ASSIGNMENT-WEEK 4

Data Warehouse and Analytics in the Cloud-Fall 2023

Submitted To:

Dr. Nayem Rehman

6 Aug 2024

Yatender Singh Suman - (yssuman)

yssuman@arizona.edu

Q1.
{ A } Using the MySQL Workbench, create a database called Customer. The database must be named "Customer".
{ B } Check if the database was created and use the same for further questions.

```
base ~/codebase/github.com/channelblend/channel_platform git:(main)
mysql -u root -p'root_password' -P 3306 --protocol=tcp
mysql> create database Customer;
Query OK, 1 row affected (0.02 sec)

mysql> show databases;
+--------------------+
| Database           |
+--------------------+
| Customer           |
| Worker             |
| information_schema |
| mysql              |
| performance_schema |
| sys                |
+--------------------+
6 rows in set (0.00 sec)

mysql> 
```

Q2.
{ A } Create a staging table, ** Customer.CustomerChurn_Stage **, in a database system, with the column list provided in the CSV file. Define the ' CustomerId ' as the Primary Key (PK). Get the table definition (DDL) from the database system and capture it in a Word document for submission.

```
mysql>
mysql> CREATE TABLE Customer.CustomerChurn_Stage (
    ->     CustomerID INTEGER NOT NULL PRIMARY KEY,
    ->     Surname VARCHAR(50),
    ->     CreditScore INTEGER,
    ->     Geography VARCHAR(50),
    ->     Age TINYINT NOT NULL,
    ->     Gender VARCHAR(10) NOT NULL,
    ->     Balance DECIMAL(13, 2) NOT NULL,
    ->     Exited TINYINT(1)
    -> );
Query OK, 0 rows affected, 1 warning (0.08 sec)
```

```
mysql> show create table Customer.CustomerChurn_Stage;
+---------------------+--------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------
---------+
| Table               | Create Table

         |
+---------------------+--------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------
---------+
| CustomerChurn_Stage | CREATE TABLE `CustomerChurn_Stage` (
  `CustomerID` int NOT NULL,
  `Surname` varchar(50) DEFAULT NULL,
  `CreditScore` int DEFAULT NULL,
  `Geography` varchar(50) DEFAULT NULL,
  `Age` tinyint NOT NULL,
  `Gender` varchar(10) NOT NULL,
  `Balance` decimal(13,2) NOT NULL,
  `Exited` tinyint(1) DEFAULT NULL,
  PRIMARY KEY (`CustomerID`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci |
+---------------------+--------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------
---------+
1 row in set (0.01 sec)

mysql> 
```

{ B } Create a persistent table, ** Customer.CustomerChurn **, with the column list provided in the CSV file + following 5 columns : << SourceSystemNm NVARCHAR(20) NOT NULL , CreateAgentId NVARCHAR(20) NOT NULL , CreateDtm DATETIME NOT NULL, ChangeAgentId NVARCHAR(20) NOT NULL , ChangeDtm DATETIME NOT NULL >> Define the ' CustomerId ' as the Primary Key (PK). Get the table definition (DDL) from the database system and capture it in a Word document for submission.

```
mysql> CREATE TABLE Customer.CustomerChurn (
    ->     CustomerId INT PRIMARY KEY,
    ->     Surname VARCHAR(50),
    ->     CreditScore INT,
    ->     Geography VARCHAR(50),
    ->     Gender VARCHAR(10),
    ->     Age TINYINT,
    ->     Balance DECIMAL(13, 2),
    ->     Exited TINYINT(1),
    ->     SourceSystemNm NVARCHAR(20) NOT NULL,
    ->     CreateAgentId NVARCHAR(20) NOT NULL,
    ->     CreateDtm DATETIME NOT NULL,
    ->     ChangeAgentId NVARCHAR(20) NOT NULL,
    ->     ChangeDtm DATETIME NOT NULL
    -> );
Query OK, 0 rows affected, 4 warnings (0.05 sec)

mysql> show create table Customer.CustomerChurn;
+---------------+-----------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------
```

```
Database changed
mysql> SHOW CREATE TABLE CustomerChurn;
+---------------+-----------------------------------------------------------------------------------------------------------------------
| Table         | Create Table
+---------------+-----------------------------------------------------------------------------------------------------------------------
| CustomerChurn | CREATE TABLE `CustomerChurn` (
  `CustomerId` int NOT NULL,
  `Surname` varchar(50) DEFAULT NULL,
  `CreditScore` int DEFAULT NULL,
  `Geography` varchar(50) DEFAULT NULL,
  `Gender` varchar(10) DEFAULT NULL,
  `Age` tinyint DEFAULT NULL,
  `Balance` decimal(13,2) DEFAULT NULL,
  `Exited` tinyint(1) DEFAULT NULL,
  `SourceSystemNm` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `CreateAgentId` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `CreateDtm` datetime NOT NULL,
  `ChangeAgentId` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `ChangeDtm` datetime NOT NULL,
  PRIMARY KEY (`CustomerId`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci |
+---------------+-----------------------------------------------------------------------------------------------------------------------
1 row in set (0.00 sec)
```

