

**Project Description**

The world football federation has decided to build a database to record detailed information about its competitions, matches, clubs, and players. This system will serve as the foundation for generating statistics, analyzing player performance, and tracking the history of football events over time. In the first phase of the project the task is to create a data model (conceptual and physical) that allows the storage of the required information to fill the requirements described below.

**Requirements**

A football competition, such as the Premier League or the Champions League, must be stored in the database with attributes such as its name, the country (or countries) in which it is organized and its type (e.g., national league, national cup, international competition). Each competition is divided into seasons (e.g., 2024/2025), and every season must include a start date and an end date.

Each football club that competes in a season must also be recorded. A club is described by its name, the country it belongs to, and the year it was founded. It should be possible to find out which titles a club won and in which seasons. Clubs are managed by coaches, referred to as managers in the database. The system must not only record which manager currently leads a club but also keep a complete historical record of all past managers, including their start and end dates of employment.

Players are central to the system. For each player, the database must store their name, date of birth, physical attributes such as height and weight, and their current market value. Since players frequently change clubs during their careers, the system must also record transfers. A transfer involves a player moving from one club to another on a given date, with a transfer fee and a contract length in years. The system must allow the reconstruction of a player's full transfer history and determine their current club based on the most recent transfer.

Matches are at the heart of football statistics. A match belongs to a specific season and always involves two clubs: the home club and the away club. Matches must include attributes such as the matchday number (e.g., round 6 of the Portuguese League), the date of the match, and the attendance (number of spectators).

For each match, the system must also store the lineups of both clubs. A lineup connects a player to a specific match and specifies the position the player played (e.g., center midfielder, striker), and whether the player was a starter or a substitute.

During a match, many actions occur. The system must record every action with details about the match in which it occurred, the player who performed it, the minutes and seconds, and the type of action. Examples of action types include (but not limited to) goals, shots, passes, fouls, and cards. In addition to actions, the system must store/calculate match statistics for each player,

such as the number of minutes played, passes completed, shots on target, and goals scored. The final result of a match could (and should) be obtainable using the goals stored in actions.

This football statistics database must therefore integrate information about competitions, seasons, clubs, managers, players, transfers, matches, lineups, actions, and player statistics. With such a structure, it will be possible to answer important questions such as which player scored the most goals in a season, which club has spent the most on transfers in the past five years, or which manager has had the longest tenure at a single club.

### **Deliverables (PART I)**

The model that will result from the answers to the requirements set out above can always be expanded as students identify new questions that enrich the model. The requirements description is obviously incomplete, and the identification of entities, attributes, or other design decisions, not explicitly mentioned, are considered in project's part I objectives. Additionally, in this first delivery, there may be restrictions, implicitly or explicitly identified in the requirements, that cannot be represented, and their implementation will have to be postponed to the business rules definition phase that will be part of the 2nd delivery. It is the students job to be able to separate what can be represented in the data model and what cannot be represented at this stage.

The first delivery consists of the design of the entity-relationship and relational models<sup>1</sup>, and the production of SQL code (DDL) instantiated to Microsoft SQL Server as indicated below:

1. Conceptual Data Model (Entity-Relationship Model) developed in Power Designer (.cdm file).
2. Physical Data Model (Relational Model) developed in Power Designer and obtained using Microsoft SQL Server specifications (.pdm file).
3. SQL script generated in Power Designer with SQL statements for database objects creation in SQL Server (.sql file).
4. Text file (.txt) with the identification (name and number) of all the group elements.

### **Additional Guidelines (PART I)**

- The project can be implemented individually or in groups of a maximum of five (5) students.
- **PART I** deadline is **October 19, 2025**. For each day of delay in the delivery of the 1st part of the final work, 2 points (in a scale from 0 to 20) will be taken from the Part I score. Three days after the deadline (October 22, 2025) the delivery is not accepted, and the group will have a grade of zero (0) in this part of the project.

---

<sup>1</sup> Power Designer Conceptual Data Model and Physical Data Model, respectively.