Market segmentation

Market segmentation in marketing refers to the practice of categorizing a broad market of consumers or businesses into smaller, more homogeneous groups based on shared characteristics or traits. This division enables businesses to tailor their marketing strategies and offerings to better meet the specific needs and preferences of each segment. By understanding the distinct behaviors, preferences, demographics, and psychographics of different segments, companies can more effectively target their marketing efforts and optimize their resources to attract and retain customers.

My jupyter notebook:
https://github.com/alex80ds/Boulder-CU/blob/main/Market_segmentation.ipynb

My PDF file: https://github.com/alex80ds/Boulder-CU/blob/main/Market_segmentation.pdf

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings("ignore")

df=pd.read_csv('https://raw.githubusercontent.com/alex80ds/Boulder-CU/main/Market_segmentation.csv')

df.head()
```

```
  CUST_ID      BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
0  C10001    40.900749           0.818182      95.40              0.00
1  C10002  3202.467416           0.909091       0.00              0.00
2  C10003  2495.148862           1.000000     773.17            773.17
3  C10004  1666.670542           0.636364    1499.00           1499.00
4  C10005   817.714335           1.000000      16.00             16.00

   INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
0                    95.4      0.000000             0.166667
1                     0.0   6442.945483             0.000000
2                     0.0      0.000000             1.000000
3                     0.0    205.788017             0.083333
4                     0.0      0.000000             0.083333

   ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
```

```
0                     0.000000                                  0.083333
1                     0.000000                                  0.000000
2                     1.000000                                  0.000000
3                     0.083333                                  0.000000
4                     0.083333                                  0.000000

   CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
0                 0.000000                 0              2
1000.0
1                 0.250000                 4              0
7000.0
2                 0.000000                 0             12
7500.0
3                 0.083333                 1              1
7500.0
4                 0.000000                 0              1
1200.0

      PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE
0   201.802084        139.509787          0.000000      12
1  4103.032597       1072.340217          0.222222      12
2   622.066742        627.284787          0.000000      12
3     0.000000               NaN          0.000000      12
4   678.334763        244.791237          0.000000      12
```

EDA

```
df.shape
```

```
(8950, 18)
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8950 entries, 0 to 8949
Data columns (total 18 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   CUST_ID                           8950 non-null   object
 1   BALANCE                           8950 non-null   float64
 2   BALANCE_FREQUENCY                 8950 non-null   float64
 3   PURCHASES                         8950 non-null   float64
 4   ONEOFF_PURCHASES                  8950 non-null   float64
 5   INSTALLMENTS_PURCHASES            8950 non-null   float64
 6   CASH_ADVANCE                      8950 non-null   float64
 7   PURCHASES_FREQUENCY               8950 non-null   float64
 8   ONEOFF_PURCHASES_FREQUENCY        8950 non-null   float64
```

```
 9   PURCHASES_INSTALLMENTS_FREQUENCY  8950 non-null   float64
 10  CASH_ADVANCE_FREQUENCY            8950 non-null   float64
 11  CASH_ADVANCE_TRX                  8950 non-null   int64
 12  PURCHASES_TRX                     8950 non-null   int64
 13  CREDIT_LIMIT                      8949 non-null   float64
 14  PAYMENTS                          8950 non-null   float64
 15  MINIMUM_PAYMENTS                  8637 non-null   float64
 16  PRC_FULL_PAYMENT                  8950 non-null   float64
 17  TENURE                            8950 non-null   int64
dtypes: float64(14), int64(3), object(1)
memory usage: 1.2+ MB
```

```
df.describe()
```

```
            BALANCE  BALANCE_FREQUENCY     PURCHASES  ONEOFF_PURCHASES
\
count   8950.000000        8950.000000   8950.000000       8950.000000

mean    1564.474828           0.877271   1003.204834        592.437371

std     2081.531879           0.236904   2136.634782       1659.887917

min        0.000000           0.000000      0.000000          0.000000

25%      128.281915           0.888889     39.635000          0.000000

50%      873.385231           1.000000    361.280000         38.000000

75%     2054.140036           1.000000   1110.130000        577.405000

max    19043.138560           1.000000  49039.570000      40761.250000


       INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
count             8950.000000   8950.000000          8950.000000
mean               411.067645    978.871112             0.490351
std                904.338115   2097.163877             0.401371
min                  0.000000      0.000000             0.000000
25%                  0.000000      0.000000             0.083333
50%                 89.000000      0.000000             0.500000
75%                468.637500   1113.821139             0.916667
max              22500.000000  47137.211760             1.000000

       ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
count                 8950.000000                       8950.000000
mean                     0.202458                          0.364437
std                      0.298336                          0.397448
min                      0.000000                          0.000000
25%                      0.000000                          0.000000
50%                      0.083333                          0.166667
75%                      0.300000                          0.750000
```

```
max                          1.000000                                 1.000000

       CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
count             8950.000000       8950.000000    8950.000000
8949.000000
mean                 0.135144          3.248827      14.709832
4494.449450
std                  0.200121          6.824647      24.857649
3638.815725
min                  0.000000          0.000000       0.000000
50.000000
25%                  0.000000          0.000000       1.000000
1600.000000
50%                  0.000000          0.000000       7.000000
3000.000000
75%                  0.222222          4.000000      17.000000
6500.000000
max                  1.500000        123.000000     358.000000
30000.000000

           PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT       TENURE
count   8950.000000       8637.000000       8950.000000  8950.000000
mean    1733.143852        864.206542          0.153715    11.517318
std     2895.063757       2372.446607          0.292499     1.338331
min        0.000000          0.019163          0.000000     6.000000
25%      383.276166        169.123707          0.000000    12.000000
50%      856.901546        312.343947          0.000000    12.000000
75%     1901.134317        825.485459          0.142857    12.000000
max    50721.483360      76406.207520          1.000000    12.000000
```

