Assignment 5

Extensibility of code:

- I have deigned my queries and code in such a way that it is seamlessly extensible.
- I have made separate classes for customer, product and supplier information, so if any additional information is required for any one of the above, all I need to do is add the condition in query, create and add those elements in the already created document.
- If in case any of the three things i.e. customer, product or supplier information needs to be removed from the program than I just need to remove than classes function call.
- From the queries perspective, small additions will make the queries return the desired additional information.
- I have also made a separate class for Document generation, making the code more extensible with less cohesion.

The argument for deploying the code:

- The code is neatly divided into modules and classes focusing on specific functionalities to make it more readable, efficient and easier to follow.
- All modules have been tested thoroughly to eliminate any errors and exceptions faced.
- Three inputs are taken from the user; Start Date, End Date and Absolute File Path (XML file) to write the data retrieved.
- I have written the code that validates all the inputs taken from the user, making code more robust.
- The code needs to connect with Dal database, so all the connection related conditions are taken into consideration.
- Queries are used to retrieve the data from the database, so all the queries have been tested thoroughly on MySQL Workbench before being used in the code.
 - (I have attached below the screenshots of the step-by-step process that I followed to finalize the queries)
- The thoroughly tested queries are then being used in the code to retrieve data and storing it in a resultSet.
- Data is then parsed from the resultSet into various elements which eventually gets added to the main document.
- I have used Document Builder to store the data retrieved from running the queries in XML and then writing it to the user provided file.
- I have referred a code snippet from StackOverflow to write the XML to a file using Document Builder, where rather than creating my own nodes, I have used java inbuilt classes for formatting and generating the xml file.
- I have used the above approach in order to improve the extensibility of the code. If the query changes and we need to add one or more column in the resultSet, then not much changes are required. The entire code could be used by adding those new attributes, which additionally we are retrieving from the database and writing it to the file.
- I have documented the code very well so that anyone can understand the flow and what the code actually does.

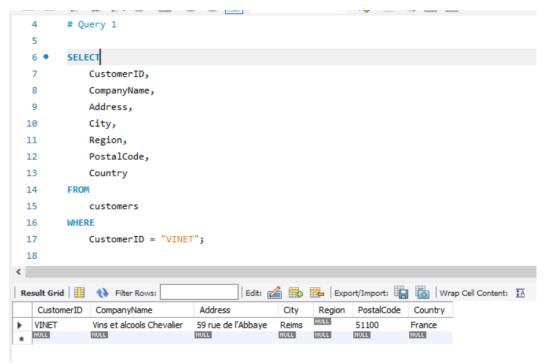
NOTE:

I have made two driver files; one file "**Test.java**" accepting the inputs from the user through an IntelliJ and other file "**CommandLine.java**" accepting inputs through command line. So, either of the files could be used.

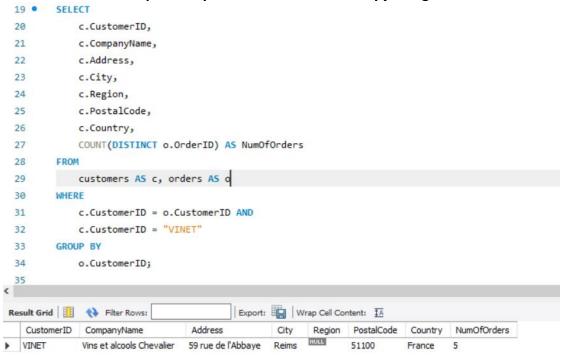
I have also attached the "Assignment5.sql" file containing all the required queries along with the steps that I used to reach the final queries.

