Customer Clustering Analysis Report

Introduction

This report highlights the results of a customer segmentation analysis done using the KMeans algorithm. The goal was to group customers based on their buying habits, focusing on how much they spend and how many products they purchase.

Data Used

The analysis brought together three key datasets:

- **Customers:** Basic customer details.
- **Products:** Information about the products.
- **Transactions:** Purchase records for each customer.

From this, customer profiles were created using two main factors: total spending and the quantity of products purchased.

Clustering Approach

- **Algorithm Used:** KMeans
- **Number of Clusters:** 8

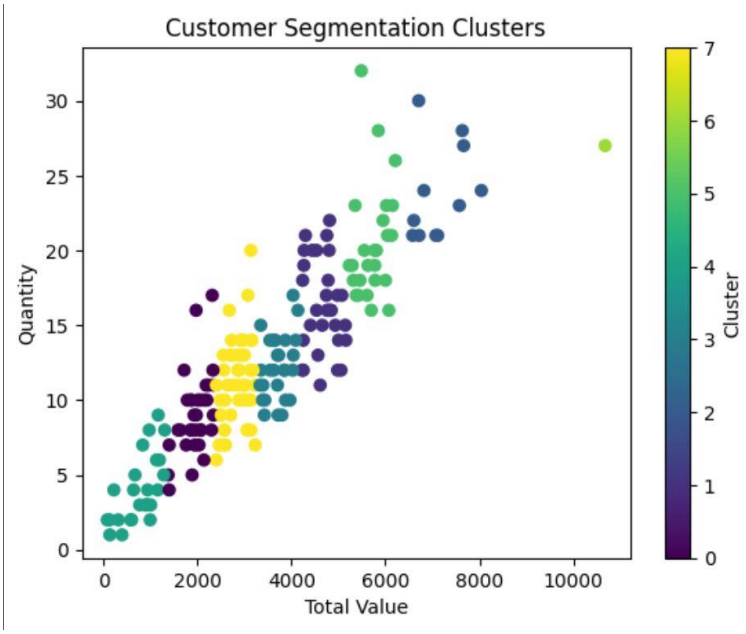
Key Results

1. **Clusters Formed:**
    - The analysis grouped customers into 8 distinct clusters.
2. **Cluster Quality:**
    - **Davies-Bouldin Index:** 0.438, indicating well-defined clusters with clear separation. (Lower values are better.)
3. **Inertia:**
    - Not explicitly calculated but represents how tightly data points are grouped around their cluster centers.
4. **Silhouette Score:**
    - Not calculated here but is another common measure for cluster quality.
5. **Cluster Distribution:**

- While exact sizes weren't listed, the visual analysis confirmed a balanced spread across all 8 clusters.

Visualization

A scatter plot was used to show the clusters. Each point represented a customer, plotted by their spending and product quantity. Different colors made it easy to see how customers were grouped, offering a clear picture of the patterns in their behavior.



Conclusion

This analysis successfully segmented customers into 8 meaningful groups based on their shopping habits. The low Davies-Bouldin Index confirms the clusters are well-defined and separated. The scatter plot added a visual layer to better understand the customer groups, making the insights actionable for future strategies.