

Task 3: Customer Segmentation / Clustering

Objective:

- Perform customer segmentation using clustering techniques to identify distinct customer groups.

Methodology:

1. Data Preparation:

- Used Customers.csv and aggregated transaction data from Transactions.csv.
- Standardized features using StandardScaler.

2. Clustering Algorithm:

- Applied **KMeans** clustering with different values of k (2 to 10).
- Evaluated clusters using the **Davies-Bouldin Index** (DBI).

3. Optimal Clusters:

- Optimal number of clusters: **4** (lowest DBI value of **1.0653**).

4. Visualization:

- Visualized clusters using PCA for dimensionality reduction.
- Plotted clusters to understand customer group characteristics.

Results:

1. Cluster Descriptions:

- **Cluster 1:** High spenders with frequent transactions.
- **Cluster 2:** Moderate spenders focused on specific categories.
- **Cluster 3:** Occasional buyers with low transaction value.
- **Cluster 4:** New customers with minimal transaction history.

2. Metrics:

- **Davies-Bouldin Index** : 1.0653
- **Silhouette Score**: 0.3414
- **Calinski-Harabasz Index**: 82.9618

Business Implications:

- Personalized marketing campaigns can target specific clusters.
- Cluster 1 customers should be prioritized for premium offerings.
- Cluster 4 customers require onboarding and engagement strategies.