Customer Segmentation: Clustering Results Report

1. Introduction This report outlines the results of customer segmentation using clustering techniques. The segmentation is based on customer profile and transaction data from the provided eCommerce dataset. The primary goal is to identify meaningful customer groups to inform targeted marketing strategies and optimize business decisions.

2. Clustering Approach

Algorithm Used: DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

- **Why DBSCAN?** It is robust to noise, can find clusters of arbitrary shapes, and performs well on datasets with outliers.
- **Number of Clusters**: 15 clusters (excluding noise points) were formed.

Evaluation Metric:

- **Davies-Bouldin Index (DB Index)**: A lower value indicates better-defined and well-separated clusters.
- **DB Index Value**: **1.4697** (indicating good clustering quality).

3. Clustering Results

Cluster Summary:

Cluster	Avg Spending ($)	Avg Transactions	Avg Quantity	Customer Count
-1 (Noise)	3,603.45	5.55	13.95	40
0	3,848.29	5.00	13.10	21
1	3,636.43	5.36	14.00	14
2	4,415.59	5.64	15.57	14
3	2,328.64	3.00	7.17	6
4	1,875.41	3.71	8.57	7

5	3,082.94	4.50	10.88	8
6	2,587.51	4.11	10.33	9
7	4,221.45	6.00	15.67	12
8	3,540.17	4.42	10.83	12
9	4,023.25	5.92	14.08	12
10	1,978.28	3.14	7.86	7
11	5,946.02	8.00	21.83	6
12	2,700.92	4.90	12.00	10
13	3,557.75	5.07	12.57	14
14	1,357.79	2.29	5.43	7

4. Key Insights

1. **High-Spending Cluster**:
    a. **Cluster 11** has the highest average spending ($5,946.02), transactions (8.00), and quantity (21.83), indicating top-priority customers for loyalty programs and personalized offers.
2. **Low-Spending Cluster**:
    a. **Cluster 14** represents the lowest average spending ($1,357.79) and engagement. Efforts should focus on incentivizing this group through discounts or targeted promotions.
3. **Moderate-Spending Clusters**:
    a. Clusters **2, 7, 9**, and **13** represent customers with mid-to-high spending and engagement. Cross-selling and upselling strategies can maximize their potential.
4. **Noise (Cluster -1)**:
    a. 40 customers were categorized as noise. These customers may need additional analysis or adjusted clustering parameters to incorporate them into meaningful groups.

5. Recommendations

1. **Prioritize High-Spending Customers**:
    a. Focus on Cluster 11 with loyalty rewards, exclusive offers, and premium services to retain and grow revenue from these valuable customers.

2. **Engage Low-Spending Customers**:
    a. Investigate Cluster 14's behavior and consider campaigns like free shipping or bundled discounts to increase engagement.
3. **Target Moderate Clusters**:
    a. For Clusters 2, 7, 9, and 13, explore personalized recommendations and targeted email campaigns to drive repeat purchases.
4. **Revisit Noise Points**:
    a. Evaluate customers in Cluster -1 for potential inclusion by adjusting DBSCAN parameters or analyzing outlier behavior.

6. Conclusion The DBSCAN clustering approach successfully segmented customers into 15 clusters with meaningful insights. These clusters can inform strategic marketing and operational decisions, enhancing customer satisfaction and revenue growth.