Databases

Assignment

This assignment requires you to answer several questions on relational database principles and SQL, and to design a database based on a case study.

The assignment addresses the following learning outcomes for the unit:

* 1. Apply your knowledge of relational database principles and theory to create effective and efficient database designs.
	2. Define, create, and manage relational database systems using SQL.

You will need to ensure that your materials reflect a **high professional standard\***. The marks allocated for the Assignment will be allotted in terms of the percentages shown in the following table.

|  |  |
| --- | --- |
| Question 1: Relational algebra | 20 |
| Question 2: SQL Select queries | 20 |
| Question 3: Further SQL | 15 |
| Question 4: Normalisation | 20 |
| Question 5: Conceptual design | 25 |
| **Total** | **100** |

## \*Formatting and presenting your assignment appropriately is important, and the assignment includes marks for the overall organisation and presentation of the document (up to 5 marks). This includes marks for things such as:

* *Standard Font Size and Colour i.e., Arial/12.*
* *Formatting and layout: double line spacing, page number, separate sections (Headings/Subheadings), aligning paragraphs.*
* *Includes a title page and table of contents. Correct use of equations, footers, and headers.*
* *Good technical style, avoiding excessive jargon, grammar, and spelling errors.*

Before you submit your assignment, make sure to check the formatting and overall presentation of your document. Make sure the file type is a **single** MS word document or PDF. If you're not sure about what’s required contact your teaching staff or UC.

# Question 1: Relational algebra (20 marks)

A local insurance company manages its data using a relational database that stores information about insurance agents, clients, policies, and areas they operate in. The company offers various policies tailored to different clients’ needs e.g., life, health, auto, and home insurance. The assignment relation records instances of when an agent is assigned to sell policies in a particular area at a given period.

The schema for the database used in this question is as follows: (note that primary keys are shown underlined, **foreign keys in bold**).

AGENT (ID, FullName, Phone, Commission-Rate) POLICY (P-Number, P-Type, Description, Premium) AREA (A-Number, Population, Description)

AREA-POSTCODE(**A-Number**, Postcode)

ASSIGNMENT (**ID, P-Number, A-Number**, StartDate, EndDate, QuantitySold)

CLIENT (C-Number, FullName, DateofBirth, Phone, Suburb, StreetNumber, **Postcode**) POLICY-RECORDS (**P-Number, ID, C-Number**, StartDate, EndDate)

Provide **Relational Algebra** (NOT SQL) queries, with **description** as necessary to find the following information. Each question is worth 2 marks. NOTE:

* You can use the symbols , , etc or the words ‘PROJECT’, ‘RESTRICT’ etc as you prefer.
* Optimising query performance (providing efficient queries) will attract higher marks.
* Where you use a join, always show the join condition.
* Provide a detailed description to explain how the query will run.
1. Show the quantity of policies sold by agent “Paige Turner” in Area “101”.
2. List the premium of each available home insurance policy sold in Area “404”.
3. Show the area population and policy type of any policy where more than 10 of the policy has been sold in a particular area.
4. Find the name of any agent who sold policies in an area that covers postcode “0810” or in an area that covers postcode “0909”.
5. Find the name of any agent who sold policies in an area that covers postcode “0810” and in an area that covers postcode “0820”.
6. List the names of agents who have sold a policy that ends on 30-12-2024 .
7. Show the descriptions of areas that have had policies sold by agent “Florence Gump” or an agent with “2%” commission rate.
8. List all details of any agent who has sold policies in an area that has a population of more than“20000” people.
9. Find the type of any policy that was assigned to an agent commencing in 2024.
10. List the type of any policy that has been sold in all areas that commenced in 2020.

Question 2: SQL – SELECT queries (20 marks)


#### Provide SQL queries and the result tables for the following (20 marks):

*Please ensure that you include the result table as well as your SQL; you can copy and paste this from either your ssh client or SQL Developer. Each query is worth 2 marks. These tables exist in Rhea and are owned by the user “tutorials”. You may, if you wish, create your own copies of the tables under your own account. If you do so, you should ensure that you copy the sample data in tutorials’ tables.*

*These queries are based on the View Ridge Gallery database you have been using in the Lab sessions. Please see Chapters 6 and 7 of Kroenke ford background to the case and table structures.*

*Marks are allocated not only for correct answers, but also for* **best practice** *in the creation of the queries and for providing efficient queries. You should also include a description along with each query to explain how it will run.*

