CPT140 Database Concepts

Study Period 4, 2011

Assignment 1

Due: 11:59pm, Sunday 8th January 2012.

Worth: 100 marks in total, which is worth 20% of the overall assessment for the course.

Submission: Submission is via the WebLearn assignment submission system.

- To access WebLearn, log in to myRMIT and go to the Studies tab. Look under Courses and find Database Concepts. Click the Weblearn link.
- In WebLearn, click on the Assignments button on the right-hand-side and follow the instructions
- Select "Ian Baker" as tutor.
- Submit only one file in pdf format or a zip file containing pdfs.
- Each time you submit, the old submission is over-written with the new submission.
- Make as many partial submissions as you like, up to the due date.
- Keep a backup of your submission.
- It is important to keep your confirmation email from WebLearn. It is your proof of submission.

Assessment: This is an individual assignment. The minimum penalty for plagiarism is loss of marks for this assignment. If this means that a hurdle requirement is not met, the student fails the course. Please read http://www.cs.rmit.edu.au/students/integrity/ for further information.

Late Submission Penalty: WebLearn will remain open for 5 days after the Due Date for late submissions. A late submission incurrs a penatly of 10% per day or part day late. As an Example, an assignment submitted at 2AM on Tuesday morning will incurr a 20% penalty. Submission more than 5 days late will not be accepted.

Extensions and Special Considerations:

To request an extension or Special Consideration, please email the details to

Email : ouacsit@rmit.edu.au

Please note the following points about extensions and special consideration;

- I am unable to (as in, not allowed to) grant extensions. The Admin people are the only ones who can.
- You must apply at least three days before the due date of the assignment.
- You must apply on the appropriate form.

- You must have supporting documentary evidence.
- The grounds for extensions are very limited and do not include excessive workload or tiredness.

Other Information : The aim of the assignment is to practice the concepts learnt from the material and to allow you to test your understanding and learn more about each topic.

The questions as presented are not meant to represent real-world situations. They are contrived in a manner that tests your knowledge. You are expected to review the text and Blackboard materials to work out the answers. There are no 'trick' questions.

You should be aware that the material required to answer some questions will not have been covered as of the release date of the assignment. It is covered in subsequent weeks. None of the material covered in the week directly prior to the Submission Date is required to do the assignment.

Not every situation in the assignment (or in real life) is presented in the references. You may have to ask questions (of me or others) in the Blackboard discussion forums while doing the assignment.

Start each answer with the number of the question you are answering.

Please set out SQL answers as you see them in the text and lectures, that is using multiple lines, without line numbers. This makes testing your query and analysing it for part marks easier.

I do not do 'reviews' of submissions before the due date. I will not answer "Is this right" type questions. I will answer requests for clarifications for any of the questions. I will also answer specific questions like, "I can get this far, but I do not know what the next step is". I will typically give a clue or hint rather than solve it for you.

The number of marks for each question is given with the question. The questions cover material in Weeks 1 to 5 (inclusive).

It is important that you make an effort to submit. This assignment is one of two assignments. If you do not submit, you will have to score 100% for the second assignment to pass the Assignment Hurdle (which is a big task). Even submitting an incomplete assignment for Assignment 1 and failing with a low mark will be better than not submitting at all.

Students often require clarification for some questions. This is done in the Discussion Forums in the course Blackboard. Please check the Discussion Forums regularly for information.

You ***MUST*** put your name and student number on your submission.

The ***ONLY*** acceptable format is PDF.

Good Luck with your assignment!

Question 1. (30 marks)

The MovieBuster online movie review service has a database as described below:

MovieInfo(mvID, title, rating, year, length, studio)

GenreInfo(mvID*, genre)

DirectInfo(mvID*, Director)

- Primary keys are underlined and foreign keys are denoted with asterisks (*).
- Movie information is kept in the MovieInfo relation, where mvID is an artificial ID for movies.
- A movie can have more than one genre and may be directed by more than one director.
- Note that tables have names different from those in the Movies database used in lectures.
- To access the MovieBuster database from SQLDeveloper, use the same techniques for connection to any Read-Only database (as shown in the video) and use ;
    - Username : romb1 *note the last character is the number one.*
    - Password : mb1 *and again last character is the number one*

Write an SQL query for each question below to extract information from the database. Each question is worth 6 marks. **Do not** supply the output of the query. Only the SQL query is required.

1. Make a list of movies in the MovieBuster database. List the title and the length as well as its rating. The output should be sorted by increasing order of the length.
2. Are there any directors whose name starts with the letter "A". Ensure that any name only appears once in the output.
3. Make a list of movies that have a G rating. In addition, these movies must be either longer than 90 minutes or made in the year 1990. Output the title of the movie.
4. When it comes to movies, some studios are more productive than others. List each studio and how many movies have been made by that studio. The most productive studios should be at the top of the list.
5. List all studios that have only made one movie. The list should be in alphabetic order.

Notes.
Some Common Issues for SQL Questions.

- Each answer should be set out with multiple lines as shown in lectures and tutes.

- Answers should not include line numbers.
- You are expected to develop your answers on RMITs Oracle system. Your queries should execute without error on this system. This is the system that will be used to test your queries.
- Logic is more important than efficiency. There are no marks for efficient coding. At this stage, it far more important to understand the logic of your query and that it produce the required result.
- Not every question needs to produce an output. Having no output from a query still tells you information about the data you did not know before and can represent a good result.
- Despite instructions, student still supply the output of their SQL query as their answer. Please do not do this.

Question 2 (40 marks).

A catering company wants to use a database to help manage part of its operations. One of the companies activities is to purchase items from supplies and on-sell them to customers. Use the following information to understand the design requirements;

- The company buys specialty items (like cheese cakes and pasteries and the like) from suppliers. The databse will record the name of each supplier, their address and phone number. No two suppliers have the same name.
- Items are identified by a type (like Lamington) There is a record of the quantity of each item, its cost and its margin (or markup).
- There are two ways items are sold. One is by orders. Orders are recorded in the database with an order number and payment method.
- The other way is by over-the-counter sales. In the database, these sales have a docket number and the employee ID of the sales person handling the sale.
- Each supplier only supplies one type of item.
- To ensure competitive pricing from suppliers, items may be supplied by multiple suppliers.
- An order typically has many items. The same can be said for sales.

Hints :

- Employee ID looks like it may be a primary key but there is no other mention of Employees in the question so this is not the case.
- Keep in mind that we will be covering how to do this type of problem in up-coming tutorials.

Draw the ER diagram for the required database. Use the symbols and notation as used in lectures and notes.

Notes.
Some Common Errors In ER Diagrams.

- If the scenario is set in say, a club, you do not need an entity for the club. The arrangement you depict in the ER diagram **is** the club.
- Entities do not connect to other entities with straight lines. Entities have relationships to/with other entities so there has to be a relationship between every pair of connected entities.
- Every entity must have a primary key.
- The names of all relationships and entities must be unique.
- There are no foreign keys in ER diagrams. They come about as part of the mapping process into a relational schema.
- Make sure you consider what multiplicity markings are required.

Question 3 (30 marks).

Convert the following ER diagram into a Relational Schema.
List the final full schema at the end showing all Primary Keys and Foreign Keys.

Proposed Car Dealership Distribution Database ER Diagram