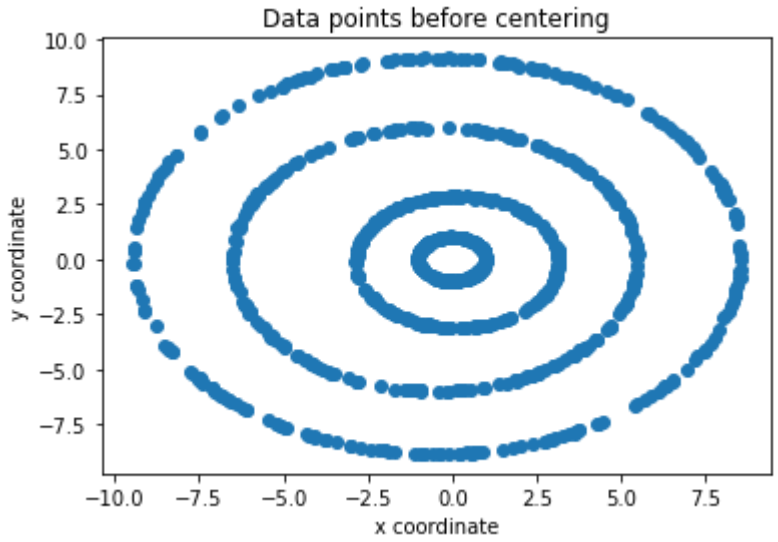
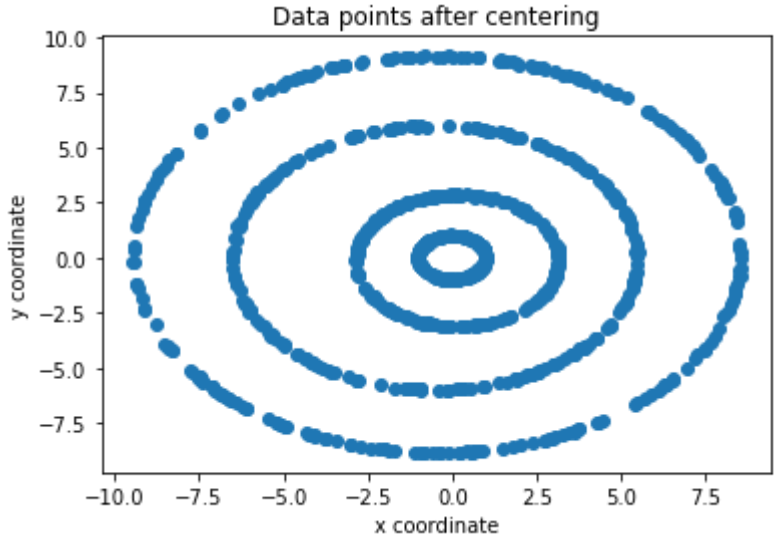
1.i

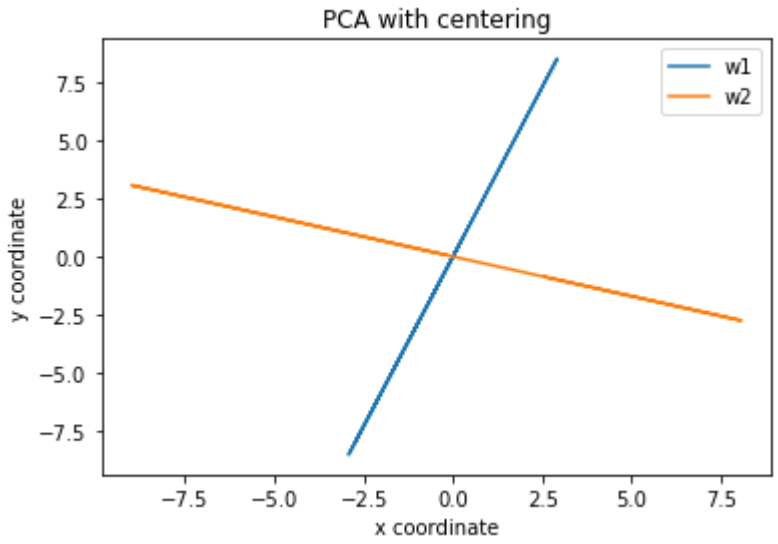
PCA Algorithm



Centering the data



PCA:



The top two eigenvalues obtained are: 17.149063465, 14.504108858

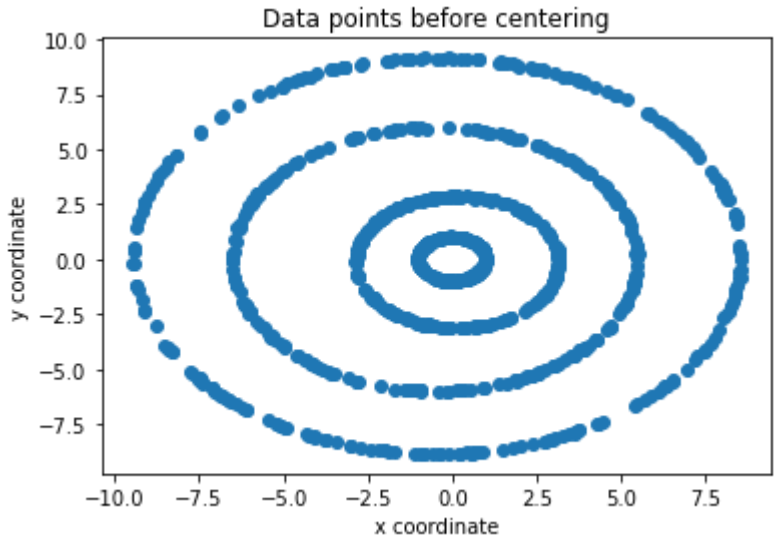
Percentage of variance given by a principal component

= corresponding eigenvalue * 100 /(sum of all eigenvalues)

Percentage of variance explained by Principal component 1 = 54.1780245288 %

Percentage of variance explained by Principal component 2 = 45.8219754711 %

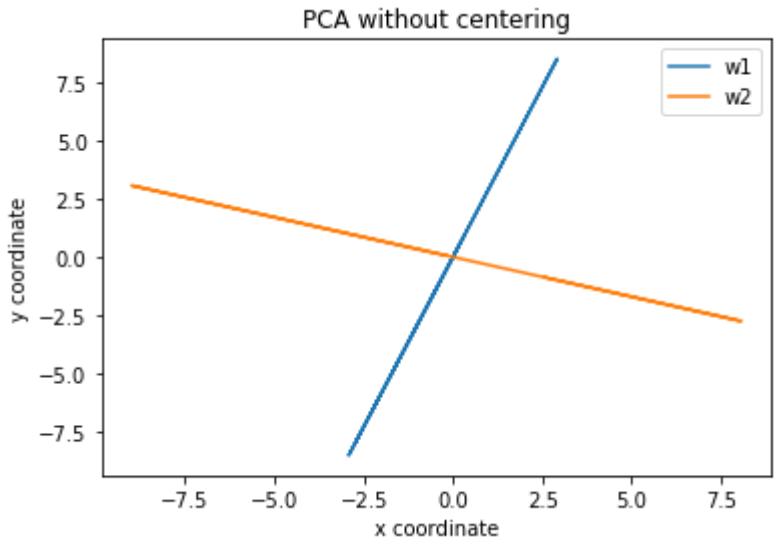
Q1.ii) PCA without centering



The given data is already centered. The mean we calculated for data is approximately the origin.

