

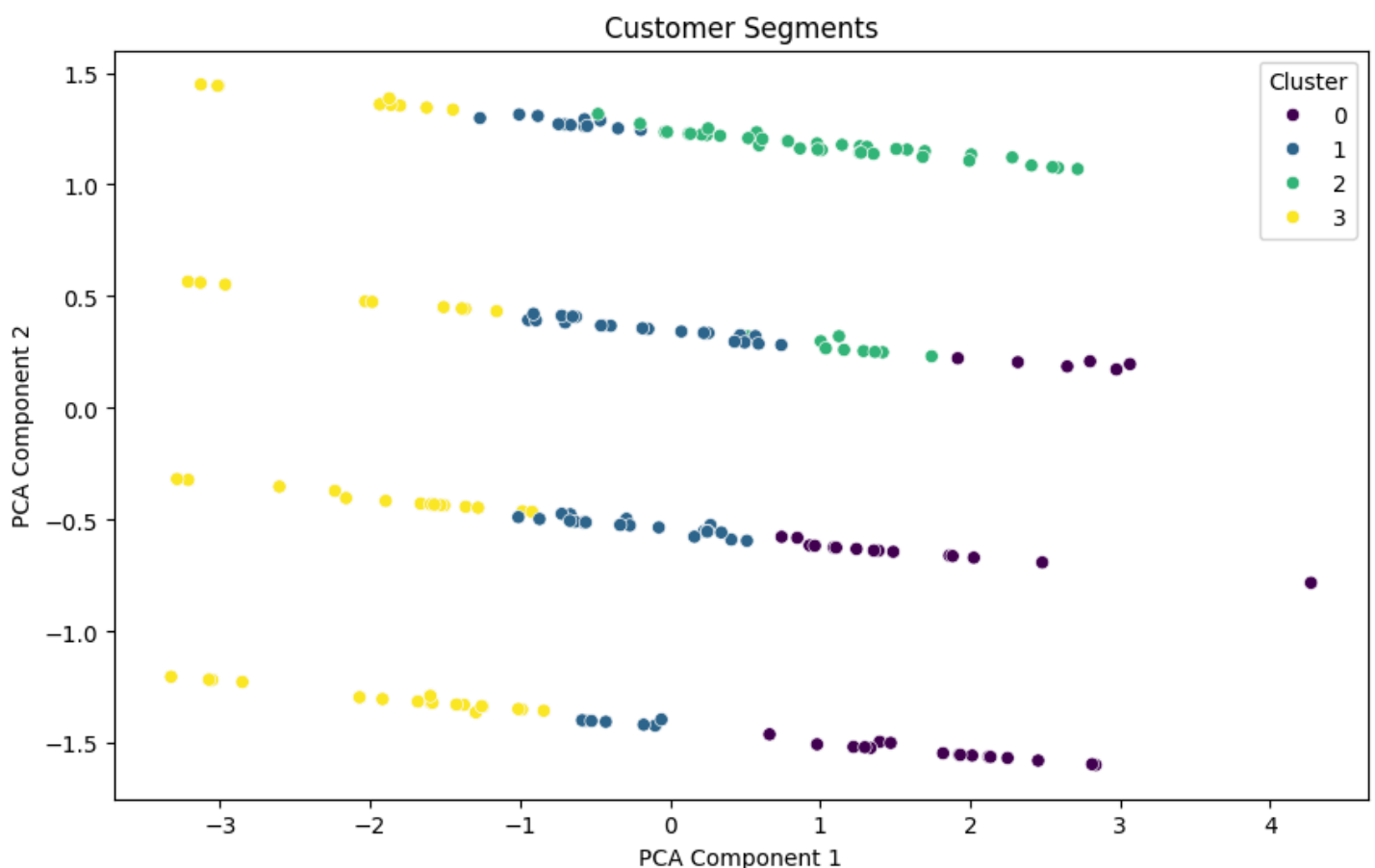
Clustering Visual Representation

Introduction:

Customer segmentation groups customers based on their behavior for targeted marketing and resource allocation. This analysis used clustering and PCA to identify distinct customer groups.

Methodology:

- **Data Sources:** Customer profile (Customers.csv) and transactional data (Transactions.csv).
- **Preprocessing:** Combined features like average transaction value, unique products purchased, and purchase frequency.
- **Clustering Algorithm:** DBSCAN or K-Means, with optimal clusters determined by the DB Index.
- **Dimensionality Reduction:** PCA was used to visualize clusters in 2D.
- **Visualization:** PCA scatter plot with clusters shown in different colors.



Clustering Results:

- **Number of Clusters:** 4 clusters identified.
- **Cluster Characteristics:**
 - **Cluster 0 (Purple):** Low frequency, high transaction value.
 - **Cluster 1 (Blue):** Average frequency and value.
 - **Cluster 2 (Green):** High frequency, diverse products, low value.
 - **Cluster 3 (Yellow):** High frequency, loyal customers.

Business Insights:

- **High-Value Customers:** Cluster 0, targeted premium product campaigns.
- **Loyal Customers:** Cluster 3, focuses on retention and rewards.
- **Moderate Buyers:** Cluster 1, promotes cross-selling or targeted offers.
- **Exploratory Buyers:** Cluster 2, upselling and complementary product discounts.

