CS181 ASSIGNMENT 3
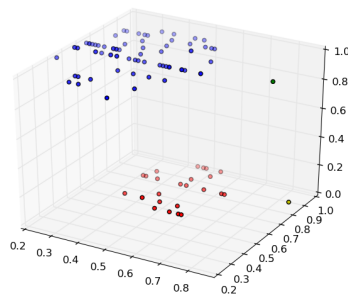
ASHOK CUTKOSKY AND TONY FENG

Problem 3.
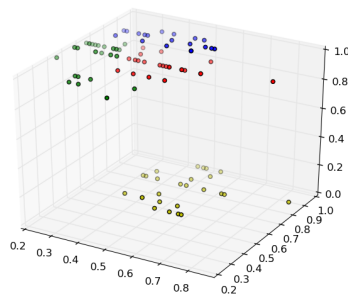
a. (a) Number of instances in each cluster by metric:

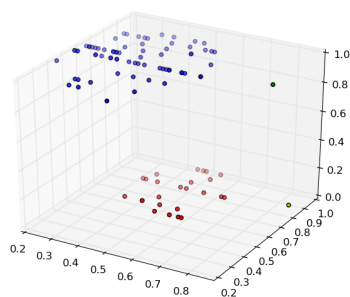| Cluster Number	Min metric	Max metric
0	73	19
1	24	23
2	1	32
3	1	26

Scatterplot of Min metric clusters:

Scatterplot of Max metric clusters:

(b) Number of instances in each cluster by metric:

Cluster Number	mean metric	centroid metric
0	73	9
1	25	89
2	1	1
3	1	1

Scatterplot of mean metric clusters:

Scatterplot of centroid metric clusters:
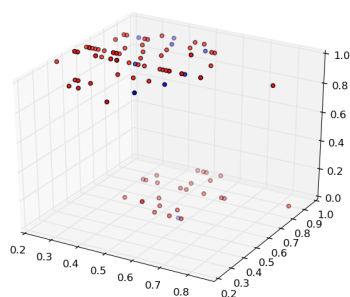
Problem 4.

a. (a)

k	Mean squared distance
1	1.97
2	1.80
3	1.81
4	1.77
5	1.52
6	1.37
7	1.42
8	1.25
9	1.31
10	1.26

(a) (b) It is difficult to tell, because the mean squared values always seem to be
improving as the number of clusters increases (which should be the case). How-
ever, the improvements seem to level off somewhat at 8 clusters, so that is about
as reasonable a guess as we can make just from the above data.