Hence it has similar results to the PCA with centering.

If we run PCA, Centering has no effect..



The top two eigenvalues obtained are: 17.149063465, 14.504108858

Percentage of variance given by a principal component

= corresponding eigenvalue * 100 /(sum of all eigenvalues)

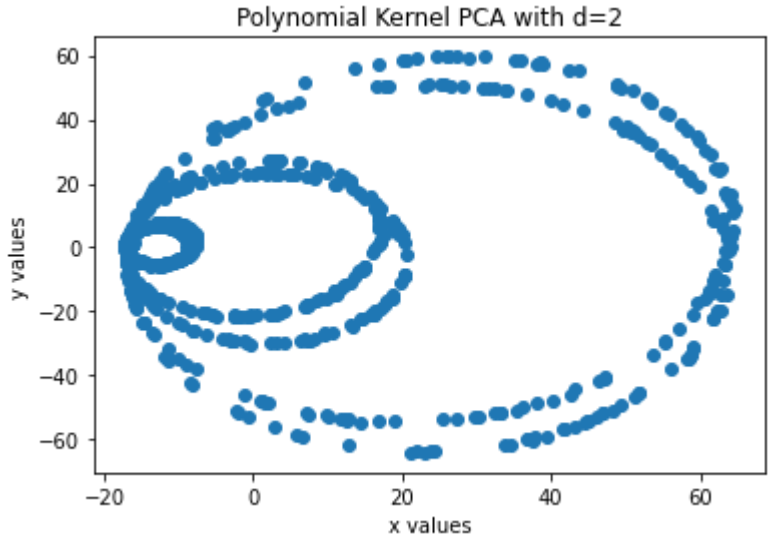
Percentage of variance explained by Principal component 1 = 54.1780245288 %

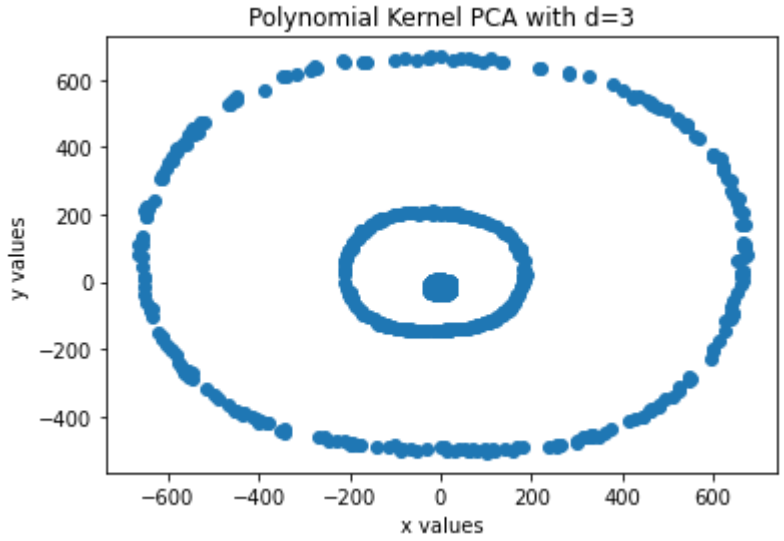
Percentage of variance explained by Principal component 2 = 45.8219754711 %

1.iii A)

We plot the projection of each point in the dataset onto the top-2 components for the polynomial kernel:

κ(x, y) = (1 + x T y) ^d for d = {2, 3}

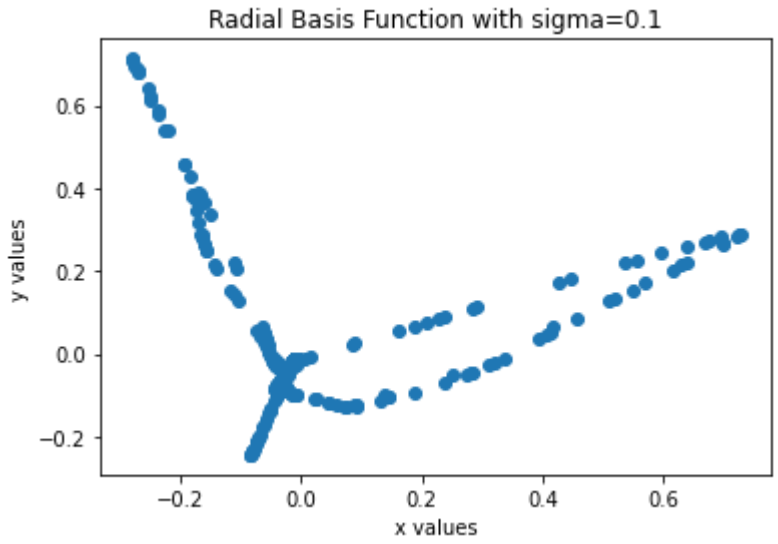


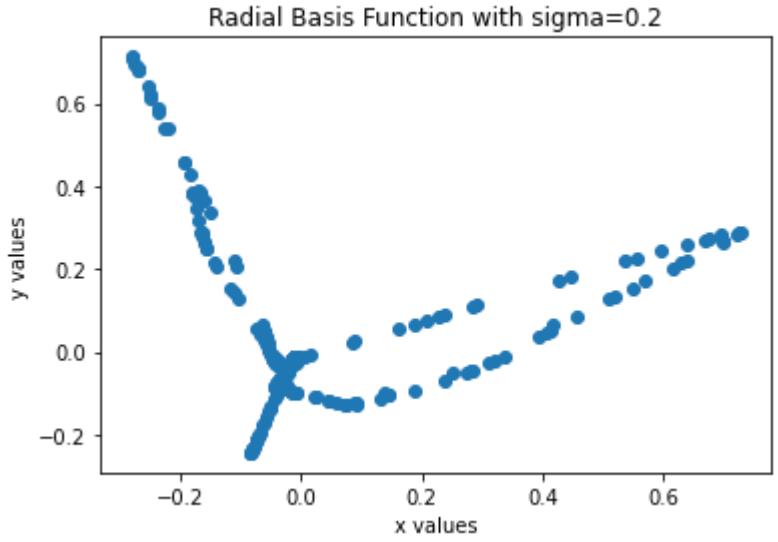


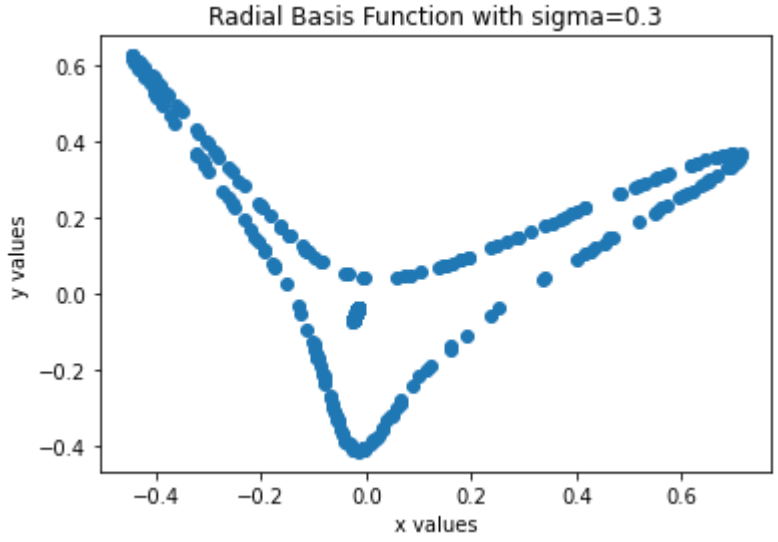
1.iii B)

