Customer Segmentation Clustering Report

The objective of this analysis is to segment customers based on their purchasing behaviour. By leveraging KMeans clustering, we aimed to identify distinct customer groups that can help businesses target specific customer segments with tailored marketing strategies. The clustering was performed on transaction data, which was aggregated at the customer level, including features like total spending and total quantity purchased.

Clustering Overview

- **Clustering Algorithm**: KMeans
- **Optimal Number of Clusters (k)**: 9
- **Davies-Bouldin Index (DBI)**: 0.6146

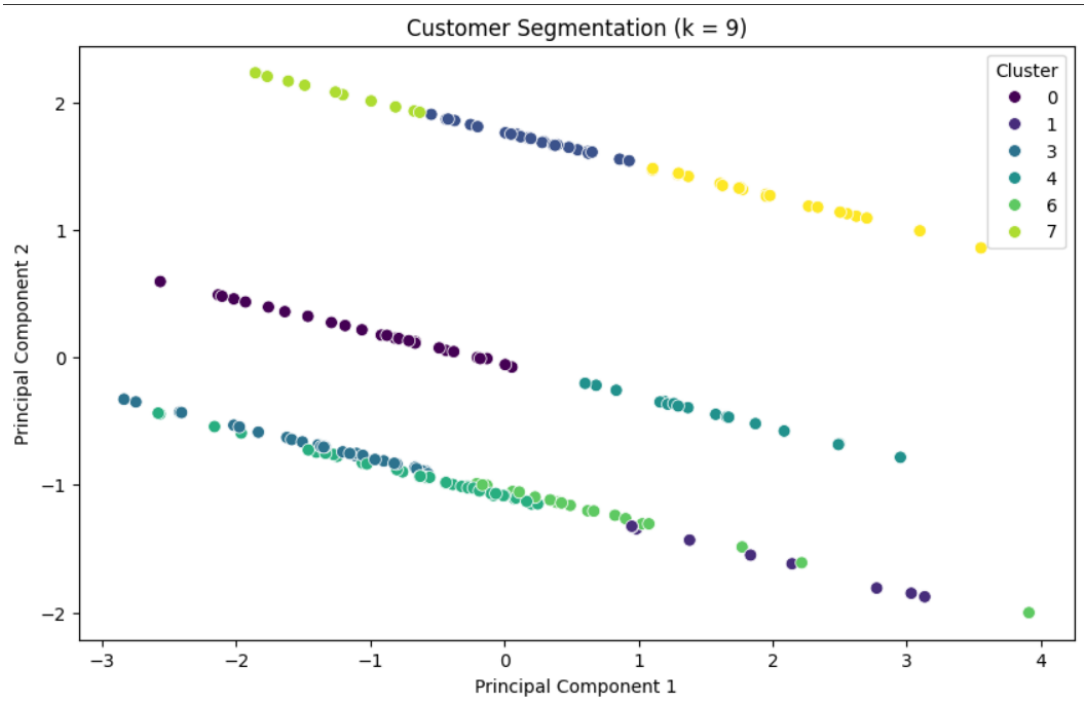
Key Clustering Metrics

- **Davies-Bouldin Index (DBI)**:
    - The **DBI** value is a measure of how well-separated the clusters are, with lower values indicating better-defined clusters. The DBI for this clustering model is **0.6146**, which suggests that the clusters are relatively well-separated and compact.
- **Cluster Compactness and Separation**:
    - The chosen **k=9** clusters show a good balance between compactness and separation, as indicated by the DBI score.
- **Optimal Number of Clusters (k=9)**:
    - The **optimal number of clusters** was determined by testing a range of cluster sizes (from 2 to 10) and selecting the number that minimized the DBI. **k=9** clusters provided the best separation and compactness for the dataset.
- **Cluster Distribution**
- Each of the nine clusters represents a group of customers with similar purchasing behavior. These clusters can be used for further analysis to understand patterns such as high-spending customers or frequent

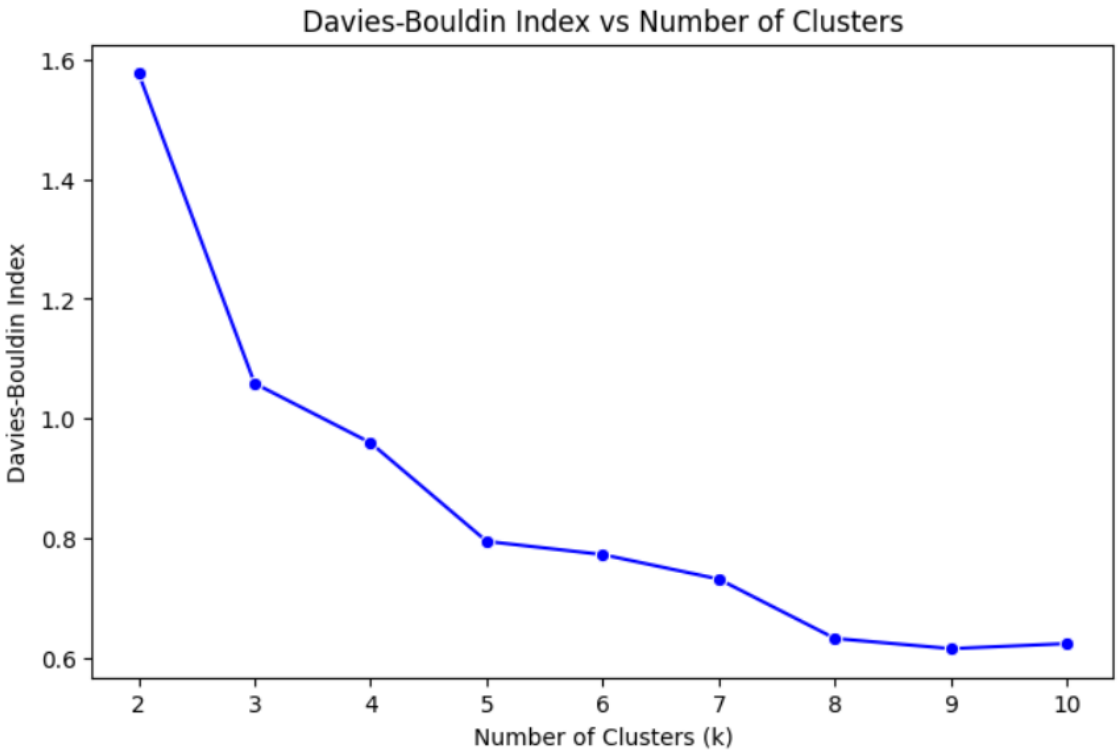
buyers. By analyzing the characteristics of each cluster, businesses can design more targeted marketing campaigns and optimize product offerings.

Cluster Visualization

The **2D PCA plot** below visualizes the clusters in two dimensions using the first two principal components. This visualization helps to understand the spatial distribution of the clusters and provides a clear view of how well-separated the clusters are.



This plot shows the Davies-Bouldin Index across different values of **k** (the number of clusters). It helps visualize the relationship between the number of clusters and the compactness/separation of clusters.



Cluster Characteristics:

- The segmentation identifies 9 distinct groups, each representing customers with similar purchasing behaviors.
- Detailed analysis of each cluster can help identify key customer profiles, such as high-value customers, frequent buyers, or customers with low spending but high transaction frequency.

Conclusion

The KMeans clustering model successfully segmented customers into **9 distinct groups**. The **Davies-Bouldin Index of 0.6146** indicates that the clusters are well-separated and compact. The results provide valuable insights into customer behaviors and can serve as a foundation for targeted marketing strategies and further analysis.