CS-3200 Homework #5

Creating and manipulating a MySQL schema for football players

Domain Description
The attached file contains information on the soccer (football) players for the 2011/2012 Premier League. The Premier league is an English professional football league. The Premier League is the most watched sports league in the world, broadcasted in 212 territories to 643 million homes and a potential TV audience of 4.7 billion people.

A game of football is played with 2 teams and each team plays with 11 players at one time. The team (club) comprises of 1 **goalkeeper**, **defenders**, **midfielders** and **forwards**. The main objective of the game is to score a goal into the other team's net. Each club generally has a squad of 33-34 players.

Data File Description
There are three different csv files. Each row within the player's file is data specific to one player. Each row within the stadium file is specific to each team's home stadium. Each team has only one home stadium. Each row within the manager's file is data specific to one manager. A team may have one to many managers.

The data fields in the player's file (EPL_dataset.csv) are the following:
Player: the player's name
Club: the name of the club the player is assigned to
Nationality: the player's nationality
Position: the position most played by the player
Appearances: number of matches the player played
Minutes_Played: the total number of minutes played by the player
Goals_Scored: number of goals scored by the player
Yellow_cards: the number of times a referee showed a yellow card or a caution card to a player. A number of yellow cards can result in a ban (or a red card) from the game or a red card for the player.
Red_Cards: the number of times a player is banned from playing. The number of matches the player is banned from varies depending on the number of red and yellow cards.
Clean_Sheets: clean sheet refers to a defender's or a goalkeeper's preventing the opposing team from scoring during an entire match. This statistics is not applicable to midfielders or forward players. A value of -1 means the statistics is not applicable to this player.

The data fields in the stadium's file (EPL_stadium.csv) are the following:

Team: the team or the name of the Premier League team
Venue: the name of the team's home stadium
Each team has one home stadium.

The data fields within the manager's file (EPL_manager.csv) are the following:
Manager: the name of the manager for the team.
Team: the name of the Premier league team
Nationality: the nationality of the manager

Assignment Description

1. Design the schema for the Premier Football League. This should be done in the MySQL workbench. Create a pdf of the EER model. (10 points)
2. Create a database for your schema named premierlastnamefirstnameinitial. You may create the tables' schemas using the SQL CREATE command or using the EER modeling tool. (20 points)
3. Import the .csv file into your tables using the Import table wizard. Compose queries to answer the following questions:
4. Generate a list of players who have scored more than 10 goals. The result should contain the player's name, team and number of goals.
5. Generate a list of managers, teams and the number of goals for each team for the season. The result should contain the manager's name, the team name, and the number of goals.
6. Determine the most popular nationality for the players. The result is a single value.
7. Determine the number of yellows cards given per team. The result should contain the team name and the number of card.
8. Determine the top 5 goalkeepers ranked by clean sheets. The result should contain the player's name, the team's name and the count of the clean sheets. The result should be ordered in descending order by the count of clean sheets.
9. Determine the number of minutes played for each of the player. The result should consist of a player's name and the number of minutes.
10. Generate a list consisting of a manager's name, a team's name and its home stadium, as well as the total number of red cards for issued to the player's of the team.
11. Determine the top 5 forwards ranked by minutes played. The result should consist of the player's name, the team's name and the total minutes played.
12. Are there any players from Ireland (either Northern or Southern) with more than 10 goals?
13. What is the average playing time for the players? The result is a single number.
14. What is the average playing time for each of the positions. The result should contain the position and the average playing time for that position.
15. Determine the players with above average playing time. The result should contain the player's name, the team's name and the player's playing time.
16. Generate a self-contained extract of your database to blackboard using the 'data export' tool from the 'Server' menu. Make sure you include the create schema as well as other objects in the database. We must be able to import your schema so please ensure the extracted file works with import.

Assignment submission

Create a zip file named hwk5lastnamefi.zip that contains 3 files :

- hwk5EERlastnamefi.pdf (picture of the model)
- hwk5footballdumplastnamefi.sql (export of the created schema)
- Hwk5footballquerieslastnamefi.sql (SQL commands for the queries)

Submit the zip file to blackboard.