Exam - Database Systems and Information Modeling INFO90002_2022_SM1

Started: Jun 10 at 15:01

Quiz Instructions

INFO90002

Database Systems and Information Modelling

EXAM

Semester 1, 2022

Reading Time: 15 minutes

Writing Time: 120 minutes (2 Hours)

Upload Time: up to 30 minutes

Authorised Materials:

While you are undertaking this assessment you are permitted to:

- make use of lecture slides (including soft copies)
- make use of any lecture notes and recommended texts
- use the subject materials and your laptop/PC in this exam

While you are undertaking this assessment you MUST NOT

- Copy material without attribution (including slide notes and shared study notes)
- Plagiarise
- Collude with any other person in any form

- Make use of any messaging or communication technology
- Make use of any world wide web or internet based resources such as Wikipedia, stackoverflow, google, social media, or any search engine services
- Act in a manner that could be regarded as providing assistance to a student who is undertaking this assessment, or in the future, will be undertaking this assessment
- Seek assistance from any other student who is undertaking this assessment, or in the future, will be undertaking this assessment
- Record, broadcast, stream or distribute in any way your exam attempt

After completing the exam, remember to upload your exam paper to OneDrive **ExamSubmission. (https://unimelbcloud-my.sharepoint.com/:f:/g/personal/lindent_unimelb_edu_au/Eoy_XeiNGr5Lr1Vt6GNdk48B816eVMZnL0epQIvC5vzGQw)** This link is also available at the bottom of the exam.

Instructions to Students:

- This exam has 8 questions. Attempt all questions and sub-questions.
- We recommend using pencil and paper for modelling questions to save time
- The total for this exam is 100 marks representing 50% of your final assessment
- This exam is a timed assessment which must be completed within **120 minutes** of the official commencement of writing time
- Questions can be answered in any order (please number your attempts)
- Start each question on a new page
- **PLEASE DO NOT USE RED font colour or RED pens**
- You must not communicate with other students whilst taking this exam, e.g. using messaging, chat rooms or email

IMPORTANT – YOU MUST READ THIS SECTION

- **Your file upload must be a single Word .docx or .pdf document before the elapsed time.**
- **Other document formats will NOT be assessed (e.g. Pages, .txt, .sql, etc).**
- **Email submissions will NOT be assessed.**
- **Every question attempt must be numbered (e.g. Q2C, Q1, Q6) to ensure it is assessed.**

- **The official exam language is English. Sections of the submission in languages other than English will NOT be assessed and will be marked as 0.**
- Before submitting your solution document, check that the diagram(s) is/are readable. It is your responsibility to ensure your answers are readable and make sense to the marker.

Download the docx answer **template**.

Rename the file to StuID_Last_FirstName_INFO90002Exam.docx, e.g. 1234567_Linden_Tanya_INFO90002Exam.docx. Make sure your student ID is correct and you spell your name correctly.

You must enter your student ID and your first name at the top of this file.

Question 1 – ER Modelling (20 Marks)

CreaSoft Software Developing Company

CreaSoft is a small software company that develops custom IT solutions for various businesses. For each business they record an Id, Business Name, Contact Person, Email and Phone No.

Each business may have one or more projects given to CreaSoft to work on. Each project has a Project Code, a description, a start date, and when the product is delivered, the completion date is also recorded. CreaSoft allocates one or more staff to work

on each project. One staff member is nominated as the Project Leader. Staff may work on many projects for different clients during the same time period, however, there may be staff members who are not allocated to any projects (e.g. while on leave).

Employees need to keep daily records of the time spent on each project (the date and number of hours). Staff may work on one or more projects on the same day.

For each staff member CreaSoft records staffID, first and last name, email, formal qualifications including awarding institution and year qualification was obtained, and skills the person possesses. The list of skills is stored in the database as skill code and skill description. For example, Alice may have the following list of skills: UX (user experience design), INTc (interviewing clients), FG (running focus groups with clients), JAVA (programming in Java), PYTH (programming in Python), whereas John may have a slightly different set of skills: JAVA (programming in Java), PYTH (programming in Python), SA&D (systems analysis and design). There may be skills that none of the current staff possess (e.g. a staff who can install and maintain MongoDB server left and a replacement has not been recruited yet).