Q3. { A } Load the staging table, ** Customer.CustomerChurn_Stage **, with data from the CSV file, CustomerChurn1.csv .
{ B } Verify data by comparing the row counts between the CSV file and the staging table, ** Customer.CustomerChurn_Stage [Data Source: CustomerChurn1.CSV] **. Provide the screenshot of last few rows using the ' SELECT * ' . Make sure the output shows all column values. The SELECT statement must use the ORDER BY ' CustomerId '.

```
mysql> LOAD DATA LOCAL INFILE '/Users/ysingh/Downloads/mysql_files_csv/customer.csv'
    -> INTO TABLE Customer.CustomerChurn_Stage
    -> FIELDS TERMINATED BY ','
    -> ENCLOSED BY '"'
    -> LINES TERMINATED BY '\n'
    -> IGNORE 1 ROWS
    -> (CustomerID, Surname, CreditScore, Geography, Gender, Age, Balance, Exited);
Query OK, 100 rows affected (0.01 sec)
Records: 100  Deleted: 0  Skipped: 0  Warnings: 0

mysql>
mysql> select * from Customer.CustomerChurn_Stage order by CustomerID limit 10 offset 90;
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
| CustomerID | Surname   | CreditScore | Geography | Age | Gender | Balance   | Exited |
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
|   15788218 | Henderson |         549 | Spain     |  24 | Female |      0.00 |      0 |
|   15788448 | Watson    |         490 | Spain     |  31 | Male   | 145260.23 |      0 |
|   15789484 | Hammond   |         751 | Germany   |  36 | Female | 169831.46 |      0 |
|   15792365 | He        |         501 | France    |  44 | Male   | 142051.07 |      0 |
|   15794171 | Lombardo  |         475 | France    |  45 | Female | 134264.04 |      1 |
|   15803136 | Postle    |         416 | Germany   |  41 | Female | 122189.66 |      0 |
|   15804771 | Velazquez |         614 | France    |  51 | Male   |  40685.92 |      0 |
|   15805254 | Ndukaku   |         652 | Spain     |  75 | Female |      0.00 |      0 |
|   15809248 | Cole      |         524 | France    |  36 | Female |      0.00 |      0 |
|   15812518 | Palermo   |         657 | Spain     |  37 | Female | 163607.18 |      0 |
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
10 rows in set (0.00 sec)

mysql>
```

Q4. Create a database stored procedure based on the template provided along with this assignment << StoredProc_Template.txt >>. Name the stored procedure name this: ** Customer.PrCustomerChurn ** . [[NOTE : This stored procedure will use the table, ** Customer.CustomerChurn_Stage ** , as the source (aka, staging table). This stored procedure will use the table, ** Customer. CustomerChurn **, as the target (aka, persistent table).]]

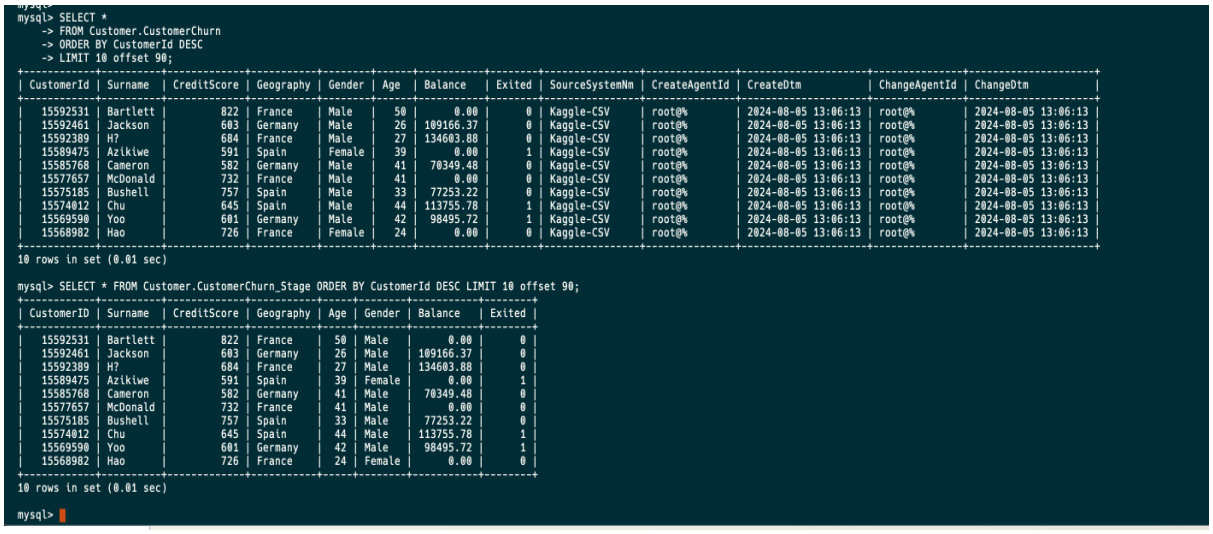
Q5. Execute the stored procedure, ** Customer.PrCustomerChurn **, that was created in Q4. After execution, the stored procedure should load data from the stage to the persistent table: ** Customer.CustomerChurn **.