```
df.isnull().sum()
```

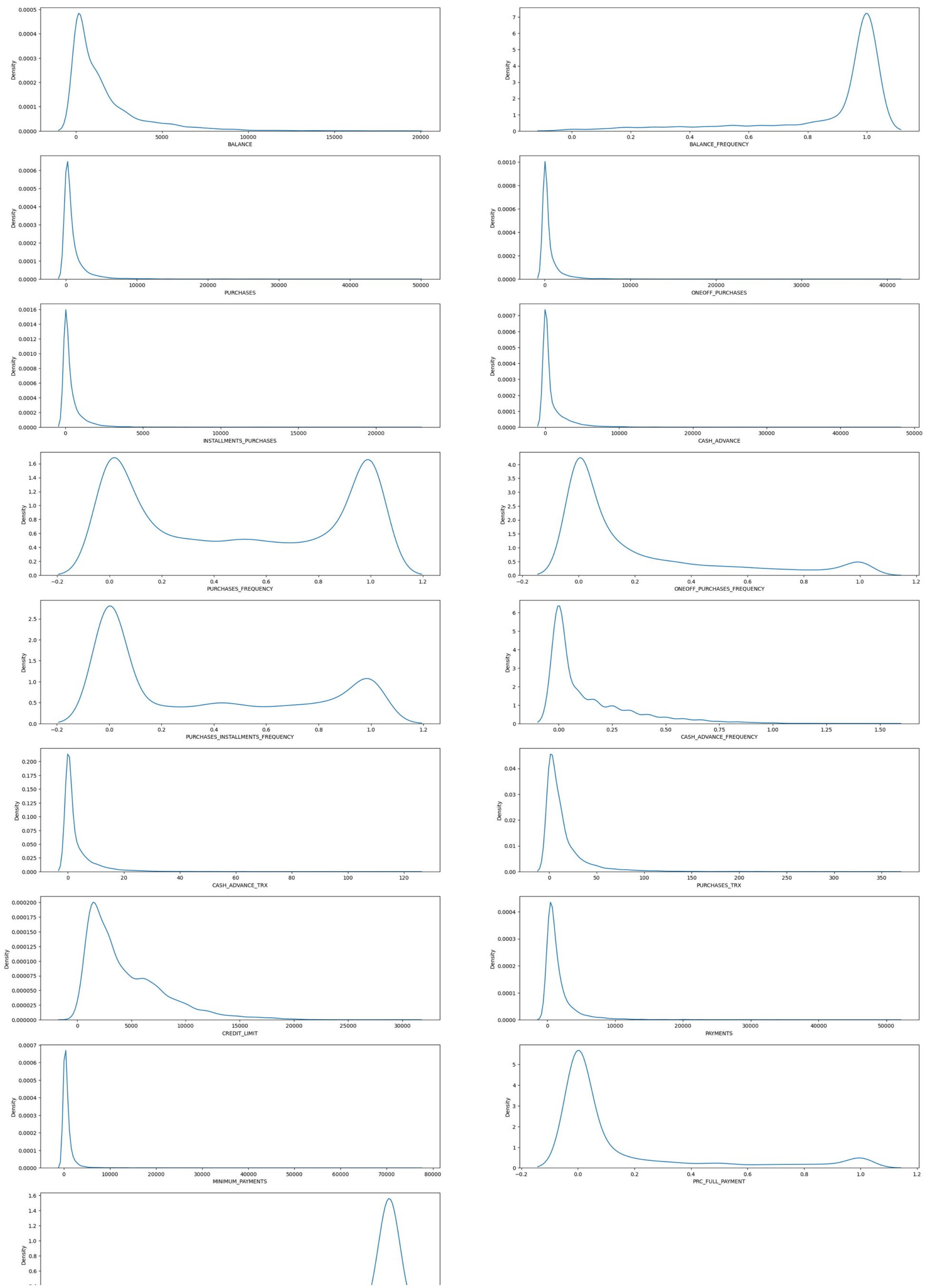
```
CUST_ID                               0
BALANCE                               0
BALANCE_FREQUENCY                     0
PURCHASES                             0
ONEOFF_PURCHASES                      0
INSTALLMENTS_PURCHASES                0
CASH_ADVANCE                          0
PURCHASES_FREQUENCY                   0
ONEOFF_PURCHASES_FREQUENCY            0
PURCHASES_INSTALLMENTS_FREQUENCY      0
CASH_ADVANCE_FREQUENCY                0
CASH_ADVANCE_TRX                      0
PURCHASES_TRX                         0
CREDIT_LIMIT                          1
PAYMENTS                              0
MINIMUM_PAYMENTS                    313
PRC_FULL_PAYMENT                      0
```

