程式碼請見: hw3-kmeans-and-tsne.ipynb

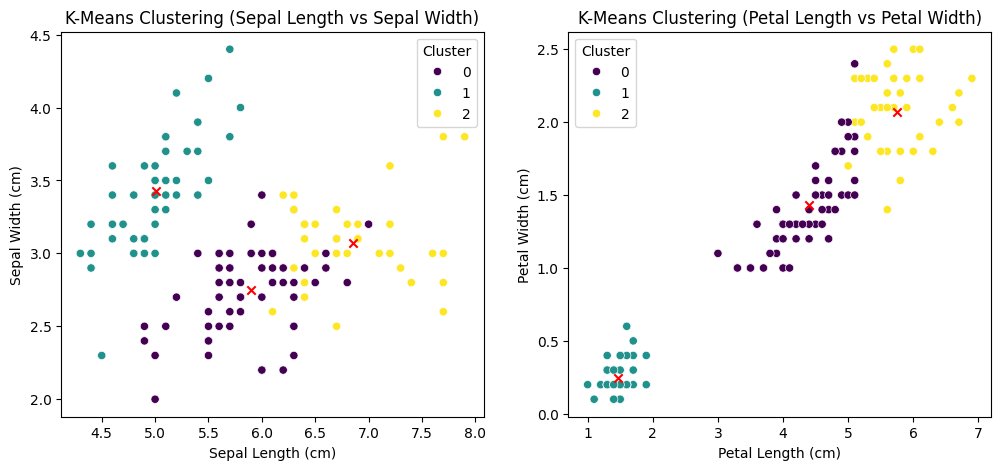
Please write a program to perform one of the following tasks on the IRIS data.

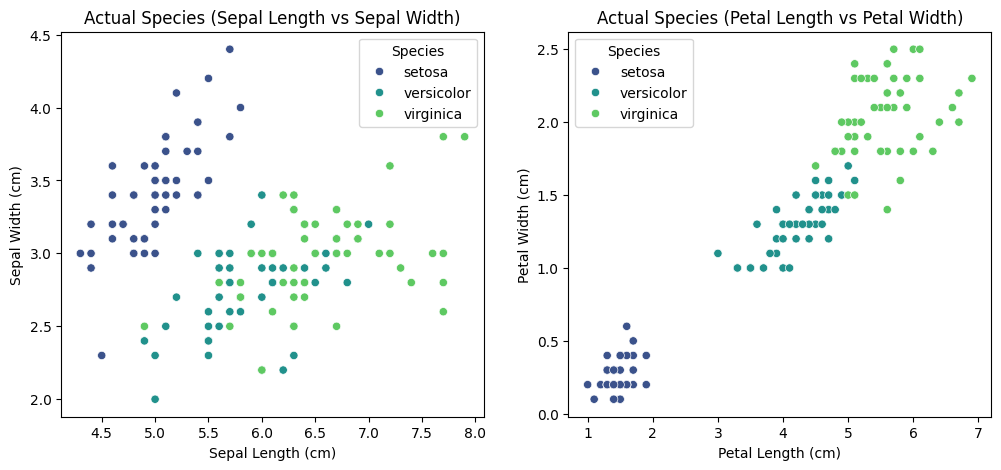
1. k-means clustering algorithm

In IRIS data every sample was labeled with a class of flower among three. Please perform clustering with three clusters, and compare the clustering result with the class labels and have a few discussions.

