Report: Key Findings and Business Recommendations from Premier League Dataset for FanDuel



By Harkanwar Chadha and Armand Debray

Our Main Key Findings from the Kaggle Premier League Dataset:

- Goal Scoring Patterns: The distribution of goals scored shows that most teams score between 1 to 2 goals per game, with few extreme outliers. Teams often cluster around a "typical" goal-scoring range, indicating that goal predictions could be made based on this pattern.
- **Possession Trends:** The possession box plot reveals that teams typically maintain around 52% possession, with few teams significantly dominating or being dominated. The data suggests that possession-heavy teams aren't always significantly outperforming in terms of goals scored, but possession is still an important indicator of match control.
- Relationship Between Shots and Goals: There is a clear positive correlation between shots on target and goals scored, meaning that teams with more shots on target are likely to score more goals. However, total shots and goals scored show a weaker correlation, indicating that efficiency (shots on target) matters more than the sheer number of shots taken.
- Correlation Analysis: The heatmap shows that shots on target have a strong positive
 correlation with goals scored, while possession and total shots show weaker but still positive
 correlations. Notably, goals against (goals conceded) do not correlate positively with shots or
 shots on target, as expected.
- Matchweek Trends: Specific matchweeks showed a higher tendency for teams to either score or concede goals, suggesting that team performance fluctuates over time, which can be factored into performance forecasting and betting strategies.

Business Recommendations:

- 1. Leverage Goal Scoring Trends for Over/Under Bets: The distribution of goals scored can be used to set accurate over/under lines. For example, betting markets could be adjusted to target games with high likelihoods of scoring 1-2 goals. Highlight games where there is a higher or lower goal-scoring tendency based on historical patterns.
- 2. **Prop Bets on Possession and Match Control:** Since possession is a useful, though not definitive, indicator of match control, you can offer prop bets on possession rates, such as

- betting on teams to exceed a certain possession threshold (e.g., over 55% possession) without directly tying it to match outcomes.
- 3. Incorporate Shot Efficiency Metrics: Use the strong correlation between shots on target and goals to create betting markets for shot efficiency (e.g., a team's shots on target to goals ratio). Users could place bets on which team will have more shots on target or have a higher conversion rate.
- **4. Provide Betting Insights Based on Team Efficiency:** Highlight teams that consistently convert shots into goals and offer users insights about these teams for more informed betting on first goal scorer or correct score bets. Custom in-app alerts can be generated when these teams play.
- **5.** Target Vulnerable Matchweeks for Upset Bets: Use the identified trends in matchweeks where teams underperform (e.g., conceding more goals) to create markets for potential upsets. Offering special bets on underdog teams during these matchweeks can drive user engagement, particularly in live betting.
- **6. Betting Markets Based on Correlations:** The positive correlation between shots on target and goals provides a predictive metric for total goals betting. Recommend bets on total goals in games where a high number of shots on target is expected. Also, use the weaker correlations from possession and total shots to guide markets related to match dynamics, such as team domination or set-piece opportunities.

Conclusion:

By leveraging these insights, FanDuel can improve its offering by providing smarter, data-driven betting options, and tailoring specific markets around key performance indicators like goal-scoring trends, shots on target, and possession. This approach can enhance user engagement, increase the volume of bets, and ensure that FanDuel remains a leader in the sports betting industry with cutting-edge predictive analytics.