National Basketball Association (NBA) Mini-Database Final Report

Group Members: Peter Mah, Jason Paul, Thomas Yee

April 12, 2020

Table of Contents

Table of Contents	1
Abstract	2
Introduction	3
Project Design	3
Implementation	5
Enhanced Entity Relationship Schema Diagram (EERD)	5
Relational Model	6
Functional Model	6
DBMS Design	7
SQL Queries	8
User Manual	10
Setup instructions	10
UI Demonstrations	16
Demonstration of Inner Join via the Height Search Page	16
Demonstration of Left Outer Join via the Team Schedule Search Page	17
Demonstration of Union All via the Team Shooting Search Page	19
Demonstration of Select via the Teams page	21
Demonstration of Full Outer Join via the Games page	22
Demonstration of Intersect via the Home Dominators page	24
Demonstration of Minus via the Miami Heat page	25
Demonstration of Right Join & Natural Join via the Player Stats Page	26
Demonstration of Right Outer Join via the Player Information Page	27
References	28

Abstract

The main goal of this project was to create a mini-database that holds statistics (stats) about the National Basketball Association (NBA) sports league. The mini-database includes the stats for teams, players, and games. A user will be able to search the database for their favourite teams or players in order to see their stats. The process of creating the application is discussed, such as the type of data needed for the database, how the data was generated, and how the data was shown. In subsequent sections, the steps taken to implement the design are explained in detail.

First, an enhanced entity relationship conceptual schema diagram (EERD) was created to display the relationship between the entities (teams, players, etc) in the database. Next, a relational model was created to represent database attributes as a collection of relations. This model shows how the data is stored in the database and how the tables relate to each other. Next, the database management system (DBMS) design is examined to discuss the software package used to facilitate the creation and maintenance of the computerized database. Finally, a functional model was created to show the functions of our system. The frontend was created using PHP, HTML, and CSS, and the database with MySQL. The functions of the website are shown alongside the corresponding queries to showcase the different queries used in the system. A user manual is provided to give additional support on how to use the website.

Introduction

In 1949, the National Basketball Association was formed in the United States. ^[1]The NBA is a professional basketball league with teams in the United States and one team in Canada. The NBA has a regular season that typically occurs during the months of October to April where each team will play 82 games during this time. ^[2]Depending on how a team performed in the regular season, they may advance to the playoffs where additional games will be played in a tournament style. ^[3]There are currently 30 teams in the NBA with each team having a maximum of 15 players.

Millions of people watch the NBA daily with a large percentage of the audience interested in stats of teams and players. Every year, hundreds of games are played, generating various recorded statistics for each game, player, coach, etc. The decision to create a database involving the NBA was due to interest in the sport, the vast amount of data generated, and the public availability of the data online.

Project Design

This section details the process of creating the application. Due to the complexity and scope of this project, careful planning and organization of the various steps were required.

The first step involved identifying the type of information and data needed for the database. A user may be interested in extracting a variety of information about the NBA. In this early phase of the project, the processes of drafting the Enhanced Entity Relationship Diagram (EERD) and gathering data from various sources were conducted in parallel by members of the team. Subsequent iterations of the EERD then reflected the type of data that was ultimately available.

The data used in the database was ultimately gathered from various websites^{[4][5][6][7][8][9]}. One complication that arises from having data from several sources is that certain attributes common to different data sources have data that are inconsistent among each other. For example, player names from one data source varied slightly in syntax with player names from a different data source. Since attributes for a given player had to be consolidated from the different sources,

player names across data sources for the same player had to be cleaned and made to be unique between all the data. This required significant data processing efforts, in order to compile the data from all sources, filter out duplicates, and properly parse the data in preparation for insertion into the database.

During the import process, some data were not correctly parsed into the resulting csv files. This required going through and deleting certain cells of data. Replacing or fixing some of these discrepancies was not considered time effective due to the large number of statistics involved. Also, a number of special characters had to be replaced throughout all relations in order to prevent issues with performing queries on the data.

After data was obtained from all sources and with the EERD in hand, the DBMS design was considered. XAMPP software was used to establish a local server and PHPMyAdmin facilitated tasks relating to creating the first version of a functional MYSQL database, such as being able to use data in CSV files. Further details of the DBMS design are discussed in a later section of this report.

With the prototype database in hand, we began implementing the front and back end components to the application. The PHP scripting language was used in conjunction with HTML and CSS to provide both front and back end functionality. Since the entire application, including server, is run on a single computer, the front and back end components are both integrated and implemented through dynamic browser web pages. Since none of the team members were familiar with PHP coming into this project, [10] one team member took the lead in getting a fully functional, albeit minimal, prototype application built. This benefitted the remaining team members in that the existing PHP code was used to get up to speed and provide good component templates for the remaining features of the application.

After iteration of the database query functionality and the user interface of the application, the functional model was created. This model serves primarily to illustrate the functionality of the application, specifically the features that are available to the user.

Implementation

Enhanced Entity Relationship Schema Diagram (EERD)

Appendix A contains the EERD for our database. The EERD is a high-level conceptual data model which displays the relationship between the entities (e.g. teams, players, etc) in the database. It is generally used, as was the case with our project, as a tool for designing a database application.

