

NBA DATABASE

CSS 475 Winter 2019 – Team Avatar

Jeremy Tandjung, Nathan Phan, Aaron Handjojo, Brandon Posey

Contents

1. INTRODUCTION	2
2. Application	2
3. Entities	2
4. Design	4
4.1. Assumptions	4
4.2. Entity Relationship (ER) Diagram	4
Figure 1: NBA ER Diagram	6
4.3. Relational Model	7
Figure 2: NBA RM Diagram	8
4.4. Constraints	8
5. Queries	13
6. Data Generation	14
7. Schedule	14
7.1. Team Formation	14
7.2. Diagram	14
7.3. Tooling Section	14
7.4. Develop Database	15
7.5. Finish Database and Presentation	15
8. Work Distribution	15
9. Tooling Assessment	15
9.1. DBMS: MySQL Workbench CE	16

9.2. UI: HTML/CSS	16
9.3. Version Control: GitHub	16
9.4. Web Server: AWS or USBWebServer	16

1. INTRODUCTION

Team Avatar intends to create a basketball database based on the properties of the National Basketball Association. The database will track team and player information, statistics, awards, etc within the current NBA season, which is the 2018-19 NBA season.

Team Avatar will implement an update meeting cycle, where every week we meet up to give each other an update on what we have done in the past week.

2. APPLICATION

The application is primarily used by NBA teams to determine whether a player is playing well or not. The database could also be used by fantasy league users to see if their current fantasy roster is playing well or not. NBA fans could also look up trivial information of their favorite team and players, such as birthdate, age, height, and many more.

3. ENTITIES

The following are the entity data types that are in the database

- TEAMS: The 30 Teams that are in the NBA as of the current NBA season; 2018-19 NBA Season
 - TeamID
 - Name
 - CoachID
 - ArenaID
 - Conference
 - Division
 - Championships

- Wins
 - Losses
- ARENAS: The current and past arenas used by the 30 NBA teams.
 - ArenaID
 - Name
 - Capacity
 - Address
 - City
 - State
- HEAD_COACHES: The current and past head coaches of the 30 NBA teams
 - CoachID
 - Name
 - Wins
 - Losses
 - Dob
- PLAYER: The current roster of the 30 NBA teams
 - PlayerID
 - Name
 - TeamID
 - FName (First Name)
 - LName (Last Name)
 - DOB (Date of Birth)
 - Height
 - Weight
 - DraftYear (The Year the player got drafted)
- STATISTICS: The player's per game statistics
 - PlayerID
 - StatsID
 - PPG (Points Per Game)
 - APG (Assist Per Game)
 - RPG (Rebound Per Game)

- SPG (Steals Per Game)
- BPG (Blocks Per Game)
- FG% (Field Goal Percentage)
- 3FG% (3-Point Field Goal Percentage)
- AWARDS: The awards given to the players for their performance in the regular season
 - AwardID
 - Name
 - Description

4. DESIGN

In this section, we will discuss the assumption, design decision, and both ER and RM diagrams that our group have chosen to work it

4.1. ASSUMPTIONS

1. This database only contains data from the current NBA season during the making of this project, which is the 2018-19 NBA season.
2. All player statistics and information are up to date right before the NBA All-Star Weekend which is Feb 16th, 2019.
3. This database will include all current and past arenas that are used by the 30 NBA teams.
4. This database will include past and current head coaches that has coached an NBA team. That means, coaches who are currently not coaching an NBA team are included in the database.

