COS10082 APPLIED ANALYTICS IN BUSINESS

Assignment 2

Total Marks: 100 (25%)

Due date & Time: Friday (22-November-2024) by 11:59PM

This assignment is to be completed **INDIVIDUALLY**.

Instructions:

- Complete Part I of this assignment using SQL Developer/Oracle Live SQL and:
  - Save Question 1, the ERD as an image file named it as **StudID_ERD.jpg**
  - Save your answers to Questions 2-12 as a **SQL** file (name the SQL file: **StudID_SQL.sql**). Please indicate the question numbers accordingly as comment in the SQL script file.
- Complete Part II of this assignment and save your star schema with design justifications as a **PDF** (name the PDF file: **StudID_Star.pdf**).
- On Canvas, upload the above three files in a zipped folder (name the zipped folder: **StudID_COS10082_Assignment 2**)

Notes:

- Student MUST ensure that the script is **error-free** (except for some questions where errors are expected).
- In this assignment, you are **NOT** allowed to use **FULL JOIN**.
- Student MUST ensure that the submitted work is their own.
- **Any attempt to copy & paste contents will result in ZERO mark given to all parties involved regardless of who copied from whom.**

Part I: SQL [60 Marks]

Digital Dreamer is a magazine retailer based in Malaysia that publishes and distributes magazines across different states in Malaysia. The company has several warehouses across three states in Malaysia: Kuala Lumpur, Sabah and Sarawak, where each of the warehouses deals with the distribution of magazines to the customers. The company has kept all of the data to keep track of their warehouse, magazine, customer and sales details in the forms of excel worksheet. The Chief Information Officer has appointed you to design a database to cater for the growth of the business.

The following are the excel worksheets with its data:

Worksheet Name: Warehouse

WID	PHONE	ADDRESS	STATECODE	STATENAME
w_001	082-991246	Jalan Astana, 93050	MY_13	SARAWAK
w_002	03-27206611	Damansara Heights, 52200	MY_14	W.P. Kuala Lumpur
w_003	088-521916	Jalan Tuaran Batu, 88450	MY_12	SABAH
w_004	03-4210886	Taman Mestika, 56100	MY_14	W.P. Kuala Lumpur
w_005	082-421433	Jalan Pending, 93450	MY_13	SARAWAK

Worksheet Name: Customer

CUSTID	CUSTNAME	EMAIL	STATECODE	STATENAME
c_01	John Tan	john@gmail.com	MY_14	W.P. Kuala Lumpur
c_02	Mary Lai	mlai@gmail.com	MY_12	SABAH
c_03	Jane Bennette	janeb@gmail.com	MY_13	SARAWAK
c_04	Mohd Nazim	mdm@gmail.com	MY_14	W.P. Kuala Lumpur
c_05	Serena Choo	schoo@gmail.com	MY_13	SARAWAK

Worksheet Name: Magazine

**** One magazine may have multiple authors**

MAGAZINEID	MAGTITLE	YEAR	PRICE	WID	AUTHORNAME
B011	Business Insight	2021	33	w_001	Paul Bennette, Johnson Tay
N022	Nature Guide	2017	57	w_003	Nazim Karim
C059	IT Gallery	2020	25	w_005	Terry Peri, Mark Tay, Terrence Tim
F124	Food Arts	2021	19	w_003	Natasya Kim
AP866	Animal Planet	2018	50	w_004	Jason Cordon
H557	Home Plant	2021	72	w_002	Tan Chee Hung

Worksheet Name: ShoppingBasket

BASKETID	MAGAZINEID	CUSTID	QUANTITY	SELLINGPRICE	PURCHASEDATE
sb11	B011	c_02	3	52	2/5/2021
sb22	C059	c_03	2	36	1/1/2021
sb33	AP866	c_04	1	85	7/2/2021
sb44	F124	c_02	2	25	3/7/2021
sb55	H557	c_05	4	105	23/4/2021
sb66	C059	c_01	3	36	2/3/2021
sb77	F124	c_01	1	25	2/3/2021
sb88	C059	c_04	2	36	24/4/2021
sb99	B011	c_01	4	52	30/6/2021

Complete the following questions:

1. **[6 Marks]** Design an ERD for the database using UML notation, ensure that all entities, relationships (including names) and attributes with the associated datatypes are included:
    a. All the tables identified are normalized
    b. Identify primary and foreign keys
    c. Include cardinality and show using UML notation.

    The ERD should be created using software tool such as Lucidchart or Draw.io and submitted as an image file. Hand-drawn diagrams will not be accepted.

