CSCI 6350 Spring 2006
Homework 9 (Due: beginning of class, Tuesday March 28th)

1. Calculate the dissimilarity between pair-wise objects using the dissimilarity measure given on page 345-350/[Han and Kamber]. List the dissimilarities in 6x6 dissimilarity table. Fever and Cough are asymmetric binary attributes, Category is ordinal attributes with the order of the three values being: stable unstable and severe.

	Gender	**Age**	**Heart-rate**	**Fever**	**Cough**	**Category**
Obj1:	M	18	120	Y	N	stable
Obj2:	F	36	89	N	Y	unstable
Obj3:	M	20	115	Y	Y	stable
Obj4:	M	3	94	Y	N	severe
Obj5:	F	28	110	N	Y	severe
Obj6:	F	44	80	N	Y	unstable

2. Applying PAM clustering approach on the above data to partition data into K=2 clusters. Assuming Obj2 and Obj4 are selected as the initial Medoids of the two clusters. Show:
    a. In which cluster would each of the data objects be assigned on the first iteration of object distribution?
    b. Should Obj 3 be used to replace Obj4 as the Medoid of one of the clusters for the next iteration of clustering? Answer the question with results obtained from computation.

3. Perform hierarchical clustering on the six objects in question 1, using any two of the following agglomerative clustering methods:
    (a) the single-link,
    (b) the complete-link, and
    (c) the average-link

Show the clustering hierarchy constructed.

4. Apply K-means clustering algorithm to group the following 6 objects into K=3 clusters. (Show all steps in all iterations before the clusters converge). Assuming obj1, obj3, and obj5 were selected initially as the cluster centroids for the three clusters.

	length	**width**	**height**
obj1:	3	18	5
obj2:	5	12	10
obj3:	1	16	7
obj4:	4	12	9
obj5:	8	5	8
obj6:	9	3	6