dbscan

# example code

|  |
| --- |
| import numpy as np  import matplotlib.pyplot as plt  import pandas as pd  data = pd.read\_csv("Mall\_Customers.csv") # importing the dataset  data.head()  print("Dataset shape:", data.shape)  # checking for NULL data in the dataset  data.isnull().any().any()  # extracting the above mentioned columns  x = data.iloc[:, [3,4]].values  print(x.shape)  from sklearn.neighbors import NearestNeighbors # importing the library  neighb = NearestNeighbors(n\_neighbors=2) # creating an object of the NearestNeighbors class  nbrs=neighb.fit(x) # fitting the data to the object  distances,indices=nbrs.kneighbors(x) # finding the nearest neighbours  # Sort and plot the distances results  distances = np.sort(distances, axis = 0) # sorting the distances  distances = distances[:, 1] # taking the second column of the sorted distances  plt.rcParams['figure.figsize'] = (5,3) # setting the figure size  plt.plot(distances) # plotting the distances  plt.show() # showing the plot  from sklearn.cluster import DBSCAN  # cluster the data into five clusters  dbscan = DBSCAN(eps = 8, min\_samples = 4).fit(x) # fitting the model  labels = dbscan.labels\_ # getting the labels  # Plot the clusters  plt.scatter(x[:, 0], x[:,1], c = labels, cmap= "plasma") # plotting the clusters  plt.xlabel("Income") # X-axis label  plt.ylabel("Spending Score") # Y-axis label  plt.show() # showing the plot |

# testing result

|  |
| --- |
|  |