

Clustering Analysis Report

Number of Clusters: Both **K-Means** and **GMM** algorithms were configured to create 4 clusters (labeled as Clusters 0, 1, 2, and 3). The visualization shows consistent cluster formation between both methods, with similar spatial distribution patterns.

Clustering Quality Metrics:

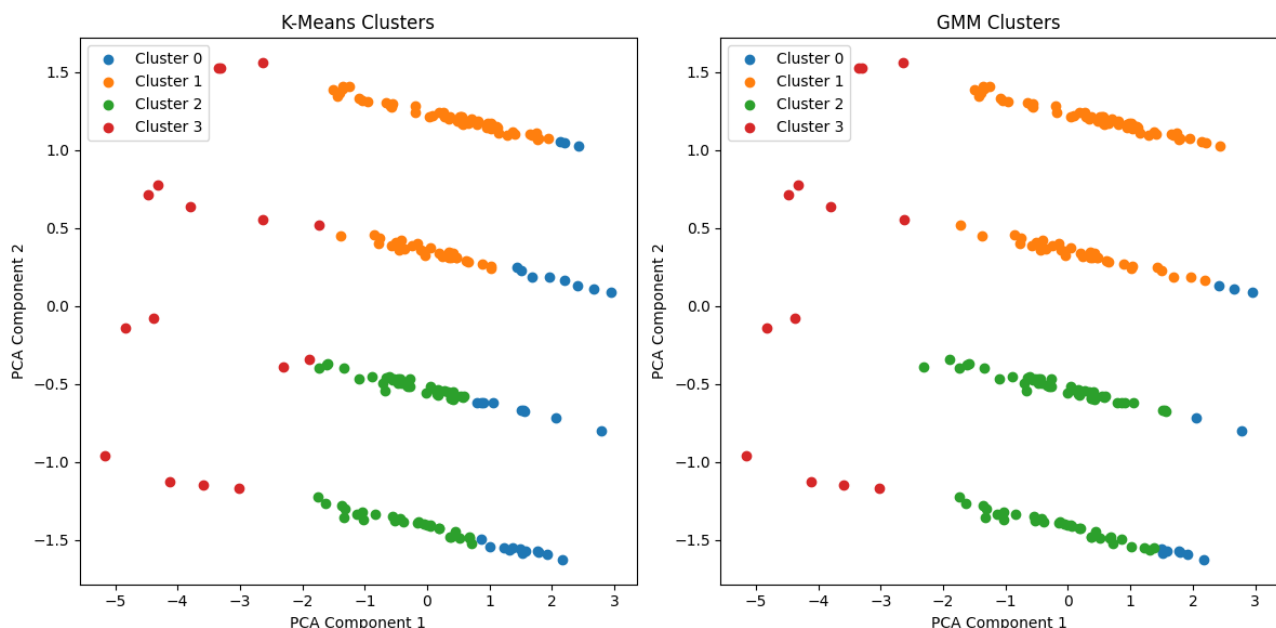
1. Davies-Bouldin (DB) Index:

- K-Means: 0.8077
- GMM: 0.8307 The DB index values are relatively similar between both methods, with K-Means performing slightly better (lower is better). These values indicate reasonably good cluster separation, as values closer to 0 indicate better-defined clusters.

2. Silhouette Score:

- K-Means: 0.3949
- GMM: 0.3686 The Silhouette scores are also comparable between methods, with K-Means showing marginally better performance. These moderate positive values suggest that points are generally well-matched to their assigned clusters, though there is room for improvement.

Visual Analysis:



The PCA plot reveals several interesting patterns:

- Four distinct linear or curved formations in the data
- Clear separation between cluster groups
- Similar cluster assignments between K-Means and GMM
- Data points appear to follow trajectory-like patterns
- Clusters show varying densities and distributions

Comparative Performance: K-Means slightly outperforms GMM according to both metrics:

- DB Index: 3% better (0.8077 vs 0.8307)
- Silhouette Score: 7% better (0.3949 vs 0.3686)

These results suggest that while both algorithms provide acceptable clustering solutions, K-Means might be marginally more suitable for this particular dataset's structure.