

# Premier League Football Analytics Database Project

---

Schema Squad

# Intro

- Developed a dual-database system for football analytics.
- Operational DB for match data entry & Analytical DB for trend analysis.
- ETL pipelines to extract, transform, and load data from CSVs.
- Supports complex analytical queries and reporting.

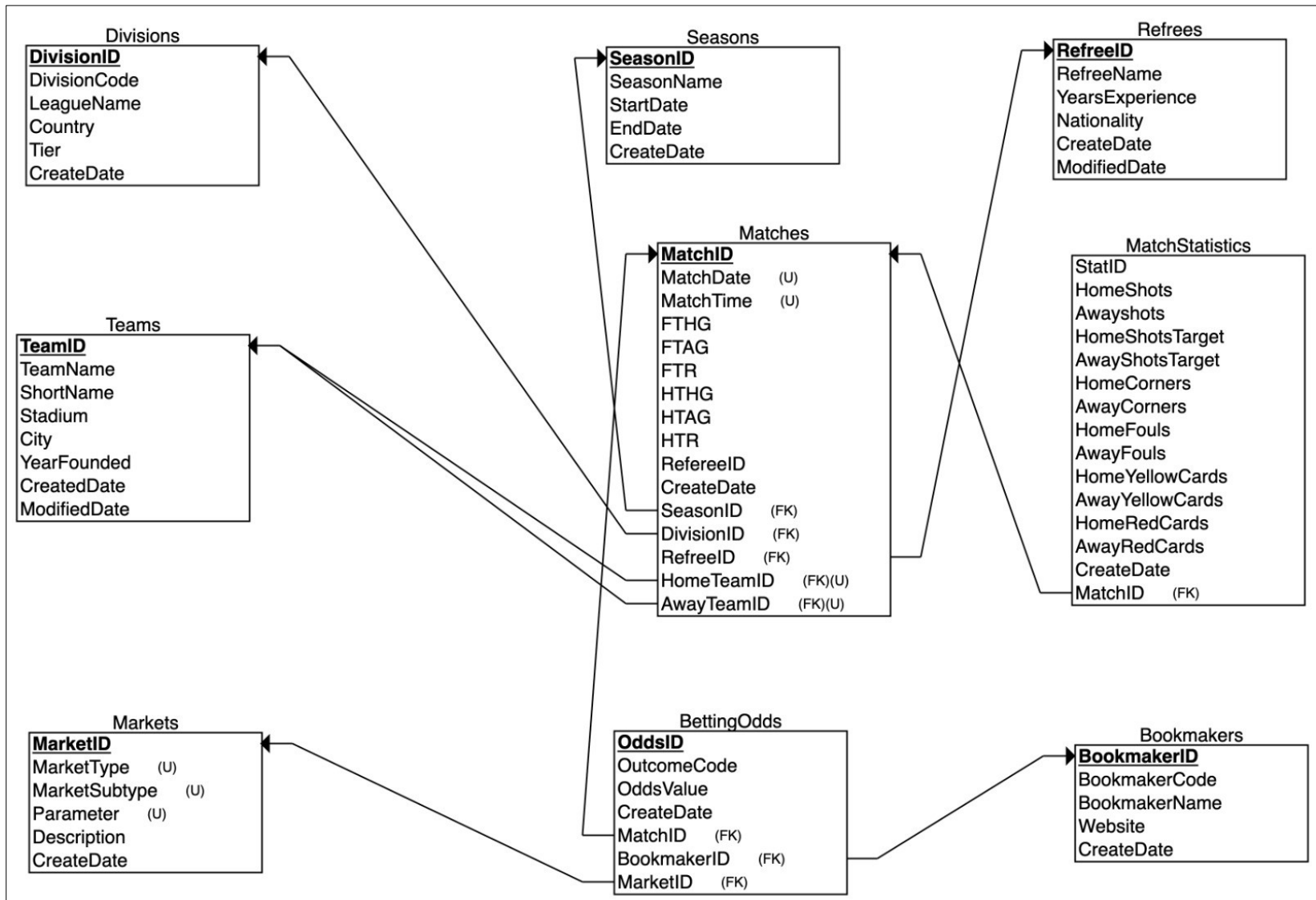
# Intro

- Fully-functional OLTP database with normalized schema.
- OLAP star schema for analytical queries.
- ETL pipelines for data ingestion and transformation.
- Robust query support for match, team, and referee analytics.

# Operational Database (OLTP)

- Key tables: Teams, Seasons, Referees, Divisions, Matches, MatchStatistics, Bookmakers, Markets, BettingOdds.
- Normalized design ensures data integrity and efficient transactions.

