import numpy as np

import pandas as pd

from matplotlib import pyplot as plt

from sklearn.mixture import GaussianMixture

from sklearn.cluster import KMeans

data = pd.read\_csv('em.csv')

print("Input Data and Shape")

print(data.shape)

print(data.head())

f1 = data['V1'].values

f2 = data['V2'].values

X = np.array(list(zip(f1, f2)))

print("X ", X)

print('Graph for whole dataset')

plt.scatter(f1, f2, c='black', s=7)

plt.show()

kmeans = KMeans(20, random\_state=0)

labels = kmeans.fit(X).predict(X)

print("labels ",labels)

centroids = kmeans.cluster\_centers\_

print("centroids ",centroids)

plt.scatter(X[:,0],X[:,1],c=labels,s=40,cmap='viridis');

print('Graph using Kmeans Algorithm')

plt.scatter(centroids[:, 0], centroids[:, 1], marker='\*', s=200, c='#050505')

plt.show()

gmm = GaussianMixture(n\_components=3).fit(X)

labels = gmm.predict(X)

probs = gmm.predict\_proba(X)

size = 10 \* probs.max(1) \*\* 3

print('Graph using EM Algorithm')

plt.scatter(X[:, 0], X[:, 1], c=labels, s=size, cmap='viridis');

plt.show()