Customer Segmentation/Clustering Report

1. Overview

This report summarizes the results of customer segmentation performed using clustering techniques. The analysis is based on customer profile data (Customers.csv) and transaction data (Transactions.csv). The goal was to group customers into meaningful segments and evaluate the clustering performance using relevant metrics.

2. Clustering Results

2.1 Number of Clusters

The optimal number of clusters was chosen to be 4.

2.2 Evaluation Metrics

Davies-Bouldin Index (DB Index): 0.8549946326430091

The DB Index measures the compactness and separation of clusters. A lower value indicates better clustering. A value of 0.8549946326430091 suggests reasonably good clustering.

3. Cluster Profiles

The customers were segmented into 4 clusters based on their behavior and characteristics. Below is a brief description of each cluster:

Cluster 0: These customers have moderate spending and a high number of transactions. They frequently engage with the platform and buy a variety of products from different categories.

Cluster 1: This group consists of low spenders with very few transactions. They are likely inactive users or customers who only make occasional purchases, primarily focusing on essential items.

Cluster 2: Customers in this cluster make high-value purchases but have a low number of transactions. They tend to buy expensive products, possibly focusing on premium or luxury goods.

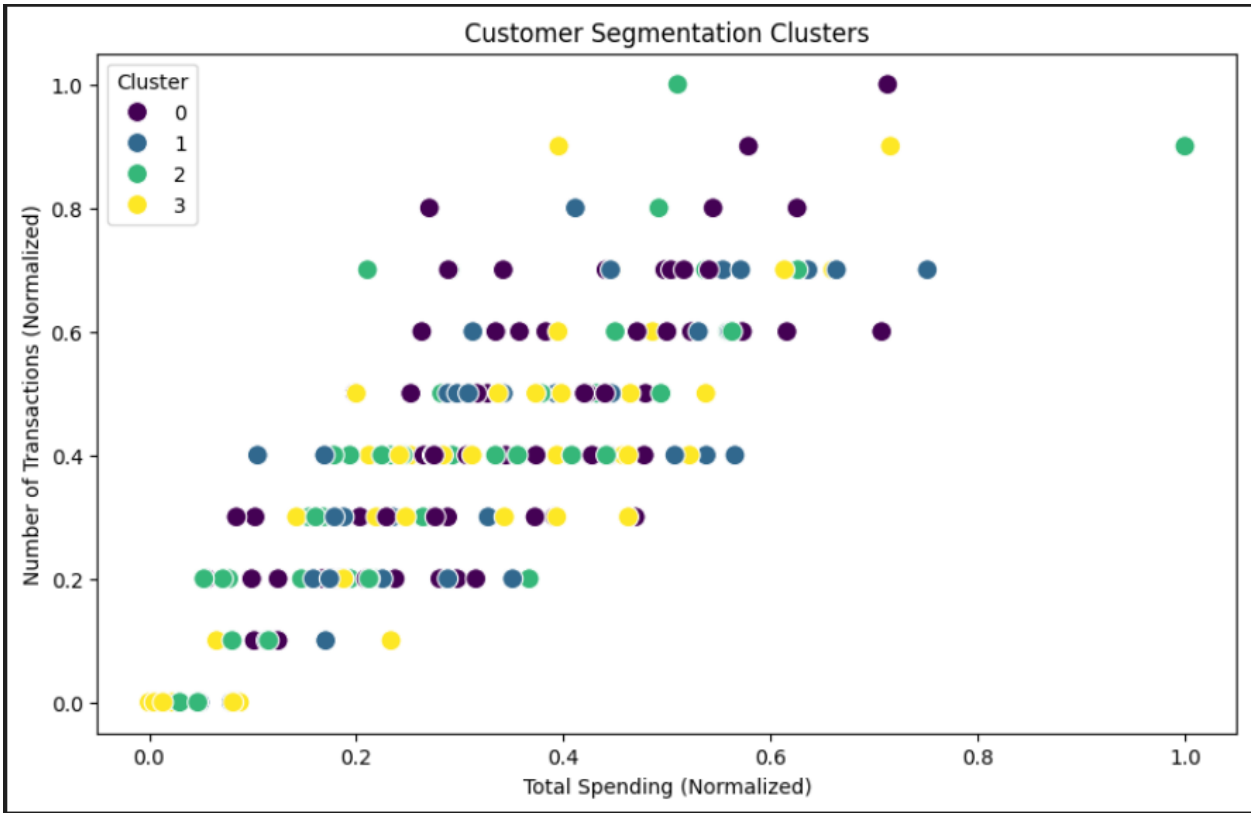
Cluster 3: This segment represents frequent buyers who also spend significantly. They are among the most engaged customers, regularly making purchases across different product categories.

This segmentation helps in understanding different customer behaviors, allowing for targeted marketing strategies, loyalty programs, and personalized recommendations.

