

Segmentation of credit card customers

Objective Defining the segments of customers to help make business strategy for marketing.

Business Problem

- Derive KPIs
- Get insights on customer profile
- Clustering
- Profiling
- Strategic insights for marketing

Data Availability

- Customer level credit usage behavioral features were given for the time period of 6 months for 9000 customers.

Approach:

1 Importing the dataset

2 Data Audits

- check for duplicate rows
- check for missing values
- check for outliers

3 Preliminary data preparation

- Missing Imputation
The missing values in the data were imputed with median value of the respective features.
- Outlier capping
The Outliers were present in the columns like purchases, payments which were then capped with the 1% and 99% percentile values

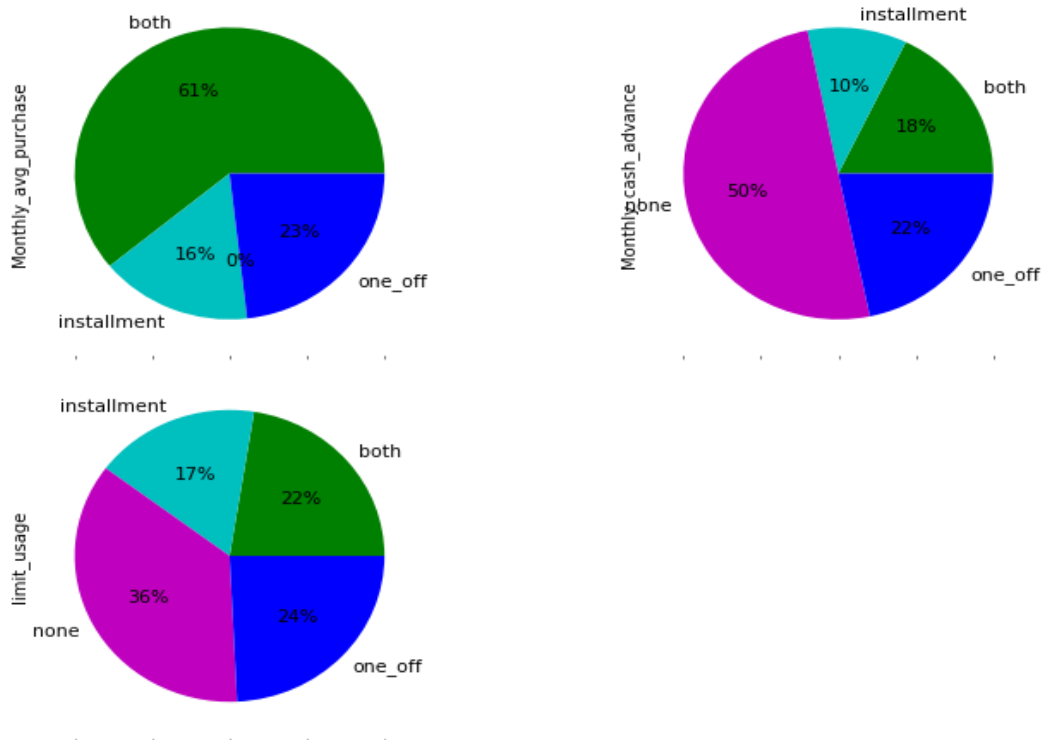
4 Identifying the key performance indicators

Based on the customer data provided some of the KPIs were derived,

- Monthly average purchase
 $\text{Monthly_avg_purchase} = \text{PURCHASES} / \text{TENURE}$
- Monthly cash advance
 $\text{Monthly_cash_advance} = \text{CASH_ADVANCE} / \text{TENURE}$
- Average amount per purchase transaction
PURCHASES_TRX - given in the dataset
- Average amount per cash-advance transaction
CASH_ADVANCE_TRX - given in the dataset
- Limit of usage
 $\text{limit_usage} = \text{BALANCE} / \text{CREDIT_LIMIT}$

- Purchase types
None, One-off, Installment, Both

Insights derived from observing and visualising the above KPIs,



From the above plots the following insights are found,

- Monthly average purchase
High → both, Low → none
- Monthly cash advance
High → none, Low → installment
- Limit of usage
High → none, Low → installment

5 Feature Reduction

- PCA (Principal Component Analysis)

By PCA, 8 PCs were able to explain 80% of the variance of all Xs. Factor loading matrix was created with those 8 PCs and the variables having high correlation with each PCs were selected.

```
selected_vars_pca = ['BALANCE', 'BALANCE_FREQUENCY', 'CASH_ADVANCE', 'CASH_ADVANCE_FREQUENCY',
'CASH_ADVANCE_TRX', 'CREDIT_LIMIT', 'INSTALLMENTS_PURCHASES', 'MINIMUM_PAYMENTS']
```

