Nick Hornung, Kevin Sutcliffe, and Rob Clancy

Dr. Weese

Database Systems and Concepts

26 April 2021

**Introduction**

The database we created has information based on past FIFA World Cups. The database holds information about each team in the tournament which includes players and managers of each nation represented in the World Cup. The database also has the statistics of each World Cup. Such as who won the tournament that specific year, the nation where it was held, and the start and end date of the tournament. Not only does it hold statistics about the big details of the tournament but has specific data for each player and each game. The database has each player’s name, age, height, nationality, and position. To go along with that we are keeping track of some of the statistics for each player such as how many goals they have scored in the competitions. The database is also keeping track of all the games played in these World Cups.­­­ For each game we are tracking which World Cup it was in, the stadium it was played in, the two teams playing, winner of the game, how many people were in attendance, date of the game, and what stage of the tournament the game was played in. Such as group stage, round of 16, quarterfinal, semi-final, and the final. Along with game statistics we are also keeping track of the statistics about the broadcaster of the game. Such as what network was broadcasting the game, what channel the network put the game, and how many views the broadcast got.

There are many potential users for this type of database. The users could range from someone who just turned on the television and the World Cup was on to a high-level soccer analyst reporting on the World Cup to an everyday sports gambler. Anyone of these people could use our database for different reasons. Even someone who does not like sports but likes history might want to use our database because of how popular this event is many people might want to see the history of the competition. We would call someone who just randomly wants some information on the World Cup as a casual user and the reason they would use our database would be just to get basic information on the past World Cups. A Sports analyst would want to use our database because of the statistics we have on the world cup. They would want the stats we have because they can go and look at that information and take it back to the teams to show them what was working and what was not. From that data they could make decisions about what they should try and do to outperform the rest of the competition in the next World Cup to win it all. The last user we mentioned was a sports gambler. They would want this information because that would give them an extra insight to try and beat the sport bookies. They could go look at who has won the most World Cups and use that information to see if they want to gamble.

**Technical Description**

*Backend*

For our database, we used multiple different tools to build and increase the efficiency of our program. We started by outlining our design using the diagram tool draw.io to create a basis for how our backend should be structured. We drew a UML based system design, and this was useful to see what data columns we needed and the key constraints needed for each component. We then implemented the backend data part of the database using the Microsoft SQL Server Management Studio integrated environment to input data and create data queries. Data was stored and hosted individually on localDBs by using SQL Server Express 2019 within our local machines. Since there was no way to send functional databases to each other and remote desktop access was restricted, local databases were most effective. The queries to first initialize and create data were written in standard SQL, which uses data manipulation commands included in the Data Definition, Data Manipulation, and Data Query languages. DDL is used to create, drop and alter data; DML is used to insert, update, and delete data; and DQL is used to select the data. For our functionality and report queries, we used more of the Transact structured query language, T-SQL. This allowed us to extend the functionality of standard SQL by giving access to ranking and aggregate functions to manipulate data. These functions were helpful to find specific statistics within our data like average goals or the top team within a certain category.

*Frontend*

After we finished designing the backend section of our database which holds and manipulates data, we moved on to the user interface. This allows users to interact with the data and find reports about past World Cups. We started by using the Visual Studio integrated development environment to create a Windows Presentation Foundation application and lay out model data and data access to be used in the interface, written in C#. The WPF or visible portion of the user interface was then built using Visual Studio Blend, Microsoft’s user interface design tool. This tool was chosen to simplify the action of adding window components to a point and click action, which helped reduce redundant actions for views. The format and specific design details for the interface were then adjusted using XAML, a markup language for .NET application designing. The library, Windows Presentation Foundation ToolKit, helped extend the functionality of the XAML because it included the most updated design features for applications, like property grids and combo boxes which were needed for the most convenient user design. A different XAML view or page was created for each component of data in our database, using C# model classes to populate the lists. To connect the front and back end, we used an ADO.NET. All of these tools were then shared and edited between us using GitHub.

