

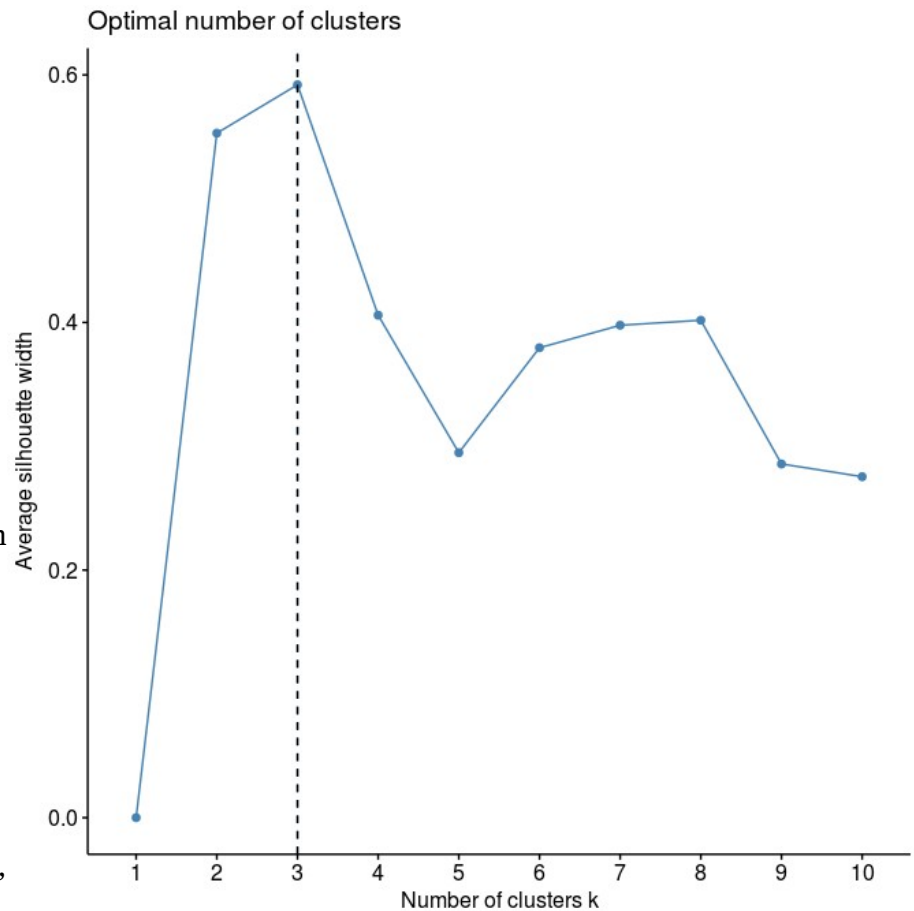
Homework 3

Q.2)

From the adjacent graph, we can observe that optimal number of clusters for the given dataset is 3. Though there are 5 clusters in the datasets.

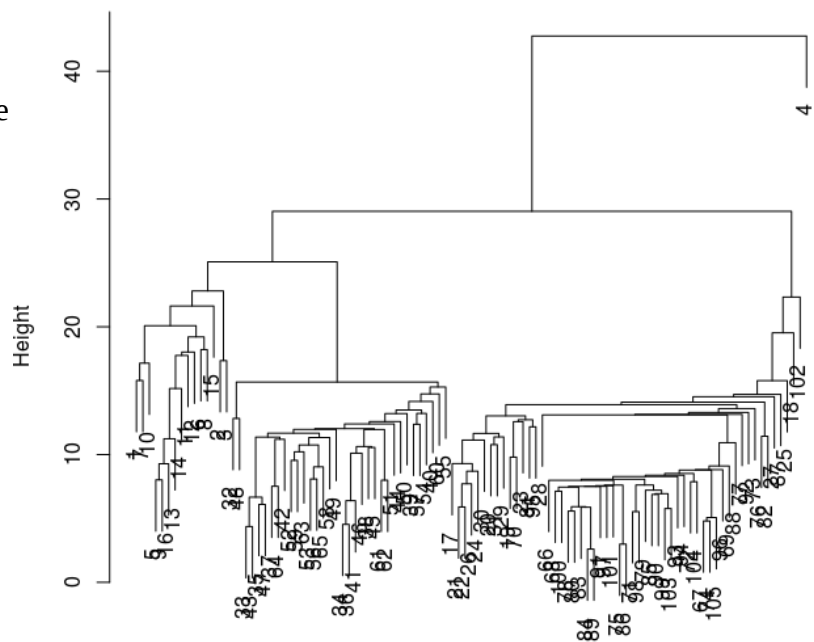
Agglomerative coefficient shows accuracy of the classification. From the results shown in the graph, we can observe that for the single linkage method of hierarchical clustering, error rate is maximum (accuracy is min). For complete linkage, misclassification rate is minimum.