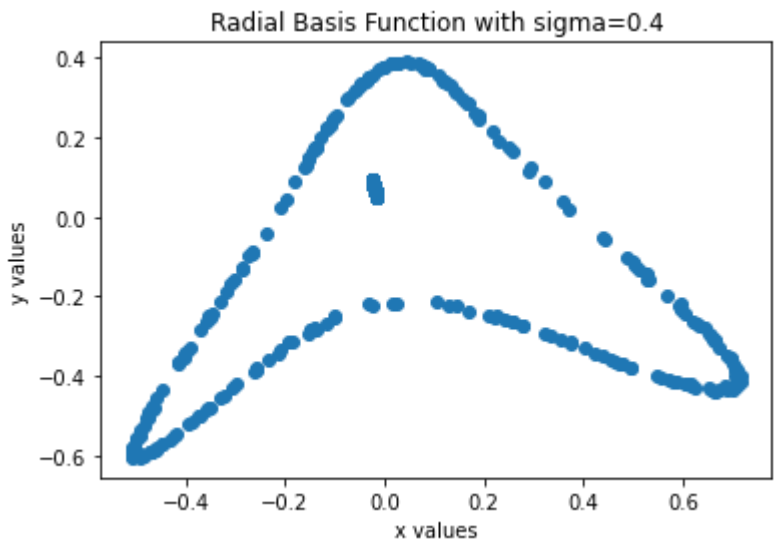
We plot the projection of each point in the dataset onto the top-2 components for the kernel(Radial Basis function):

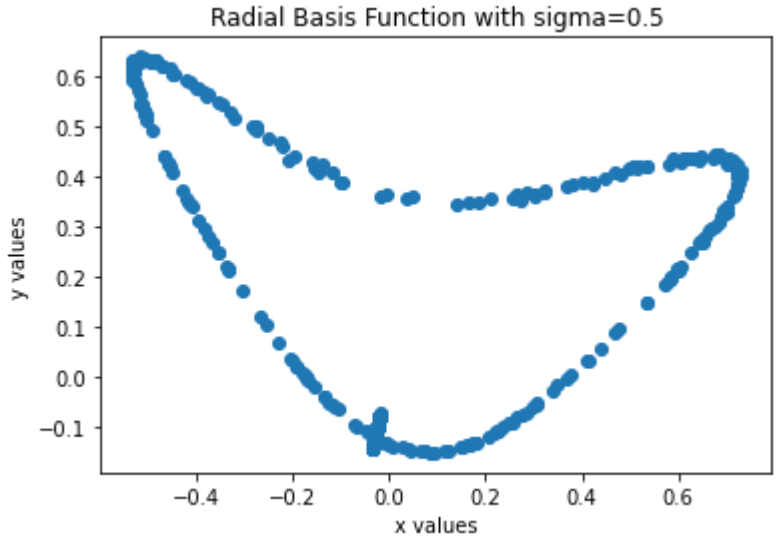
$\kappa(x, y) = \exp(-(x-y)^T (x-y)/2\sigma^2)$ for $\sigma = \{0.1, 0.2, \ldots, 1\}$

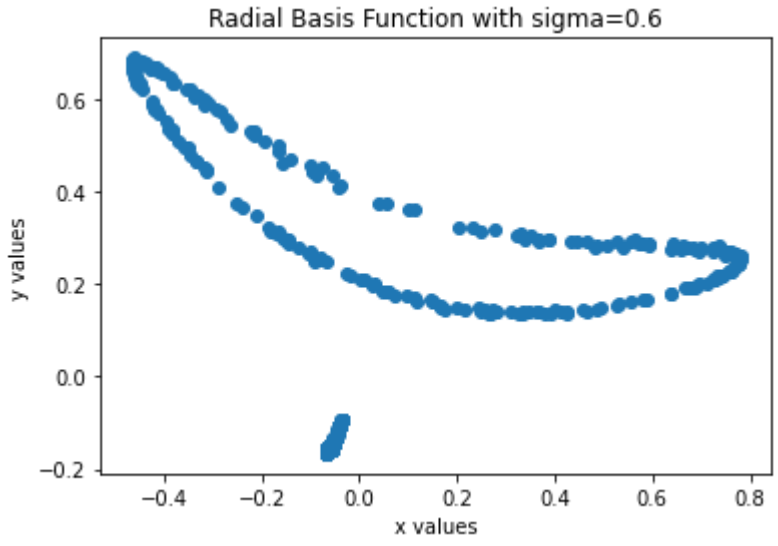

Radial Basis Function with sigma=0.1
y values
x values
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6

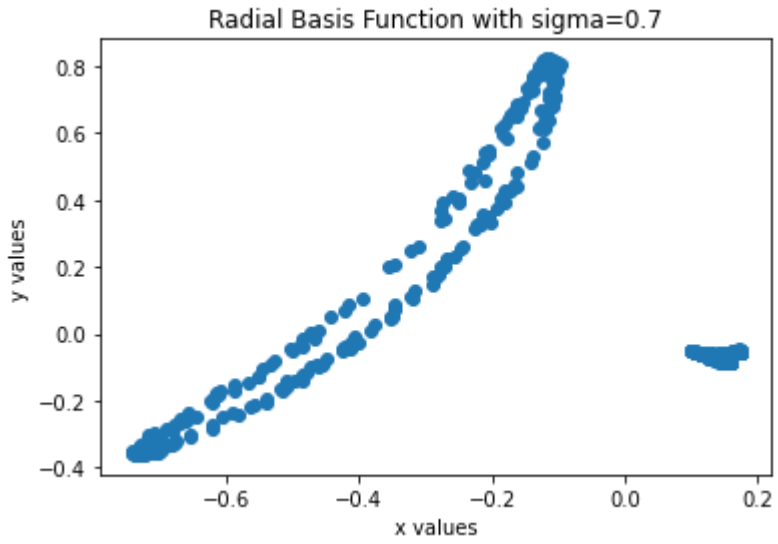

Radial Basis Function with sigma=0.2
y values
x values
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6

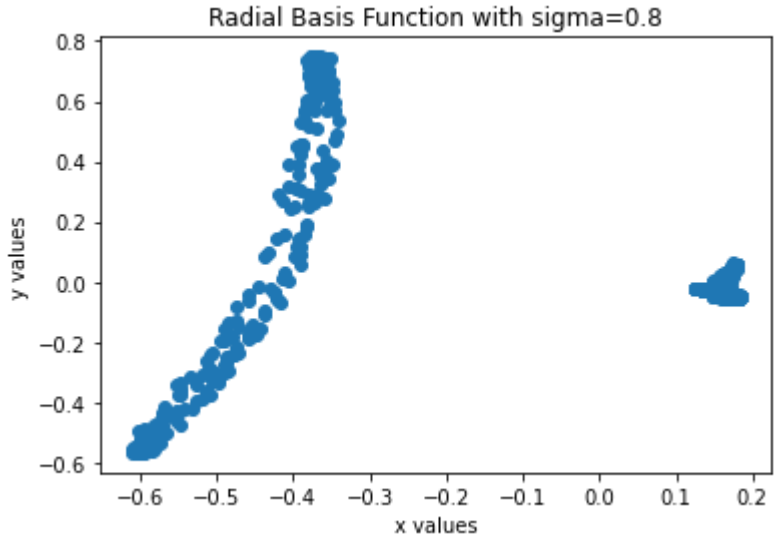

Radial Basis Function with sigma=0.3
y values
x values
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.4
−0.2
0.0
0.2
0.4
0.6

