

# Customer Clustering Report

The goal of this analysis was to segment customers into distinct clusters based on their purchasing behavior, product preferences, and regional attributes. Clustering helps identify patterns in customer profiles, enabling tailored marketing strategies and better business decision-making.



# Dataset Overview

- **Customers Dataset:** Contains 200 unique customers, their names, regions, and signup dates.
- **Products Dataset:** Includes 100 unique products with categories such as Books, Electronics, Clothing, and Home Decor.
- **Transactions Dataset:** Comprises 1,000 transaction records with details like transaction value, quantity, and product purchased.

# Data Preparation

**Merged Datasets:** Transactions were merged with customer and product datasets to create a comprehensive view of customer activity.

## **Aggregated Features:**

- **TotalValue:** Total revenue generated by each customer.
- **Quantity:** Total number of products purchased.
- **Category:** Most frequently purchased product category.
- **Region:** Region of the customer.

**Encoded Features:** Categorical variables (Region, Category) were one-hot encoded for numerical processing.



# Clustering Process

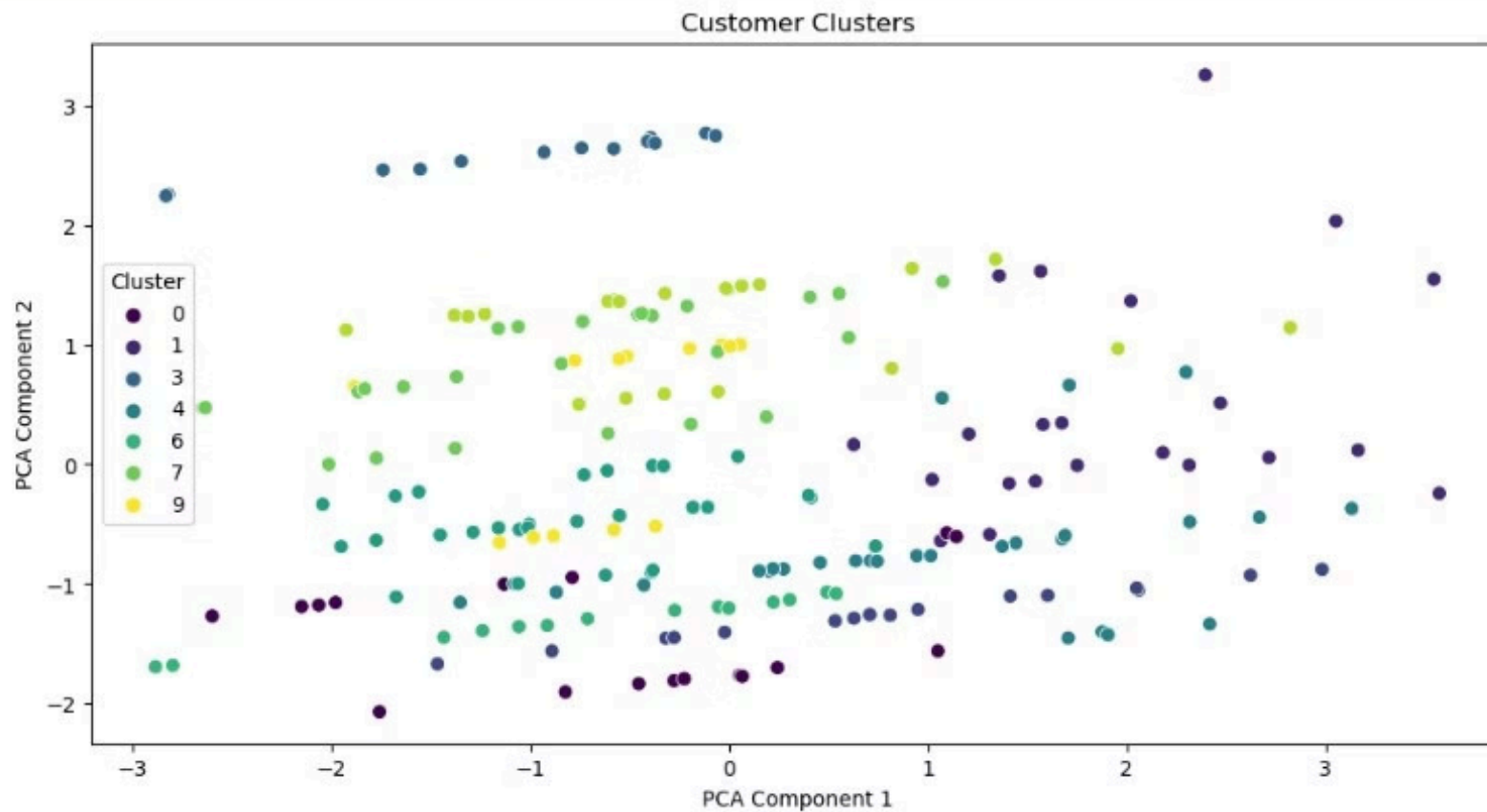
**Standardization:** Features were standardized using `StandardScaler` to ensure equal weightage across all variables.

## KMeans Clustering:

- The number of clusters ( $k$ ) was tested in the range of 2 to 10.
- The **Davies-Bouldin Index (DB Index)** was used to evaluate cluster quality. A lower DB Index indicates better-defined clusters.
- The optimal number of clusters was determined to be **10**, with a DB Index of **1.15**.

# Clustering Results and Visualization

- **Clusters:** Customers were segmented into **10 distinct clusters**.
- **Key Characteristics:**
  - a. **Spending Patterns:** High-value vs. low-value customers.
  - b. **Regional Preferences:** Variations in product preferences by region.
  - c. **Product Categories:** Clusters favor specific categories like Books or Electronics.
- **Visualization:**
  - **PCA Scatter Plot:**
    - Each point represents a customer.
    - Clusters are color-coded to highlight groupings.



# Key Insights from Clustering

1

## Customer Segmentation

The analysis identified distinct customer segments with varying needs, preferences, and buying behaviors. These segments can be used to tailor marketing strategies.

2

## Targeted Marketing

By understanding the characteristics of each cluster, we can tailor marketing campaigns and messages to resonate with specific customer segments.

3

## Personalized Experiences

Personalized recommendations, promotions, and communication can enhance customer satisfaction and drive repeat purchases by addressing individual needs.