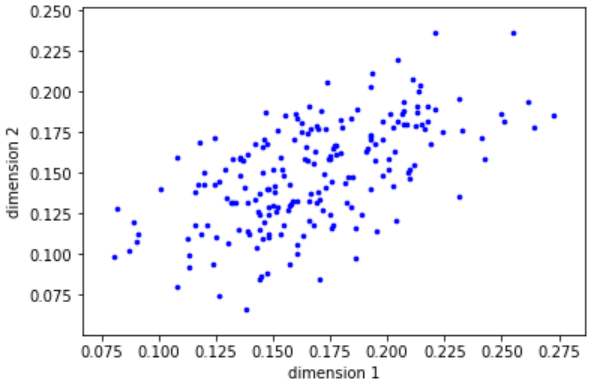
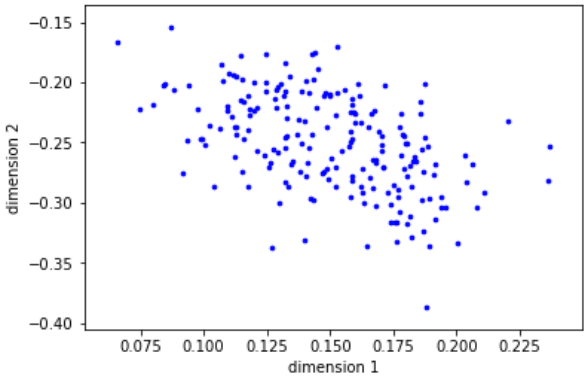
Results and Analysis for Q7:

PartA)

(i)

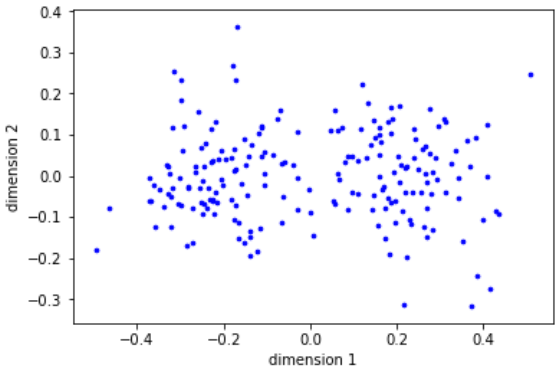


(ii)



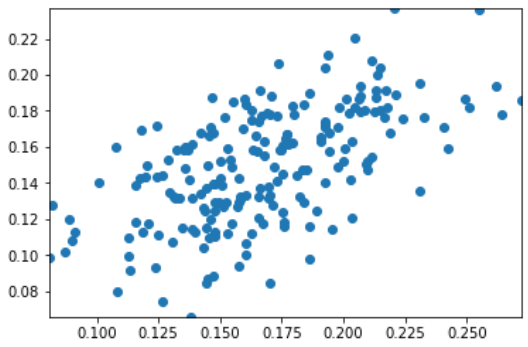
No specific structure or pattern is visible in parts (i) and (ii) because only a part of the input data (i.e. 2 out of 40 features) have been plotted.

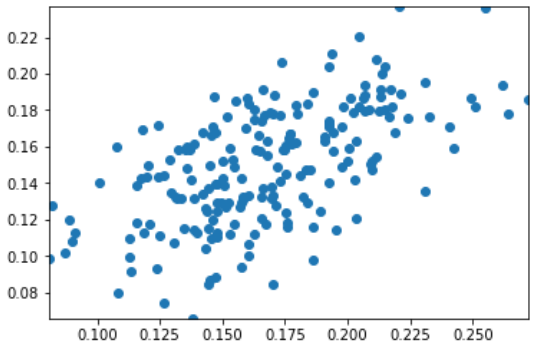
(iii)



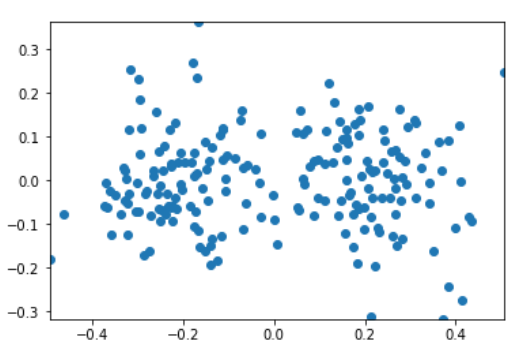
The data is present in 2 clusters approximately. Though it contains most of the original input by selecting the principal components.

