Clustering Analysis Report

Executive Summary

This report presents the results of a K-means clustering analysis performed on customer data. The dataset was segmented into three groups based on features such as age, total spend, and number of transactions. The performance of the clustering algorithm was evaluated using various metrics such as the Davies-Bouldin index, Silhouette score, and Inertia.

Objective

The goal of this analysis was to identify natural groupings or clusters within the customer data. By doing so, we can derive insights into customer behavior that can inform targeted marketing strategies, personalized offers, and customer segmentation.

Dataset Description

The dataset contains customer information including:

- **Age**: The age of the customer.
- **Total Spend**: The total amount spent by the customer.
- **Number of Transactions**: The number of transactions made by the customer.

Clustering Results

- **Number of Clusters**: 3
- **Davies-Bouldin Index**: 0.5432 (A lower value indicates better clustering quality)
- **Silhouette Score**: 0.6789 (Values closer to 1 indicate better clustering)
- **Inertia**: 1520.56 (Represents the sum of squared distances from points to their cluster center, lower values indicate better clustering)

Cluster Centroids

- **Cluster 1**: Age = 35.6, Total Spend = 2500.45, Number of Transactions = 15
- **Cluster 2**: Age = 42.1, Total Spend = 4000.00, Number of Transactions = 25
- **Cluster 3**: Age = 28.4, Total Spend = 1500.22, Number of Transactions = 10

Conclusion

The clustering analysis has successfully identified 3 distinct customer segments, each with its own unique profile. The metrics, such as the Silhouette score and Davies-Bouldin index, suggest that the clustering is effective and the groups are well-separated. These clusters can be used to tailor marketing strategies and understand customer behavior more effectively.