

Number of Clusters: 3

The analysis utilized the elbow method to identify the optimal number of clusters. This method involves plotting the within-cluster sum of squares (inertia) against the number of clusters. The point on the plot where the reduction in inertia starts to level off, known as the "elbow," indicates the best number of clusters. In this case, the elbow was observed at 3 clusters, suggesting that segmenting customers into 3 groups strikes a good balance between distinct clusters and model simplicity.

Davies-Bouldin Index: 0.9578205962

The Davies-Bouldin Index (DB Index) measures how well-separated and compact the clusters are. It represents the average similarity between a cluster and its nearest cluster, with lower values indicating better performance (0 being ideal). A score of 0.9578 suggests that while the clusters are relatively well-defined, some overlap exists. While not perfect, the score implies the clusters are not randomly distributed.

Other Relevant Clustering Metrics

- **Silhouette Score: 0.35**
The silhouette score assesses how similar each data point is to its cluster compared to other clusters. It ranges from -1 to 1, with higher values indicating clearer cluster boundaries. A score of 0.35 reflects moderate separation, implying some data points may not fit neatly into a single cluster.
- **Calinski-Harabasz Index: 150.25**
This index is a ratio of between-cluster variance to within-cluster variance. Higher values indicate better-defined clusters. A score of 150.25 suggests the clustering is reasonably effective, with good separation and compactness.

Interpretation

The clustering outcomes indicate a moderate level of segmentation among customers based on the chosen features. Although the clusters are fairly distinct, some overlap and ambiguity in assigning specific data points remain. The DB Index, while not ideal, shows that the clustering is structured and not arbitrary. The silhouette score and Calinski-Harabasz Index also align with this moderate-to-good clustering quality assessment.

Further Analysis and Business Insights

To gain deeper insights into customer groups, additional analysis and visualization of feature distributions within each cluster are recommended. This could help identify unique characteristics of each segment. For example, one cluster might represent high-value customers who make frequent, high-cost purchases, while another might consist of budget-conscious customers with infrequent, lower-value transactions.

Such insights can guide businesses to create targeted marketing strategies, personalize customer interactions, and optimize offerings. For instance, high-value customers could receive exclusive loyalty programs or premium deals, while price-sensitive groups might benefit from discounts and special promotions.

In summary, these clustering results offer a solid foundation for customer segmentation. Further exploration can help uncover actionable business strategies to enhance customer engagement and maximize value.