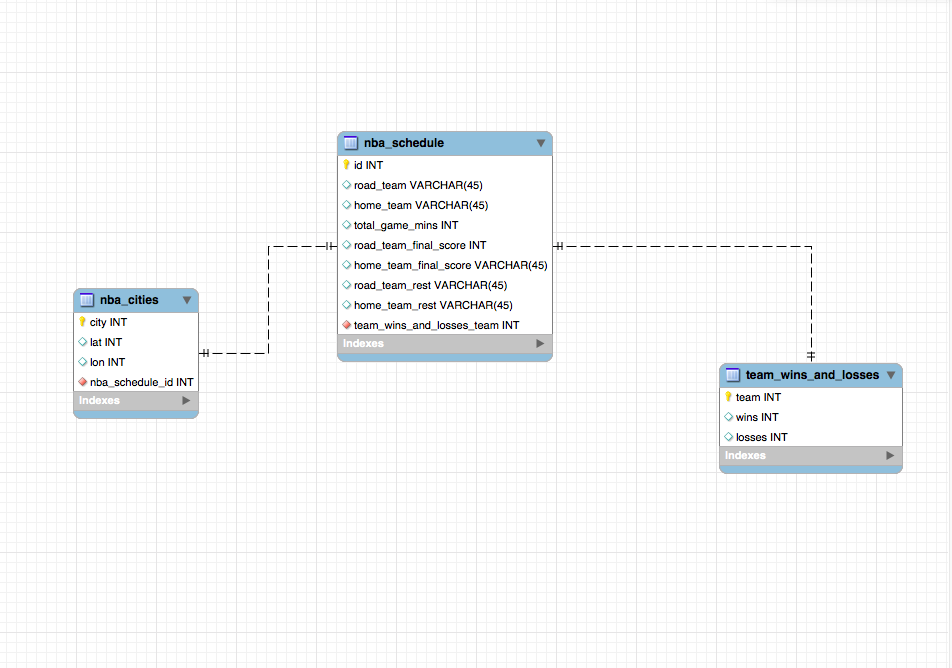
Stephen Kay

CSCI-3287 Final Project

Beating the system:

As previously described the goal of my project was to make a multi-relational database and then use it to run queries against multiple tables to determine if there is any insight into NBA regular season data that could help a sports better make more statistically supported bets.

According to professionalgambler.com a professional sports bettor rarely sustains a long-term winning percentage of higher than 55%. Thus, the goal of my project was to see if there were long term bets across an NBA season that an individual could make in order to beat a winning percentage of 55% and therefore come out in the positive for their season total bets. In order, to be able to run my queries and join different table in an efficient way I created a multi-relational data base with the structure shown in the figure below.



As you can see my multi-relational database consisted of three table. Team and city names were then used to run various queries and joins across these tables to answer questions in my problem space.

**In summary the tools used for the project were:**

* A PostgresSQL database using Elephant SQL as the remote host
* mySQL workbench for the creation of the EER-Diagram
* A Jupyter notebook for data analysis and as a data-base connection interface
* SQLAlchemy and pandas for a database connection interface and pandas for creating the initial data formats
* Postgis in order to create Geography points for my queries to then include team travel distance in my analysis

**Learning Goals:**

* Create a multi-relational database with three tables and access it with a Jupyter Notebook
* Perform a variety of unique joins and table queries to gain insight into the data

**Key Results:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Results When the Home Teams Previous Season Winning Percentage was NOT Taken Into Account** | | | |
| **Win Total** | **Road Team Rest** | **Home Team Rest** | **Win Total** |
| **15** | **B2B** | **1** | **31** |
| **5** | **B2B** | **2** | **7** |
| **26** | **3IN4** | **1** | **37** |
| **9** | **3IN4** | **2** | **11** |
| **36** | **3IN4-B2B** | **1** | **58** |
| **11** | **3IN4-B2B** | **2** | **26** |
| **Sum Total of Road Team Wins** | **…** | **…** | **Sum Total of Home Wins** |
| **102** | **…** | **…** | **170** |

In summary when the home team is in a rest advantageous situation they win 170/272 or 62.5% of the time while when the road team is in a rest disadvantageous situation they win 102/272 37.5% of the time

|  |  |  |  |
| --- | --- | --- | --- |
| **Results when the Home Team Won at Least 43 or More Games (I.E., they were good)** | | | |
| **Win Total** | **Road Team Rest** | **Home Team Rest** | **Win Total** |
| **4** | **B2B** | **1** | **23** |
| **3** | **B2B** | **2** | **6** |
| **11** | **3IN4** | **1** | **28** |
| **5** | **3IN4** | **2** | **5** |
| **12** | **3IN4-B2B** | **1** | **39** |
| **3** | **3IN4-B2B** | **2** | **20** |
| **Sum Total of Road Team Wins** | **…** | **…** | **Sum Total of Home Wins** |
| **38** | **…** | **…** | **121** |

Ok now we have some data that we can bet on... if the home team is good and the road team is not well rested then the home team wins ~69 % of the time and the road team wins ~31 % of the time % SHOW ME THE MONEY I SUPPOSE

**In Summary:**

The two above Tables show the results of numerous queries and table joins that I describe in my video. Then end results is that we’ve created a budget situation where we can win at either a 62% or 69% win rate which is much better than the long term goal of a professional gambler at 55%. Also the tables were successfully related and queried from a remote Elephant SQL database, thus completing my intial learning goals.