

Clustering

1. Plot the datasets in the code below with a different color for each clustX .
2. Describe the dataset using Pandas.
3. Get the clusters using kmeans and DBSCAN
4. Plot the each cluster with a different color (2 images, 1 for kmeans, 1 for DBSCAN)
5. Export the data to MS Excel
6. Generate a Jupyter notebook

```
import numpy as np
import matplotlib.pyplot as plt

from sklearn.metrics import silhouette_score
from sklearn import cluster, datasets, mixture
from sklearn.neighbors import kneighbors_graph

np.random.seed(844)
clust1 = np.random.normal(5, 2, (1000,2))
clust2 = np.random.normal(15, 3, (1000,2))
clust3 = np.random.multivariate_normal([17,3], [[1,0],[0,1]], 1000)
clust4 = np.random.multivariate_normal([2,16], [[1,0],[0,1]], 1000)
dataset1 = np.concatenate((clust1, clust2, clust3, clust4))

# we take the first array as the second array has the cluster labels
dataset2 = datasets.make_circles(n_samples=1000, factor=.5, noise=.05)[0]
```