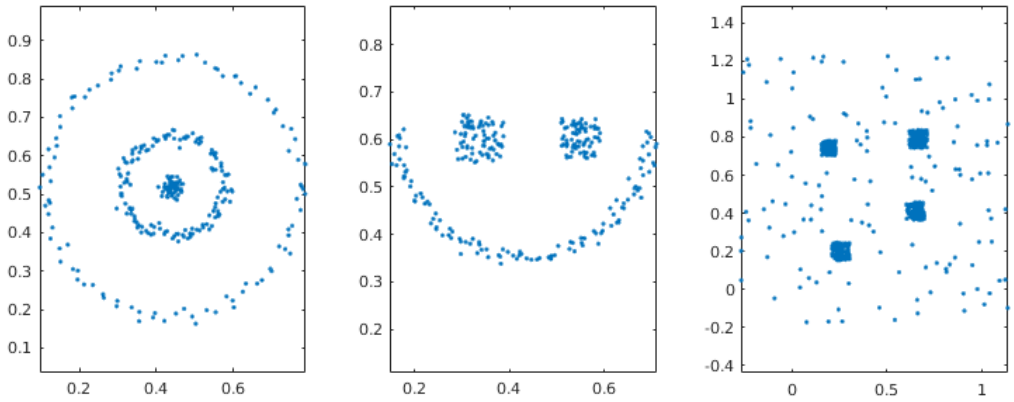
Clustering Computer Lab: March 31 2025

For the lab, we want to provide a clustering of some "toy" data. The data is stored in the Matlab file `HWMar26data.mat` as matrix A, B, and C (each is stored as "number of points" by 2), and the data for the last problem is described below.

Problem 1 (data set A) is a worked example, so you only need to turn in your "solutions" to 2, 3 and 4. The lab is due (uploaded to Canvas) **Sunday, Apr 6** at 11:59PM.



1. Problem 1: For data set A, the following clustering examples have been provided, so you don't need to turn in anything for Problem 1.
   (a) Run $k-$means algorithm with three cluster centers, and plot the results.
   (b) Run Neural Gas with 20 clusters and plot the result (again, a sample will be provided).
   (c) Run DBSCAN. I won't give you the right parameters, but I'll provide some "ball park" estimates you should work with to get a good clustering.

2. Problem 2: (Only DBSCAN)

   For data set B: Just run the DBSCAN algorithm, and see if you can find good parameters so that the clustering is what you would want.

3. Problem 3: (k-means, Neural Gas and DBSCAN):

   For data set C: The point of this example is to have very different densities in the data. I'd like you to be thinking of our claim that $k-$means does not really depend on the density, and Neural Gas does. With that in mind,

   - Use 30 cluster centers in the k-means algorithm. Plot the data with the cluster centers. If the data is in matrix X and the cluster centers in matrix C, then assuming X is $p \times 2$ and C is 30×2, then in Matlab, to plot:

     ```
     plot(X(:,1),X(:,2),'k.',C(:,1),C(:,2),'r*');
     ```

   - Cluster using 30 centers in Neural Gas. The plotting routine should work without changes.
   - Cluster using DBSCAN (you might play around with ϵ and `MinPts` to see if you can get a clustering that looks good). The plotting routines should work without change.

   Upload the scripts you used and either the images or screenshots of the plots.

4. Problem 4: (Neural Gas only) This is the homework problem in the text. We want to find a path through an "obstacle course" using the Neural Gas clustering. Download the data file `obstacle1.mat` from the class website, and modify the previous example- We'll use 20 cluster centers.