NYU Tandon School of Engineering
Computer Science and Engineering
CS6083-INET, Spring 2018

February 8, 2017

Homework #1 (due February 18)

Note: In this homework, you do not need to create tables and execute queries using an actual DBMS. This will however be needed for the second homework set. A written solution is sufficient for now. Also, while the problems talk about relational schemas for websites, you do not have to worry about how to build such a site, but only need to think about suitable underlying schemas.

Problem 1 (40 points): Consider the following relational schema for a website for fans of live music that keeps track of artists (individual musicians or groups), concerts by artists, and the times and venues (places such as bars, concert halls, etc.) where concerts take place:

Artist (aid, aname, adesc)
ArtistGenre (aid, genre)
Concert (cid, vid, cdate, cstarttime, cendtime, cdesc)
ConcertArtist (cid, aid)
Venue (vid, vname, vaddress, vcity, vphone)
TicketType (cid, ticketclass, ticketcost)

Thus, for each artist we have an aid, a name, and a short description of their music. An artist can also play one or more genres, such as Jazz, Americana, Soul, etc. For each concert, we store where it takes place (the venue), the date, the start and end time, and a short description. A concert may involve several artists. For each venue we store its name, address, and contact phone number. Finally, we store information about the prices of different categories of tickets for each event. Note that the site is not keeping track of individual tickets and who purchased them; it only provides information about upcoming concerts and their ticket prices (probably with hyperlinks to the actual sites selling tickets).

(a) (2 points) Why does TicketType and ConcertArtist have two primary keys? What would be the problem of using only one primary key? Explain.

(b) (2 points) Identify suitable foreign key representing the tables.

(c) (12 points) Write statements in SQL for the following queries.

1. Output the names of the venues in city ‘New York’.
2. Output all the concerts that were played in ‘Chicago’ in year 2017.
3. For each venue, output the number of artists played in year 2017.
4. Give the cid and ticket class where the cost of ticket exceeded $100.
5. Output the aids of all artists who have appeared together with “Bruno Mars” in at least two concerts during 2017.
6. Give the names of the artist who have done the concerts at all the venues at least once.

(d) (12 points) Write relational algebra statements for the above queries if it is possible, or give a reason why is it not possible.

(e) (12 points) Write statements in (Domain or Tuple) Relational Calculus for the above queries or explain why it is not possible.

Problem 2 (16 points): In this problem you need to design a relational schema for a baseball stadium booking agency. This website is used to let various baseball teams book the stadiums on demand.

For each baseball teams, it has information like their name, no. of players, home city, coach name. You also need to store the unique ids for each of the teams. For each stadium, there would be different kinds of information like name, city, area, capacity, built in (year), manager name and would also have a unique id.

Each stadium has a schedule on the website, so that it is possible to search which stadiums are available at a particular time. For simplicity, you may assume that the schedule consists of slots lasting 3 hours. Teams may also upload stadium requests, specifying the time and date.

When a team and stadium manager comes to an agreement about the requested time, a booking is confirmed. All the necessary information regarding the booking need to be stored including the slots, booking fee, etc. A team can also rate the stadium between 1 (very bad) to 5 (very good).

(a) (8 points) Design a relational database schema for the given problem, with suitable table, attributes, primary keys, and foreign keys. Discuss any assumptions you are making in your design! (Do not submit an E-R diagram, we will cover E-R diagrams later in the class!)

(b) (8 points) Write SQL statements for the following queries. If your schema does not support these, you need to modify it appropriately.

1. Output the names of all stadiums available between 9 am to 6 pm on Feb 10th, 2018. (They should be available all the hours, not just part of the time.)
2. For each stadium, output their ids, and the number of distinct team they have been booked to in 2017.
3. Output the stadium(s) who earned the most rating overall during 2017.
4. Output the names of teams who have booked at least 5 stadiums, but who have never given a rating of 4 stars or higher.