Project Description
English Premier League Performance Analysis (2018–2023)

Project Overview:

This project analyzes the performance of English Premier League (EPL) players and clubs over five seasons, from 2018 to 2023. The goal is to provide a comprehensive, data-driven understanding of how individual players and teams have evolved over time. By leveraging historical data, the project aims to support decision-making for club management, coaches, analysts, and fans.

Business Problem:

Club management often lacks clear insights into player performance trends and team strengths or weaknesses. Decisions such as player transfers, training focus, and performance evaluation are sometimes made without sufficient quantitative evidence. This project addresses that gap by providing structured, actionable insights.

Objectives:

- Track and compare player performance season by season.
- Identify top scorers, key playmakers, and consistent defenders.
- Analyze team performance from offensive and defensive perspectives.
- Deliver an interactive, easy-to-understand dashboard for non-technical users.

Data Collection and Structure:

- Historical EPL data collected from official sources, including player stats, match results, and club information.
- Data stored in relational SQL tables covering:
    - **Players:** personal information and match stats
    - **Clubs:** seasonal rankings and team statistics
    - **Seasons:** yearly aggregation of matches and outcomes
    - **Stats:** goals, assists, cards, and minutes played
- CSV/Excel files were used initially, then cleaned and transformed into structured SQL tables.

Data Cleaning & Preparation:

- Removed duplicates and handled missing values.
- Converted data types and standardized dates.
- Prepared aggregated tables for analysis of both individual players and clubs.

Analysis and Methods:

- Python (Pandas, Matplotlib, Seaborn) for data cleaning, analysis, and visualization.
- SQL for querying, joining tables, and calculating aggregated metrics.
- Power BI for building dashboards and interactive visualizations.
- Delta tables created to track season-over-season performance changes.

Team Roles and Contributions:

- **Data Extraction:** SQL queries to extract and join relevant tables.
- **Data Cleaning:** Handling missing data, duplicates, and preparing final datasets.
- **Season Summary:** Calculating totals and deltas for player metrics.
- **Club Performance:** Analyzing top clubs' offensive and defensive performance.
- **Player Performance & Visualization:** Tracking trends for top players and creating visual charts.
- **Documentation & Presentation:** Compiling insights, writing reports, and preparing the final presentation.

Deliverables:

- Fully documented GitHub repository with all project code.
- Interactive dashboards visualizing player and club performance trends.
- Insights into top performers, clubs, and season-to-season performance variations.

Impact and Value:

- Provides club managers and analysts with actionable insights for strategic decisions.
- Makes EPL data accessible to non-technical audiences.
- Highlights performance trends and exceptional players to guide scouting and coaching strategies.