EERD Assumptions

The following assumptions and clarifications pertain to our EERD:

- All teams play the same amount of games within a particular season
- Each game has a winner (i.e. the table only contains historical data on games, does not include data on future games)
- In the CoachStats relation, the certain attribute names are somewhat ambiguous upon first reading it. The following serves to provide clarification:
 - CurrentGamesCoached refers to the number of games a coach has acted in any coaching role within the coach's current tenure with the team
 - CurrentGamesWon refers to the number of games won by the coach during their current tenure with the team
 - CurrentGamesLost refers to the number of games lost by the coach during their current tenure with the team
- Each member in the league has a unique FirstName and LastName combination.
 However, the primary key for all relations for members, players, coaches, and their associated statistics, is a unique identifier called the MemberID.

Relational Model

A relational model was constructed and is located in Appendix B. The textbook [11] Fundamentals of Database Systems, 7th Ed., and course notes were used in helping prepare the model. In particular, the rules for mapping an EERD (Enhanced Entity Relationship Diagram) to a relational model proved very useful. Several iterations were required prior to achieving the final version. The preliminary model resulted in multiple tables in various normalized forms (i.e. tables were in 2NF (second normal form), 3NF (third normal form), and BCNF (Boyce-Codd normal form)). Subsequent iterations decomposed all tables into BCNF subtables and added a MemberID attribute as the primary key for several relations. This allowed removal of a composite primary key that was based on members' first and last names. On the diagram, arrows indicate the connection between relations (tables) where the tail of the arrow is the foreign key (FK) and the head is the primary key (PK). Underlined attributes are the primary key for the relationship. The final model consists of 11 relations.

Functional Model

The functional model design for the application is shown in Appendix C. It's intent is primarily to illustrate the functionality of the application, specifically the features that are available to the user.

Upon launching the application, the user should visit the home page titled *Home*. The only functionality of the home page is to display the navigation toolbar on the top of the display such that the user can choose what types of information to obtain. Each page in the navigation toolbar displays a unique information set.

DBMS Design

This section details the software package used to facilitate the creation and maintenance of the computerized database.

The NBA Mini-Database was created via the PHPMyAdmin web interface. Each table in the database was populated from predefined .csv data files.

The NBA Mini-Database system utilizes a number of software packages to interface with the database. Users interact with the database system via a web-based front end interface called *nba.php*. The web-based user interface obtains user input for page requests, search criteria, and filter criteria to determine what SQL queries to execute on the NBA Mini-Database. Apache hosts the local server in which it is able to interact with the MySQL API. SQL queries are executed by the MySQL API which interacts directly with the database. The NBA Mini-Database then returns results from the query, where the results are passed through the same tech stack, and finally displayed on the web-based front end for the user to view.

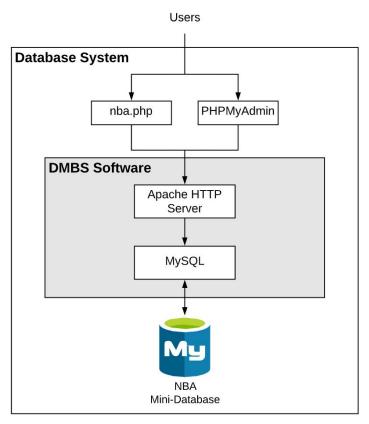


Figure 1: Simplified Database System Environment

SQL Queries

Our database is based on static historical data and, as such, is not continuously updated. Because of this, and because write access to the database has been restricted from the user, the user interface does not support any data manipulation functionalities (i.e. update, insert, delete).

It is our understanding that the queries we have implemented represent approximately 80% of the queries concepts that were covered in the lectures and meet/exceed the minimum requirements for the project. The following table *Table A* summarizes the query concepts implemented into the NBA Mini-Database system.

Table A: NBA Mini-Database Supported Functionalities

Function	Query Type	Query	
Height Search	INNER JOIN ON	SELECT m.FirstName, m.LastName, m.TeamName, p.Height FROM member AS m INNER JOIN player AS p ON (m.MemberID = p.MemberID) WHERE (p.height LIKE '%\$height%' AND m.TeamName LIKE '%\$team%')	
Team Schedule Search	LEFT OUTER JOIN	SELECT d.Date, g.VisitorTeamName, g.HomeTeamName FROM date AS d LEFT OUTER JOIN (SELECT * FROM game WHERE (game.VisitorTeamName LIKE '%\$teamName%' OR game.HomeTeamName LIKE '%\$teamName%')) AS g ON g.DateId = d.Id WHERE d.Date BETWEEN '\$startDate' AND '\$endDate' ORDER BY d.date	
Team Shooting Search	UNION ALL	SELECT TeamName, Season, `FG%`, `3P%` FROM team_stats WHERE `FG%` BETWEEN '\$fgmin' AND '\$fgmax' UNION ALL SELECT TeamName, Season, `FG%`, `3P%` FROM team_stats WHERE `3P%` BETWEEN '\$p3min' AND '\$p3max' ORDER BY TeamName, Season	
Team Stats	SELECT	SELECT * FROM team_stats	
1. Game stats	FULL OUTER JOIN (MySQL equivalent using LEFT JOIN and	SELECT d.Season, g.GameDate, g.GameStartET, g.VisitorTeamName, g.VisitorPoints, g.HomeTeamName, g.HomePoints, g.Attendance, g.ArenaName FROM `date` AS d LEFT JOIN game AS g ON (g.DateId = d.Id) WHERE (g.ArenaName LIKE '\$stadiumName%')\n	