Qualifications are recorded as qualification code and description, e.g. B.Math - Bachelor of Mathematics, M.IT – Master of IT, M.Eng - Master of Engineering. An employee may have multiple qualifications. Several employees may have the same qualification (although they may be obtained at different institutions/universities). For example, Alice may have M.IT from the University of Melbourne and John have M.IT from Swinburne University of Technology. Note, there may be a qualification on the list that none of the current staff members possess.

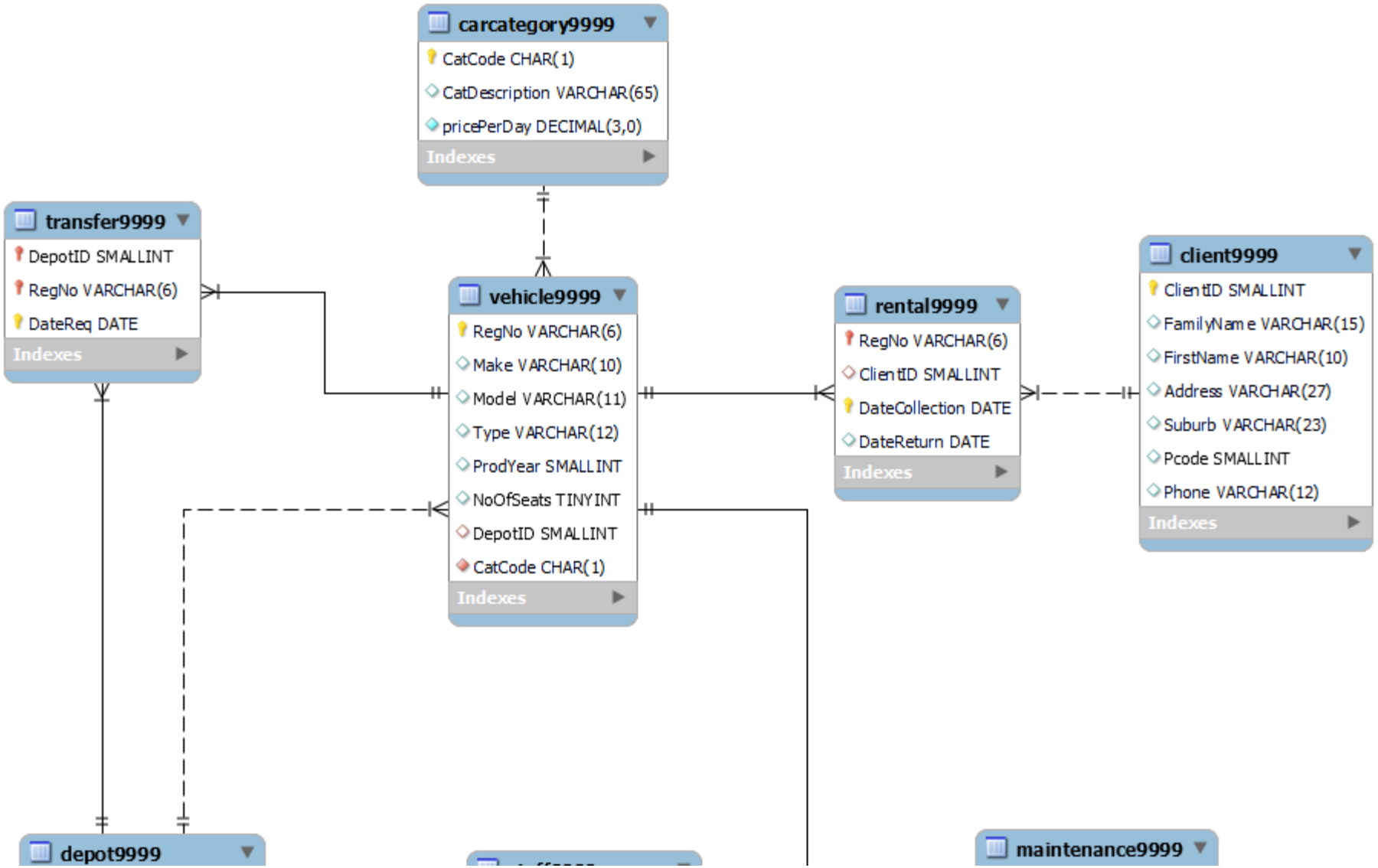
Q1. Develop a **physical** Entity Relationship model in Crows foot notation for the CreaSoft case study.

