Date: 03-30-2025

Name: Abdikarim Jimale

HW07 – Clustering

My program clustering the result to 3 groups of songs based on Liveness, Energy (how intense), Loudness (volume level), using the elbow method, and the best value for k was 3.

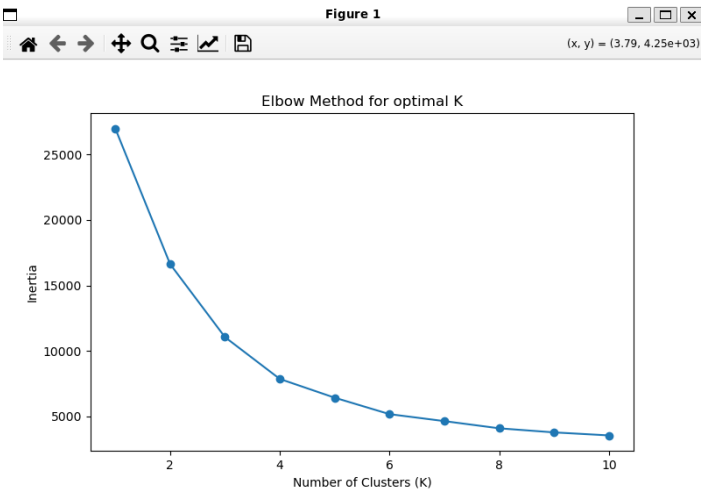
Clusters	Liveness	Energy	Loudness	Contains	result
0	low	medium	Medium	6,318 songs (71%)	Regular songs professionally produced
1	Very low	low	Very low	1,903 songs (21%)	Soft, chill song low volume and intensity
2	high	High	High	778 songs (8%)	Live / concert loud, with crowd noise

Result:

```
(DS5110) (base) abdikarim@DESKTOP-K3E9SL8:~/hw7-ajimale$ /home/abdikarim/miniconda3/envs/DS5110/bin/python /home/abdikari
m/hw7-ajimale/Data/K-means.py
MESA: error: ZINK: failed to choose pdev
glx: failed to create drisw screen
Cluster distribution: {np.int32(0): np.int64(6318), np.int32(1): np.int64(1903), np.int32(2): np.int64(778)}
Cluster Centers:
 [[-0.23010237  0.39331397  0.39137933]
 [-0.31631719 -1.47010501 -1.41027803]
 [ 2.63464039  0.39467767  0.26461831]]
Inertia (SSE): 11051.307105797765
```

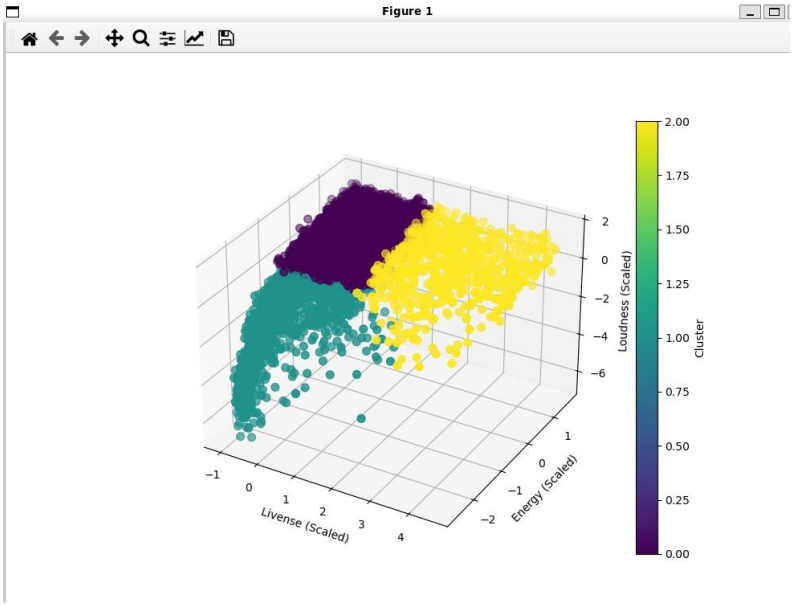
Elbow plot:

Show that 3 clusters is the best choice – after k=3, the decrease in error slows down.



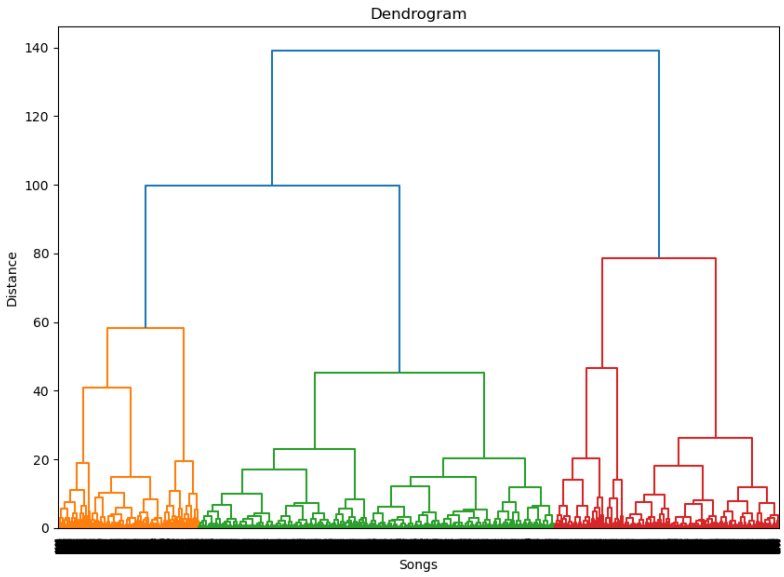
3D Scatter Plot:

The graphs show 3 clusters using different colors. Songs are grouped based on similarity.



Dendrogram (Hierarchical Clustering)

Confirms that the dataset naturally splits into 3 groups, consistent with the KMean results.



What this result might mean to me:

This helps me to understand and organize music by how it sounds, rather than just by genre artist. If I want loud energy music I can go to cluster 2 music. If I want to chill or relax music, I'd pick from cluster 1. For the general listening a mix energy and loudness, I'd pick from cluster 0. This analysis help make better playlists, recommend songs based on sound, and understand artists style and trends.