	RIGHT JOIN)	SELECT d.Season, g.GameDate, g.GameStartET, g.VisitorTeamName, g.VisitorPoints, g.HomeTeamName, g.HomePoints, g.Attendance, g.ArenaName FROM `date` AS d RIGHT JOIN game AS g ON (g.DateId = d.Id) WHERE (g.ArenaName LIKE '\$stadiumName%' AND d.Id is null)	
2. Home Dominators	INTERSECT (MySQL equivalent using Inner Join)	SELECT tableA.GameDate, tableA.GameStartET, tableA.VisitorTeamName, tableA.VisitorPoints, tableA.HomeTeamName, tableA.HomePoints, tableA.Overtime, tableA.Attendance, tableA.ArenaName FROM game as tableA INNER JOIN (SELECT * FROM game WHERE HomePoints>=100) as tableB ON tableA.GameDate=tableB.GameDate AND tableA.HomeTeamName=tableB.HomeTeamName WHERE tableA.VisitorPoints<=80	
3. Miami Heat Information	MINUS (MySQL equivalent using Right Outer Join)	SELECT FirstName, LastName, TeamName, tableB.Position, tableB.JerseyNo, tableB.Height, tableB.Weight, tableB.BirthDate, tableB.College FROM member as tableA RIGHT OUTER JOIN (SELECT * FROM player WHERE NOT player.Position='G') as tableB ON tableA.MemberID=tableB.MemberID WHERE TeamName='Miami Heat	
Player Stats	RIGHT JOIN	SELECT m.FirstName, m.LastName, m.TeamName, p.Season, p.Salary, p.GP, p.GS, p.MIN, p.FG, p.`FG%`, p.3PT, p.`3P%`, p.FT, p.`FT%`, p.OR, p.DR, p.REB, p.AST, p.AST, p.BLK, p.STL, p.PF, p.TO, p.PPG FROM player_stats AS p RIGHT JOIN member AS m ON m.MemberID=p.MemberID	
Player Stats (drop-down-team.php)	NATURAL JOIN	SELECT * FROM member as tableA NATURAL JOIN (SELECT * FROM player_stats WHERE Season='2017-2018') AS tableB WHERE tableA.TeamName LIKE '\$teamName%'	
Player Information	RIGHT OUTER JOIN	SELECT m.LastName, m.FirstName, m.TeamName, p.Position, p.JerseyNo, p.Height, p.Weight, p.BirthDate, p.College FROM member AS m RIGHT JOIN player AS p ON (m.MemberID = p.MemberID) WHERE (m.LastName LIKE '\$letter%')	

Notes:

- 1. FULL OUTER JOIN is not a supported keyword in MySQL. To obtain equivalent results, the following queries were used to emulate the FULL OUTER JOIN.
 - Left join would get all the records from the left table regardless of whether or not they had a match in the right table.
 - Right join would get all the records from the right table regardless of whether or not they had a match in the left table.
 - Union to add the results of both of these queries.
 - Added a where clause to the right join query to get the unmatched in the left table.
- 2. INTERSECT is not a supported keyword in MySQL. To obtain equivalent results, an INNER JOIN was used.
- 3. MINUS is not a supported keyword in MySQL. To obtain equivalent results, a RIGHT OUTER JOIN was used.

User Manual

This section provides a comprehensive guide into the installation of the software required to run the application and subsequent use of the application. The application functionality described below is organized according to the individual query options available to the user.

Setup instructions

1. Installing XAMPP, Apache HTTP Server, and MySQL

Note: if you have an existing version of MySQL already installed on your computer, uninstall the existing version before continuing.

XAMPP is an application containing software modules required to host the database and interact with it. The XAMPP application installer can be downloaded at the following link: https://www.apachefriends.org/download.html

Upon installation of the latest version of XAMPP, ensure XAMPP installs Apache HTTP Server and MySQL.

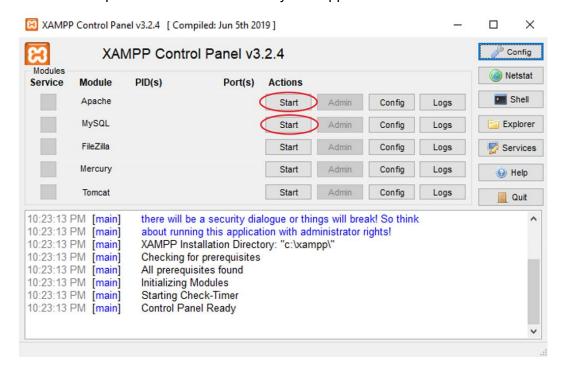
2. Download relevant NBA Mini-Database files

Download the *nbadbfiles* folder containing all files relevant to the NBA Mini-Database. Save the *nbadbfiles* folder into the following folder inside the xampp root folder: C:\xampp\htdocs

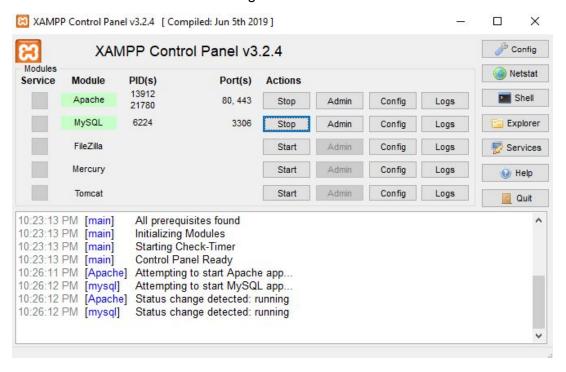
Note: nbadbfiles also contains the database file called nba.sql