{A} Verify data by comparing the row counts between the staging table, ** Customer.CustomerChurn_Stage [Data Source: CustomerChurn1.CSV] ** and the persistent table: ** Customer.CustomerChurn **.

```
mysql>
mysql> CALL Customer.PrCustomerChurn();
+----------------------------+
| DebugMessage               |
+----------------------------+
| Debug 1: Starting procedure |
+----------------------------+
1 row in set (0.01 sec)

+----------------------------+-------------------+-------------------+
| DebugMessage               | VarSourceRowCount | VarTargetRowCount |
+----------------------------+-------------------+-------------------+
| Debug 2: After counting rows |               100 |                 0 |
+----------------------------+-------------------+-------------------+
1 row in set (0.01 sec)

+-----------------------------+-----------------+
| DebugMessage                | VarThresholdNbr |
+-----------------------------+-----------------+
| Debug 3: Calculated threshold |               0 |
+-----------------------------+-----------------+
1 row in set (0.01 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 5: Before DELETE operation |
+---------------------------------+
1 row in set (0.01 sec)

ERROR 1093 (HY000): You can't specify target table 'TrgtTbl' for update in FROM clause
mysql> CALL Customer.PrCustomerChurn();
+----------------------------+
| DebugMessage               |
+----------------------------+
| Debug 1: Starting procedure |
+----------------------------+
1 row in set (0.01 sec)

+----------------------------+-------------------+-------------------+
| DebugMessage               | VarSourceRowCount | VarTargetRowCount |
+----------------------------+-------------------+-------------------+
| Debug 2: After counting rows |               100 |                 0 |
+----------------------------+-------------------+-------------------+
1 row in set (0.01 sec)

+-----------------------------+-----------------+
| DebugMessage                | VarThresholdNbr |
+-----------------------------+-----------------+
| Debug 3: Calculated threshold |               0 |
+-----------------------------+-----------------+
1 row in set (0.01 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 5: Before DELETE operation |
+---------------------------------+
1 row in set (0.01 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 6: Before UPDATE operation |
+---------------------------------+
1 row in set (0.01 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 7: Before INSERT operation |
+---------------------------------+
1 row in set (0.02 sec)

+-----------------------------+
| DebugMessage                |
+-----------------------------+
| Debug 8: Procedure completed |
+-----------------------------+
1 row in set (0.03 sec)

Query OK, 0 rows affected (0.03 sec)
```

```
mysql> show tables;
+----------------------+
| Tables_in_Customer   |
+----------------------+
| CustomerChurn        |
| CustomerChurn_Stage  |
+----------------------+
2 rows in set (0.02 sec)

mysql> select count(*) from CustomerChurn;
+----------+
| count(*) |
+----------+
|      100 |
+----------+
1 row in set (0.00 sec)

mysql> select count(*) from CustomerChurn_stage;
ERROR 1146 (42S02): Table 'Customer.CustomerChurn_stage' doesn't exist
mysql> select count(*) from CustomerChurn_Stage;
+----------+
| count(*) |
+----------+
|      100 |
+----------+
1 row in set (0.01 sec)
```

{ B } Provide the screenshot of last few rows using the SELECT *. Make sure the output shows all column values. The SELECT statement must use the ORDER BY CustomerId.



```
mysql> SELECT *
    -> FROM Customer.CustomerChurn
    -> ORDER BY CustomerId DESC
    -> LIMIT 10 offset 90;
+------------+----------+-------------+-----------+--------+-----+-----------+--------+--------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname  | CreditScore | Geography | Gender | Age | Balance   | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+----------+-------------+-----------+--------+-----+-----------+--------+--------------+---------------+---------------------+---------------+---------------------+
|   15592531 | Bartlett |         822 | France    | Male   |  50 |      0.00 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15592461 | Jackson  |         603 | Germany   | Male   |  26 | 109166.37 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15592389 | H?       |         684 | France    | Male   |  27 | 134603.88 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15589475 | Azikiwe  |         591 | Spain     | Female |  39 |      0.00 |      1 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15585768 | Cameron  |         582 | Germany   | Male   |  41 |  70349.48 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15577657 | McDonald |         732 | France    | Male   |  41 |      0.00 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15575185 | Bushell  |         757 | Spain     | Male   |  33 |  77253.22 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15574012 | Chu      |         645 | Spain     | Male   |  44 | 113755.78 |      1 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15569590 | Yoo      |         601 | Germany   | Male   |  42 |  98495.72 |      1 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15568982 | Hao      |         726 | France    | Female |  24 |      0.00 |      0 | Kaggle-CSV   | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
+------------+----------+-------------+-----------+--------+-----+-----------+--------+--------------+---------------+---------------------+---------------+---------------------+
10 rows in set (0.01 sec)

mysql> SELECT * FROM Customer.CustomerChurn_Stage ORDER BY CustomerId DESC LIMIT 10 offset 90;
+------------+----------+-------------+-----------+-----+--------+-----------+--------+
| CustomerID | Surname  | CreditScore | Geography | Age | Gender | Balance   | Exited |
+------------+----------+-------------+-----------+-----+--------+-----------+--------+
|   15592531 | Bartlett |         822 | France    |  50 | Male   |      0.00 |      0 |
|   15592461 | Jackson  |         603 | Germany   |  26 | Male   | 109166.37 |      0 |
|   15592389 | H?       |         684 | France    |  27 | Male   | 134603.88 |      0 |
|   15589475 | Azikiwe  |         591 | Spain     |  39 | Female |      0.00 |      1 |
|   15585768 | Cameron  |         582 | Germany   |  41 | Male   |  70349.48 |      0 |
|   15577657 | McDonald |         732 | France    |  41 | Male   |      0.00 |      0 |
|   15575185 | Bushell  |         757 | Spain     |  33 | Male   |  77253.22 |      0 |
|   15574012 | Chu      |         645 | Spain     |  44 | Male   | 113755.78 |      1 |
|   15569590 | Yoo      |         601 | Germany   |  42 | Male   |  98495.72 |      1 |
|   15568982 | Hao      |         726 | France    |  24 | Female |      0.00 |      0 |
+------------+----------+-------------+-----------+-----+--------+-----------+--------+
10 rows in set (0.01 sec)

mysql>
```