*Backend/Database:*

Languages - Standard SQL, Transact - SQL

Tools - SQL Server Management Studio

*Frontend/UserInterface:*

Languages - C#, XAML

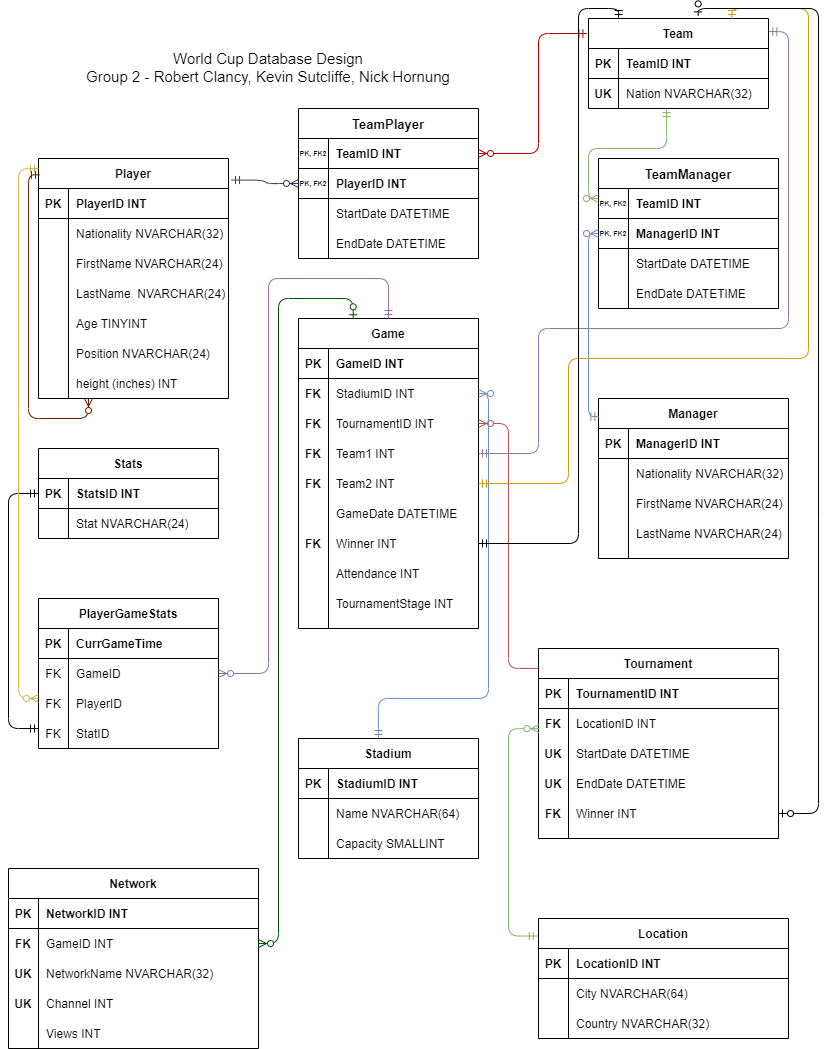
Tools - Visual Studio, Blend,

Libraries - WPF ToolKit

*Code/Repository Sharing:*

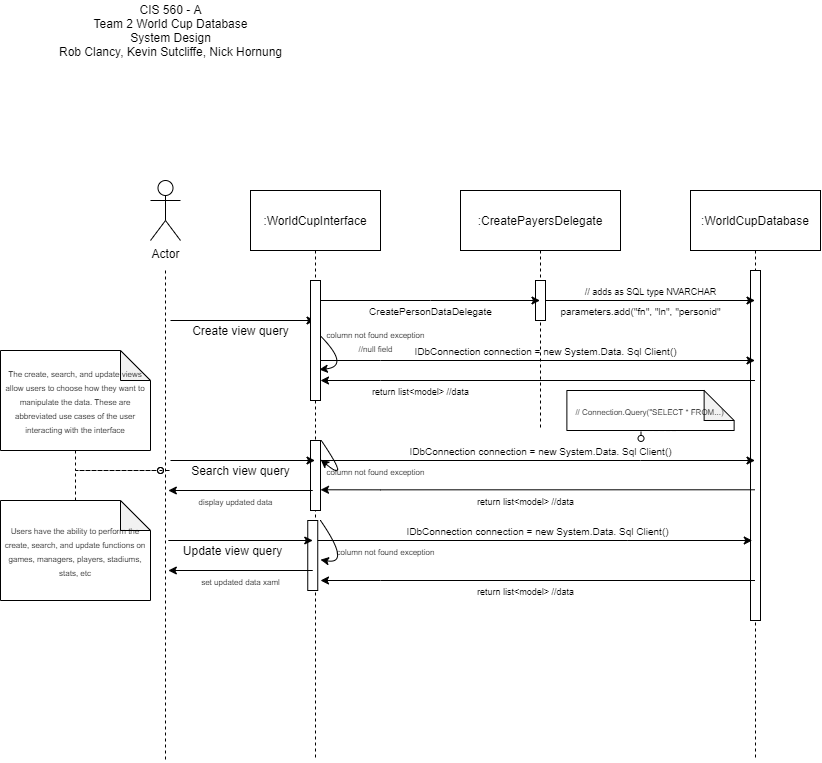
Github

**Database Design**

****

**System Design**

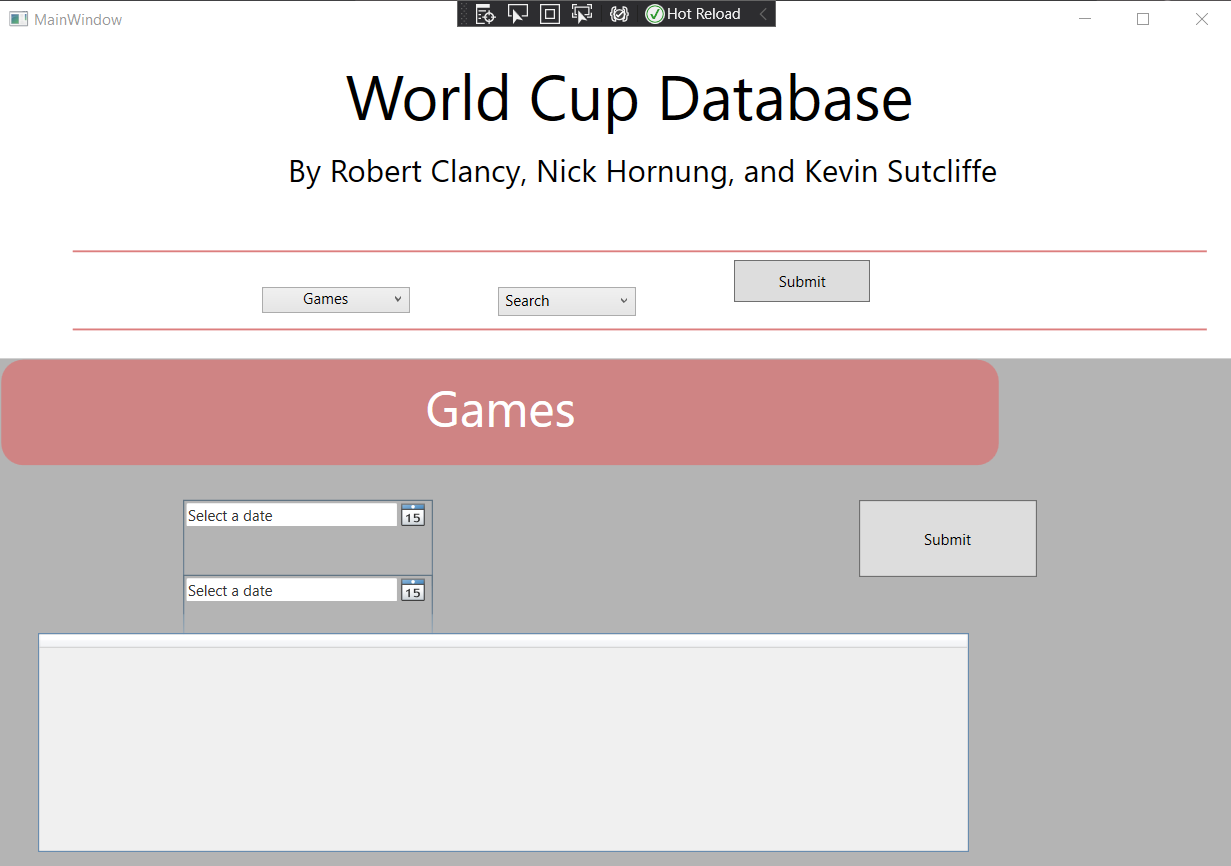
Below is a call sequence diagram to help show the interaction between the interface and SQL database:Users should be able to interact with a main interface to select whether they want to create, search, or update data through an SQL query. The request travels through data access layers and through a SQL server connection to perform the function and receive an output directly back.



