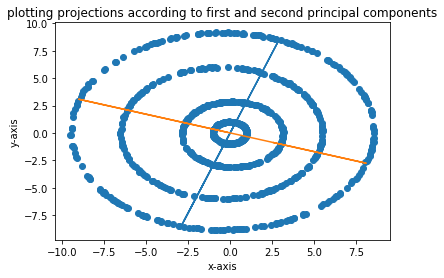
**Q1:**

**i.** code to run PCA on dataset:



Variance explianed by first principal component: 54.17802452885223

Variance explianed by second principal component: 45.82197547114777



**ii.** Observation: running PCA without centring does not effect much.  
**mean along x-axis**: 4.074999999685858e-07   
**mean along y-axis**: 2.2270000000723655e-07

So the mean is very close to zero. Therefore subtracting it form every data point does not have much effect.

Also, variance explained by first eigen vector: 54.17802452885223 and Variance explained by second eigen vector: 45.82197547114777

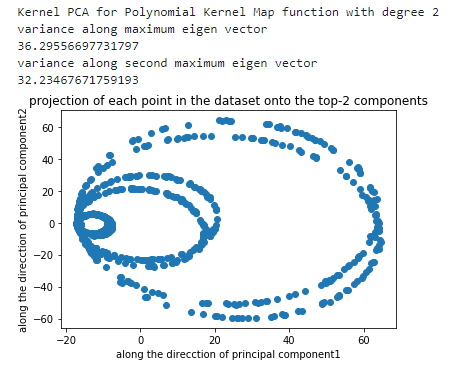
Which turns out to be exact same as when done without centering.

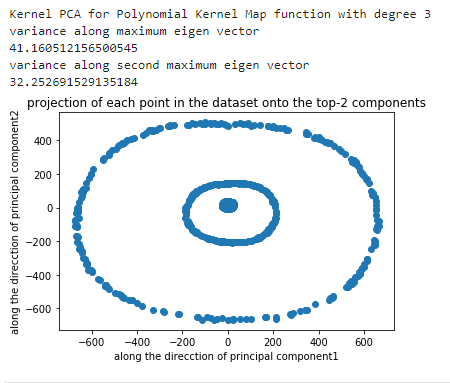
Therefore, the help obtained by centering is not much significant.

