SILVERTREE BRANDS BI DEVELOPER ASSESSMENT

# Objective:

The purpose of this assessment is to evaluate your ability to apply RFM (Recency, Frequency, Monetary) analysis to customer datasets, demonstrating both your technical proficiency in SQL and your analytical reasoning in customer segmentation.

The goal is for you to create a set of customer classification columns for each of the customer datasets provided that allocates a behavioural description to each customer. Each customer needs to be grouped according to 3 dimensions namely:

* Recency: Number of days since last order
* Frequency: Number of lifetime orders
* Monetary: Customer lifetime value

The score for each of these three dimensions need to be calculated for each customer and then combined to produce a final column, namely RFM group. You need to use your own intuition to group customer behaviours. The logic used needs to be transparently communicated in a separate doc. The summarised outcome also needs to be shown.

Finally, consider the question sheet provided and answer the questions.

# Resources:

* Customer Dataset 1 & Customer Dataset 2: Two transactional datasets. Each containing customer transaction activity.
* An empty google drive folder for you to provide save your answers to, named “Submissions [Your Name]”

# Desired outcome:

It is expected of you to produce the following materials and add them to your “Submissions” folder in Google Drive.

## Section A:

* One or multiple SQL scripts that were used to group customer behaviour.
* Excel file:
  + One or multiple copies of the resulting table(s) after the SQL scripts were applied.
  + A data dictionary describing the resulting table. Please include a column description, data type & description.

## Section B:

* A Document containing:
  + An aggregated display of the outcome (i.e. number of customers by group)
  + An intuitive demonstration / description of how the rules were applied
  + Answers to the additional questions

# Instructions:

For each dataset provided:

1. Read the input files provided using SQL version or IDE of your choice
   * You will likely need to upload the csv’s to some local or open source DB.
2. Create your own SQL scripts that:
   * Assign each customer a ‘Score’ for R,F & M, either a percentile or absolute value.
   * Bucket combinations of R, F & M and based on these buckets assign each customer one of the groups below. Either use set X, or set Y.
   * NB! - You need to set thresholds for the buckets using your own discretion and reasoning. Likewise you need to choose whether to use Set X or Set Y of customer group names using your own discretion. You will later be asked to justify your choice.

|  |  |
| --- | --- |
| **Set X** | **Set Y** |
| At\_risk | Newbies |
| Cant\_Lose\_Them | Lost\_one\_and\_done |
| Champions | 151 |
| Customer\_Needs\_Attention | Champs |
| Hibernating | Loyals |
| High\_Value\_Sleeping | Recents |
| Loyal\_Customers | Lost\_low |
| New Customer | Lost\_med |
| Ork ami Done | Lost\_high |
| Potential\_Loyalists | Promising |
| Recent Customers | Sleepy\_low |
|  | Sleepy\_med |
|  | Sleepy\_high |
|  | Small\_spenders |
|  | Hibernating\_low |
|  | Hibernating\_med |
|  | Hibernating\_high |
|  | Potential\_loyals |
|  | Sleepy\_one\_and\_done |
|  | Hibernating\_one\_and\_done |

# Additional Questions:

1. Explain why you set the thresholds to the values that you did.
2. Say which set(s) of customer group descriptions you decided to use and explain your reasoning for why you picked that set of descriptions.
3. Do you observe any behavioural differences between Customer Dataset 1 and Customer Dataset 2? Please explain.