CS586/486 Introduction to Databases

Winter 2016 Quarter

Graduate Project Database Implementation

Due Dates:

1. Domain Description and Data Source – Monday, 1 Feb. 2016, before class
2. ER Diagram and Relational Schema – Monday, 22 Feb, 2016, before class
3. Final Write-up with Queries and Results – Friday, 10 Mar. 2016 (due electronically), by 5p

You may do this assignment individually or you may work with one partner. That is, this assignment is to be completed by individuals or by teams of two students. If you work with a partner, then you should turn your assignments jointly, with both of your names on the paper. You should only talk to the instructor, the TA and your partner about this assignment. You may also post comments and questions on Piazza.

Please turn in the first two submissions on paper in class. The third one should be submitted electronically, either as a PDF document or as a link to a web page.

Project Overview

The goal of this project is to gain experience with database design and implementation. You need to choose a small domain (for example, food carts or campus sports teams) for which to implement a database, along with populating it with real data and producing example queries to execute over it.

Submissions

There are three submissions for this project.

1. Domain Description and Data Source: Describe in English the domain you intend to build a database for, along with 20 questions (in English) that someone might want to ask about the domain. (You will be permitted to revise these questions later if needed.) Also describe what source you intend to use for data, and *how you intend to ingest the data into your database*. You should choose a domain where you can easily get several hundred rows of data.

2. ER Diagram and Relational Schema: Produce an ER diagram for your domain, and its translation into a relational schema. You should aim for a database with 6 – 10 tables. You should also submit evidence that you have created at least one table from your schema and populated it with at least one row.

3. Final Write-up: You are to implement your schema in a relational database and populate it with data. The preferred DBMS is Postgres. If you want to use a different DBMS, you must get advanced approval. You then need to translate your 20 questions into SQL and execute them on your database. Your write-up of these activities should include the following:

- Your ER diagram, showing any changes you made during the implementation process
- The CREATE TABLE statements for your database
- A brief description of how you populated the database
- A listing of the contents of all of your tables
- For each of your 20 questions, the question in English, its translation to SQL and the (full) answer to the query. (If you needed to change any of your original questions, also list the originals and why you needed to change or replace them.)

Please submit your write-up as a single PDF or as a link to a web page, which can in turn link to the different parts of the assignment.

Grading

Grading on all submissions will include a component for spelling and grammar.

Submission 1 is worth 10 points.

Submission 2 is worth 20 points. Part of the score will be on how realistically your design represents the corresponding real-world domain.

Submission 3 is worth 70 points, and the score will include the following factors, in addition to general correctness:

- Were you able to avoid large amounts of manual data entry with your approach to population?
- Does the data used do a good job of demonstrating the adequacy of your schema and the correctness of your queries?
- Does your set of queries make use of all the parts of your schema and a range of SQL features?

The points breakdown for this submission is
I. Grammar and completeness: 15
II. Data entry – description and automation: 15
III. Data suitability: 10
IV: Queries – correct translation of English, coverage of schema and SQL features: 30