Q6. After data verification is completed, in Q5 ,
{ A } create table, ** Customer.CustomerChurn_Version1 **, with data from ** Customer.CustomerChurn ** (that was already loaded from Customer.CustomerChurn_Stage via the stored procedure).

{ B } Show table definition of Customer.CustomerChurn_Version1 and show the row count of the table, ** Customer.CustomerChurn_Version1 **:

```
mysql> CREATE TABLE Customer.CustomerChurn_Version1 AS
    -> SELECT * FROM Customer.CustomerChurn;
Query OK, 100 rows affected (0.05 sec)
Records: 100  Duplicates: 0  Warnings: 0

mysql> show tables;
+------------------------+
| Tables_in_Customer     |
+------------------------+
| CustomerChurn          |
| CustomerChurn_Stage    |
| CustomerChurn_Version1 |
+------------------------+
3 rows in set (0.01 sec)

mysql> show create table CustomerChurn_Version1;
+------------------------+--------------------------------
| Table                  | Create Table
                                 |
+------------------------+--------------------------------
| CustomerChurn_Version1 | CREATE TABLE `CustomerChurn_Version1` (
  `CustomerId` int NOT NULL,
  `Surname` varchar(50) DEFAULT NULL,
  `CreditScore` int DEFAULT NULL,
  `Geography` varchar(50) DEFAULT NULL,
  `Gender` varchar(10) DEFAULT NULL,
  `Age` tinyint DEFAULT NULL,
  `Balance` decimal(13,2) DEFAULT NULL,
  `Exited` tinyint(1) DEFAULT NULL,
  `SourceSystemNm` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `CreateAgentId` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `CreateDtm` datetime NOT NULL,
  `ChangeAgentId` varchar(20) CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci NOT NULL,
  `ChangeDtm` datetime NOT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci |
+------------------------+--------------------------------
1 row in set (0.00 sec)
```

{ C } Provide the screenshot of last few rows for ** Customer.CustomerChurn_Version1 ** [Originally data came from: CustomerChurn1.CSV]. Make sure the output shows all column values. The SELECT statement must use the ORDER BY CustomerID.

```
mysql> select count(*) from CustomerChurn_Version1;
+----------+
| count(*) |
+----------+
|      100 |
+----------+
1 row in set (0.01 sec)

mysql> select * from CustomerChurn_Version1 order by CustomerID LIMIT 10 OFFSET 90;
+------------+-----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname   | CreditScore | Geography | Gender | Age  | Balance   | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+-----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
|   15788218 | Henderson |         549 | Spain     | Female |   24 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15788448 | Watson    |         490 | Spain     | Male   |   31 | 145260.23 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15789484 | Hammond   |         751 | Germany   | Female |   36 | 169831.46 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15792365 | He        |         501 | France    | Male   |   44 | 142051.07 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15794171 | Lombardo  |         475 | France    | Female |   45 | 134264.04 |      1 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15803136 | Postle    |         416 | Germany   | Female |   41 | 122189.66 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15804771 | Velazquez |         614 | France    | Male   |   51 |  40685.92 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15805254 | Ndukaku   |         652 | Spain     | Female |   75 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15809248 | Cole      |         524 | France    | Female |   36 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15812518 | Palermo   |         657 | Spain     | Female |   37 | 163607.18 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
+------------+-----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
10 rows in set (0.00 sec)

mysql> select * from CustomerChurn_Version1 order by CustomerID DESC LIMIT 10 OFFSET 90;
+------------+----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname  | CreditScore | Geography | Gender | Age  | Balance   | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
|   15592531 | Bartlett |         822 | France    | Male   |   50 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15592461 | Jackson  |         603 | Germany   | Male   |   26 | 109166.37 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15592389 | H?       |         684 | France    | Male   |   27 | 134603.88 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15589475 | Azikiwe  |         591 | Spain     | Female |   39 |      0.00 |      1 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15585768 | Cameron  |         582 | Germany   | Male   |   41 |  70349.48 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15577657 | McDonald |         732 | France    | Male   |   41 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15575185 | Bushell  |         757 | Spain     | Male   |   33 |  77253.22 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15574012 | Chu      |         645 | Spain     | Male   |   44 | 113755.78 |      1 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15569590 | Yoo      |         601 | Germany   | Male   |   42 |  98495.72 |      1 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
|   15568982 | Hao      |         726 | France    | Female |   24 |      0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 13:06:13 | root@%        | 2024-08-05 13:06:13 |
+------------+----------+-------------+-----------+--------+------+-----------+--------+----------------+---------------+---------------------+---------------+---------------------+
10 rows in set (0.01 sec)

mysql>
```

