ECS650/ECS789 Semi-structured Data and Advanced Data Modelling

Coursework 1 Database development and tuning

This coursework should be carried out in groups of no more than four students. Ideally you should work in a group of three, although groups of between two and four students are acceptable. You can choose who to work with, but your group must consist of either MSc or BSc students. MSc and BSc students cannot be together in a group. If you have problems joining a group, contact me at t.stockman@qmul.ac.uk

Requirements

You have been commissioned to design and build a database for a taxi company. The system should be built using the Oracle database management system. The system is required to store information about the Company according to the requirements below.

1) Drivers: including details of where they live, their other contact details and information about their employment record with the Company.

2) Cars: details of the registration number, age, date of last MOT test, status of cars: e.g. roadworthy, in for service, awaiting repair, written off, details of who owns the car.

3) Operators: the taxi firm employs 8 operators who take bookings and allocate them to drivers.

4) Bookings: details of bookings taken over the telephone.

5) Payments: you need to decide how the details of payments are collected and associated with the correct booking.

6) Shifts: the drivers work a shift system in order to ensure the Company is effective over 24 hours. You need to be able to track which drivers are on which shift. Similarly the operators work a shift system to cover each 24 hour period.

7) Clients: in addition to taking whatever bookings come in via the phone, the Company has a number of registered clients who have regular bookings. There are two types of client, corporate and private. Clients can make regular bookings, the details of which must be stored. For example, these may be for daily bookings, once weekly bookings and so on.

8) Revenue: you need to be able to record the amount of money earned by drivers and the amount paid by them to the Company (to cover costs of car maintenance, salaries of operators, lighting, heating and general taxi office costs). Drivers can be employed by the Company either on a fixed-fee or percentage-of-receipts basis.

Deliverables

You are required to produce the following:

1. A list of any assumptions you have made during the development of your system (half a page).
2. A UML class, ER or other diagram showing the different entities and relationships in your design (maximum 1 page).
3. A description of how you mapped your design into a set of tables to be implemented (maximum 2 pages)
4. A relational schema listing the tables and attributes in your design in the form:  
   Table\_name (attribute1, attribute2, ….., attributen) identifying primary and foreign keys (maximum one page).
5. One or more SQL scripts to create the tables with appropriate use of declarative constraints and to populate the tables with data.
6. A set of 12 SQL queries to extract information from the system. You should aim to use a good range of language constructs to demonstrate your understanding of the language. Each query should include:
7. A brief statement of what the query is intended to do
8. The SQL code
9. The output produced by the query
10. A set of 6 triggers to perform validation or other processing you believe to be useful in the system. Each trigger should include:  
    I) A brief statement of what the trigger is intended to do  
    II) The trigger code
11. A demonstration of the use of the trigger

H) Examples of performance tuning experiments to test the efficiency of queries on the system. To undertake these experiments you are likely to need to substantially increase the volume of data in one or more tables in your system. Your experiments might investigate changes to the implemented schema, the use of indexes and/or query rewriting to examine tuning alternatives. There is no specific number of experiments to be submitted because some experiments require more effort than others, but we suggest that you use a minimum of 4 different interventions (schema changes, use of indexes and/or query rewriting). The description of each experiment should be no more than one page, if you wish to include further details then please put these in an appendix.

1. A design document describing how you might migrate part of the design of your relational schema to a MongoDB database, describing the collections you would create and how you would exploit the mechanisms available in MongoDB such as arrays and embedded documents for the modelling of semi-structured data. It is suggested you choose two suitable subsets of the schema that could contain semi-structured data to model. Two possible suggested subsets include 1) the details (including addresses and other contact details) of employees and 2) their employee histories. If you prefer to use other parts of your schema for this exercise that is ok providing a case can be made for why they could contain semi-structured data. (maximum 3 pages).

Marking schemes for coursework 1

The following notes are provided to assist you in judging the standard of work required for your coursework submission.

Undergraduate marking scheme:

To obtain a mark above 70% (First class Honours): all parts of the required work must be completed to an excellent standard, including a high level of detail and accuracy. Evidence of an ability to address issues beyond the given specification, and use of materials beyond those in the taught course will normally be expected for this grade.

To obtain a mark between 60-69% (2:1 class Honours): Most parts of the work will be completed to a high standard of detail and accuracy, though the work may include a small number of minor errors. The work should demonstrate a good a consistent level of ability across the full range of requested deliverables. The work may at times demonstrate the characteristics of work deserving of a mark over 70%, but at the same time include flaws which prevent it from being rated at that level.

To obtain a mark between 50-59 (2:2 class Honours)% Most deliverables should be developed in a competent manner, though a higher number of minor errors, or a small number of more significant errors may be present.

To obtain a mark between 40-49% (3rd class Honours): the work must demonstrate some real understanding and competence by the students, provide a serious attempt at all of the requested aspects of the work, though parts of it will contain significant errors of design and/or implementation.

Marking scheme for postgraduates:

To obtain a mark above 70% (Distinction): all parts of the required work must be completed to an excellent standard, including a high level of detail and accuracy. Evidence of an ability to address issues beyond the given specification, and use of materials beyond those in the taught course will normally be expected for this grade.

To obtain a mark between 60-69% (Merit): Most parts of the work will be completed to a high standard of detail and accuracy, though the work may include a small number of minor errors. The work should demonstrate a good a consistent level of ability across the full range of requested deliverables. The work may at times demonstrate the characteristics of work deserving of a mark over 70%, but at the same time include flaws which prevent it from being rated at that level.

To obtain a mark between 50-59% (Pass): Most deliverables should be developed in a competent manner, though a higher number of minor errors, or a small number of more significant errors may be present.