03_customer_seg_bankdata_report

December 22, 2022

1 Customer Segmentation Report

2 Import Data

```
    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
```

3 Load Model Artifacts

4 Transform Dataset

```
   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  cluster
0  201.802084        139.509787          0.000000      12        1
1  4103.032597       1072.340217          0.222222      12        0
2  622.066742        627.284787          0.000000      12        0
3    0.000000               NaN          0.000000      12        0
4  678.334763        244.791237          0.000000      12        1
```
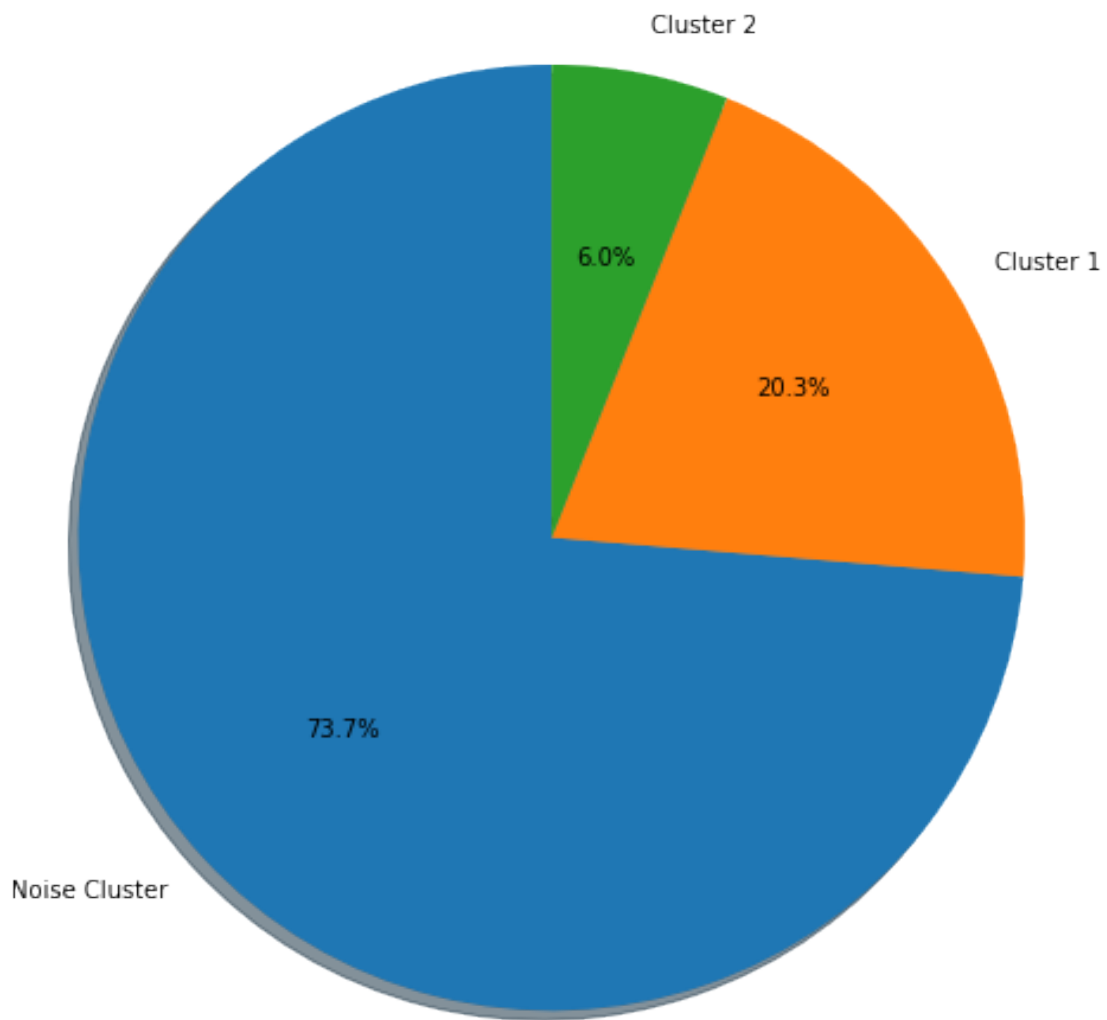
5 Data Visualization

```
0    6594
1    1816
2     540
Name: cluster, dtype: int64
```

5.1 What does the Cluster Distribution Lool Like?

```
['Noise Cluster', 'Cluster 1', 'Cluster 2']
```

6 Determine Cluster Centers

cluster	BALANCE	BALANCE_FREQUENCY	PURCHASES	ONEOFF_PURCHASES
0	937.806150	1.0	582.895	178.815
1	958.125169	1.0	0.000	0.000
2	73.222225	1.0	431.965	0.000

cluster	INSTALLMENTS_PURCHASES	CASH_ADVANCE	PURCHASES_FREQUENCY
0	163.48	0.000000	0.583333
1	0.00	96.134078	0.000000
2	419.94	0.000000	1.000000

cluster	ONEOFF_PURCHASES_FREQUENCY	PURCHASES_INSTALLMENTS_FREQUENCY
0	0.1	0.333333
1	0.0	0.000000
2	0.0	1.000000

cluster	CASH_ADVANCE_FREQUENCY	CASH_ADVANCE_TRX	PURCHASES_TRX
0	0.000000	0.0	9.0
1	0.083333	1.0	0.0
2	0.000000	0.0	12.0

| cluster	CREDIT_LIMIT	PAYMENTS	MINIMUM_PAYMENTS	PRC_FULL_PAYMENT	TENURE
0	4000.0	1232.78107	357.086194	0.000000	12.0
1	1950.0	442.85716	315.054229	0.000000	12.0
2	2000.0	408.75430	174.121900	0.166667	12.0

