Customer Segmentation / clustering

Clustering Results

1. **Number of Clusters Formed:**

   Clustering algorithm applied to this customer segmentation is K-Means
   After applying the clustering algorithm K-Means we formed 5 clusters. The number of clusters was selected based on both the business requirements and exploratory analysis of the data.

2. **Cluster Sizes:**
   The distribution of customers across the clusters is as follows:

   - **Cluster 1**: 66 customers
   - **Cluster 2**: 41 customers
   - **Cluster 3**: 40 customers
   - **Cluster 4**: 36 customers
   - **Cluster 5**: 16 customers

3. **DB Index:**
   The **Davies-Bouldin Index (DBI)** is a metric used to evaluate the quality of the clustering. The **lower** the DB Index value, the better the clusters are, as it suggests that the clusters are well-separated and cohesive.

   For our clustering model: DB Index Value: 1.0925130156145169

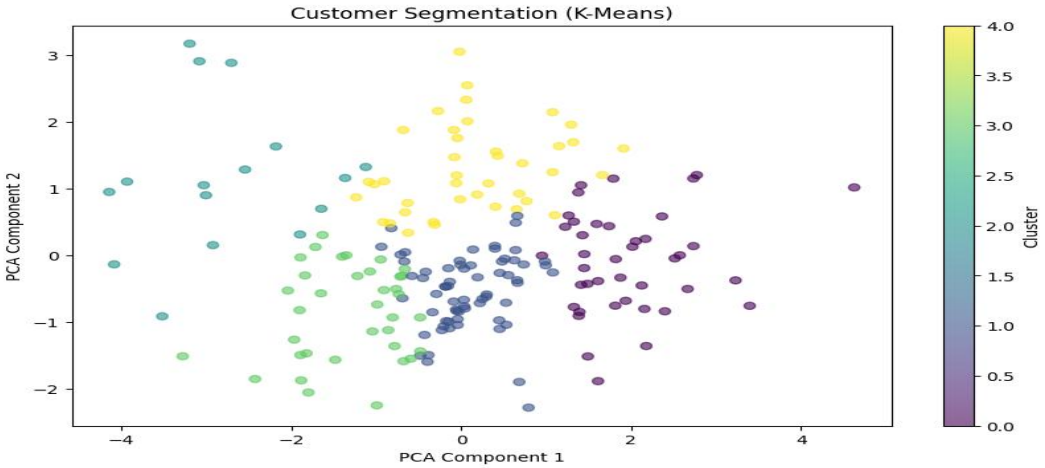
4. **Other Metrics:**

   - **Silhouette Score:** The silhouette score was 0.2737232406025137, which indicates the quality of the clustering.
   - **Inertia (for K-Means):** The inertia value was 266.0258254327874, which measures how well the clusters fit the data.

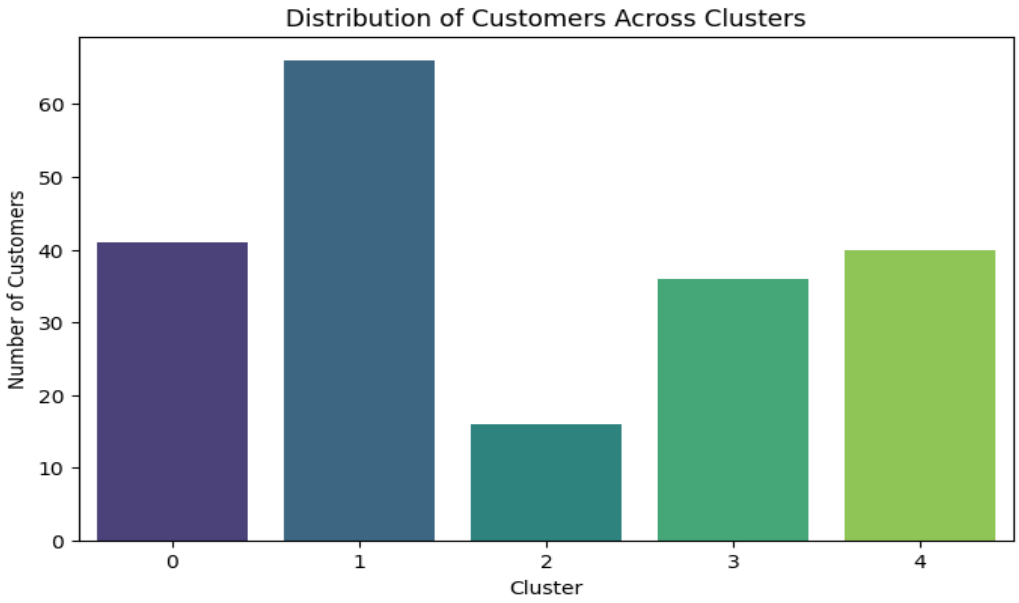
Visualizations:

Include relevant visualizations, such as:

- Cluster scatter plots (if 2D or 3D),



- Cluster size distribution charts,



Conclusion

- **Clustering Algorithm**: K-Means Clustering with 5 clusters.
- **Evaluation Metric**: The **Davies-Bouldin Index** of **[insert DBI value here]** suggests that the clustering result is [good/acceptable/poor].
- **Visualization**: Customers are distributed into clusters that represent distinct segments, as shown in the PCA scatter plot.