



The University of Melbourne
School of Computing and Information Systems

INFO90002

Database Systems and Information Modelling

EXAM

End of Semester 1, 2021

Reading Time: 30 minutes

Writing Time: 120 minutes (2 Hours)

Authorised Materials:

While you are undertaking this assessment you are permitted to

- make use of any textbook, lecture slides (including soft copies)
- Any lecture notes and recommended texts
- You are free to use the course materials and your laptop/PC in this exam.

While you are undertaking this assessment you MUST NOT

- Directly copy material without attribution
- Plagiarise
- Collude with any other person in any form
- make use of any messaging or communication technology
- record, screenshot, stream, upload or in any known format duplicate this document
- record, screenshot, stream, upload or in any known format duplicate your solutions
- make use of any world wide web or internet based resources such as wikipedia, github, stackoverflow, google, Weichat or any known search engine / messaging 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

Instructions to Students:

- This exam is in 8 sections. Attempt all questions in all sections.
- Start a new question on a new page
- We recommend using pencil and paper/ e-pen tablet for modelling questions to save time
- The total for this exam is 100 marks representing 50% of your final assessment
- Attempt **all** questions which are of unequal marks value
- This exam is a timed assessment which must be completed within **150 minutes** of official commencement time
- Questions can be answered in any order
- Clearly identify the question you are answering
- Unnumbered attempts will not be assessed as we cannot determine which question you are answering.
- **PLEASE DO NOT USE RED font / ink colour**
- You must not communicate with other students whilst taking this exam, e.g. using messaging, chat rooms or email
-

IMPORTANT

- Your file upload must be a single PDF document before the elapsed time.
- No other document format will be assessed (e.g. Pages, doc, 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.
- If you choose to **upload photos**, ensure that they are clear, **compressed to < 1mb**, and importantly, check them once you have uploaded. It is your responsibility to ensure your file and any images has been submitted successfully.

The work you submit **must be based on your own knowledge and skills** and without the assistance of any other person. You **MUST NOT** directly copy work that you have not authored (e.g. slide notes, websites, other student's study notes)

Question 1. – ER Modelling

(25 Marks)

Nighty Night Stores

Nighty Night is a bedding specialist store that stocks beds, bedroom furniture, and mattresses. Nighty Night owns twelve warehouses and 140 stores located in Australia and New Zealand. For the goods stocked by Nighty Night we need to store an identifier, brand name, item type, bed/mattress size - e.g. single, king single, double, queen, king, and the primary material used in its construction (e.g. melamine, timber, steel, foam, wool), the country of manufacture, the cost price, profit margin and recommended retail price in Australian dollars.

About each warehouse we store its identifier (e.g. GOR2906), its physical address (structure is important), warehouse manager name, warehouse manager mobile number, warehouse phone and warehouse email.

Stores are located in various towns and suburbs in Australia and New Zealand. We store the trading name, address (structure is important), total floor space in square meters. For each store we need to store the manager's first and last name, email, and phone number.

For each supplier we store their trading name, registered name, address, and the name, email, and phone number of the sales manager who handles the Nighty Night account. We need to know what goods each supplier stocks.

Each Warehouse organises deliveries of goods from the warehouse to many stores. However a store receives its goods from the same warehouse. A delivery from that single warehouse can contain many different goods destined for many different stores. Nighty Night need to know the warehouse, the goods, and how many of each good is delivered to each store. Every delivery has an 8 digit consignment in hexadecimal format (0-9,A-F), the consignment date, the driver's name and driver's mobile phone number. Each warehouse delivery specifies the priority order in which the stores will receive their goods.

Q1. You have been asked to provide a logical Entity Relationship model in Crows foot notation of this case study.

(25 Marks)

Question 2. – SQL

(20 Marks)

St Clement's Grocery

St Clement's Grocery is an online only grocery store that specialises in fresh produce delivery (primarily fruit and vegetables) to local restaurants ('trade') in the Brunswick, Carlton, Parkville, Coburg and Essendon area. During the Covid19 pandemic they also branched out into contactless delivery for residential ('retail') customers.

