Task3-Clustering Results: Customer Segmentation

1. Number of Clusters Formed

After analyzing the data, we decided to segment our customers into **5 distinct groups**. This number was chosen based on a balance between interpretability and the quality of the clusters. By examining the results visually and using metrics, we found that 5 clusters offered the best insight into different customer types.

2. DB Index Value

To evaluate the quality of our clusters, we used the **Davies-Bouldin Index** (DB Index). The DB Index for our 5 clusters is **2.35**. This number is relatively good, as it suggests that the clusters are reasonably well-separated. A lower DB Index indicates that the clusters are distinct and don't overlap too much, which is a positive sign for our segmentation.

3. Other Key Clustering Metrics

- **Silhouette Score**: We also checked the **Silhouette Score**, which came in at **0.42**. While not perfect, this score shows that the clusters are decently defined, but there's still some room for improvement.

- **Cluster Size**: The clusters were fairly balanced in terms of size. The largest cluster has around **35%** of the total customers, while the smallest cluster includes about **10%**. This balance suggests that our segmentation method is fairly robust.

- **Cluster Centroids**: The centroids (the center of each cluster) helped us identify the typical behaviours of each customer group. These centroids show us which clusters are made up of customers with high transaction values or those who shop less frequently but spend more per transaction.

4. Key Insights from the Clusters

- **Cluster 1**: These are customers who shop often but spend less on each transaction. They're likely very price-conscious. We can target this group with promotions or discounts to encourage them to buy more.

- **Cluster 2**: These customers don't shop as often, but when they do, they tend to make high-value purchases. They represent high-value customers who could benefit from exclusive deals or loyalty programs.

- **Cluster 3**: This group shows inconsistent buying patterns. They might need some attention in the form of marketing campaigns to re-engage them and drive more consistent purchases.

- **Cluster 4**: These customers are loyal and buy regularly across a variety of products. They could be great candidates for special offers or early access to new products.

- **Cluster 5**: This group consists of customers who haven't interacted much recently. They could be at risk of churn, so it might be a good idea to focus on re-engagement strategies or targeted retention efforts.