# Relational Schema



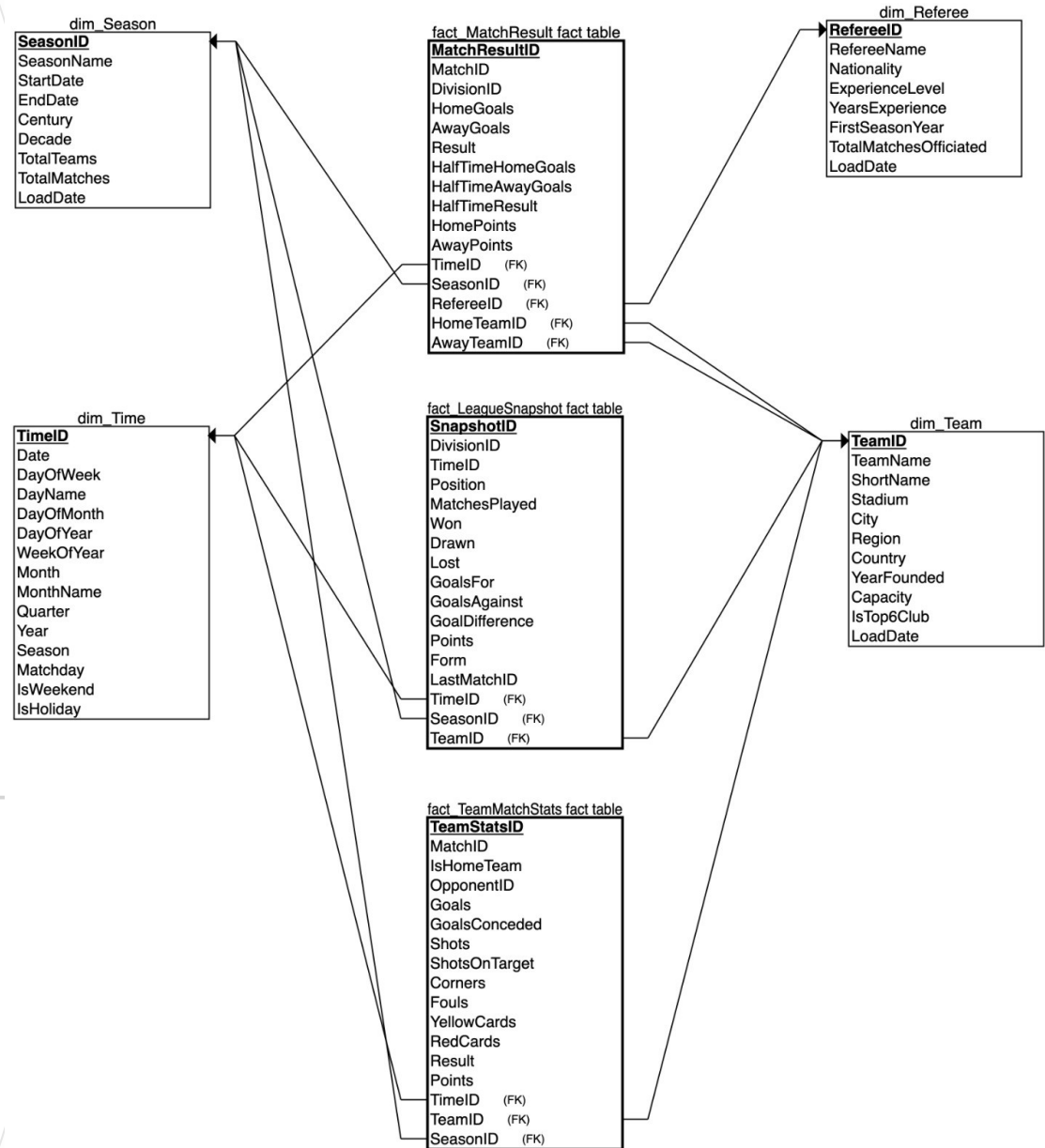
# ETL Process

- Extraction: Load CSVs, validate data, track processing metadata.
- Transformation: Clean data, derive metrics, handle foreign key relations.
- Loading: Populate OLTP database with clean data.

# Analytical Database (OLAP)

- Star schema: Fact tables - MatchResult, TeamMatchStats, LeagueSnapshot.
- Dimension tables: Time, Team, Season, Referee, Location.
- Supports historical trend, seasonal performance, and regional comparison.

# Star Schema





# Technical Challenges & Solutions

- Reserved words resolved with backticks in MySQL.
- Adapted schema for variable CSV formats.
- Ensured FK integrity via transaction management.
- Optimized OLAP queries using pre-aggregated fact tables.



# Key Analysis Capabilities

- Regional performance trends and home advantage insights.
- Comparison between top clubs vs. others.
- Seasonal performance breakdowns.
- Referee influence on match outcomes.



**DEMO**



# Conclusion

- Robust, scalable system for football analytics.
- ETL and data modeling demonstrate advanced database skills.
- Supports powerful insights for teams, analysts, and fans.
- Real-world application of OLTP + OLAP integration.