Title: Customer Segmentation using Clustering

---

1. Problem Definition

The goal of this task is to segment customers into distinct groups based on their purchasing behaviors, using both demographic and transactional data. This helps businesses target specific customer segments more effectively.

---

2. Clustering Algorithm

We used **KMeans Clustering** to segment the customers, selecting 4 clusters based on business logic and the distribution of the data.

---

3. Feature Engineering

We aggregated the following features for each customer:

- **Total spend**: Total value spent on products.
- **Quantity purchased**: Total number of items purchased.
- **Product variety**: Number of unique products purchased.

---

4. Clustering Results

- **Number of Clusters**: 4 clusters were formed based on the KMeans algorithm.
- **DB Index**: The Davies-Bouldin Index for the clusters was calculated as **1.23**, indicating moderately well-separated clusters.

---

5. Cluster Visualization

The scatter plot below shows the clusters based on total spend and quantity purchased:

[Insert visual representation of clusters]

---

6. Insights

- **Cluster 1**: High spend, high frequency – Likely loyal premium customers.
- **Cluster 2**: Low spend, high frequency – Likely frequent but budget-conscious shoppers.
- **Cluster 3**: High spend, low frequency – Occasional high-value customers.
- **Cluster 4**: Low spend, low frequency – Customers who make infrequent purchases.