Objective:
The goal of this project is to segment customers based on their transaction behavior and signup information using K-Means clustering.

Data Description:

1. Customers Dataset: Contains CustomerID, CustomerName, and SignupDate.
2. Transactions Dataset: Contains TransactionID, ProductID, TransactionDate, Quantity, TotalValue, and Price.
3. Merged Dataset: Combined the Customers and Transactions datasets on CustomerID for feature engineering.

Feature Engineering:
Derived the following features for each customer:

- TotalTransactions: Total number of transactions made.
- TotalValue: Total monetary value of all transactions.
- TotalQuantity: Total quantity of products purchased.
- Recency: Days since the most recent transaction.

Data Preprocessing:

1. Datetime Conversion: Converted TransactionDate and SignupDate to datetime format for accurate computations.
2. Feature Normalization: Used StandardScaler to normalize features for unbiased clustering.

Clustering Methodology:

1. Clustering Algorithm: Performed clustering using K-Means with cluster numbers ranging from 2 to 10.
2. Evaluation Metrics:

- Davies-Bouldin Index (DB Index): Lower values indicate better clustering.
- Silhouette Score: Higher values indicate better clustering.

3. Optimal Number of Clusters: Determined the optimal number of clusters based on the lowest DB Index.

Results:

1. Optimal Number of Clusters: The optimal number of clusters is X (replace with actual value).
2. Cluster Summary: Below is a summary of the clusters based on mean values of key features.

Cluster	TotalTransactions	TotalValue	TotalQuantity	Recency
Cluster 0	7.14	5220.89	18.96	60.58
Cluster 1	3.72	2383.80	8.91	87.92

Visualizations:

1. DB Index and Silhouette Score: A line plot shows how these metrics vary with the number of clusters.
2. t-SNE Visualization: A scatter plot demonstrates clear separation of clusters after dimensionality reduction.

Key Insights:

1. Customer Behavior:
   - Cluster X: High transaction frequency and high total value suggest loyal, high-spending customers.
   - Cluster Y: Lower transaction frequency but higher total value indicates occasional high spenders.

- Cluster Z: Low spending and infrequent transactions suggest disengaged customers.

2. Actionable Recommendations:
    - Cluster X: Focus on retention strategies (e.g., loyalty programs).
    - Cluster Y: Encourage frequent purchases through targeted discounts.
    - Cluster Z: Re-engage with personalized promotions or reminders.

Conclusion:

The clustering analysis effectively segmented customers into meaningful groups based on transaction behaviour and recency, providing insights for targeted marketing strategies and improved customer engagement.