{ D } Empty the staging table, ** Customer.CustomerChurn_Stage **, and load it with data from the CSV file, "CustomerChurn2.csv ". Verify data by comparing the row counts between the CSV file and the staging table, ** Customer.CustomerChurn_Stage ** [Data Source: CustomerChurn2.CSV]. Provide the row count of ** Customer.CustomerChurn_Stage ** that you loaded from CustomerChurn2.csv file. Provide the screenshot of last few rows using the SELECT *. Make sure the output shows all column values. The SELECT statement must use the ORDER BY CustomerId.

```
mysql> TRUNCATE TABLE Customer.CustomerChurn_Stage;
Query OK, 0 rows affected (0.04 sec)

mysql> select * from Customer.CustomerChurn_Stage;
Empty set (0.01 sec)

mysql> LOAD DATA LOCAL INFILE '/Users/ysingh/Downloads/mysql_files_csv/CustomerChurn2.csv'
    -> INTO TABLE Customer.CustomerChurn_Stage
    -> FIELDS TERMINATED BY ','
    -> ENCLOSED BY '"'
    -> LINES TERMINATED BY '\n'
    -> IGNORE 1 ROWS
    -> (CustomerID, Surname, CreditScore, Geography, Gender, Age, Balance, Exited);
Query OK, 101 rows affected (0.02 sec)
Records: 101  Deleted: 0  Skipped: 0  Warnings: 0
```

```
mysql> SELECT * FROM Customer.CustomerChurn_Stage ORDER BY CustomerID LIMIT 10 OFFSET 90;
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
| CustomerID | Surname   | CreditScore | Geography | Age | Gender | Balance   | Exited |
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
|   15780961 | Cavenagh  |         735 | France    |  21 | Female | 178718.19 |      0 |
|   15788218 | Henderson |         549 | Spain     |  24 | Female |      0.00 |      0 |
|   15788448 | Watson    |         490 | Spain     |  31 | Male   | 145260.23 |      0 |
|   15789484 | Hammond   |         751 | Germany   |  36 | Female | 169831.46 |      0 |
|   15792365 | He        |         501 | France    |  44 | Male   | 142051.07 |      0 |
|   15794171 | Lombardo  |         475 | France    |  45 | Female | 134264.04 |      1 |
|   15803136 | Postle    |         416 | Germany   |  41 | Female | 122189.66 |      0 |
|   15804771 | Velazquez |         614 | France    |  51 | Male   |  40685.92 |      0 |
|   15805254 | Ndukaku   |         652 | Spain     |  75 | Female |      0.00 |      0 |
|   15809248 | Cole      |         524 | France    |  36 | Female |      0.00 |      0 |
+------------+-----------+-------------+-----------+-----+--------+-----------+--------+
10 rows in set (0.00 sec)

mysql>
```

Q7. Execute the stored procedure, Customer.PrCustomerChurn, that was created in Q4. After execution, the stored procedure should load data from the stage to the persistent table: Customer.CustomerChurn. CALL `customer`.`PrCustomerChurn`(); This time, the table will be refreshed via DELETE, UPDATE, and INSERT/SELECT statements in the stored procedure. Show the row count results of both Customer.CustomerChurn_Version1 table [Data Source: CustomerChurn1.CSV] and the persistent table: Customer.CustomerChurn. Compare the rows between the Customer.CustomerChurn_Version1 [Data Source: CustomerChurn1.CSV] table and the persistent table: Customer.CustomerChurn [Data Source: CustomerChurn2.CSV]. Show the rows that are available in the Customer.CustomerChurn_Version1 table but not in the Customer.CustomerChurn table (implementation of brand-new row DELETE