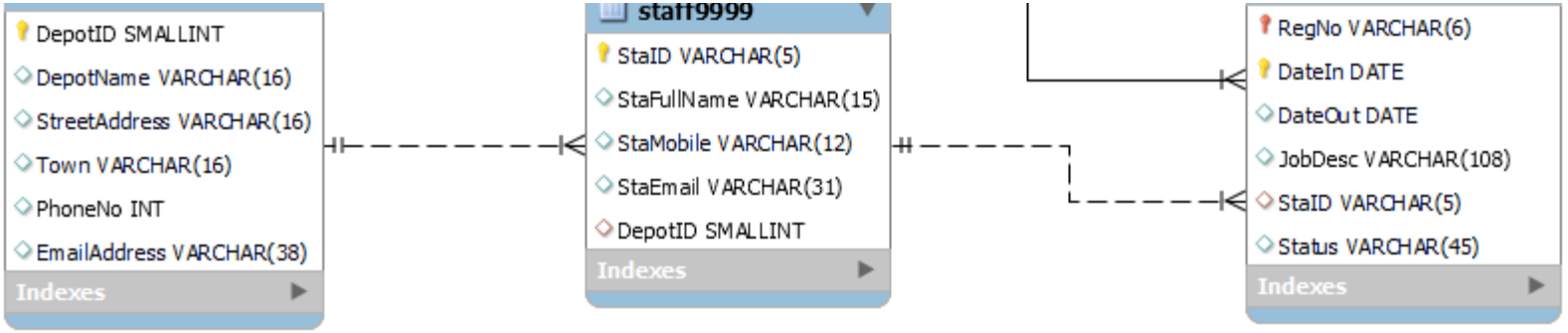
Question 2 – SQL (20 Marks)

WesternVic Car Rentals is a medium size car rental company that currently has 5 depots in the Western Suburbs of Melbourne and Western Victoria. Customers can hire a vehicle from any depot and if necessary vehicles are transferred between depots.

The company keeps track of vehicles, their transfers, their rentals and their maintenance. The rent per day is determined by car category. All staff members are associated with a specific depot where they either perform admin duties or conduct maintenance of vehicles.

The ER diagram below represents the current database model. Note it is *different* from what you were presented in Assignment 2. Since then the company kept working on improving their database model.





When answering questions, refer to the tables as you have renamed them, replacing 9999 with the last 4 digits of your student ID.

Do **not** use inline views / schema on read or views. Queries with inline views / schema on read or views will get 0 marks.

Create the queries to address the requirements below.

2A. List the number of cars in each car category. Show the category description, and the number of cars. Order the result by the number of cars from highest to lowest.

2B. List the total number of transfers per depot. Show the depotID, depot name, and transfer count. Order the result by depot name alphabetically.

2C. List client ID and Client Full Name for those who rented a car at least once in the previous year (i.e. in 2021). Order the results by client family name. Each client should appear on the list only once. The query should still be valid in future years.

2D. List vehicle registration, make, model, category description, dates in and out, and job description for all vehicles that had a maintenance job involving a windscreen. Order the results by make/model and registration.

Question 3 – Normalisation (10 Marks)

Parkville Vet Clinic needs to normalise records of appointments. An owner brings their pet to the appointment at a specified date/time with a specific vet and they pay a specified amount for the appointment.

PetType means cat, dog, guinea pig, etc.

Booking channel name can be phone, website, Facebook message, email

Below is the Appointment relation

Appointment (**vetID**, vetName (**PetID**, PetName, PetType, **OwnerMobile**, OwnerFullName, **channelID**, channelName, **date**, **time**))

One of the candidate keys for this relation is PetID, OwnerMobile, vetID, channelID, date, time

The following functional dependencies hold:

PetID –> PetName, PetType

OwnerMobile –> OwnerFullName

vetID –> vetName

channelID **–>** channelName

The task is to convert the non-normalised Appointment relation into 3rd normal form (3NF). You do not need to identify the intermediate steps (1NF, 2NF)

Important: you **must specify the legend** for **primary key (PK), foreign key (FK)** and **primary foreign key (PFK), e.g. PK - bold…**

Question 4 – Data Warehousing (13 Marks)

South East Medical Group (SEMG) operates several clinics in South Eastern suburbs of Melbourne. They want to analyse administration of different vaccines via their medical clinics. SEMG needs a data warehouse to report information about usage of vaccines over time. For each clinic they want to store the clinic details (clinic ID, clinic name, address, Manager's first and last name, phone number, email). Vaccine details stored are batch number, vial number, vaccine group (e.g. pneumococcal, polio, rabies, tetanus, etc), vaccine name (e.g. PCV13 (Prevnar13), PPSV23 (Pneumovax 23)), dose, retail price, purchase price). Batch number and vial number identify a vaccine dose uniquely. When patients get vaccinated, the clinics record patient ID, name, address, mobile, date of birth, and timestamp (date and time of vaccination). SEMG management wants to discover information about the number of vaccines sold, the revenue from vaccinations as per different clinics and per groups of patients as well as based on time periods (such as months, *seasons*, etc).

