## Algorithmic Logic and Mathematical correlation:

## Naive Based K means Implementation:

```
import random
import math

#Euclidian Distance between two d-dimensional points
def eucldist(p0,p1):
    dist = 0.0
    for i in range(0,len(p0)):
        dist += (p0[i] - p1[i])**2
    return math.sqrt(dist)

#K-Means Algorithm
def kmeans(k,datapoints):
    # d - Dimensionality of Datapoints
```

```
d = len(datapoints[0])
  #Limit our iterations
  Max Iterations = 1000
  i = 0
  cluster = [0] * len(datapoints)
  prev_cluster = [-1] * len(datapoints)
  #Randomly Choose Centers for the Clusters
  cluster_centers = []
  for i in range(0,k):
     new cluster = []
     #for i in range(0,d):
     # new cluster += [random.randint(0,10)]
     cluster_centers += [random.choice(datapoints)]
     #Sometimes The Random points are chosen poorly and so there ends up being empty
clusters
     #In this particular implementation we want to force K exact clusters.
     #To take this feature off, simply take away "force_recalculation" from the while conditional.
     force recalculation = False
  while (cluster != prev_cluster) or (i > Max_Iterations) or (force_recalculation):
     prev_cluster = list(cluster)
     force_recalculation = False
     i += 1
     #Update Point's Cluster Alligiance
     for p in range(0,len(datapoints)):
       min_dist = float("inf")
       #Check min_distance against all centers
       for c in range(0,len(cluster_centers)):
          dist = eucldist(datapoints[p],cluster centers[c])
          if (dist < min dist):
            min dist = dist
            cluster[p] = c # Reassign Point to new Cluster
```

```
#Update Cluster's Position
    for k in range(0,len(cluster_centers)):
       new center = [0] * d
       members = 0
       for p in range(0,len(datapoints)):
          if (cluster[p] == k): #If this point belongs to the cluster
            for j in range(0,d):
               new_center[j] += datapoints[p][j]
            members += 1
       for j in range(0,d):
          if members != 0:
            new center[j] = new center[j] / float(members)
          #This means that our initial random assignment was poorly chosen
          #Change it to a new datapoint to actually force k clusters
          else:
            new center = random.choice(datapoints)
            force_recalculation = True
            print "Forced Recalculation..."
       cluster centers[k] = new center
  print "====== Results ======"
  print "Clusters", cluster_centers
  print "Iterations",i
  print "Assignments", cluster
#TESTING THE PROGRAM#
if __name__ == "__main__":
  #2D - Datapoints List of n d-dimensional vectors. (For this example I already set up 2D
Tuples)
  #Feel free to change to whatever size tuples you want...
  datapoints = [(3,2),(2,2),(1,2),(0,1),(1,0),(1,1),(5,6),(7,7),(9,10),(11,13),(12,12),(12,13),(13,13)]
  k = 2 # K - Number of Clusters
  kmeans(k,datapoints)
```