Clustering Results Report

1. Objective

To perform customer segmentation using clustering techniques, evaluate clustering performance using the Davies-Bouldin Index (DB Index), and visualize the resulting clusters. The segmentation will help in understanding customer behavior and identifying distinct customer groups for targeted marketing and strategy.

2. Methodology

Data Preparation

- **Three datasets were used:**
  - Customers.csv
  - Transactions.csv
  - Products.csv

- **The datasets were merged to create a comprehensive customer profile:**
  - Total spend (total_spend)
  - Number of transactions (num_transactions)
  - Recency of last transaction (recency)
  - Product preferences (quantities of product categories)

Feature Standardization

Features were standardized using StandardScaler to normalize scales for clustering.

Clustering Algorithms

Four clustering algorithms were tested:
1. KMeans
2. Agglomerative Clustering
3. DBSCAN
4. Gaussian Mixture Model

Evaluation Metrics

The following metrics were used to evaluate the clustering results:
- Davies-Bouldin Index (DB Index): Lower values indicate better clustering.

Visualization

The following visualizations were used:
- DB Index Plot: Evaluates clustering performance for 2 to 10 clusters.
- PCA Visualization: Projects clusters into 2D space for visual interpretation.

3. Clustering Results

Optimal Model

The KMeans algorithm with 10 clusters achieved the best Davies-Bouldin Index of 1.33.

DB Index Values

Algorithm	Number of cluster	Db index	
K-means	10	1.33	
Agglomerative	7	1.54	
Gaussian Mix	3	2.01	

Key Insights

1.Number of Clusters Formed
- The KMeans algorithm segmented customers into 10 distinct clusters based on their transaction history, spending behavior, and product preferences.
2. Cluster Characteristics
- Each cluster represents customers with unique profiles, such as high spenders, frequent buyers, or specific product category preferences.
3.Evaluation Metrics
- DB Index: 1.33 (KMeans, k=10)

5. Recommendations

1. Targeted Marketing Campaigns: Use customer segments to design tailored campaigns (e.g., loyalty programs for high-spending clusters, upselling for low-spending clusters).
2. Product Recommendations: Analyze product category preferences within clusters for personalized products recommendations.
3. Region-Specific Strategies: Examine regional distribution within clusters to optimize marketing strategies across geographies.

6. Deliverables

1. Clustering Results Report:
- Number of clusters formed: 10
- DB Index value: 1.33