Q4. Draw a *star schema* to support the design of this data warehouse, showing the attributes in each table. You do not need to select data types. Clearly display the legend for Primary Key, Foreign Key and Primary Foreign Key.

Question 5 – Transactions (12 Marks)

5A

A designer brands retailer has the following tables in their database: Product, ProductInStock, Action. Note there are other tables which are not relevant to this question. The Action table keeps track of all sales and returns.

Product

ProdID	Brand	Description	RetailPrice	PurchasePrice
G43546	Gucci	Leather mid-heel pump	1080.00	450.00

ProductInStock

ProdID	QtyInStock
G43546	12

Action

ActionID	ActionDateTime	Action	ProdID	ProdQty	ProdCost
1008	21/01/2021 10:24	Purchase	G43546	2	2160.00
1026	23/01/2021 13:28	Return	G43546	-1	1080.00

Create an SQL transaction that records a sale of the product G43546 with the quantity being 2.

Hint: it is a better practice to use variables for calculations. Syntax example: SET @name='John';

(8 marks)

5B

Transactions must be **atomic** and **consistent**. Explain the term **atomic** and the term **consistent** using task 5a to illustrate your explanation.

(2+2=4 marks)

Question 6 – Distributed Databases and NoSQL databases (11 Marks)

'WeCollaborate" is a start-up offering researchers all over the world an opportunity to share their projects to find contributors and collaborators who are interested in joining their projects. Researchers need to keep their profiles up to date, including employment details, publications, funded projects they work on. Although the main language is English, the participants can opt to duplicate their profiles in French, Spanish, Mandarin. Other languages could be added in the future. Having those profiles in different languages will help improve target advertising from universities and research institutions to users based on the language as well as other profile features.

Developers decided that 'WeCollaborate" need a *distributed* database, and they are inclined to use a Graph database.

Answer the questions below providing **clear detailed explanation**. Make sure you are providing the explanation in the context of WeCollaborate, *not a generic list of advantages/ disadvantages/ features*.

6A

Why is a distributed database more suitable than a centralised database for WeCollaborate?

(5 marks)

6B

Provide a convincing explanation to WeCollaborate management as to why a Graph database would be most suitable for WeCollaborate. Draw a diagram illustrating how WeCollaborate data could be stored in a graph database. Come up with some sample data for your diagram.

(6 marks)

Question 7 –Database Security (7 Marks)

An SQL injection is a hacker's technique converting an SQL statement into malicious code. This is done by entering certain expressions instead of expected data into the form fields.

Imagine that a hacker knows that the company has a table Customers but as an ordinary user this hacker only has access to the table Products and can enter search words into a text box as in the figure below.

Consider the JSP code below

```
prodDetails = "SELECT prodName, description, price FROM products WHERE description LIKE ' %" +request.getParameter("input")+ " %' " ;
```

7A. What does this code do? Use an example of user input to illustrate your explanation.

(2 marks)

7B. What would a hacker put in the textbox to get all customer details from the table Customers. Explain how this input will work. Show the resulting SQL statement. Hint: use UNION

(3 marks)

7C. Discuss 2 measures that developers could take to prevent SQL injections. Make sure you provide details what each approach involves.

(2 marks)

Question 8 – Database Architecture (7 Marks)

Explain the **four states** that a *database buffer* in the buffer pool can be in. To illustrate your answer, use the example of executing an SQL Update statement using either products and sales database from question 5 or car rentals database from question 2.

Question 1 100 pts

Check again that your Word doc is named StuID_Last_FirstName_INFO90002Exam.docx

First upload your answer document to this section.

When done upload it to OneDrive **ExamSubmission** (https://unimelbcloud-my.sharepoint.com/:f:/g/personal/lindent_unimelb_edu_au/Eoy_XeiNGr5Lr1Vt6GNdk48B816eVMZnL0epQlvC5vzGQw)

Upload Choose a File

Saved at 13:02 Submit Quiz