For kmeans clustering of 3 clusters, misclassification rate is 0.45 whereas, for 5 groups, error rate is 0.55.



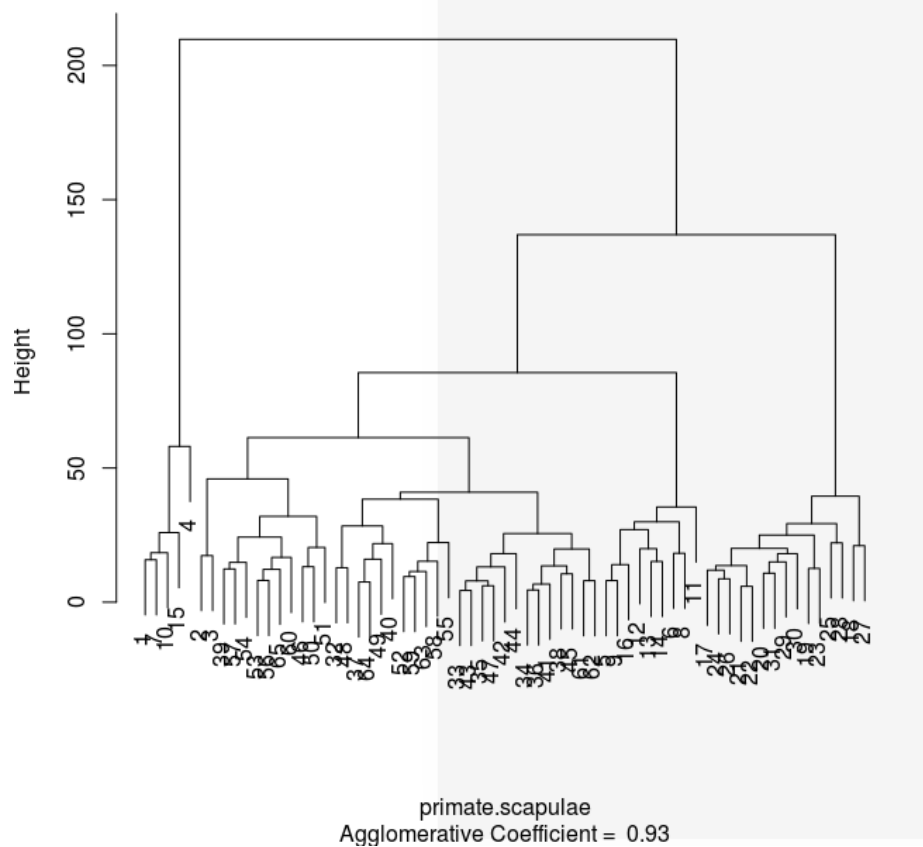
Dendrogram of `agnes(x = primate.scapulae, method = "single")`

So we can conclude that, hierarchical clustering with complete linkage is the best method for classification.