3. Launch Apache HTTP Server and MySQL

Launch the Apache HTTP Server and MySQL applications via the XAMPP Control Panel



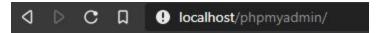
Successful launch of the Apache and MySQL modules should result in an XAMPP Control Panel similar to the following



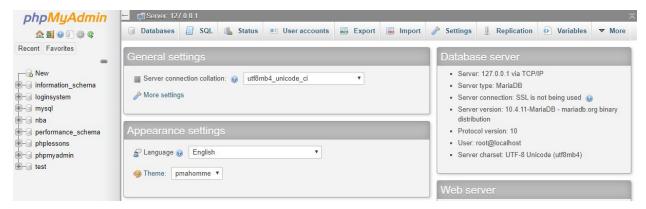
4. Importing the NBA Mini-Database

In this step we will be creating a database on your local server, and importing the actual NBA Mini-Database data into your local server.

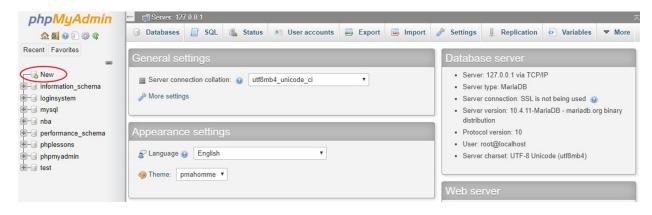
- 4a. Open a web browser
- 4b. In the address bar, go to the following url: localhost/phpmyadmin



4c. If Apache and MySQL are running properly, the url in step 4b should bring you to a page similar to the following



4d. Create a new empty database called nba





- Click the *nba* database from the list of databases on the left hand side panel.

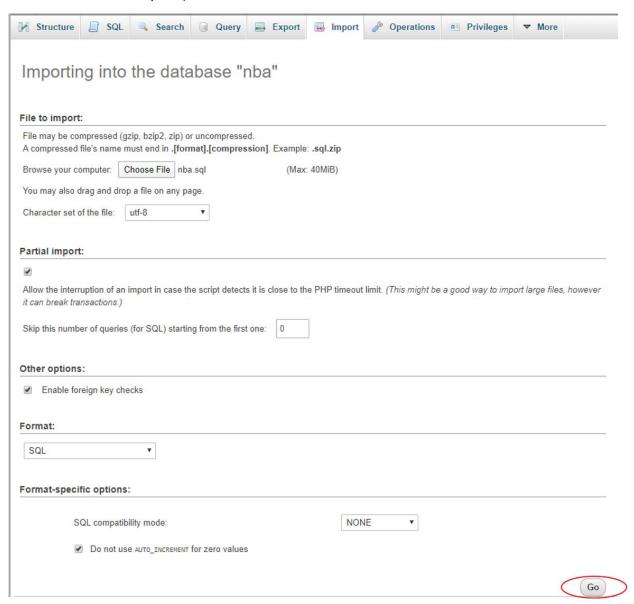


- Choose the *Import* tab in the top panel



Press the Choose File button and navigate to the nba.sql database file located in
 C:\xampp\htdocs\nbadbfiles

- In the *Import* tab in MyPHPAdmin, press the *Go* button in the bottom right corner to import the database onto your local server (the other options selected should be adequate)

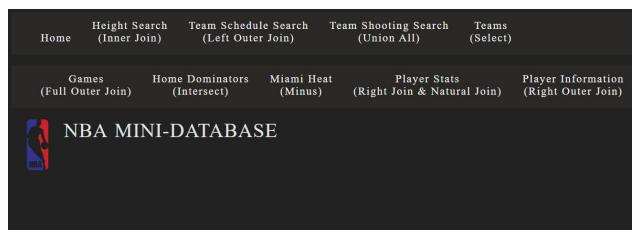


- The NBA Mini-Database has been successfully imported to your local server

5. Access the NBA Mini-Database application's home page

In your web browser, access the following url in the address bar: localhost/nbadbfiles/nba.php

If all steps have setup steps have been completed successfully, the web browser should display the homepage of the NBA Mini-Database application



NBA official logo image sourced from^[12]

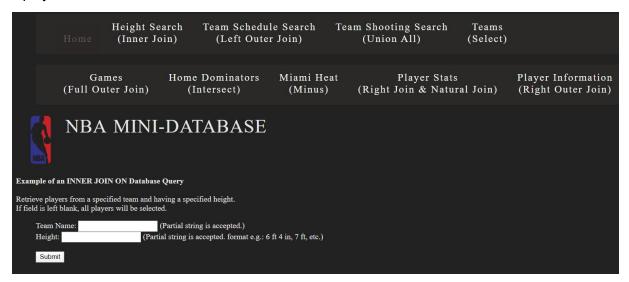
Comments:

All pages in the NBA Mini-Database application contain a navigation toolbar pinned to
the top of the webpage. There are a total of 10 page accessing buttons in the toolbar.
Each button is associated with a hyperlink that will navigate the application to a new
page where user input is obtained, and where the results of SQL queries on the NBA
Mini-Database are displayed.

