

## Performance using different clustering techniques on various parameters

Parameters	No Data Processing			Using Normalization			Using Transform			Using PCA			Using T+N			T+N+PCA		
	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5
Silhouette	0.5711	0.5578	0.5169	0.5479	0.5377	0.5489	0.5711	0.5587	0.5175	0.5711	0.5669	0.5640	0.5711	0.5573	0.5239	0.5711	0.5587	0.5489
Calinski-Harabasz	561.81	704.92	710.01	498.86	702.20	787.04	561.81	702.67	686.93	561.81	696.83	703.27	561.81	595.80	689.79	561.81	702.67	787.04
Davies-Bouldin	0.5342	0.5441	0.5542	0.5531	0.5453	0.5450	0.5342	0.5460	0.5306	0.5342	0.5486	0.4834	0.5342	0.5026	0.5174	0.5342	0.5460	0.5450

Table 1: Using K-Means Clustering

Parameters	No Data Processing			Using Normalization			Using Transform			Using PCA			Using T+N			T+N+PCA		
	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5
Silhouette	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075
Calinski-Harabasz	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22
Davies-Bouldin	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513

Table 2: Using Agglomerative Clustering

Parameters	No Data Processing			Using Normalization			Using Transform			Using PCA			Using T+N			T+N+PCA		
	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5
Silhouette	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075	0.5645	0.5607	0.5075
Calinski-Harabasz	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22	552.85	670.63	684.22
Davies-Bouldin	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513	0.5357	0.5536	0.5513

Table 3: Using Birch Clustering

Parameters	No Data Processing			Using Normalization			Using Transform			Using PCA			Using T+N			T+N+PCA		
	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5	c=3	c=4	c=5
Silhouette	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803	0.2803
Calinski-Harabasz	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37	3.37
Davies-Bouldin	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477	0.4477

Table 4: Using Spectral Clustering

The reason why the values in **BIRCH clustering** and **Agglomerative Clustering** are the same becaues they both are hierarchical in nature. Some other reacons may include usage of same distance metrics, number of clusters and preprocessing techniques.

---

<sup>0</sup>T : Transformation    N : Normalization    PCA : Principal Component Analysis