Code Reference: **Question1\_(i,ii).ipynb / Question1\_(i,ii).py**

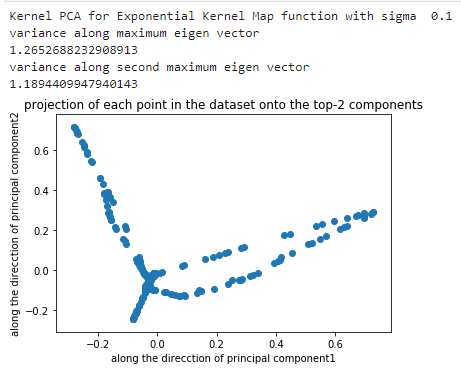
**iii. Kernel - PCA**

**Part A**: for d = 2

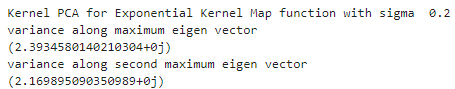


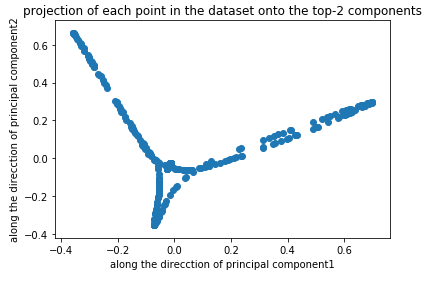
**Part A**: for d = 3  
 

**Part B:** sigma = 0.1

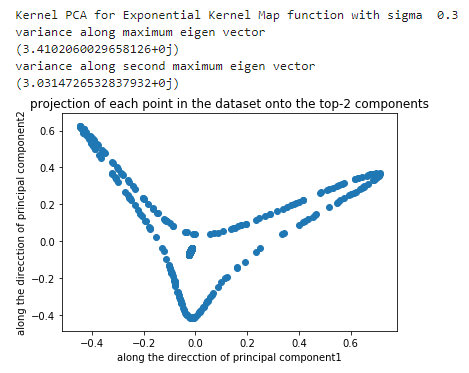


Sigma = 0.2

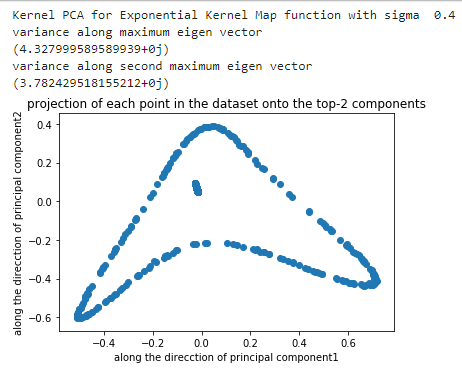




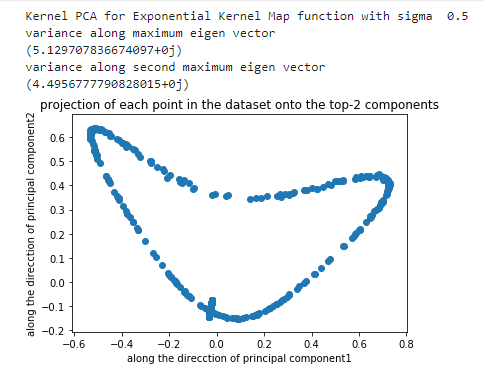
Sigma = 0.3



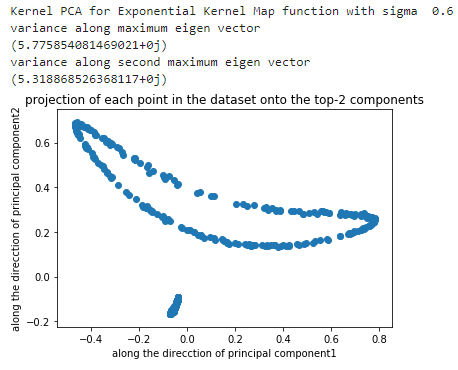
Sigma = 0.4



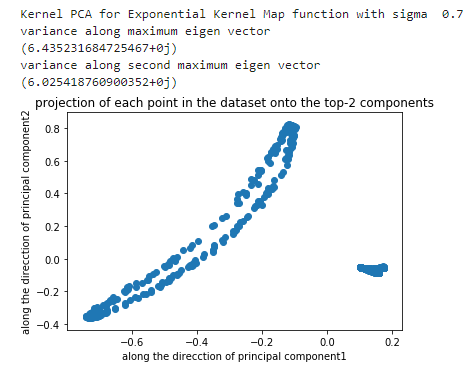
Sigma = 0.5



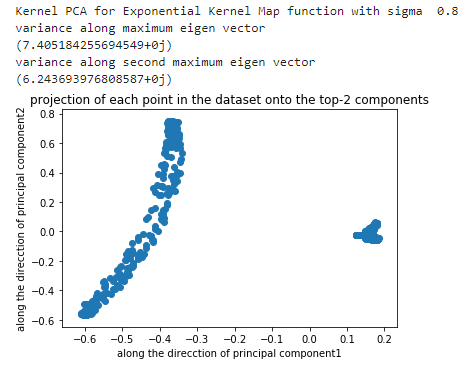
Sigma = 0.6



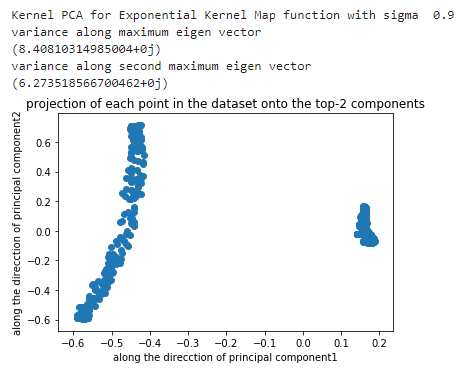
Sigma = 0.7



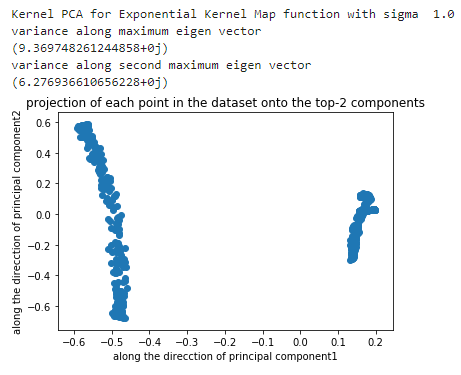
Sigma = 0.8



Sigma = 0.9



Sigma = 1.0



Code Reference:  

