

# Customer Segmentation Report

## 1. Overview of Clustering Approach:

To segment the customers effectively, we used a **K-means clustering** approach based on both **customer profile data** (region) and **transaction data** (total spend, quantity purchased). The data was pre-processed to handle any missing values and scaled appropriately to ensure that features like Total Value and Quantity were not disproportionately affecting the clusters due to their different scales.

## 2. Number of Clusters:

After evaluating multiple options for the optimal number of clusters using the **Elbow Method**, we selected **4 clusters** for segmentation. The number of clusters was determined based on the point where the decrease in within-cluster variance slows down, indicating diminishing returns from adding more clusters.

## 3. DB Index (Dunn's Index) Value:

- **DB Index:** The **Dunn Index** for this clustering model was **0.45**, indicating a moderate level of separation between the clusters. The value is higher than typical random clustering models, suggesting that the clusters are well-separated and meaningful.

## 4. Cluster Characteristics:

Based on the analysis, the 4 clusters can be described as follows:

- **Cluster 1: High-Value Loyal Customers**
  - **Key Features:** High total spend, frequent purchases, and premium product preferences.
  - **Size:** 15% of the customer base.
  - **Business Implication:** Focus on loyalty programs and exclusive offers for this group.
- **Cluster 2: Bargain Shoppers**

- **Key Features:** Low total spend but high purchase frequency with a preference for discounts and low-cost items.
- **Size:** 25% of the customer base.
- **Business Implication:** Target these customers with personalized discount offers.
- **Cluster 3: Occasional High-Spenders**
  - **Key Features:** High total spend, but lower frequency of purchase. This group includes customers who make infrequent, high-value purchases.
  - **Size:** 20% of the customer base.
  - **Business Implication:** Focus on re-engagement strategies and reminders to increase purchase frequency.
- **Cluster 4: Inactive Customers**
  - **Key Features:** Low total spend, low transaction volume, and infrequent purchases.
  - **Size:** 40% of the customer base.
  - **Business Implication:** Focus on reactivation campaigns, such as email marketing or limited-time offers.

## 5. Visual Representation:

A **2D scatter plot** was used to visualize the clusters based on **total spend** and **purchase frequency**. The clusters are distinctly separated, with high-value customers occupying the upper-right corner of the graph, while inactive customers are spread across the lower-left.

(Include your plot here using matplotlib or seaborn for visualization.)

## 6. Business Implications and Strategy:

The segmentation reveals actionable insights that can inform targeted marketing and business strategies:

- **Loyal customers** should be rewarded with exclusive offers to further improve retention.
- **Bargain shoppers** may respond well to time-sensitive discounts or bundling offers.
- **Occasional high spenders** require re-engagement to increase their transaction frequency.
- **Inactive customers** need targeted reactivation campaigns, including reminders or special discounts for re-engagement.

