Peter Mensah, Cameron Jackson, James Miller

INST327 - Section 0102

Final Report

Nov 29, 2022

Introduction:

Our team decided to create a database relating to the European soccer league with the focus being on teams in the English Premier League from the year 2016. As huge soccer fans we thought this idea would be great to formulate a final project on incorporating what we have learned this semester. Oftentimes it is hard to compare players so using a database seemed like a great idea.

Our database is sampled from using FUTHEAD ratings where anyone who uses the database will be able to compare data from the 2015-2016 and compare players, teams, team seasons, records, positions, and games. The purpose of our database is to cater to soccer fans and to emphasize that you can always incorporate your favorite sport into a database and do some great database designing.

Database Description:

Our database was created around the 2015-2016 Premier League season. It was also created to house brief information on the 2016-2017 season as well. The Premier League is the top flight of soccer in England and is considered by many to be the best league in the world for competition. Our database is a tool for people to look primarily into the 2015-2016 season and see various information and statistics about the season. Our database consists of tables that provide teams and their correlating Fifa ratings, the standings of the league at the end of the season, players for each club and their correlating Fifa ratings.

Logical Design:

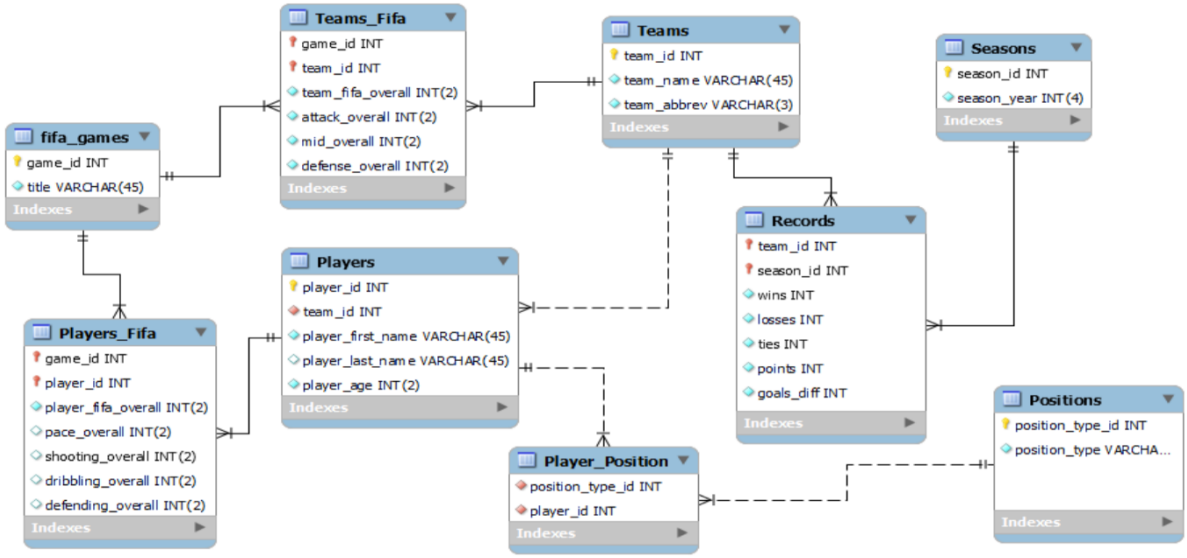

Image of our Entity Relationship Diagram for European Soccer League database

The Entity Relationship Diagram design of our database includes nine tables. Each table is designed to hold European Soccer League data particularly to a team, player, season, and Fifa video game. Relationships are maintained among tables via linking tables, consisting of one-to-many relationships. Primary keys from their perspective tables are added into linking tables as foreign/composite keys in order to maintain relationships among two tables. While inside of the linking tables, the foreign/composite keys also serve as primary keys, making other attributes inside dependent upon them. By repeating the process of incorporating linking tables for every two tables where it is appropriate, we were able to maintain proper one-to-many relationships for every table in our design, thus allowing for efficient creation of queries.

The group put much emphasis into building a well organized and efficient design for our European Soccer League database. With the amount of effort added into the Entity Relationship Diagram design, we were able to create a structured design of nine tables. By implementing nine tables into the design, we established a solid foundation to being able to capture essential data records regarding the European Soccer League that many fans seek to locate. In addition to being able to capture essential data records, the ERD design made for fluent creation in queries.

Physical Design:

For the physical design of the European Soccer League database, our goal was to acquire sample data that most efficiently would represent our foundation previously created in the ERD design. With the acquired sample data we were able to accomplish that goal, by implementing 20 records of European Soccer teams in the Premier League and 520 records of each player recorded on those teams. The primary season of focus consisted of the 2015-2016 season, but we additionally added brief 2016-2017 season data in order to maintain one-to-many relationships. Each player is matched with one or more positions and one or more Player Fifa profiles

depending on the title of the Fifa video game. In comparison, each team is matched with one or more season records and one or Team Fifa profiles depending on the title of the Fifa video game. Therefore, this physical design of our European Soccer League database gives accurate and ethical data records of the 2015-2016 and 2016-2017 seasons for fans who seek to obtain information regarding the subject.

