Dataset info:

RangeIndex: 6340 entries, 0 to 6339

Columns: 219 entries, group to interest217

dtypes: float64(217), int64(1), object(1)

After filling the Nan values with 0.

A graph with a line

Description automatically generated

Silhouette Score for K-means clustering on PCA-reduced training data: 0.4184891089875287

A graph of a diagram

Description automatically generated with medium confidence

For DBSCAN:

Best epsilon: 0.5

Best min\_samples: 9

Best silhouette score on train data: 0.42203735409656

Silhouette Score for DBSCAN clustering on PCA-reduced training data: 0.42203735409656

Silhouette Score for K-means clustering on test set: 0.4009797508064607

A screen shot of a graph

Description automatically generated