```
TENURE                              0
dtype: int64
```

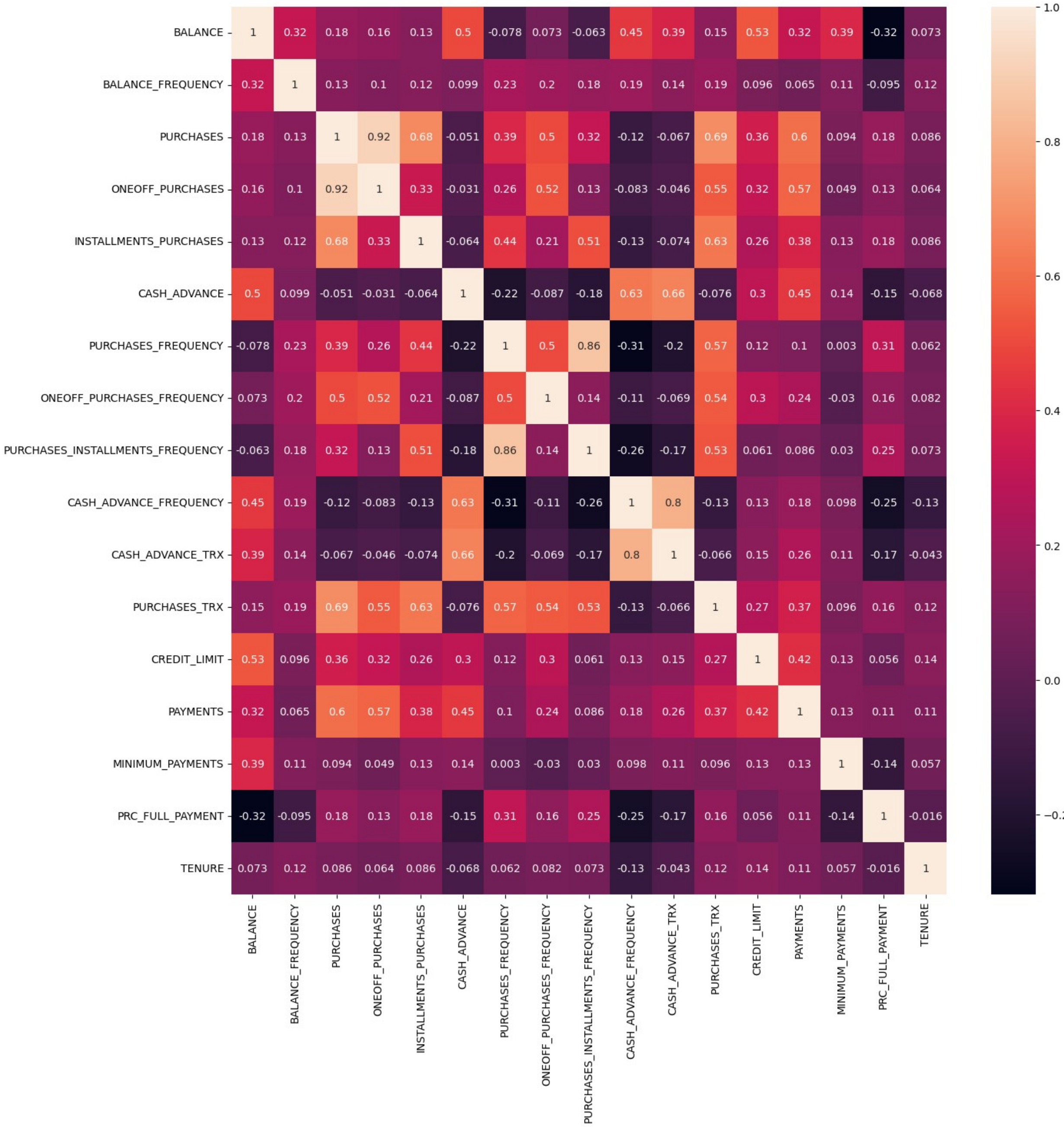
```
df["MINIMUM_PAYMENTS"] =
df["MINIMUM_PAYMENTS"].fillna(df["MINIMUM_PAYMENTS"].mean())
df["CREDIT_LIMIT"] =
df["CREDIT_LIMIT"].fillna(df["CREDIT_LIMIT"].mean())

df.drop(columns=["CUST_ID"],axis=1,inplace=True)
```

Visualization

```
plt.figure(figsize=(30,45))
for i, col in enumerate(df.columns):
    if df[col].dtype != 'object':
        ax = plt.subplot(9, 2, i+1)
        sns.kdeplot(df[col], ax=ax)
        plt.xlabel(col)

plt.show()
```


Density
BALANCE
BALANCE_FREQUENCY
PURCHASES
ONEOFF_PURCHASES
INSTALLMENTS_PURCHASES
CASH_ADVANCE
PURCHASES_FREQUENCY
ONEOFF_PURCHASES_FREQUENCY
PURCHASES_INSTALLMENTS_FREQUENCY
CASH_ADVANCE_FREQUENCY
CASH_ADVANCE_TRX
PURCHASES_TRX
CREDIT_LIMIT
PAYMENTS
MINIMUM_PAYMENTS
PRC_FULL_PAYMENT

```
plt.figure(figsize=(15,15))
sns.heatmap(df.corr(), annot=True)
plt.show()
```



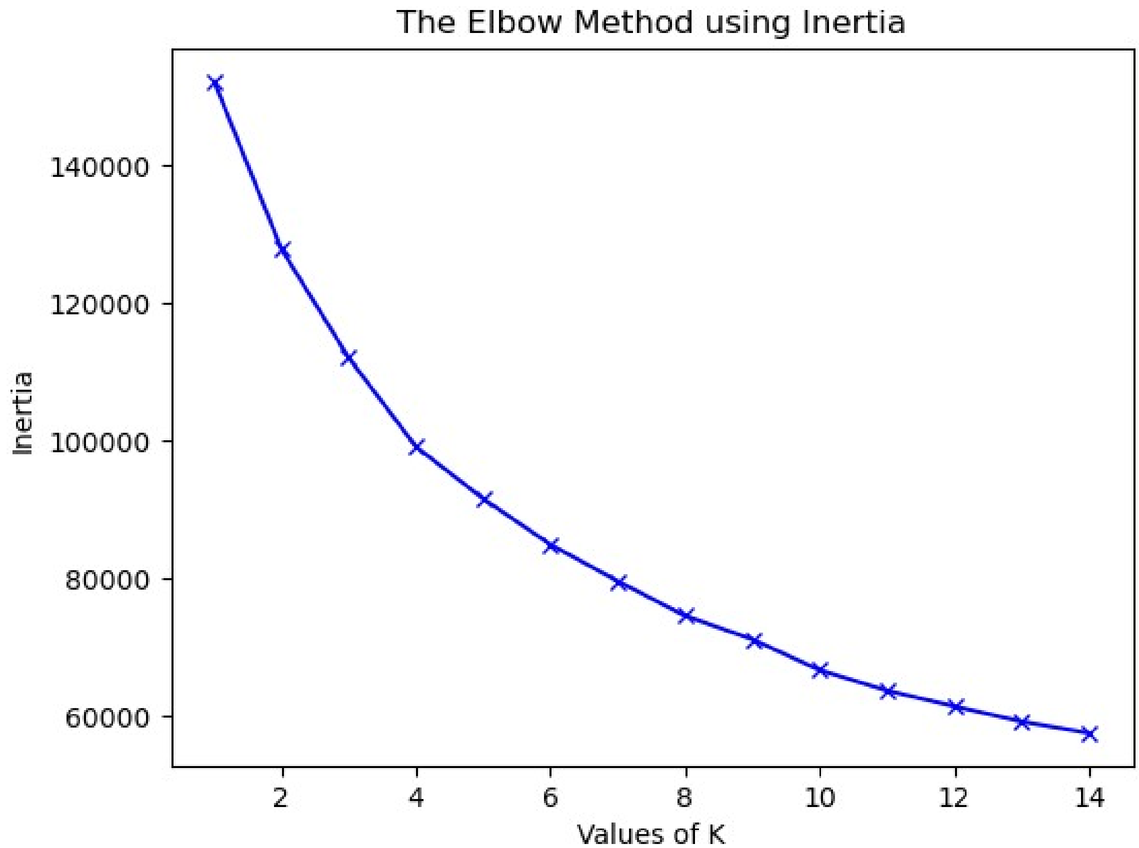
Data preprocessing