Sample Data:

Originally we believed that we could retrieve all the data from a CSV file; however, due to our database design and usability we had to move away from this idea. In order to find the data for this database we used two websites: football reference to find which players played for our team and futhead in order to find the fifa ratings. We had to add all the data manually into the database because of this. Teams also had players who played no games or very few for them or transferred halfway through the season so we had to take in account these variables. It was a hassle for our team to find all the pertinent data but in order to create a database which is usable and has correct data it was important. Below is an example of our players table:

player_id	team_id	player_first_na...	player_last_name	player_age
1	1	Cesar	Azpilicueta	25
2	1	Branislav	Ivanovic	31
3	1	Cesc	Fabregas	28
4	1	Willian		26
5	1	Nemania	Matic	26

Views / Queries:

Query Name	Join	Filter	Aggregate	Linking	Subquery
greatTeams	**X**				
playersOlderThanThirty		**X**			
fifa16DefensiveTeams	**X**	**X**			
aboveAverageAttackers	**X**	**X**	**X**	**X**	**X**

Descriptions of of queries:

Query 1: Creates a view all teams considered "great" (won more than 20 games in the season as well as scoring more than 30 points)
Query 2: Creates a view of all players that are over 30 years old and orders them in alphabetical order by their first name
Query 3: Creates a view of teams that have a defense that's overall is greater 75 from the season '16
Query 4: Creates a view of all offensive players who are considered "above average" attackers by joining positions and player_id alongside a subquery that shows from the specific team Chelsea

Changes from Original Design:

Ultimately we truly did enjoy doing this project and felt that it better helped us understand SQL and to gain the actual experience to feel confident using SQL. Originally we had to make a few changes to our database such as deciding to use smaller sample data and to only focus on the English Premier league. We learned that the amount of data in a database is a huge factor, especially when conducting proper data ethics and integrity. Futbol/Soccer has 11 players on the pitch but roster sizes have ranges of anywhere from 20 to 40 players and with that our team decided it would be best to just stick with the English Premier League at the time. In addition, we also substituted using a matches table with a records table in order to hold team statistic totals at the end of the season. By not including a matches table, we were able to bypass the process of needing two primary keys in the teams table such as, 'home_team_id' and 'away_team_id'. In an attempt to maintain one-to-many relationships, our team changed our ERD to include a seasons and fifa games tables in order to prevent one-to-one relationships. By incorporating these extra tables we were able to maintain one-to-many relationships via linking tables.

Database Ethics Considerations:

With our database having to relate to a professional sports league we had to ensure that we were being fair and transparent with the data that was included in our database. We want users to trust our database so we ensured to use FUTHEAD that provided fair and honest ratings for players and teams instead of providing our personal opinions on what the ratings would be. Soccer has decades and decades of history and trends so we ensured our sample data was in a good range of years.

Lessons Learned:

This project taught us a lot about how important collaboration and time management are. From our first group meeting, having to figure out a weekly plan to always meetup and communicate in our group chats was no easy task but we all had a common goal so we were able to complete a project we were all proud of. We played to each others strengths and rather that being one person being more comfortable with coding with SQL, one person good with communicating with the AMP's, TA's, and Professor Duffy, or one person who just really likes making reports, we found a way to have great balance, communication and respect for each other.

We would say collectively the hardest part was importing our data into a spreadsheet mostly due to having to make sure information matched from spelling to the actually data itself was very time consuming and took us the longest to complete but as we learned hard work always pays off and we were able to stay on top of things and not have to worry about not meeting the deadlines. Lastly, we learned that starting early is the best way to make sure you are doing things right. We often completed tasks ahead of time in order to receive feedback or comments from the instructional team that helped us recreating tables and linking to create a great ERD and database. In conclusion, we worked well together and enjoyed the learning process

Potential Future Work:

We felt as a team that this was only the beginning of what could be explored with our database, adding all the major soccer leagues in one database would help a company like FIFA have an all international collection in order to pull whatever information they would need for any given season of their choosing. Secondly, adding women soccer leagues would also provide greater usage of our database where views could be created on the desired user to see more information on players, teams, etc. Lastly, even for the casual user could use this as a way to educate themselves of the game of soccer and all the leagues.

Works Cited:

Futhead. (n.d.). *FIFA 16 Ultimate Team Database, players & squad builder*. Futhead. Retrieved December 10, 2022, from https://www.futhead.com/

***Football statistics and history*. FBref.com. (n.d.). Retrieved December 10, 2022, from https://fbref.com/en/**