Task 3: Customer Segmentation / Clustering

Objective:

The goal is to perform clustering using K-Means, Gaussian Mixture Model (GMM), and DBSCAN, and evaluate the results based on Davies-Bouldin Index (DBI.

Key Findings:

- **Optimal Clusters**:
  - **K-Means**: 8 clusters
  - **GMM**: 5 clusters
  - **DBSCAN**: 5 clusters (Optimal Epsilon: 0.4)
- **Performance Evaluation**:
  - **K-Means DB Index**: **0.852** (Higher values indicate poorer clustering)
  - **GMM DB Index**: **1.116** (Indicates suboptimal clustering)
  - **DBSCAN DB Index**: **0.431** (Best clustering result with well-separated clusters)

Conclusion:

- **Best Clustering Algorithm**: **DBSCAN**, with the lowest DBI, indicating it provides the best clustering performance.
- **K-Means** and **GMM** are less optimal based on the DBI.

Final Clustering Metrics:

- **K-Means**:
  - DB Index: 0.852
  - Optimal Number of Clusters: 8
- **GMM**:
  - DB Index: 1.116
  - Optimal Number of Clusters: 5
- **DBSCAN**:
  - DB Index: 0.431
  - Optimal Number of Clusters: 5 (Optimal Epsilon: 0.4)