In this assignment, we analyzed audio characteristics from the Spotify/YouTube dataset by focusing on three key features: liveness (recording type), energy (track intensity), and loudness (volume level). Both K-means clustering (with 3D visualization) and hierarchical clustering (individual dimensions) were utilized to identify meaningful patterns in the data.

For K-Means Clustering, the dataset was pre-processed and scaled for optimal k determination. After the elbow graph suggested K=5 as the optimal number of clusters, a 3D scatter plot was implemented with cluster centers and interpretable labels. Here is the semantic cluster interpretation based on their characteristics:

1. Moderate Studio Recording: Average values across all dimensions
2. High-Energy Studio Banger: High energy and loudness with moderate liveness
3. Quite Ambient: Low values across all dimensions
4. Live High-Energy Performance: High liveness and energy
5. Hybrid Live-Studio Track: Moderate liveness with high energy

For hierarchical clustering, we performed on each individual dimension (liveness, energy, and loudness): started with the dendrogram analysis to visualize potential clusters, then visualized the cluster distribution with 3 clusters (based on the dendrogram) and 5 clusters (consistent with K-MEANS) respectively.