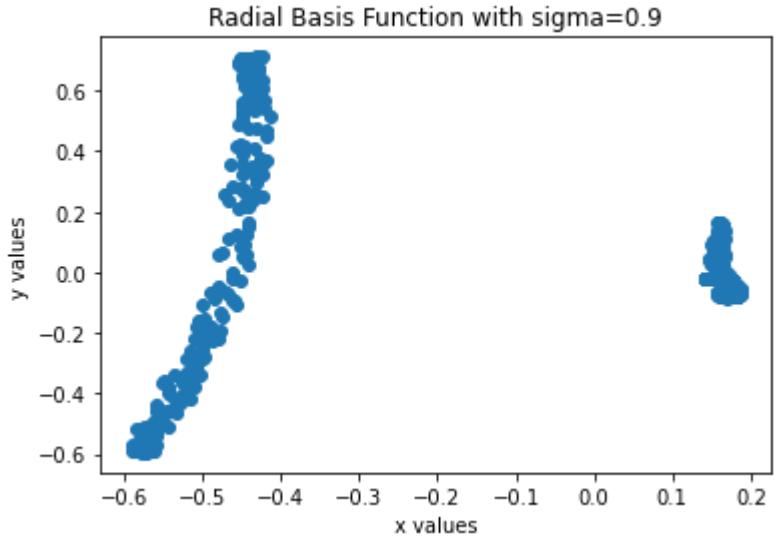

Radial Basis Function with sigma=0.4
y values
x values

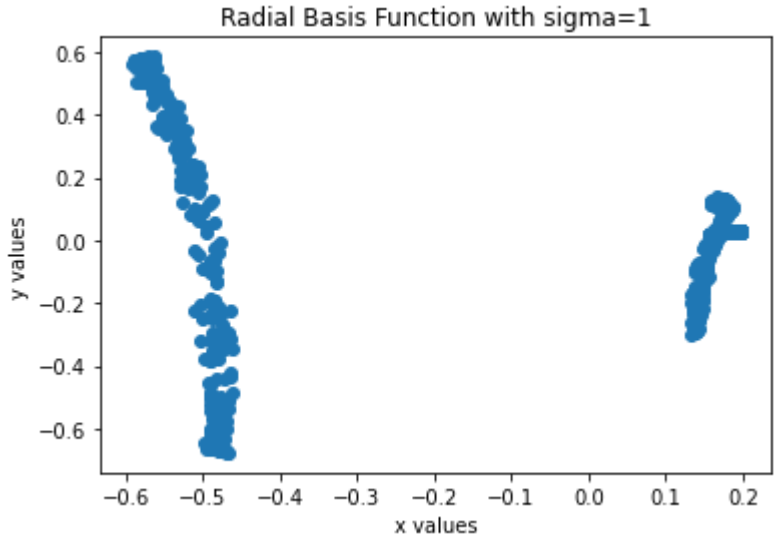

Radial Basis Function with sigma=0.5
y values
x values


Radial Basis Function with sigma=0.6
y values
x values


Radial Basis Function with sigma=0.7
y values
x values
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.4
−0.2
0.0
0.2


Radial Basis Function with sigma=0.8
y values
x values
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2





Q 1 iv

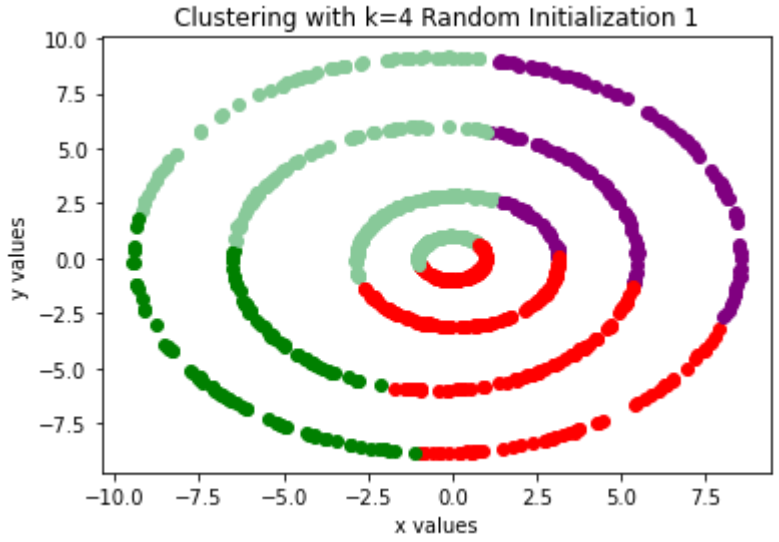
We choose the below polynomial kernel function:

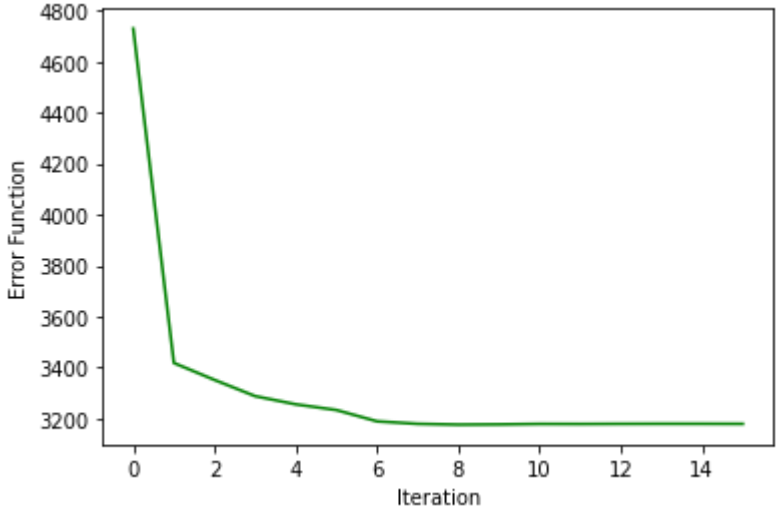
$$K(x, y) = (1 + x^T y)^d \text{ for } d = \{2\}$$

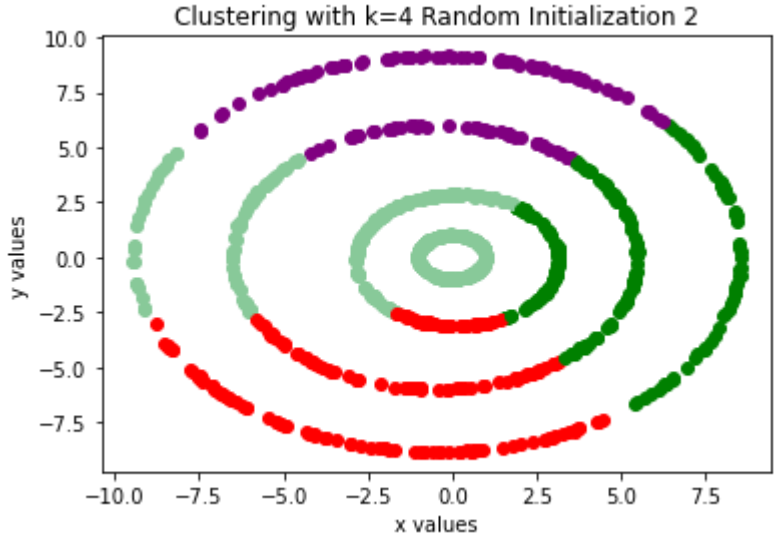
This is because among the given Kernel functions, the clusters formed after running the spectral clustering algorithm with k=4 are clearly distinct from each other.