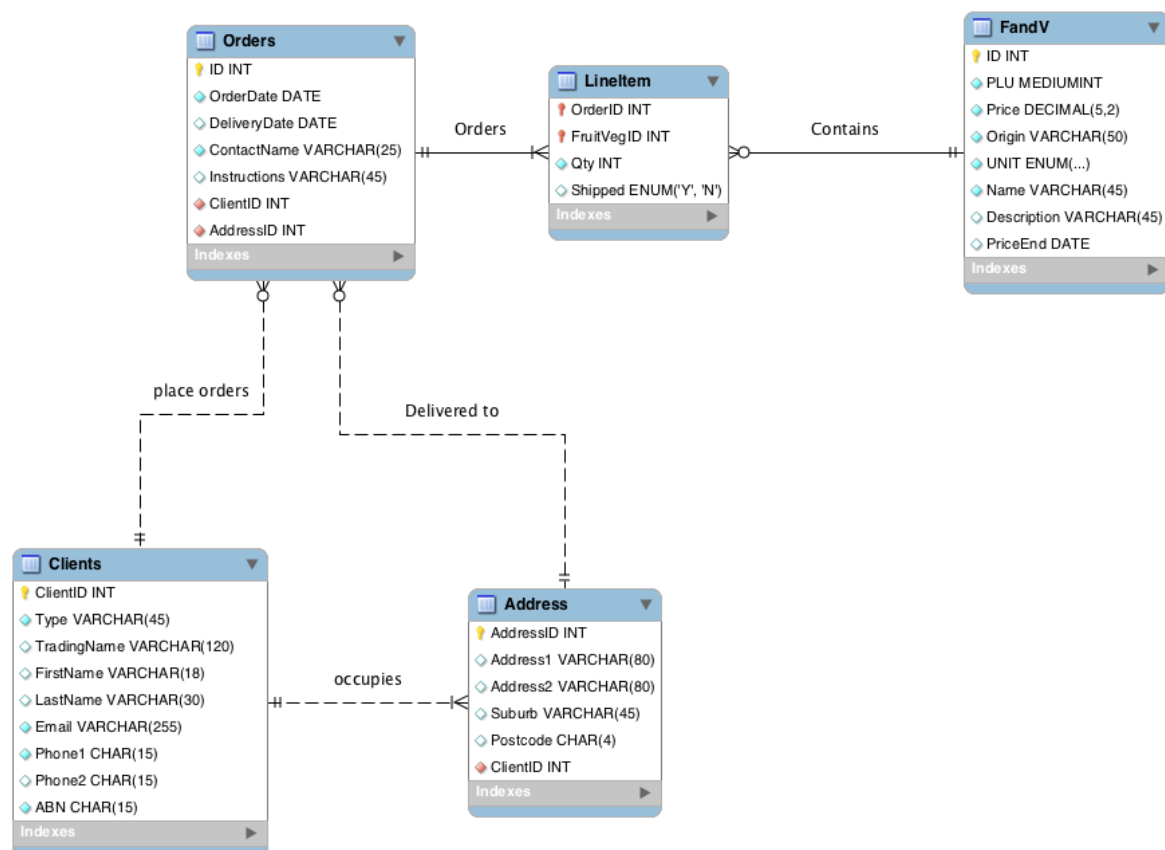


Figure 1. The Entity Relationship Model for St Clement's Grocery

Formatting requirements

For each question, present an answer in the following format:

- Show the question number and question in **black** text.
- Show your answer (**a single SQL statement**) (**DO NOT use a screen shot**)
- Ensure you use **straight quotes** ' ' NOT curly quotes for String and Date types
- Do not use schema on read, views, or inline views as part of your answer

Example:

Q.XX List all Addresses on Sydney Road

```
SELECT Address1, Address2, Suburb, Postcode
FROM Address
WHERE address1 like '%Sydney road%'
OR address2 like '%Sydney road%';
```

Answer each question below with a single SQL statement. **Do not use inline views, or views of any type.** All queries must be schema on write. Marks will be deducted for poorly formatted and excessively long and complex SQL.

Q2A. List the total number of trade orders per postcode. List the postcode and order count. Order the result by postcode alphanumerically.

(2 marks)

Q2B List the total quantity ordered of Oranges, Limes, Lemons, Clemintines (sic) and Grapefruits. List the name, quantity and unit.

(3 marks)

Q2C. In the FandV table the fruit Clementine is spelled incorrectly (as Clemintine). Write **one** SQL UPDATE statement to correct the spelling of Clementine for all occurrences in the FandV table

(3 marks)

Q2D. List all months where the total invoice amount exceeds \$100000. List the month name, year, and total invoice amount.

