

Analytical SQL Case Study

Background:

Customers has purchasing transaction that we shall be monitoring to get intuition behind each customer behavior to target the customers in the most efficient and proactive way, to increase sales/revenue , improve customer retention and decrease churn.

You will be given a dataset, and you will be required to answer using SQL Analytical functions you have learnt in the course.

Q1- Using OnlineRetail dataset

- write at least **5 analytical SQL** queries that tells a story about the data
- write small **description** about the business meaning behind each query

Q2- After exploring the data now you are required to implement a Monetary model for customers behavior for product purchasing and segment each customer based on the below groups

Champions - Loyal Customers - Potential Loyalists – Recent Customers – Promising - Customers Needing Attention - At Risk - Cant Lose Them – Hibernating – Lost

The customers will be grouped based on 3 main values

- **Recency** => how recent the last transaction is (**Hint**: choose a reference date, which is the most recent purchase in the dataset)
- **Frequency** => how many times the customer has bought from our store
- **Monetary** => how much each customer has paid for our products

As there are many groups for each of the R, F, and M features, there are also many potential permutations, this number is too much to manage in terms of marketing strategies.

For this, we would decrease the permutations **by getting the average scores of the frequency and monetary** (as both of them are indicative to purchase volume anyway)

Expected outcome

CustomerID text	recency double precision	frequency bigint	monetary double precision	r_score integer	fm_score integer	cust_segment text
13256	14	1	0	5	1	Recent Customers
15118	134	1	0.17	3	1	Promising
12875	143	2	0.34	2	1	About to Sleep
13366	50	1	0.39	5	1	Recent Customers
17752	359	1	0.42	1	1	Lost
14609	72	4	0.49	5	2	Potential Loyalists
15753	304	1	0.55	1	1	Lost
16881	67	1	0.72	5	1	Recent Customers

Label each customer based on the below values

Group name	Recency score	AVG(Frequency & Monetary) score
Champions	5	5
	5	4
	4	5
Potential Loyalists	5	2
	4	2
	3	3
	4	3
Loyal Customers	5	3
	4	4
	3	5
	3	4
Recent Customers	5	1
Promising	4	1
	3	1
Customers Needing Attention	3	2
	2	3
	2	2
At Risk	2	5
	2	4
	1	3
Cant Lose Them	1	5
	1	4
Hibernating	1	2
Lost	1	1

Q3- You are given the below dataset, Which is the daily purchasing transactions for customers.

Cust_Id	Date	Amount
145272	11/5/2019	1.59
145272	11/6/2019	2.98
145272	11/7/2019	2.19
145272	11/8/2019	8.74
1026223	11/3/2019	2
1026223	11/7/2019	33
1026223	11/8/2019	25.5
1767267	11/1/2019	132.69
1767267	11/2/2019	18.64
1767267	11/3/2019	0.4
1767267	11/4/2019	126.33
1767267	11/6/2019	1.92
1767267	11/7/2019	10.07

You are required to answer two questions:

- a- What is the maximum number of consecutive days a customer made purchases?

Expected Output

Cust_Id	Max_Consecutive_Days
145272	4
1026223	2
1767267	4

- b- On average, How many days/transactions does it take a customer to reach a spent threshold of 250 L.E?

Delivery format:

1. File/s (.txt, .sql, .docx) contain your Queries
2. File(.doc) contains your story (it may contain your Queries as well) written in a **well-formatted manner**