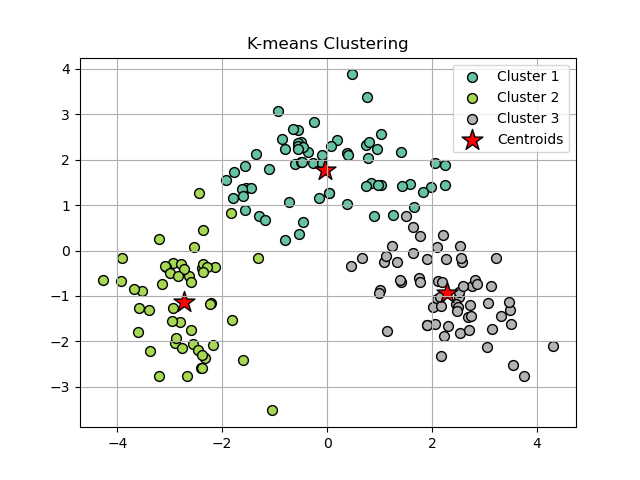
Clustering Comparison Report

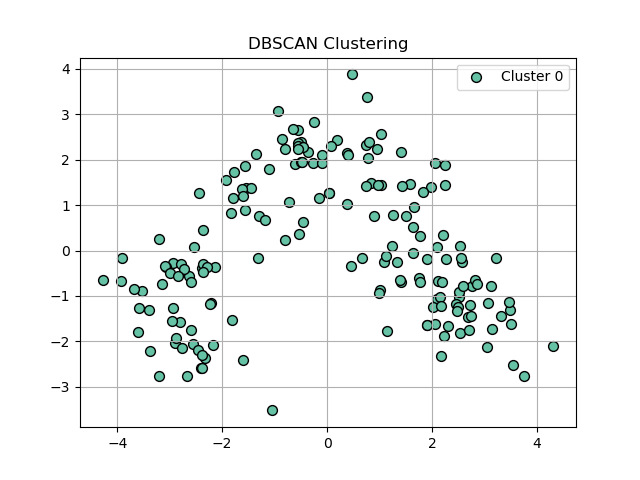
# K-Means Clustering

K-means clustering was performed with 3 clusters. The results, as visualized in the figure below, show distinct clusters with centroids marked.



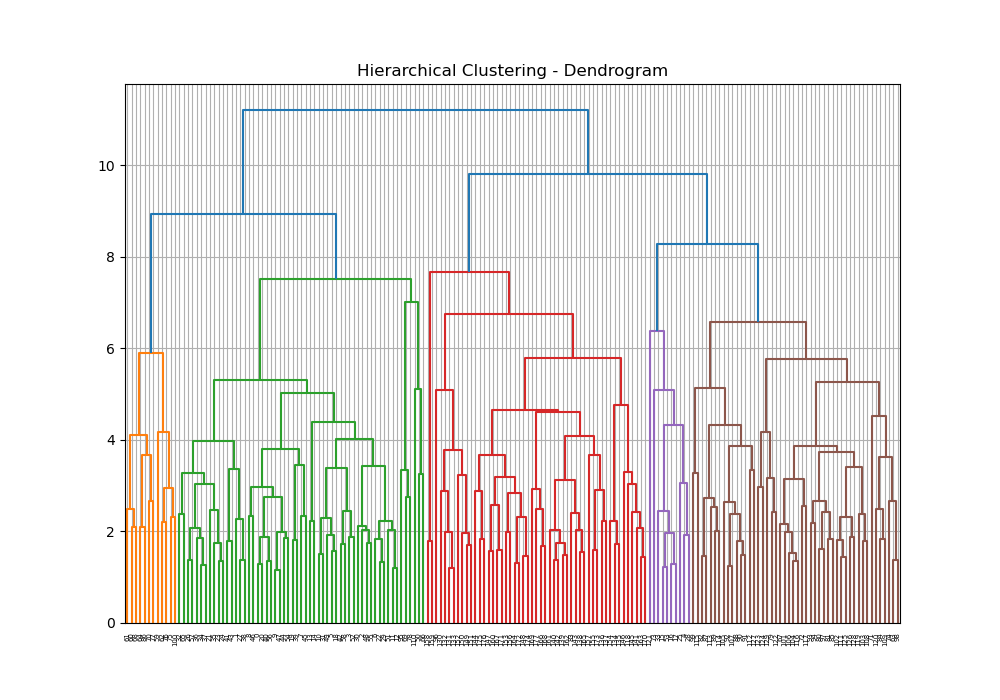
# DBSCAN Clustering

DBSCAN clustering identified clusters based on density. The figure below demonstrates the identified clusters.