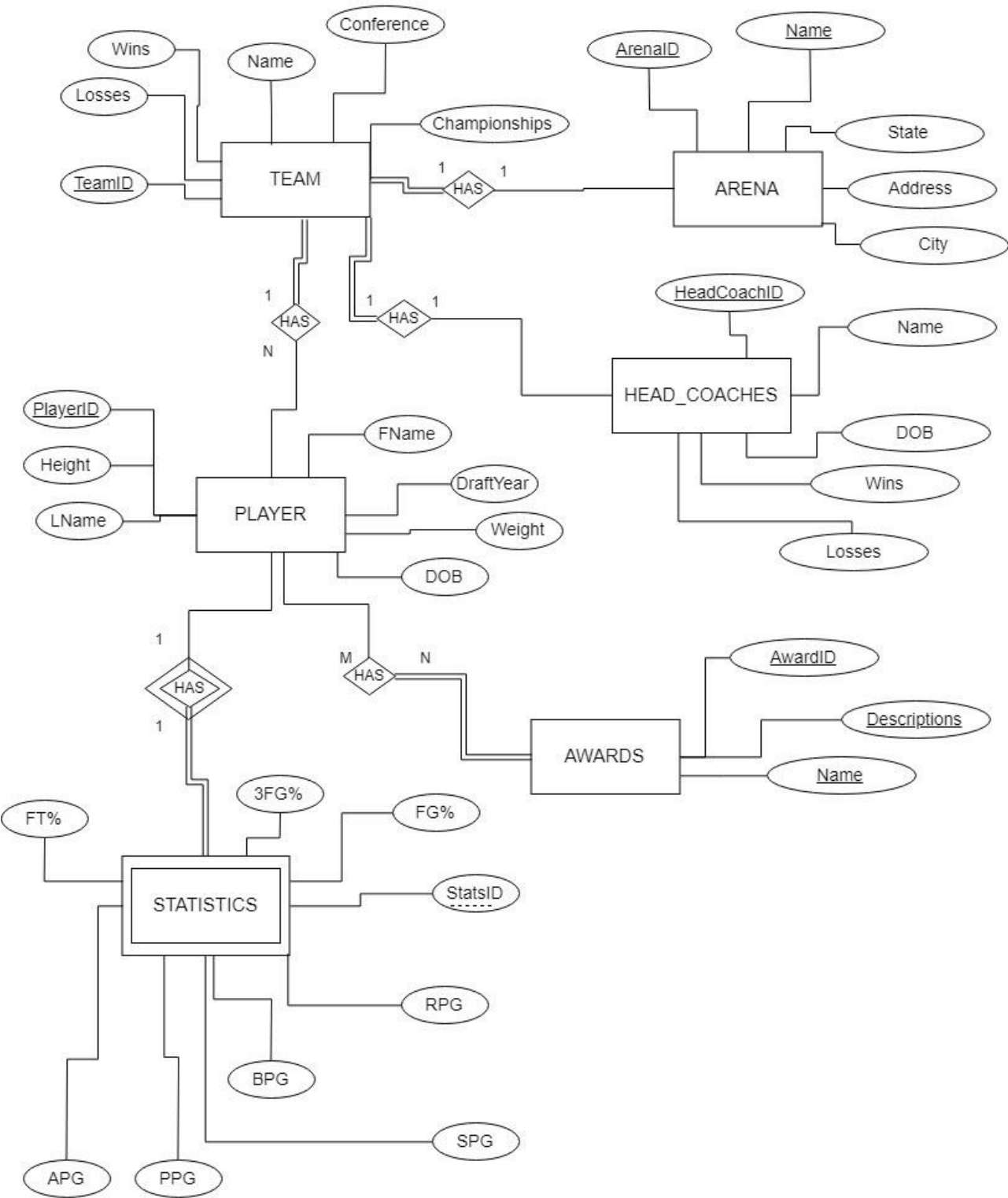
4.2. ENTITIY RELATIONSHIP (ER) DIAGRAM

The Entity Relationship Diagram provides an abstract overview of the various relationships between the entities derived from the requirements of the NBA Database. The strong entities include the following entities: TEAM, ARENA, PLAYER, HEADCOACHES, and AWARDS. The only weak entity is STATISTICS, which has an identifying relationship with the entity PLAYER. The attributes that correspond with each entity are clearly connected. In addition,

attributes that serve as keys are underlined to indicate their status as candidate keys, or partially underlined to indicate their status as partial keys. Each relationship between entities has its cardinality and participation status clearly labeled. Refer to Figure 1 for further details regarding the diagram.

