UNIVERSITY OF FORT HARE

IFS215 & IFS215E

DATABASES

JULY 2023

Time: 3 Hours **Marks: 120**

This paper consists of SIX (6) pages including the cover page

Internal Examiner
Prof R Piderit
Mr V Funda
Dr E Chindenga

Internal Moderator
Mr D Boucher

GENERAL INSTRUCTIONS TO CANDIDATES

1. This paper consists of **four (4)** *SECTIONS.*
   a. Answer **TWO** questions in SECTION A. Only the first two questions will be marked.
   b. Answer **ALL** the questions in SECTION B.
   c. Answer **ALL** the questions in SECTION C.
   d. Answer **TWO** questions in SECTION D. ***Answer in a new answer book.***

2. Questions can be answered in any order but must be numbered correctly. Any answers you do not want marked must be clearly cancelled.

3. Time management is very important. The value of the mark for each question should be used as a rough guide to the amount of time allocated to answer the question. (120 marks in 180 minutes)

4. It is in the candidate's interest to write neatly.

5. At the end of the examination place all answer books inside the first answer book used.

6. Pencil may be used for the modelling questions.

SECTION A [ANSWER **TWO** QUESTIONS FROM THIS SECTION] = 30 MARKS

Question 1 **[15 marks]**

a. Define data, information and knowledge by highlighting the difference between them. (5)

b. What are three advantages and two disadvantages of database management systems. *[Include a description of each to show you understand what they mean.]* (5)

c. Consider a company that sells products online. One important business rule for this company might be that *a customer can place multiple orders, but each order can only have one customer.* Describe and draw this relationship using **UML notation.** (5)

AND/OR

Question 2 **[15 marks]**

a. Name and describe the five methods of classifying a database. (5)

b. What does it mean to say that a database displays both *entity integrity* and *referential integrity* and why are *entity integrity* and *referential integrity* important in a database? (5)

c. Consider the STUDENT and CLASS tables below and provide the relational schema. *[Note you will need to identify a suitable primary key for each table.]* (5)

Students

Student ID	StudFirst Name	StudLast Name	StudStreet Address	StudCity	StudState	StudZipcode
60001	Zachary	Erlich	1204 Bryant Road	Seattle	WA	98125
60002	Susan	McLain	101 C Street, Apt. 32	Redmond	WA	98052
60003	Joe	Rosales	201 Cherry Lane SE	Redmond	WA	98073
60004	Diana	Barlet	4141 Lake City Way	Woodinville	WA	98072
60005	Tom	Wickerath	2100 Mirenda Avenue	Bellevue	WA	98005

Classes

Class ID	Class Name	Class Category	Credits	Instructor ID	Classroom
900001	Advanced Calculus	Math	5	220087	2201
900002	Advanced Music Theory	Music	3	220039	7012
900003	American History	History	5	220146	3305
900004	Computers in Business	Computer Science	2	220387	5115
900005	Computers in Society	Computer Science	2	220387	5117
900006	Introduction to Biology	Biology	5	220498	3112
900007	Introduction to Database Design	Computer Science	5	220516	5105
900008	Introduction to Physics	Physics	4	220087	2205
900009	Introduction to Political Science	Political Science	5	220337	3308

AND/OR

[TURN OVER THE PAGE FOR THE NEXT QUESTION]

Question 3 **[15 marks]**

a. Identify and describe five characteristics of data quality. (5)

b. Describe (with examples) the difference between a *disjointed*, *union* and *null set*. (5)

Use the following to answer (c) and (d):

I = {1;2;3;4;5;6;7;8;9;10}
II = {3;4;5;6}
III = {1;2;3}
IV = {11;12;13}

c. Do II and IV have any common members? Write your answer in the proper notation. (3)

d. If I want to join II and III (make them a union), how would I show this with proper notation? (2)

SECTION B [ANSWER **ALL** THE QUESTIONS FROM THIS SECTION] = 30 MARKS

Question 4 **[10 marks]**

Data modelling is the first step in the database design journey, serving as a bridge between real-world objects and the database that resides in the computer.

a. What makes data modelling important? (5)

b. What is a business rule, and what is its purpose in data modelling? (5)

Question 5 **[20 marks]**

Scenario: A music streaming service *(such as Spotify)*

Grace has been asked by her boss to draw the Entity Relationship Diagram (ERD) to set up a new database for a music streaming service. She has managed to find some of the information she needs below from interviewing a few people in the company. But she knows she is missing information about Primary and Foreign Keys. She is worried as her boss said this is a big client and her job is "on the line".

Grace has collected the following information -

The music streaming service has multiple users who can create playlists and listen to songs. Each song has a unique title, artist, and album, and belongs to a specific genre.

Grace has collected the following entity information:

- User (with attributes: username and email)
- Playlist (with attributes: name and description)
- Song (with attributes: title, artist, album, and genre)

She also thinks she knows what the relationships are between these entities:

- A user can create multiple playlists
- A playlist can contain multiple songs, and a song can belong to multiple playlists
- A song belongs to a single playlist, and a playlist can have multiple songs

Help Grace keep her job by assisting her with creating an ERD using UML notation and indicate all primary and foreign keys.

