Customer Segmentation Report:

1. Number of Clusters Formed

- **Clustering Algorithm**: KMeans
- **Number of Clusters**: 3
  The number of clusters ($k=3$) was selected as an optimal balance between simplicity and performance, based on evaluation using the Davies-Bouldin Index.

2. Davies-Bouldin (DB) Index

The **DB Index** for the clustering solution is:

$$\text{DB Index} = 0.64$$

This value reflects the quality of clustering, where a lower value indicates better compactness and separation of clusters. The achieved score of **0.64** suggests good-quality clusters, appropriate for segmentation.

3. Other Relevant Clustering Metric:

- **Cluster Sizes**:
  - Cluster 0: 132 customers
  - Cluster 1: 143 customers
  - Cluster 2: 125 customers
- **Inertia (Within-Cluster Sum of Squares)**:

  $$\text{Inertia} = 116.85$$

  This is a measure of how tightly data points are grouped within their respective clusters. Lower inertia indicates well-defined clusters.

- **Silhouette Score**:
  - The silhouette score measures how similar a data point is to its own cluster compared to other clusters.

o Calculated score: **0.52**

This indicates moderate clustering quality, which is acceptable given the real-world nature of the data.

4. Visual Representation of Clusters

Using PCA to reduce the dimensions to two, the clusters were visualized in a 2D scatter plot. The plot shows distinct groups, with minimal overlap between clusters, indicating meaningful segmentation

5. Summary of Cluster Characteristics

Cluster	Description	Key Characteristics
Cluster 0	High spenders with frequent transactions	High total spend, high transaction count
Cluster 1	Moderate spenders with diverse regions	Moderate spending, average transaction count
Cluster 2	Low spenders with specific preferences	Low spend, low transaction count

Conclusion

- **Number of Clusters**: 3
- **DB Index**: 0.64
- **Key Insights**:
  o **Cluster 0**: Represents the top customers (high-value segment). Recommended for premium offers and loyalty rewards.
  o **Cluster 1**: Represents mid-tier customers. Target for cross-selling and personalized marketing campaigns.
  o **Cluster 2**: Represents low-value customers. Focus on retention strategies and enhancing product offerings.

.