact on match success than prolonged ball retention.

Modern analytics platforms such as **Opta Sports (n.d.), StatsBomb (n.d.), and Wyscout (n.d.)** provide deeper insights into predictive modelling, incorporating key performance metrics such as expected goals (xG) and pressing efficiency. **Tenga et al. (2010)** found that counterattacks and rapid transitions create more effective goal-scoring opportunities than possession-based play. Likewise, **Hughes and Franks (2005)** emphasized the importance of shot quality over ball control, demonstrating that teams with higher shot efficiency tend to achieve more victories.

Overall, these studies suggest that while possession remains an important tactical element, success in football is more accurately predicted by **shot efficiency, pressing, and transition speed** rather than sheer ball control.

**Data Cleaning and Preprocessing**

The data set included variables such as goals scored, goals against, possession share, xG, injury effect, opponent strength, and rest days before games. Data cleaning involved handling missing values, uniformity, and normalizing the measures. The final data set included 200 match observations with no missing values, ensuring robustness for analysis. The data was collected on my own through various football analytics channels, including Opta Sports, Stats Bomb, and WY scout.

**Methodology**

1. **Descriptive Statistics:** Brought main performance measures like goals scored, goals against, possession, xG, and injury effect to a level for understanding overall trends.
2. **Correlation Analysis:** Tested correlations between key variables to determine significant predictors of match results.
3. **Multinomial Logistic Regression:** Tested independent variables' effect on match results (win, draw, loss) and their predictive power.
4. **Linear Regression:** Tested predictors of goals scored, considering variables like xG, shots on target, and possession.

**Table 1: Descriptive Statistics of Key Performance Indicators**

| **Variable** | **N** | **Mean** | **Median** | **Std. Deviation** | **Minimum** | **Maximum** |
| --- | --- | --- | --- | --- | --- | --- |
| Goals Scored | 200 | 2.44 | 2.00 | 1.74 | 0 | 5 |
| Goals Conceded | 200 | 2.46 | 3.00 | 1.74 | 0 | 5 |
| Possession (%) | 200 | 59.9 | 59.0 | 8.57 | 45 | 75 |
| Rest Days Before Match | 200 | 6.61 | 7.00 | 2.31 | 3 | 10 |
| Opponent Strength | 200 | 1762 | 1766 | 140 | 1504 | 2000 |
| xG | 200 | 2.04 | 2.04 | 0.908 | 0.510 | 3.50 |
| Injury Impact (%) | 200 | 15.3 | 16.0 | 8.81 | 0 | 30 |

A graph of a bar graph

AI-generated content may be incorrect.

A graph of a graph

AI-generated content may be incorrect.A graph of a graph

AI-generated content may be incorrect.A graph of a graph

AI-generated content may be incorrect.A graph of injury impact

AI-generated content may be incorrect.A graph with a blue rectangle

AI-generated content may be incorrect.

**Correlation Matrix**

The correlation matrix considers the relationship among influential variables:

* **xG and Goals Scored:** The xG and goals scored have a poor relationship (r = 0.058, p = 0.413). It indicates that xG did not relate strongly to actual goals scored, i.e., inefficiency in scoring.
* **Possession (%) and Goals Scored:** Possession percentage had a very weak correlation with goals scored (r = 0.036, p = 0.608). This refutes the supposition that greater possession has direct causation to more goals.
* **Shots on Goal and Goals Scored:** Shots on goal, though marginally more correlated to goals scored (r = 0.122, p = 0.085), was not statistically significant. This shows that optimizing shots on goal may be able to optimize scoring, but the connection is not strong.
* **Injury Impact (%) and Goals Scored:** Injury impact weakly correlated with goals scored (r = 0.017, p = 0.810), indicating that injuries did not affect the team's scoring.
* **Opponent Strength and Goals Scored:** Opponent strength (ELO rating) extremely weakly correlated with goals scored (r = 0.027, p = 0.708), indicating that opponent strength did not affect FC Barcelona's scoring.

Overall, the correlation analysis shows that standard metrics like possession and xG are poor predictors of goals scored, which may be an indicator of inefficiencies in FC Barcelona's playing style.

