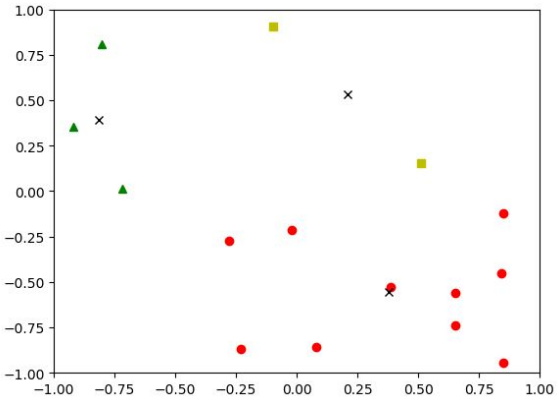
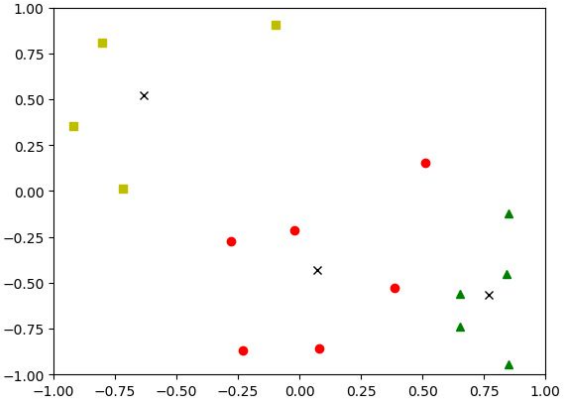
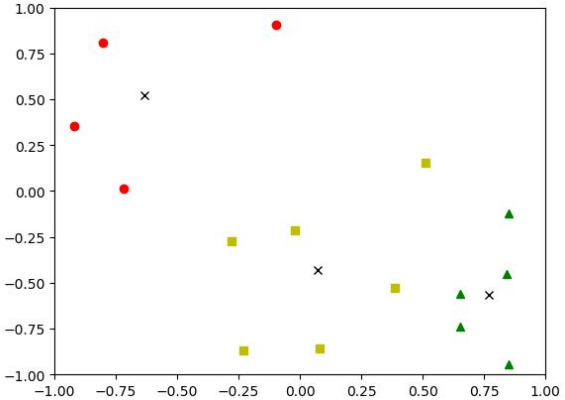
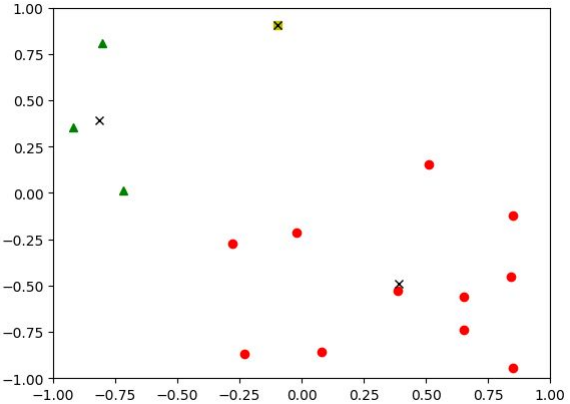
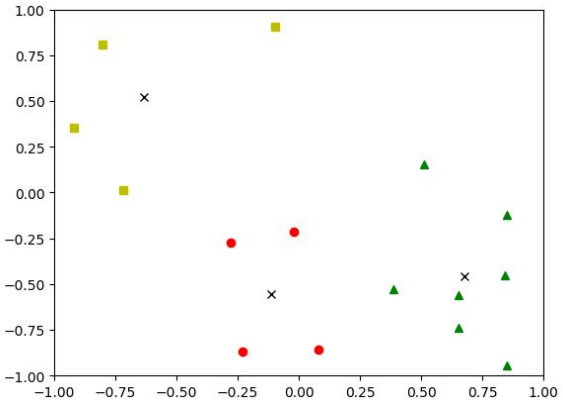
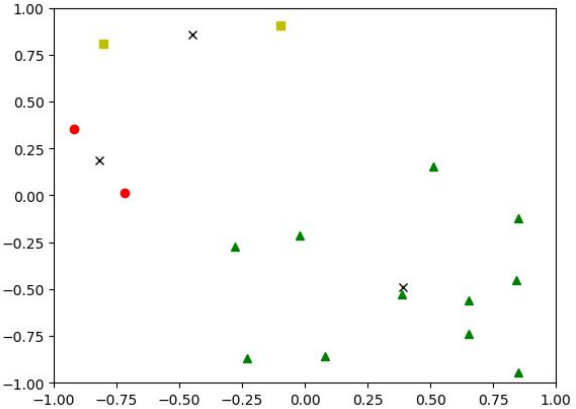
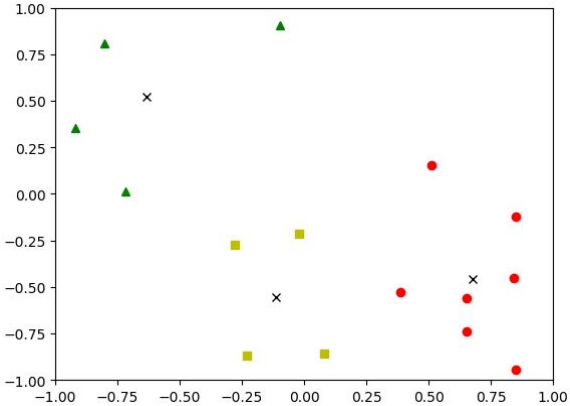
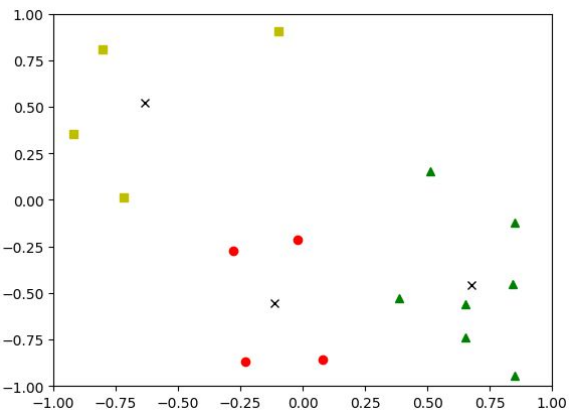
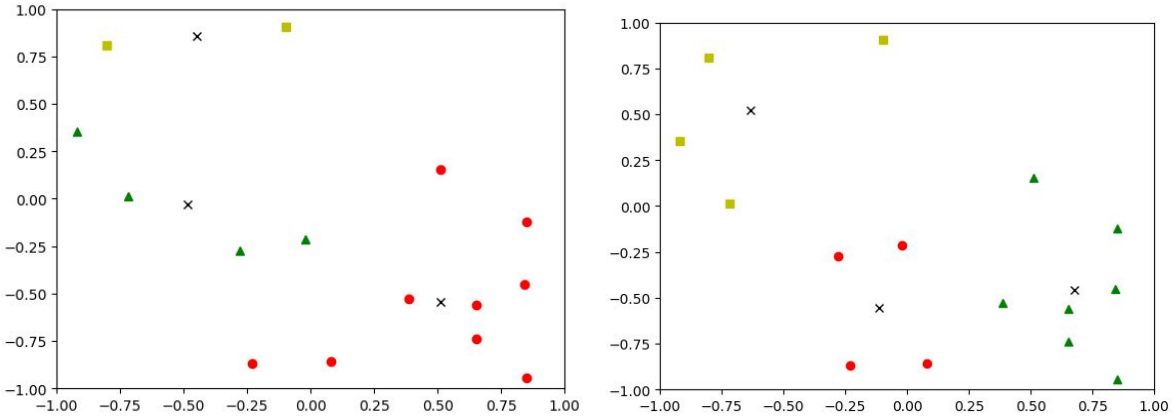
Question1.