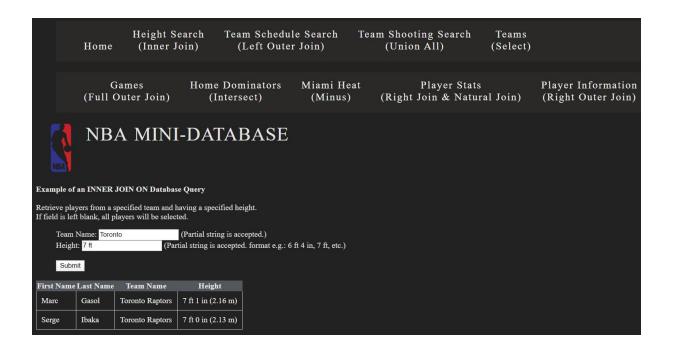
UI Demonstrations

Demonstration of Inner Join via the Height Search Page

After clicking Height Search (Inner Join) in the navigation bar, the following page will be displayed.

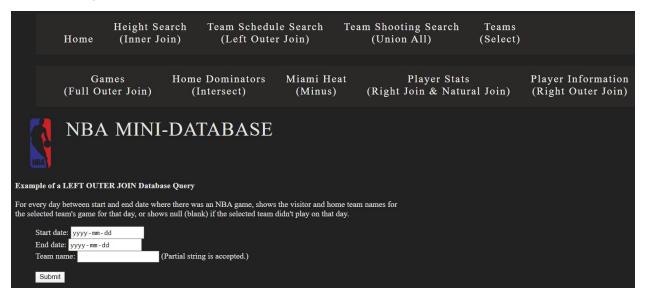


The purpose of this page is to allow the user to specifically search for players within a team that have desired height characteristics. The team name can be entered as a partial string to simplify or shortcut the search, or can be left empty if the user desires to search all teams for players with the desired height characteristics. Similarly, player height can be entered as a partial string as a way to expand a search to include more results. For example, entering 7 ft into the Height cell will search for players with heights of between 7 ft 1 in and 7 ft 11 in, whereas entering 6 ft 5 in will search for players with that specific height. Also, the Height field can be left blank if the user desires to show all players and their heights. The following shows an example query.



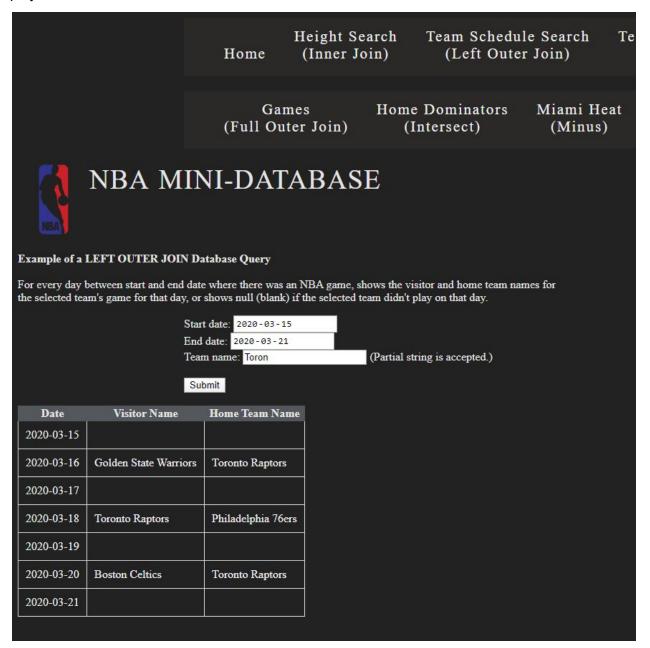
Demonstration of Left Outer Join via the Team Schedule Search Page

After clicking Team Schedule Search (Left Outer Join) in the navigation bar, the following page will be displayed.



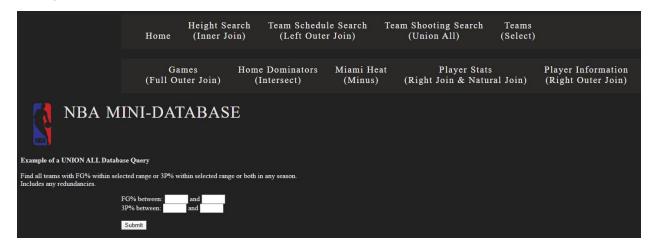
The purpose of this page is to allow the user to search for a team's game schedule for the range of selected dates. It will show both the selected team and the opposing team in their respective columns as home team or visiting team. Start and end dates are entered with the assistance of a date picker. Team name can be entered as a partial string as a means to shortcut or to simplify the search. All dates within the range will be shown, and any days where the team

didn't play will be displayed as blanks. The following example query shows Totonto games played in the third week of March 2020.



Demonstration of Union All via the Team Shooting Search Page

After clicking Team Shooting Search (Union All) in the navigation bar, the following page will be displayed.



The purpose of this page is to allow the user to search for all teams (and associated seasons) where their field goal percentage (FG%) and three point percentage (3P%) are within specific ranges of values. These statistics show a team's annual average throwing attempt accuracy. The following example query shows the teams/seasons where the FG% is higher than 48% or their 3P% is higher than 38%. Also, this is a "union all" query, which means that if a team got greater than the specified FG% *and* greater than the specified 3P%, that row result would show up twice. In the example below, in the 2017-2018 season the Golden State Warriors had FG% *and* 3P% within the selected ranges, so it shows twice in the guery result.

