For HW2 question 3, there is one subquestion about judging which confusion matrix corresponds to a better cluster result. I wonder how to judge if the two confusion matrices both have a low diagonal sum.Confusion matrix 1 =   
 10 4 8 8 6 7 2 4 3 4 1 2 4 1 5 7 2 8 9 7 5 3 7 1 7 12 0 0 16 14 10 3 9 0 7 4 3 7 5 9 2 1 1 4 4 1 0 3 12 0 6 3 1 7 22 17 3 5 7 6 4 3 1 4 4 1 4 1 1 6 3 3 3 1 4 3 1 2 7 5 0 2 7 1 4 1 13 6 1 2 0 0 0 0 1 2 0 2 0 0Confusion matrix 2 =  
 23 0 0 13 16 0 0 0 3 1 16 0 0 9 20 1 0 0 0 0 3 1 0 32 23 2 0 0 4 0 1 0 0 39 11 2 1 0 3 0 1 0 0 12 15 0 0 0 0 0 4 0 1 31 33 0 0 1 5 2 12 0 0 7 9 0 0 0 1 0

9 0 0 11 12 0 0 0 0 0

15 0 0 9 13 0 0 0 0 0

0 0 0 3 2 0 0 0 0 0