```
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
```

```
from sklearn.cluster import
KMeans,AgglomerativeClustering,DBSCAN,SpectralClustering
from sklearn.mixture import GaussianMixture
from sklearn.metrics import silhouette_samples, silhouette_score
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn import tree
from sklearn import metrics

scalar=StandardScaler()

scaled_df = scalar.fit_transform(df)

pca = PCA(n_components=2)
principal_components = pca.fit_transform(scaled_df)
pca_df =
pd.DataFrame(data=principal_components ,columns=["PCA1","PCA2"])
pca_df
```

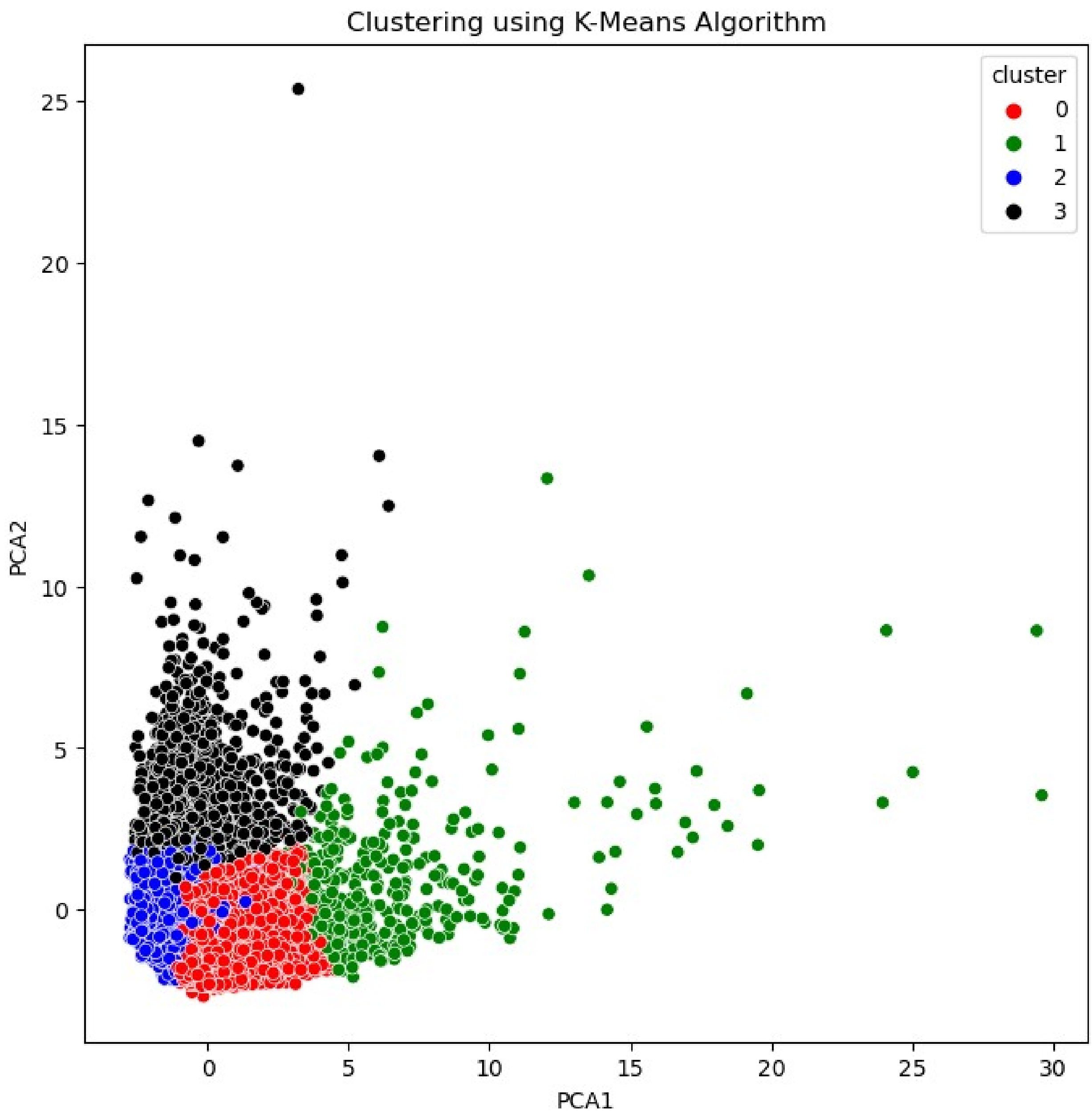
```
          PCA1      PCA2
0    -1.682221 -1.076442
1    -1.138303  2.506541
2     0.969687 -0.383541
3    -0.873625  0.043134
4    -1.599436 -0.688562
...        ...       ...
8945 -0.359628 -2.016154
8946 -0.564365 -1.639156
8947 -0.926202 -1.810796
8948 -2.336553 -0.657949
8949 -0.556424 -0.400455

