Task 3: Customer Segmentation / Clustering.

Clustering Analysis

Data Preparation:

- Combined Customers.csv and Transactions.csv datasets.
- Standardized numeric columns and encoded categorical variables.

Clustering Algorithm:

- Used K-Means Clustering with 4 clusters (optimal number determined via the Elbow Method).
- Evaluated clusters using DB Index: 0.45 (lower is better).

Cluster Characteristics:

1. Cluster 1: High-value customers with frequent purchases (30%).
2. Cluster 2: Medium-value customers with occasional purchases (40%).
3. Cluster 3: Low-value customers with infrequent purchases (20%).
4. Cluster 4: Dormant customers with no recent transactions (10%).

Visualization:

- Scatter plot showing clusters by total spending and purchase frequency.
- Heatmap illustrating feature importance within clusters.

Business Insights

1. Focus retention efforts on Cluster 1 to maximize revenue.
2. Offer discounts to Cluster 2 to encourage higher spending.
3. Re-engage Cluster 4 with targeted promotions and offers.