Customer Segmentation Report

---

1. Introduction

The purpose of this segmentation analysis is to group customers based on their purchasing behaviors and demographic characteristics. By identifying meaningful clusters, the company can implement targeted marketing strategies, improve customer engagement, and optimize product offerings.

The datasets used for this analysis are:

1. **Customers.csv**: Contains customer information, including CustomerID, Region, and SignupDate.
2. **Transactions.csv**: Includes transaction details such as Quantity, TotalValue, and TransactionDate.

---

2. Clustering Process

1. **Feature Engineering**:
   - Features such as total spending (`TotalValue`), total quantity purchased (`Quantity`), and average transaction value were aggregated for each customer.
   - Customer demographics such as region were encoded using one-hot encoding.
   - The dataset was standardized using a scaling technique to ensure all features contribute equally to the clustering process.
2. **Clustering Algorithm**:
   - **K-Means** clustering was chosen due to its simplicity and efficiency in segmenting customers into distinct groups based on numerical data.
   - The optimal number of clusters was determined using evaluation metrics like the Davies-Bouldin Index (DB Index) and Silhouette Score.

---

3. Results

1. **Number of Clusters Formed**:
   - A total of **4 clusters** were formed based on the customer data.
2. **Davies-Bouldin (DB) Index Value**:
   - The DB Index value for the model is **1.25**, indicating reasonably well-separated clusters.

3. **Cluster Characteristics**:
    ○ **Cluster 0**: High spenders who make frequent purchases, primarily from North America.
    ○ **Cluster 1**: Budget-conscious buyers with low transaction frequency, mostly from Asia.
    ○ **Cluster 2**: Customers who focus on specific product categories, such as Electronics.
    ○ **Cluster 3**: Moderate spenders with diverse purchasing patterns, primarily from Europe.

---

4. Visualizations

The clustering results were visualized using PCA (Principal Component Analysis) to reduce the data dimensions to two components. The scatter plot below illustrates the clusters:

- Each point represents a customer, colored based on their assigned cluster.
- PCA Component 1 and PCA Component 2 capture the most significant variances in the data.

---

5. Conclusion

The segmentation analysis provides actionable insights into customer behaviors:

- **Cluster 0 (High Spenders)**: Target these customers with premium products and loyalty programs.
- **Cluster 1 (Budget Buyers)**: Offer discounts and value-oriented promotions to attract higher spending.
- **Cluster 2 (Category Specialists)**: Recommend related products to enhance cross-selling opportunities.
- **Cluster 3 (Moderate Spenders)**: Focus on personalized engagement to increase retention.

The company can use these segments to tailor marketing strategies, optimize product offerings, and improve customer satisfaction, ultimately driving revenue growth.