**Analytics project**

Raw data: Ref. **Kaggle** – Customer data of a “X” shopping mall located in Australia

**Problem statement:**

Let’s say the shopping mall named “X” is thinking to launch their next store in Australia. The problem is in which state in Australia the next store to be opened. For this, we are going to find out the state in Australia that has large number of customers and sales for store “X” based on its historical data using SQL.

**Raw data set:**





**Aggregate functions/Clauses/joins used:**

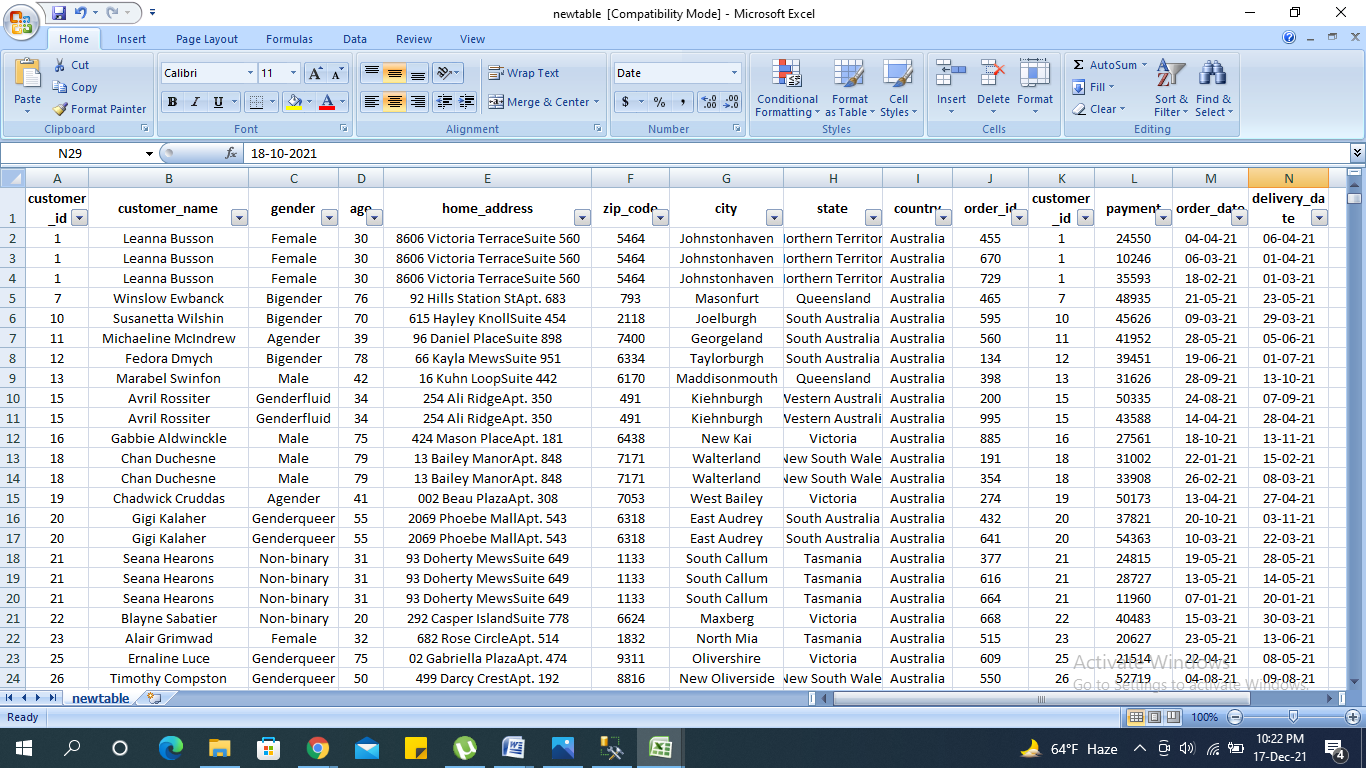
* Count (\*)
* Group by
* Order by
* desc
* Sum
* Inner join

The data set (project1) contains details of 1000 customers of different ages, different genders from different cities and states and ‘orders’ contains payments made by the customers past one year within Australia where our interest lies in finding the state in which large group of customers are present and revenue generated which in turn increases the sales from that area if a store is launched.

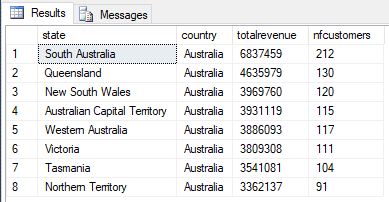
**Steps involved in the process:**

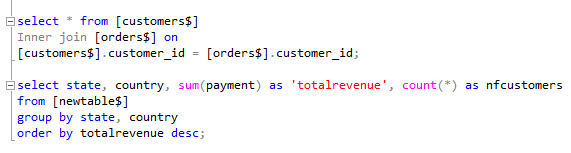
* Import the data from source into SQL server management studio in the form of tables.
* As the data required is in two tables, and customer Id is the common column that can be seen, merge the common data using Inner join. A new table will be created in the results. Use export and import options to add new table into the database.
* Now that we merged the two tables using Inner join, use Group by clause to group required data from the new table, here I’ve grouped state, country, total revenue made from each state(aggregate function ‘sum’ was used) and no. of customers in each state (count(\*) was used) in Australia
* Then it was arranged in descending order (used order by clause and desc aggregate function) to find out which state had generated highest revenue and luckily the state which got highest revenue has largest group of customers for store “X”

**This is how new data set (#newtable) looks like after merging using Inner join**

****

**Query: Results:**





**Conclusion:**

Based on the data analysis, Store “X” can plan their next store launch at South Australia having highest revenue generated (~6838K AUD) and larger number of customers. (21% of total customers)