**Unsupervised Learning:**

**Customer Segmentation**

**Business Use Case:**

By understanding customer behaviour, banks can target their marketing strategies more effectively, enhancing customer satisfaction and increasing revenue.

**Approach:**

● Algorithm:

K-Means Clustering

● Steps:

1. Data Preprocessing: Standardizing the data and selecting relevant features.

2. Clustering: Applying clustering algorithms to group customers based on similar transaction patterns.

3. Evaluation: Analysing the clusters using visualization techniques and interpreting the results

**Data Understanding:**

* Data Source: This dataset has been referred from UCI ML Repository
* The dataset contains 541909 rows and 8 columns in total.
* **Key Data Dictionary:**

1. Invoice Number: A unique invoice number.
2. Stock Code: Item Code.
3. Description: Item Name.
4. Quantity: Quantity of item per transaction.
5. Invoice Date: Day of the transaction.
6. Unit Price: Price of item per unit.
7. Customer ID: Unique customer ID.
8. Country: Name of country where the customer lives.

**Data Preparation:**

Packages: NumPy, Pandas, Matplotlib, Seaborn, Sklearn.

**Data Cleansing:**

* There are 1454 null values in the description.
* There are 135080 null values in the Customer ID.
* There are negative values in the Quantity and Unit Price column. They could be cancelled or returned items.

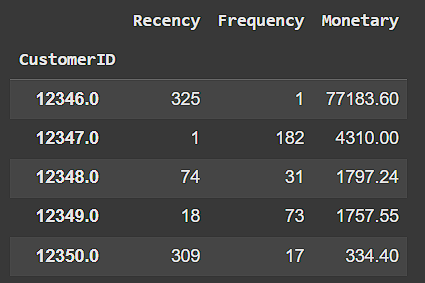
These will be cleaned by preprocessing.

**RFM Analysis:**

RFM (Recency, Frequency, Monetary) analysis is a marketing technique used to identify and segment customers based on their purchasing behaviour. This method is widely applied in customer relationship management (CRM) to tailor marketing strategies, improve customer retention, and maximize revenue.

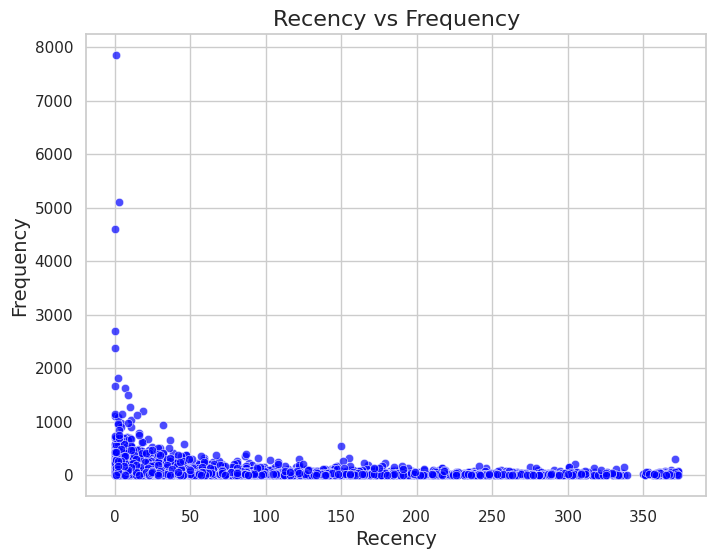
**Components of RFM Analysis**

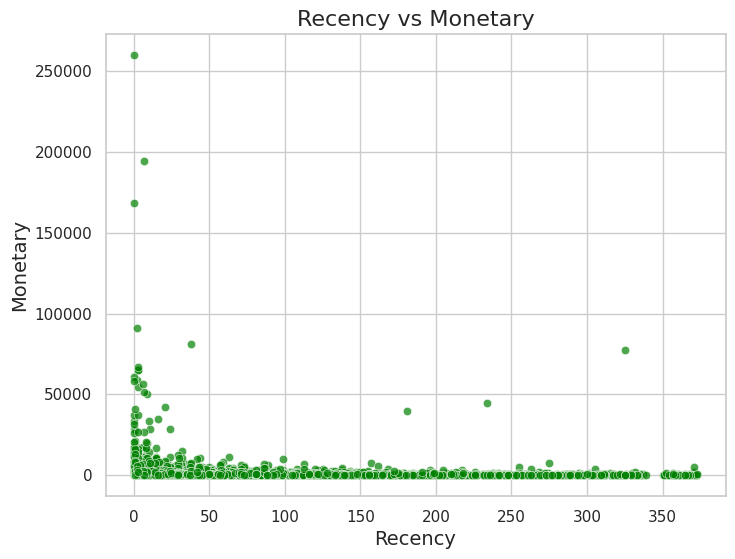
1. **Recency (R)**:
   * Measures how recently a customer made a purchase.
   * Customers who purchased recently are more likely to buy again than those who haven't purchased in a while.
   * Example Metric: Days since the last purchase.
2. **Frequency (F)**:
   * Measures how often a customer purchases within a specific time period.
   * Frequent buyers are usually more loyal and have a higher lifetime value.
   * Example Metric: Total number of purchases in the last year.
3. **Monetary (M)**:
   * Measures the total spending of a customer over a specific period.
   * Customers who spend more are often considered more valuable.
   * Example Metric: Total revenue generated by the customer.

