

Report on Clustering

1. Introduction

Customer segmentation is a critical process in understanding the diverse customer base of a business. By grouping customers based on shared characteristics and behaviors, businesses can tailor their marketing strategies, improve customer experience, and drive sales. This report details the results of customer segmentation performed using clustering techniques on customer profile and transaction data.

2. Methodology

2.1 Data Preparation

Two datasets were utilized:

- **Customers.csv:** Contains customer profile information including CustomerID, CustomerName, Region, and SignupDate.
- **Transactions.csv:** Contains transaction details including TransactionID, CustomerID, ProductID, TransactionDate, Quantity, TotalValue, and Price.

2.2 Clustering Approach

- The K-Means algorithm was chosen for clustering due to its efficiency and scalability.
- The number of clusters was varied between 2 and 10.
- Clustering metrics, including the Davies-Bouldin (DB) Index and Silhouette Score, were calculated to evaluate the quality of the clusters.

2.3 Dimensionality Reduction

- Principal Component Analysis (PCA) was used to reduce the feature dimensions to 2 for visualization purposes.

3. Results

3.1 Clustering Metrics

The table below summarizes the clustering metrics:

Number of Clusters	Davies-Bouldin Index	Silhouette Score
2	0.91	0.54
3	0.84	0.48
4	0.78	0.46
5	0.80	0.45
6	0.85	0.44
7	0.92	0.41
8	0.95	0.39
9	1.01	0.37
10	1.10	0.35

- The optimal number of clusters was determined to be **4**, as it resulted in the lowest DB Index (0.78).

4. Conclusion

Key Insights:

- Customers in **Cluster 3** exhibit high spending with recent transactions, suggesting they may represent loyal, high-value customers.

- Customers in **Cluster 0** spend significantly but have fewer transactions, which may represent infrequent but high-value buyers.
- **Cluster 2** customers are characterized by high transaction counts but low spending and older transactions, potentially indicating price-sensitive or dormant customers.

Recommendations:

1. **Cluster 3:** Implement loyalty programs to retain these high-value customers.
2. **Cluster 0:** Encourage more frequent purchases with targeted promotions.
3. **Cluster 2:** Re-engage dormant customers with discounts or tailored offers.