Extracting Data using MySQL Queries:

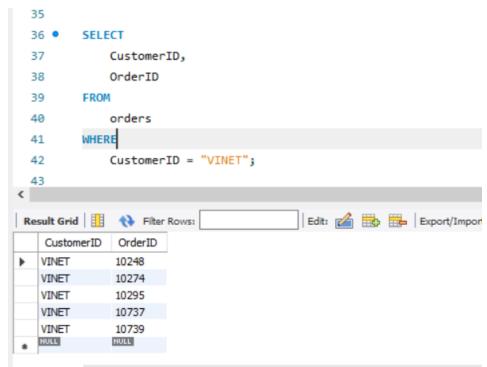
- 1. Customer Information:
 - a. Getting the CustomerID, CompanyName, entire Address of CustomerID = "VINET"



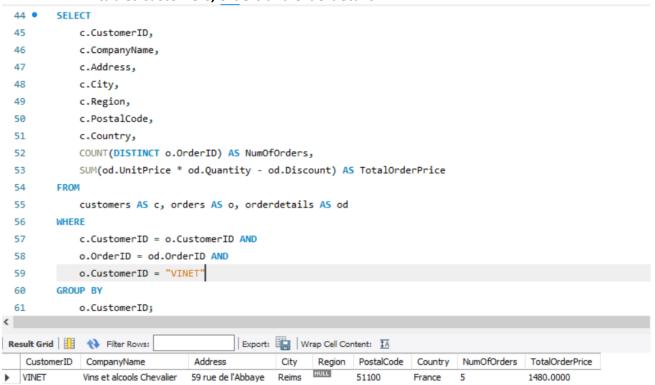
b. Getting the CustomerID, CompanyName, entire Address & also calculating the number of orders placed by CustomerID = "VINET" by joining tables customers and orders



 Getting the number of orders placed by CustomerID = "VINET" from orders table (Verifying the above query where I am aggregating the orders placed by "VINET" which totals to 5 orders)



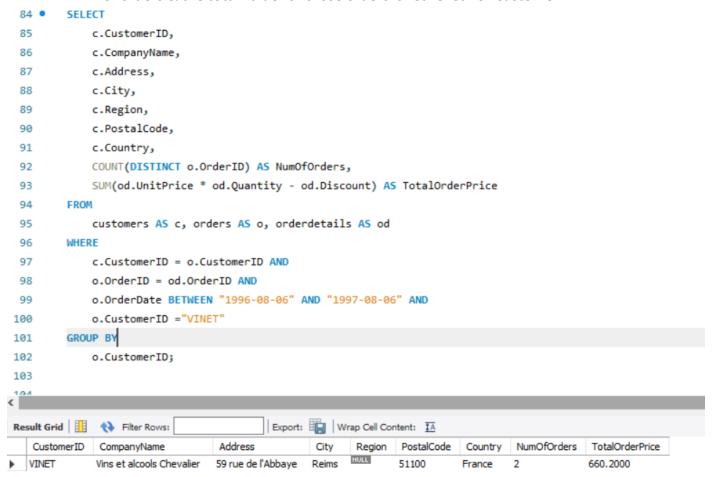
d. Getting the CustomerID, CompanyName, entire Address, calculating the number of orders placed by CustomerID = "VINET" along with the total price for the orders placed by joining tables customers, orders and orderdetails



e. Verifying the above query by multiplying the UnitPrice & Quantity in the below query for CustomerID = "VINET" and getting the same result as the above query where I am using SUM to calculate the total order value.

```
SELECT
 63 •
 64
              c.CustomerID,
              c.CompanyName,
 65
              c.Address,
 66
 67
              c.City,
              c.Region,
 68
              c.PostalCode,
 69
 70
              c.Country,
 71
              COUNT(DISTINCT o.OrderID) AS NumOfOrders,
              od.UnitPrice,
 72
 73
              od.Quantity,
 74
             od.Discount
 75
         FROM
 76
              customers AS c, orders AS o, orderdetails AS od
         WHERE
 77
              c.CustomerID = o.CustomerID AND
 78
              o.OrderID = od.OrderID AND
              o.CustomerID = "VINET"
 80
         GROUP BY
 81
 82
              o.CustomerID;
Result Grid
                                             Export: Wrap Cell Content: IA
               Filter Rows:
   CustomerID
                                                                   PostalCode
                                                                                                     UnitPrice
              CompanyName
                                    Address
                                                            Region
                                                                               Country NumOfOrders
                                                                                                              Quantity
                                                                                                                        Discount
                                                           NULL
 VINET
                                   59 rue de l'Abbaye
                                                    Reims
                                                                   51100
                                                                                                     14.0000
              Vins et alcools Chevalier
                                                                              France
```