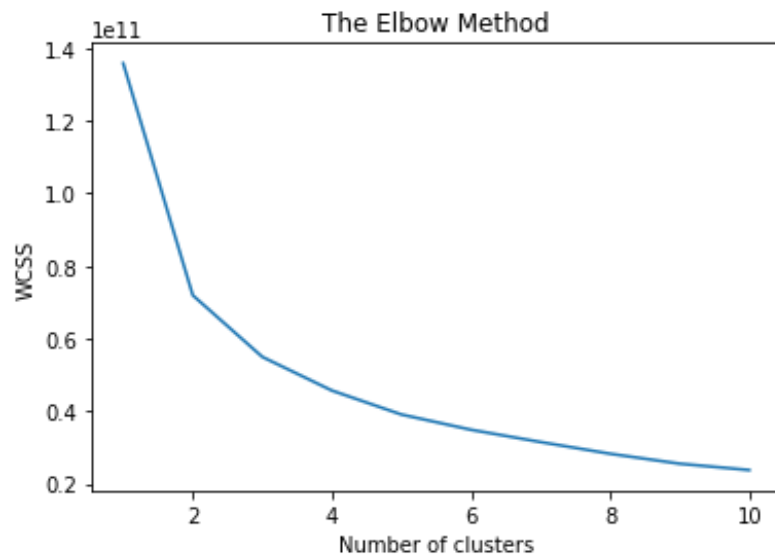
The above features were able to explain the maximum variance in the data.

6 Finding Optimum #Clusters

- Elbow method (wcscs)

Within Cluster Sum of Squares – minimum value of wcscs will result in maximizing the distance between the clusters.

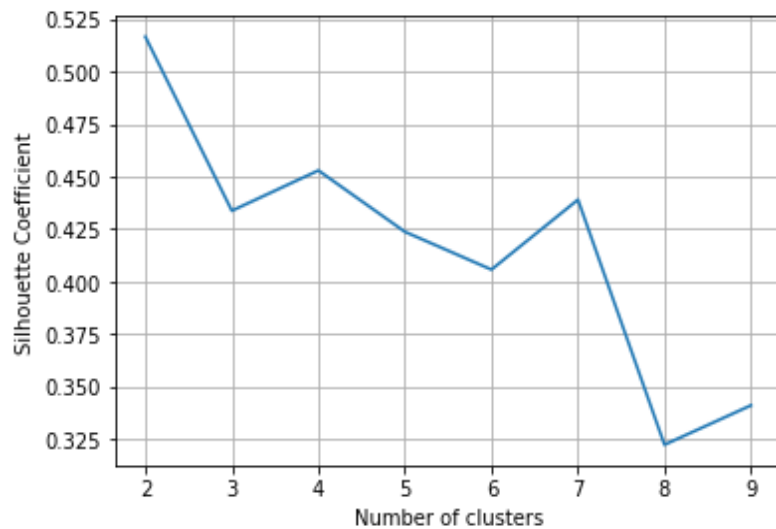
The Elbow plot was created for the range of 1 to 10 number of Clusters,



From the above plot it is found that 3 cluster solution was the Optimum as the graph showing a sudden decrease around the 3 cluster point.

- Silhouette analysis

The Silhouette score was calculated for the range of 2 to 10 number of clusters,



The Silhouette value is high for the #3, #4 and #7 cases.

From the above two methods,

It is found that the data can be well segmented if the number of clusters is either 3 or 4.

7 Model Implementation

- K Means clustering

The K Means was applied on the data for 3, 4 and 7 number of Clusters as identified from the above plots.

8 Profiling

The Profiling was done for the 3, 4 and 7 Cluster models outputs

The clusters were grouped and the mean of the other features were taken to create the profile out.

The values for each feature less than 0.8 times their overall mean was highlighted by red,

The values for each feature greater than 1.2 times their overall mean was highlighted by green.

