SQL CASE STUDY – Basic

Business Context

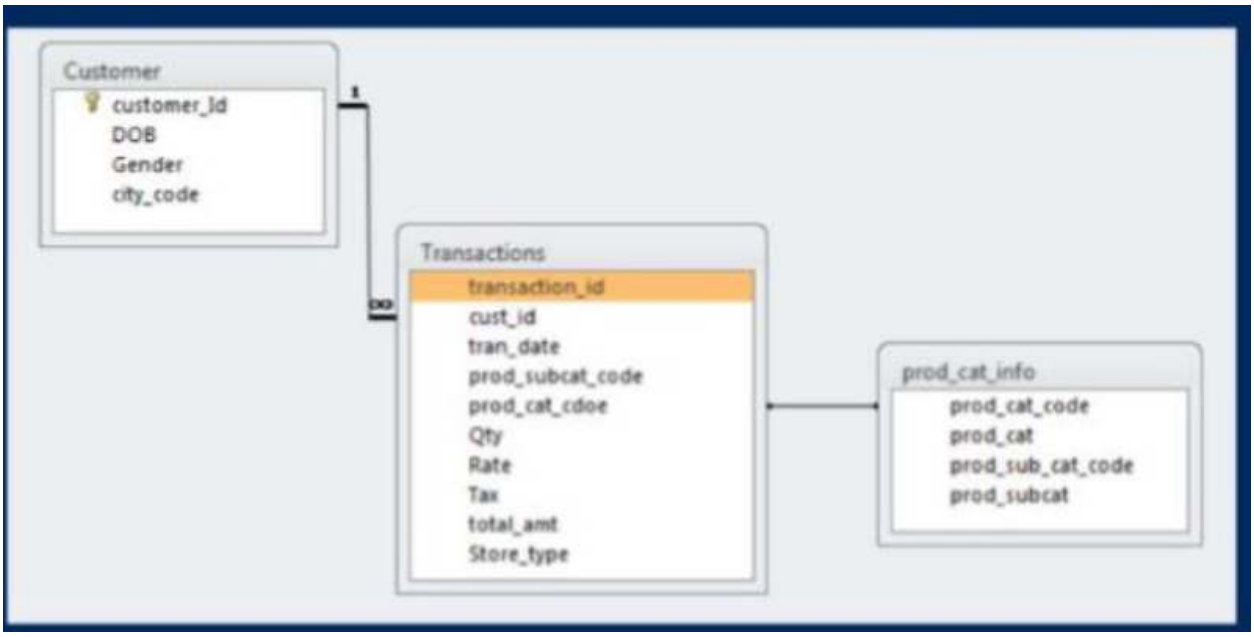
A retail store would like to understand customer behavior using their point of sale data (POS).

Data Availability

The data set we will be using for this exercise comprises of 3 tables:

Customer: Customer demographics
Transactions: Customer transaction details
Product category: Product category and sub category information

The following diagram explains the relationship between tables.



Write SQL queries to get the required output for following business scenarios.

Notes:

- *Only one SQL query should be written for each question.*
- *SQL queries should be written keeping in mind that you have read only access.*
- *You can insert 1000 records only using the INSERT in one INSERT statement in MS SQL Server. To insert more records, use INSERT statement again.*

DATA PREPARATION AND UNDERSTANDING

1. What is the total number of rows in each of the 3 tables in the database?
2. What is the total number of transactions that have a return?
3. As you would have noticed, the dates provided across the datasets are not in a correct format. As first steps, pls convert the date variables into valid date formats before proceeding ahead.
4. What is the time range of the transaction data available for analysis? Show the output in number of days, months and years simultaneously in different columns.
5. Which product category does the sub-category "DIY" belong to?

DATA ANALYSIS

1. Which channel is most frequently used for transactions?
2. What is the count of Male and Female customers in the database?
3. From which city do we have the maximum number of customers and how many?
4. How many sub-categories are there under the Books category?
5. What is the maximum quantity of products ever ordered?
6. What is the net total revenue generated in categories Electronics and Books?
7. How many customers have >10 transactions with us, excluding returns?
8. What is the combined revenue earned from the "Electronics" & "Clothing" categories, from "Flagship stores"?
9. What is the total revenue generated from "Male" customers in "Electronics" category? Output should display total revenue by prod sub-cat.
10. What is percentage of sales and returns by product sub category; display only top 5 sub categories in terms of sales?
11. For all customers aged between 25 to 35 years find what is the net total revenue generated by these consumers in last 30 days of transactions from max transaction date available in the data?

12. Which product category has seen the max value of returns in the last 3 months of transactions?
13. Which store-type sells the maximum products; by value of sales amount and by quantity sold?
14. What are the categories for which average revenue is above the overall average.
15. Find the average and total revenue by each subcategory for the categories which are among top 5 categories in terms of quantity sold.