Data Preparation

- **Datasets Used**:
- **`Customers.csv`**: Contains customer profile information (e.g., Age, Income).
- **`Transactions.csv`**: Contains transaction details (e.g., Transaction Amount, Transaction ID).
- **Key Features Engineered**:
- **Total Spend**: The total amount spent by each customer.
- **Transaction Count**: The total number of transactions made by each customer.

Clustering Analysis

- **Clustering Algorithm Used**: K-Means Clustering
- **Optimal Number of Clusters**: 4 (determined using the Elbow Method)

Clustering Metrics

-

Davies-Bouldin Index (DB Index):

-
- **Value**: [Insert DB Index Value Here]
- **Interpretation**: The Davies-Bouldin Index is a measure of clustering quality. A lower DB Index indicates better clustering, as it suggests that clusters are well-separated and compact.
-

Silhouette Score:

-
- **Value**: [Insert Silhouette Score Here]
- **Interpretation**: The Silhouette Score measures how similar an object is to its own cluster compared to other clusters. A score close to +1 indicates that the object is well clustered, while a score close to -1 indicates that the object may have been assigned to the wrong cluster.

Visualizations

- **Elbow Method Plot**:
- The plot shows the inertia (sum of squared distances from each point to its assigned cluster center) for different numbers of clusters. The "elbow" point indicates the optimal number of clusters.
- **Cluster Visualization**:
- A scatter plot visualizing the clusters in a 2D space using PCA (Principal Component Analysis) for dimensi