7 Explain Clusters

8 Custom Functions

8.1 Average Feature Importance on Cluster Building
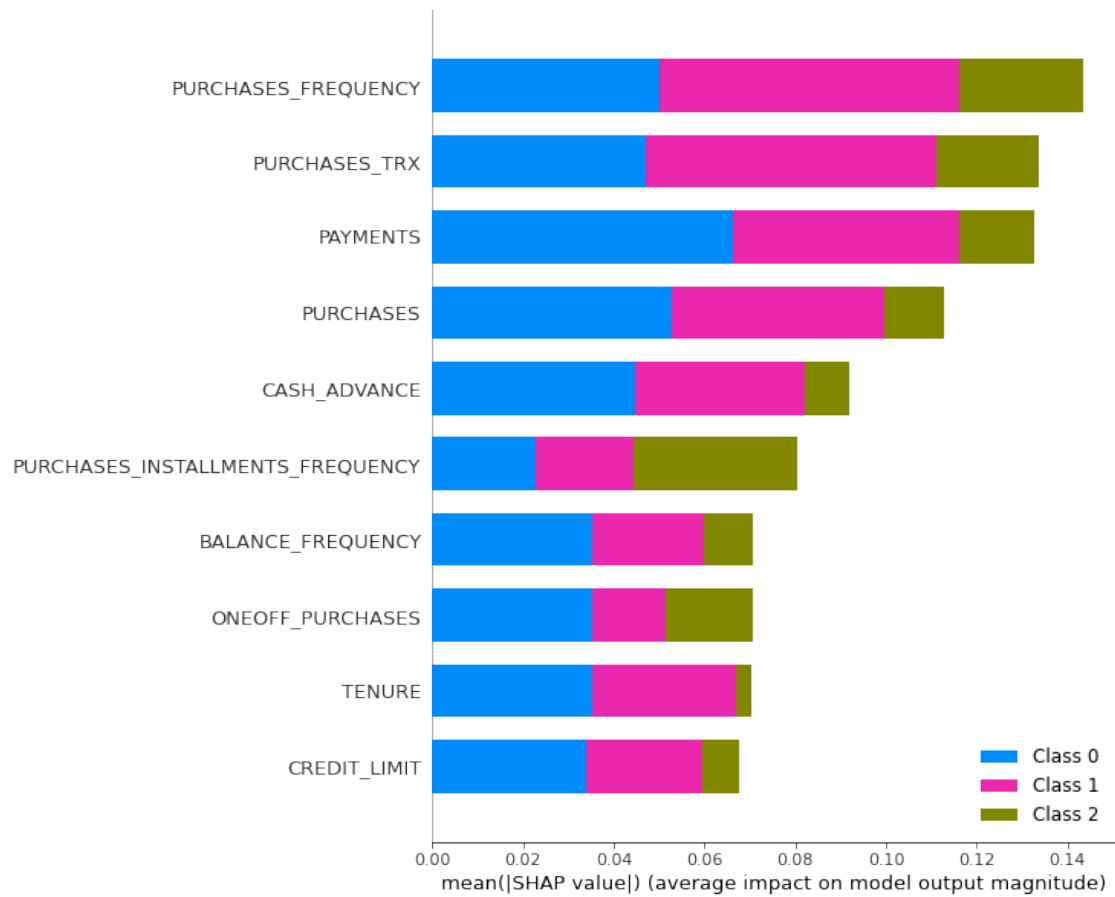
8.2 Explain Single Cluster vs Rest

