**Name: Neha Antony**

**Roll No: 23**

**Batch: MCA-B**

**Date: 14/11/2022**

**DATA SCIENCE LAB**

**Experiment No.: 12**

**Aim**

K-mean clustering Algorithm.

**Procedure**

from sklearn.cluster import KMeans

import pandas as pd

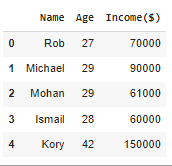
from sklearn.preprocessing import MinMaxScaler

from matplotlib import pyplot as plt

%matplotlib inline

df = pd.read\_csv("/content/income.csv")

df.head()



scaler = MinMaxScaler()

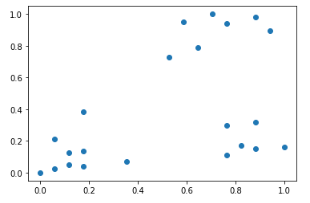
scaler.fit(df[['Income($)']])

df['Income($)'] = scaler.transform(df[['Income($)']])

scaler.fit(df[['Age']])

df['Age'] = scaler.transform(df[['Age']])

plt.scatter(df.Age,df['Income($)'])



km = KMeans(n\_clusters=3)

y\_predicted = km.fit\_predict(df[['Age','Income($)']])

y\_predicted



km.cluster\_centers\_



df1 = df[df.cluster==0]

df2 = df[df.cluster==1]

df3 = df[df.cluster==2]

plt.scatter(df1.Age,df1['Income($)'],color='green')

plt.scatter(df2.Age,df2['Income($)'],color='red')

plt.scatter(df3.Age,df3['Income($)'],color='black')

plt.scatter(km.cluster\_centers\_[:,0],km.cluster\_centers\_[:,1],color='purple',marker='\*',label='centroid')

plt.legend()

