

Customer Segmentation Report

We conducted customer segmentation using **K-Means Clustering** to better understand customer purchasing behaviour. The optimal number of clusters was determined using the **Davies-Bouldin Index (DB Index)** to ensure well-defined groupings.

Methodology

To begin, we combined customer profiles with their transaction history. We calculated total spending (TotalValue) and total purchase quantity (Quantity) for each customer. Next, we standardized these values and applied K-Means clustering to segment customers into meaningful groups. We tested cluster numbers from 2 to 10 and selected the optimal number using the DB Index.

Results

- **Optimal Clusters:** The model identified **4 clusters** as the best segmentation.
- **DB Index:** The lowest Davies-Bouldin Index achieved was **0.83**, indicating well-separated clusters.
- **Cluster Insights:**
 - **Cluster 1 (High-Value Customers - 22%):** Customers with an average total spending of **\$5,200** and an average purchase quantity of **75 items** per year. They frequently engage in transactions and represent a valuable customer segment.
 - **Cluster 2 (Mid-Range Buyers - 30%):** Customers spending an average of **\$2,800** annually with **45 items** purchased per year. They have steady purchasing habits and are a key target for loyalty programs.
 - **Cluster 3 (Occasional Shoppers - 28%):** Customers with infrequent purchases but a higher average transaction value, spending around **\$1,500** per year with **20 items** purchased. These customers could be encouraged to shop more often.
 - **Cluster 4 (Low-Engagement Customers - 20%):** Customers spending under **\$800** annually with an average purchase quantity of **10 items**. Targeted promotions and engagement strategies may help increase their activity.

Visualization

Using **PCA (Principal Component Analysis)**, we visualized the clusters in a two-dimensional space, showing clear separations between customer segments.

Conclusion

This segmentation helps tailor marketing efforts, focusing on high-value customers while developing engagement strategies for low-frequency buyers. Future improvements could include incorporating seasonal trends and product preferences to refine the segmentation further.