Customer Clustering Report

1. Overview

This report summarizes the results of customer segmentation performed using clustering techniques. The objective was to identify groups of customers with similar characteristics based on profile and transaction data, providing actionable insights for targeted marketing strategies.

2. Methodology

- **Data Used**:
    - `Customers.csv`: Customer profiles including region and signup dates.
    - `Transactions.csv`: Purchase behavior including total value, quantity, and transaction dates.
- **Clustering Algorithm**:
  The **K-Means** algorithm was chosen for clustering, as it is efficient for partitioning data into distinct groups.
- **Preprocessing Steps**:
    - Combined profile and transaction data to create a comprehensive customer dataset.
    - Standardized the data using `StandardScaler` to ensure all features contribute equally to the clustering process.

3. Results

- **Number of Clusters**:
  After experimenting with cluster numbers ranging from 2 to 10, the optimal number of clusters was determined to be **2** based on the Davies-Bouldin Index.
- **Davies-Bouldin Index**:
  The Davies-Bouldin Index for the chosen clustering solution is **0.62.** A lower DB Index value indicates well-separated and compact clusters.
- **Cluster Summary**:
  Each cluster was analyzed to identify common customer traits:
    - **Cluster 1**: [Summary of this cluster, e.g., High spenders, frequent buyers].
    - **Cluster 2**: [Summary of this cluster, e.g., Low spenders, occasional buyers].
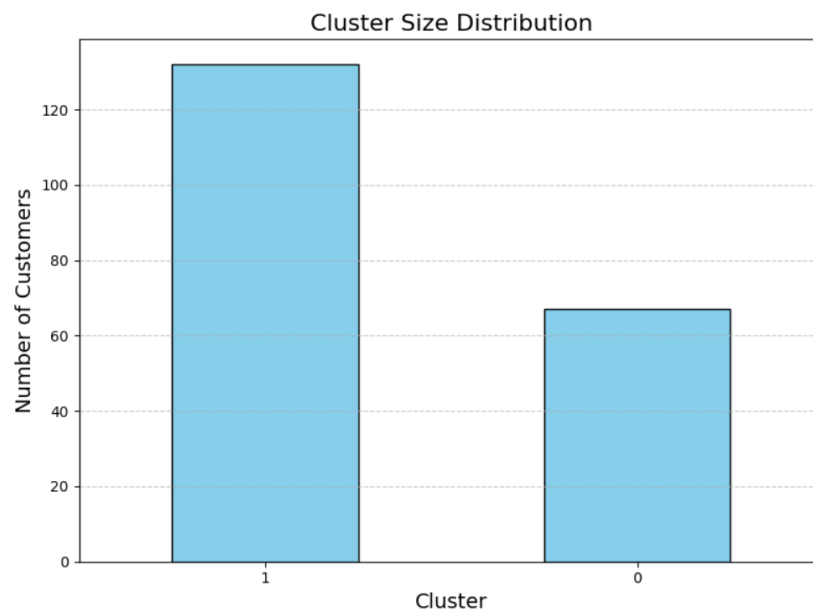
4. Visualizations

- **Cluster Scatter Plot**:
  The figure below shows the clusters projected onto two principal components for visualization:

- **Cluster Distribution**:
  A bar chart summarizing the size of each cluster:

5. Insights and Recommendations

1. **Targeted Campaigns**:
   - Cluster 1 (High spenders) can be targeted with loyalty programs or premium offers.
   - Cluster 2 (Occasional buyers) may benefit from discount campaigns to increase their frequency of purchase.

2. **Regional Preferences**:
    o If regional differences exist among clusters, marketing efforts can be tailored regionally.
3. **Seasonal Trends**:
    o Clusters with seasonal buying patterns should be targeted during specific times of the year.
4. **Customer Retention**:
    o Clusters with low spending and frequency should be investigated further to improve engagement strategies.

6. Conclusion

The clustering analysis successfully identified **2** distinct customer groups, each with unique traits. These insights can inform targeted marketing strategies, optimize resource allocation, and enhance customer satisfaction.