Customer Segmentation Report

Your Name

January 27, 2025

Introduction

This report details the customer segmentation analysis conducted using KMeans and DBSCAN clustering algorithms. The analysis utilized three key features: Region, Total Value, and Quantity. The goal was to identify distinct customer segments to enhance targeted marketing strategies and improve customer engagement.

KMeans Clustering

Methodology

The KMeans algorithm was employed to partition the dataset into clusters. The Elbow method was used to determine the optimal number of clusters, which was found to be 5.

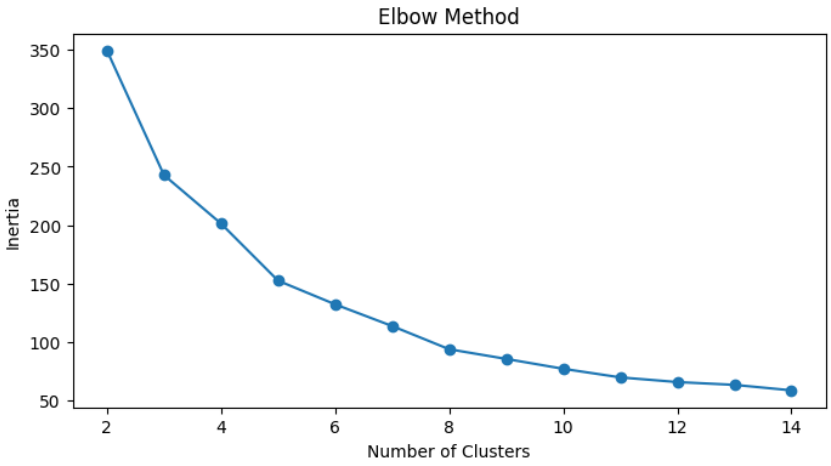

Figure 1: Elbow Method for Determining Optimal Clusters

Clustering Metrics

The clustering performance was evaluated using the Davies-Bouldin Index and Silhouette Score:

- **Davies-Bouldin Index:** 0.9155
- **Silhouette Score:** 0.3640

These metrics suggest moderate cluster separation, indicating potential overlap between clusters.

Visualization

The 3D visualization below illustrates the customer segmentation achieved with KMeans, highlighting the distribution of customers across the identified clusters.