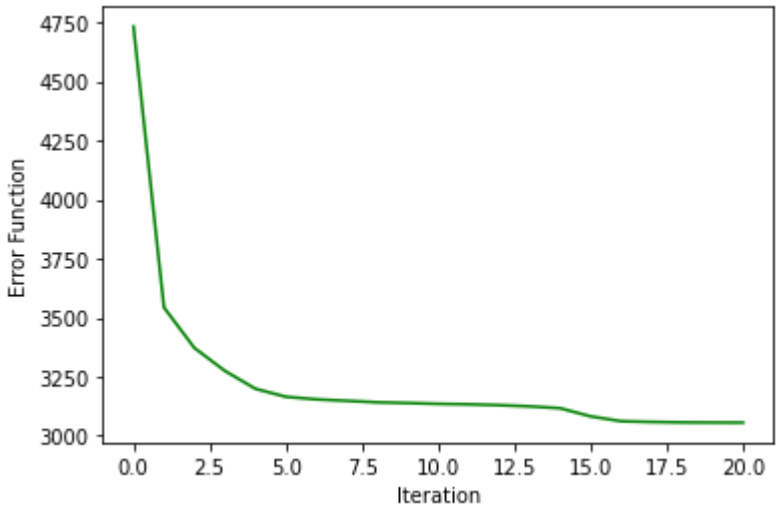
Q 2 i)

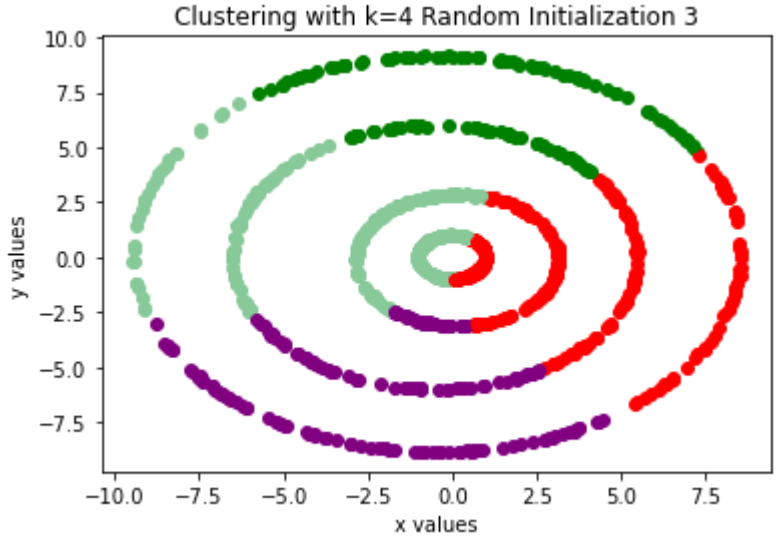
We plot the clusters obtained for 5 different random initialization with k=4