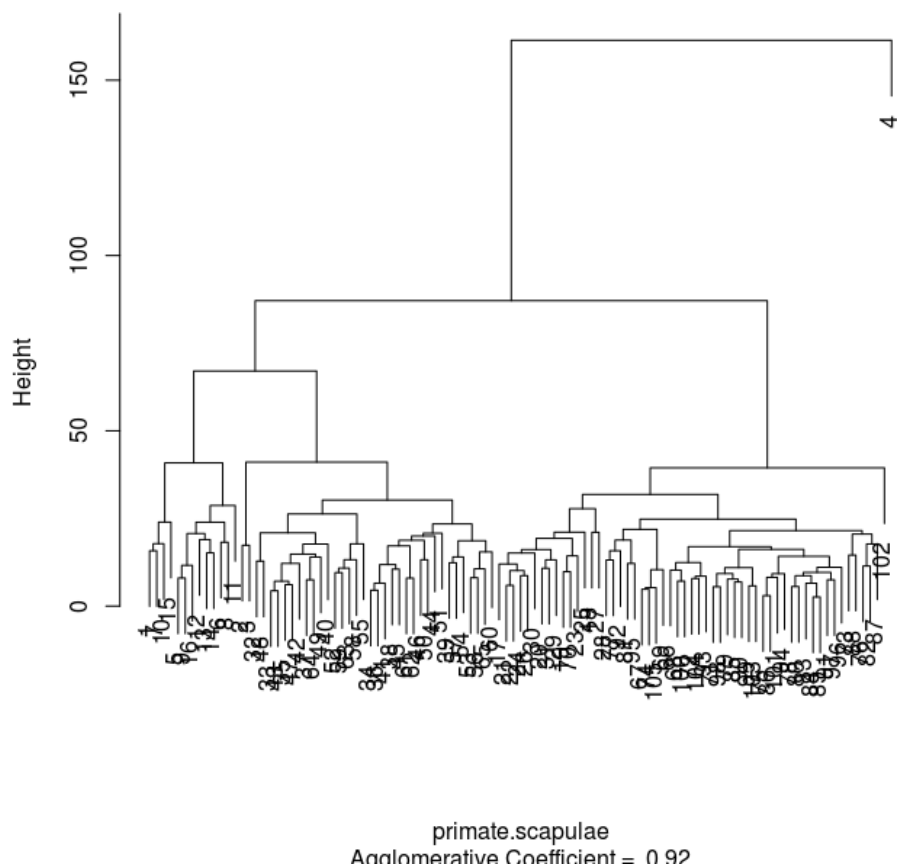


primate.scapulae
Agglomerative Coefficient = 0.76

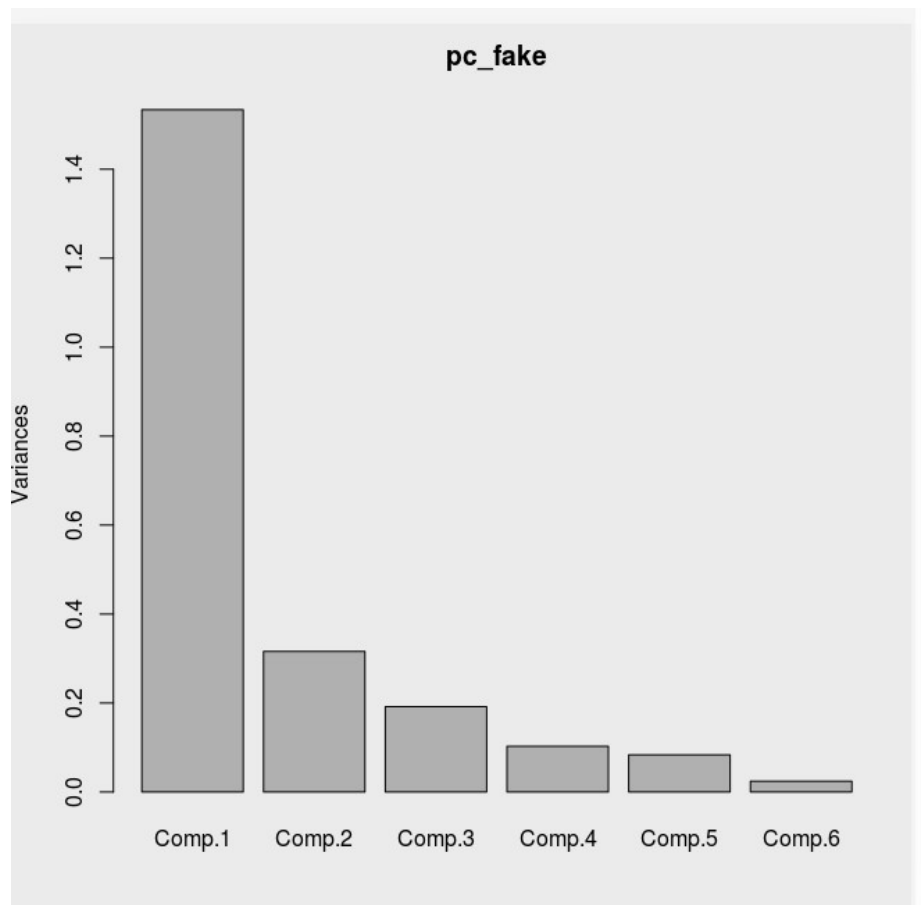