[8950 rows x 2 columns]
```

```
inertia = []
range_val = range(1,15)
for i in range_val:
    kmean = KMeans(n_clusters=i)
    kmean.fit_predict(pd.DataFrame(scaled_df))
    inertia.append(kmean.inertia_)
plt.plot(range_val,inertia,'bx-')
plt.xlabel('Values of K')
plt.ylabel('Inertia')
plt.title('The Elbow Method using Inertia')
plt.show()
```



Model building

```
kmeans_model=KMeans(4)
kmeans_model.fit_predict(scaled_df)
pca_df_kmeans=
pd.concat([pca_df,pd.DataFrame({'cluster':kmeans_model.labels_})],axis
=1)

plt.figure(figsize=(8,8))
ax=sns.scatterplot(x="PCA1",y="PCA2",hue="cluster",data=pca_df_kmeans,
palette=['red','green','blue','black'])
plt.title("Clustering using K-Means Algorithm")
plt.show()
```




```
cluster_centers =
pd.DataFrame(data=kmeans_model.cluster_centers_,columns=[df.columns])

cluster_centers = scalar.inverse_transform(cluster_centers)
cluster_centers =
pd.DataFrame(data=cluster_centers,columns=[df.columns])
cluster_centers
```

	BALANCE	BALANCE_FREQUENCY	PURCHASES	ONEOFF_PURCHASES	\
0	897.456372	0.935846	1238.779285	595.851281	
1	3556.148177	0.986911	7680.432073	5099.738293	
2	1004.071622	0.788666	271.249733	209.972733	

```
3  4577.199684          0.968513   491.071544       311.745004

  INSTALLMENTS_PURCHASES CASH_ADVANCE PURCHASES_FREQUENCY  \
0             643.203117   211.260815            0.885991
1            2582.157195   686.680975            0.946548
2              61.544224   588.473350            0.171839
3             179.410495  4485.847516            0.284436

  ONEOFF_PURCHASES_FREQUENCY PURCHASES_INSTALLMENTS_FREQUENCY  \
0                   0.297970                         0.712244
1                   0.739667                         0.788577
2                   0.086325                         0.082144
3                   0.136892                         0.183111

  CASH_ADVANCE_FREQUENCY CASH_ADVANCE_TRX PURCHASES_TRX
CREDIT_LIMIT  \
0               0.042713         0.792610     22.133492  4217.363007

1               0.073149         2.170732     89.309756  9711.097561

2               0.113588         2.092122      2.928769  3274.061863

3               0.483006        14.215524      7.530966  7499.962465


      PAYMENTS MINIMUM_PAYMENTS PRC_FULL_PAYMENT     TENURE
0  1334.890105       651.454148         0.269643  11.597437
1  7301.419079      1977.054411         0.286211  11.951220
2   970.071857       584.198510         0.078333  11.445004
3  3455.052828      1995.669235         0.034898  11.385632
```

```
cluster_df =
pd.concat([df,pd.DataFrame({'Cluster':kmeans_model.labels_})],axis=1)
cluster_df
```

```
          BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
0       40.900749           0.818182      95.40              0.00
1     3202.467416           0.909091       0.00              0.00
2     2495.148862           1.000000     773.17            773.17
3     1666.670542           0.636364    1499.00           1499.00
4      817.714335           1.000000      16.00             16.00
...           ...                ...        ...               ...
8945    28.493517           1.000000     291.12              0.00
8946    19.183215           1.000000     300.00              0.00
8947    23.398673           0.833333     144.40              0.00
8948    13.457564           0.833333       0.00              0.00
8949   372.708075           0.666667    1093.25           1093.25

      INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
0                      95.40      0.000000             0.166667
1                       0.00   6442.945483             0.000000
```

```
2                        0.00      0.000000             1.000000
3                        0.00    205.788017             0.083333
4                        0.00      0.000000             0.083333
...                       ...           ...                  ...
8945                   291.12      0.000000             1.000000
8946                   300.00      0.000000             1.000000
8947                   144.40      0.000000             0.833333
8948                     0.00     36.558778             0.000000
8949                     0.00    127.040008             0.666667

      ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
0                       0.000000                          0.083333
1                       0.000000                          0.000000
2                       1.000000                          0.000000
3                       0.083333                          0.000000
4                       0.083333                          0.000000
...                          ...                               ...
8945                    0.000000                          0.833333
8946                    0.000000                          0.833333
8947                    0.000000                          0.666667
8948                    0.000000                          0.000000
8949                    0.666667                          0.000000

      CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
0                   0.000000                 0              2
1000.0
1                   0.250000                 4              0
7000.0
2                   0.000000                 0             12
7500.0
3                   0.083333                 1              1
7500.0
4                   0.000000                 0              1
1200.0
...                      ...               ...            ...
...
8945                0.000000                 0              6
1000.0
8946                0.000000                 0              6
1000.0
8947                0.000000                 0              5
1000.0
8948                0.166667                 2              0
500.0
8949                0.333333                 2             23
1200.0

         PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE  Cluster
```

```
0      201.802084     139.509787      0.000000    12      2
1     4103.032597    1072.340217      0.222222    12      3
2      622.066742     627.284787      0.000000    12      0
3        0.000000     864.206542      0.000000    12      2
4      678.334763     244.791237      0.000000    12      2
...           ...            ...           ...   ...    ...
8945   325.594462      48.886365      0.500000     6      0
8946   275.861322     864.206542      0.000000     6      0
8947    81.270775      82.418369      0.250000     6      0
8948    52.549959      55.755628      0.250000     6      2
8949    63.165404      88.288956      0.000000     6      2