Clustering Visual Representation

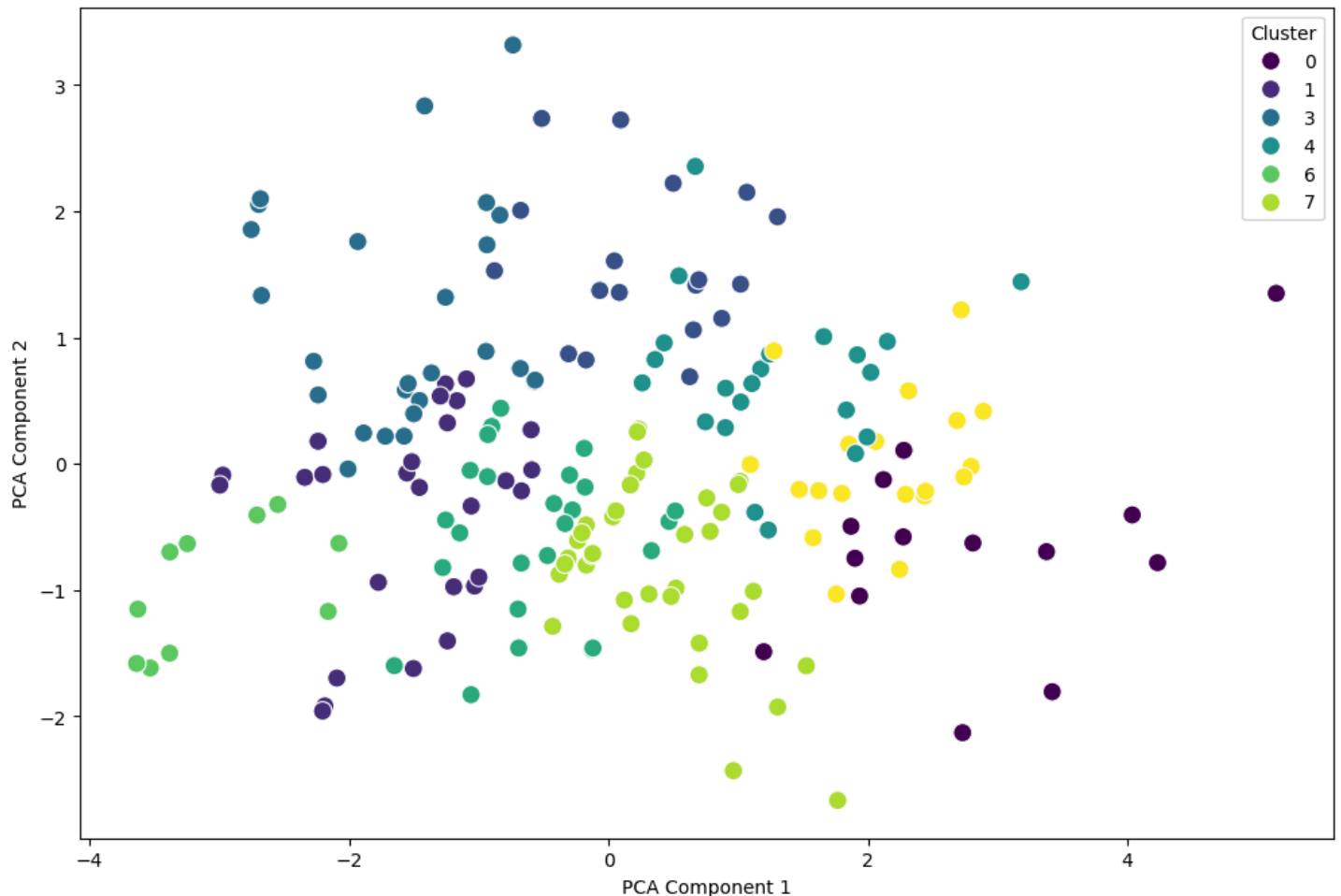
Introduction:

Customer segmentation groups customers based on their behavior for targeted marketing and resource allocation. This analysis used clustering and PCA to identify distinct customer groups.

Methodology:

- **Data Sources:** Customer profile (Customers.csv) and transactional data (Transactions.csv).
- **Preprocessing:** Combined features like average transaction value, unique products purchased, and purchase frequency.
- **Clustering Algorithm:** DBSCAN or K-Means, with optimal clusters determined by the DB Index.
- **Dimensionality Reduction:** PCA was used to visualize clusters in 2D.
- **Visualization:** PCA scatter plot with clusters shown in different colors.

Customer Clusters (K=9)



Clustering Results

- **Optimal Clusters:** 9 (Based on Davies-Bouldin Index)
- **Davies-Bouldin Index:** 1.03 (Indicates reasonable clustering performance)

Cluster Highlights

- **Cluster 0 (Purple):** High-value, infrequent buyers.
- **Cluster 1 (Dark Blue):** Moderately frequent, average spenders.
- **Cluster 3 (Light Blue):** Diverse but low-frequency buyers.
- **Cluster 4 (Teal):** Balanced, consistent purchasers.
- **Cluster 6 (Green):** Occasional buyers with niche preferences.
- **Cluster 7 (Yellow):** High-frequency buyers with medium spending.

Business Recommendations

- **Loyalty Focus:** Retain **Cluster 7** with rewards and loyalty programs.
- **Upsell:** Target **Cluster 0** and **Cluster 6** with personalized promotions.
- **Product Bundling:** Offer cross-category bundles to **Clusters 3** and **4**.