COMP1350 Exam
Introduction to Database Design and Development

Exam Instructions:

1. You should be able to complete the exam within 3 hours. However, for the exam window, you will have up to 6 hours to submit the PDF file of your exam answers to Turnitin. The exam window will automatically close after the 6 hours.
2. This exam is graded out of 100 marks and is worth 50%. This is not a hurdle exam.
3. It is an individual take-home exam which means you will have access to resources of your choice to answer the questions. However, you would not need access to any resources other than what has been provided to you on iLearn.
4. There are three sections. Attempt ALL questions in the Three (3) sections. The first section (Section-A) is about database modelling, the second section (Section-B) is about database development, and the third section has some questions about database concepts (Section-C).
5. Please use the answer template file provided for you in iLearn to write your answers for all three sections. DO NOT change the order of the questions in the template provided.
6. Make sure your answers are in the same order as the questions listed in this question paper.
7. Convert your answer file to a PDF document before submitting it to the exam Turnitin link on iLearn.

Section-A: Question 1 in Section A will require you to use lucid charts/draw.io to draw diagrams to present your answer. Please ensure you sign in using your student email account to login and work on these diagrams. To add these images in your document, you will have to export your diagrams as Images and add them to your answer document. Illegible diagrams will not be considered and no second attempts will be provided. It is your responsibility to ensure that the diagrams are legible when zoomed in. You may choose to hand draw and upload the image in the word document. Question 2,3 can be typed-up.

Section-B: An sql script for Section B is provided on iLearn; You should be able to run the script on MySQL Workbench to double-check your queries. You may choose to use the tables listed under Section B to write your queries. Code execution is not compulsory. But highly recommended.

Section-C: Based on your understanding of the concepts, you will have to answer the questions. You do not have to cite/use any external sources. Please refrain from defining terms as you would be tested based on your answer.

Case Study

Gaz Oakley is the founder of *Avant-Garde*. His business is based in Sydney to provide top quality organic foods delivered to customers in and around Sydney area. He is passionate about providing maximum benefits for the farmers as well as other local businesses who sell their produce/products through his business. As his business is growing rapidly, Gaz is interested in developing a IT system that requires a relational database. Below are his statements about the business and requirements of the system.

Businesses & Supplies: Every business is required to register with their ABN, this acts as a unique identifier. Other details of the registered businesses include the name of the business and its address. There are two types of businesses that may choose to register with Avant-Garde: Farmers and Local shops. Farmers have their farm location stored, whereas the local retail shops have their establishment date and each of their suburban locations recorded. Note, some of the farmers may also register as a local shop as well.

Every product has a unique identifier. Other details of the products include the name of the product and its description which are stored. There are two types of products that are advertised through Avant-Garde: they are perishable goods and non-perishable goods. Perishable goods (like vegetables/fruits) are supplied by the farmers. A farmer may supply multiple perishable goods and the same perishable goods may also be supplied by multiple different farmers. Months during which the perishable goods are supplied are stored for every perishable good that is supplied. Details of non-perishable goods such as the shelf life is stored. Non-perishable goods could be supplied by either of the businesses.

Customers: Every customer needs to be uniquely identified in the system. Other details which the system needs to store about a customer include their name, address, and phone numbers. Some customers may also choose to become loyalty members. In this case, the sign-up date and membership status (like active/suspended) is recorded.

Product Listing: Every week a product listing gets created (Think of this like the product list that the customers can see to place an order). Every product listing has a unique identifier. Also, the date when the product listing is created is also stored.
A product listing has multiple products listed in them and a product could be in multiple listings. Product price and weights (quantity) for each of the product is stored for every product displayed on the listing.

Orders: Every order that is placed has a unique identifier. Other details such as the date of the order and the customer who placed the order are stored. Orders can have multiple products listed in the listing and a product could be in multiple orders. For each of the products in the order, the quantity is stored. Multiple payments can be made for an order. Hence, for every order, a payment number is generated (Order 1 can have Payment 1, Payment 2, …, and Order 2 can also have Payment 1, Payment 2, ...). The date of the payment and payment amount are also stored.

