Task 3: Customer Segmentation (Clustering)

Deliverables:

1. **Customer Clusters**:

   Using KMeans clustering, customers are divided into 4 groups based on transaction and profile data.

   Each cluster represents a unique segment of customers with similar behaviors:

   Cluster 1: High spenders with frequent transactions.

   Cluster 2: Medium spenders with occasional transactions.

   Cluster 3: Low spenders with infrequent transactions.

   Cluster 4: Region-based grouping (e.g., customers from similar regions with specific spending habits).

2. **Davies-Bouldin Index**:

   A low Davies-Bouldin Index value indicates well-separated clusters, suggesting the clustering algorithm performed effectively.

   At the first Davies-Bouldin Index: 0.988066754535579 is high it means clusters are not well-separated

   Using PCA to reduce dimensionality before clustering to eliminate noise and irrelevant features.

   Davies-Bouldin Index (PCA + KMeans) : 0.754997112380128 is low it means clusters are well-separated

3. **Visualization of Clusters**:

   Using PCA, the high-dimensional data is reduced to 2 dimensions for visualization.

   The scatter plot shows the distinct clusters, making it easy to interpret how customers are grouped.

4. **Business Impact**:

   Customer segmentation allows the company to:

   Tailor marketing strategies for each segment.

   Allocate resources effectively (e.g., focus on high-value segments).

   Improve customer satisfaction through personalized offerings.