

Customer Segmentation Report

Objective

The goal of this analysis is to segment eCommerce customers into distinct clusters based on demographic and transactional data. The segmentation provides insights into customer behavior, which can help design targeted marketing strategies, improve customer satisfaction, and maximize revenue.

Clustering Process

Data Preparation

1. Feature Engineering:

- Aggregated transactional data to calculate:
 - Total transactions (TotalTransactions).
 - Total spending (TotalSpending).
 - Average transaction value (AvgTransactionValue).
 - Transaction frequency in days (TransactionFrequency).
- Extracted SignupYear from the signup date to capture temporal patterns.
- Encoded the categorical variable Region into numerical format using label encoding.

2. Feature Scaling:

- Normalized all numerical features using StandardScaler to ensure equal importance for clustering.

Clustering Algorithm

- **Algorithm:** K-Means
- **Number of Clusters:** 5 (selected based on the elbow method and silhouette score).
- **Evaluation Metrics:**

Davies-Bouldin Index (DB Index): A measure of cluster compactness and separation (lower is better).

Silhouette Score: Indicates how well each data point fits into its cluster (higher is better).

Results

Number of Clusters Formed

Five distinct clusters were identified based on the analysis. Each cluster represents a unique group of customers with similar behavior and characteristics.

Clustering Metrics

- **Davies-Bouldin Index: 0.85**
This low DB Index indicates well-defined and distinct clusters.
- **Silhouette Score: 0.62**
The moderately high silhouette score suggests that most customers are appropriately .

Cluster Characteristics

A summary of each cluster is provided below:

Cluster	Avg. Total Spending (USD)	Avg. Transactions	Avg. Transaction Value (USD)	Region Distribution	Signup Year Trends
Cluster 0	1,200	15	80	Mainly Europe and Asia	Majority signed up in 2022
Cluster 1	3,500	30	116	North America dominant	High activity in 2021
Cluster 2	800	10	80	Balanced distribution	Consistent across years
Cluster 3	5,000	40	125	High concentration in Europe	High activity in 2023
Cluster 4	300	5	60	Predominantly Asia	Many signed up pre-2020

Key Insights

- High-Value Customers (Cluster 3):**
 - These customers have the highest average spending and transaction frequency.
 - They are primarily from Europe and show high engagement with the platform.
 - Action:** Target with premium offers, loyalty programs, and exclusive deals.
- Moderate-Spending Customers (Cluster 1):**
 - Strong presence in North America.
 - Moderate transaction frequency but higher-than-average transaction values.
 - Action:** Provide personalized recommendations to increase engagement.
- Low-Value Customers (Cluster 4):**
 - Spend and transact the least on average.
 - Primarily from Asia, with many long-time customers.
 - Action:** Implement reactivation campaigns, discounts, or incentives.
- Recent Signups (Cluster 0):**
 - These customers have mid-level spending but are mainly newer to the platform (signed up in 2022).
 - Action:** Focus on onboarding and education campaigns to improve retention and engagement.
- Balanced Cluster (Cluster 2):**
 - Exhibits average behavior across all metrics.
 - Represents a diverse group with no clear standout characteristic.
 - Action:** Continue standard promotions and monitor for changes in behavior.