Section-A (40 marks)

1. Based on the business rules and data needs outlined in the case study, construct an **EER** model. This model should include entities, attributes, primary identifiers, and the relationships among entities with cardinalities and constraints. If any inheritance relationships are used, they must show total/partial and overlap/disjoint constraints. You must **explain** your choice of total/partial and overlap/disjoint constraints for each of the relationship. You may need to introduce additional entities also to ensure every data requirement is captured. Data types for attributes are **not required**. State any assumptions if you think something is not clear or has not been addressed.

   Also, to demonstrate your understanding, explain the following concepts with an example from your diagram:
   a. Strong and weak entities
   b. An associative entity
   c. A composite attribute
   d. A multi-valued attribute
   e. A total / partial constraint

   (25 marks)

2. Transform your diagram (only the entities, attributes and relationships **relating to the case study sections** of **Product Listing** and **Orders**) into logical tables using the steps outlined in lectures. For step 8, use the 8a approach. Show what is done in each of the steps with inclusion of primary and foreign keys and a final list of tables.

   (9 marks)

3. In your own words, using the table below and your own examples, explain the three different types of anomalies.

   (6 marks)

FarmerID	ProduceID	FarmName	FarmLocation	ProduceName	Price per kilo
F1	P1	Claire's Farm	NSW	Tomato	$7.99
F1	P2	Claire's Farm	NSW	Zucchini	S2.00
F2	P1	Star World	VIC	Tomato	$6.99
F2	P3	Star World	VIC	Spinach	$20.75
F2	P4	Star World	VIC	Orange	$1.99
F1	P4	Claire's Farm	NSW	Orange	$3.99
F3	P1	Where Pigs Fly	VIC	Tomato	$5.50

Section-B (35 marks)

A '.sql' file has been provided for you to download so you can write and test your sql queries for Section-B. Using the tables and data descriptions provided below, answer the questions below. Put your final answers into the answer document you use to answer Section-A and Section-C. You do not have to show the execution results of your queries. Each question in this section is worth 5 marks.

Table description for: Farmer

Column-Name	Comments/Description
FarmerID [PK]	A unique identifier for the farmer
FarmerName	The name of the farmer
FarmName	The name of the farmer
FarmLocation	The area and state of the business

Table description for: ProduceType

Column-Name	Comments/Description
ProdTypeID [PK]	A unique identifier for the product type
ProdTypeName	The name of the product type

Table description for: Produce

Column-Name	Comments/Description
ProduceID [PK]	A unique identifier for the product type
ProduceName	The name of the product type
ProdTypeID [FK]	FK referencing ProdTypeID in ProduceType Table

Table description for: Supplies

Column-Name	Comments/Description
ProduceID [PK,FK]	PK, FK referencing ProduceID in Produce Table
FarmerID [PK,FK]	PK, FK referencing FarmerID in FarmerTable
Weight	Weight of the produce for which the cost is computed
Metric	The metric in which the weight is calculated
PricePerWeight	Cost per weight of the product supplied by the farmer

Please note you are not allowed to use 'Natural Join' keyword for any of the questions

1. Write a query that uses wildcard operators to print the names of the Farmers, Farms along with the location for those farms from NSW.

2. Write a query to print the names of the Produce and the name of their types.

3. How many types of produce are each of the farmers producing? Your result should include the names of farmers and the total number of produces. Also, include farmers who do not supply any produce as well. Sort the results by the highest number of produce.

4. Extend question 3 above to include only the farmers who have the highest number of produces. (Note - You cannot use LIMIT function.)

5. What is the cost of each vegetable supplied by a farmer with first name 'Alex'? Display the result with a $ sign before the cost.

6. Using a subquery, print the names of the farmers who produce Fruits.

7. Using a join (instead of a subquery), rewrite question 6.

Section- C (25 marks)

1. List three different applications that can be created for *Avant-Garde* where the database that is created for the case study can be utilised. For each of these applications, will thin/thick client be better? Justify your choice.

   (5 marks)

2. How can *Avant-Garde* use a data warehouse? Using a star-schema, justify any one of your scenarios.

   (10 marks)

3. Explain in your own words, why Twitter might not use a relational database to store tweet data. What is an alternative solution for Twitter to store their tweet data?

   (10 marks)

That's it! End of the exam questions!

Be sure to double check you have attempted all the questions and included all of you answers (in question order) for Section-A, Section-B, and Section-C in the same answer document.

Remember to convert your document to a PDF file before submitting it to the Turnitin link in the exam section on iLearn.