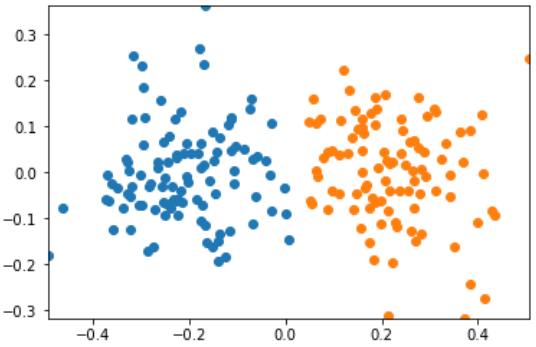
iv) For k=2: Original Graph and clustered Data:





For k=2: Original Graph and clustered Data:



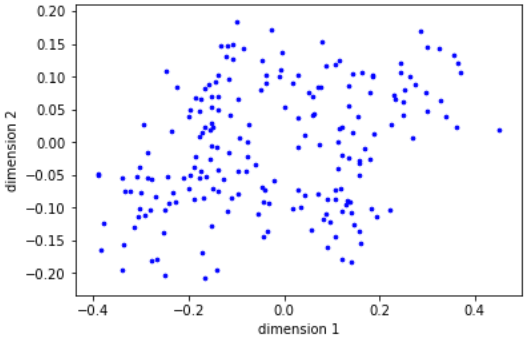


K-means without PCA on the original input returns the data as it is, whereas when applied on the data reduced by PCA does successful clustering. Both get the general structure of the original data

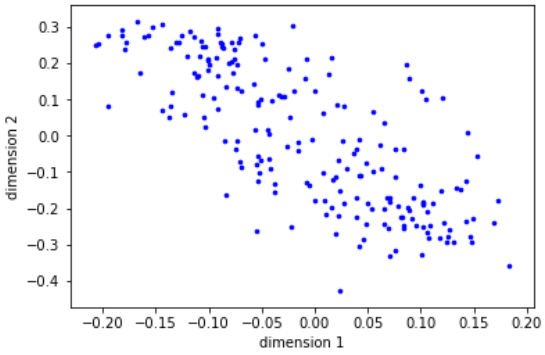
v) In general, PCA reduces the dimension of the data and provides us with principal clusters or reveal the clusters

Part B)

(i)

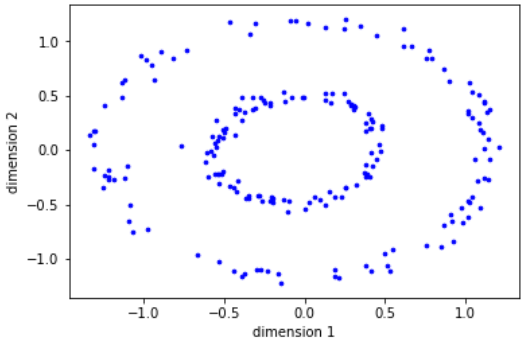


(ii)



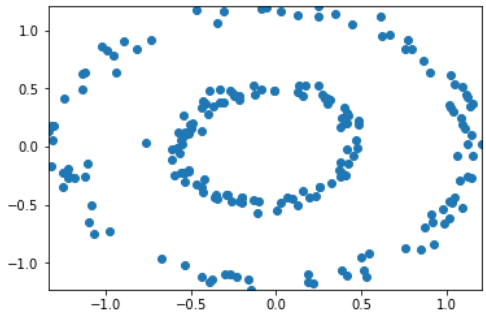
No specific structure or pattern is visible in parts (i) and (ii) because only a part of the input data (i.e. 2 out of 40 features) have been plotted.

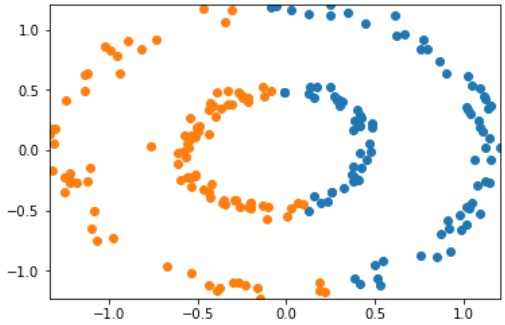
(iii)



The data is present in 2 clusters approximately. Though it contains most of the original input by selecting the principal components.

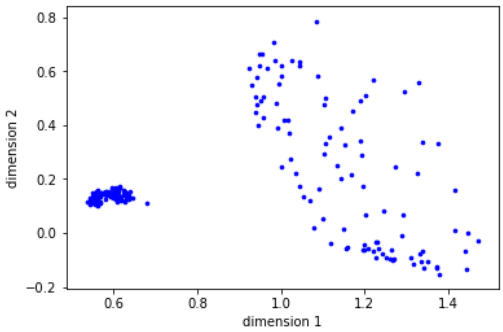
iv) For k=2: Original Graph and clustered Data:





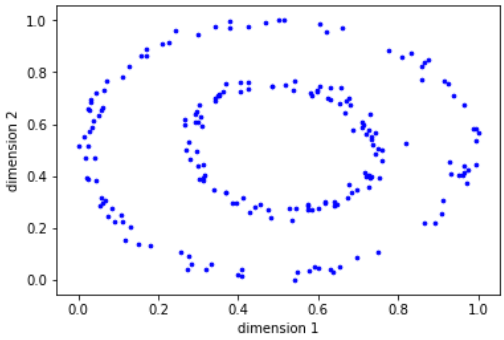
K-means fails after PCA because the two clusters are non-linearand not linearly separable.

v) KPCA and K-means on the reduced data:



K-means is successful after KPCA because the two clusters are now separable.

vi) Spectral Clustering and K-means on the reduced data:



K-means is successful after Spectral clustering because the two clusters are now separable.