Customer Segmentation Report

Objective:
The goal of this analysis was to segment customers based on their profile and transaction information using clustering techniques. This segmentation helps businesses understand customer behavior, tailor marketing strategies, and optimize customer experiences.

Methodology:

1. **Data Aggregation:**

   o Data was grouped at the customer level.

   o Key metrics calculated included:

      ▪ **Average Price (Price_y):** The mean price of items purchased by each customer.

      ▪ **Total Transaction Value (TotalValue):** The sum of all purchase amounts per customer.

      ▪ **Total Quantity Purchased (Quantity):** The total number of items bought by each customer.

2. **Data Scaling:**

   o Since clustering algorithms are sensitive to scale, the aggregated metrics were scaled to ensure uniformity.

3. **Clustering:**

   o KMeans clustering was applied with 4 clusters. This algorithm groups customers into clusters based on their purchasing patterns and transaction details.

4. **Evaluation Metric:**

   o The **Davies-Bouldin Index (DBI)** was used to assess cluster quality. A lower DBI indicates better-defined clusters with less overlap between groups.

Results:

- The calculated **Davies-Bouldin Index** was **1.1218**.

- This score suggests moderately well-defined clusters, indicating that customers within the same cluster exhibit similar purchasing behaviors.

Visualization:

- A scatter plot was generated to visually represent the clustering results. Each data point represents a customer, with colors indicating their assigned cluster.

- The visualization highlights the distinct groupings and helps interpret customer behavior patterns across clusters.

Observations:

1. **Cluster 0 (Purple):**
   - Concentrated around the center with slight dispersion.
   - Represents customers with moderate spending and transaction frequency.

2. **Cluster 1 (Teal/Green):**
   - Spread more towards the negative side of the X-axis.
   - Likely includes customers with lower spending and fewer purchases.

3. **Cluster 2 (Blue):**
   - Positioned centrally with a moderate spread.
   - These customers might have average spending and balanced purchasing patterns.

4. **Cluster 3 (Yellow):**
   - Distributed towards the upper-right region.
   - Represents high-value customers with frequent purchases and higher transaction values.

Insights & Business Implications:

- **Targeting High-Value Customers (Cluster 3):**
   - Businesses can focus on retaining these customers through loyalty programs and exclusive offers.

- **Engaging Low-Spending Customers (Cluster 1):**
   - Strategies like discounts, promotions, or personalized recommendations can encourage them to spend more.

- **Understanding Moderate Spenders (Clusters 0 & 2):**
   - These customers can be converted into high-value clients with targeted marketing.

Conclusion:

The clustering process provided valuable insights into customer segmentation. This analysis can guide targeted marketing campaigns, improve customer retention, and enhance overall business strategies.