Clustering Results Report

1. Number of Clusters Formed

The optimal number of clusters for the dataset was determined by evaluating the Davies-Bouldin Index (DB Index) across different values of k(number of clusters). Through this analysis, the most efficient number of clusters was found to be **9**.

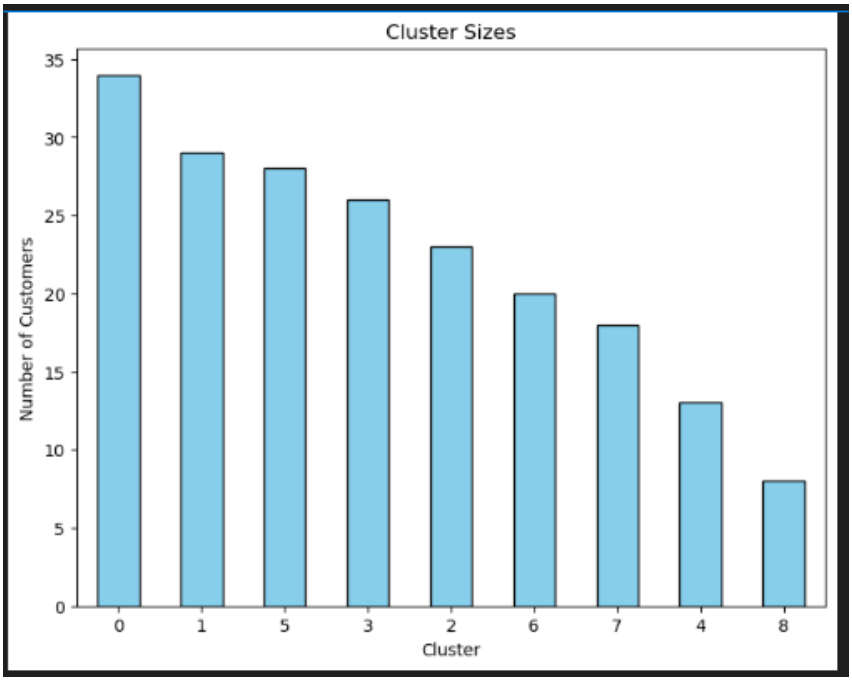
2. Davies-Bouldin Index (DB Index)

The Davies-Bouldin Index was used to assess the quality of the clustering solution. The DB Index measures the average similarity ratio of each cluster with the cluster that is most similar to it. Lower values of the DB Index indicate better clustering, with more distinct and well-separated clusters.

After performing the clustering with 9 clusters, the final Davies-Bouldin Index value was calculated (1.0289431154491853).

3. Cluster Visualization

The distribution of customers across the 9 clusters was visualized using a pairplot. The pairplot provides a visual understanding of how the clusters differ across various features. The clusters are represented by different colors, and the diagonal plots show the distribution of individual features within each cluster. This visualization helps identify any noticeable separations between clusters.



4. Cluster Sizes

The sizes of the clusters were visualized with a bar plot, which showed the number of customers assigned to each cluster. The cluster sizes give insight into the distribution of customers within the clusters.

The bar plot suggests that the clusters are somewhat evenly distributed, but with some clusters having significantly more customers than others. This could indicate that some clusters represent more general customer behaviors, while others might represent more niche groups.

5. Cluster Breakdown

The number of customers in each of the 9 clusters can be reviewed at the end, these are as follows:

CLUSTER_LABELS	NUMBER OF CUSTOMERS
0	34
1	29
5	28
3	26
2	23
6	20
7	18
4	13
8	8