

National Basketball Association (NBA) Mini-Database Progress Report 2

Group Members: Peter Mah, Jason Paul, Thomas Yee

Project Update

A relational model was constructed and is located in Appendix B. The textbook *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. 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)). Following a discussion and strategy group meeting, the latest iteration decomposed those tables in 2NF or 3NF into subtables. The latest iteration shown in Appendix B of this report is completely in BCNF. 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 relation; several relations have multi-attribute PKs. There are 11 relations in total.

Completion of the Functional Model may require further changes to the relational model. Any additional changes will be reported in the next update.

The project schedule is provided in Appendix A. Raw data was obtained from various sources. As such, and due to the large number of tuples in some of the relations, data cleaning took longer than anticipated

Examples of data that required cleaning include:

- Names of players in the different sources of data were slightly different. Since join operations require that attributes are identical, these differences had to be fixed.
- Misplaced cell data. 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.)
- Special characters. A number of special characters had to be replaced throughout all relations in order to prevent issues with performing queries on the data.

We are currently wrapping up all work on the database back-end, while investing significant resources on the browser-based front-end of the application. We plan on implementing this database as a web app using PHP. Our next target is to have the database and back-end components initial builds complete by April 4.

Appendix A: Project Schedule

Task	Deadline	Responsible	Completion Status
Scrape data from web	March 13, 2020	Peter, Jason, Thomas	Complete
Entity Relationship Diagram	March 15, 2020	Peter, Jason, Thomas	Complete
Clean raw data	March 18, 2020	Thomas, Jason	Complete
Create functional database	April 4, 2020	Jason, Thomas	In progress
Backend completed	April 4, 2020	Peter	In progress
Relational Model	March 29, 2020	Thomas, Jason	Complete
Frontend completed	April 10, 2020	Peter, Jason, Thomas	In progress
Functional Model	April 12, 2020	Peter, Jason	Pending
Final Report	April 15, 2020	Peter, Jason, Thomas	Pending

Blue font indicates changes made since the previous project update provided on Mar 15, 2020.

Appendix B: Relational Model - NBA Mini-Database

Peter Mah, Thomas Yee, Jason Paul

