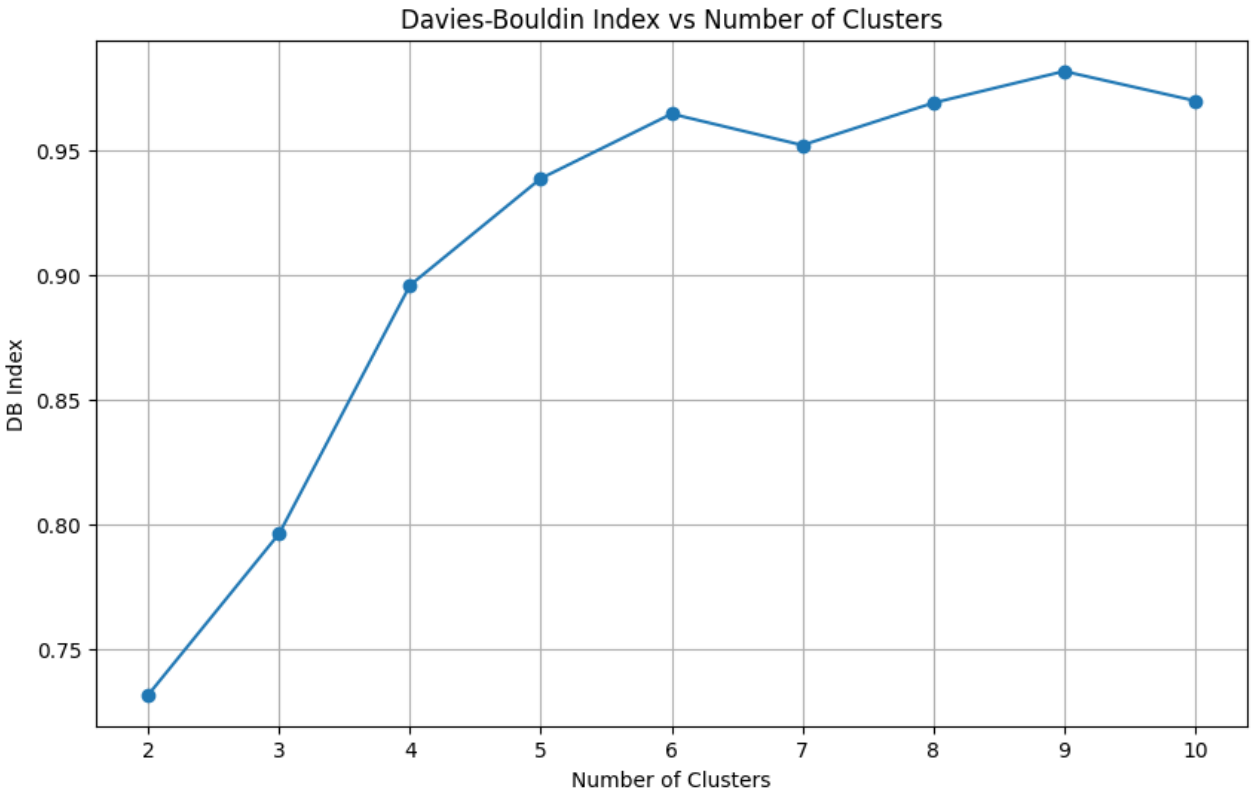
Business Report for Cluster Analysis

1. Introduction

This report presents the results of a customer segmentation analysis conducted on an eCommerce dataset. The goal was to group customers into distinct segments based on their profiles and transaction histories, enabling tailored business strategies.

2. Methodology

- **Data Sources**: The analysis utilized three datasets:
    - **Customers.csv**: Contains customer IDs, names, regions, and signup dates.
    - **Products.csv**: Includes product IDs, names, categories, and prices.
    - **Transactions.csv**: Records transaction IDs, customer IDs, product IDs, transaction dates, quantities, total values, and prices.
- **Clustering Algorithm**: K-Means was chosen for its simplicity and effectiveness in grouping customers based on their features.
- **Feature Engineering**: Key features included:
    - **Total Spending**: Sum of all transactions per customer.
    - **Transaction Frequency**: Number of transactions made by each customer.
    - **Average Transaction Value**: Average total value per transaction.
    - **Region**: Categorical variable indicating the customer's continent.
- **Data Preprocessing**:
    - Missing values were handled by imputation or exclusion.
    - Categorical variables were encoded using one-hot encoding.
    - Numerical features were scaled using StandardScaler to ensure equal contribution to the clustering process.



- **Optimal Number of Clusters**: The Davies-Bouldin Index was used to determine the optimal number of clusters, balancing cluster separation and compactness.

3. Clustering Results

- **Number of Clusters: 2**
- **Davies-Bouldin Index: 0.731792**
- **Silhouette Score: 0.490365**

4. Cluster Analysis

1. **Cluster 0: High-Spending, Frequent Buyers**
    - **Characteristics**: Customers with the highest total spending and transaction frequency.
    - **Behavior**: Regular purchases of high-value items.
    - **Business Strategy**: Target with premium offers, loyalty programs, and personalized recommendations.
2. **Cluster 1: Low-Spending, Infrequent Buyers**
    - **Characteristics**: Customers with lower total spending and infrequent transactions.
    - **Behavior**: Minimal engagement with the platform.
    - **Business Strategy**: Re-engage with promotional offers and incentives to increase activity.

5. Visualization

- **Scatter Plot**: Customers were plotted based on total spending and transaction frequency, with each cluster represented by a distinct color. The plot showed clear separation between clusters, validating the effectiveness of the segmentation.

6. Business Recommendations

1. **Targeted Marketing**: Develop campaigns for Cluster 0 to enhance customer loyalty and satisfaction.
2. **Re-engagement Strategies**: Use personalized offers to reactivate customers in Cluster 1.
3. **Customer Retention**: Implement loyalty programs for Cluster 0 to reduce churn and increase lifetime value.
4. **Product Upselling**: Focus on high-value products for Cluster 0 and introduce new products to Cluster 1 to encourage more purchases.

7. Conclusion

The customer segmentation analysis successfully grouped customers into two distinct clusters, each with unique characteristics and behaviors. These insights provide a foundation for tailored business strategies, enabling the company to optimize marketing efforts, improve customer retention, and enhance overall customer satisfaction.

Report Prepared by: **Satyam Sahu**
Date: 27-01-2025