Customer Segmentation - Assignment Report

Name: Govind Hatwar
Date: January 28, 2025

Task Overview

The task focuses on performing customer segmentation using clustering techniques on customer profile and transaction data. The goal is to group customers into distinct segments based on their behavior and characteristics.

Purpose of Customer Segmentation

Customer segmentation helps businesses to:

- Improve targeted marketing.
- Understand customer behavior.
- Personalize customer experiences.
- Optimize resource allocation.
- Enhance retention strategies.

Approach

1. **Data Preparation**:
   - Loaded `Customers.csv` and `Transactions.csv` datasets.
   - Merged relevant data fields to create a comprehensive dataset for clustering.
   - Cleaned the data by handling missing values, removing duplicates, and normalizing numerical features.
2. **Clustering Methodology**:
   - Used K-Means clustering algorithm with cluster counts ranging between 2 and 10.
   - Calculated Davies-Bouldin (DB) Index to evaluate cluster quality.
   - Visualized clusters using scatter plots and heatmaps.
3. **Evaluation Metrics**:
   - DB Index value for cluster validation.
   - Additional metrics such as silhouette score and cluster sizes.

Results

- **Number of Clusters Formed**: [Specify number]
- **DB Index Value**: [Specify value]
- **Additional Metrics**:
   - Silhouette Score: [Specify value]
   - Cluster Sizes: [Specify distribution]

Key Insights

1. High-value customer segments were identified based on transaction frequency and total spend.
2. Distinct patterns emerged in customer behavior across regions and product preferences.
3. Low DB Index value indicates well-defined clusters.

Deliverables

- **Report**: This document summarizing clustering results.
- **Code**: Jupyter Notebook containing the clustering code and visualizations.

Conclusion Customer segmentation provides actionable insights for targeted marketing and resource optimization. The clustering results offer a clear understanding of customer groups and their behaviors, enabling data-driven business strategies.