

## Assignment

### 1. Case Study: Customer Data Analysis

You are a data analyst at a retail company, and you've been tasked with analyzing the customer data to gain insights that can improve marketing strategies and customer satisfaction.

Dataset:  Customer Data set

#### Dataset Description:

customerid: Unique identifier for each customer.

name: First name of the customer.

surname: Last name of the customer.

age: Age of the customer.

gender: Gender of the customer.

job: Occupation of the customer.

balance: Account balance of the customer.

region: Region where the customer resides.

#### Questions:

1. List all customers sorted by their age in ascending order.
2. Display the top 10 customers with the highest balances.
3. List the average balance of customers in each region, only for regions where the average balance is above 60000.
4. Display the top 3 jobs with the highest total balance.
5. Display the top 5 regions with the highest total balance.
6. Find the total balance held by customers aged between 25 and 35 (inclusive).
- 7. List all male customers whose balance is greater than 50000.**

## 2. Case Study: TechMart Data Creation

**You are a data analyst working for a fictional company called "TechMart." TechMart is an online marketplace that sells various tech products. As part of your role, you are tasked with managing the company's MySQL database, which stores information about products, customers, and orders.**

### **MySQL Database Schema:**

**TechMart's database consists of the following tables:**

#### Products

product\_id (INT, Primary Key): Unique identifier for each product.

product\_name (VARCHAR): Name of the product.

price (DECIMAL): Price of the product.

category (VARCHAR): Category of the product.

#### Customers

customer\_id (INT, Primary Key): Unique identifier for each customer.

customer\_name (VARCHAR): Name of the customer.

email (VARCHAR): Email address of the customer.

phone (VARCHAR): Phone number of the customer.

#### Orders

order\_id (INT, Primary Key): Unique identifier for each order.

customer\_id (INT, Foreign Key): Identifier of the customer who placed the order.

product\_id (INT, Foreign Key): Identifier of the product ordered.

quantity (INT): Quantity of the product ordered.

order\_date (DATE): Date when the order was placed.

### **Questions:**

1. Create a SQL script to create the above tables in the TechMart database.
2. Insert 5 records into the Products table with sample data.
3. Insert 3 records into the Customers table with sample data.
4. Insert 10 records into the Orders table with sample data, ensuring that each order is associated with a random customer and product.

### **3. Case Study: Data Analysis on Customer Database**

**You are a data analyst at a retail company. The company maintains a database of customers which contains information such as customer ID, name, gender, age, region, job, and balance. Your task is to perform data analysis on this database to extract valuable insights.**

#### **Database Schema:**

The customers table has the following schema:

customerid (INT): Unique identifier for each customer.

name (VARCHAR): First name of the customer.

surname (VARCHAR): Last name of the customer.

gender (ENUM 'Male', 'Female', 'Other'): Gender of the customer.

age (INT): Age of the customer.

region (VARCHAR): Region where the customer resides.

job (VARCHAR): Occupation of the customer.

balance (DECIMAL): Account balance of the customer.

### **Questions:**

1. Update the job of the customer with customerid 200000008 to 'Data Scientist'.
2. Increase the account balance of all customers by 10%.
3. Set the balance of customers aged below 30 to 0.
4. Update the region of the customer named 'Thomas Lawrence' to 'Wales'.
5. Set the job of all male customers to 'Engineer'.
6. Rollback the last update operation.
7. Commit the changes made so far.
8. Update the balance of customers whose names start with 'J' to double their current balance.
9. Update the job of the customer with the highest balance to 'Manager'.
10. Set the balance of customers in the 'England' region to 50000.

### **4. Case Study: Customer Database Management**

**You are working as a data analyst for a retail company that wants to optimize its customer database management. The company has a MySQL database containing a table named customers with the following columns: customerid, name, surname, gender, age, region, job, and balance.**

### **Questions:**

1. Write a SQL query to delete all records of customers who are older than 50 years.
2. Delete the customer record with customerid 300000812 from the table.
3. Delete records of customers whose job title is 'Other'.
4. Remove records of male customers older than 40 years.
5. Remove records of customers who are from the 'Scotland' region and have a balance less than

### **5. Case Study: Analyzing Customer Data**

**You are working as a data analyst for a retail company. Your task is to analyze the customer data to provide valuable insights to the management team. MySQL database containing a table named customers, which includes the following columns:**

customerid (unique identifier for each customer)  
name (first name of the customer)  
surname (last name of the customer)  
gender (gender of the customer)  
age (age of the customer)  
region (region where the customer resides)  
job (occupation of the customer)  
balance (current balance in the customer's account)

### **Questions:**

1. Determine the total number of male and female customers.
2. Find the maximum and minimum age of customers.
3. Identify the region with the highest number of customers.
4. Find the oldest and youngest customers in terms of age.
5. Calculate the total balance of male and female customers separately.

6. Find the job titles with the highest and lowest average balance.
7. Find the most common job among customers.
8. Calculate the average balance of customers in each region.
9. Identify the top 3 regions with the highest total balance.