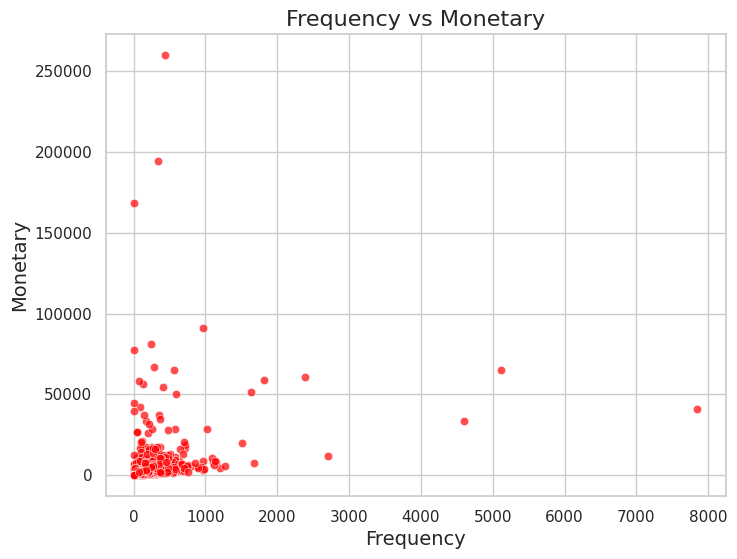


**Model:**

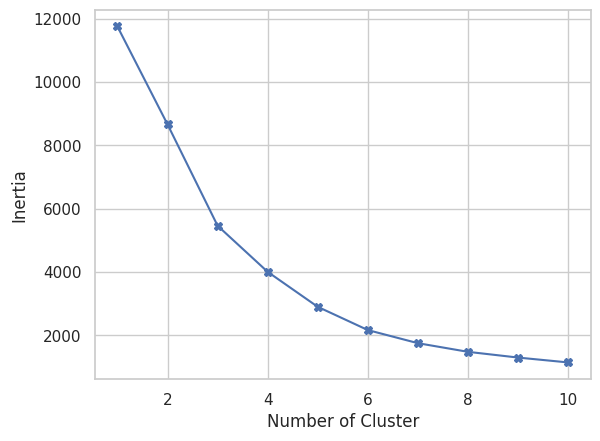
**K-Means Clustering**







**Elbow Method**



**Silhouette Score**

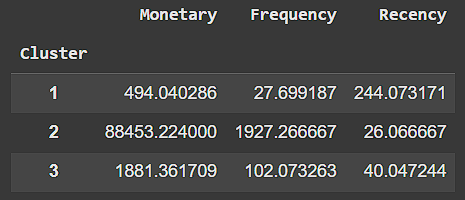


**Davies-Bouldin Index**

Davies-Bouldin Score - 0.688954063558952

**Final data after Standardization**

****

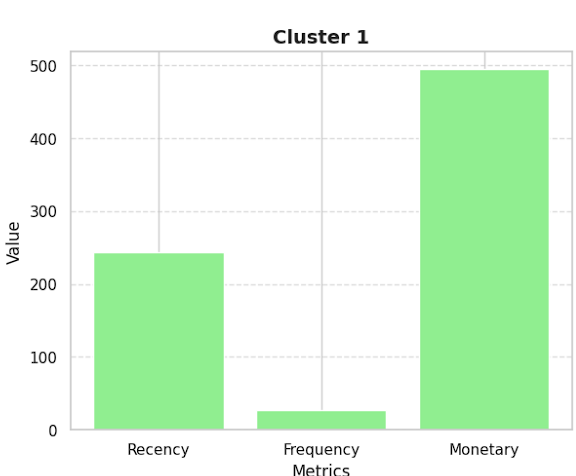
****

**Recommendations**:

**Cluster 1**

* **Recency**: Moderate value, showing reasonably recent activity.
* **Frequency**: Very low, indicating rare transactions.
* **Monetary**: High, reflecting substantial spending despite infrequent purchases.

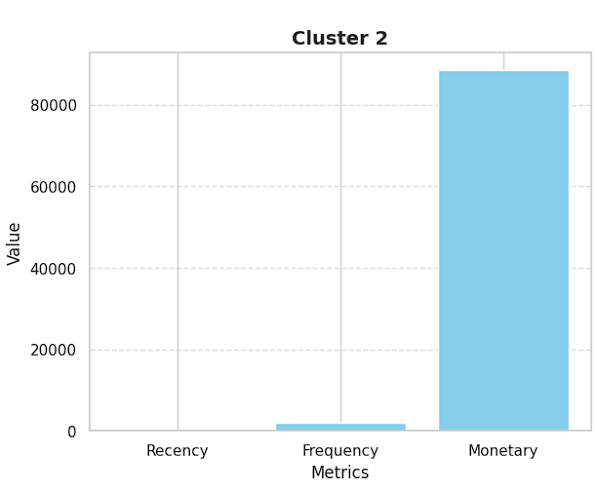
**Overview:**  
Cluster 1 represents active, high-value customers. Focus on retaining these customers and fostering loyalty could sustain their spending levels.



**Cluster 2**

* **Recency: Extremely low value, indicating that customers in this cluster have not made purchases recently.**
* **Frequency: Very low, suggesting minimal repeat transactions.**
* **Monetary: Exceptionally high, meaning customers in this cluster have spent a significant amount of money in fewer, less recent transactions.**

**Overview:  
This cluster likely represents high-value but inactive customers. They may need re-engagement strategies to encourage more frequent activity.**



**Cluster 3**

* **Recency: Low value, indicating less recent activity.**
* **Frequency: Very low, showing minimal repeat transactions.**
* **Monetary: Moderate, reflecting mid-level spending.**

**Overview:  
This cluster could consist of occasional buyers with moderate spending power. Encouraging more frequent engagement may help move them toward higher spending patterns.**