(3 marks)

Q2E. Currently the FandV table in the St Clement's database is not normalised. *Describe* how you would fix the FandV table so that the data is normalised, and yet still retain previous price history of produce?

HINT: There is no requirement to write any SQL or normalisation for this question

(3 marks)

Q2F. St Clement's has decided to add St David's Dairy products to its current price list. Write **one** SQL INSERT statement to add these rows to the St Clement's database

ID	PLU	PRICE	Origin	Unit	Name
1201	5802	6.49	Fitzroy	ea	St David Dairy Full Cream Milk 2L
1202	5801	3.25	Fitzroy	ea	St David Dairy Full Crem Milk 1L
1203	5902	6.40	Fitzroy	ea	St David Dairy Low Fat Milk 2L
1204	5901	3.25	Fitzroy	ea	St David Dairy Low Fat Milk 1L

Table 1. St David's Dairy produce information

(6 marks)

Question 3. – Normalisation

(15 Marks)

Can't Stop the Music is a streaming company that streams music up until the year 1990. It specialises in punk and disco music. While it offers different Plans of plans (freemium with advertising, monthly or annual subscription) it must keep a record of every time a user streams a song to pay royalties to the royalty owner. You have been asked to normalise their Royalties relation.

Consider the following relation:

Royalties(**userID**, username, email, plan, cost, (**playID**, **song**, artist, album, genre, playdate))

userID, playID, song is the key attribute for the Royalties relation

The following functional dependencies hold:

userID, playID, song → playdate

userID → username, email

plan → cost

song → artist, album, genre

Q3. Please convert the non-normalised Royalties relation into 3rd normal form (3NF). Please identify the functional dependencies and intermediate steps (1NF, 2NF) .

HINT: Clearly identify the legend which identifies primary key, foreign key and primary foreign key

(15 Marks)

Question 4. – Data Warehousing (10 marks)

Australian Consolidated Publishing

Q4A. Australian Consolidated Publishing (ACP) wants to track information about the selling of its magazines via bookshops and newsagents ("shops") located throughout Australia. ACP needs a data warehouse to report information about sales of books over time. You need to store the shop (shop ID, Manager's first name, Manager's last name, phone number, address), and magazine (ISBN, title, description, type, category, retail price, wholesale price). Managers want to discover information about the number of magazines sold, the revenue earned, and profit made. The information needs to be accessible by shop, by book and for different times (day, week, month, quarter and year).

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

(8 marks)

Q4B. *Describe* how you would change the schema if ACP would like to see the same measurements, broken down according to a bookshop's city, state & country?

(2 mark)

Question 5. – Transactions (8 marks)

Q5A. Explain how user defined transactions work. Use at least 3 SQL statements in your explanation.

(5 marks)

Q5B. What are the possible purposes of a transaction log?

(3 marks)

Question 6. – No SQL (8 marks)

Q6A. Describe a "document database", explaining how it differs from a relational database

(3 marks)

Q6B. Describe a "graph database", explaining how it differs from a relational database.

(5 marks)

Question 7. Applications

(4 marks)

Q7A. Describe the advantages of JSON over XML for dynamic web pages.

(2 marks)

Q7B. Explain what is meant by distributed processing in 2-tier client server architecture.

(2 marks)

Question 8. - Storage & Indexes

(10 marks)

This is the relation for the OrderItem table

OrderItem (**OrderID**, *ItemID*, Quantity)

Where PK = **Bold** FK = *Italic* PFK = **Bold + Italic**

The file containing an OrderItem table is sorted by its composite primary key in the following order :
orderID itemID (Table 2).

ORDERID	ITEMID	Quantity
101	05	3
101	09	9
101	11	3
102	23	3
103	05	3
103	11	7
103	23	1

Table 2: File storage for the OrderItem table

Q8A. What would be an example of the column order for an unclustered index? List the columns

(2 marks)

Q8B. If the database administrator created an index on the OrderItem table using only the OrderID column - what would be the index classification? Justify your answer.

(3 marks)

Q8C. Explain what a sparse index is, and how it is different to a dense index

(3 marks)

Q8D. Given the hash bucket number $n=4$ sort the following values into their appropriate hash bucket. Be sure to label the bucket number.

Buckets $n=4$

{1,3,7,8,13,14,15,19,23,34,52,79, 88,101}

(2 marks)

GOOD LUCK!

END OF EXAM