Report of Customer Segmentation / Clustering

1. Clustering Results:
The objective of this clustering analysis was to segment the customers based on their transactional and profile information.

- Optimal Number of Clusters:
  The optimal number of clusters, determined using the Davies-Bouldin Index (DB Index), as well as the KElbowVisualizer and clustering metrics, was found to be **4**.

2. Clustering Metrics:
- DB Index (Davies-Bouldin Index): **1.0068**
- Silhouette Score: **0.3158**
- Calinski-Harabasz Index: **99.80**

3. Analysis and Interpretation:
- The clustering results show that the data naturally divides into 4 groups, with some overlap, as observed from the Silhouette Score and PCA scatter plot.
- Cluster 0: High-Value Customers
  - Total spent:
    - Mean: 5587.84 (significantly higher than other clusters)
    - Spread: Moderate standard deviation (1261.08), indicating some variation in spending
    - 25th percentile: 4722.36, suggesting that a quarter of customers spend above this amount.
    - Max: 10,673.87
  - Transaction Count:
    - Mean: 7.55 transactions per customer, shows fairly frequent purchases.
    - Max: 10673 transactions, shows a few customers might be very active.
  - Average Order Value (AOV):
    - Mean: 842.71, showing that the orders are reasonably high.
    - Range: 0 to 10673.87, indicating that a few purchases are significantly large.
  - This cluster is highly engaged with the platform and customers are relatively consistent in their spending.
  - Can be segmented into premium buyers (high AOV) and regular buyers (moderate AOV).
  - Includes both loyal customers who are regularly purchasing and some who may have become less active

- Cluster 2: Moderate-Value Customers
    - Total Spent:
        - Mean: 3788.11, which is lower than Cluster 0 but still significant.
        - Spread: The standard deviation of 1133.80 indicates moderate variability.
    - Transaction Count:
        - Mean: 3.93 transactions per customer, which is lower than Cluster 1, indicating that customers in this group are less frequent purchasers.
    - Average Order Value (AOV):
        - Mean: 1069.80, which is higher than Cluster 1, suggesting these customers tend to spend more per order.
    - Cluster 2 consists of customers who spend moderately. Strategies to increase their spending through targeted promotions and discounts could be effective.
    - Though they purchase less often, when they do, their order values are relatively high.
- Cluster 3: Low-Value, Infrequent Customers
    - Total Spent:
        - Mean: 1132.33, significantly lower than Clusters 1 and 2.
        - Spread: High standard deviation (848.43), suggesting that although the average is low, there is some variability.
    - Transaction Count:
        - Mean: 1.81 transactions, indicating infrequent purchases.
    - Average Order Value (AOV):
        - Mean: 758.52, lower than other clusters.
    - Cluster 3 consists of customers who spend less overall.
    - This group might consist of one-time buyers or customers who make occasional purchases.
    - Customers in this cluster likely purchase smaller items.
- Cluster 4: Low-Value, Moderate Engagement
    - Total Spent:
        - Mean: 2306.72, which is lower than Clusters 1 and 2, but still notable.
        - Spread: High standard deviation (891.93), indicating variability in spending.
    - Transaction Count:
        - Mean: 4.46 transactions per customer, indicating moderate engagement.
    - Average Order Value (AOV):
        - Mean: 625.99 USD, which is the lowest among all clusters.
    - Cluster 4 represents customers with moderate total spending.
    - This consists of customers who purchase relatively regularly.
    - Customers in this group tend to buy low-priced items.