Shap Feature Importance for Cluster 0

Shap Feature Importance for Cluster 1

Shap Feature Importance for Cluster 2
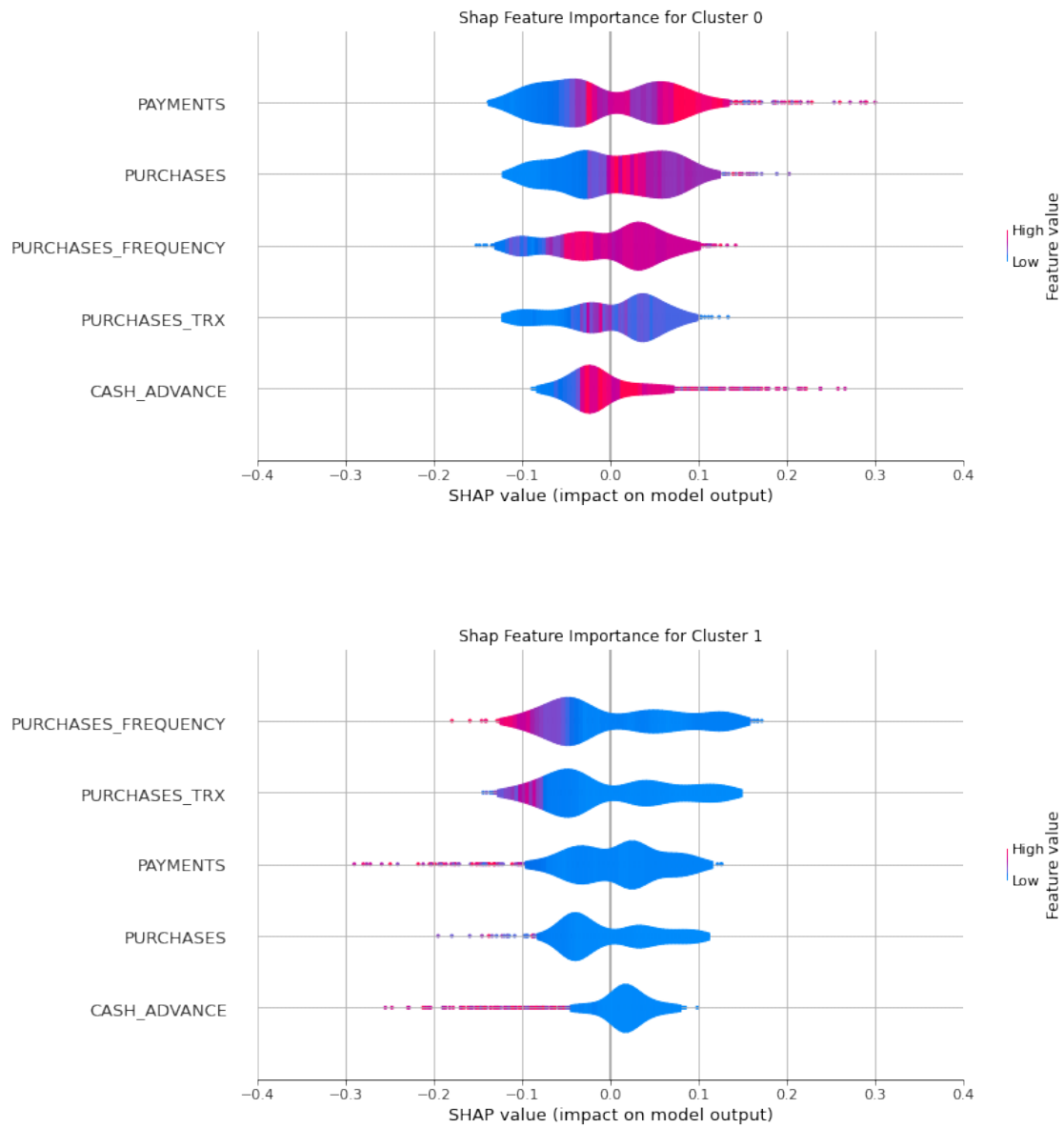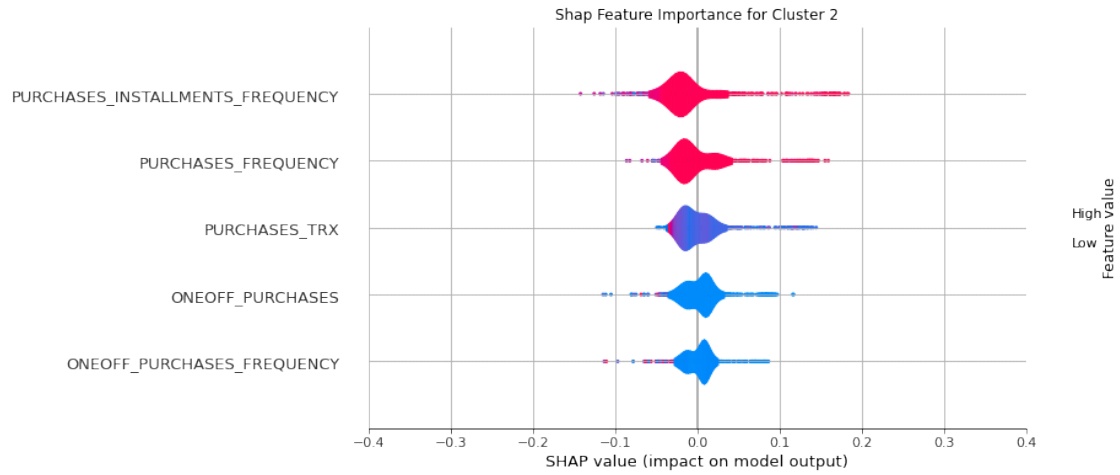
```
['ONEOFF_PURCHASES',
 'CASH_ADVANCE',
 'PAYMENTS',
 'PURCHASES_TRX',
 'PURCHASES_INSTALLMENTS_FREQUENCY',
 'PURCHASES',
 'PURCHASES_FREQUENCY',
 'ONEOFF_PURCHASES_FREQUENCY']
```