Height Search Home (Inner Join) Team Schedule Search
(Left Outer Join)

Games (Full Outer Join) Home Dominators (Intersect)

Miami Heat (Minus)

Te



NBA MINI-DATABASE

Example of a UNION ALL Database Query

Find all teams with FG% within selected range or 3P% within selected range or both in any season. Includes any redundancies.

FG% between: 48 and 70 3P% between: 39 and 70

Submit

Team Name	Season	FG%	3P%
Golden State Warriors	2017-2018	50.3	39.1
Golden State Warriors	2017-2018	50.3	39.1
Golden State Warriors	2018-2019	49.1	38.5
Los Angeles Lakers	2019-2020	48.5	35.5
New Orleans Pelicans	2017-2018	48.3	36.2
San Antonio Spurs	2018-2019	47.8	39.2

Demonstration of Select via the Teams page

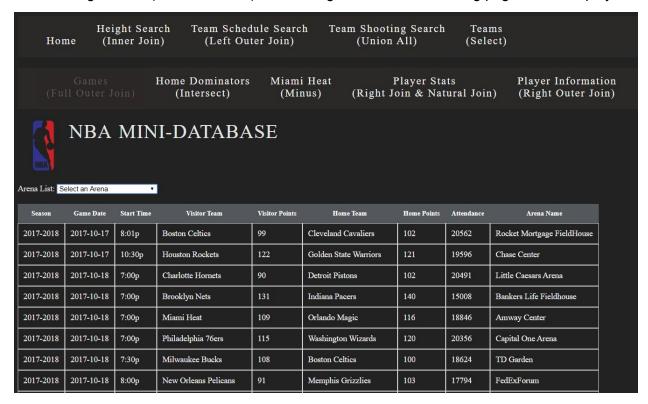
After clicking Teams (Select) in the navigation bar, the following page will be displayed.



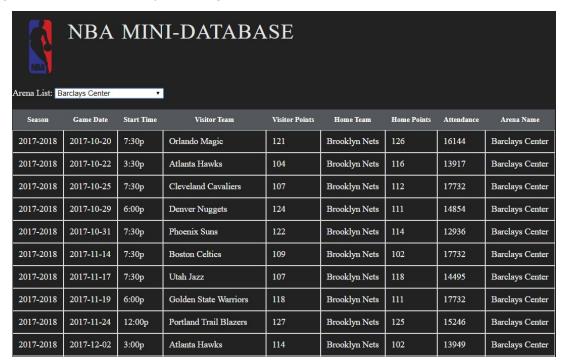
This page is a demonstration of the utilization of a simple Select all query on the team_stats table in the database.

<u>Demonstration of Full Outer Join via the Games page</u>

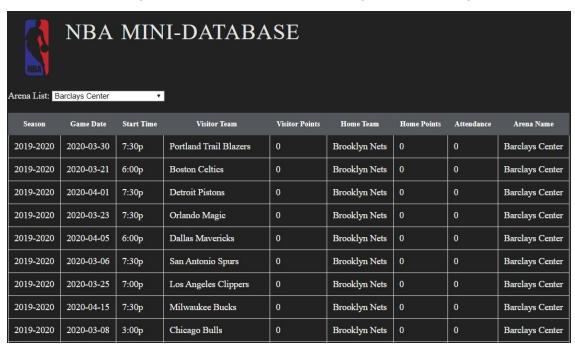
After clicking Games (Full Outer Join) in the navigation bar, the following page will be displayed.



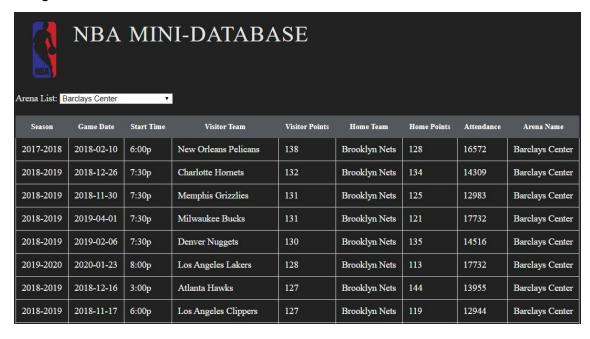
This page displays all of the games that have already ended and games that have yet to start. The games can be filtered by selecting an arena from the Arena List as shown below:



The games can also be sorted by pressing on a column header. For example, if you were to select Visitor Points, the games will be shown in descending order according to Visitor Points.



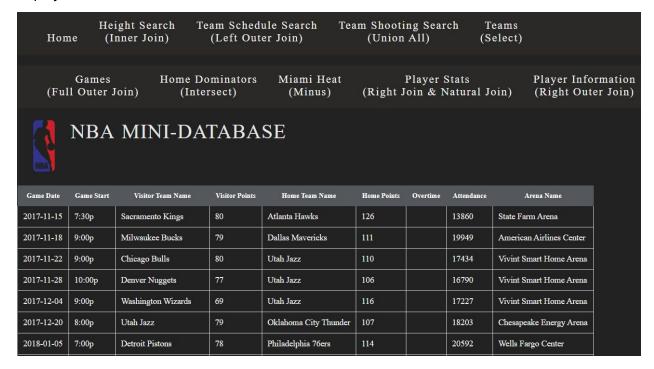
If the Visitor Points header is selected again, the games will be shown in ascending order according to Visitor Points.



Note: the Full Outer Join keyword is not supported in MySQL. To obtain the same results as a Full Outer Join query available in other database management systems, a left join, union, and right join must be used.

Demonstration of Intersect via the Home Dominators page

After clicking Home Dominators (Intersect) in the navigation bar, the following page will be displayed.

