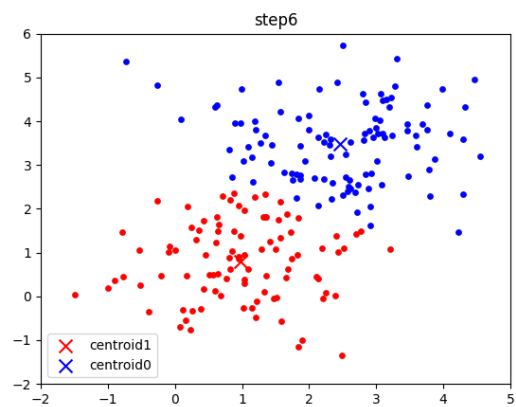
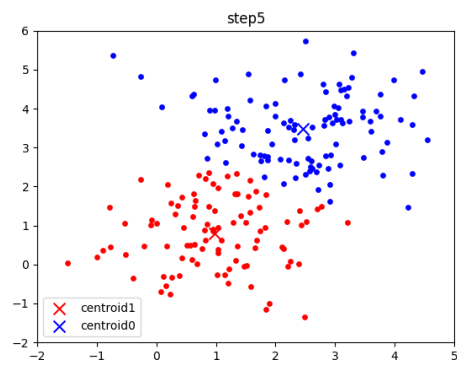
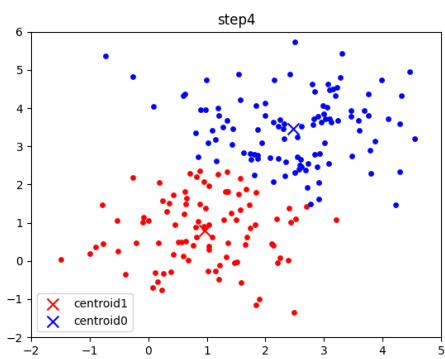
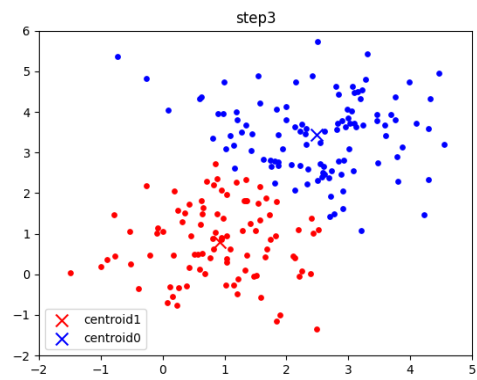
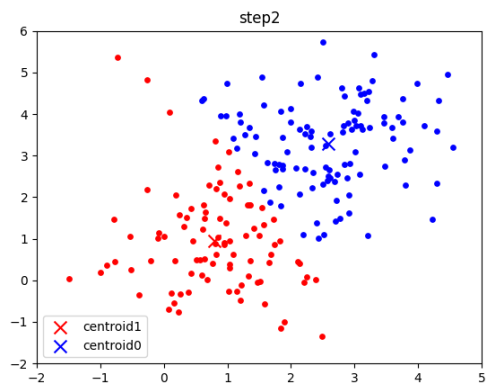
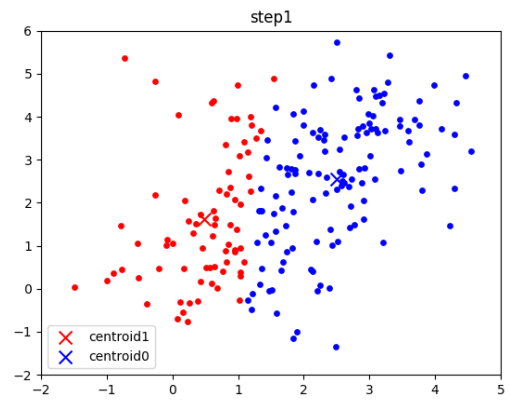
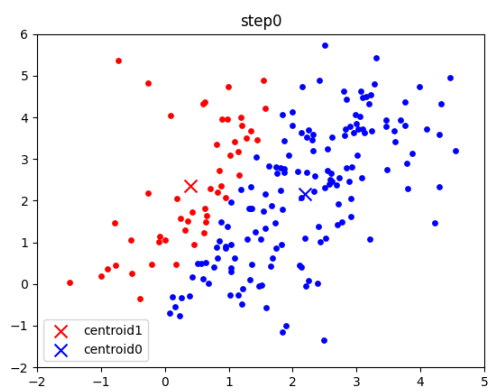


