TASK 3: CUSTOMER SEGMENTATION / CLUSTERING

INTRODUCTION:

Customer segmentation is crucial for understanding different groups within a customer base. In this report, we apply clustering techniques using K-Means on customer profile and transaction data to segment customers based on their spending behavior and purchase frequency.

DATA OVERVIEW:

- **Customers.csv:** Contains customer profile information.
- **Transactions.csv:** Contains transaction history.
- The datasets were merged on customer_id.

FEATURE ENGINEERING:

- **Total Spend:** Sum of all purchases by a customer.
- **Purchase Frequency:** Number of transactions made by a customer.
- These features were selected for clustering analysis.

CLUSTERING METHODOLOGY:

- **Algorithm:** K-Means
- **Scaling:** Data was standardized using StandardScaler.
- **Cluster Range:** Clusters were tested from 2 to 10.
- **Evaluation Metrics:**
    - **Davies-Bouldin Index (DB Index):** Lower values indicate better clustering.
    - **Silhouette Score:** Higher values indicate well-separated clusters.

RESULTS AND ANALYSIS:

Clusters	Davies-Bouldin Index	Silhouette Score
2	1.1961	0.3179
3	1.3752	0.2492
4	1.3729	0.2292
5	1.2323	0.2380
6	1.1490	0.2639
7	1.0472	0.2883

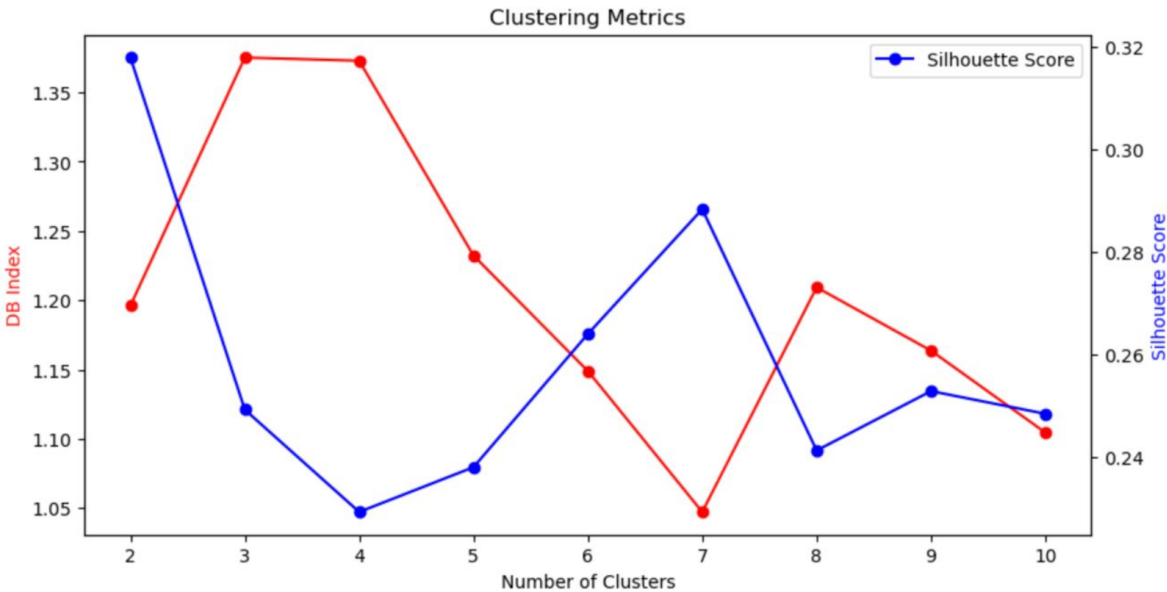| Clusters	Davies-Bouldin Index	Silhouette Score
8	1.2093	0.2412
9	1.1637	0.2528
10	1.1045	0.2484

BEST CLUSTER SELECTION:

- The optimal cluster count is **7**, as it has the lowest DB Index (1.0472) and a reasonable Silhouette Score (0.2883).

