**Report on RFM Analysis of Sales Data**

**Introduction to RFM Analysis**

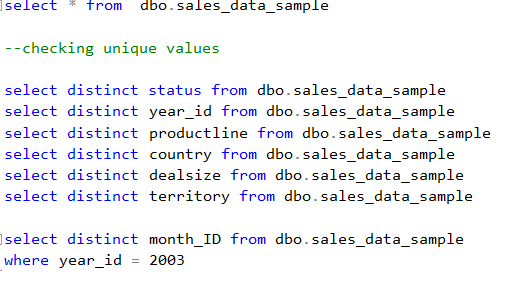
RFM (Recency, Frequency, Monetary) analysis is a data-driven customer segmentation technique that allows businesses to identify which customers are more likely to respond to promotions and which ones are at risk of churn. It categorizes customers based on their transaction history - how recently, how often, and how much did they buy.

RFM (Recency, Frequency, Monetary Value) is crucial for businesses because it segments customers based on their purchase history, enabling personalized marketing strategies. By analyzing how recently and frequently customers buy, and how much they spend, companies can identify and prioritize valuable customers. This insight allows for targeted communication and promotions, enhancing customer loyalty, increasing retention, and maximizing profit. RFM also helps in identifying under-engaged customers, providing opportunities to re-engage them effectively.

**Data Source and Preparation**

The analysis is based on a sales dataset (**dbo.sales\_data\_sample**), which was first explored for unique values across different attributes like status, year, product line, country, deal size, and territory. This preliminary step is essential for understanding the dataset's structure and ensuring data quality.

**Overview of Sales Data**

1. **Basic Data Exploration:**
   * **Queries to check unique values in columns like status, year\_id, productline, country, dealsize, and territory.**
   * **These queries are useful for understanding the distinct categories in the data, which is essential for any further detailed analysis.**
   * ****
2. **Grouping Sales by Product Line, Sales Analysis by Year, Sales Analysis by Deal Size, Best Month for Sales in a Specific Year and Sales in November: Product Line Analysis**

**A screenshot of a computer code

Description automatically generated**

1. **Customer Analysis with RFM (Recency, Frequency, Monetary Value) Model:**
   * **This complex query creates an RFM model to segment customers based on their purchase history.**

**A screenshot of a computer program

Description automatically generated**

**Things done in this SQL quires.**

* **Temporary Table Creation and Cleanup:**
  + **The script starts by dropping the temporary table #rfm if it exists to ensure a clean slate for the analysis.**
* **RFM Calculation:**
  + **The WITH clause creates two Common Table Expressions (CTEs): rfm and rfm\_calc.**
  + **In the rfm CTE:**
    - **The script calculates RFM metrics for each customer.**
    - **MoneteryValue is the total sales per customer.**
    - **AvgMonetaryValue is the average sales per order per customer.**
    - **Frequency is the count of orders per customer.**
    - **Recency is calculated as the number of days since the customer's last order.**
  + **In the rfm\_calc CTE:**
    - **The script uses the NTILE function to divide customers into quartiles (4 groups) based on recency, frequency, and monetary value.**
* **RFM Segmentation:**
  + **The final SELECT statement computes the RFM score and assigns each customer a segment based on their RFM score.**
  + **rfm\_recency+ rfm\_frequency+rfm\_monetary as rfm\_cell calculates a composite score for RFM.**
  + **rfm\_cell\_string concatenates the quartile rankings of recency, frequency, and monetary.**
  + **The CASE statement classifies customers into segments like 'Lost\_Customer', 'New\_customer', 'Potential\_customer', 'Active', and 'Loyal', based on their rfm\_cell\_string.**
  + **This segmentation is based on predefined criteria where each combination of RFM quartile scores is mapped to a customer segment.**
* **Output:**
  + **The result is a list of customers with their RFM scores and segments, which provides insights into customer behavior and value. This information is useful for tailored marketing strategies, customer retention, and overall business strategy.**

1. **Products Sold Together:**
   * **This query identifies which products are often purchased together, which is useful for marketing and sales strategies like bundling or cross-promotion.**

**A screenshot of a computer code

Description automatically generated**

**Things done on this quires.**

* **Initial Setup:**
  + **The query starts with a SELECT statement that aims to list distinct ORDERNUMBERs alongside a string of PRODUCTCODEs representing products sold in each order.**
* **Nested Subqueries and String Aggregation:**
  + **The STUFF function combined with a subquery and FOR XML PATH('') is used to concatenate PRODUCTCODEs into a single string for each order.**
  + **Within the STUFF function, there's a nested subquery that selects PRODUCTCODEs from the dbo.sales\_data\_sample table (p alias) where certain conditions are met.**
  + **These conditions are defined in another nested subquery that filters orders based on the STATUS 'Shipped' and ensures that the order contains exactly 2 items (rn=2). This is done by grouping by ORDERNUMBER and counting the number of occurrences, which is stored in rn.**
* **Final Output:**
  + **The STUFF function then removes the first character (a comma) from the concatenated string of PRODUCTCODEs, making the string neatly formatted.**
  + **The outer query (FROM dbo.sales\_data\_sample s) runs through the sales data sample table, matching each ORDERNUMBER with its corresponding string of PRODUCTCODEs created by the inner queries.**
* **Ordering Results:**
  + **Finally, the results are ordered by PRODUCTCODES in descending order.**

1. **City with Highest Sales in a Specific Country, Best Product in the United States:**
   * **This query identifies the city in the UK with the highest sales, providing a geographical insight into market performance.**
   * **It analyzes sales in the USA, broken down by year and product line, to identify which products are most popular or generate the most revenue in the United States.**

**A screenshot of a computer code

Description automatically generated**

**Conclusion**

The RFM analysis of the sales data reveals critical insights into customer behaviors and preferences. This understanding enables businesses to make strategic decisions in marketing, sales, and customer relationship management to maximize revenue and growth.