Project Report: Group 59

Comp 3380

Adam Pazdor

December 20, 2021

Juan A., Nelson Y., Saksham B., Ryan Loeppky

**Summary of the data**

Our data was chosen from a list data that our group compiled from an open-source website. We made sure that the data that we had selected for the project had a Creative Commons license. We choose this set of data from our list as it had very large table of data, and was well connected throughout the data and that could make for some interesting queries. This set of data contains information from the NBA of over 60000 games, 30 teams as well as 4500 players with attributes included for all three. We didn’t take every table available from the data set of the information, but the original file size was 811.36 MB of raw data dating back to the first NBA season in 1946-1947. We selected the columns that we agreed was essential and that contained valid data. The columns that we decided to set aside contained a lot of NULL values, therefore making the database unnecessary large.

**Discussion of the data model**

Juan had drawn up a EER diagram over our reading week and after he posted it to our group server in Discord. We all took a look at this diagram and tried to make sure that it was broken down into the correct groups and made sure that the cardinalities for the diagram were chosen in a way that would make sense. The data model was broken down into tables based on the diagram that we agreed as final. There had been a couple of cardinalities that had been up to debate on if they should have been 1 to many or many to many, but the original EER model was done up well and there was little that needed to be tweaked. Overall, the final tables that we had made sense and worked well for the diagram and could be used with interesting queries.

**Summary of the database**

(Insert summary here)

List each of the final tables, along with its cardinality and arity

List of the queries implemented in Part 2 (and thus available through the interface in Part 3)

Students do not need to write out the SQL code, just a text explanation of wht the query returns

**Summary of the interface**

(Insert summary here)

This should include both a description of how it was created (what language was it written in, including any specific libraries) and instructions on how to use it.