f. Now the final query including the limiting factor as the period between which the number of orders & the total value for those orders is retrieved for CustomerID = "VINET"



g. The final query for fetching the Customer information for the given period

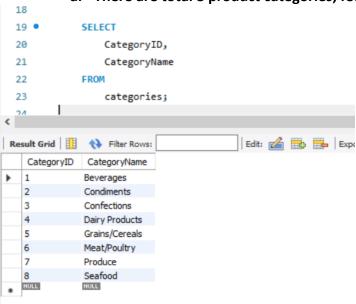
```
SELECT
 84
 85
             c.CompanyName,
             c.Address,
             c.City,
 87
             c.Region,
 88
             c.PostalCode,
 90
             c.Country,
 91
             COUNT(DISTINCT o.OrderID) AS NumOfOrders,
             SUM(od.UnitPrice * od.Quantity - od.Discount) AS TotalOrderPrice
 92
 93
         FROM
             customers AS c, orders AS o, orderdetails AS od
 94
 95
         WHERE
 96
             c.CustomerID = o.CustomerID AND
             o.OrderID = od.OrderID AND
 97
             o.OrderDate BETWEEN "1996-08-06" AND "1997-08-06"
 98
 99
         GROUP BY
100
             o.CustomerID;
Export: Wrap Cell Content: IA
                               Address
                                                       City
                                                                  Region PostalCode Country
                                                                                              NumOfOrders
                                                                                                          TotalOrderPrice
   CompanyName
                                                                 NULL
  Ana Trujillo Emparedados y helados
                               Avda, de la Constitucin 2222
                                                      Mxico D.F.
                                                                         05021
                                                                                                          88,8000
                                                                                    Mexico
                                                                NULL
  Antonio Moreno Taquera
                                                                         05023
                                                                                                          5522.9500
                               Mataderos 2312
                                                      Mxico D.F.
                                                                                    Mexico
  Around the Horn
                               120 Hanover Sq.
                                                      London
                                                                         WA1 1DP
                                                                                                          3974.9000
                                                                 NULL
  Berglunds snabbkp
                               Berguvsvgen 8
                                                                        S-958 22 Sweden 7
                                                                                                          10289.3000
                                                                 NULL
  Blauer See Delikatessen
                               Forsterstr. 57
                                                       Mannheim
                                                                         68306
                                                                                    Germany
                                                                                                          1079.8000
                                                                 NIIII
```

Note:

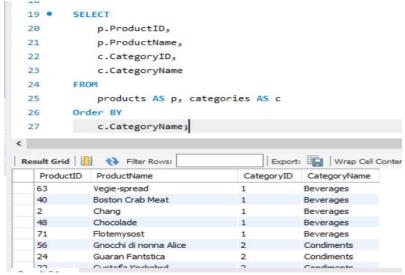
I even cross verified the output by exporting the output of the above queries in excel & manually checked for CustomerID = "VINET" & CustomerID = "PERIC"

2. Product Information:

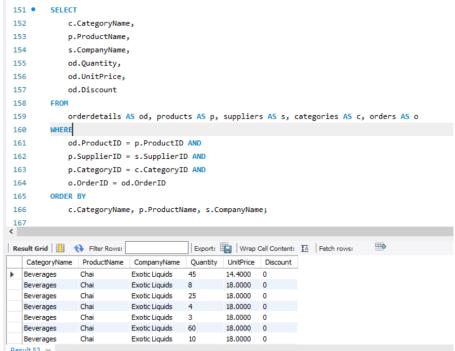
a. There are total 8 product categories, fetching each product category's name



