Conclusions

Based on the provided table and considering the different categories of evaluation metrics (silhouette score, Calinski-Harabasz index, Davies-Bouldin index), here are the best-performing combinations for each category which I have concluded:

1. Best for Silhouette Score:

The combination of PCA transformation followed by normalization (T+N) seems to consistently yield the highest silhouette scores across different numbers of clusters and clustering algorithms.

2. Best for Calinski-Harabasz Index:

The combination of PCA transformation followed by normalization (T+N) also tends to perform well in terms of the Calinski-Harabasz index, consistently providing higher values compared to other preprocessing techniques and combinations.

3. Best for Davies-Bouldin Index:

The best-performing combination for the Davies-Bouldin index varies slightly depending on the number of clusters and clustering algorithm used. However, in general, PCA transformation alone (T) or combined with normalization (T+N) tends to produce lower Davies-Bouldin index values compared to other preprocessing techniques and combinations.

- For the K-Means, PCA transformation alone (T) or combined with normalization (T+N) generally produces lower values, indicating better clustering qual
- For Hierarchical Clustering,it slightly varies here PCA transformation alone (T) or combined with normalization (T+N) generally leads to lower values.
- K-Mean Shift, PCA transformation alone (T) or combined with normalization (T+N) generally produces lower values.

Overall, based on these observations, the combination of PCA transformation followed by normalization (T+N) appears to be the most effective across different evaluation metrics. However, it's essential to consider the specific characteristics of the dataset and the clustering goals when determining the best approach.

From observations K-Mean shift is best algorithm and best number of clusters is 3 with best silhouette 0.6577