

Business Insights

1. **Feature Engineering:** The script aggregates total revenue and quantity purchased for each customer and converts categorical data (region) into numerical format using one-hot encoding.
2. **Standardization:** The features are standardized to ensure that they are on the same scale, which is important for clustering algorithms.
3. **Clustering:** The script uses the K-Means algorithm to segment customers into clusters. It also calculates inertia and silhouette scores to help determine the optimal number of clusters.
4. **DB Index Calculation:** The Davies-Bouldin Index is calculated to evaluate the quality of the clustering.
5. **Visualization:** PCA is used to reduce the dimensionality of the data for visualization purposes, and a scatter plot is created to visualize the clusters.
6. **Reporting:** The number of clusters formed and the DB Index value are printed.