Clustering Lab Report

AKSH PATEL

September 26, 2025

1 Analysis and Questions

Optimal Clusters: Both K-Means and Hierarchical Clustering suggest $k = 5$ as the optimal number of clusters, supported by the Elbow Method and the dendrogram cut.

Cluster Comparison: K-Means and Hierarchical Clustering produce almost identical results with 5 clear groups. DBSCAN also finds 5 main clusters but marks some points as noise, giving slightly different boundaries in sparse regions.

DBSCAN Performance: DBSCAN successfully detected 5 clusters and about 22 noise points, effectively identifying outliers between clusters. Unlike K-Means and Hierarchical Clustering, it leaves anomalous points unassigned, resulting in cleaner clusters.

Algorithm Suitability: K-Means is best suited here due to roughly spherical, evenly sized clusters. DBSCAN is helpful for spotting outliers, while Hierarchical Clustering gives similar results but is slower.

Real-World Use: Mall marketers can use these groups for promotions. For instance, customers with high income but low spending could get luxury discounts or personalized offers to boost spending.

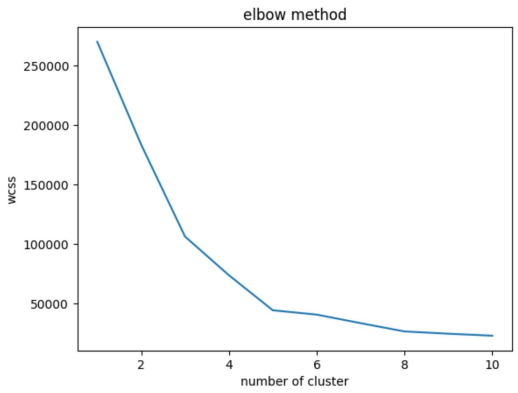

Figure 1: Elbow plot

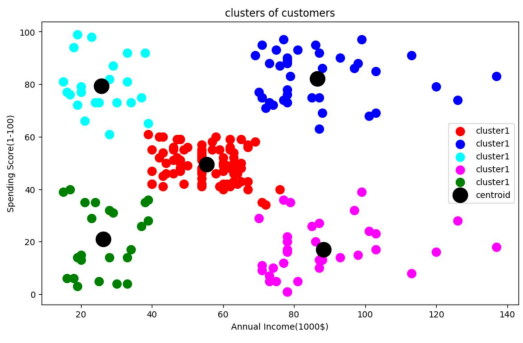

Figure 2: K-mean Clustering

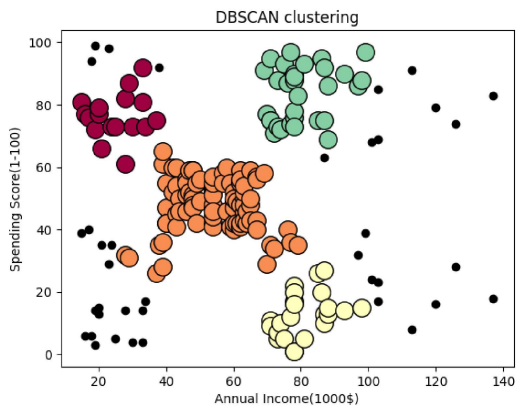

Figure 3: DBSCAN