**Technical Report: Premier League Club Metrics and Efficiency Analysis**

**1. Outline**

* Introduction
* Story of Data
* Data Splitting and Preprocessing
* Pre-Analysis
* In-Analysis
* Post-Analysis and Insights
* Data Visualizations & Charts
* Recommendations and Observations
* Conclusion
* References & Appendices

**2. Introduction**

**Objective of the Project:**

To evaluate the performance efficiency of Premier League football clubs across multiple metrics, including goals, match outcomes, attendance, and top scorers.

**Problem Being Addressed:**

With vast match data and team performance statistics available, clubs need effective performance analytics to assess competitiveness, optimize strategy, and enhance fan engagement. This analysis aims to identify patterns, top performers, and efficiency gaps among clubs.

**Key Datasets and Methodologies:**

Data was analysed using Power BI, drawing on a structured dataset of Premier League team metrics (e.g., matches played, goals scored/conceded, attendance, top scorers). Analytical tools include dynamic dashboards, DAX calculations, and interactive charts.

**3. Story of Data**

**Data Source:**  
Compiled from Premier League match statistics (official fixtures, club reports, and publicly available sports databases).

**Data Collection Process:**  
Team-level data was collected for 20 clubs, encompassing 38 matches per team for a full season.

**Data Structure:**

* Rows represent club-season entries.
* Columns include squad name, goals, wins, losses, draws, points, attendance, and scorers.

**Important Features and Their Significance:**

* **Goals For / Against**: Primary metric for offensive and defensive performance.
* **Goal Difference**: Indicates team balance.
* **Attendance**: Reflects fan engagement.
* **Top Scorers**: Showcases individual contribution.
* **Rank / Points**: Determines team success.

**Data Limitations or Biases:**  
The dataset does not account for injuries, fixture difficulty, or mid-season transfers which could influence performance context.

**4. Data Splitting and Preprocessing**

**Data Cleaning:**  
The dataset was well-organized within Power BI and required minimal cleaning. Uniform naming conventions and missing value checks were conducted.

**Handling Missing Values:**  
All records were complete. In case of any null entries, Power BI default logic was applied to exclude or calculate averages.

**Data Transformations:**  
Calculated columns and measures were created using DAX to evaluate average points per match, goal differences, attendance averages, and player scoring distribution.

**Data Splitting:**  
Dependent variables: Team rank, total points.  
Independent variables: Goals for/against, attendance, match outcomes.

**Industry Context:**  
This project is grounded in sports analytics, focusing on football (soccer) performance assessment.

**Stakeholders:**  
Football club managers, analysts, fan engagement teams, broadcasters, and sports journalists.

**Value to the Industry:**  
Offers performance benchmarking and competitive insights to optimize coaching, marketing, and ticketing strategies.

**5. Pre-Analysis**

**Identify Key Trends:**

* Liverpool identified as the best-performing team.
* Average goals scored across clubs: 55.8
* Manchester United had the highest average attendance.

**Potential Correlations:**

* Higher attendance correlates with better club ranking (e.g., Manchester United, Arsenal).
* Teams with consistent goal differences trend higher in the ranking.

**Initial Insights:**  
Offensive strength and home attendance may contribute positively to total points.

**6. In-Analysis**

**Unconfirmed Insights:**

* Attendance might influence morale and match outcomes, but this needs qualitative backing.
* Mid-ranked teams (e.g., Brentford, Brighton) show efficient goal use despite lower budgets.

**Recommendations:**

* Improve squad depth for teams with poor goal difference.
* Clubs with high attendance but low rank (e.g., West Ham) should review tactical execution.

**Analysis Techniques Used in Power BI:**

* DAX for average and total metrics
* Donut chart for top scorer distribution
* Line charts for goal difference vs. rank
* Bar charts for attendance comparison

**7. Post-Analysis and Insights**

**Key Findings:**

* Mohamed Salah led scoring with 29 goals.
* Average matches played per team: 38
* Arsenal had one of the best point totals (74) with a strong win record.
* Goal difference sharply declines past rank 10, indicating performance gaps.

**Comparison with Initial Findings:**  
Insights confirmed expectations: high-performing clubs maintain both strong scoring records and consistent attendance.

**8. Data Visualizations & Charts**

* **Line Graph:** Goal difference by team rank shows performance stratification.
* **Donut Chart:** Goal contribution of top 5 scorers across the league.
* **Bar Charts:** Attendance figures and average goals per squad.
* **Dashboard:** Combines key KPIs (e.g., avg. points, matches played, best team) for executive overview.

Each visualization supports a specific aspect of team efficiency, aiding in quick comparative analysis.

A screenshot of a football game

AI-generated content may be incorrect.

**9. Recommendations and Observations**

**Actionable Insights:**

* Focus recruitment on defensive gaps for bottom 10 teams.
* Reinforce attacking support for top scorers to maintain momentum.
* Use attendance insights to enhance matchday fan experience and club revenue.

**Optimizations or Business Decisions:**

* Marketing teams should leverage top scorers and high-attendance clubs for branding.
* Coaching staff can target clubs with unstable goal differences for strategic advantage.

**Unexpected Outcomes:**

* Some mid-table teams outperformed higher-ranked squads in attendance, suggesting strong fan loyalty despite performance.

**10. Conclusion**

**Key Learnings:**  
Power BI can reveal key patterns in sports data highlighting player impact, club consistency, and fan engagement drivers.

**Limitations:**  
The report doesn't account for off-field variables like injuries, weather, or internal management decisions.

**Future Research:**

* Include player injury and transfer data.
* Explore multi-season performance for predictive modeling.
* Integrate financial metrics (e.g., club spend, revenue).

**11. References & Appendices**

**References:**

* Premier League official statistics
* Internal Power BI data model and measures

**Appendices:**

* Full Power BI .pbix file
* Dashboard screenshot (included)
* KPI matrix and scoring table exports