HW-4 Data Warehouse

Overview

- This homework is worth 15 points (15%) toward your final grade.
- It is due on 4/9/2020, 11:59 pm.
- Late policy applies (grace period of 2 days with 20% late penalty). After that time, no late work will be accepted.
- Save your results in a document using tools such as word, excel, text or similar format and convert into pdf.
- Submit your assignment in PDF on Canvas.
- You must turn in BOTH your SQL and your ANSWER SET (unless otherwise specified).

This homework will give you hands-on practice in working with the structure of a dimensional model /star schema data warehouse. You will first create a data warehouse and insert sample data by using scripts provided. The data warehouse will then be used to answer the query problems given below.

You will be using MySQL workbench to write SQL for this assignment, the tool we 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. Windows based installation files are also available in Canvas.

Creating the Database

You need to download the "**HW_4_SQL_Sales_DW.sql**" script file from Canvas, and execute it. It will create sample database and tables, and then load the tables with data. The script runs fine as-is without any modification. The script begins with a command to DROP the table before creating it. This allows you to run the script over and over as 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 for this data warehouse in the Canvas called "**HW_4 DW_ERD.pdf**". You should download ERD and keep it handy while doing the homework. It is very helpful to have table, column names and their relationship in front of you when writing SQL queries.

HW-4 Data Warehouse

Your **HW_4_SQL_Sales_DW** database consists of the following tables:

- Dim_Product
- Dim_Store
- Dim_Customer
- Dim_Date
- Dim_SalesPerson
- Fact_ProductSales

After running scripts, run the 'Verify' script, as below. You should see the following five dimension and one fact tables and row counts for each

<mark>Note: Please let instructor know if your tables/counts are different.</mark>
<mark>-- Verify row counts</mark>

> **USE HW_4_SQL_sales_dw;**
> SELECT 'Table', 'Rows' FROM dim_customer
> UNION
> SELECT 'dim_customer', COUNT(*) FROM dim_customer
> UNION
> SELECT 'dim_date', COUNT(*) FROM dim_date
> UNION
> SELECT 'dim_product', COUNT(*) FROM dim_product
> UNION
> SELECT 'dim_salesperson', COUNT(*) FROM dim_salesperson
> UNION
> SELECT 'dim_store', COUNT(*) FROM dim_store
> UNION
> SELECT 'fact_productsales', COUNT(*) FROM fact_productsales;

Table	Rows
Table	Rows
dim_customer	5
dim_date	91
dim_product	5
dim_salesperson	6
dim_store	3
fact_productsales	297

HW-4 Data Warehouse

Data Warehouse Query Problems

- You must submit in <mark>BOTH your SQL and your ANSWER SET</mark> (unless otherwise specified).
- Make sure to use table/column aliases, where applicable. Also, use derived column names, e.g. Sum (SalesPrice * Quantity) = Total Revenue so derived column should name as Total Revenue
- **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).

----

1. Show a list of Customer Name, Gender, Sales Person 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 by their Total Revenue for 2017, i.e. Top 2 sales person with HIGHEST Total Revenue.

4. Display a 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.

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 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.

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 ☺☺☺