

Digital Egypt Pioneers Initiative

Final Project Proposal

**Project Name / Premier League Performance
Analytics (2018–2023)**

MNF3_DAT1_S1

Business Problem:

- The team's management lacks a clear understanding of how player performance has evolved over the seasons. They need data-driven insights to support decisions such as player transfers, training focus, and performance evaluation.

Proposed Solution: We will use historical player data from 2018 to 2023 to perform a comprehensive analysis.

The project will:

- Compare player performance season by season.
- Identify top scorers, key playmakers, and consistent defenders.
- Track performance trends across positions and clubs.
- Deliver a dashboard that visually summarizes the key insights for decision-makers.

Objectives:

- Collect and clean Premier League player and team data (2018–2023).
- Build season-by-season comparisons to identify key performance changes.
- Extract and visualize top scorers, assist leaders, and best defensive teams.

- Develop interactive dashboards for club-level and player-level performance.
- Generate data-driven insights and actionable recommendations for decision-makers.

Project Description

- This project aims to analyze **English Premier League performance data from 2018 to 2023**, extracting meaningful KPIs, patterns, and insights for decision-makers, including coaches, analysts, and sporting directors.
 - The analysis focuses on player and team performance over time, comparing offensive and defensive metrics across seasons. The final deliverable will include interactive dashboards and reports summarizing key findings and actionable recommendations.
-

Group Members & Roles

1.Hanen Hosam EL-Khishen (Team Leader)(hanenhosam16@gmail.com)

- Documentation & Presentation - Visualization - checking the team's work - Data Extraction

2.Hager Ayman

- Season Summary & Delta Tables - Data Extraction

3.Malak Ehab

- Data Cleaning & Preparation -Data Extraction

4.Abdullah Atef Elshazly

- Club Performance Analysis - Data Extraction

5.Ahmed Mahmoud Alhawata

- Player Performance & Visualization - Data Extraction

6.Ahmed Mosad

- Data Extraction

Business Questions:

1. Who are the team's top scorers for each season?
 2. Which players are potential sell/buy candidates based on their contribution and market value?
 3. Which teams have improved or declined the most in their offensive or defensive performance season over season?
 4. Which players deserve to be signed based on consistent and effective performance?
 5. Which players show a significant decline in performance from one season to another?
-

Tools & Technologies:

1 - Data Processing

- Python (Numpy-Pandas) / Data cleaning, aggregation, and KPI calculation

2- Storage

- Excel - CSV - SQL / Storing and Organizing Datasets

3- Visualization

- Power Bi - Tableau / Interactive dashboards and insights

4- Analysis

- DAX - Matplotlib - Seaborn / Trend and performance analysis

5- Collaboration

- GitHub / Version control and teamwork coordination

Milestones & Deadlines:

The project will follow a structured timeline where each phase takes 4 days, followed by a one-week full review of the entire project to ensure accuracy and quality before submission.

- Data Collection (4 days – until 16 Oct 2025):
Gathering all player and match data from seasons 2018 to 2024.
- Data Cleaning & Preprocessing (4 days – until 20 Oct 2025):
Handling missing values, removing duplicates, and ensuring data consistency.
- Exploratory Data Analysis (4 days – until 24 Oct 2025):
Exploring data patterns, distributions, and identifying early insights or anomalies.
- Business Question Mapping (4 days – until 28 Oct 2025):
Defining the key business questions and linking them with the appropriate data columns.
- Data Analysis & Insights Generation (4 days – until 1 Nov 2025):
Performing season-to-season comparisons and identifying major performance trends.
- KPI Calculation (4 days – until 5 Nov 2025):
Computing key performance indicators such as goals, assists, clean sheets, and player rankings.
- Dashboard Design & Visualization (4 days – until 9 Nov 2025):
Building interactive dashboards to visualize results clearly and intuitively.
- Final Documentation & Report Writing (4 days – until 13 Nov 2025):
Compiling all findings into a structured final report with

- Full Project Review & Quality Check (1 week – until 20 Nov 2025):
Reviewing the full project, refining visualizations, validating results, and preparing the final presentation.

KPIs:

To evaluate the project's performance and ensure high-quality outcomes, the following KPIs will be used:

- Data Cleaning & Processing:
 - 100% of missing and duplicate data handled accurately to ensure dataset integrity.
 - Full consistency between different seasons' data formats and structures.
- Analysis & Insights:
 - At least 90% of business questions answered through data analysis.
- Visualization & Reporting:
 - Dashboard load time should be less than 3 seconds for optimal usability.
 - At least 80% of users should be able to navigate the dashboard without external help.
 - Clear and interactive visuals that highlight top scorers, assists, and clean sheet trends.
- Final Documentation & Presentation:
 - 100% completeness of the final report and presentation materials.
 - At least 3 actionable recommendations based on data insights.

Expected Outcome:

A comprehensive analytical system summarizing (2018-2023) Premier League seasons, featuring interactive dashboards and insight reports.

It will allow decision-makers to:

- **Identify consistent top performers.**
- **Monitor performance trends per club or player.**
- **Improve recruitment and training strategies based on data.**