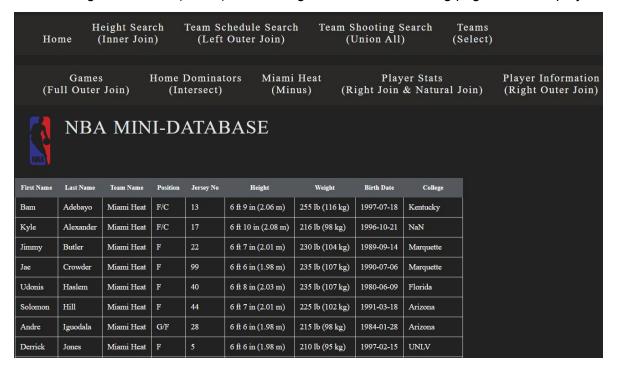


This page displays the games where the home team scored more than 100 points and the visiting team scored less than 80 points.

Note: the Intersect keyword is not supported in MySQL. To obtain the same results as an Intersect query available in other database management systems, an Inner Join must be used.

Demonstration of Minus via the Miami Heat page

After clicking Miami Heat (Minus) in the navigation bar, the following page will be displayed.

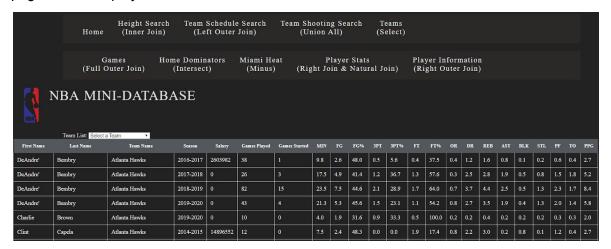


This page displays the information for the players who are not solely a guard (Position = G) on the Miami Heat team.

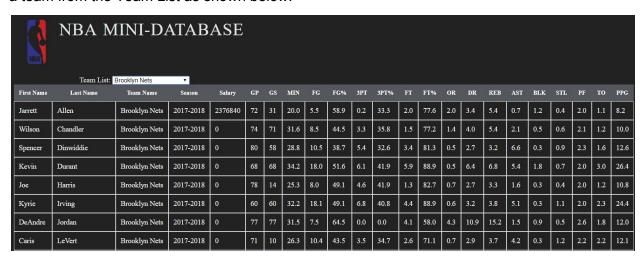
Note: the Minus keyword is not supported in MySQL. To obtain the same results as a Minus query available in other database management systems, an Right Outer Join must be used.

Demonstration of Right Join & Natural Join via the Player Stats Page

After clicking the Player Stats (Right Join & Natural Join) in the navigation bar, the following page will be displayed.

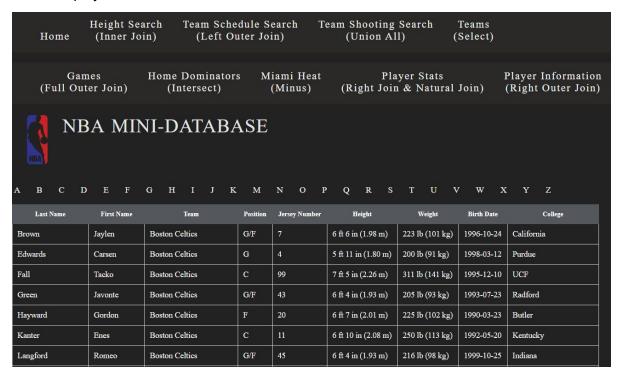


This page displays the stats of all players in the league. The players can be filtered by selecting a team from the Team List as shown below.

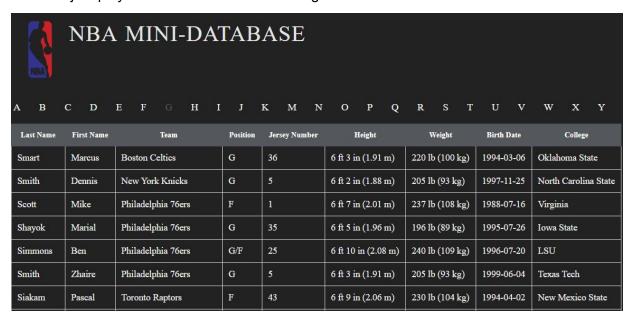


Demonstration of Right Outer Join via the Player Information Page

After clicking the Player Information (Right Outer Join) in the navigation bar, the following page will be displayed.