Team Avatar chose to create “HAS” relationships between TEAM and the entities of ARENA, HEAD_COACHES, and PLAYER. The reasoning for this decision is that TEAMS in the NBA have direct relationships with those entities wherein the latter entities are identified as belonging to the TEAM. Since an NBA team needs to have a relationship with those entities, it was decided to impose total participation on the TEAM side of those relationships. In addition, the cardinality ratio of the relationships are 1:1 in the case of TEAM and ARENA, and TEAM, and HEAD_COACHES

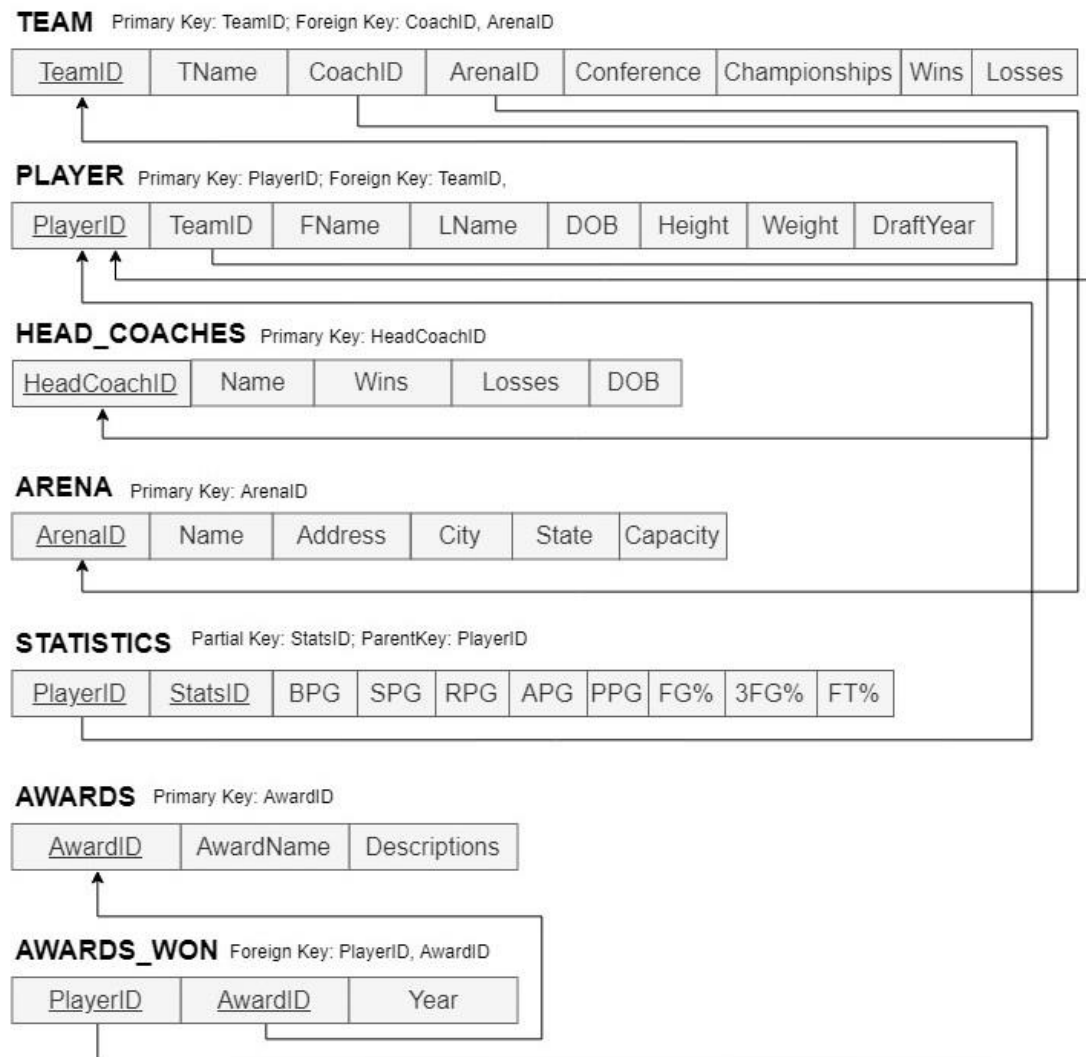
Figure 1: NBA ER Diagram



4.3. RELATIONAL MODEL

The Relational Model provides an abstract overview of the different relations derived from the requirements. Each relation will include its entity's attributes including keys such as primary keys, foreign keys, and partial keys. The keys that are referenced from other relations are modeled using arrows to demonstrate the origin of the foreign key. In this case, the relation AWARDS_WON represents the relationship between PLAYER and AWARDS. The relation itself will have no distinctive key as the relationship between PLAYER and AWARDS with the foreign keys of PlayerID, AwardID, and the relation's attribute of Year providing meaning and uniqueness for the AWARDS_WON relation. Refer to Figure 2 for further details regarding the Relational Model schema.

