Customer Segmentation Summary Report

1. Overview

This report reports the results of customer segmentation, applying clustering to the eCommerce transactions dataset. The purpose is to segment the customers into distinct groups based on the behaviour in making transactions for the meaning to be deduced.

2. Methodology

- The customer profile data were first merged with their transaction history data.
- Features employed for the task of clustering were:
    - Total Spent
    - PurchaseCount
- The data was standardized by employing StandardScaler
- K-Means was applied with varying number of clusters as 2-10.
- The optimal number of clusters was decided using the Davies-Bouldin (DB) Index.

3. Results:

- The best clustering solution was achieved at K = X clusters (which is when the DB Index was at its lowest value Y).
- DB Index Value: Y. Lower the values, the better the clustering.
- Other Metrics
    - Within-Cluster Sum of Squares (WCSS): It measures the compactness of clusters.
    - Silhouette Score: It evaluates the separation of clusters.

4. Key Business Insights:

- High-Value Customers: Customers in this category often make purchases with a high amount of overall spending. Targeted loyalty programs and premium membership offers can drive them even more.
- Moderate Spenders: These spenders shop regularly but don't spend as much per purchase. Offering them personal bundles or discounts will help them increase their purchase value.
- Low-Income Customers: Such customers have a low or minimal transaction. Tailor-made recommendations, discounts, and advertisements may help the business retain and change over low-income customers.
- Dormant Users: Some users signed up but not performing a lot of transactions or at all. Re-engagement campaigns with welcome offers, reminders, etc. may re-engage them.
- Growth Opportunities: By considering the trends related to high-value customers and moderately spent users, businesses may identify ways to grow mid-tier customers to spend more, thus driving higher revenues.

5. Conclusion & Next Steps:

- The clusters can be used to personalize promotions, enhance customer retention, and optimize business strategies.
- Further refinement can be done by incorporating additional features like product categories or regional demographics.

Appendix

- Visualization: A scatter plot of TotalSpent vs PurchaseCount was generated to illustrate customer distribution across clusters.
- Cluster assignments are stored in Customer_Clusters.csv for further analysis.