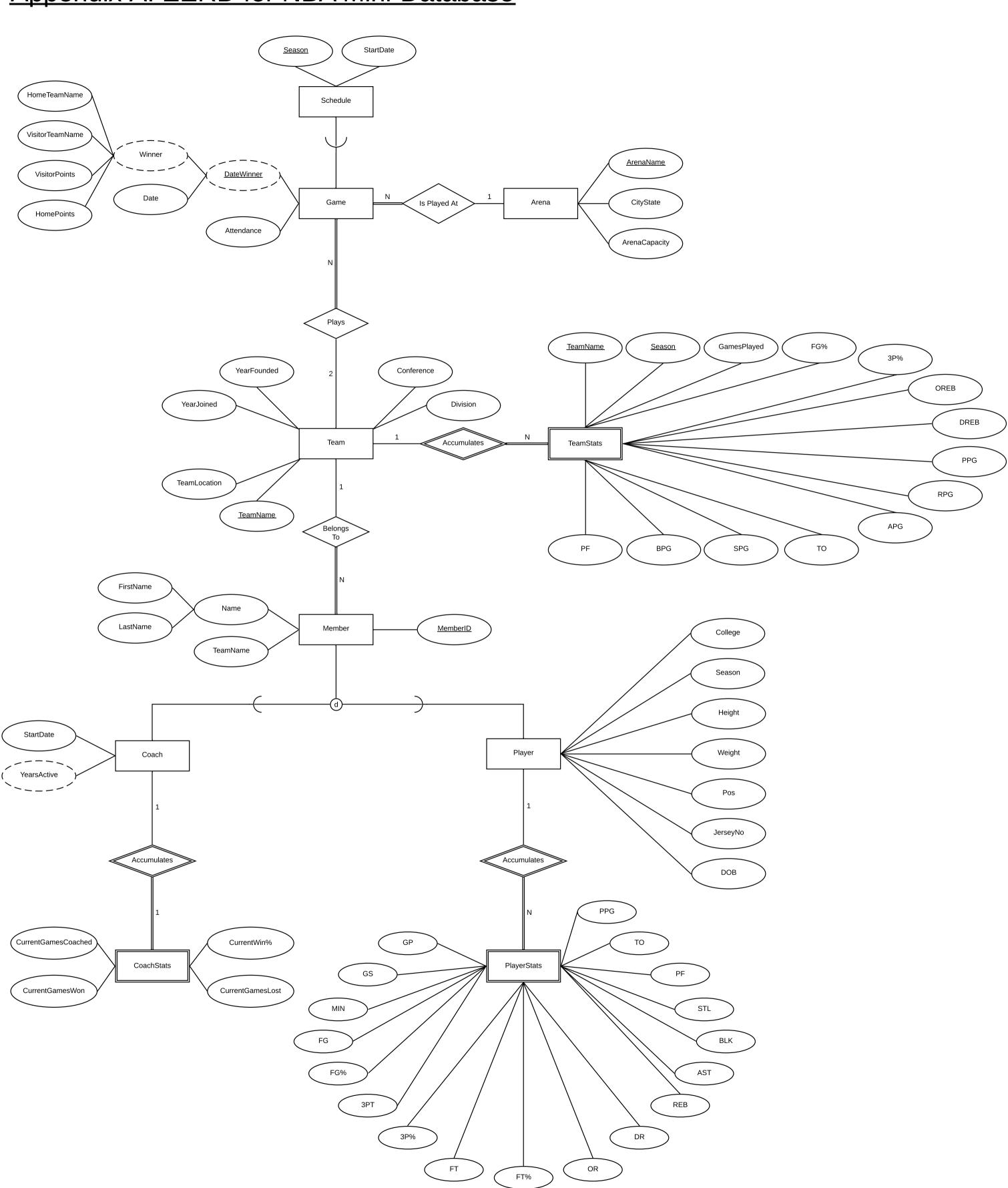
This page displays the personal information of all players in the league. The table can be filtered by selecting a letter located above the table to show all the players having the last name that starts with the letter chosen by the user. For example, if the letter S is chosen, the following table with just players with the last name starting with S will be shown.



References

- [1] The Editors of Encyclopaedia Britannica, "National Basketball Association," *Encyclopædia Britannica*, 01-Nov-2019. [Online]. Available: https://www.britannica.com/topic/National-Basketball-Association. [Accessed: 12-Apr-2020].
- [2] "National Basketball Association," *New World Encyclopedia*. [Online]. Available: https://www.newworldencyclopedia.org/entry/National_Basketball_Association. [Accessed: 12-Apr-2020].
- [3] NBA, "NBA Frequently Asked Questions," *NBA*, 2019. [Online]. Available: https://www.nba.com/news/faq. [Accessed 11 April 2020].
- [4] Wikipedia, "National Basketball Association," [Online]. Available: https://en.wikipedia.org/wiki/National Basketball Association.
- [5] ESPN, "NBA Stat Leaders," *ESPN*, 2020. [Online]. Available: https://www.espn.com/nba/stats. [Accessed 6 March 2020].
- [6] Sports Reference LLC, "Basketball Reference," *Sports Reference LLC*, 2020. [Online]. Available: https://www.basketball-reference.com/. [Accessed 6 March 2020].
- [7] Basketball Reference, "NBA Games," [Online]. Available: https://www.basketball-reference.com/leagues/NBA 2020 games-february.html.
- [8] Wikipedia, "List of National Basketball Association head coaches," [Online]. Available: https://en.wikipedia.org/wiki/List of National Basketball Association head coaches.
- [9] Wikipedia, "List of current NBA team rosters," [Online]. Available: https://en.wikipedia.org/wiki/List_of_current_NBA_team_rosters.
- [10] D. Nielsen, "How to Create a MySQL Database for PHP | PHP Tutorial | Learn PHP Programming | PHP Lesson," mmtuts, Copenhagen, 2015.
- [11] S. B. N. Ramez Elmasri, Fundamentals of Database Systems, 7th Ed., Pearson, 2016.
- [12] "NBA Logo PNG Transparent & SVG Vector," *Freebie Supply*. [Online]. Available: https://freebiesupply.com/logos/nba-logo/. [Accessed: 30-Mar-2020].

Appendix A: EERD for NBA Mini-Database



Appendix B: Relational Model - NBA Mini-Database

Peter Mah, Thomas Yee, Jason Paul