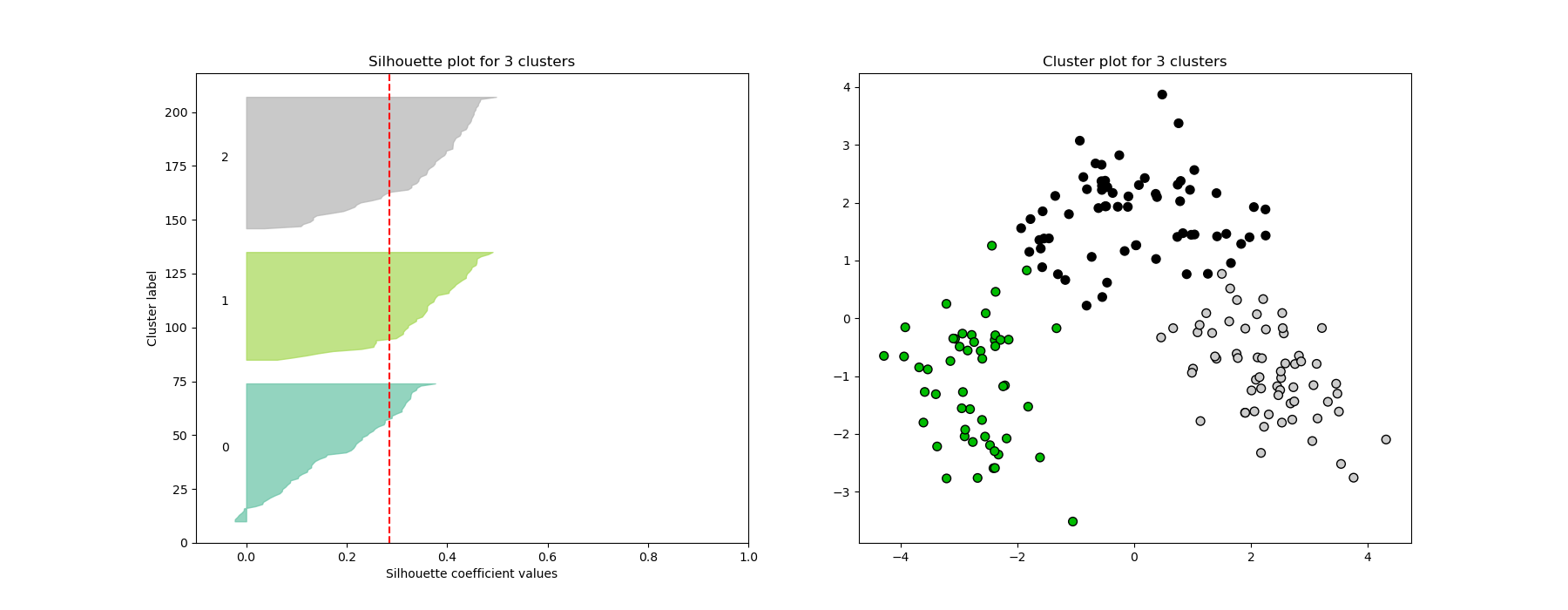
# Hierarchical Clustering

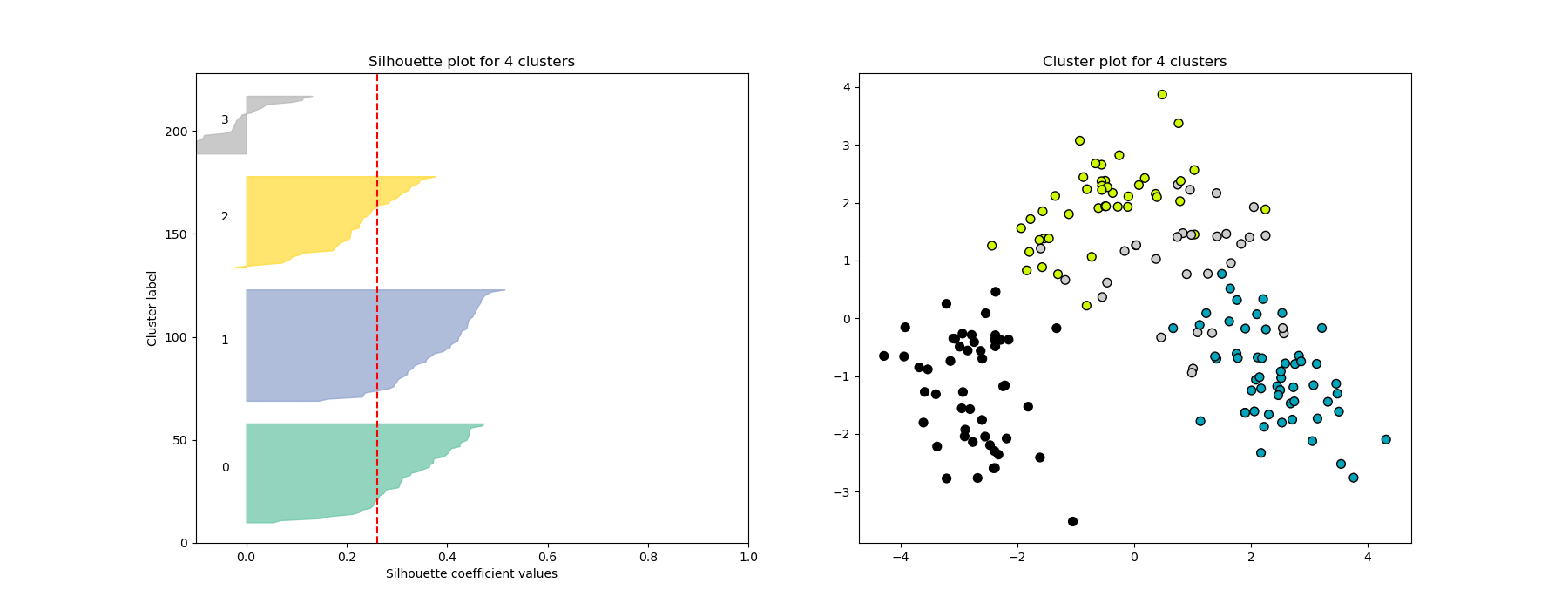
Hierarchical clustering was performed with 3 clusters. The dendrogram below illustrates the clustering hierarchy.