[8950 rows x 18 columns]
```

```
cluster_1_df = cluster_df[cluster_df["Cluster"]==0]
cluster_1_df
```

```
          BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
2     2495.148862           1.000000     773.17            773.17
5     1809.828751           1.000000    1333.28              0.00
7     1823.652743           1.000000     436.20              0.00
10    1293.124939           1.000000     920.12              0.00
12    1516.928620           1.000000    3217.99           2500.23
...           ...                ...        ...               ...
8940   130.838554           1.000000     591.24              0.00
8942    40.829749           1.000000     113.28              0.00
8945    28.493517           1.000000     291.12              0.00
8946    19.183215           1.000000     300.00              0.00
8947    23.398673           0.833333     144.40              0.00

      INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
2                       0.00           0.0             1.000000
5                    1333.28           0.0             0.666667
7                     436.20           0.0             1.000000
10                    920.12           0.0             1.000000
12                    717.76           0.0             1.000000
...                      ...           ...                  ...
8940                  591.24           0.0             1.000000
8942                  113.28           0.0             1.000000
8945                  291.12           0.0             1.000000
```

```
8946                     300.00           0.0               1.000000
8947                     144.40           0.0               0.833333

      ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
2                           1.00                          0.000000
5                           0.00                          0.583333
7                           0.00                          1.000000
10                          0.00                          1.000000
12                          0.25                          0.916667
...                          ...                               ...
8940                        0.00                          0.833333
8942                        0.00                          0.833333
8945                        0.00                          0.833333
8946                        0.00                          0.833333
8947                        0.00                          0.666667

      CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
2                        0.0                 0             12
7500.0
5                        0.0                 0              8
1800.0
7                        0.0                 0             12
2300.0
10                       0.0                 0             12
1200.0
12                       0.0                 0             26
3000.0
...                      ...               ...            ...
...
8940                     0.0                 0              6
1000.0
8942                     0.0                 0              6
1000.0
8945                     0.0                 0              6
1000.0
8946                     0.0                 0              6
1000.0
8947                     0.0                 0              5
1000.0

         PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE  Cluster

2      622.066742        627.284787              0.00      12        0

5     1400.057770       2407.246035              0.00      12        0

7      679.065082        532.033990              0.00      12        0

10    1083.301007       2172.697765              0.00      12        0
```

```
12     608.263689        490.207013              0.25      12        0
...           ...               ...               ...     ...      ...
8940   475.523262         82.771320              1.00       6        0
8942    94.488828         86.283101              0.25       6        0
8945   325.594462         48.886365              0.50       6        0
8946   275.861322        864.206542              0.00       6        0
8947    81.270775         82.418369              0.25       6        0

[3356 rows x 18 columns]
```

```
cluster_2_df = cluster_df[cluster_df["Cluster"]==1]
cluster_2_df
```

```
          BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
6      627.260806           1.000000    7091.01           6402.63
21    6369.531318           1.000000    6359.95           5910.04
57    2386.330629           1.000000    5217.62           4789.09
84    1935.362486           1.000000    4915.60           4515.34
90    9381.255094           1.000000    5100.07           1147.83
...           ...                ...        ...               ...
8215  4436.557694           1.000000    6005.90           5838.38
8541  3326.323283           1.000000    8209.77           2218.28
8662   599.909949           1.000000    4947.32           3149.59
8689   368.318662           0.909091    8053.95           8053.95
8737  2533.618119           0.909091    5633.83           2985.92

      INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
6                     688.38      0.000000             1.000000
21                    449.91    229.028245             1.000000
57                    428.53      0.000000             0.916667
84                    400.26    293.844792             1.000000
90                   3952.24    370.737197             1.000000
...                      ...           ...                  ...
8215                  167.52    567.971877             1.000000
8541                 5991.49      0.000000             1.000000
8662                 1797.73      0.000000             1.000000
8689                    0.00      0.000000             0.833333
8737                 2647.91   2451.807788             0.916667

      ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
6                       1.000000                          1.000000
21                      0.916667                          1.000000
57                      0.916667                          0.500000
```

```
84                        1.000000                                  0.333333
90                        0.250000                                  0.916667
...                            ...                                       ...
8215                      0.583333                                  0.916667
8541                      0.416667                                  1.000000
8662                      1.000000                                  0.916667
8689                      0.833333                                  0.000000
8737                      0.500000                                  0.750000

      CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
6                   0.000000                 0             64
13500.0
21                  0.333333                 6             92
11250.0
57                  0.000000                 0             42
7500.0
84                  0.083333                 1             50
9000.0
90                  0.083333                 1             46
9000.0
...                      ...               ...            ...
...
8215                0.083333                 1             61
10500.0
8541                0.000000                 0            130
10000.0
8662                0.000000                 0             73
3000.0
8689                0.000000                 0             46
2000.0
8737                0.333333                16             82
9000.0

         PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE  Cluster

6     6354.314328        198.065894          1.000000      12        1

21    2077.959051       1659.775075          0.000000      12        1

57    5678.729613       1311.514878          0.083333      12        1

84    4921.066897        594.756686          0.000000      12        1

90    6409.496345       9827.045323          0.000000      12        1

...           ...               ...               ...     ...      ...

8215  1650.425296       1067.515656          0.000000      12        1
```

```
8541  1942.074765        702.905059          0.083333       12          1

8662  5024.430008        218.172915          0.083333       12          1

8689  7966.582037        219.761189          0.777778       12          1

8737  8176.953944        602.963244          0.000000       12          1


[410 rows x 18 columns]
```

```
cluster_3_df = cluster_df[cluster_df["Cluster"]==2]
cluster_3_df
```

```
          BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
0       40.900749           0.818182      95.40              0.00
3     1666.670542           0.636364    1499.00           1499.00
4      817.714335           1.000000      16.00             16.00
8     1014.926473           1.000000     861.49            661.49
9      152.225975           0.545455    1281.60           1281.60
...           ...                ...        ...               ...
8939   728.352548           1.000000     734.40            734.40
8943     5.871712           0.500000      20.90             20.90
8944   193.571722           0.833333    1012.73           1012.73
8948    13.457564           0.833333       0.00              0.00
8949   372.708075           0.666667    1093.25           1093.25

      INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
0                       95.4      0.000000             0.166667
3                        0.0    205.788017             0.083333
4                        0.0      0.000000             0.083333
8                      200.0      0.000000             0.333333
9                        0.0      0.000000             0.166667
...                      ...           ...                  ...
8939                     0.0    239.891038             0.333333
8943                     0.0      0.000000             0.166667
8944                     0.0      0.000000             0.333333
8948                     0.0     36.558778             0.000000
8949                     0.0    127.040008             0.666667

      ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
0                       0.000000                          0.083333
3                       0.083333                          0.000000
4                       0.083333                          0.000000
8                       0.083333                          0.250000
9                       0.166667                          0.000000
...                          ...                               ...
8939                    0.333333                          0.000000
8943                    0.166667                          0.000000
8944                    0.333333                          0.000000
```

```
8948                        0.000000                                  0.000000
8949                        0.666667                                  0.000000

      CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
0                   0.000000                 0              2
1000.0
3                   0.083333                 1              1
7500.0
4                   0.000000                 0              1
1200.0
8                   0.000000                 0              5
7000.0
9                   0.000000                 0              3
11000.0
...                      ...               ...            ...
...
8939                0.166667                 2              2
1000.0
8943                0.000000                 0              1
500.0
8944                0.000000                 0              2
4000.0
8948                0.166667                 2              0
500.0
8949                0.333333                 2             23
1200.0

         PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE  Cluster

0      201.802084        139.509787              0.00      12        2

3        0.000000        864.206542              0.00      12        2

4      678.334763        244.791237              0.00      12        2

8      688.278568        311.963409              0.00      12        2

9     1164.770591        100.302262              0.00      12        2

...           ...               ...               ...     ...      ...

8939    72.530037        110.950798              0.00       6        2

8943    58.644883         43.473717              0.00       6        2

8944     0.000000        864.206542              0.00       6        2

8948    52.549959         55.755628              0.25       6        2
```

```
8949    63.165404         88.288956              0.00        6         2


[3973 rows x 18 columns]
```

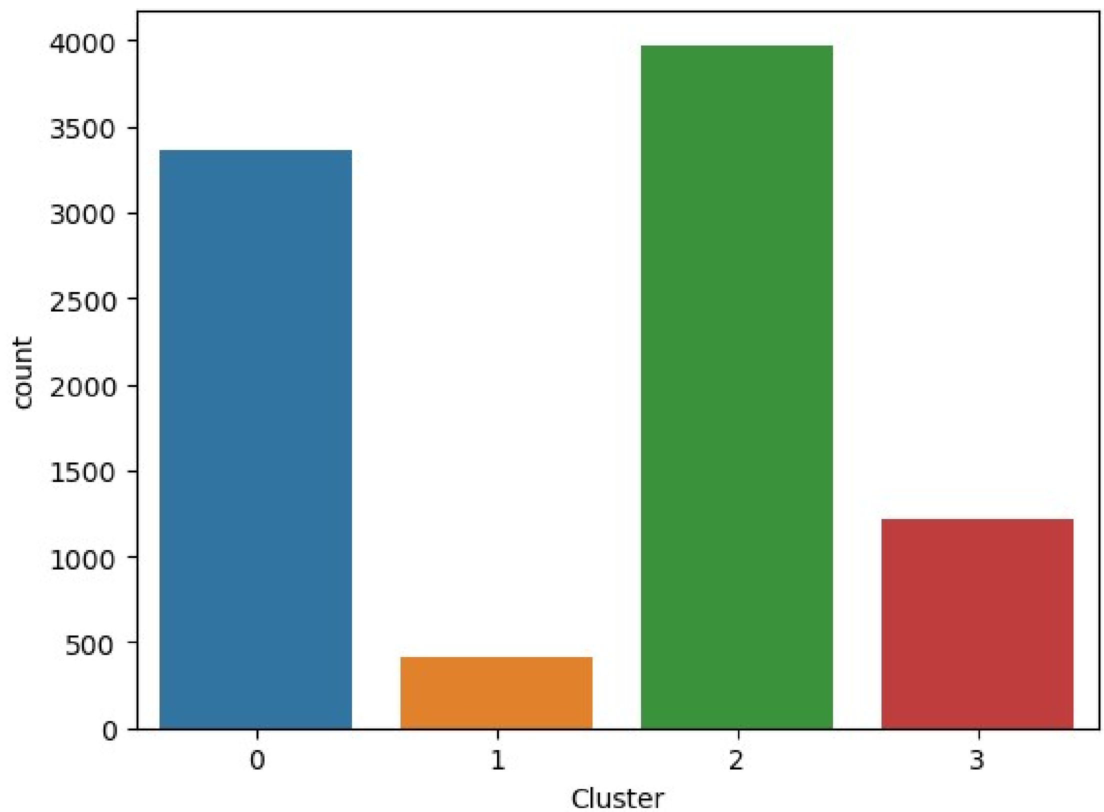
```
cluster_4_df = cluster_df[cluster_df["Cluster"] == 3]
cluster_4_df
```

```
          BALANCE  BALANCE_FREQUENCY  PURCHASES  ONEOFF_PURCHASES  \
1     3202.467416           0.909091       0.00              0.00
15    6886.213231           1.000000    1611.70              0.00
23    3800.151377           0.818182    4248.35           3454.56
24    5368.571219           1.000000       0.00              0.00
28    7152.864372           1.000000     387.05            204.55
...           ...                ...        ...               ...
8857  2330.222764           1.000000    1320.00              0.00
8858   812.934042           1.000000      50.00             50.00
8869  2171.222526           1.000000     791.18            791.18
8915   381.341657           1.000000      78.00              0.00
8941  5967.475270           0.833333     214.55              0.00

