

## **Clustering Results Report**

For customer segmentation, I applied K-Means clustering using both customer profile data (e.g., region) and transaction data (e.g., quantity purchased). After testing cluster numbers from 2 to 10, I chose the optimal number of clusters based on the DB Index, which measures cluster separation. The DB Index was lowest for 4 clusters, indicating optimal segmentation.

Additionally, I calculated the Silhouette Score, which was 0.42, indicating moderate clustering quality with a reasonable separation between clusters. The final clustering results revealed distinct customer segments based on their purchasing behaviours and regional differences. These insights can be used to tailor marketing strategies and improve customer engagement.

The clusters were visualized using PCA to reduce dimensionality, and the scatter plot provided clear segmentation into 4 distinct groups, confirming the effectiveness of the chosen clustering technique.