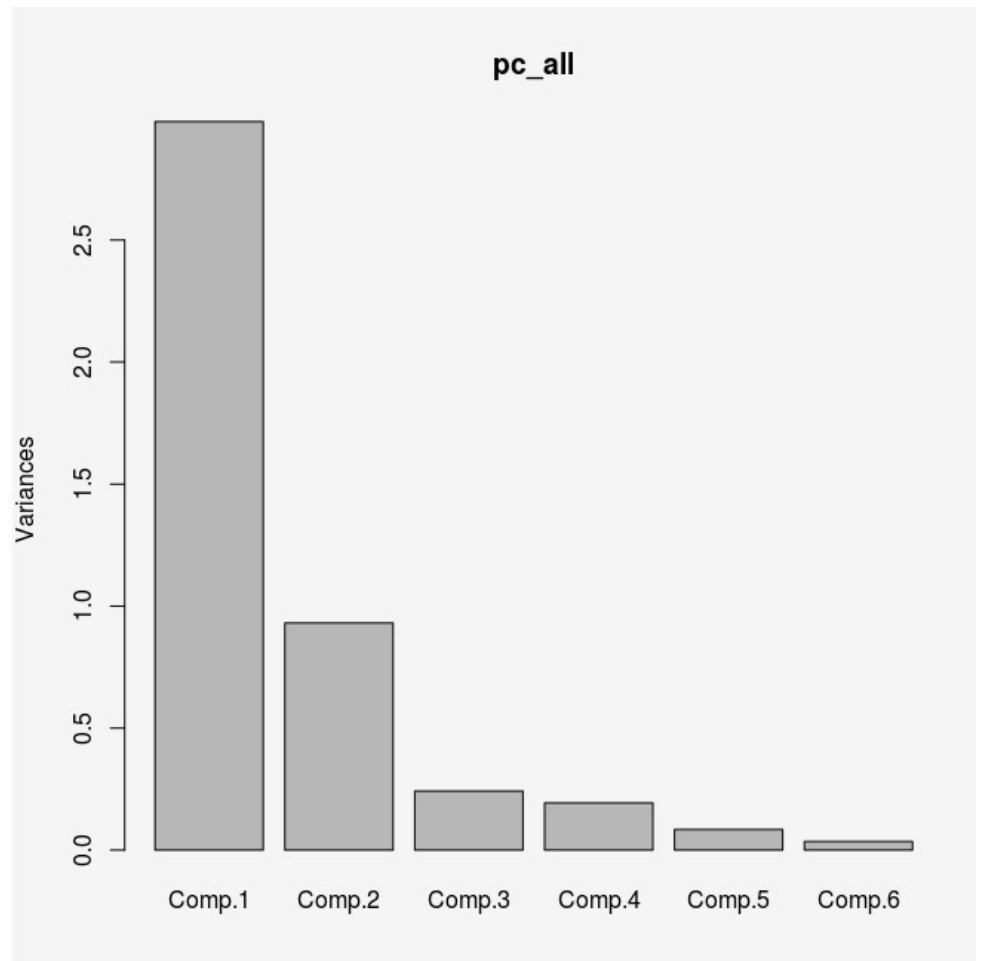
Dendrogram of agnes(x = primate.scapulae, method = "complete")

