Máster en Big Data, Inteligencia Artificial e Ingeniería de Datos

Modulo 5: Data Analytics

Classroom Practical Exercise I

Part I: Conceptual and Logical Design

1. **Develop a logical design from a Data Mart with the star schema, add the attributes in each table suggesting if the attributes are a primary or external key. Also, show the hierarchy levels that you consider for each dimension:**

   The Spanish football team organization has decided to set up a Data Warehouse to analyze trends of teams and try to improve their results. To do this, management wants to analyze the information available on games played in the league.

   Information analysts want to know relevant data about matches such as players in each match, the goals scored by the player, and the goals that she has avoided. They also want to know information regarding the yellow and red cards received by each player in a given match. Finally, they want to know the time spent by each player playing.

   In addition to the information about the performance of players, analysts want to collect other data, such as the name of the player, his position, his age and his height. Likewise, they want to group the players in teams to analyze the performance of the different teams in the season. For each team, it is known their name and the value the team is worth in euros.

   On the other hand, we want to know information about the stadium of the match, in order to analyze if the location influences the result. The stadium has a maximum capacity, a name, the quality of the playground, year of inauguration, city and the country where it is located.

   In addition to these data, the coach has indicated that she is interested in knowing data regarding referee who supervise each match to be able to reclaim a fair arbitration. To do this, she wants to include the name of the referee, his age, his country of origin and the organization which she belongs to.

   Finally, management has decided that it is also important to store the date of the match, including information about the day, month and year in which it is celebrated and the weather.

   Note: Remember that multiple roles can be assigned to the same dimension, especially, if it participates several times in each instance of the event. If there is no degenerate dimension to convert the relation into a type 1-m, you will need to use bridge tables to solve m-m relationships between facts and dimensions.