**Lab Submission-4**

**KMeans Algorithms**

**Task : Simple Program to demonstrate the KMeans Algorithm.**

**Source Code:**

from sklearn.cluster import KMeans

import pandas as pd

from matplotlib import pyplot as plt

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

# print(df.head())

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

plt.xlabel('Age')

plt.ylabel('Income($)')

km = KMeans(n\_clusters=3)

y\_predicted = km.fit\_predict(df[['Age','Income($)']])    #Use fit in predict for 2 columns

print(y\_predicted)

df['cluster']=y\_predicted

# print(df.head())

print(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')

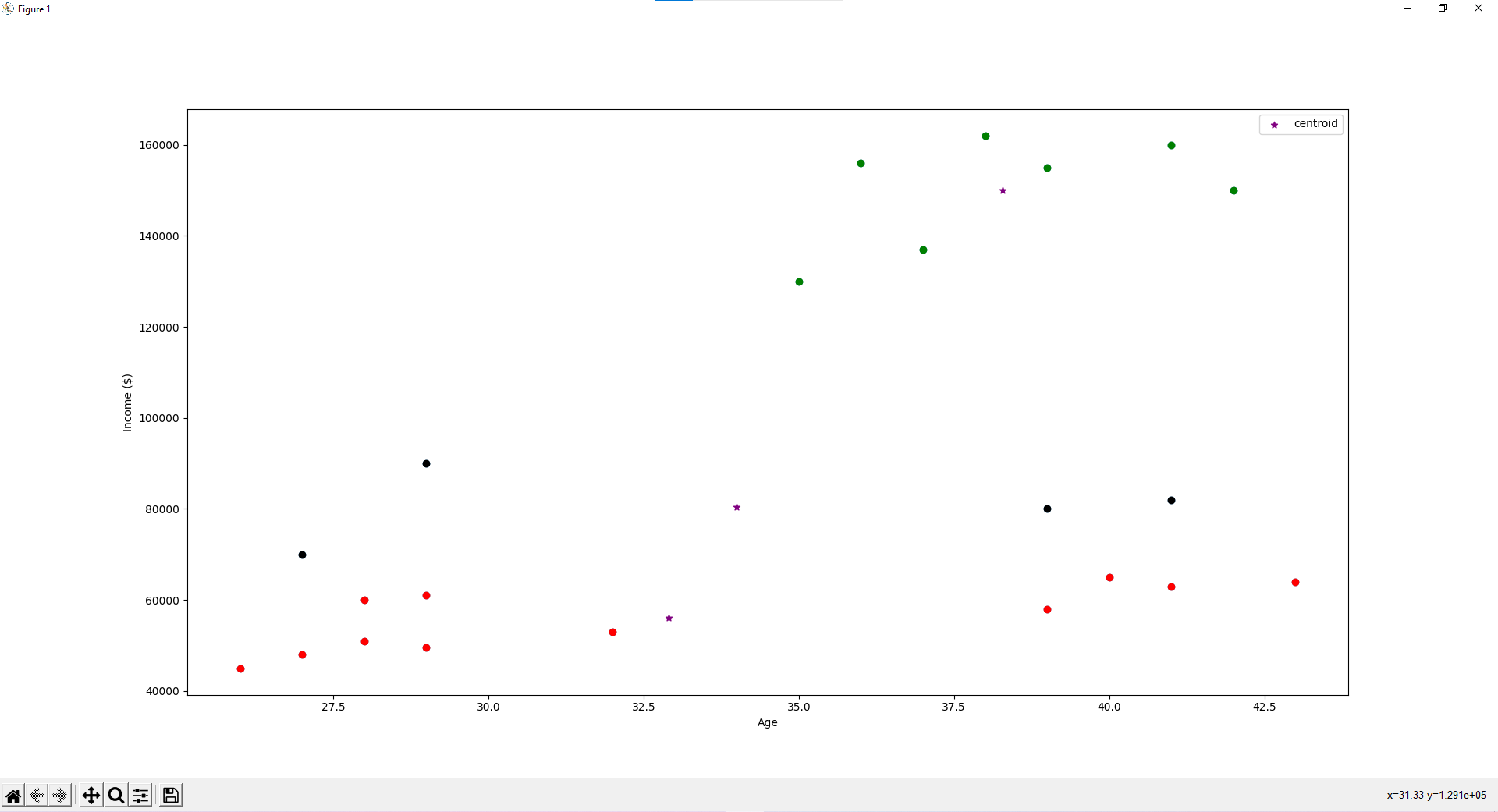
plt.xlabel('Age')

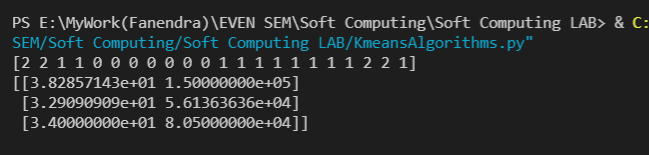
plt.ylabel('Income ($)')

plt.legend()

plt.show()

**Output:**

****

****