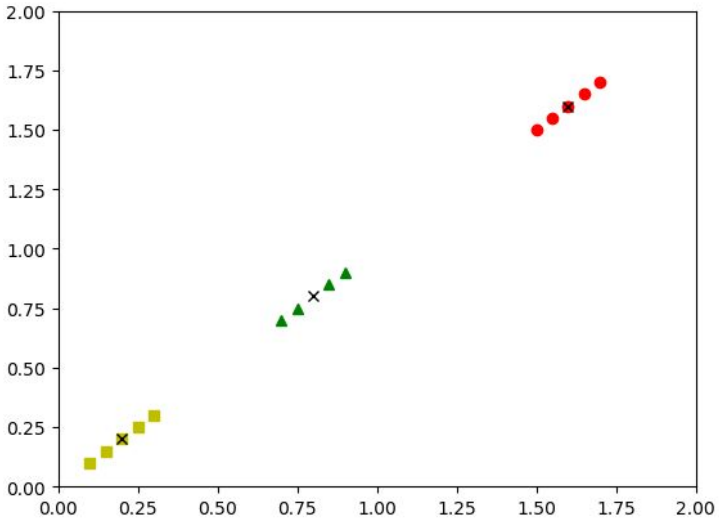


From the result, it can be seen running k-means algorithm several times, sometimes it will get different results and sometimes are the same, like the first run and the second, et. Because the initial centroids are chosen randomly, K-means will likely give different results each time it is run. Ideally these differences will be slight, but it is still important to run the algorithm several times and choose the result which yields the best clusters.

Question2.

I created a data set with very obvious three clusters.

```
data = np.array([[0.1, 0.1], [0.15, 0.15], [0.2, 0.2], [0.3, 0.3], [0.25, 0.25], [0.7, 0.7], [0.75, 0.75], [0.9, 0.9],[0.85, 0.85], [1.5, 1.5], [1.55, 1.55], [1.6, 1.6], [1.65, 1.65], [1.7, 1.7]])
```



For this data set with very obvious clustering characteristics, no matter how many times I run, the result is the same, because the k-means algorithm can always find the most reasonable classification scheme after iteration, even if the initialization centroid is random.

Question 3.

The initial centroids are chosen randomly, K-means will likely give different results each time it is run. It converges to local minimum, not the global optimum.

Many times, it is most appropriate to know in advance how many categories a given data set should be divided into. By automatically merging and splitting the class, a reasonable number of types K is obtained, such as the ISODATA algorithm.

If the cluster contains anomalous points, it will cause the mean deviation to be severe (ie, sensitive to noise and isolated point data). For this situation, changing to the median can be solved.

References.

1. Bezdek, J.C., 1980. A convergence theorem for the fuzzy ISODATA clustering algorithms. *IEEE transactions on pattern analysis and machine intelligence*, (1), pp.1-8.
2. Park, H.S. and Jun, C.H., 2009. A simple and fast algorithm for K-medoids clustering. *Expert systems with applications*, *36*(2), pp.3336-3341.