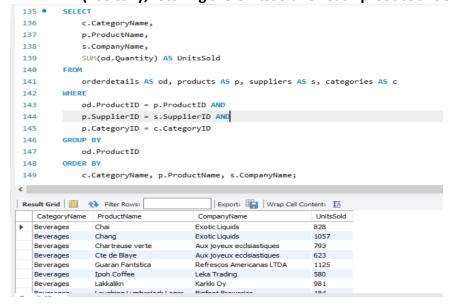
b. Fetching all the products by joining the products & categories tables (Each category has 77 products, totalling to 616 products across all categories)



c. Fetching the category name, corresponding product name, supplier name & their respective quantities, unit price & discount in the placed orders by joining products, categories, suppliers & orderdetails tables



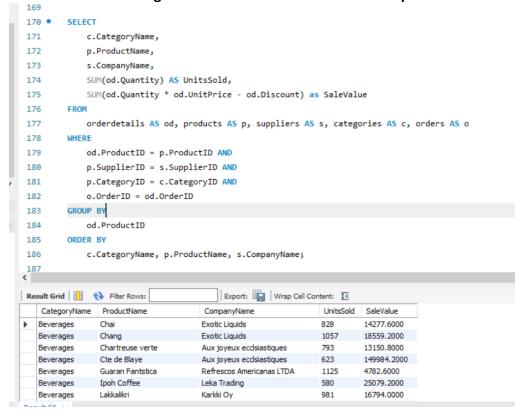
d. Now fetching the sum of the quantities for the above query & grouping it by ProductID (Basically, fetching the UnitsSold for each product under each category)



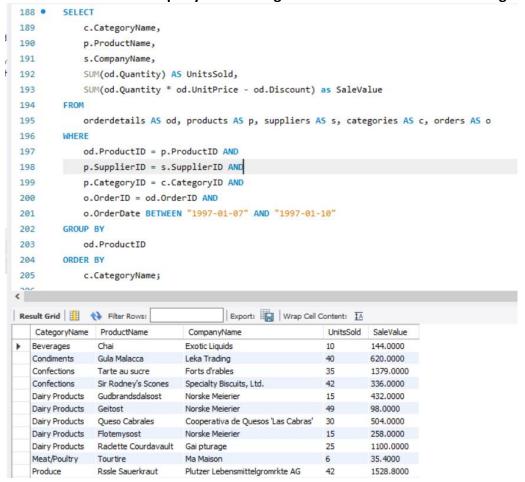
Note:

I calculated in excel the UnitsSold for each product & it matches with the above query's result.

e. Fetching the units sold & sale value for each products without the period constraint



f. The final query for fetching the Product Information for the given period

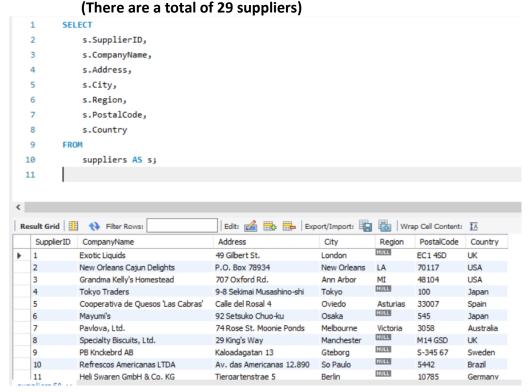


Note:

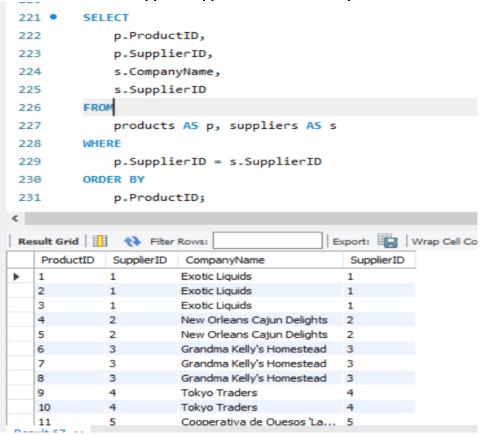
I calculated in excel the SaleValue for ProductName = "Chai" & for ProductName = "Chang", it matches with the above query's result.