4. Visual Representation of Clusters

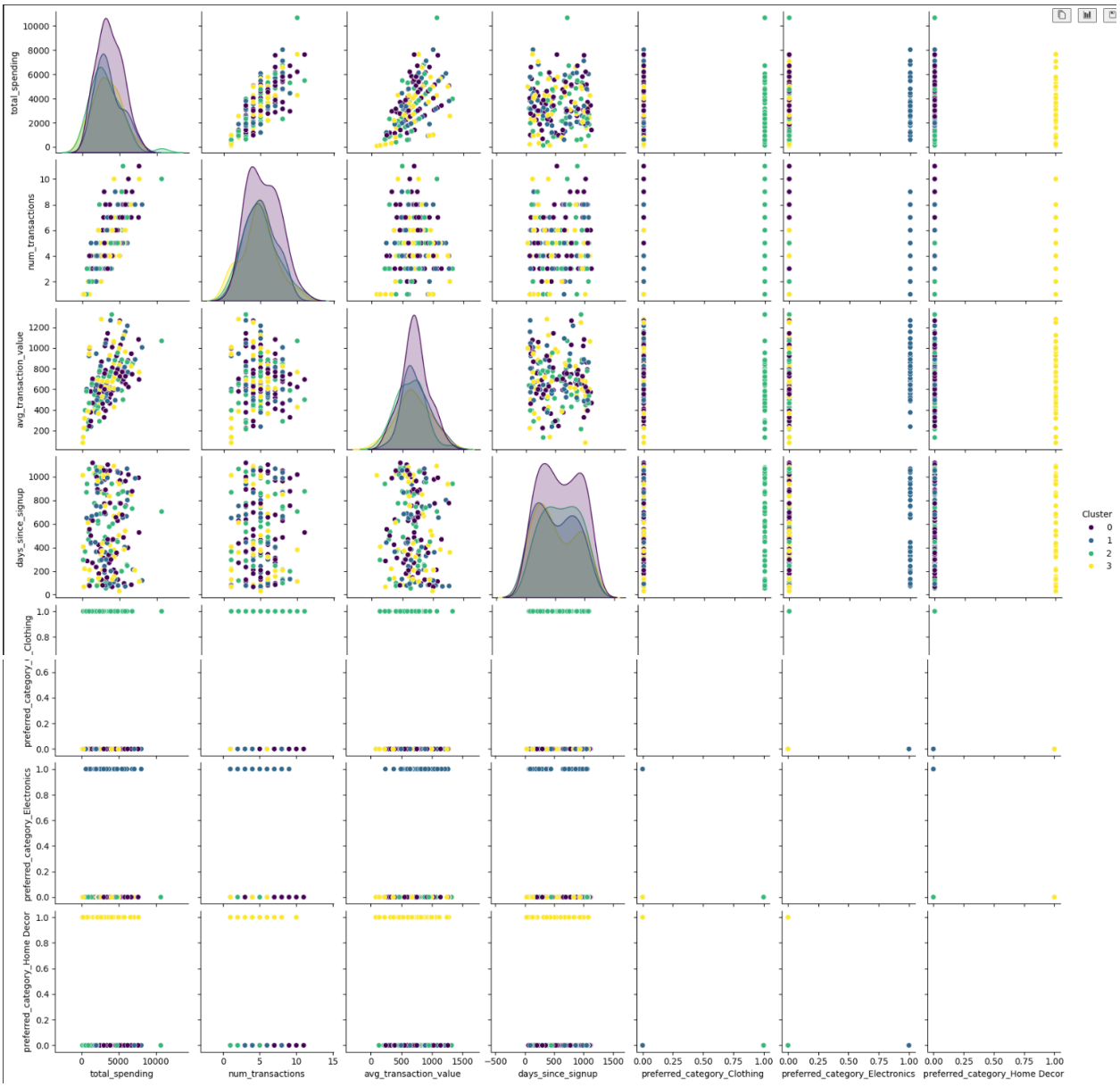
4.1 Scatter Plot of Clusters

A scatter plot of the first two features (e.g., total_spending and num_transactions) was used to visualize the clusters.



4.2 Pair Plot

A pair plot was created to visualize the relationships between features and clusters.



5. Deliverables

5.1 Clustering Results

The clustering results, including the CustomerID and their assigned cluster, are saved in the file: Aaditya_Vijayvargiya_Clustering.csv.

5.2 Jupyter Notebook / Python Script

The complete code for data preparation, clustering, evaluation, and visualization is provided in the Jupyter Notebook: Aaditya_Vijayvargiya_Clustering.ipynb.

5.3 Customer Segmentation/Clustering Report

A report is also provided defining clusters, showing plots etc. Named as Aaditya_Vijayvargiya_Clustering.pdf

6. Conclusion

The customer segmentation analysis successfully grouped customers into 4 clusters based on their behavior and characteristics. The clustering performance was evaluated using the Davies-Bouldin Index (0.85), indicating reasonably good clustering. The results and visualizations provide actionable insights for targeted marketing and customer engagement strategies.