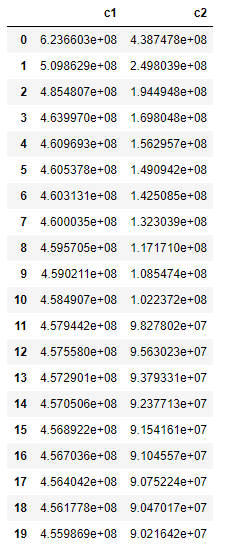
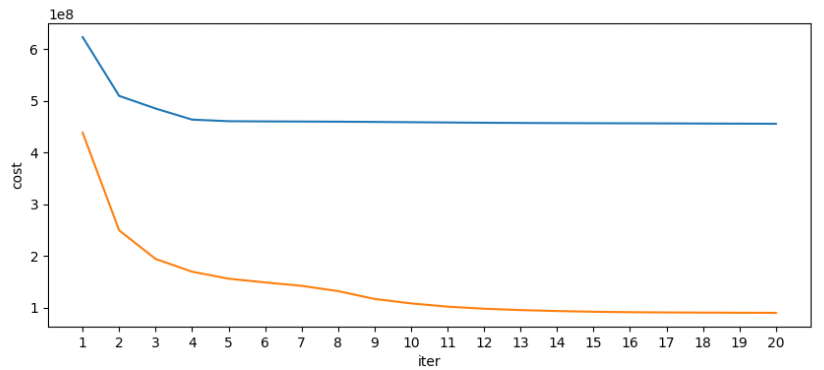
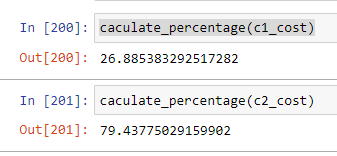
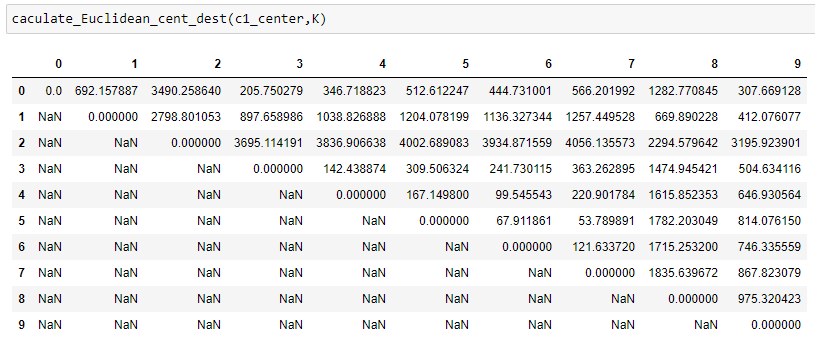
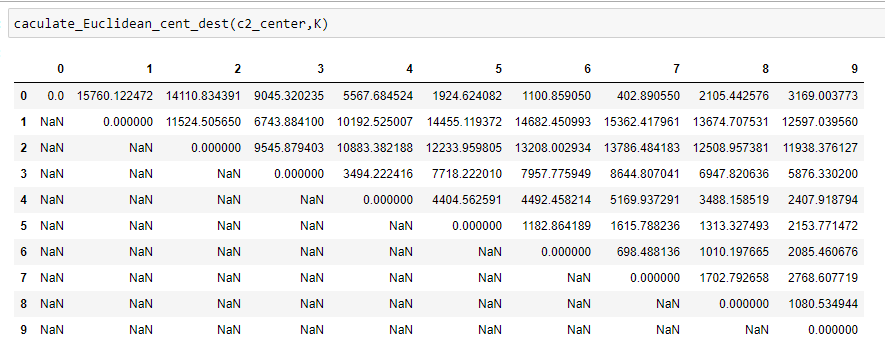
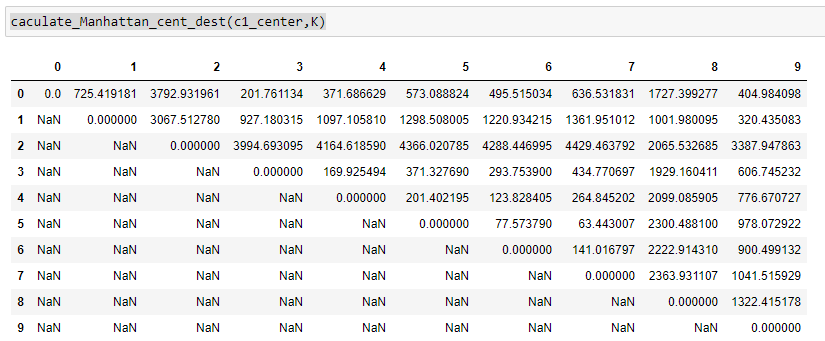
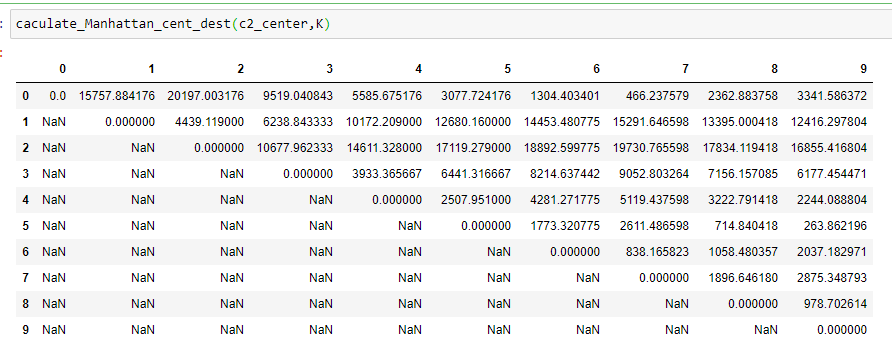
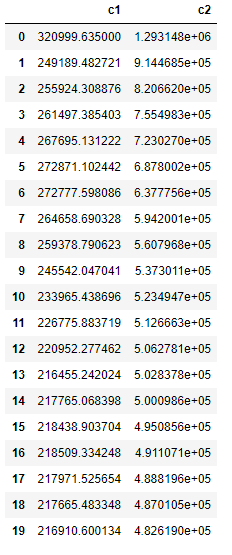
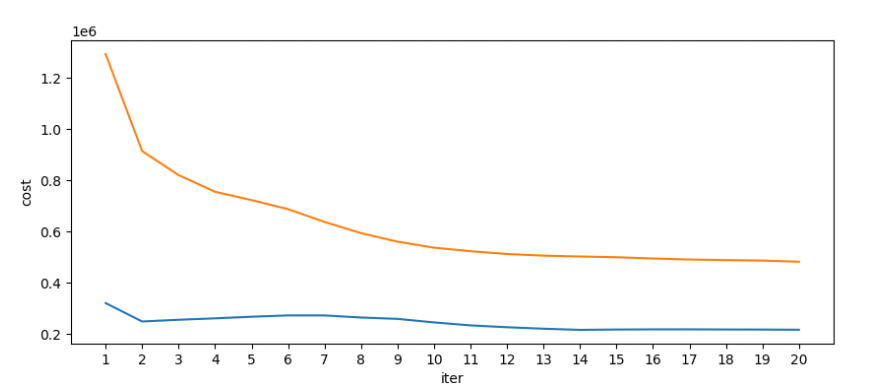
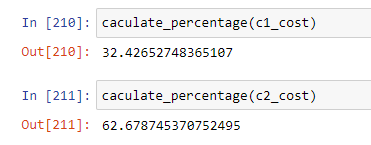
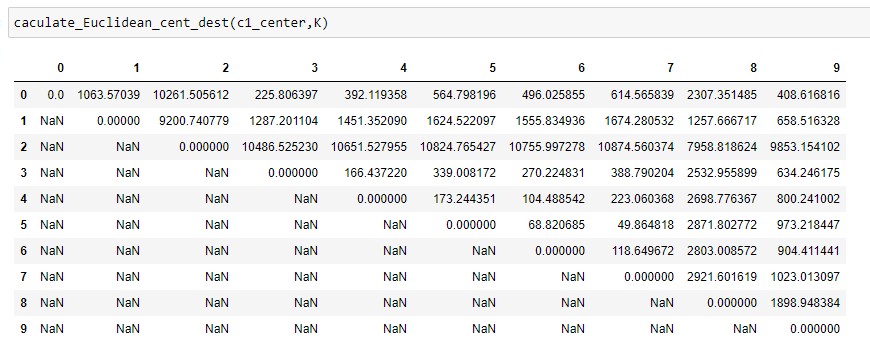
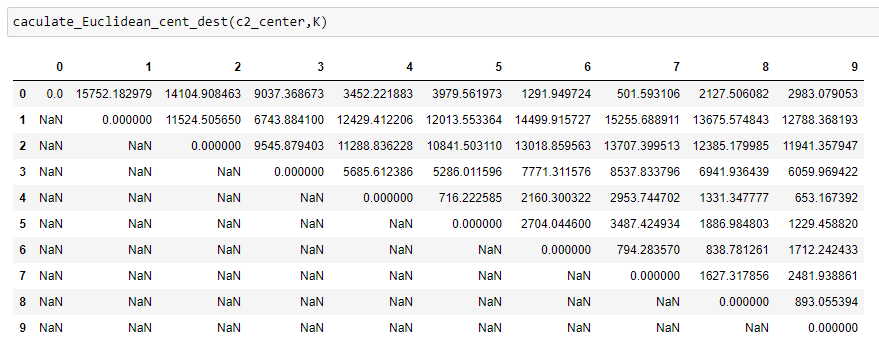
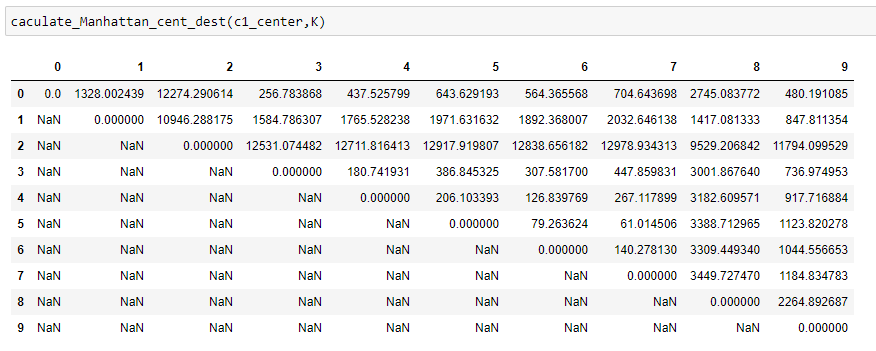
