E-Commerce Data and Customer Retention Analysis with SQL

An e-commerce organization demands some analysis of sales and delivery processes. Thus, the organization hopes to be able to predict more easily the opportunities and threats for the future.

You are asked to make the following analyzes for this scenario by following the instructions given.

Introduction

- You can benefit from the ERD diagram given to you during your work.
- You have to create a database and import into the given csv files. (https://www.youtube.com/watch?v=14FpoXKTEJw)
- During the import process, you will need to adjust the date columns. You need to carefully observe the data types and how they should be.In our database, a star model will be created with one fact table and four dimention tables.
- The data are not very clean and fully normalized. However, they don't prevent you from performing the given tasks. In some cases you may need to use the string, window, system or date functions.
- There may be situations where you need to update the tables.
- Manually verify the accuracy of your analysis.

OPTIONAL: You can clean and normalize the data, change the data types of some columns, clear the id columns, and assign them as keys. Then you can create the data model.

Analyze the data by finding the answers to the questions below:

1. Join all the tables and create a new table with all of the columns, called combined_table. (market_fact, cust_dimen, orders_dimen, prod_dimen, shipping_dimen)
2. Find the top 3 customers who have the maximum count of orders.
3. Create a new column at combined_table as DaysTakenForDelivery that contains the date difference of Order_Date and Ship_Date.

4. Find the customer whose order took the maximum time to get delivered.
5. Retrieve total sales made by each product from the data (use Window function)
6. Retrieve total profit made from each product from the data (use windows function)
7. Count the total number of unique customers in January and how many of them came back every month over the entire year in 2011
8. Write a query to return for each user the time elapsed between the first purchasing and the third purchasing, in ascending order by Customer ID.
9. Write a query that returns customers who purchased both product 11 and product 14 as well as the ratio of these products to the total number of products purchased by the customer.

Find month-by-month customer retention rate[i] since the start of the business (using views).

1. Create a view where each user's visits are logged by month, allowing for the possibility that these will have occurred over multiple years since whenever business started operations.
2. Identify the time lapse between each visit. So, for each person and for each month, we see when the next visit is.
3. Calculate the time gaps between visits.
4. Categorise the customer with time gap 1 as retained, >1 as irregular and NULL as churned.
5. Calculate the retention month wise.

- [i] **Customer retention refers to the ability of a company or product to retain its customers over some specified period.**

  https://en.wikipedia.org/wiki/Customer_retention
  https://www.shopify.com/blog/customer-retention-strategies
  https://ecommercefastlane.com/importance-ecommerce-customer-retention-strategies-keep-customers-coming-back/