使用kmeans根據['Sepal.Length', 'Sepal.Width', 'Petal.Length', 'Petal.Width']

4個維度分成3群，然後兩個維度兩個維度視覺化。



真實label兩個維度兩個維度視覺化。

Kmeans的跟真實label比較的混淆矩陣、正確率。

Confusion Matrix:

[[50 0 0]

[ 0 48 2]

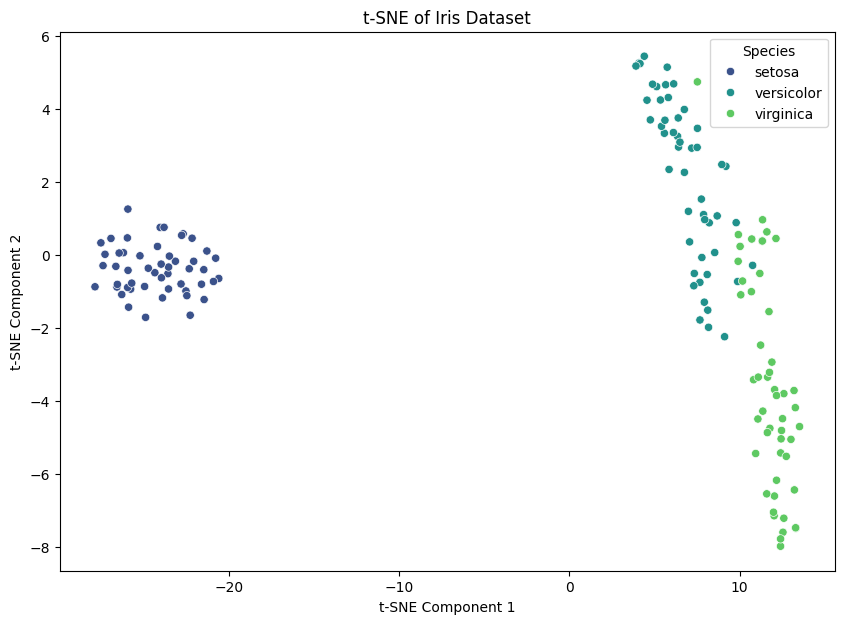
[ 0 14 36]]

Accuracy: 0.8933

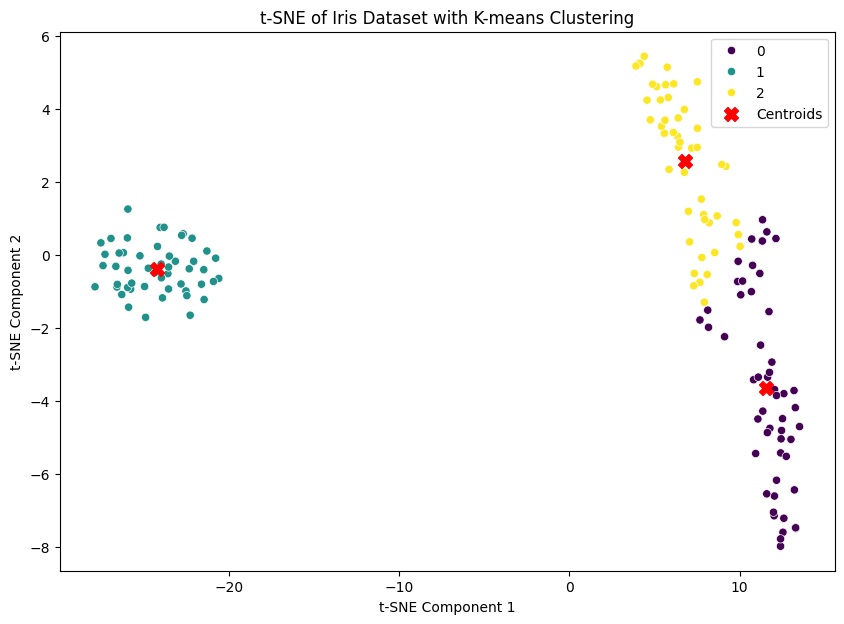
1. t-SNE for reducing every samples as 2D points

The 4D samples of IRIS dataset could be converted into 2D points through t-SNE algorithm, which may be further visualized on a plane. Please perform t-SNE to produce 2D points, and use three colors corresponding to the class labels to display the 2D points.

把4D資料降維到2D，使用tsne方法:



降維到2D的資料用K-means分3群:



並將分群結果與真實label比較:  
Confusion Matrix:

[[50 0 0]

[ 0 44 6]

[ 0 3 47]]

Accuracy: 0.9400