Task 3: Customer Segmentation using Clustering techniques:

Data Preprocessing:

The segmentation was performed using the following key attributes:

- **Customer Profile Information**: Extracted from Customers.csv such as region and customer-related details.
- **Transaction Information**: Extracted from Transactions.csv such as total transaction value (TotalValue), quantity of items purchased (Quantity), and average price (Price).

These features were combined and aggregated as follows:

- **TotalValue**: The sum of all purchase values for each customer.
- **Quantity**: The total quantity of products bought by each customer.
- **Price**: The average price of products purchased by each customer.

The data was then scaled using **StandardScaler** for normalization and standardization to ensure that each feature contributes equally to the clustering process.

Clustering Algorithm:

We used the **K-Means clustering** algorithm, which is well-suited for customer segmentation as it can identify natural groupings within the data.

- **Number of clusters (k)**: 2 to 10 clusters were tested to ensure optimal value of db index.
- **Initialization**: K-Means was initialized with 10 different centroids (n_init=10) to ensure robustness.

Clustering Results:

- **Number of Clusters**: We chose **5 clusters** based on the performance and behavior of the clustering algorithm.
- **DB Index**: The Davies-Bouldin Index for the 5 clusters was calculated as **0.8381**, indicating that the clusters are reasonably well separated.
- **Inertia**: It is the sum of squared distances of samples to their closest cluster center. Lower inertia values indicate more compact clusters. Inertia=162.57