Task 3: Customer Segmentation

Clustering Overview

Customer segmentation was performed using the **KMeans clustering algorithm**. Features used for clustering include:

- **TotalValue**: Total revenue generated by a customer.
- **Quantity**: Total quantity of products purchased.
- **Price**: Average price of purchased products.

Key Results

- **Number of Clusters**: 5
- **Davies-Bouldin Index**: 0.57 (lower is better, indicating well-separated clusters).

Visual Representation

Clusters were visualized using **PCA (Principal Component Analysis)** to reduce dimensions to two. The plot illustrates distinct clusters, demonstrating the segmentation.

Business Insights

1. **Cluster Profiles**:
   o Cluster 1: High-value customers with frequent purchases.
   o Cluster 2: Moderate spenders who buy mid-priced products.
   o Cluster 3: Low-spending customers with infrequent transactions.
   o Cluster 4: Customers preferring low-priced items in bulk.
   o Cluster 5: Occasional high-spenders.
2. **Actionable Recommendations**:
   o Focus on retaining high-value customers in Cluster 1 through loyalty programs.
   o Design targeted promotions to encourage repeat purchases for customers in Clusters 2 and 3.
   o Upsell or cross-sell to occasional high-spenders in Cluster 5.