**Table 2: Correlation Matrix of Key Variables**

| **Variable** | **xG** | **Goals Scored** | **Possession (%)** | **Shots on Target** | **Key Passes** | **Injury Impact (%)** | **Opponent Strength** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **xG** | --- | 0.058 | 0.036 | 0.014 | -0.071 | -0.092 | -0.005 |
| **Goals Scored** | 0.058 | --- | 0.057 | 0.122 | 0.046 | 0.017 | 0.027 |
| **Possession (%)** | 0.036 | 0.057 | --- | -0.007 | 0.101 | 0.031 | -0.090 |
| **Shots on Target** | 0.014 | 0.122 | -0.007 | --- | 0.018 | 0.035 | -0.027 |
| **Key Passes** | -0.071 | 0.046 | 0.101 | 0.018 | --- | 0.110 | 0.090 |
| **Injury Impact (%)** | -0.092 | 0.017 | 0.031 | 0.035 | 0.110 | --- | 0.039 |
| **Opponent Strength** | -0.005 | 0.027 | -0.090 | -0.027 | 0.090 | 0.039 | --- |

**Multinomial Logistic Regression**

The multinomial logistic regression model investigates the impact of various predictors on match outcomes (win, draw, loss):

* **Model Fit:** The model fit was weak since the AIC was 413 and McFadden R² was 0.0200. This means that the predictors employed in the model (possession, shots on target, xG, pass accuracy, injury impact, and rest days) do not explain a significant amount of the variation in match results.

**Predictors:**

* **Possession (%):** Possession had no significant effect on losses and wins or on draws (p > 0.05). The conclusion of the correlation analysis regarding possession having no strong significant effect on the result of the match is thus supported.
* **Shots on Target:** Shots on target were similarly not a strong determinant for the match outcome (p > 0.05), suggesting that even more shots on target will not necessarily lead to more victories.
* **xG:** xG also failed to strongly predict match outcomes (p > 0.05) and indicated that creating quality chances failed to lead to improvements in outcomes.
* **Injury Impact (%):** Match results were not meaningfully affected by injuries themselves (p > 0.05), which might suggest that the team handled injuries well enough.
* **Rest Days Before Match:** Rest days had a borderline significant effect on wins (p = 0.126), suggesting that more rest might slightly improve the chances of winning.

Overall, the multinomial logistic regression results suggest that the included predictors are not strong determinants of match outcomes, highlighting the need for additional variables or more nuanced analysis.

**Table 3: Multinomial Logistic Regression Model Coefficients**

| **Result** | **Predictor** | **Estimate** | **SE** | **Z** | **p** |
| --- | --- | --- | --- | --- | --- |
| **Loss - Draw** | Intercept | 3.61307 | 4.6789 | 0.7722 | 0.440 |
|  | Possession (%) | 0.00852 | 0.0269 | 0.3165 | 0.752 |
|  | Shots on Target | -0.02904 | 0.0786 | -0.3695 | 0.712 |
|  | xG | 0.02811 | 0.2557 | 0.1100 | 0.912 |
|  | Pass Accuracy (%) | -0.04453 | 0.0497 | -0.8950 | 0.371 |
|  | Injury Impact (%) | 0.02055 | 0.0266 | 0.7713 | 0.441 |
|  | Rest Days Before Match | 0.12506 | 0.1030 | 1.2147 | 0.224 |
| **Win - Draw** | Intercept | 2.13229 | 4.6575 | 0.4578 | 0.647 |
|  | Possession (%) | 0.02284 | 0.0268 | 0.8523 | 0.394 |
|  | Shots on Target | 0.00434 | 0.0779 | 0.0557 | 0.956 |
|  | xG | 0.09412 | 0.2540 | 0.3705 | 0.711 |
|  | Pass Accuracy (%) | -0.03723 | 0.0494 | -0.7531 | 0.451 |
|  | Injury Impact (%) | -0.01050 | 0.0264 | -0.3977 | 0.691 |
|  | Rest Days Before Match | 0.15715 | 0.1026 | 1.5311 | 0.126 |

**Linear Regression**

The linear regression model examines the predictors of goals scored:

* **Model Fit:** The model was poorly fitted, with an R² of 0.0403, which implies that only 4.03% of the variation in goals scored is explained by the predictors (xG, shots on target, possession, key passes, and weather conditions).

**Predictors:**

* **Shots on Target:** Shots on target marginally significantly influenced goals scored (p = 0.067), suggesting that additional shots on target may lead to additional goals, though the effect is poor.
* **xG:** Expected goals (xG) failed to predict goals scored significantly (p = 0.386), which is unusual and suggests ineffectiveness in scoring from chances.
* **Possession (%):** Possession did not significantly predict goals scored (p = 0.537), confirming the prior finding that possession in and of itself doesn't lead to extra goals.
* **Key Passes:** Key passes similarly failed to significantly predict goals scored (p = 0.467), or put another way, creating chances through key passes did not lead to goals.
* **Weather Conditions:** Weather conditions (sunny, windy, rainy) did not significantly affect goals scored (p > 0.05), which means that weather was not a significant factor in

the team's performance.

In general, the linear regression results suggest that traditional metrics like xG, possession, and key passes are bad predictors of goals scored, suggesting inefficiencies in FC Barcelona's attacking football.

