HW7 - K-Means (100 points)

1 Introduction (Due 3/31/2025)

Take the code that was written in class for K-means and make it work for a 3D data sets. Using the Spotify_ YouTube.csv data set, read in the following three columns: Liveness, Energy, Loudness. Using an elbow graph, find the optimal number of K and use that to visualize the data and groups based on that K. **Graphs should be appropriately labeled with an x, y, and z axis along with a title and legend. Then write what your results might mean to you.**

Then, run hierarchical clustering on the same three variables. Are there any distinct groups? If so, how would you define each group? **Graphs should be appropriately labeled. Then write what your results might mean to you.**

2 Grading (Out of 100 points)

- 60 points: Cover what you did to update the code to work for 3D data and visualization along with what number of K you found to be optimal and the graphs showing the results of running K-means and what your results mean.
- 40 points: Report your findings for running Hierarchical clustering.

3 How to turn in

Turn in the final report and code that you wrote to Githbub. Then, put a link to your github submission to Canvas.

Good Luck Cat