

ENGLISH PREMIER LEAGUE ANALYSIS



Premier League

PREMIER LEAGUE

LEAGUE TABLE

POS	NAME	PL	W	D	L	GD	PTS
1	Youssef Rabee	20	15	4	1	30	49
2	Mohamed Ashraf	20	14	3	3	20	45
3	Mariam Mohamed	20	12	5	3	18	41
4	Rana Tarek	20	10	6	4	5	36



Premier League

Agenda

1. PROBLEM STATEMENT
2. PRESENTED BY AND TASKS ASSIGNED TO EACH ONE
3. KEY MILESTONES IN DETAILS
4. -ASKING QUESTIONS
5. -EDA AND COLLECTING DATA
6. -CLEANING AND PREPARING
7. -ANSWERING QUESTIONS AND EDA
8. CONCLUSION
9. RECOMMENDATIONS AND POINTS OF IMPROVEMENT



Premier League

PROBLEM STATEMENT

The English Premier League is one of the most competitive football leagues globally. With vast amounts of historical data available, this project aims to uncover meaningful patterns, team strategies, and factors influencing match outcomes. Our goal is to leverage data analysis to provide insights that can guide teams, coaches, and sports analysts.



Premier League

PRESENTED BY AND TASKS ASSIGNED:

-Youssef Rabee: Data Collection, Statistical Modeling

-Mohamed Ashraf:: Data Cleaning and Feature Engineering

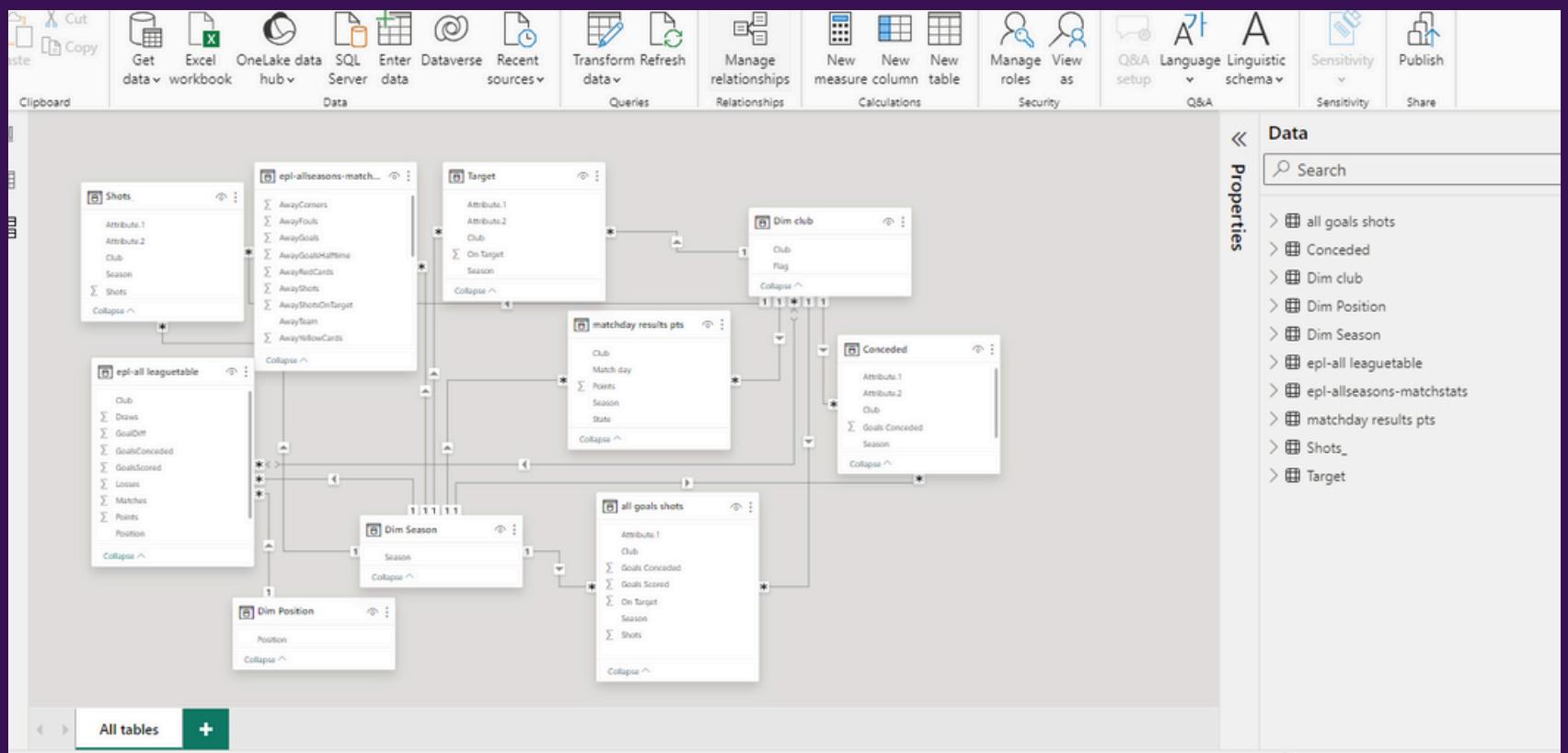
-Mariam Mohamed: Predictive Modeling, Visualization