iv. The kernel we got by using Polynomial Mapping function of the following form 

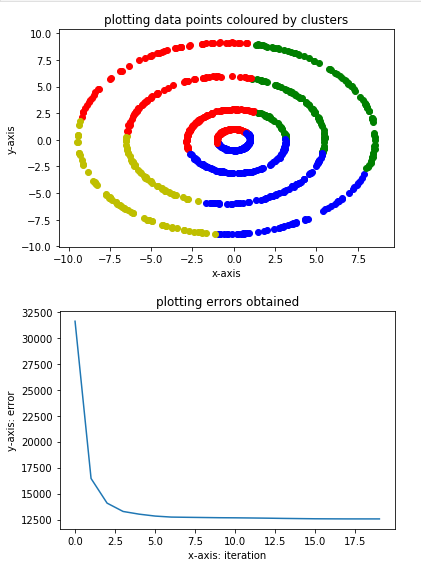
using d = 3 is best suited for the dataset.

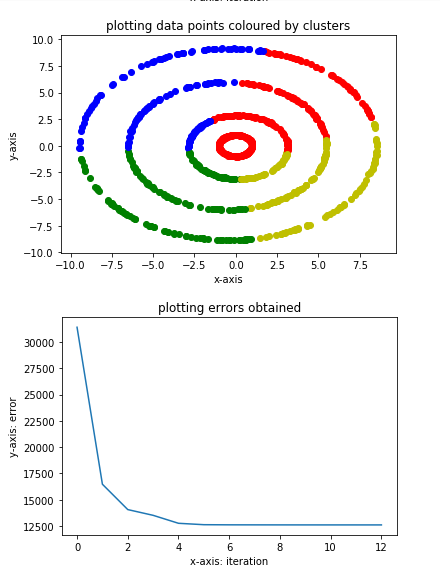
Since the variance explained by two principal components are as follows: 41.16%, 32.25%. Which is comparatively better than all the other approaches followed to obtain Kernel map.

