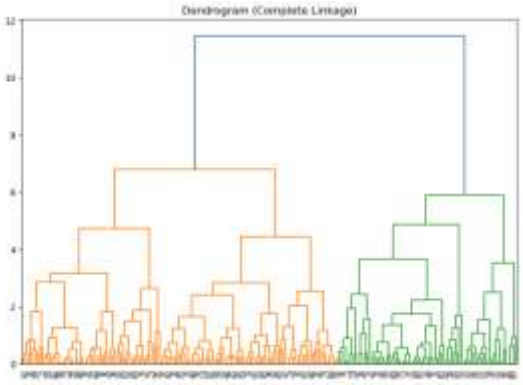
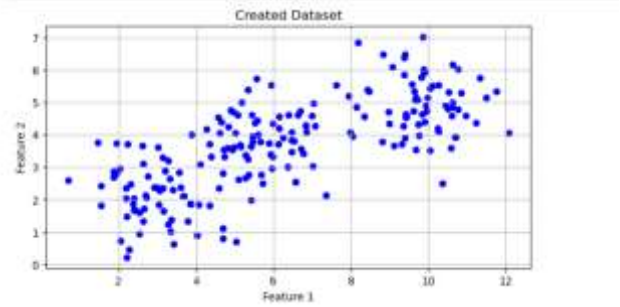
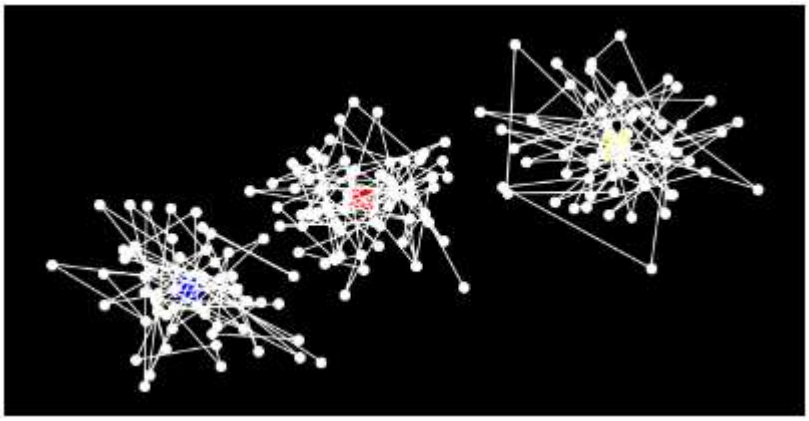
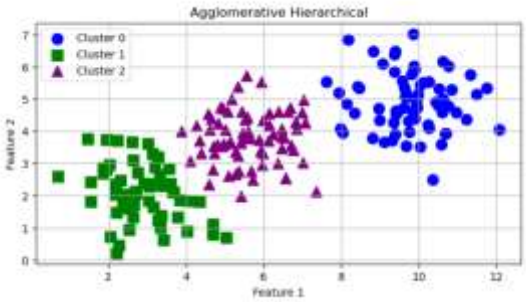
Pass Task 7.1P: K-Means and Hierarchical Clustering









Two crucial unsupervised learning algorithms—K-Means and Agglomerative Hierarchical clustering—are applied practically in this Python code. The method first uses `make blobs` to create a synthetic dataset, from which it generates 200 data points distributed among three different clusters. The dataset is subsequently split into three clusters using the K-Means algorithm, with the centroids prominently shown. This illustrates how K-Means iteratively improves cluster centers. The distinct visual representation, including colored clusters on a black background, highlights the algorithm's capacity to efficiently group related data points.

Afterwards, Agglomerative Hierarchical clustering is implemented by the code, which merges the nearest clusters one after the other to create a hierarchical tree. The whole linkage approach used to compute the distance matrix is projected on a dendrogram, providing a detailed perspective of the clustering hierarchy. The resulting clusters are visualized with unique markers. The relevance of data visualization in evaluating machine learning results is emphasized by this activity, which gives a practical grasp of both clustering algorithms. The understanding of these potent analytical tools is further improved by reinforcing the ideas of cluster centroids, hierarchical linkages, and the significance of linkage criteria in hierarchical clustering.