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

February 22, 2018

Homework #2 (due March 4)

Problem 1 (26 Points):

You are hired to create a database design for a company called EasyStore. The company has a huge amount of data which becomes difficult to manage. EasyStore has contracts with different cloud service providers. The companies approach EasyStore with their data to help them manage it through cloud services. The companies make a deal with EasyStore to manage X amount of data to be stored by some cloud service. Each deal will be associated with a single service provider. Also, the transaction for the deal is done between a company's bank account and EasyStore. For simplicity consider each company has only 1 bank account and each deal is valid only for a month.

For database design, you need to consider information for each company, like CompanyId, CompanyName, CompanyCity, CompanyRevenue, etc. Bank account information related to each company like AccNo, BankName, Balance, RoutingNo, etc. You can consider different kinds of cloud services for storing data like AWS, Microsoft Azure, Dropbox, Google, etc. Manage the details of a deal made by a company and EasyStore like the quotation at which the deal was closed, the service provider they opt for, amount of data management agreed in the deal, etc. A company can even rate the deal (1-5) based on the service provided.

a) (10 points) Design an ER diagram that can model the above scenario. Identify any weak entities, suitable keys, and the cardinalities of the relationships. Discuss any assumptions that you are making in your design.

b) (6 points) Convert your ER diagram into a relational schema. Show primary key and foreign key constrains.

c) (10 points) Write statements in SQL for the following queries. Note that if your schema does not allow you to answer a query, you may have to go back and change your design.

1. For each company output the amount they have spent in total for all their transactions.
2. Give the total amount of data managed by each service provider for deals closed during 2017.
3. For each cloud service providers, output the avg rating they have received during 2016.
4. Output the companies who paid more than $1,000 per GB for each of the service providers.
5. Output the companies who have opted at least four times for same service provider in same year.

Problem 2 (24 Points):

In this problem, you have to write SQL and RA queries for a database modeling the short-term leasing of houses or apartments, in a system somewhat similar to Airbnb. Here are the tables:

Customer (cid, cname, cphone, ccity);
Landlord (lid, lname, lphone, lcity)
Residence (rid, rname, rstate, rcity, raddr, rtype, rarea, lid);
Leases (cid, rid, startdate, enddate, price);
Rating (cid, rid, rtime, score);

Customers are identified by a cid, and we also store their name, phone number, and the city they live in. Landlords are identified by a lid, and have a name, phone number, and city. One landlord could own multiple houses/apartments, but a house could only have one owner. Residences are identified by a rid, and have a rname, and rstate, rcity, and raddr to store the precise address of the house (e.g. rcity = 'Brooklyn', raddr = '3rd floor, 308 45st, 6 Avenue'), an rtype that stores the type of residence (e.g studio or 2BR-1BA), and rarea indicating the square footage. Leases are identified by cid, rid, lid, and startdate, and we also store enddate and price. The startdate and enddate attributes store both time and date information. When the lease ends, customers could give a rating to this house, where a rating contains cid, hid, rtime, and a score. rtime is a timestamp which indicates the date and time when the customer made the rating; scores are ranging between 1 star (Terrible!) and 5 stars (Awesome!).

a) (6 points) Create the above schema in your database system of choice (MySQL, SQLite, etc). Choose appropriate attribute types and define primary keys, foreign keys, and other constraints. Submit your SQL DDL statements.
b) (10 Points) Write SQL Statements for below queries:
    1. List the pairs of landlord and residence where landlord lives in Brooklyn and residence is in 'Queens'.
    2. During 2017, for each month when lease started, output the total number of residences in 'Chicago'.
    3. Output the residence type that is owned by the largest number of distinct landlords.
    4. Output the customer(s) who took the most expensive residence(s).
    5. Output the average amount of rent earned by landlord.
c) (8 points) Write Relational Algebra Expressions for queries 1 to 4.