* 1. Find the details of any works of art along with the full name of the Artist who created the work that have at least two copies recorded in the database (i.e., a work that is listed in the database more than two times).
	2. List the details of any work of art (including the Artist who created the work) that has an Expressionist style.
	3. List the details of the works of art (including the Artist who created the work, and the acquisition and asking price details) currently held in the gallery (*i.e., works of art that have not been sold).*
	4. List the sales for each customer made in year 2011 (*i.e., when a customer purchases a work of art from the Gallery, a transaction line is created. For a purchase, there will be values in the DateSold and SalesPrice columns).* The query should include the details of the customer, the transaction and the work of art purchased.
	5. List the names of the deceased artists who lived over 90 years of age (*for example, an artist born in 1950 and deceased in 2001 has an age of 51).*
	6. The sum of the acquisition price of works of art for each year (f*or example, if there were two works of art purchased for $1500 and $1000 in 2019, and one work of art purchased for*

*$500 in 2020, then the sums would be $2500 and $500, for 2019 and 2020 respectively*).

* 1. Calculate the profit made on works of art that have been sold in year 2008 (*i.e., the profit/loss on an individual work of art is the difference between the acquisition price and the sales price*).
	2. Which artist has had the most works of art sold, and how many of the artist’s works have

been sold?

* 1. Sales of which artist’s works have resulted in the highest average profit (*i.e., the average of the profits made on each sale of works by an artist*), and what is that amount?
	2. Customer name of any customers who have an interest in ALL artists.

# Question 3: Further SQL (15 marks)

1. **Provide ALL of the SQL statements** required to insert the details of the following: A customer “Ann Doe”, of 45 Boring Street, Pert, WA, 6000, Australia (email: Ann.Doe@somemailservice.com) has sold a work of art called “Mona Misa” by the renowned French artist, Francois Smith (b. 2000) to the Gallery (*i.e., the Gallery* ***purchased it from him*)**. It is a unique Oil and collage and is 5.5 x 6.1 inch signed by the Artist. The purchase price was $52, and the transaction took place on 2nd Jan 2024. (7 marks)
2. You have been given the following specifications of a simple database for a local Gym that keeps record of Trainer, Sessions, and Customers.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | TrainingSession |  | Customer |
|  | SessionID (PK) |  |
| SessionName |  |
| RoomName | CustomerID (PK) |
|  |  | DateStarted |  | Name |
| Status | Address |
| **LeadTrainerID (FK)** | ContactPhone |
| **CustomerID (FK)** |  |

**Give the SQL to create the Training session table.** You may assume that the Trainer and Customer tables have already been created, and that the TrainerID and CustomerID columns are of the data type VARCHAR2(10). The status of the Session refers to the current state of the Session and can only be Active or Dropped. (5 marks)

|  |  |
| --- | --- |
| Trainer |  |
| TrainerID (PK) FirstName LastName Address Phone DateOfBirth **SessionID (FK)** |
|  |  |
|  |

1. The Session table must also include a record of the Budget which is allocated each year to cover its running expenses. **Provide the SQL to amend the original table design to allow for this change in requirement.** It is most unlikely (impossible) that a budget would ever be more than 99999 annually. (3 marks)

# Question 4: Normalisation (20 marks)

The following question is based upon a Car-Service relation which records the details of transactions occurring in an Automotive Mechanic and Repair business. You may assume the data are representative.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CustName** | **CustDOB** | **ServiceDATE** | **CarRego** | **CarMade** | **CarModel** | **Service FEE** | **Staff NAME** | **Staff speciality** |
| John Doe | 8/09/1998 | 19/5/2023 | 1y568 | Toyota | Camry | 199 | Ali | Service tech |
| Justina Thyme | 10/10/2000 | 8/4/2022 | 1y576 | Honda | CRV | 400 | Chip | Auto body |
| Jack Supp | 10/10/2000 | 9/5/2023 | 1t567 | Toyota | Camry | 199 | Ali | General |
| Jack Pott | 29/03/1981 | 2/10/2022 | 1t158 | Ford | Focus | Fifty | Holly | Service tech |
| John Doe | 8/09/1998 | 1/8/2022 | 1y458 | Tesla | Model X | 199 | Ali | Auto body |
| Luke Smith | 03/01/1991 | 15/9/2022 | 1i132 | Holden | Astra | 190 | Chip | Auto body |
| Shane Joe | 30/01/2004 | 15/9/2022 | 1l884 | BMW | Z4 | One twenty | Ali | 99999 |
| Luke Smith | 03/01/1991 | 10/8/2019 | 1i123 | Black car | Astra | 269 | Ali | General |

You have been asked to design a relational database for this system. You know that there are problems with the current design and that it will need to be modified in order to work effectively.

