

ADVANCED DATABASES – LAB. 2, NORMALIZATION

26th September 2016

Dr. Pierpaolo Dondio

Exercise 1

The Gill Art Gallery wishes to maintain data on their customers, artists and paintings. They may have several paintings by each artist in the gallery at one time. Paintings may be bought and sold several times. In other words, the gallery may sell a painting, then buy it back at a later date and sell it to another or the same customer (but one transaction per painting per day!).

Produce a 1NF, 2NF, 3NF to store the data of the following form. The form contains all the paintings sold by the gallery to its customers (it does not contain the paintings that were bought by the gallery)

Gallery Customer History Form

Customer Details

Jackson, Elizabeth
Phone (206) 284-6783
123 – 4th Avenue
Fonthill
Zip Code: L3J 4S4

Purchases Made

Code	Artist	Painting code and Title	Purchase Date	Sales Price
03 -	Carol Channing	P1 - Laugh with Teeth	09/17/2000	7000.00
15 -	Dennis Frings	P2 - Toward Emerald Sea 0	5/11/2000	1800.00
03 -	Carol Channing	P9 - At the Movies	02/14/2002	5550.00
15 -	Dennis Frings	P2 - Toward Emerald Sea	07/15/2003	2200.00

In the above example, the galley sold 4 paintings to the customer Jackson, the paintings were made by two artists. The painting "Toward Emerald Sea" was sold twice (since the gallery can sell, rebuy and resell). Note that each artist has a unique code, while each painting is numbered for each artist (i.e. the painting code is not unique, it is unique only for a specific artist).

Store the customer address using the following fields: street, city and zip code. The zip code identifies a city. Store the phone as a single (text) field.

Exercise 2

You are required to:

1. Create the table **Apps_NOT_Normalized** running the sql script applications.sql (from webcourses).
2. Normalize the following **Apps_NOT_Normalized** table respecting all the assumptions listed below
3. Create the normalized tables with sql developer and move data into the new tables
4. The normalized tables are usually more storage-efficient. What is the gain in storage efficiency (=the size in bytes of the normalized tables divided by the size of the starting non normalized table)?

(In order to estimate the size of the table, assume that a varchar(X) has a size of X bytes, a date has a size of 8 bytes, an integer has a size of 4 bytes and a number(2) has a size of 4 bytes).

Table Apps_NOT_Normalized:

The following data describes information stored about students applying to a graduate school. Initially they are stored in one single table with no data normalization.

Application-No + StudentID+ StudentName + Street + State + Zip-Code + ApplicYear + Reference-Name + RefInstitution + Reference-Statement + Prior-School-Id + Prior-School-Addr + GPA

The table contains information about the application number (unique for each year), the student ID and name, the student address (street, state, Zip-Code) the year of the application, the name of one or more referees, their institutions and the reference statement written by each referee (so one reference statement for each referee for a specific application), and a list of prior schools the student attended. For each prior school the DB stores the address and the GPA (final score obtained by the student in that school).

Assume the following:

- Applications are submitted for each year, applications numbers are reset every year and therefore they are unique only inside each year.
- Student ID is unique for all the years (is assigned once for all once the student submit her/his first application)
- A student might move to another address and the database has to store all the students addresses
- An applicant can only apply once during each year.
- Reference-Name and RefInstitution together are unique (but are not unique as separate attributes).

- Prior-School-Id uniquely identifies a university or college.
- A student might have many prior-schools **and** if she/he sends an applications in different years, it might have added a new prior school (therefore prior school lists might change across applications).
- An applicant has only one GPA from any specific prior school.
- For each application there could be one or more referees
- Referees may write a reference statement for more than one applicants.
- However, the reference statement for different applicants is different.
- Reference statements are attached to an application. If a student submit a second application, the reference statements (even from the same referee) could be different.