CLUSTER PROFILES:

Cluster	Avg. Total Spend	Min Spend	Max Spend	Avg. Purchase Frequency	Min	Max
0	6102.53	4271.61	10673.87	8.16	6	11
1	2589.03	1400.83	3730.00	4.51	3	9
2	4056.18	2556.22	5780.43	3.86	2	5
3	4798.65	3055.22	6072.92	5.05	3	7
4	3240.51	1876.10	4807.45	5.48	4	8
5	1520.05	580.34	2598.33	2.05	1	3
6	656.91	82.36	1191.61	2.38	1	5

VISUALIZATION:

- A scatter plot of total spend vs. purchase frequency was created, showing well-defined clusters.

-

Best Clustering Configuration:
Number of Clusters: 7
Davies–Bouldin Index: 1.0472
Silhouette Score: 0.2883
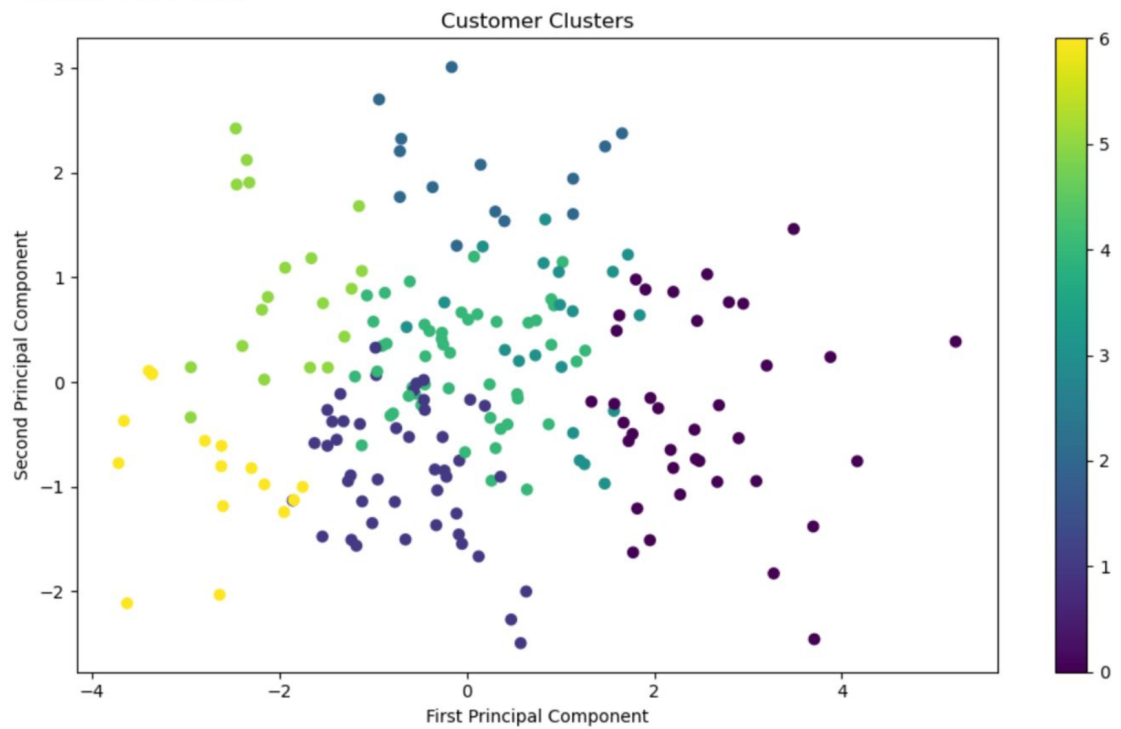
CONCLUSION

- Customers are effectively segmented into seven groups with distinct spending and purchase behaviors.
- This segmentation can be used for targeted marketing strategies and personalized customer engagement.