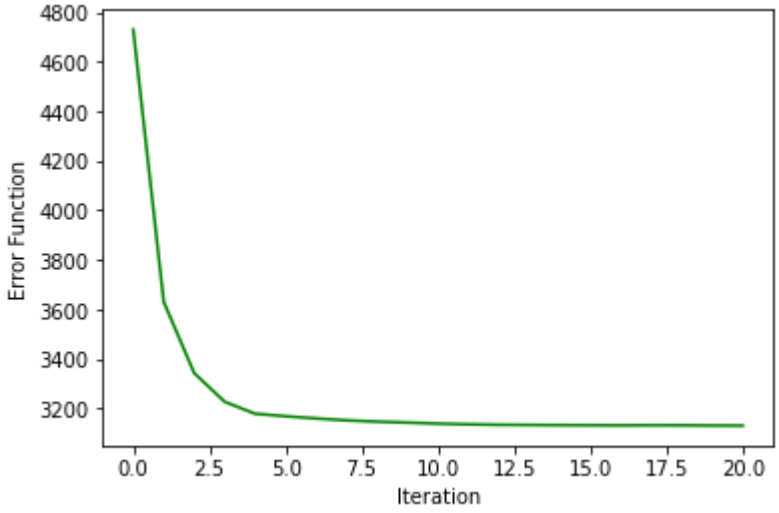

Clustering with k=4 Random Initialization 1
y values
x values

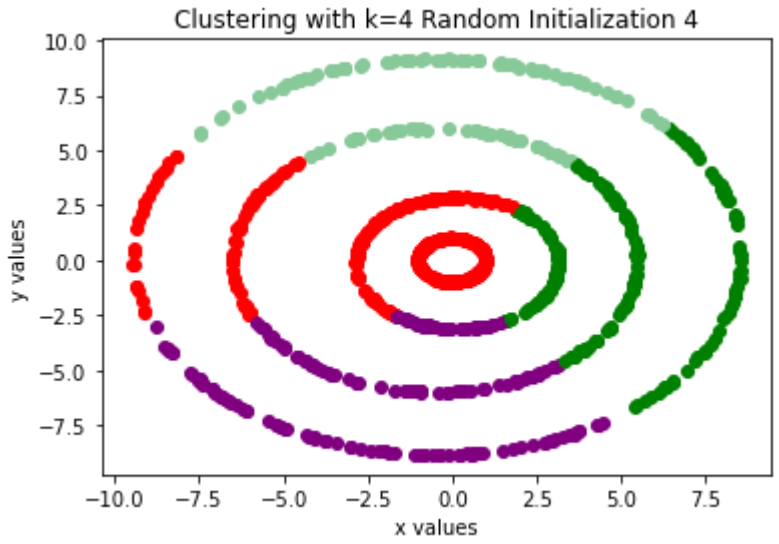

Error Function
Iteration

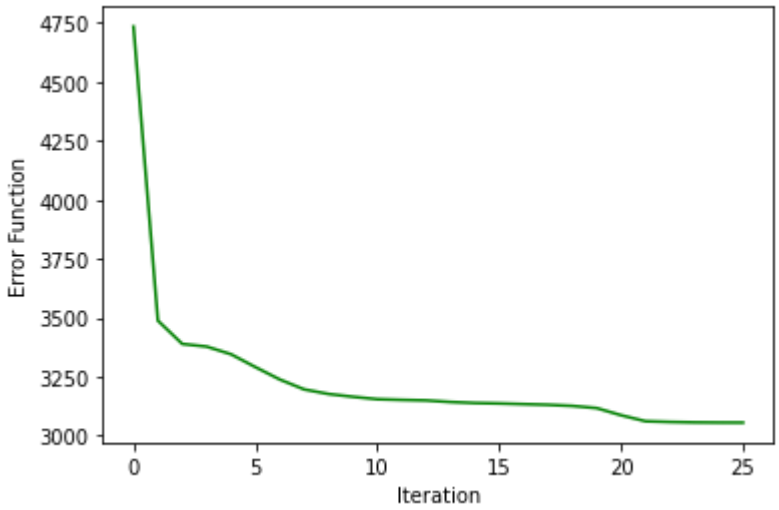

Clustering with k=4 Random Initialization 2
y values
x values

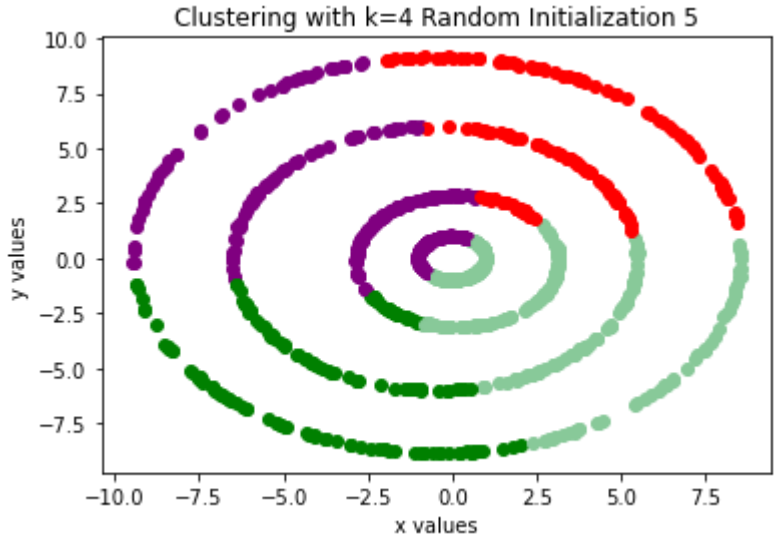

Error Function
Iteration

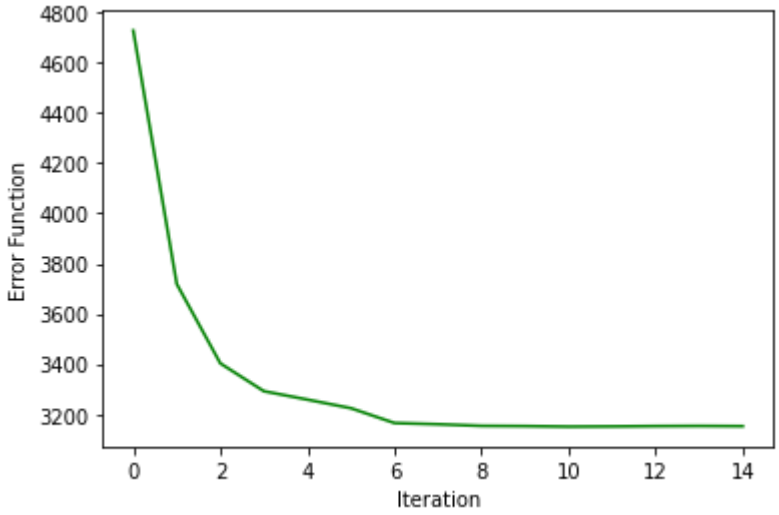

Clustering with k=4 Random Initialization 3
y values
x values


Error Function
Iteration


Clustering with k=4 Random Initialization 4
y values
x values
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5