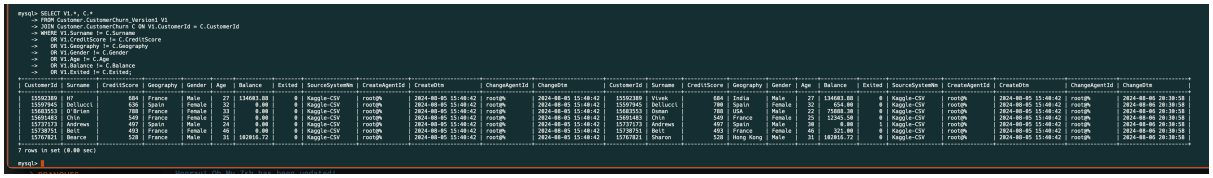
statement of the stored procedure).

```
mysql> CALL `customer`.`PrCustomerChurn`();
+-----------------------------+
| DebugMessage                |
+-----------------------------+
| Debug 1: Starting procedure |
+-----------------------------+
1 row in set (0.00 sec)

+-----------------------------+------------------+------------------+
| DebugMessage                | VarSourceRowCount | VarTargetRowCount |
+-----------------------------+------------------+------------------+
| Debug 2: After counting rows |              101 |              100 |
+-----------------------------+------------------+------------------+
1 row in set (0.00 sec)

+-----------------------------+-----------------+
| DebugMessage                | VarThresholdNbr |
+-----------------------------+-----------------+
| Debug 3: Calculated threshold |              20 |
+-----------------------------+-----------------+
1 row in set (0.00 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 5: Before DELETE operation |
+---------------------------------+
1 row in set (0.00 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 6: Before UPDATE operation |
+---------------------------------+
1 row in set (0.03 sec)

+---------------------------------+
| DebugMessage                    |
+---------------------------------+
| Debug 7: Before INSERT operation |
+---------------------------------+
1 row in set (0.04 sec)

+-----------------------------+
| DebugMessage                |
+-----------------------------+
| Debug 8: Procedure completed |
+-----------------------------+
1 row in set (0.04 sec)

Query OK, 0 rows affected (0.04 sec)

mysql>
```

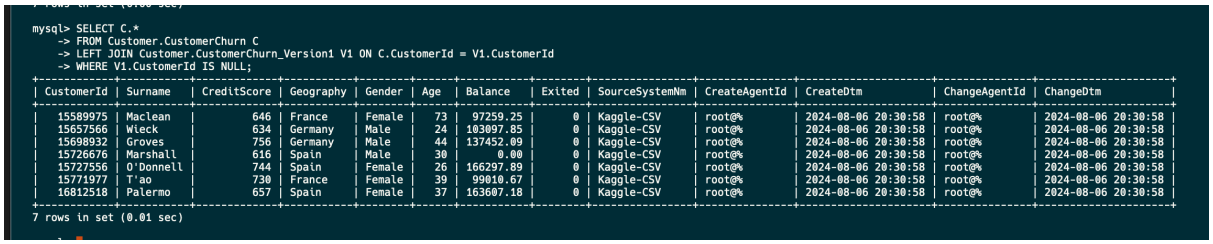
```
mysql> SELECT 'CustomerChurn_Version1' AS TableName, COUNT(*) AS RowCount FROM Customer.CustomerChurn_Version1
    -> UNION ALL
    -> SELECT 'CustomerChurn' AS TableName, COUNT(*) AS RowCount FROM Customer.CustomerChurn;
+------------------------+----------+
| TableName              | RowCount |
+------------------------+----------+
| CustomerChurn_Version1 |      100 |
| CustomerChurn          |      101 |
+------------------------+----------+
2 rows in set (0.01 sec)

mysql>
```

```
mysql> SELECT V1.*
    -> FROM Customer.CustomerChurn_Version1 V1
    -> LEFT JOIN Customer.CustomerChurn C ON V1.CustomerId = C.CustomerId
    -> WHERE C.CustomerId IS NULL;
+------------+-----------+-------------+-----------+--------+-----+----------+--------+----------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname   | CreditScore | Geography | Gender | Age | Balance  | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+-----------+-------------+-----------+--------+-----+----------+--------+----------------+---------------+---------------------+---------------+---------------------+
|   15604348 | Allard    |         710 | Spain     | Male   |  22 |     0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15687946 | Osborne   |         556 | France    | Female |  61 | 117419.35 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15701164 | Onyeorulu |         506 | France    | Female |  34 |  90307.62 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15725737 | Mosman    |         669 | France    | Male   |  46 |     0.00 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15755648 | Pisano    |         675 | France    | Female |  21 |  98373.26 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15812518 | Palermo   |         657 | Spain     | Female |  37 | 163607.18 |      0 | Kaggle-CSV     | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
+------------+-----------+-------------+-----------+--------+-----+----------+--------+----------------+---------------+---------------------+---------------+---------------------+
6 rows in set (0.00 sec)

mysql>
```