2. **[6 Marks]** Implement the above database and tables using SQL, create all tables with the appropriate data types (refer to tables when deciding on the data types, and make sure the order is correct).

3. **[4 Marks]** Insert the data in the above tables to each table created accordingly.

4. **[2 Marks]** Display all columns in the Magazine table for all magazine that are published after the year of 2018. Sort the result in ascending order by year.

5. **[2.5 Marks]** Using a **sub-query**, display the warehouse state name of "Sarawak" based on any valid condition in the Warehouse table.
   (Note: Your output must be the same as the given output)

STATENAME
SARAWAK

6. **[3 Marks]** For each of the magazines in the ShoppingBasket table, display the magazine ID, purchase date, and the total sales. You need to sort the output by the purchase date in descending order.
   (Note: Your output must be the same as the given output)

Magazine ID	Purchase Date	Total Sales
F124	03/07/2021	50
B011	30/06/2021	208
B011	02/05/2021	156
C059	24/04/2021	72
H557	23/04/2021	420
C059	02/03/2021	108
F124	02/03/2021	25
AP866	07/02/2021	85
C059	01/01/2021	72

7. **[3 Marks]** For each of the magazine, display the magazine id and its total number of authors. You need to sort the output by the total number of authors in descending order.
   (Note: Your output must be the same as the given output)

Magazine ID	Number of Authors
C059	3
B011	2
H557	1
N022	1
AP866	1
F124	1

8. **[10.5 Marks]** Write the SQL statement that would display the following output

State Name	Number of Warehouse	Number of Customer
W.P. Kuala Lumpur	2	2
SABAH	1	1
SARAWAK	2	2

9. **[5 Marks]** Write the SQL statement that would display the magazine name and calculate the total profit and the percentage of total profit based on the magazine sold.

   (Note 1: The percentage should have maximum of 2 decimal points)

   (Note 2: Your output's heading and result should be as follow:)

   Hint on Formula needed:

   Total Profit = quantity * (sellingprice – price)

MAGTITLE	Total Profit (RM)	Profit Percentage
Home Plant	132	45.83
Business Insight	133	57.58
Animal Planet	35	70
IT Gallery	77	44
Food Arts	18	31.58

10. **[7.5 Marks]** Write the SQL statement that would display the total sales for each of the magazine based on the state that the customer is in. The output should be ordered from highest to lowest based on the total sales for each state.

    (Note 1: Sales for each magazine refers to the selling price * quantity sold)

    (Note 2: Your output's heading and result should be as follow:)

State Name	Magazine Name	TOTALSALE
SABAH	Business Insight	156
SABAH	Food Arts	50
SARAWAK	Home Plant	420
SARAWAK	IT Gallery	72
W.P. Kuala Lumpur	Business Insight	208
W.P. Kuala Lumpur	IT Gallery	108
W.P. Kuala Lumpur	Animal Planet	85
W.P. Kuala Lumpur	IT Gallery	72
W.P. Kuala Lumpur	Food Arts	25

11. **[5.5 Marks]** Using **Rollup**, write the SQL statement that would display the magazine and the total quantity sold as shown in the following.

    (Note 1: Your output's heading and result should be as follow:)

Magazine	Total Quantity Sold
Animal Planet	1
Business Insight	7
Food Arts	3
Home Plant	4
IT Gallery	7
All magazines	22

12. **[2 Marks]** Write any SQL statement that would produce an error on unique constraints violations such as the error given below:

```
Error report -
ORA-00001: unique constraint (SYSTEM.SYS_C007303) violated
```

13. **[3 Marks]** Write the SQL statement to alter the ShoppingBasket table to enforce having a sellingprice value for each row.

Part II: Data Warehouse Design [40 Marks]

You are hired by Digital Dreamer to help them to bring their data from the database to a single data warehouse where the management could mainly query the data warehouse looking for data such as total sales in volume and revenue for a specific region for a particular magazine for a certain period this year compared to the same period last year. You may design the data warehouse based on the ERD in Q1:

- **Your design:**
  - Must follow the **star schema**
  - Must have **one fact table** and **multiple dimension tables**
  - Must demonstrate the use of **surrogate key**
  - Provide an explanation with design justification in 500 words (±100words). ***Note:*** Marks will be deducted if the word count is not met.
  - Provide a discussion on scalability and future consideration in 500 words (±100words). ***Note:*** Marks will be deducted if the word count is not met.
- **Marking will be as follows**:

Criteria	Mark
Understanding of requirements	5
Star Schema Design	15
Justification of Design Choices	10
Scalability and Future Consideration	10

--End of document--