Customer Segmentation Analysis Report

Section 1: Introduction

This report provides an analysis of customer segmentation using clustering techniques. The objective is to group customers based on similarities in their purchasing behavior and demographic features, enabling targeted marketing and improved customer service.

Section 2: Data Preparation

The data preparation involved loading three datasets: Customers, Products, and Transactions. Key steps included converting dates to datetime objects, aggregating transaction data by customer to calculate total spend and quantity, and merging these datasets to form a unified view suitable for analysis.

Section 3: Feature Engineering

Feature engineering enhanced the dataset by encoding categorical data such as 'Region' and adding new features like 'Purchase Frequency' and 'Product Diversity'. These features are critical for clustering as they provide multiple dimensions that reflect customer behaviors and preferences.

Section 4: Clustering Analysis

Clustering was performed using the KMeans algorithm. Data was scaled using StandardScaler to ensure uniformity. The optimal number of clusters was determined through the Elbow Method, and cluster quality was evaluated using the Davies-Bouldin Index, ensuring that the clusters are well-separated and compact.

Section 5: Visualization

Clusters were visualized using scatter plots based on scaled features and PCA for dimensionality reduction. These visualizations help in understanding the distribution of clusters and are instrumental in interpreting the clustering results effectively.

Section 6: Conclusion

The clustering analysis successfully segmented the customer base into distinct groups. This segmentation allows for targeted marketing campaigns and can help in optimizing resource allocation and enhancing customer satisfaction.