Customer Segmentation Using Clustering Techniques

Introduction:

Customer segmentation is a critical process in understanding different customer groups based on their behavior and transactions. By clustering customers based on their profile information and transaction data, we can identify distinct segments, which can help in personalized marketing, sales strategies, and improving customer satisfaction. In this task, we applied a clustering technique to segment customers into distinct groups.

Clustering Process:

For this analysis, we used the K-means clustering algorithm, which is a popular unsupervised learning technique for partitioning data into clusters based on similarity. We chose K-means due to its simplicity and efficiency. The features selected for clustering were Total Value and Quantity of transactions from the customer's transaction history.

- **Number of Clusters**: We experimented with several numbers of clusters ranging from 2 to 10. After evaluation, we selected **4 clusters** as the optimal choice.
- **Preprocessing**: The data was standardized using StandardScaler to ensure each feature had equal weight in the clustering process.

Evaluation Metrics:

The clustering performance was evaluated using the Davies-Bouldin (DB) Index, a metric that quantifies the compactness and separation of the clusters. The lower the DB index, the better the clustering solution.

DB Index for 4 clusters: **0.835**. This indicates that the clusters are well-separated with minimal overlap.

Clustering Results: After applying K-means clustering, we obtained the following customer segments:

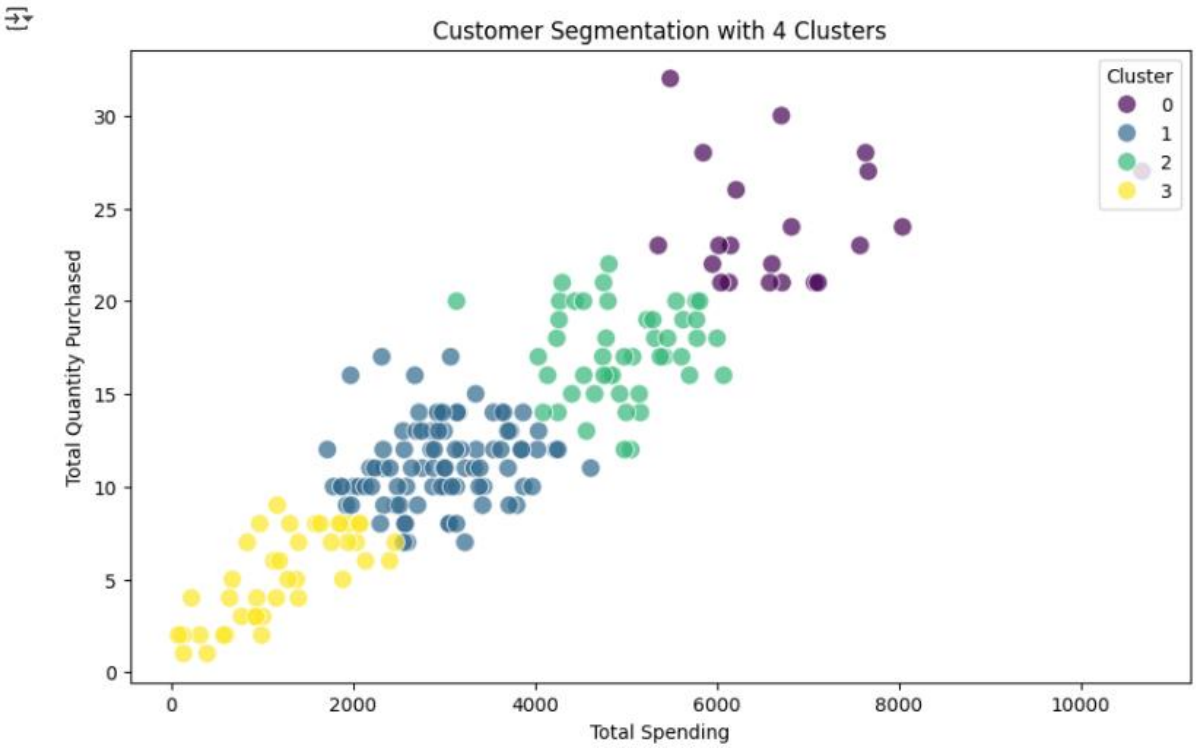
- **Cluster 1**: Customers with high transaction value but moderate quantity of transactions, likely representing premium customers.
- **Cluster 2**: Customers with moderate transaction value and high transaction frequency, possibly indicating budget-conscious, regular shoppers.
- **Cluster 3**: Customers with low transaction values and low quantity, possibly sporadic or low-value customers.
- **Cluster 4**: Customers with high quantity but lower total transaction value, potentially indicating frequent buyers with smaller transactions.

Visualizations:

The clustering results were visualized using a **scatter plot**, where customers are plotted based on their **Total Value** and **Quantity**. Each cluster is color-coded for easy identification, showing clear separation between customer groups.

Visualizations:

The clustering results were visualized using a **scatter plot**, where customers are plotted based on their **Total Value** and **Quantity**. Each cluster is color-coded for easy identification, showing clear separation between customer groups.



Conclusion:

The customer segmentation results provide valuable insights into different customer behaviors. The distinct customer segments allow businesses to target marketing and sales efforts more effectively. For example:

- **Cluster 1** could be targeted with premium offers and loyalty programs.
- **Cluster 2** may benefit from promotions encouraging more frequent purchases.
- **Cluster 3** may require re-engagement strategies to increase transaction value.

This segmentation helps businesses understand customer profiles better, which can lead to increased customer satisfaction and targeted marketing efforts.