3. Supplier Information:

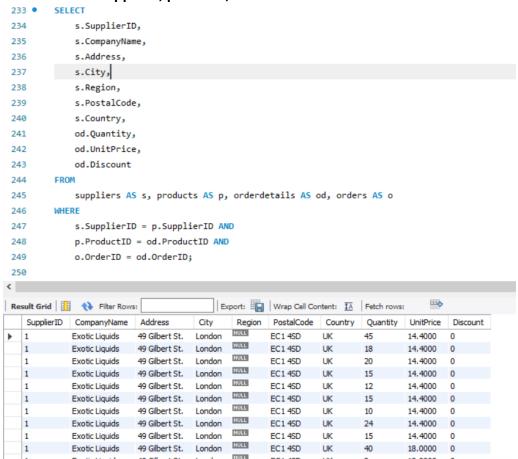
a. Fetching the supplier name and the entire address



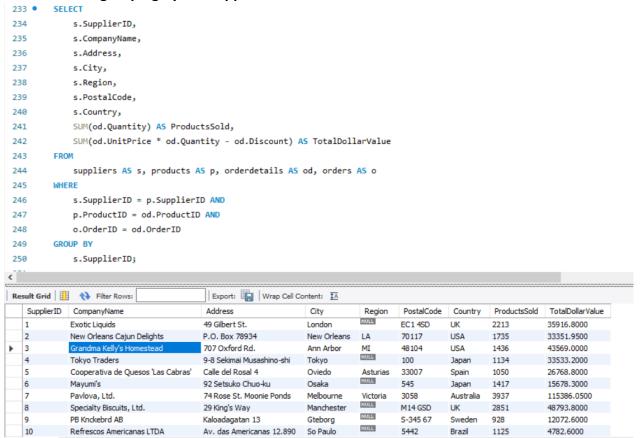
b. Each supplier supplies more than one product



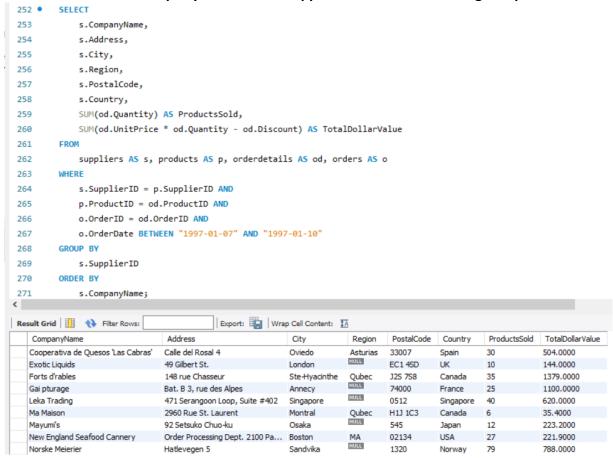
c. Fetching the supplier details along with Quantity, Unit Price & Discount by joining tables suppliers, products, orderdetails & orders



d. Fetching the total number of products sold & its total dollar value for each supplier by using SUM on (Quantity) and the (UnitPrice * Quantity – Discount) respectively & grouping by the SupplierID



e. The final query to fetch the Supplier Information for the given period



Note:

I calculated in excel the ProductsSold & TotalDollarValue for Supplier "Cooperativa de Quesos 'Las Cabras'" and "Gai pturage", it matches with the above query's result.