

# Clustering Results

## Data Science Assignment: eCommerce Transactions Dataset

### Task 3: Customer Segmentation / Clustering

#### Clustering Results:

##### 1. Number of Clusters:

- The optimal number of clusters was determined to be 5, based on the elbow method and silhouette score analysis.

##### 2. Clustering Metrics:

Davies-Bouldin Index: 0.8581876573884915

Silhouette Score: 0.4827623130400012

##### 3. Cluster Characteristics:

- Cluster 1: High-value customers with frequent purchases.
- Cluster 2: New customers with low transaction history.
- Cluster 3: Moderate-value customers from specific regions.
- Cluster 4: Seasonal buyers with large transaction values during holidays.
- Cluster 5: Low-value customers with sporadic activity.

#### Visualizations:

##### 1. PCA Plot for Cluster Visualization:

- A 2D PCA scatter plot was used to visualize customer clusters, with distinct separation between the clusters.

## 2. Cluster Size Distribution:

- A chart highlighted the distribution of customers across the 5 clusters.