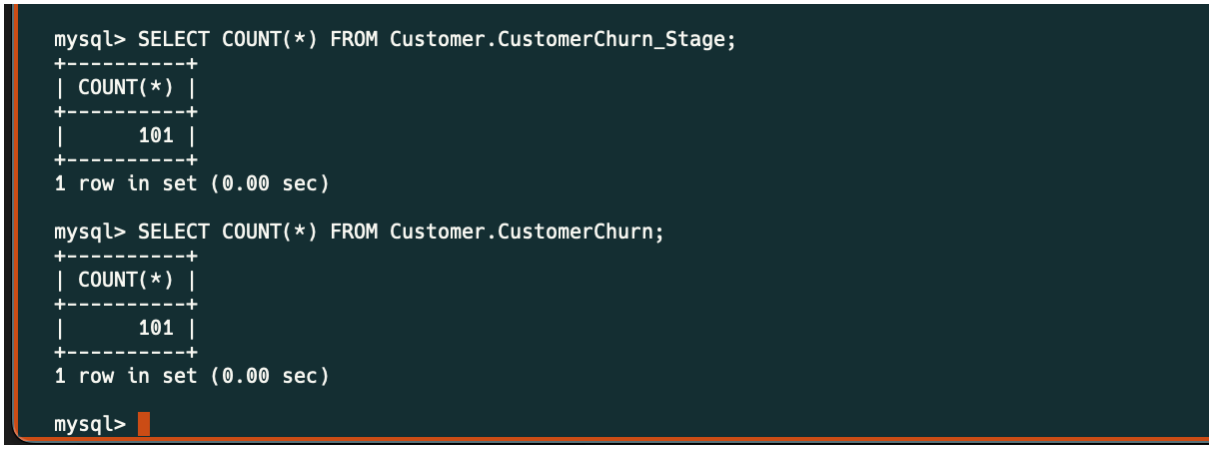
Note: Additional Analysis to see Updated rows



Analysis: Newly inserted rows



```
mysql> SELECT C.*
    -> FROM Customer.CustomerChurn C
    -> LEFT JOIN Customer.CustomerChurn_Version1 V1 ON C.CustomerId = V1.CustomerId
    -> WHERE V1.CustomerId IS NULL;
+------------+-----------+-------------+-----------+--------+-----+-----------+--------+---------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname   | CreditScore | Geography | Gender | Age | Balance   | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+-----------+-------------+-----------+--------+-----+-----------+--------+---------------+---------------+---------------------+---------------+---------------------+
|   15589975 | Maclean   |         646 | France    | Female |  73 |  97259.25 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15657566 | Wieck     |         634 | Germany   | Male   |  24 | 103097.85 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15698932 | Groves    |         756 | Germany   | Male   |  44 | 137452.09 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15726676 | Marshall  |         616 | Spain     | Male   |  30 |      0.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15727556 | O'Donnell |         744 | Spain     | Female |  26 | 166297.89 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15771977 | T'ao      |         730 | France    | Female |  39 |  99010.67 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   16812518 | Palermo   |         657 | Spain     | Female |  37 | 163607.18 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
+------------+-----------+-------------+-----------+--------+-----+-----------+--------+---------------+---------------+---------------------+---------------+---------------------+
7 rows in set (0.01 sec)
```



```
mysql> SELECT COUNT(*) FROM Customer.CustomerChurn_Stage;
+----------+
| COUNT(*) |
+----------+
|      101 |
+----------+
1 row in set (0.00 sec)

mysql> SELECT COUNT(*) FROM Customer.CustomerChurn;
+----------+
| COUNT(*) |
+----------+
|      101 |
+----------+
1 row in set (0.00 sec)

mysql>
```

Q8. Show the rows (SELECT *) that changed (one or many non-Primary Key columns), in the Customer.CustomerChurn table (implementation of UPDATE statement of the stored procedure). You need to perform a comparison between Customer.CustomerChurn table [Data Source: CustomerChurn2.CSV] and Customer.CustomerChurn_Version1 table [Data Source: CustomerChurn1.CSV] in terms of non-PK columns (Excluds: SourceSystemNm, CreateAgentId, CreateDtm, ChangeAgentId, ChangeDtm), and with a join condition using the PK column(s). You must do ORDER BY CustomerId. The output of this query should show different values for the CreateDtm and ChangeDtm columns in Customer.CustomerChurn table for the changed rows. Take a screenshot and capture it in the Word document. Make sure all columns including CreateDtm and ChangeDtm of CustomerChurn table are displayed.

```
mysql> SELECT
    ->     C.CustomerId,
    ->     C.Surname,
    ->     C.CreditScore,
    ->     C.Geography,
    ->     C.Gender,
    ->     C.Age,
    ->     C.Balance,
    ->     C.Exited,
    ->     C.SourceSystemNm,
    ->     C.CreateAgentId,
    ->     C.CreateDtm,
    ->     C.ChangeAgentId,
    ->     C.ChangeDtm,
    ->     V1.Surname AS V1_Surname,
    ->     V1.CreditScore AS V1_CreditScore,
    ->     V1.Geography AS V1_Geography,
    ->     V1.Gender AS V1_Gender,
    ->     V1.Age AS V1_Age,
    ->     V1.Balance AS V1_Balance,
    ->     V1.Exited AS V1_Exited
    -> FROM
    ->     Customer.CustomerChurn C
    -> JOIN
    ->     Customer.CustomerChurn_Version1 V1 ON C.CustomerId = V1.CustomerId
    -> WHERE
    ->     C.Surname != V1.Surname OR
    ->     C.CreditScore != V1.CreditScore OR
    ->     C.Geography != V1.Geography OR
    ->     C.Gender != V1.Gender OR
    ->     C.Age != V1.Age OR
    ->     C.Balance != V1.Balance OR
    ->     C.Exited != V1.Exited
    -> ORDER BY
    ->     C.CustomerId;
+------------+---------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+------------+----------------+--------------+--------+-----------+------------+
| CustomerId | Surname | CreditScore | Geography | Gender | Age | Balance  | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           | V1_Surname | V1_CreditScore | V1_Geography | V1_Gender | V1_Age | V1_Balance | V1_Exited |
+------------+---------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+------------+----------------+--------------+--------+-----------+------------+
|   15592389 | Vivek   |         684 | India     | Male   |  27 | 134603.88 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | H?         |            684 | France       | Male      |     27 |  134603.88 |         0 |
|   15597945 | Dellucci |         708 | Spain     | Female |  32 |   654.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | Dellucci   |            636 | Spain        | Female    |     32 |       0.00 |         0 |
|   15683553 | Osman   |         708 | USA       | Male   |  22 | 75888.30 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | O'Brien    |            708 | France       | Female    |     33 |       0.00 |         0 |
|   15691483 | Chin    |         549 | France    | Female |  25 | 12345.50 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | Chin       |            549 | France       | Female    |     25 |       0.00 |         0 |
|   15737173 | Andrews |         497 | Spain     | Male   |  30 |     0.00 |      1 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | Andrews    |            497 | Spain        | Male      |     24 |       0.00 |         0 |
|   15738751 | Bett    |         493 | France    | Female |  46 |   321.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | Bett       |            493 | France       | Female    |     46 |       0.00 |         0 |
|   15767821 | Sharon  |         528 | Hong Kong | Male   |  31 | 102816.72 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-06 20:30:58 | Bearce     |            528 | France       | Male      |     31 |  102816.72 |         0 |
+------------+---------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+------------+----------------+--------------+--------+-----------+------------+
7 rows in set (0.01 sec)
```