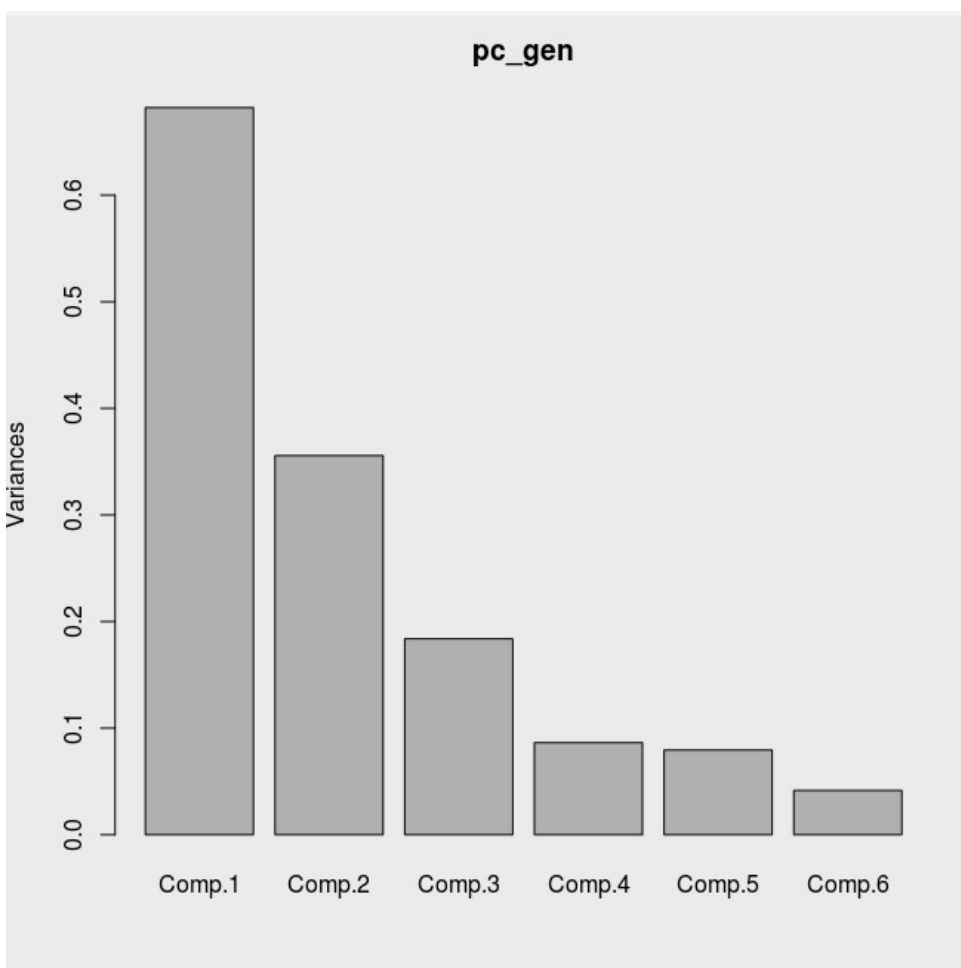


Dendrogram of agnes(x = primate.scapulae, method = "average")