You need to write a 1-2 page report that addresses the following:

1. What are the specific problems associated with the current design and why do they arise?
2. How would you change the current design and how does your new design address the problems you have identified with the current design.

In order to receive high marks for this question, you will need to demonstrate an understanding of the theories discussed in Topics 1, 2 and 3, how they apply to this problem, and justify the changes you are making to the system. **Simply providing the amended design (even if it is correct) will only attract a small percentage of the marks for this question.**

# Question 5: Conceptual Design (25 marks)

“ToursNow” is a new family-owned business offering transport options for coach (Bus) tours in the Perth metropolitan area and selected regional areas. ToursNow has experienced significant growth in recent months due to a growing customer base from all different types. As the business expands its operations and services, they have recognized the need for an efficient relational database to keep track of their fleet, sales, vehicle maintenance, and information about passengers and staffing. ToursNow prides itself as having an extensive tour selection option which are available for charter 24 hours a day, seven days a week. More specific notes about ToursNow are listed below:

* Information about a route includes a unique code, its departing depot, its destination, and estimated departure and arrival times. To reduce costs, ToursNow only has non-stop charter tours with a single origin and destination.
* Tours are timetabled for a route on one or more dates with a Bus and staff assigned to each Tour, and the available seats (i.e., seats not taken) noted. In a staff deployment, the staff ID and the Role are recorded. It is a business requirement that the number of hours that Tour staff (i.e., Driver, Tour Guide, Tour Coordinator, etc…) are on a Tour must be noted. There is no such requirement for non-tour staff (e.g., maintenance crew, admin officers, operations staff, etc..).
* Buses have a registration number (rego), make, model, year, a capacity (number of passengers the bus can accommodate), and a scheduled-maintenance date.
* The maintenance record of a bus includes a unique number, a date, a description, the bus rego, and the staff (maintenance crew) responsible for the maintenance/repairs.
* Staff have a unique staff number, full name (First and Last name), a contact number (mobile phone) , and a job position.
* Passengers have a unique number, a contact number (mobile phone), and a name (it does not have to be their real name).
* A record is kept for tour bookings including a booking number, a route code, a passenger number, a date, a fee and the payment method (which can be in cash, cheque, crypto currency, or credit card). If the payment is by credit card, a credit card number and an expiration date are part of the booking record.

### What you have to do:

1. Create an entity-relationship diagram showing the data requirements of the system. Your ERD should be able to be implemented in a relational DBMS. You should use the ERD notation we have been using in the lectures, and should include a legend to explain the notation. You should include attributes in the ERD. The use of a drawing tool such as Visio will make this task easier. However, whichever tool you use, you must copy and paste the ERD into a word-processed document. This is because your tutor might not have access to the tools you have used. **Please note that hand-drawn ERDs are not acceptable.**
2. List and explain any assumptions you have made in creating the data model.
3. Follow the best practices taught in this unit! Refer to the lecture slides and learning materials.

### Some important things to note:

#### Part of understanding a system at sufficient enough detail to model well, involves asking appropriate questions. If you are not sure about some detail of the case study, you should ask on the Discussion Forum in LMS.

1. The University email server strips out any Visio (.vsd) files that are sent; even if they are included in a zip archive. So, if you want to send a draft of your design to your lecturer by email, you will need to change the extension to something other than .vsd (.blah works well) or paste the diagram into a word document.
2. Marks will be allocated to each of the following functional areas:
	1. Legend
	2. Entities
	3. Have you included all required entities in the design?
	4. Have you included specialisation hierarchies where appropriate?
	5. Relationships
		1. Have you designed relationships between the entities that will support the functional requirements?
		2. Are your relationships correctly annotated?
	6. Will the overall design support the enterprise requirements?
	7. Following the best practices taught in the unit e.g., correct naming formats (entities, attributes, others), etc..

As Assignment 2 will require you to implement the OWO database, you will need to take into account the feedback you receive on your conceptual design when commencing your logical and physical designs.