# question a

* 1. A plot of cost vs. iteration for 2 initialization strategies(c1 and c2) for (a):
     1. 
     2. 
  2. Percentage improvement values and your explanation for (a):
     1. C1 這個沒有特別挑過起始點的群集，最後改善率只有26.8%  
        而C2盡可能挑相鄰比較遠的點則有79%的改善率  
        這顯示了起始點對於 K-MEAN演算法有非常大的影響  
        若沒有優化過的話，這個演算法在歐基里德距離下可能會收斂在一個不是非常好的區間。



* 1. The Euclidean and Manhattan Distances for all pairs of centroids, with 2 initialization strategies. (共有四個表格)
     1. 
     2. 
     3. 
     4. 

# question b

* 1. A plot of cost vs. iteration for 2 initialization strategies(c1 and c2) for (a)
     1. 
     2. 
  2. Percentage improvement values and your explanation for (a)
     1. 在曼哈頓距離下，其結果與歐基里德距離類似，  
        C1 這個沒有特別挑過起始點的群集，最後改善率只有32.4%  
        而C2盡可能挑相鄰比較遠的點則有62.6%的改善率  
        這顯示了起始點對於 K-MEAN演算法有非常大的影響  
        若沒有優化過的話，這個演算法在歐基里德距離下可能會收斂在一個不是非常好的區間。
     2. 
  3. The Euclidean and Manhattan Distances for all pairs of centroids, with 2 initialization strategies. (共有四個表格)
     1. 
     2. 
     3. 
     4. 