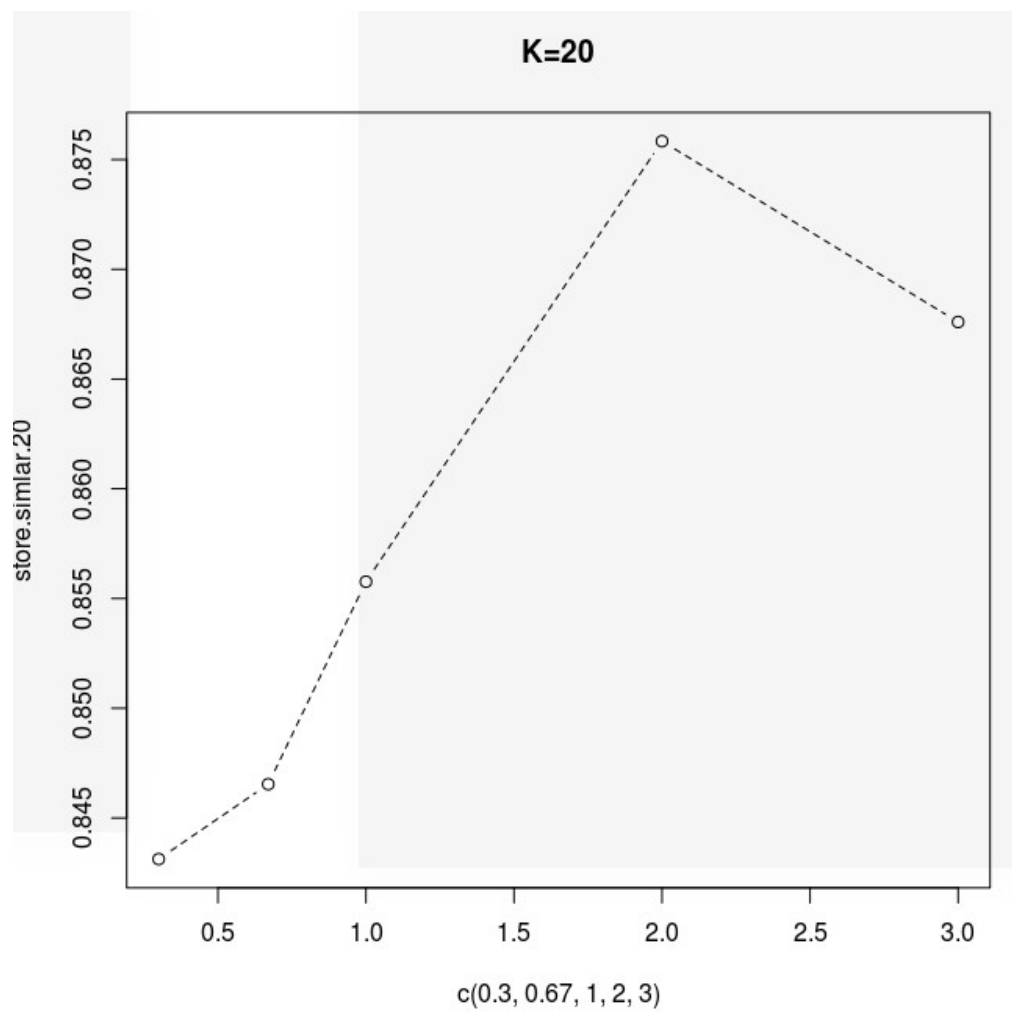
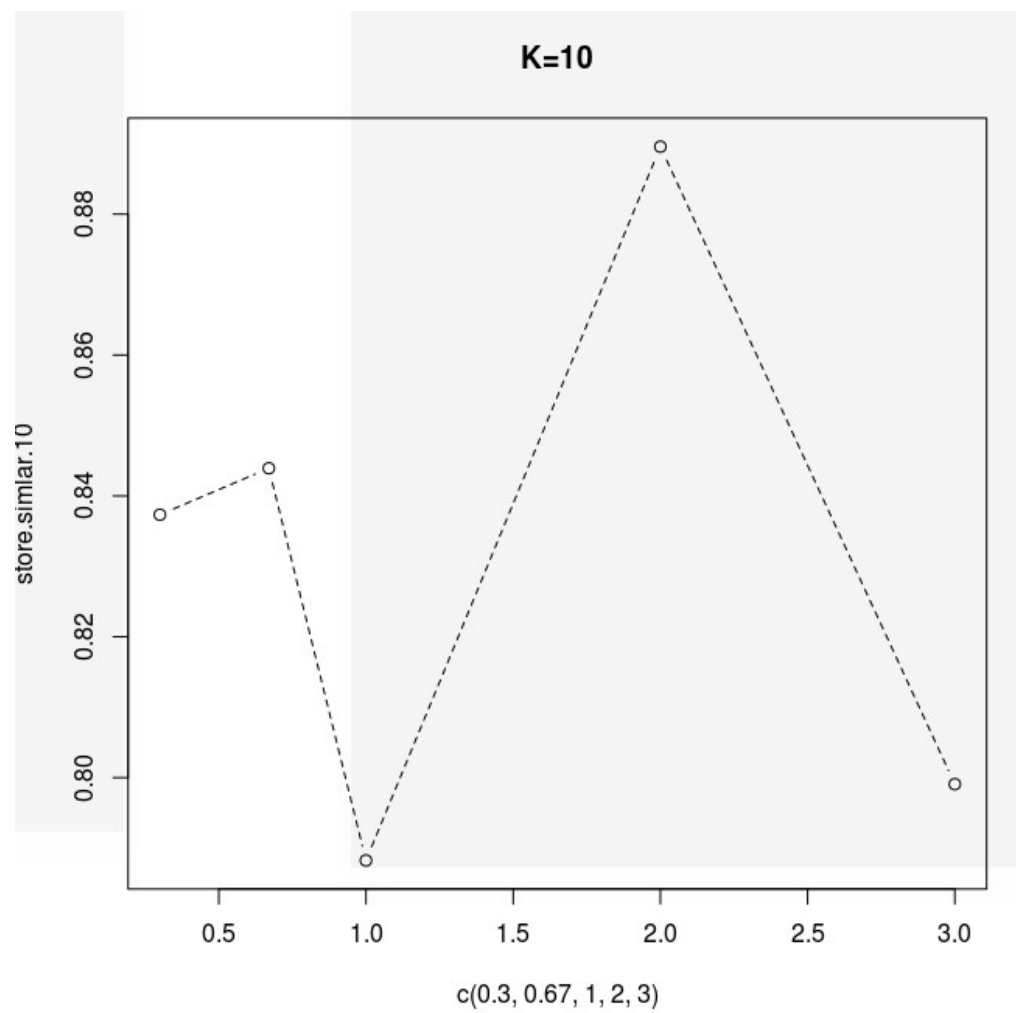