Figure 2: NBA RM Diagram



4.4. CONSTRAINTS

- TEAM
 - TeamID:
 - PRIMARY KEY
 - Not NULL

- Domain: All positive integers
- TName:
 - Not NULL
 - Domain: VarChar(40)
- Conference:
 - Not NULL
 - Domain: VarChar(40)
- Championship:
 - Not NULL
 - Domain: All non-negative integers
- Wins:
 - Not NULL
 - Domain: All non-negative integers
- Loses:
 - Not NULL
 - Domain: All non-negative integers
- CoachID:
 - FOREIGN KEY from HEAD_COACHES
 - Not NULL
 - Domain: All non-negative integers
- ArenaID:
 - FOREIGN KEY from ARENA
 - Not NULL
 - Domain: All non-negative integers
- PLAYER
 - PlayerID:
 - PRIMARY KEY
 - Not NULL
 - Domain: All positive integers
 - FName:
 - Not NULL

- Domain: All non-numerical VarChar(40)
- LName:
 - Not NULL
 - Domain: All non-numerical VarChar(40)
- DOB:
 - Not NULL
 - Domain: integers in the format mm/dd/yy where: $1 \leq mm \leq 12$, $1 \leq dd \leq 31$ (for months with 31 days) or $1 \leq dd \leq 30$ (for months with 30 days) or $1 \leq dd \leq 28$ (for February in a non-leap year) or $1 \leq dd \leq 29$ (for February in a leap year), and yy being the last two digits of years between 1900 and 2099.
- Height:
 - Not NULL
 - Domain: All non-negative integers (in cm)
- Weight:
 - Not NULL
 - Domain: All non-negative integers (in kg)
- DraftYear:
 - Not NULL
 - Domain: 4-digit positive integer between 1900 and 2099
- TeamID:
 - FOREIGN KEY from TEAM
 - Not NULL
 - Domain: All non-negative integers
- HEAD_COACHES
 - HeadCoachID:
 - PRIMARY KEY
 - Not NULL
 - Domain: All positive integers
 - Name:
 - Not NULL
 - Domain: All non-numerical VarChar(40)
 - Wins:

- Not NULL
 - Domain: All non-negative integers
- Loses:
 - Not NULL
 - Domain: All non-negative integers
- DOB:
 - Not NULL
 - Domain: integers in the format mm/dd/yy where: $1 \leq \text{mm} \leq 12$, $1 \leq \text{dd} \leq 31$ (for months with 31 days) or $1 \leq \text{dd} \leq 30$ (for months with 30 days) or $1 \leq \text{dd} \leq 28$ (for February in a non-leap year) or $1 \leq \text{dd} \leq 29$ (for February in a leap year), and yy being the last two digits of years between 1900 and 2099.
- ARENA
 - ArenaID:
 - PRIMARY KEY
 - Not NULL
 - Domain: All positive integers
 - Name:
 - Not NULL
 - Domain: All non-numerical VarChar(40)
 - Address:
 - Not NULL
 - Domain: VarChar(40)
 - City:
 - Not NULL
 - Domain: VarChar(40) of a name of a valid city in the US within the State.
 - State:
 - Not NULL
 - Domain: Char(2) of a name of a US State in abbreviation
 - Capacity:
 - Not NULL

- Domain: All non-negative integers
- STATISTICS
 - StatsID:
 - PARTIAL KEY
 - Not NULL
 - Domain: All positive integers
 - PlayerID:
 - PARENT KEY from PLAYER
 - Not NULL
 - Domain: All positive integers
 - BPG:
 - Not NULL
 - Domain: All non-negative double
 - SPG:
 - Not NULL
 - Domain: All non-negative double
 - RPG:
 - Not NULL
 - Domain: All non-negative double
 - APG:
 - Not NULL
 - Domain: All non-negative double
 - PPG:
 - Not NULL
 - Domain: All non-negative double
 - FG%:
 - Not NULL
 - Domain: All non-negative double
 - 3FG%:
 - Not NULL
 - Domain: All non-negative double

