PRACTICAL:7

AIM:

Use K-Means Clustering and Hierarchical Clustering algorithm for following datasets.

CODE:

import numpy as np

import pandas as pd

from sklearn.cluster import KMeans

from sklearn.cluster import AgglomerativeClustering

import matplotlib.pyplot as plt

df = pd.read\_csv("train.csv")

df

df = df.iloc[:5000, :50]

df = df.iloc[:5000, :50]

df.describe()

k\_means\_score = []

for i in range(1,11):

  kmeans=KMeans(n\_clusters=i)

  kmeans.fit(df)

  k\_means\_score.append(kmeans.inertia\_)

plt.plot(range(1,11), k\_means\_score)

plt.show()

kmeans = KMeans(n\_clusters=2)

y\_kmeans = kmeans.fit\_predict(df)

df = np.array(df)

plt.scatter(df[y\_kmeans==0, 0], df[y\_kmeans==0, 1], c="red")

plt.scatter(df[y\_kmeans==1, 0], df[y\_kmeans==1, 1], c="blue")

plt.scatter(kmeans.cluster\_centers\_[:,0], kmeans.cluster\_centers\_[:,1],c="yellow")

plt.show()

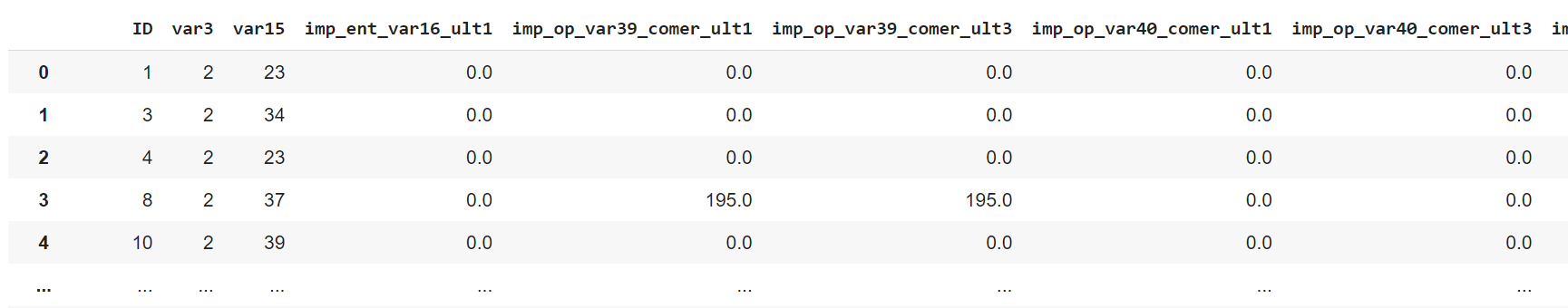
cluster = AgglomerativeClustering(n\_clusters=2)

cluster.fit\_predict(df)

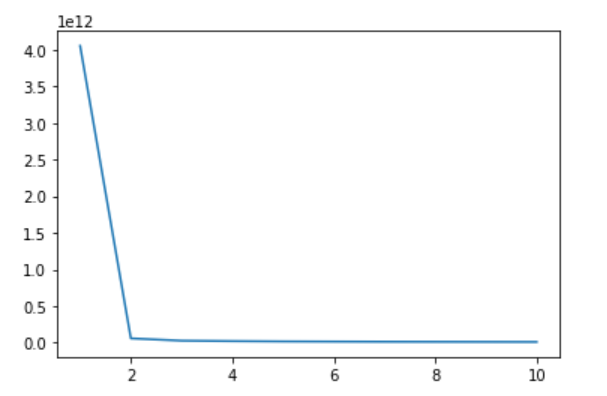
plt.scatter(df[:,0], df[:,1], c=cluster.labels\_, cmap="rainbow")

plt.show()

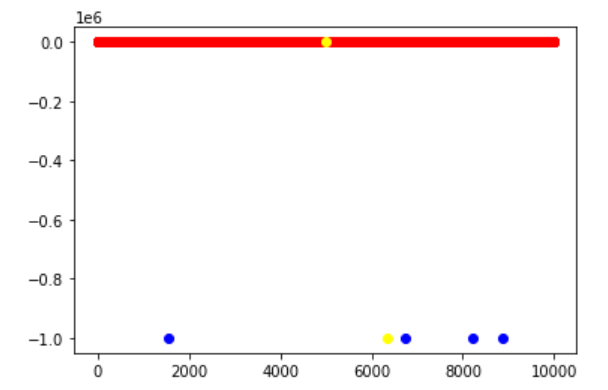
OUTPUT:



Glimpses of dataset



Plot of k\_means\_score



Scatter graph of k-means cluster

