

Instituto Superior de Engenharia do Porto MSc in Informatics Engineering

Armazéns de Dados

The Portuguese Football Association owns an information system that stores data about the main football championship of this season. More specifically, about a player is stored: the inscription number in the Association, the full name, the nickname by which the player is known in the football world, its position, the address, the zip code, the date in which the last contract made, the contract duration (in months), and the value of the termination clause. Besides these data, it is also stored the player's current club. About the club is stored: the registration number in the Association, the name by which it is known, the address, the zip code, and the foundation date. Each postal code corresponds to a city.

The football championship is organized into journeys in each season. Each journey happens around the weekend, being characterized by a description (e.g., 8th journey), by the number of football matches happening on Friday, on Saturday, on Sunday and on Monday. Each season is characterized by a unique identifier, by a description (e.g., 2014/15 season), by the number of journeys, and by the official dates of start and end of season.

The Portuguese Football Association's information system stores data about the several events related with each player, in the scope of each journey in which he participates/plays and the time of the match (e.g., 67 minutes and 23 seconds) when they happened. Each event is characterized by a unique identifier, a description (e.g., fault committed; fault suffered, goal scored; missed penalty; assistance for goal; missed pass; yellow/red card), and by the information if the event represents an infraction (e.g., a fault) or not (e.g., a missed pass) to the football rules.

Following *Kimball* methodology, develop the dimensional analysis process in order to define and create the conceptual schema for a data mart that supports different data analysis to the football events. Present all the facts, dimensions, granularity and all the relevant issues for the data mart project.

As examples, consider the kinds of analysis to be performed:

- Total of goals scored, throughout the several journeys, already in extra time (i.e., after minute 45 or minute 90).
- Total of suffered faults in the 1st, 2nd or 3rd quarter of hour, of the 1st or 2nd part, by a given player, in a given journey.