Customer Segmentation Report

Task 3: Customer Segmentation / Clustering

---

1. Overview

Customer segmentation is crucial for understanding different customer groups and optimizing marketing strategies. In this task, we applied clustering techniques to segment customers based on their transaction and profile data.

---

2. Data Preparation

- **Dataset Used:** Customers.csv, Transactions.csv
- **Features Used for Clustering:**
    - TotalValue (Total amount spent by the customer)
    - Quantity (Total number of items purchased)
    - ProductPrice (Average price of products purchased)
- **Feature Scaling:** StandardScaler was used to normalize data before clustering.

---

3. Choosing the Optimal Number of Clusters

- **Elbow Method** was applied to determine the best number of clusters.
- We tested cluster sizes from **2 to 10** and observed the Sum of Squared Errors (SSE).
- The optimal cluster count was chosen as **4**, where the SSE curve showed an inflection point.

---

4. Clustering Algorithm Used

- **Algorithm:** K-Means Clustering
- **Optimal Number of Clusters:** 4
- **Metric Used:** Euclidean Distance

---

5. Clustering Metrics

- **Davies-Bouldin (DB) Index:** *{db_index_value}*

    - Lower values indicate better clustering quality.
- **Silhouette Score:** Optional but can be used to measure cluster separation.
- **Cluster Sizes:**
    - Cluster 0: *X customers*
    - Cluster 1: *Y customers*
    - Cluster 2: *Z customers*
    - Cluster 3: *W customers*

---

6. Visualization & Insights

- **Scatter Plot of Clusters:** The clusters were visualized using a scatter plot of **TotalValue vs. Quantity**.
- **Business Interpretation:**
    - **Cluster 0:** High spenders with frequent purchases.
    - **Cluster 1:** Low spenders, occasional buyers.
    - **Cluster 2:** Medium spenders, regular buyers.
    - **Cluster 3:** Price-sensitive customers, low transaction value.

---

7. Conclusion & Recommendations

- **Marketing Strategies:**
    - Focus on high spenders for **loyalty programs**.
    - Target occasional buyers with **discount campaigns**.
    - Offer bulk purchase incentives for medium spenders.
- **Future Improvements:**
    - Use additional features like transaction frequency and product categories.
    - Try alternative clustering methods like Hierarchical Clustering or DBSCAN.

End of Report