**RFM Analysis Homework**

In this assignment you will utilize the AdventureWorks2019 database. The database diagram is attached with the assignment.

As part of an ongoing CRM program, you have been asked to prepare a report identifying who the valuable customers are. You are responsible for customers living in France, so you are interested in retrieving information about orders originated from that country.

Although, multiple different options were available, you have decided to implement RFM analysis to get some basic understanding about the customers in your segment of the company. RFM method is a basic tool allowing for a quick assessment of the customer value. RFM abbreviation represents three dimensions of the technique: recency, frequency, and monetary value. These dimensions can be described as:

* Recency – how recently did the customer make a purchase
* Frequency – how frequently does the customer purchase
* Monetary value – how much money did the customer spend

RFM analysis works in the following way:

* Customers are assigned a recency score based on date of most recent purchase or time interval since most recent purchase. This score is based on a simple ranking of recency values into a small number of categories. For example, if you use five categories, the customers with the most recent purchase dates receive a recency ranking of 5, and those with purchase dates furthest in the past receive a recency ranking of 1.
* In a similar fashion, customers are then assigned a frequency ranking, with higher values representing a higher frequency of purchases. For example, in a five category ranking scheme, customers who purchase most often receive a frequency ranking of 5.
* Finally, customers are ranked by monetary value, with the highest monetary values receiving the highest ranking. Continuing the five-category example, customers who have spent the most would receive a monetary ranking of 5.

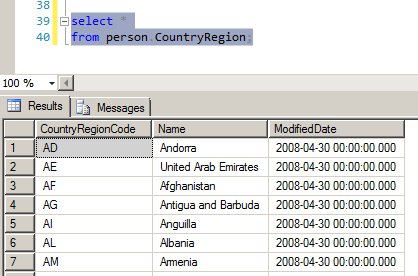
The result is four scores for each customer: recency, frequency, monetary, and combined RFM score, which is simply the three individual scores concatenated into a single value. The "best" customers are those with the highest combined RFM scores.

**Your tasks are as follows:**

1. Identify (list) the tables containing the relevant data.
2. Produce a list of customers living in France. The list should include: CustomerID, FirstName, LastName, Address, City, Zip, Province Name, and country.
3. Produce a list of customers with number of orders placed and dollars spent. The list can include only the CustomerID, Number of Orders, and Total Money Spent columns. Use appropriate aggregate functions.
4. Produce a list of customers that can be utilized for your RFM analysis. The list needs to include CustomerID, first and last name, city, country, number of orders, total dollars spent, and last order date. The list needs to be limited to the search criteria specified above.
5. Export the list to Excel (as shown below). Then, assign scores for each of the RFM dimensions. Sort the output according to the total RFM score (highest first).
   1. HINT: In order to obtain RFM scores you will need to transform your continuous data into bins. The method below describes equal frequency binning, but you are more than welcome to use other methods should you decide to do so. The PERCENTRANK function can be used to calculate the RFM scores:

=ROUNDDOWN(PERCENTRANK.EXC($A$1:$A$10, $A1)\*5, 0)+1

**How to export query results to Excel:**



Click here to select all records in the output. Then right-click and from the context menu select **Copy with Headers**. You can then open Excel and paste the results in.

**Submission guidelines:**

All the queries that you create should be saved as a “.sql” file directly from SSMS. The list of tables for task 1 should be saved in the same .sql file within blocked comments (enclosed within **/\*** and **\*/** symbols).

The results of the RFM analysis should be saved in Excel.

Both files (.sql and .xlsx) should be submitted in D2L.