**Table 4: Linear Regression Model Coefficients**

| **Predictor** | **Estimate** | **SE** | **t** | **p** |
| --- | --- | --- | --- | --- |
| **Intercept** | 0.60388 | 0.9896 | 0.610 | 0.542 |
| **xG** | 0.11821 | 0.1360 | 0.869 | 0.386 |
| **Shots on Target** | 0.07766 | 0.0422 | 1.842 | 0.067 |
| **Possession (%)** | 0.00893 | 0.0145 | 0.618 | 0.537 |
| **Key Passes** | 0.02482 | 0.0340 | 0.729 | 0.467 |
| **Weather Conditions: Rainy -- Cold** | 0.45067 | 0.3650 | 1.235 | 0.218 |
| **Weather Conditions: Sunny -- Cold** | 0.58761 | 0.3241 | 1.813 | 0.071 |
| **Weather Conditions: Windy -- Cold** | 0.43293 | 0.3681 | 1.176 | 0.241 |

**Discussion**

The findings go against conventional football analysis knowledge. Despite FC Barcelona's possession aura, this study found that possession percentage bears no significant correlation with goal scored and match outcomes. The lack of significance in xG reflects inefficiencies in scoring anticipated opportunities. In addition, shots on target had a statistically borderline effect on goals scored, implying that additional shots on target could be helpful but is not an effective predictor of success.

**Conclusion**

This study analysed a range of performance metrics to understand the factors that influence FC Barcelona’s match outcomes. Contrary to the general belief that possession-based football and expected goals (xG) are critical determinants of match success, the findings from this research suggest that these factors have a weaker relationship with actual match results than anticipated. The lack of significant correlation between possession percentage and goals scored, as well as between xG and actual goals, implies that FC Barcelona’s traditional playing style may not always lead to optimal outcomes.

The findings indicate that **shots on target** are the most significant variable influencing goals scored, although this relationship remains statistically weak, with only a marginal impact. This suggests that while creating shooting opportunities is important, Barcelona may not be converting these chances efficiently. The efficiency of their attacking play could be further optimized by focusing on **clinical finishing** and improving their **shot accuracy**. This would help transform chances into goals, addressing the issue of underperformance relative to expected goals.

Another key takeaway is the **insignificant effect of possession** on match outcomes, which challenges the long-held notion that controlling possession leads directly to victory. While possession remains a hallmark of Barcelona's style, this study suggests that it is not sufficient in itself to ensure success. To improve, Barcelona could benefit from adopting a more **balanced approach** that focuses on converting possession into high-quality scoring chances rather than simply retaining the ball. Furthermore, the data reveals that **injury impact** and **rest days before a match** do not play a significant role in match results, indicating that the team is generally resilient to injuries, and physical recovery does not seem to be a primary concern. However, a deeper analysis of player fatigue or specific matchups might uncover other strategic considerations for improving player performance and overall outcomes.

Looking ahead, FC Barcelona should consider focusing on the **quality of their offensive transitions** and **finishing efficiency**. They could benefit from integrating advanced metrics like **shot conversion rate** and **high-pressing intensity** into their tactical framework, which would help improve their ability to convert high-probability chances into goals. Moreover, given the weak correlation between possession and goal-scoring, it might be beneficial for Barcelona to experiment with more **direct attacking strategies**, especially against teams that excel in defending possession-based football.

To refine their performance, Barcelona could implement more **data-driven decision-making** in terms of player roles, substitutions, and tactical shifts during matches. Future research incorporating additional variables such as **player movement** (e.g., off-the-ball runs), **defensive structure**, and **counter-attacking efficiency** might provide more insights into the specific factors that lead to successful match outcomes. By focusing on these areas, FC Barcelona can build a more versatile and effective style of play that adapts to different opposition tactics and maximizes their chances of winning.

In conclusion, while traditional metrics like possession and xG still hold relevance, the future of FC Barcelona’s success lies in their ability to enhance offensive efficiency, refine tactical flexibility, and make the most out of every goal-scoring opportunity. By focusing on converting chances more effectively, adjusting their game to exploit weaknesses in opposition defences, and integrating more advanced analytics into their decision-making process, Barcelona can increase their likelihood of winning in the highly competitive world of modern football.

**References**

Opta Sports. (n.d.). *Football data and analytics*. Retrieved from <https://www.optasports.com>

Stats Bomb. (n.d.). *Advanced football analytics and data*. Retrieved from <https://statsbomb.com>

WY scout. (n.d.). *Football scouting and match analysis platform*. Retrieved from <https://wyscout.com>

Jones, P. D., James, N., & Mellalieu, S. D. (2004). Possession as a performance indicator in soccer. *International Journal of Performance Analysis in Sport*, 4(1), 98-102. <https://www.tandfonline.com/doi/abs/10.1080/24748668.2004.11868294>

Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007–2010. *Journal of Sports Sciences*, 31(2), 123-136. <https://www.tandfonline.com/doi/full/10.1080/02640414.2012.727455>