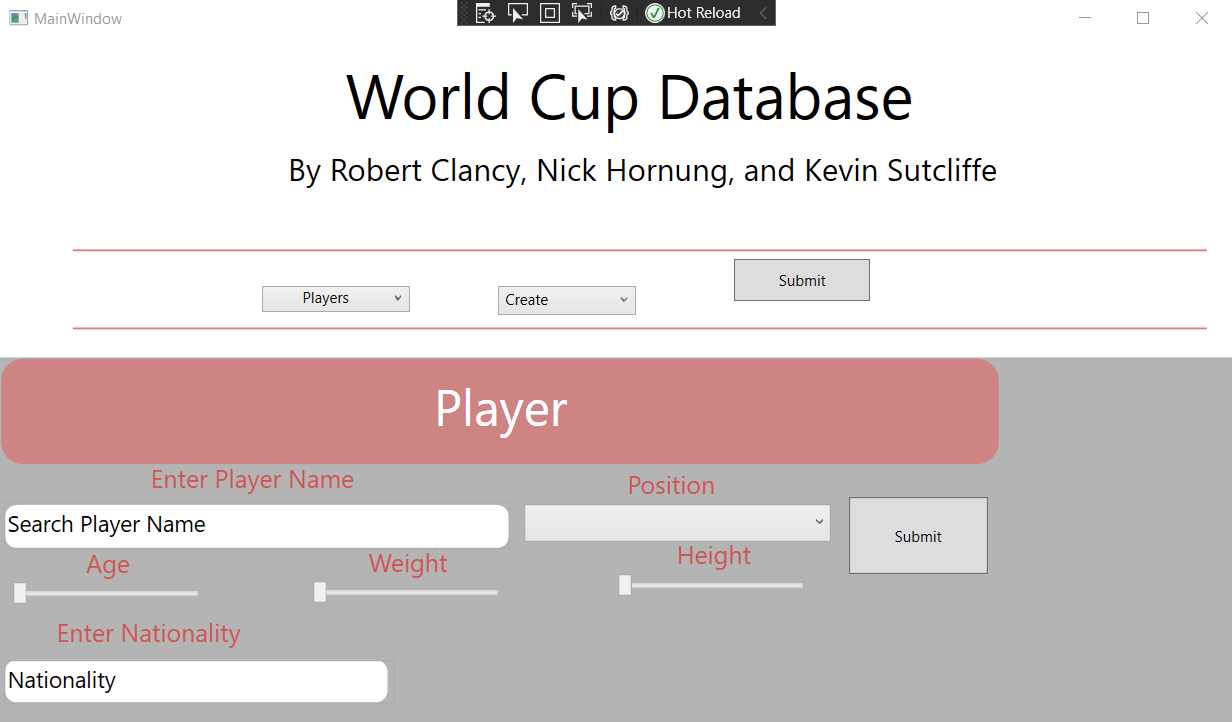
**System Features and Usage**

The system we created to use our database is a very easy to use user interface. When we set out to do this application we knew that we wanted something that was very easy to understand and use. In doing it this way our application is pretty limited. There are three main ways you can use our application with our database. The three commands that are usable are to search, modify, or create data within our database system.

The first command we will talk about is our search command within our application. Within the search command there are subcategories that you can search from by selecting the drop down menu shown in the picture below. Once you select the subcategory you want to query it will take you to the specified page. This is where we run into our first major limitation within our application. This limitation is that you can only search one table at a time for example if you are about to search for a game the query will only return games. Which means you cannot do very complex queries since you can only search one table at a time.



The second command we have within our application system is a create section. This is where you can create new data for the database. You can create new players, managers, games, and stadiums. This is also pretty limited because you can only create new data for those 4 subsections. But, this is still a very simple to use and operate application. It is very straightforward because you just have to input the information into the textboxes, sliders, combobox, or datepicker that appear once on the desired input screen.

****

The third and last command we have within application is the modify section. This also has a pretty limited functionality because we wanted to focus on making sure everything worked as it is supposed to. There are also four sections within our database system that you can modify and they are players, managers, games, and stadiums. You can look below to see what the modified data view looks like.



**Report Queries**

1. List the all-time top 5 goal scorers of each national team that has played in at least 3 world cup finals.
2. List all players that have scored a hat-trick since 1994.
3. What is the average TV viewership of the games in which more than 6 goals were scored
4. List all the teams that were in the top 50% of goals scored, saves, and tackles in each World Cup tournament and list the tournament year ( redundancy allowed; if Germany met criteria in 1978, 1994, 1998 then list Germany 3 times ).

**Summary and Discussion**

Our project finished quite well compared to where we began. When we originally set out on this project we did not know exactly how much data we would need so we tried to plan on having a decent sized database that we could manage and learn from. This worked out very well for us because focusing just on the World Cups and not all of the top soccer competitions around the world made it much more manageable. The project has not changed very much from our original plan. We took our time when planning and talking about what we wanted to be able to keep track. There was a lot of back and forth on what we wanted to do and what we thought we could manage will still create a nice clean database. We did not really have to change much on our diagram besides some of the relationships and some of the attributes we were planning on using. For example we changed some of the attributes in our player table because when we were scrapping/collecting data we realized that some of the players did not have a first and last name. The players only had one singular name so we decided to cut having a first and last name and decided to just keep track of their full name. These were most of the corrections and things we had to fix on our diagram.

If we were to continue going forward working on this database We think the next step would be trying to collect more data and add in other soccer competitions such as the European Championships, Concacaf Gold Cup, and others. When looking at these ideas for the future we see a chance to utilize clusters. Using clusters would be perfect for our database. It would be great because it would make our queries faster once we have a bunch of extra information. The way we would most likely utilize clusters is by giving each tournament its own cluster index. Each cluster index would never change since we cannot change history. Each would be unique so that once the tournament happens again in a different year we could easily add that data to the cluster it is supposed to be in.