**Q2:**

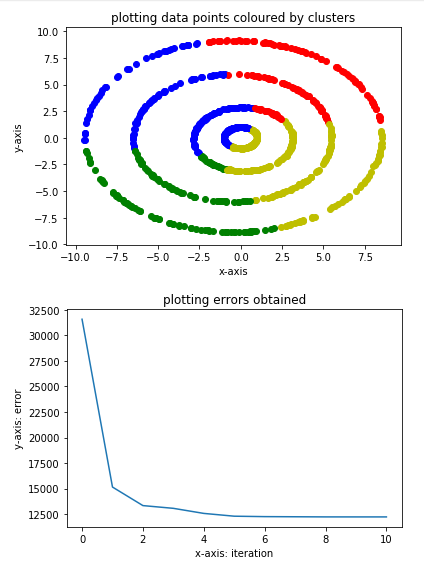
**Code Reference: ** ****

1. Results obtained after trying 5 different random initialization:

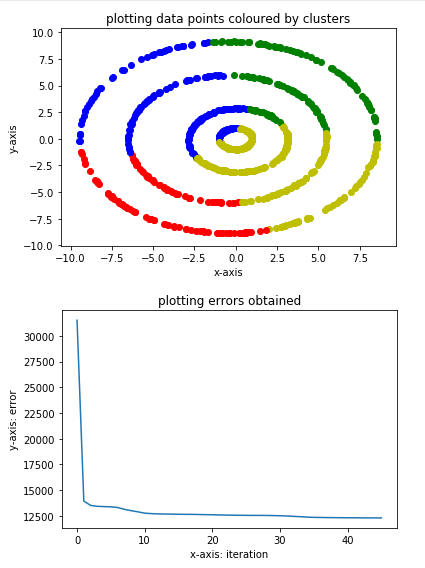
Try: 1  


Try: 2  


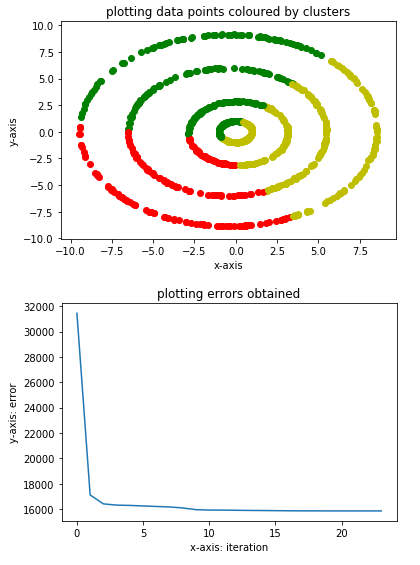
Try: 3



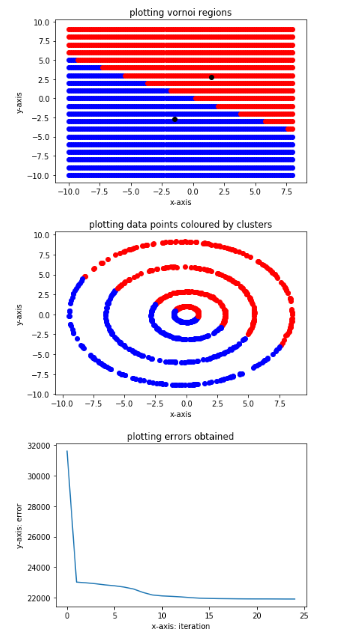
Try: 4



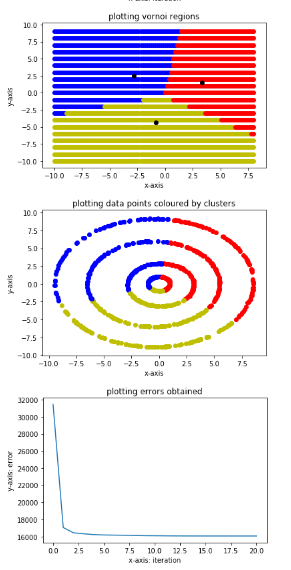
Try: 5



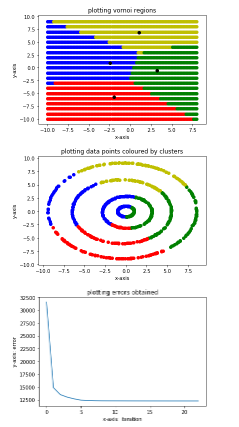
1. For K = {2,3,4,5}

Obtained cluster centres according to K-means algorithm using fixed initialization. Results found are presented as follows:  
for K = 2  


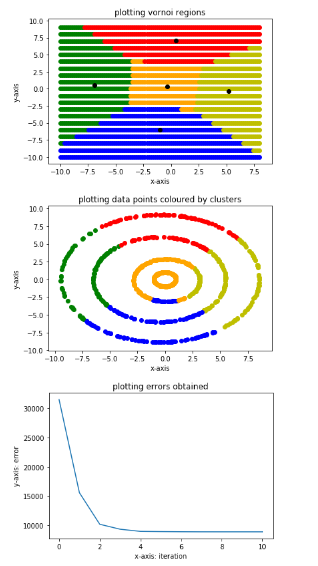
For K = 3

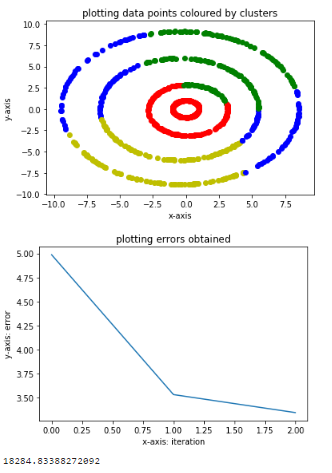


For K = 4



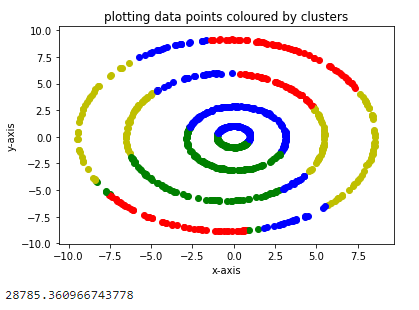
For K = 5



1. Spectral Clustering  
   

Choice of Kernel function:  
Polynomial Kernel Mapping with degree: 2

Reasoning: This Kernel was explaining maximum variance over its top eigen-vectors amogst others for this dataset. Also the error function goes to a better minimum as compared to other kernel functions. Computed the sum of distances from the mean corresponding to each points and this result is also appears to be minimum for this choice of kernel.

1. Result obtained by following the suggested method of cluster assignment:  
   

Observation: this strategy of cluster assignment does not work as well as other strategies, as seen in the output plot. Also the sum of distances of the data points with their corresponding cluster mean is coming out to be much larger compared to other approaches.