CSC 401, Weekly Coding Assignment #11, Fall 2024   Name_____

Instructions: For these weekly coding assignments, you will be asked to extend the examples from class to create custom code to answer the questions below. You will create an R code file that uses built-in datasets as the data sources. You will write the code, and an example showing that the code works. Be sure to include any packages in the code that are required for the functions to run (you may want to clear the environment in RStudio before your final check to make sure nothing is missing). The instructor will run the file to ensure that it works with no errors. Clearly label your code so it's clear which question/task is being responded to.

Submission:
A word document with any explanations (if needed), and a clearly labeled R code file.

Tasks/Questions:
1. Choose 3 distance metrics, other than Euclidean, and customize the k-means algorithm on the iris dataset. How do the different metrics perform relative to the Euclidean example from class?

2. Update the multiple runs (k-means) example to work on all three species instead of just the binary case.

3. Redo the spectral clustering example from lecture with three eigenvalues instead of 2 to see if this improves the fit on the iris dataset.

4. Apply spectral clustering, but with a different clustering algorithm than k-means (such as KNN for classification, or agglomerative clustering for true clustering, you can use a package for this part). Describe the results and compare to the original algorithm without first reducing the variable dimensions.