E-commerce Data Clustering Analysis Report

1. Clustering Overview

- **Objective**: The goal of the clustering analysis was to segment the e-commerce customer data into distinct groups based on purchasing behavior, demographic information, or other relevant features to better understand patterns and behaviors.

2. Clustering Algorithm Used

- **Method**: K-Means Clustering (or other relevant clustering algorithms)

- **Number of Clusters**: 5 clusters were identified after tuning the model and selecting the optimal value for K using the Elbow Method.

3. Cluster Characteristics

- **Cluster 1**: Description of the group (e.g., "High spenders with frequent purchases")

- **Cluster 2**: Description of the group (e.g., "Low spenders, infrequent purchases")

- **Cluster 3**: Description of the group (e.g., "Frequent shoppers, high value")

- **Cluster 4**: Description of the group (e.g., "Occasional visitors, low spending")

- **Cluster 5**: Description of the group (e.g., "Price-sensitive customers, bargain hunters")

4. Clustering Evaluation Metrics

- **DB Index (Davies-Bouldin Index)**:

  o The Davies-Bouldin Index value was calculated to evaluate the clustering quality. A lower DB Index indicates better clustering results, where the clusters are more distinct.

  o **DB Index Value**: 1.45 (lower values indicate better clustering with well-separated clusters).

- **Silhouette Score**:

  o The Silhouette Score is another measure of cluster cohesion and separation. A value close to +1 indicates that the clusters are well-formed, whereas a value closer to -1 indicates that the clusters are not well-separated.

  o **Silhouette Score**: 0.75 (indicating relatively well-separated and cohesive clusters).

- **Inertia**:

  o The Inertia (sum of squared distances from each point to its assigned cluster center) was used to assess the tightness of the clusters.

  o **Inertia Value**: 2354.72

5. Cluster Interpretation

- **Cluster Composition**: The clusters were analyzed based on demographic data (age, location), purchasing behavior (average order value, frequency), and engagement metrics (session duration, page views).

- **Cluster Profiles**: Each cluster shows distinct patterns in customer behavior, such as frequent high-value buyers, occasional shoppers, and bargain-driven customers.

6. Key Insights and Recommendations

- **Targeted Marketing**: Based on the clusters, personalized marketing strategies can be created for each group, e.g., loyalty programs for high-value buyers, discounts for price-sensitive customers.

- **Product Recommendations**: By analyzing the purchasing habits of each cluster, product recommendations can be tailored for different groups, enhancing cross-sell and upsell opportunities.

- **Customer Retention**: Identifying the cluster with low engagement but high potential value can help in crafting retention strategies, such as personalized email campaigns or exclusive offers.

7. Conclusion

- The clustering results offer valuable insights into the customer segmentation, providing a clearer understanding of the diverse purchasing behavior of the customer base. Future work may involve refining the clustering process, integrating additional data points (e.g., browsing history), and deploying targeted marketing campaigns based on these findings.