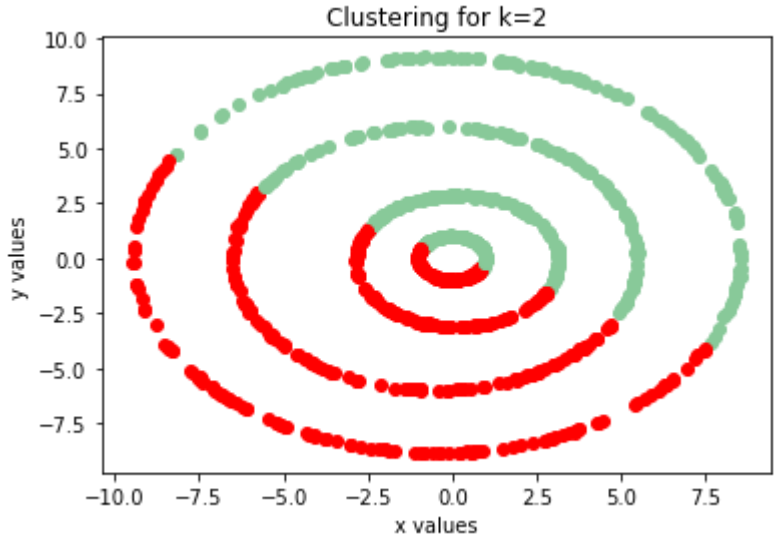

Error Function
Iteration
4750
4500
4250
4000
3750
3500
3250
3000
0
5
10
15
20
25

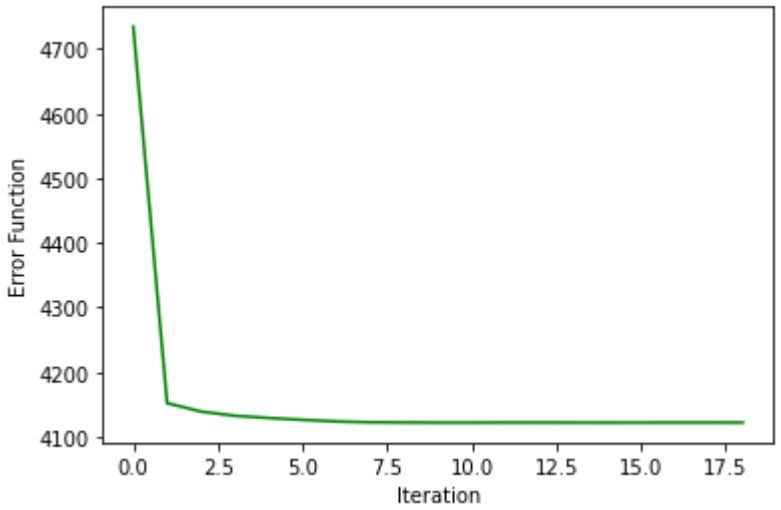


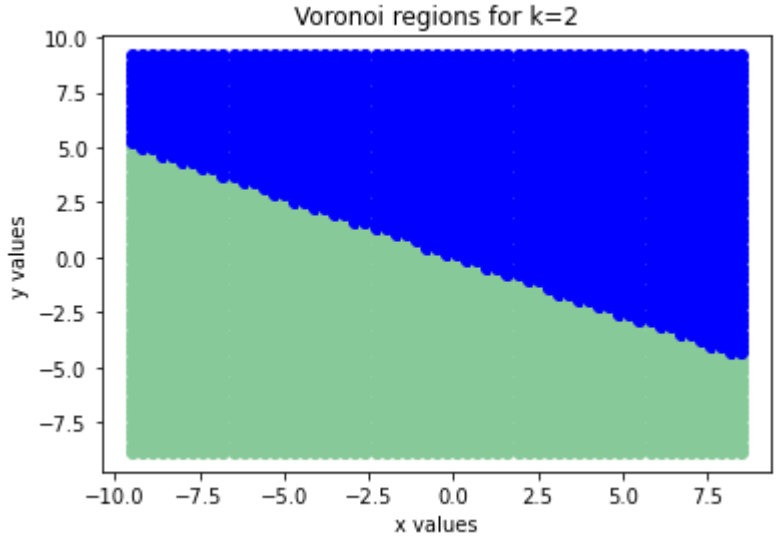


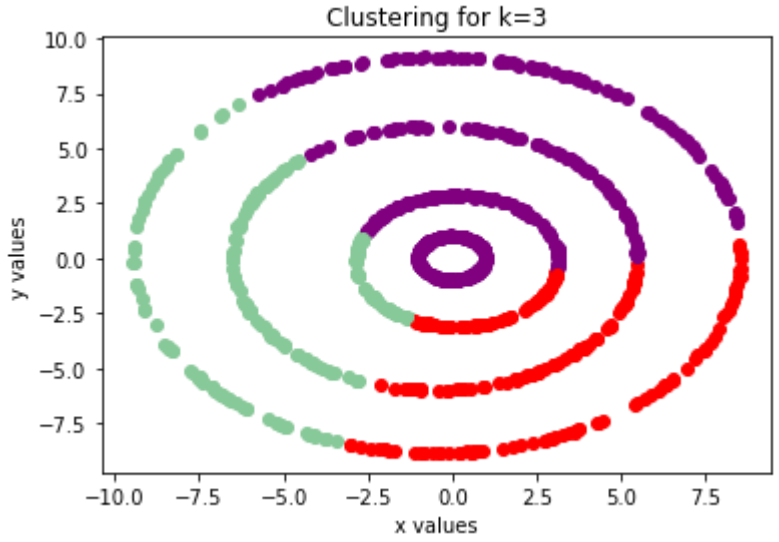
Q 2 ii)

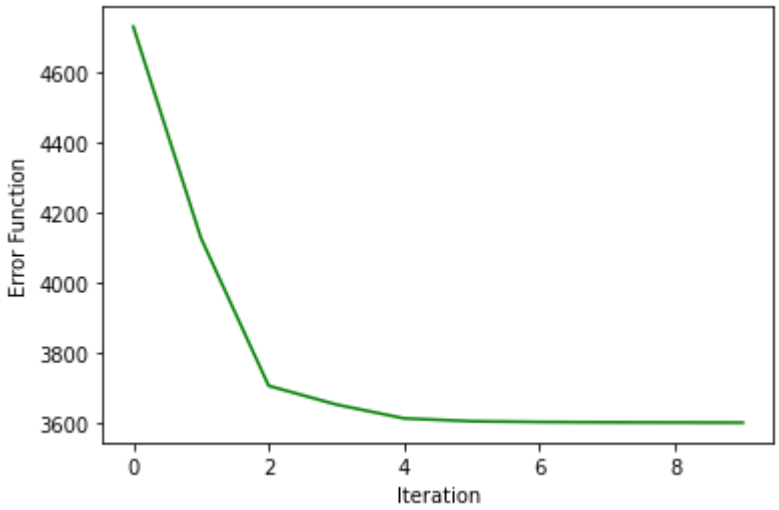
We plot the Voronoi regions associated with each cluster center:

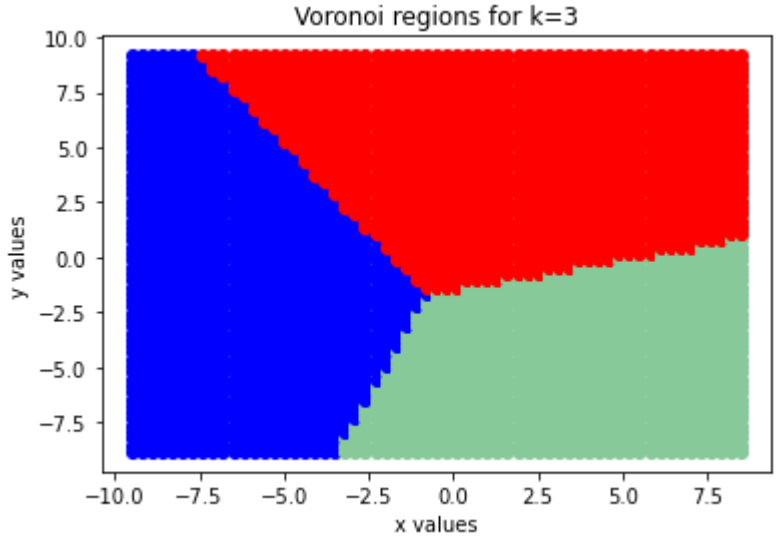

Clustering for k=2
y values
x values

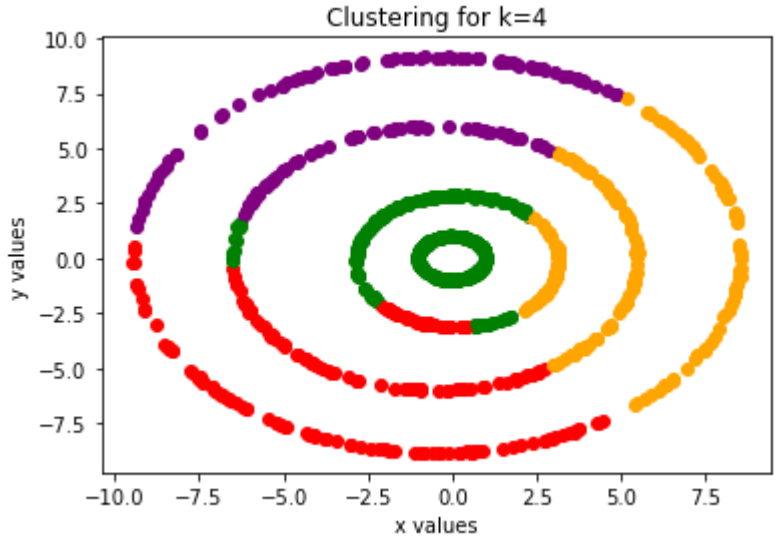

Error Function
Iteration

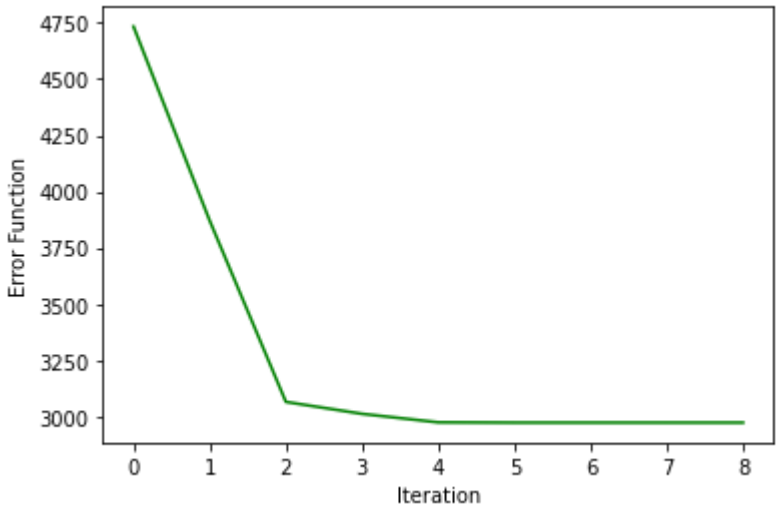

Voronoi regions for k=2
y values
x values

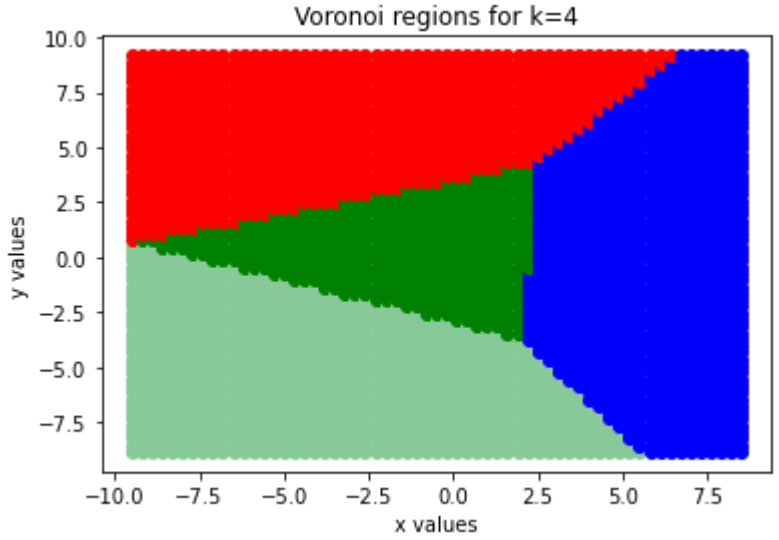

Clustering for k=3
y values
x values
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5

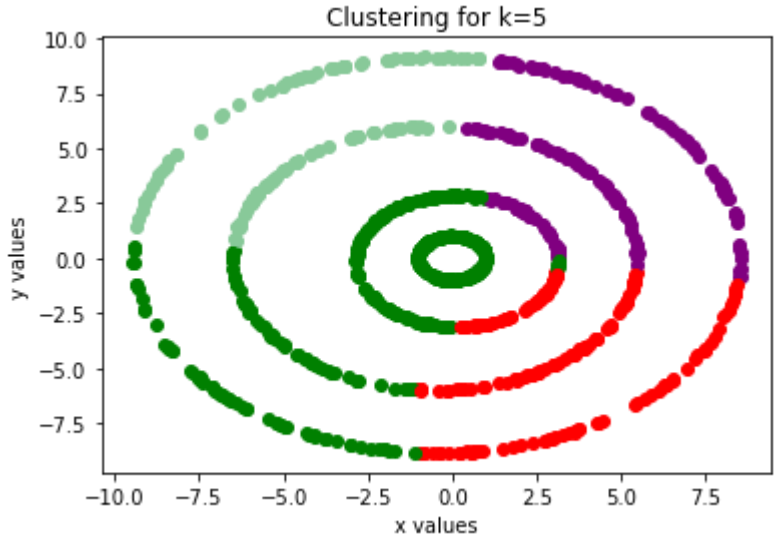

Error Function
Iteration
4600
4400
4200
4000
3800
3600
0
2
4
6
8

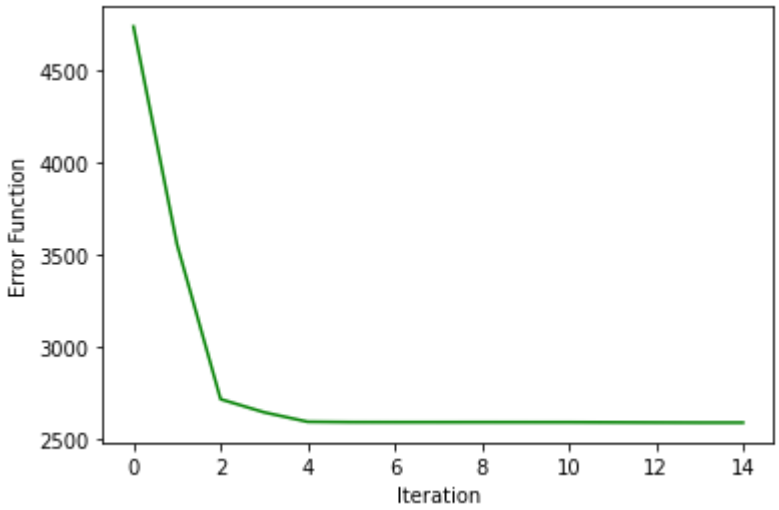

Voronoi regions for k=3
y values
x values
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5

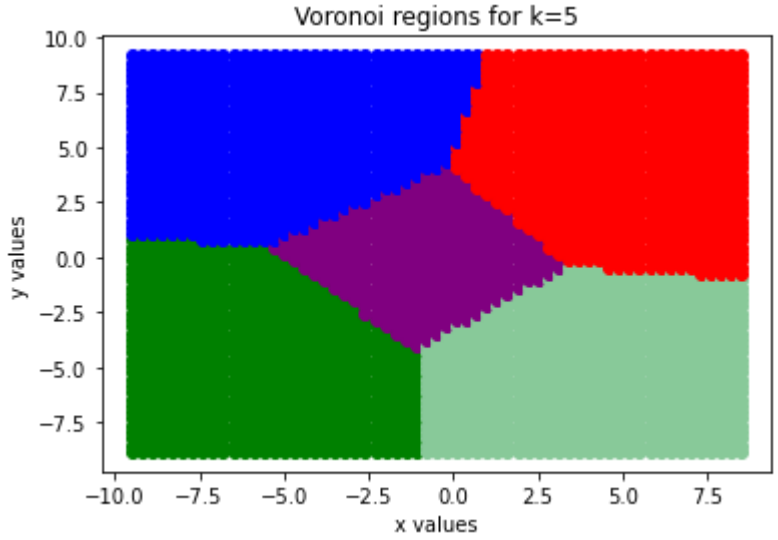

Clustering for k=4
y values
x values


Error Function
Iteration


Voronoi regions for k=4
y values
x values







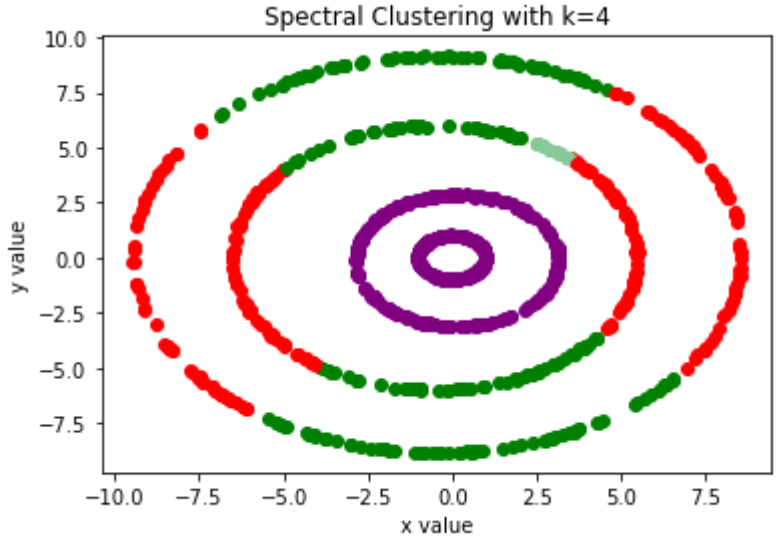
Q 2 iii)

We choose the below polynomial kernel function:

$$K(x, y) = (1 + x^T y)^d \text{ for } d = \{2\}$$

This is because the clusters formed after running the spectral clustering algorithm with k=4 are clearly distinct from each other

Following are results of spectral clustering.



Q 2 iv)The mapping does not perform good for the dataset since it is not able cluster properly i.e. the clusters formed are not distinct from each other