- FT%:
 - Not NULL
 - Domain: All non-negative double
- AWARDS
 - AwardID:
 - PRIMARY KEY
 - Not NULL
 - Domain: All positive integers
 - AwardName:
 - Not NULL
 - Domain: VarChar(40)
 - Description:
 - Not NULL
 - Domain: VarChar(400)
- AWARDS_WON
 - AwardID:
 - PARENT KEY from AWARDS
 - Not NULL
 - Domain: All positive integers
 - PlayerID:
 - PARENT KEY from PLAYER
 - Not NULL
 - Domain: All positive integers
 - Year:
 - Not NULL
 - Domain: 4-digit positive integer between 1900 and 2099

5. QUERIES

The NBA database will allow any interested party, namely NBA fantasy players, to query information about players and teams to see trends and make predictions about team, and individual performance for the 2018-2019 season.

For example, an NBA fantasy player may query the database to search for every player who currently plays for the Boston Celtics. The Fantasy player may then search for players who are not listed as out or suspended for the season.

The Fantasy player can then search the database with the existing filters to determine which of the players have a higher fantasy point average. The fantasy player can also narrow down the results by querying the database to search for players who have a higher than average projected points for the year.

6. DATA GENERATION

Team Avatar will use public information provided by basketball-reference.com to populate the database with NBA teams and players. The statistics that are available for each team and player will populate the team and player records. Team Avatar will use basketball-reference.com as their primary source of data.

7. SCHEDULE

Milestones for this project largely mirror the project iteration for each of the following assignments.

7.1. TEAM FORMATION

1. Set team by laws
2. Brainstorm project ideas
3. Check feasibility
4. Decide project idea

7.2. DIAGRAM

1. Create Entity Relationship Diagram
2. Create Relational Model Schema

7.3. TOOLING SECTION

1. Choose DBMS
2. Choose front-end framework tool

3. Choose web-hosting service
4. Choose data source

7.4. DEVELOP DATABASE

1. Set up web server
2. Write up database generation code
3. Populate database with data

7.5. FINISH DATABASE AND PRESENTATION

1. Develop website to access database
2. Make poster for presentation
3. Present final project to class

8. WORK DISTRIBUTION

With regards to selecting a database idea, our group discussed what level of interest we had in several different ideas to provide input on entity types and selection. During the brainstorming phase, we all provided ideas and input regarding what entities should exist inside an NBA database. We discussed at length regarding what relationships and references should exist between the entities that we identified. When it comes to implementing the database, Jeremy used his experience with MySQL to write code for the database and developed the ER Diagram and RM Schema on draw.io. When it comes to creating the documents for the project, Aaron was able to compile team ideas and record them with proper formatting. With regards to ensuring group collaboration, Brandon provided the framework for pushing group documents to Github and helped practice version control. Nathan worked with the group on setting up the web hosting service for the project. All members of Team Avatar will contribute equally to the voter database application. It's anticipated that each team member will produce 25% of the work overall; however, discrepancies within individual deliverables are expected.

9. TOOLING ASSESSMENT

In this section, we will outline the tools we chose to use for this project and its reasons behind the decision.

9.1. DBMS: MYSQL WORKBENCH CE

Team Avatar has chosen this DBMS because of its ease of use, popularity, and the fact that one of our members has experience with MySQL Workbench CE.

9.2. UI: HTML/CSS

For the front-end side, we have chosen a simple html page with buttons and multiple pages for each table would be sufficient enough for this project as the main focus is the database, not the UX.

9.3. VERSION CONTROL: GITHUB

GitHub is a very well-known platform for collaborators to share their code and it's free to use. Besides that, most of our members have prior experience in using GitHub as a version control platform.

9.4. WEB SERVER: AWS OR USBWEBSERVER

For now, our team is still exploring the possibilities of two web serving services; Amazon Web Service and USBWebServer. Both are free, and our final decision will depend on ease of use.