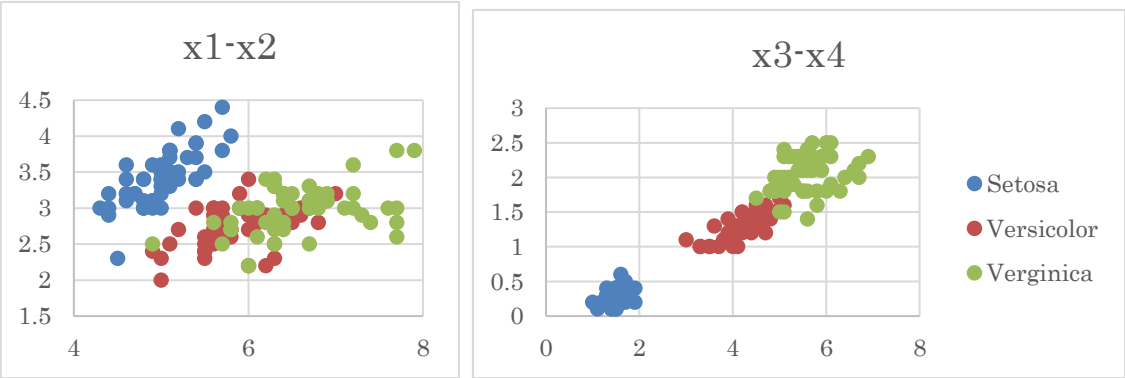
● Exercise 1

k-Means is the most basic technique for unsupervised clustering.
Develop the program source of k-Means by C language and apply it to Iris data set, which is composed of 4-D observations on 150 objects drawn from three Iris classes with 50 objects each: Iris Setosa, Iris Versicolour and Iris Virginica.
The details of the data set is available at: http://archive.ics.uci.edu/ml/datasets/Iris



The detail of k-Means is available on web sites such as Wikipedia.

"Note"
k-Means often suffer from "local minima" and brings different clustering results with different initialization.
In this exercise, you should apply your k-Means program with 100 different initial cluster centers. Then, you should report "how many different cluster partitions you found in the 100 trials" and "how often each of the partitions appeared".
Additionally, the values of the cost functions (objective functions) should also be compared.

"Deadline"
You should finish this exercise in 3 weeks. (by 31, May 2022)