Customer Segmentation Analysis for eCommerce Dataset

Introduction:

This document provides a detailed analysis of the customer segmentation performed on the eCommerce dataset, along with key insights derived from the exploratory data analysis and clustering. The segmentation aims to identify customer groups with similar purchasing behaviour, enabling actionable strategies for personalized marketing, inventory management, and revenue optimization.

Exploratory Data Analysis Insights:

1. **Customer Distribution Across Regions:**

Customers are distributed across different continents, with a notable concentration in specific regions. Targeted marketing campaigns can focus on regions with higher customer density.

2. **Spending Behaviour:**

A small percentage of customers account for a significant portion of total revenue. These high-spending customers could be targeted for loyalty programs or exclusive offers.

3. **Product Performance:**

Certain product categories consistently drive high revenue, while others have low demand. Optimizing inventory to focus on top-performing categories can increase profitability.

4. **Transaction Trends:**

Peak transaction periods were identified, often aligning with promotional campaigns or holiday seasons. Planning promotions during these times could maximize sales.

5. **Signup Behaviour:**

Analysis of signup dates revealed periods with higher customer acquisition rates. Understanding the factors contributing to these spikes could improve customer onboarding strategies.

Clustering Process:

Objective:

To segment customers into distinct groups based on their purchasing behavior and profile data, providing actionable insights for business strategies.

Methodology:

1. **Data Aggregation:**
    - Customer data was aggregated to calculate key metrics:
        - **Total Spending**: Sum of all transaction values for a customer.
        - **Total Quantity**: Total quantity of products purchased.
        - **Unique Products**: Number of unique products purchased.
2. **Data Normalization:**
    - Metrics were scaled using StandardScaler to ensure uniformity and eliminate biases caused by varying ranges in the data.
3. **Clustering Algorithm:**
    - **K Means Clustering** was applied with an optimal number of clusters determined through experimentation.
    - Four clusters were selected for segmentation.
4. **Evaluation:**
    - The clustering was evaluated using the **Davies-Bouldin Index (DBI)**, with a score of **0.90**. A lower DBI score indicates better-defined clusters.

Cluster Profiles:

Cluster 0: High-Value Customers

- **Characteristics:**
  - High total spending and high transaction frequency.
  - Purchase a wide variety of products.
- **Actionable Strategy:**
  - Prioritize for loyalty programs, early access to sales, and personalized recommendations.

Cluster 1: Occasional Shoppers

- **Characteristics:**
  - Moderate spending and infrequent purchases.
  - Focus on specific product categories.
- **Actionable Strategy:**
  - Engage with targeted offers or discounts to encourage repeat purchases.

Cluster 2: Budget Buyers

- **Characteristics:**
  - Low total spending and low product variety.
  - Sensitive to pricing.
- **Actionable Strategy:**
  - Focus on discounts, bundling offers, and promoting low-cost products.

Cluster 3: Niche Shoppers

- **Characteristics:**
  - Purchase specific products with moderate frequency.
  - Loyal to certain categories.
- **Actionable Strategy:**
  - Promote related products and upselling within their preferred categories.

Visualization and Results:

1. **Cluster Distribution:**

- A bar chart shows the number of customers in each cluster, highlighting the dominant customer groups.

2. **Spending vs. Quantity:**
    - A scatter plot visualizes customer segmentation based on spending and quantity, revealing distinct group behaviours.
