HW-5 Data Warehouse

Overview

- Assignment points: 10 points or **10%** toward your final grade.
- Due date: **11/14/2023, 11:59 pm**.
- Late submission: grace period of 2 days with 20% late penalty. After that time, assignment will NOT be accepted.
- <mark>Submission: submit **SQL scripts** and **output (result set)** in **Canvas in pdf or word doc.**</mark>
- <mark>Make sure to mark your answers with the question number</mark>.
- This is open book exam, and any kind of resource materials are allowed.

This homework will give you hands-on practice in working with a dimensional model /star schema data warehouse. You will be using MySQL workbench to write SQL for this assignment, the tool downloaded during the class. If you haven't downloaded yet, you will need to download and install MySQL workbench on your computer. Downloads are available at: https://dev.mysql.com/downloads/workbench/

Choose the download file that matches your computer's OS and version. Installation files are also available in Canvas

Creating the Database

You first need to download the "<mark>HW-5 DW.sql</mark>" script file from Canvas and execute it. It will create a database with tables and then load data into tables. The script runs fine as-is without any modification. The script begins with a command to DROP the table before creating it. This will allow you to run scripts over and over, if needed.

Before executing scripts, make sure your MySQL instance is running and Safe Updates is unchecked (Edit - SQL Editor - Uncheck Safe Updates).

There is an ERD **HW-5 DW ERD.pdf** for this database in the Canvas. You should download and keep it handy while doing the homework. It will be helpful to have table and column names, and their relationship in front of you when writing SQL queries.

After running scripts, you should see <mark>HW_5_DW</mark> database with following 6 tables and row counts.

HW-5 Data Warehouse

TABLE_NAME	TABLE_ROWS
dim_customer	5
dim_date	91
dim_product	5
dim_salesperson	6
dim_store	3
fact_productsales	297

Notes:

- Submit **SQL scripts** and **output (result set)** in **Canvas in pdf or word doc.**
- Make sure to mark your answers with the question number.
- Display **ONLY** columns / attributes that are asked to show in questions.
- Make sure to use table / column aliases, where applicable.
- Use a meaningful derived column names as asked in question, e.g., SalesPrice * Quantity = **Revenue** so your result column should display as **Revenue**.
- Make sure to mark your answers with the question number.
- **All Questions are equally weighted**.

Formulas:
Revenue = SalesPrice * Quantity
Cost = ProductCost * Quantity
Profit = Revenue – Cost

Note: If question is asking for Total then it is asking to use aggregate function, e.g., Total Sales Price = Sum (SalesPrice).

Write SQL scripts for below questions:

HW-5 Data Warehouse

1. Show a list of **Customer Name, Gender, Sales Person's Name and Sales Person's City** for all products sold on September 2015, whose Sales Price is more than 20 and Quantity sold is more than 8.

2. Show a list of **Store Name, Store's City and Product Name** for all products sold on March 2017, whose Product Cost is less than 50 and store located in 'Boulder'.

3. Show a list of Top 2 Sales Person's Name by their Total Revenue for 2017, i.e. Top 2 sales person with HIGHEST Total Revenue. **Display Sales Person's Name and Total Revenue**.

4. Display **Customer Name and Total Revenue** who has LOWEST Total Revenue in 2017.

5. Show a list of **Store Name** (in alphabetical order) **and their 'Total Sales Price'** for the year between 2010 and 2017.

6. Display a list of **Store Name, Product Name and their Total Profits** from product name like 'Jasmine Rice' for 2010.

7. Display **Total Revenue** from 'ValueMart Boulder' Store **for each Quarter** during 2016, sort your result by Quarter in chronological order. **Display Quarter as well as Total Revenue**.

8. Display **Customer Name and Total Sales Price** for all items purchased by customers Melinda Gates and Harrison Ford.

9. Display **Store Name, Sales Price and Quantity** for all items sold in March 12th 2017.

10. Display **Sales Person's Name and Total Revenue** for the best performing Sales Person, i.e., the Sales Person with the HIGHEST Total Revenue.

11. Display the Top 3 Product Name by their HIGHEST Total Profit. **Display product name as well as total profit.**

12. Display **Year, MonthName and Total Revenue** for the 1st 3 months (i.e. January, February and March) of 2017.

13. Display **Product Name, average product cost and average sales price** for the products sold in 2017. Show averages rounded to 2 decimal places.

14. Display **Customer Name, average sales price and average quantity** for all items purchased by customer Melinda Gates. Show averages rounded to 2 decimal places.

15. Display **Store Name, Maximum sales price and Minimum sales price** for store located in 'Boulder' city. Show MIN / MAX rounded to 2 decimal places.

☺☺☺ The End ☺☺☺