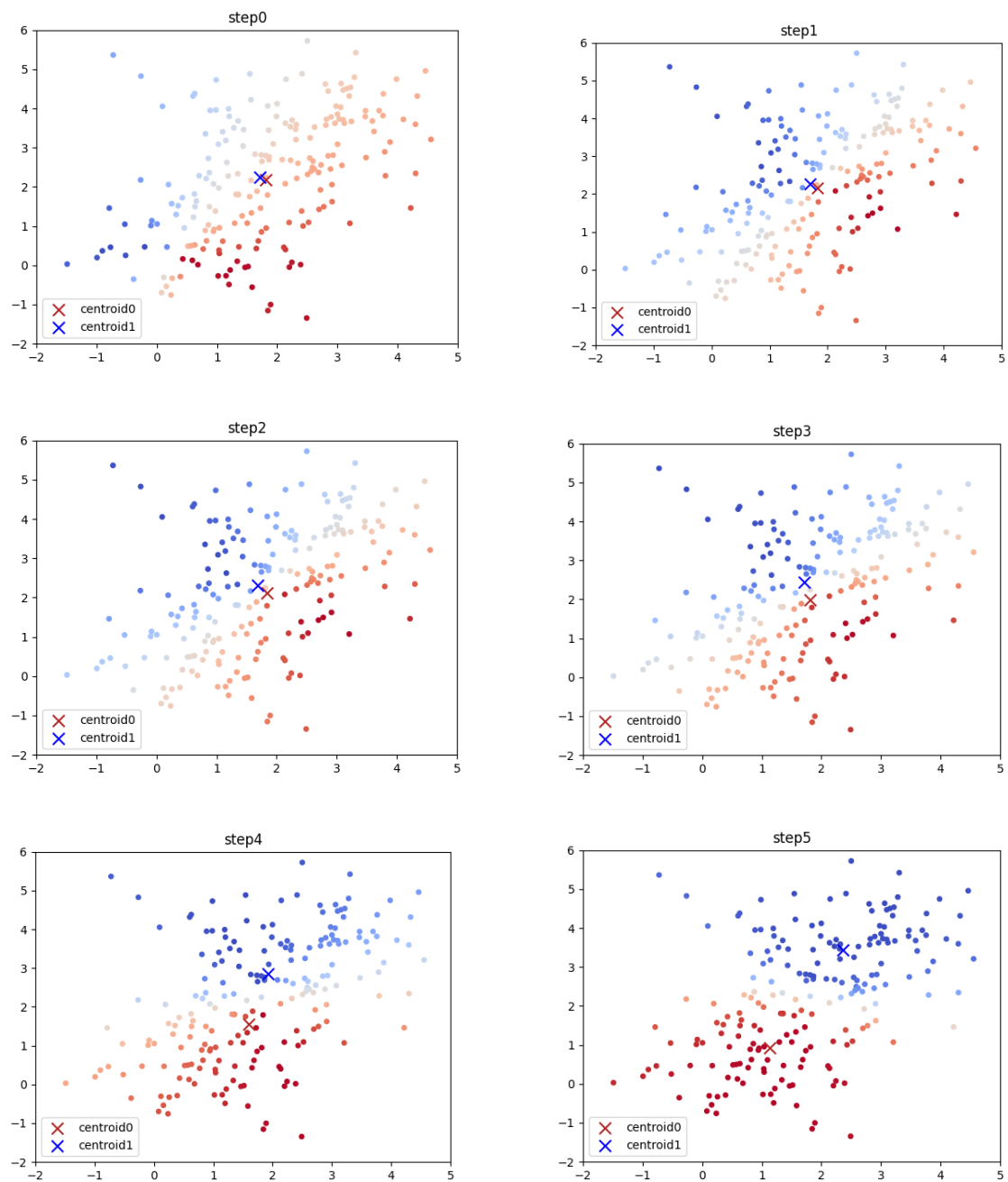
## Problem 1e

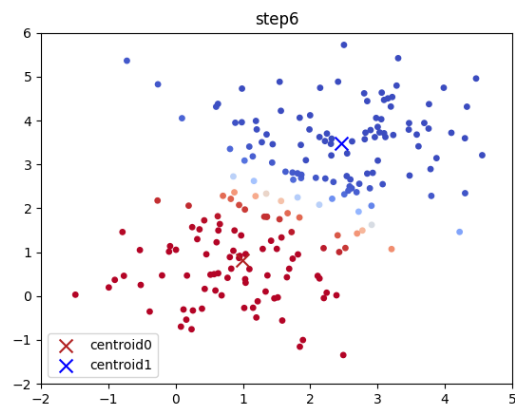


### Problem 1f

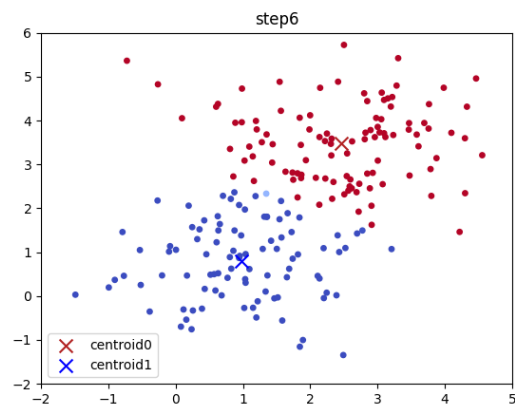
우리는 이번 과제에서 centroid를 각 그룹의 center of mass로 정의했다. 따라서 같은 숫자를 나타내는 여러 데이터들에서 흰 부분은 거의 비슷한 곳에 위치할 것이다. 그러므로 이 부분들의 center of mass를 얻는 과정을 반복하면, 결국 원래 숫자와 비슷한 모습이 될 것이라고 생각한다.

### Problem 2d





problem 1e와 다르게, centroid로부터 멀리 떨어진 data point들은 색깔이 연해진 것을 확인할 수 있다. centroid로부터 멀리 떨어져 있기 때문에 이 centroid가 대표하는 cluster에 포함되어 있을 확률이 낮다는 것을 의미한다.



centroid로부터 멀리 떨어진 data point이더라도 centroid 주변의 data point처럼 진한 색깔을 가지는 경우가 증가했다. 이것이 의미하는 바는 soft k-means clustering의 unassigned에 따른 효과가 줄어든 것이라고 생각한다. 따라서 나는 hyper-parameter beta가 커질수록 결과가 점점 hard k-means clustering에 가까워질 것이라고 생각한다.