-Rana Tarek: Final Report, Presentation Preparation



Premier League

DATA MODELING



Premier League

KEY MILESTONES:

1. FRAMING THE RESEARCH QUESTIONS

- Which clubs were the most consistent in terms of finishing in the top from 2010 to 2020?
- Which teams showed the most dramatic improvement or decline in their performance across the decade?
- How did goal differences (goals scored vs. goals conceded) impact the final league standings for each club?
- How did a club's defensive vs. offensive focus impact its overall league performance?
- Which teams had the best defensive records (fewest goals conceded) and how did that correlate with their league position?



Premier League

2-DATA COLLECTION AND EXPLORATORY DATA ANALYSIS (EDA)

- Collect data from multiple sources: match results, player statistics, team performance metrics, manager records, financial investments, etc.
- Conduct an initial EDA to understand the data structure, detect trends, and answer:
- What are the overall trends in goals, assists, and match outcomes across seasons?
- Are there any interesting patterns in the data (e.g., a team improving dramatically under a new coach)?



Premier League

3-CLEANING DATA AND FEATURE ENGINEERING

- Review the dataset for any missing values in important columns such as goals scored, match results, player stats, etc.
- Check for Duplicates: Ensure no match or player appears more than once in the dataset unless valid.
- Resolve Duplicate Entries: In case of multiple records for a match , consolidate them by averaging or merging relevant data.
- Date Formatting: Ensure that all match seasons are formatted uniformly
- Team Names: Standardize team names to ensure consistency



Premier League

4-ADVANCED EXPLORATORY DATA ANALYSIS AND ANSWERING QUESTIONS

- Analyze the effect of home advantage on match outcomes.
- Investigate whether player transfers during the season affect performance trends.
- Determine the top-performing players based on their contribution to team success.
- Use statistical tools (correlations, regression) to find links between team investment and performance.
- Examine if certain teams consistently perform better in high-pressure matches (against top-ranked opponents).



Premier League

5-CONCLUSION AND INSIGHTS

- Summarize the findings from the analysis.
- Highlight the most influential factors affecting team success and match outcomes.
- Discuss how the analysis can be used to predict future match results and identify potential trends for upcoming seasons.



Premier League

RECOMMENDATIONS AND POINTS OF IMPROVEMENT

- Suggest data-driven strategies for teams, such as focusing on specific game strategies based on historical success.
- Recommend areas for future research or data collection, such as gathering more detailed injury or player fitness data.
- Highlight potential improvements in the data pipeline, like integrating real-time analytics for live match predictions.



Premier League

ANALYZE DATA USING POWER BI