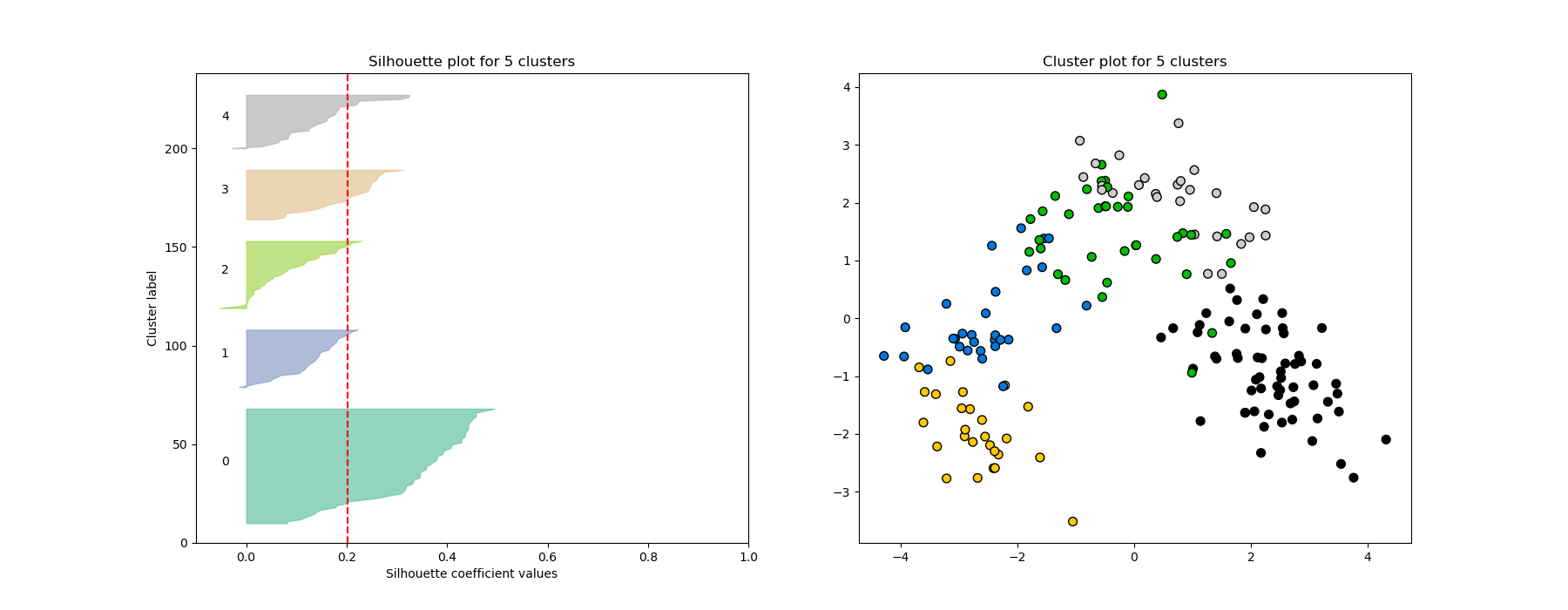


# Silhouette Analysis

Silhouette analysis for K-means clustering with different numbers of clusters (3, 4, and 5) was performed. The figures below show the silhouette plots.







Clustering Algorithm Comparison Report

# K-Means Clustering

K-Means clustering is a distance-based method that partitions the data into k clusters, where each data point belongs to the nearest centroid. In this experiment, K-Means was able to identify clear clusters in the Wine dataset. The algorithm is computationally efficient and performs well on datasets with spherical clusters.

## K-Means Clustering Result (PCA Projection)

The figure below shows the clustering results after applying K-Means with 3 clusters, projected onto two principal components.

# DBSCAN Clustering

DBSCAN is a density-based clustering algorithm. It struggled to form meaningful clusters in this experiment, indicating that the Wine dataset may not exhibit clear density differences that DBSCAN requires. Despite being good at handling noise and irregularly shaped clusters, DBSCAN could not identify useful clusters in this case.

## DBSCAN Clustering Result (PCA Projection)

The figure below shows the clustering results after applying DBSCAN with adjusted parameters, projected onto two principal components.

# Hierarchical Clustering

Hierarchical clustering creates a hierarchy of clusters by recursively merging or splitting them. In this experiment, it produced clear clusters in the Wine dataset and was visualized effectively using a dendrogram. However, as with many clustering techniques, it can become computationally expensive for large datasets.

## Hierarchical Clustering Dendrogram

The figure below shows the hierarchical clustering as a dendrogram.

## Hierarchical Clustering Result (PCA Projection)

The figure below shows the clustering results after applying Hierarchical clustering with 3 clusters, projected onto two principal components.

# Conclusion

## K-Means Clustering:

K-Means is effective for datasets with well-separated and compact clusters. In this case, it was able to partition the Wine dataset into three clusters with good results. However, K-Means assumes the clusters are spherical and of equal size, which might not always be true in real-world datasets.

## DBSCAN Clustering:

DBSCAN was unable to produce meaningful clusters for the Wine dataset, likely due to the lack of clear density variations. While DBSCAN is excellent for finding arbitrary-shaped clusters and handling noise, it requires careful parameter tuning and may struggle in high-dimensional or densely packed datasets like Wine.

## Hierarchical Clustering:

Hierarchical clustering effectively identified clusters in the Wine dataset and provided a clear visual representation of cluster relationships through the dendrogram. It is useful when understanding the hierarchical structure between clusters is important, but it may become computationally expensive for larger datasets.

## Recommendation for Real-World Use:

For datasets like the Wine dataset, K-Means is recommended due to its simplicity and effectiveness at finding clear, well-separated clusters. Hierarchical clustering is useful when insight into the relationships between clusters is needed. DBSCAN should be considered for datasets with noise or when irregular cluster shapes are expected, but it may struggle without the right density parameters.