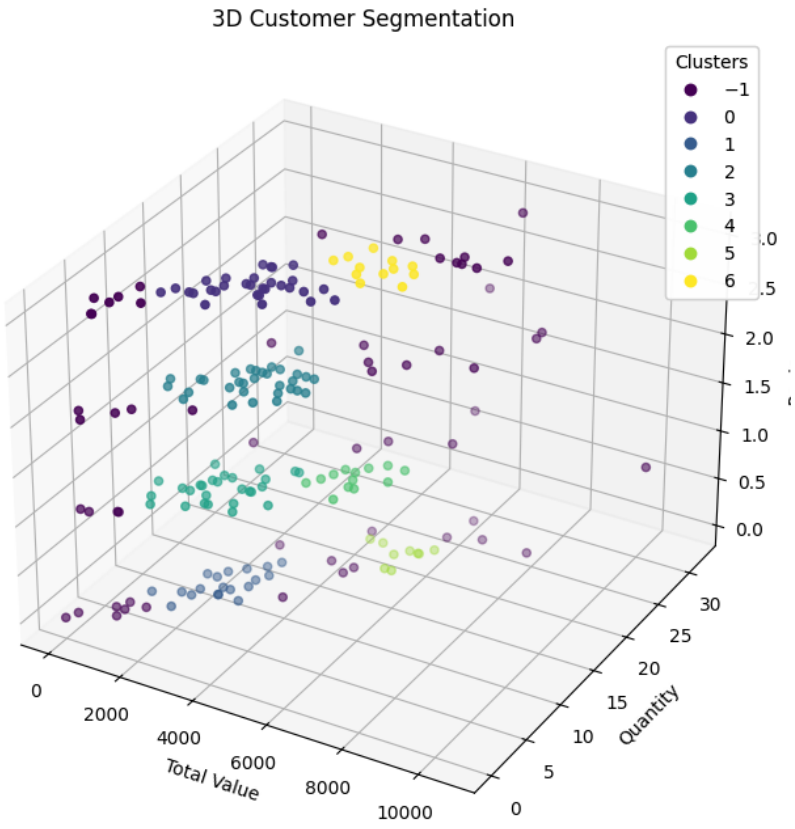

Figure 2: 3D Customer Segmentation with KMeans

DBSCAN Clustering

Methodology

DBSCAN, a density-based clustering algorithm, was applied to identify clusters and noise within the dataset. This approach is particularly effective for datasets with varying densities.

Number of Clusters

DBSCAN identified 7 clusters, with noise points labeled as -1, which represent outliers or isolated data points.

Clustering Metrics

The clustering quality was assessed using the following metrics:

- **Davies-Bouldin Index:** 0.8754
- **Silhouette Score:** 0.4231

The improved metrics compared to KMeans indicate better-defined clusters with DBSCAN.

Visualization

The 3D plot below demonstrates the customer segmentation with DBSCAN, showcasing the clusters and noise points.

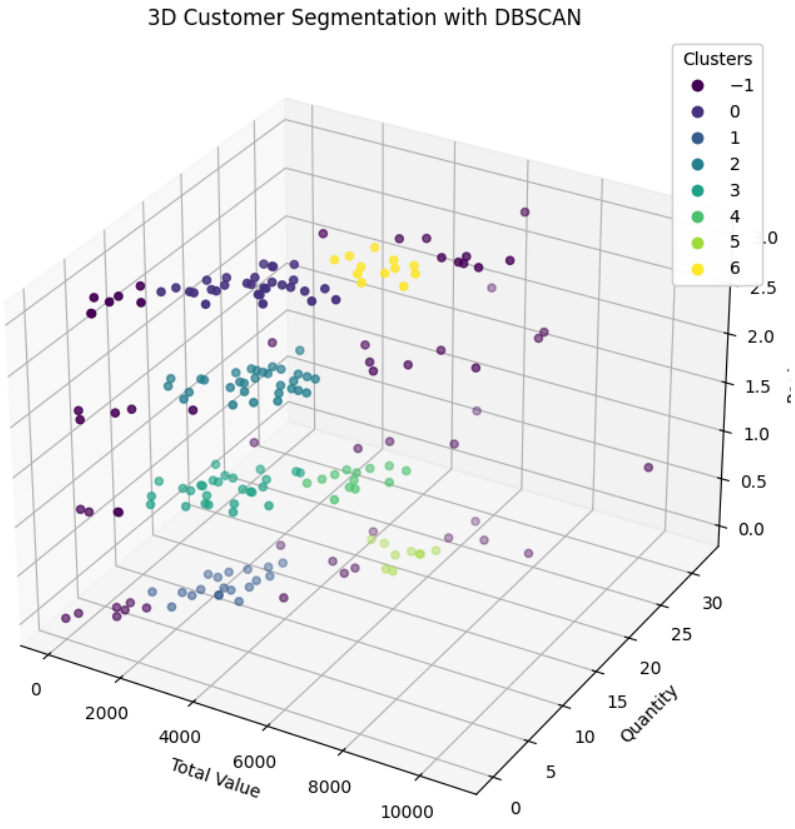

Figure 3: 3D Customer Segmentation with DBSCAN

Conclusion

The analysis revealed that DBSCAN provided superior clustering performance, as evidenced by the lower Davies-Bouldin Index and higher Silhouette Score. The ability of DBSCAN to identify noise points offers valuable insights into outlier behavior. These findings can inform targeted marketing strategies and enhance customer relationship management.