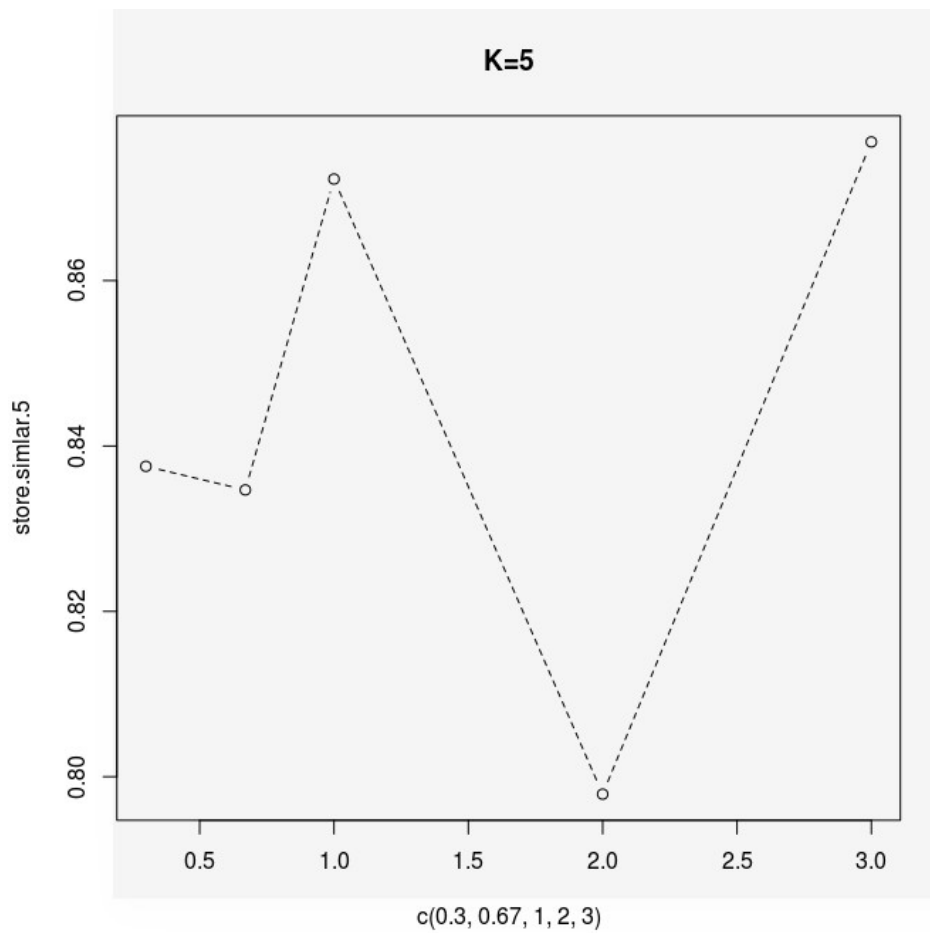
Q.1)





Q.3)





As we can see from the graphs and code, as we decrease the size of SOM neighborhood, the solution tends towards the k-means solution.

