ASSIGNMENT – 10

```
import numpy as np

from sklearn.cluster import KMeans

from sklearn.preprocessing import StandardScaler

from sklearn.model_selection import cross_val_score

df = pd.read_csv('C:\Users\Dell\Downloads\CC_GENERAL.csv')

df['BALANCE_FREQUENCY'] = df['BALANCE_FREQUENCY'].fillna(0)

df['PURCHASES_FREQUENCY'] = df['PURCHASES_FREQUENCY'].fillna(0)

df['ONEOFFPURCHASESFREQUENCY'] = df['ONEOFFPURCHASESFREQUENCY'].fillna(0)

df['PURCHASESINSTALLMENTSFREQUENCY'] =
df['PURCHASESINSTALLMENTSFREQUENCY'].fillna(0)

df['CASHADVANCEFREQUENCY'] = df['CASHADVANCEFREQUENCY'].fillna(0)

# Create a StandardScaler object

scaler = StandardScaler()

# Fit the StandardScaler object to the data

scaler.fit(df.drop('CUST_ID', axis=1))

df_scaled = scaler.transform(df.drop('CUST_ID', axis=1))

# Create a KMeans object

kmeans = KMeans(n_clusters=3)

# Fit the KMeans object to the data

kmeans.fit(df_scaled)
```

OUTPUT

CUST_ID BALANCE_FREQUENCY PURCHASES_FREQUENCY PURCHASESINSTALLMENTSFREQUENCY CASHADVANCEFREQUENCY CLUSTER	0 C0000000 01 0.0 0.0 ... 0.0	1 C0000000 02 0.0 0.0 ... 0.0	2 C0000000 03 0.0 0.0 ... 0.0	3 C0000000 04 0.0 0.0 ... 0.0	4 C0000000 05 0.0 0.0 ... 0.0	... ...
995 C000000996 0.0 0.0 ... 0.0	996 C0000009 97 0.0 0.0 ... 0.0	997 C0000009 98 0.0 0.0 ... 0.0	998 C0000009 99 0.0 0.0 ... 0.0	999 C0000010 00 0.0 0.0 ... 0.0	[1000 rows x 12 columns]	