

In [2]:

```

import random
import math
import numpy as np

def random_centers(data,k):
    data_ = np.array(data)
    centres_list = []
    for x in range(0,k):
        temp_rand = np.random.randint( 0,high = 1999)
        temp = data_[temp_rand]
        centres_list.append(temp)
        data_ = np.delete(data_,temp,axis = 0)
    return((centres_list))

def assign_data_centers(data, k,centres):
    thePartition = [[] for _ in range(0,k)] # list of k empty lists

    c = np.array(centres)
    dp = np.array(data)

    for a in dp:
        temp_norm = np.linalg.norm((c-a),axis =1 )
        minn = temp_norm.min()
        temp_list = list(temp_norm)
        thePartition[temp_list.index(minn)].append(a)

    return thePartition

def revalaute_centres(data_with_centres,k):
    data = data_with_centres
    new_centres=[]
    for x in data:
        temp = np.array(x)
        temp_m = temp.mean(axis = 0)
        new_centres.append(temp_m)
    return(new_centres)

```

In [3]:

```

import pandas as pd
data = pd.read_csv('clustering.csv')

```

[illegible]

In [6]:

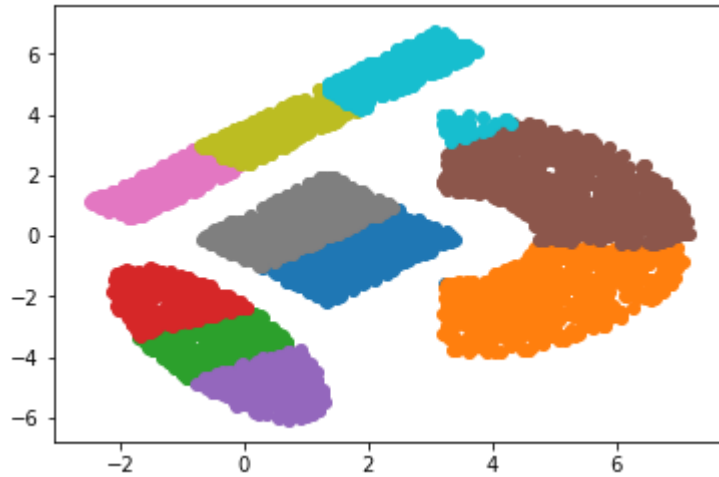
```
import pandas as pd
data = pd.read_csv('ShapedData.csv')
```

[illegible]

2/3

In [12]:

```
import matplotlib.pyplot as plt
plt.plot()
for i in D_C_2:
    plt.scatter(pd.DataFrame(i)[0],pd.DataFrame(i)[1])
plt.show()
```

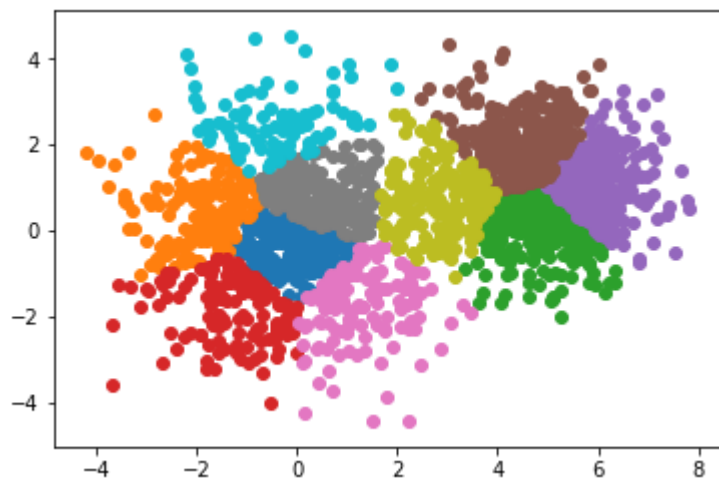


In [9]:

```
import pandas as pd
data = pd.read_csv('ShapedData.csv')
```

In [10]:

```
import matplotlib.pyplot as plt
plt.plot()
for i in D_C_2:
    plt.scatter(pd.DataFrame(i)[0],pd.DataFrame(i)[1])
plt.show()
```



In []: