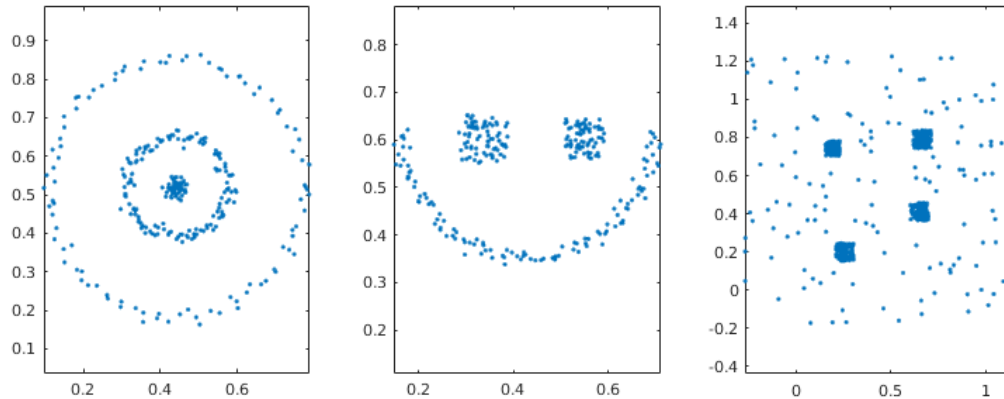


# Clustering Computer Lab: March 26 2021

For the lab, we'll be using the following data sets. They will be stored in the Matlab file `HWMar26data.mat` as matrix  $A$ ,  $B$ , and  $C$  (each is stored as “number of points” by 2).



1. For data set  $A$ :
  - (a) Run  $k$ -means clustering with three clusters. In the starter file that will be prepared for you, you'll see that after clustering, we'll plot the three clusters using different colors. Is the output what you expect?
  - (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. 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. 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,
  - Load the matrix  $C$  from the homework data file to a matrix named  $X$  (we'll use  $C$  for the cluster centers). That is, the first few lines should be:

```
load HWMar26data
X=C;
clear A B C
```
  - Cluster using  $k$ -means and 30 cluster centers. Change the plotting commands to just plot the data (in  $X$  and the cluster centers (in  $C$ )).

```
plot(X(:,1),X(:,2),'.',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  $\epsilon$  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.