Customer Segmentation Analysis Documentation

Objective

The goal of this analysis is to perform customer segmentation by leveraging both profile and transaction data. This allows us to identify distinct customer groups, enabling targeted marketing strategies and personalized services.

Steps of Analysis:

1. Loading Datasets

- The analysis begins with importing two key datasets: customer profiles and transaction data. These datasets contain essential information such as demographic details, purchase history, and product preferences.

2. Data Preprocessing and Merging

- Transaction data is aggregated at the customer level to derive features such as total spend, average spend, total quantity purchased, and the number of unique products.
- The aggregated transaction data is merged with customer profile information to create a comprehensive dataset.
- Categorical variables (e.g., Region and Category) are encoded using label encoding, and numerical features are normalized to ensure uniform scaling.

3. Clustering Algorithm Application

- The KMeans clustering algorithm is applied to segment customers based on their normalized profiles.
- The optimal number of clusters is determined by evaluating the Davies-Bouldin Index across a range of clusters (2 to 10).

4. Evaluation Metrics

- Key metrics such as the Davies-Bouldin Index and Silhouette Score are used to assess the quality of clustering.
- The Davies-Bouldin Index is minimized to identify the optimal clustering configuration.

5. Visualization

- Clusters are visualized using pair plots to explore relationships between features and assess the separation of clusters.

- Bar plots of cluster centers provide insights into the characteristics of each cluster, highlighting distinguishing features.

6. Results and Output

- The analysis identified an optimal number of clusters, and each customer was assigned to a specific cluster.
- The clustered data is saved to a file for further analysis and decision-making.
- Key metrics, including the optimal number of clusters, Davies-Bouldin Index, and Silhouette Score, are reported.

Conclusion

This segmentation analysis provides actionable insights into customer behavior by grouping similar customers based on their profiles and transaction patterns. These insights can be leveraged for personalized marketing, improved customer experience, and strategic decision-making.