[TURN OVER THE PAGE FOR NEXT SECTION]

SECTION C [ANSWER **ALL** OF THE QUESTIONS FROM THIS SECTION] = 30 MARKS

Question 6 **[10 marks]**

Having good relational database software is not enough to avoid the data redundancy.

a. Why is normalisation important for database design? (3)

b. Describe the difference between 1NF, 2NF and 3NF. (3)

c. Describe the process/steps to convert a table to 1NF. (4)

Question 7 **[20 marks]**

Attribute Name	Sample Value	Sample Value	Sample Value	Sample Value	Sample Value
INV_NUM	211347	211347	211347	211348	211349
PROD_NUM	AA-E3422QW	QD-300932X	RU-995748G	AA-E3422QW	GH-778345P
SALE_DATE	15-Jan-2019	15-Jan-2019	15-Jan-2019	15-Jan-2019	16-Jan-2019
PROD_LABEL	Rotary sander	0.25-in. drill bit	Band saw	Rotary sander	Power drill
VEND_CODE	211	211	309	211	157
VEND_NAME	NeverFail, Inc.	NeverFail, Inc.	BeGood, Inc.	NeverFail, Inc.	ToughGo, Inc.
QUANT_SOLD	1	8	1	2	1
PROD_PRICE	€34.46	€2.73	€31.59	€34.46	€69.32

Using the INVOICE table structure shown above:

a. Write the relational schema. (4)

b. Draw its dependency diagram, and identify all dependencies (including all partial and transitive dependencies). You can assume that the table does not contain repeating groups and that an invoice number references more than one product. (Hint: This table uses a composite primary key.) (6)

c. Using the answer to (b), remove all partial dependencies, write the relational schema and draw the new dependency diagrams. (6)

d. Identify the normal forms for each table structure you created in (c). (4)

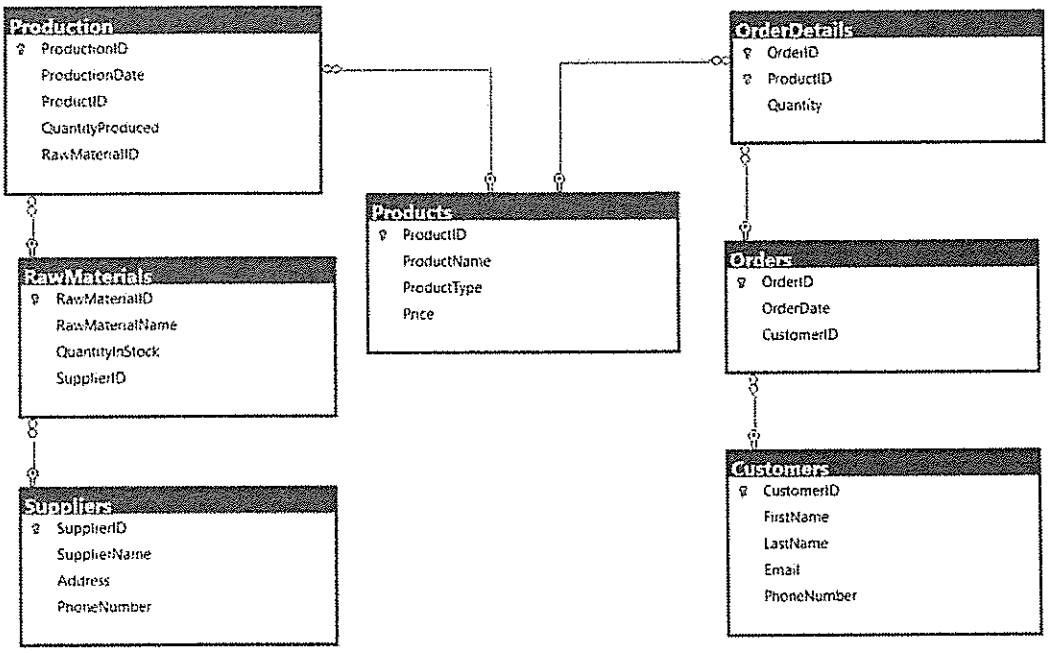
[TURN OVER THE PAGE FOR NEXT SECTION]

ANSWER THIS SECTION IN A NEW BOOK

SECTION D [ANSWER **TWO** QUESTIONS FROM THIS SECTION] = 30 MARKS

Question 8 **[Marks 15]**

Study the Database schema below and answer the following questions.



a. Write a SQL query to create the Orders table. (3)

b. Write a SQL query to retrieve the names of all products and their prices from the Products table. (2)

c. Write a SQL query to retrieve the names and email addresses of all customers who have placed an order. (3)

d. Write a SQL query to add a new column called "Colour" to the Products table to store the colour of each product; (2)

e. What is a foreign key constraint, explain by means of an example from the given schema, how is the foreign key constraint used to enforce referential integrity. (3)

f. Write a SQL query to delete the Suppliers table. (2)

AND/OR

Question 9 **[15 marks]**

a. How does the UNION operator differ from the JOIN operator in SQL, and when should you use the UNION operator? (2)

b. State and explain two benefits of Stored Procedures in SQL. (2 + 2)

c. What does the acronym DBLC mean, and what does a DBLC portray? (2)

d. Explain the concept of a deadlock in a database system? (2)

e. How can backup and recovery procedures contribute to database security? (3)

f. Explain the concept of Transaction Isolation in a multi-user database environment. (2)

[TURN OVER THE PAGE FOR NEXT QUESTION]

Question 10 **[Marks 15]**

Study the query below and answer the following questions.

```
SELECT DISTINCT CUS_CODE, CUS_LNAME, CUS_FNAME
FROM CUSTOMER JOIN INVOICE USING (CUS_CODE)
JOIN LINE USING (INV_NUMBER)
JOIN PRODUCT USING (P_CODE)
WHERE P_CODE IN (SELECT P_CODE FROM PRODUCT
WHERE P_DESCRIPT LIKE '%HAMMER%' OR P_DESCRIPT LIKE '%SAW%');
```

a. Identify the table(s) that will be joined in this query. (4)

b. Explain the purpose of the DISTINCT function in the query. (3)

c. Explain the purpose of the LIKE special operator in the query. (2)

d. Identify and list all the attributes from the CUSTOMER table that will be used in this query? (3)

e. Write out the sub-query used in the SQL query. (3)

[END OF PAPER]