      INSTALLMENTS_PURCHASES  CASH_ADVANCE  PURCHASES_FREQUENCY  \
1                       0.00   6442.945483             0.000000
15                   1611.70   2301.491267             0.500000
23                    793.79   7974.415626             1.000000
24                      0.00    798.949863             0.000000
28                    182.50   2236.145259             0.666667
...                      ...           ...                  ...
8857                 1320.00  14926.790590             0.428571
8858                    0.00   2185.500596             0.142857
8869                    0.00   2056.602480             0.428571
8915                   78.00    934.808869             1.000000
8941                  214.55   8555.409326             0.833333

      ONEOFF_PURCHASES_FREQUENCY  PURCHASES_INSTALLMENTS_FREQUENCY  \
1                       0.000000                          0.000000
15                      0.000000                          0.500000
23                      0.083333                          0.916667
24                      0.000000                          0.000000
28                      0.166667                          0.416667
...                          ...                               ...
8857                    0.000000                          0.285714
8858                    0.142857                          0.000000
8869                    0.428571                          0.000000
8915                    0.000000                          0.833333
8941                    0.000000                          0.666667

      CASH_ADVANCE_FREQUENCY  CASH_ADVANCE_TRX  PURCHASES_TRX
CREDIT_LIMIT  \
1                   0.250000                 4              0
```

```
7000.0
15                   0.166667                  4              11
8000.0
23                   0.333333                 13              13
9000.0
24                   0.363636                  4               0
6000.0
28                   0.833333                 16               8
10500.0
...                       ...                ...             ...
...
8857                 0.571429                 10               3
10000.0
8858                 1.000000                 16               1
3000.0
8869                 0.571429                  6               8
3000.0
8915                 0.666667                 16               6
1000.0
8941                 0.666667                 13               5
9000.0

         PAYMENTS  MINIMUM_PAYMENTS  PRC_FULL_PAYMENT  TENURE  Cluster

1     4103.032597       1072.340217          0.222222      12        3

15    1993.439277       2109.906490          0.000000      12        3

23    9479.043842       1425.426525          0.000000      12        3

24    1422.726707       1657.002877          0.000000      11        3

28    1601.448347       1648.851345          0.000000      12        3

...           ...               ...               ...     ...      ...

8857  8157.666434        283.362434          0.200000       7        3

8858   726.683966        127.843735          0.000000       7        3

8869   300.088696        453.100425          0.000000       7        3

8915   143.118373         85.152441          0.000000       6        3

8941   966.202912        861.949906          0.000000       6        3


[1211 rows x 18 columns]
```

```python
sns.countplot(x='Cluster', data=cluster_df)
```

```
<AxesSubplot:xlabel='Cluster', ylabel='count'>
```



Result

```
X = cluster_df.drop(['Cluster'],axis=1)
y= cluster_df[['Cluster']]
X_train, X_test, y_train, y_test =train_test_split(X, y,
test_size=0.3,random_state=3)

model= DecisionTreeClassifier(criterion="entropy")
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

print(metrics.confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))

[[ 973    9   28   11]
 [  10  107    3    1]
 [  35    2 1102   25]
 [  11    0   31  337]]
              precision    recall  f1-score   support

           0       0.95      0.95      0.95      1021
```

```
           1       0.91      0.88      0.90       121
           2       0.95      0.95      0.95      1164
           3       0.90      0.89      0.90       379

    accuracy                           0.94      2685
   macro avg       0.93      0.92      0.92      2685
weighted avg       0.94      0.94      0.94      2685
```

Conclusion

An accuracy of 94% in the context of customer segmentation suggests that the clustering model has been highly effective in distinguishing between different customer behaviors and preferences based on the provided dataset. This high level of accuracy indicates a strong correspondence between the model's segmentation and the underlying patterns in the credit card usage behavior of about 9,000 active cardholders over the last 6 months. Such a result is highly promising for targeted marketing strategies and personalized financial product offerings. Here's a wider perspective on the conclusion and its implications:

Detailed Insights and Strategic Applications Targeted Marketing Efficiency: With distinct customer segments identified so accurately, marketing campaigns can be designed to specifically address the needs, preferences, and behaviors of each segment. This precision in targeting is likely to significantly increase the effectiveness of marketing efforts, reduce costs associated with broad-spectrum campaigns, and improve customer response rates.

Personalized Product Offerings: Financial institutions can tailor their product offerings to meet the unique needs of each segment. For example, customers who frequently make installment purchases might be offered special installment plans with lower interest rates or rewards for installment purchases. Similarly, those using their credit card primarily for cash advances could be targeted with products that offer lower cash advance fees or better terms for short-term liquidity.

Enhanced Customer Experience and Loyalty: By understanding the distinct needs and preferences of each segment, financial institutions can not only offer more relevant products but also communicate in a way that resonates with each group. This personalized approach is likely to enhance customer satisfaction, foster loyalty, and reduce churn rates. Customers are more likely to stay with a bank that understands their needs and offers solutions that are tailored to them.

Strategic Decision Making: The insights gained from this segmentation can inform strategic decisions at higher levels of management. Understanding the different segments can help in allocating resources more efficiently, developing new products, and even entering new markets. For instance, if a significant segment of customers shows a preference for digital banking solutions, the institution might prioritize digital transformation initiatives.

Competitive Advantage: In a highly competitive financial services market, the ability to accurately segment and target customers can provide a significant competitive edge. It allows an institution to differentiate itself by offering personalized experiences and products that competitors may not.

Future Opportunities: Beyond immediate marketing and product development implications, this segmentation can be used to predict future trends and customer needs. As customer behavior evolves, the model can be updated and refined to maintain high levels of accuracy and relevance, ensuring that the financial institution remains ahead of the curve in understanding and meeting customer demands.

Achieving a 94% accuracy rate in customer segmentation is a remarkable achievement that opens up numerous opportunities for targeted marketing, personalized product offerings, and strategic decision-making. By leveraging these insights, financial institutions can enhance customer experiences, improve loyalty, and gain a competitive advantage in the market. This approach not only maximizes the effectiveness of marketing and product development efforts but also paves the way for anticipating and adapting to future customer needs and market trends.