Customer Segmentation and Clustering Report

eCommerce Transactions Dataset

Introduction

- This report aims to segment customers into meaningful groups based on their transaction history and profile data using clustering techniques. The insights derived will help the business enhance its marketing strategies and customer experience.

Methodology

1. **Data Preprocessing:**
    - Combined Transactions.csv and Customers.csv datasets.
    - Aggregated customer data to calculate features like total spending (TotalValue), items purchased (Quantity), and region.
    - Normalized numerical features for clustering.
2. **Clustering Algorithm:**
    - Applied **K-Means Clustering** to segment customers.
    - Determined the optimal number of clusters using the **Elbow Method**.
3. **Evaluation Metric:**
    - Used the **Davies-Bouldin Index (DB Index)** to evaluate cluster quality.

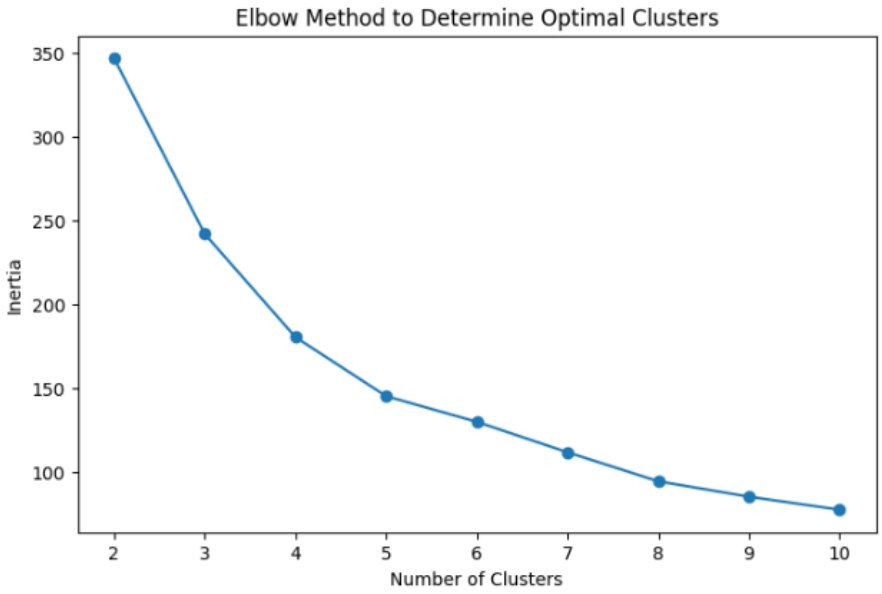
Results

1. **Number of Clusters Formed:**The Elbow Method identified **4 clusters** as optimal.
2. **DB Index Value:** The DB Index value was 0.932037310279212, indicating well separated clusters.

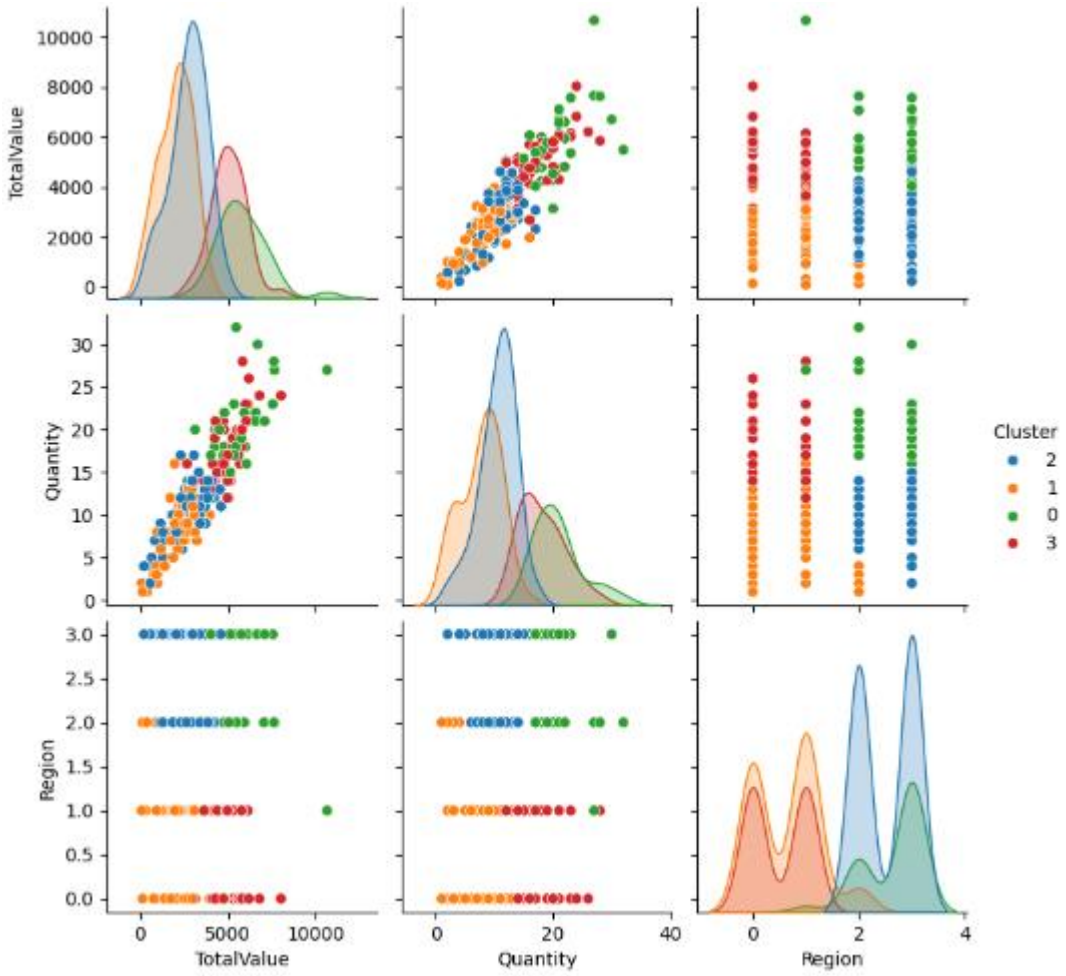
3. **Visualizations:**

   Add plots generated from your analysis:

   - **Elbow Curve:** To show the optimal number of clusters.



   - **Cluster Scatterplot or Pairplot:** To visualize customer segmentation.



Conclusion

- The clustering revealed distinct customer segments based on spending and regional distribution. These insights can guide targeted promotions and loyalty programs.
- Future analysis could integrate additional data, such as product preferences, to further refine segmentation.

Report By:

Jaswanth Kumar Param