Q9. Provide the screenshot of last few rows using the SELECT * FROM Customer.CustomerChurn. Make sure the output shows all column values. The SELECT statement must use the ORDER BY CustomerId. Show the rows that are available in the Customer.CustomerChurn table [Data Source: CustomerChurn2.CSV] but not in the Customer.CustomerChurn_Version1 table (implementation of brand-new rows INSERT by the stored procedure). Do a SELECT * along with ORDER BY CustomerId. Take a screenshot and capture it in the Word document.

```
mysql> SELECT C.*
    -> FROM Customer.CustomerChurn C
    -> LEFT JOIN Customer.CustomerChurn_Version1 V1 ON C.CustomerId = V1.CustomerId
    -> WHERE V1.CustomerId IS NULL
    -> ORDER BY C.CustomerId;
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname  | CreditScore | Geography | Gender | Age | Balance  | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
|   15589975 | Maclean  |         646 | France    | Female |  73 |  97259.25 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15657566 | Wleck    |         634 | Germany   | Male   |  24 | 103097.85 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15698932 | Groves   |         756 | Germany   | Male   |  44 | 137452.09 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15726676 | Marshall |         616 | Spain     | Male   |  30 |     0.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15727556 | O'Donnell |         744 | Spain     | Female |  26 | 166297.89 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15771977 | T'ao     |         730 | France    | Female |  39 |  99010.67 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   16812518 | Palermo  |         657 | Spain     | Female |  37 | 163607.18 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
7 rows in set (0.01 sec)

mysql>
mysql>
mysql> SELECT *
    -> FROM Customer.CustomerChurn
    -> ORDER BY CustomerId DESC
    -> LIMIT 10;
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
| CustomerId | Surname  | CreditScore | Geography | Gender | Age | Balance  | Exited | SourceSystemNm | CreateAgentId | CreateDtm           | ChangeAgentId | ChangeDtm           |
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
|   16812518 | Palermo  |         657 | Spain     | Female |  37 | 163607.18 |      0 | Kaggle-CSV    | root@%        | 2024-08-06 20:30:58 | root@%        | 2024-08-06 20:30:58 |
|   15809248 | Cole     |         524 | France    | Female |  36 |     0.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15805254 | Ndukaku  |         652 | Spain     | Female |  75 |     0.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15804774 | Velazquez |         614 | France    | Male   |  51 |  40685.92 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15803136 | Postle   |         416 | Germany   | Female |  41 | 122189.66 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15794171 | Lombardo |         475 | France    | Female |  45 | 134264.04 |      1 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15792365 | He       |         501 | France    | Male   |  44 | 142051.07 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15789484 | Hammond  |         751 | Germany   | Female |  36 | 169831.46 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15788448 | Watson   |         490 | Spain     | Male   |  31 | 145260.23 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
|   15788218 | Henderson |         549 | Spain     | Female |  24 |     0.00 |      0 | Kaggle-CSV    | root@%        | 2024-08-05 15:40:42 | root@%        | 2024-08-05 15:40:42 |
+------------+----------+-------------+-----------+--------+-----+----------+--------+---------------+---------------+---------------------+---------------+---------------------+
10 rows in set (0.00 sec)
```