	Overall	KM3_1	KM3_2	KM3_3	KM4_1	KM4_2	KM4_3	KM4_4	KM7_1	KM7_2	KM7_3	KM7_4	KM7_5	KM7_6	KM7_7
Seg_size	6265.0	4145.0	1146.0	974.0	448.0	3879.0	977.0	961.0	108.0	941.0	2974.0	944.0	361.0	498.0	439.0
Seg_Pct	1.0	0.7	0.2	0.2	0.1	0.6	0.2	0.2	0.0	0.2	0.5	0.2	0.1	0.1	0.1
BALANCE	1532.1	1110.1	4253.7	126.0	4541.1	988.8	3698.4	120.3	3916.7	117.8	821.3	1762.0	4885.9	1523.8	5550.9
BALANCE_FREQUENCY	0.9	1.0	1.0	0.4	1.0	1.0	1.0	0.4	1.0	0.4	1.0	1.0	1.0	1.0	1.0
CASH_ADVANCE	919.2	322.3	3574.7	335.0	795.4	327.8	3918.3	315.2	819.4	300.3	182.8	1905.0	5898.1	261.1	1791.4
CASH_ADVANCE_FREQUENCY	0.1	0.1	0.4	0.0	0.1	0.1	0.5	0.0	0.1	0.0	0.0	0.4	0.6	0.0	0.2
CASH_ADVANCE_TRX	3.0	1.2	11.4	0.8	2.0	1.2	12.9	0.7	2.9	0.6	0.6	8.3	18.0	0.7	3.4
CREDIT_LIMIT	4477.8	3911.6	7225.2	3655.0	8556.5	3688.6	6532.7	3672.8	4852.3	3641.2	3373.3	3313.8	8438.9	7399.7	9592.5
INSTALLMENTS_PURCHASES	385.9	445.2	366.1	156.8	1581.0	354.8	180.3	163.3	835.2	148.1	272.5	111.6	290.3	2059.2	323.8
MINIMUM_PAYMENTS	748.0	578.4	1843.5	180.5	3348.5	454.4	1278.5	181.5	7820.3	179.6	411.7	662.5	1763.7	771.6	1826.7
ONEOFF_PURCHASES	549.6	619.9	573.5	222.4	1519.8	563.7	370.2	223.0	787.5	221.9	523.6	248.7	536.5	1504.8	944.1
ONEOFF_PURCHASES_FREQUENCY	0.2	0.2	0.2	0.1	0.4	0.2	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.4	0.3
PAYMENTS	1645.0	1349.1	3182.2	1095.8	3747.2	1204.3	2963.0	1104.1	2722.5	1076.2	1026.2	1538.5	4482.9	3506.8	2575.3
PRC_FULL_PAYMENT	0.2	0.2	0.0	0.2	0.1	0.2	0.0	0.2	0.0	0.2	0.2	0.0	0.0	0.3	0.0
PURCHASES	953.2	1082.9	972.5	378.5	3275.6	926.5	552.5	385.6	1858.3	369.3	802.3	362.1	837.6	3651.2	1309.3
PURCHASES_FREQUENCY	0.5	0.6	0.4	0.3	0.8	0.6	0.3	0.3	0.5	0.3	0.6	0.3	0.4	0.9	0.5
PURCHASES_INSTALLMENTS_FREQUENCY	0.4	0.4	0.3	0.2	0.7	0.4	0.2	0.2	0.5	0.2	0.4	0.2	0.2	0.9	0.3
PURCHASES_TRX	14.4	16.6	14.1	5.0	42.3	14.9	8.6	5.1	24.3	5.0	13.5	6.1	12.3	48.8	18.0
TENURE	11.5	11.6	11.5	11.3	11.9	11.6	11.3	11.3	11.8	11.3	11.6	11.1	11.4	11.9	11.8
Monthly_avg_purchase	82.0	92.8	84.2	33.7	275.6	79.9	49.0	34.1	161.8	32.8	69.5	33.2	73.1	307.0	109.6
Monthly_cash_advance	83.8	29.5	324.2	32.1	68.6	30.1	357.5	29.3	76.9	28.1	16.5	179.6	532.4	21.9	156.3
limit_usage	0.4	0.4	0.6	0.0	0.6	0.4	0.6	0.0	0.9	0.0	0.4	0.6	0.6	0.3	0.6
installment	0.3	0.3	0.1	0.4	0.2	0.3	0.1	0.4	0.3	0.4	0.3	0.1	0.1	0.2	0.1
none	0.2	0.2	0.4	0.2	0.1	0.2	0.5	0.2	0.3	0.2	0.1	0.5	0.4	0.0	0.3
one_off	0.2	0.2	0.2	0.3	0.1	0.2	0.2	0.3	0.2	0.3	0.2	0.2	0.2	0.0	0.2
Balance		Med	High	Low	High	Low	High	Low	High	Low	Low	Med	High	Med	High
Credit Limit		Med	High	Med	High	Med	High	Med	Med	Med	Low	Low	High	High	High
Installment Purchase		High	Med	Low	High	Med	Low	Low	High	Low	Low	Low	Low	High	Med
Purchase / Installment Frequency		Med	Med	Low	High	Med	Low	Low	High	Low	Med	Low	Low	High	Med
Monthly_avg_purchase		Med	Med	Low	High	Med	Low	Low	High	Low	Med	Low	Med	High	Med
Monthly_cash_advance		Low	High	Low	Med	Low	High	Low	Med	Low	Low	High	High	Low	High
limit_usage		Med	High	Low	High	Med	High	Low	High	Low	Med	High	High	Low	High

9 Insights and marketing strategy

3 #Clusters

	KM3_1	KM3_2	KM3_3
Seg_size	4145.0	1146.0	974.0
Seg_Pct	0.7	0.2	0.2
BALANCE	1110.1	4253.7	126.0
BALANCE_FREQUENCY	1.0	1.0	0.4
CASH_ADVANCE	322.3	3574.7	335.0
CASH_ADVANCE_FREQUENCY	0.1	0.4	0.0
CASH_ADVANCE_TRX	1.2	11.4	0.8
CREDIT_LIMIT	3911.6	7225.2	3655.0
INSTALLMENTS_PURCHASES	445.2	366.1	156.8
MINIMUM_PAYMENTS	578.4	1843.5	180.5
ONEOFF_PURCHASES	619.9	573.5	222.4
ONEOFF_PURCHASES_FREQUENCY	0.2	0.2	0.1
PAYMENTS	1349.1	3182.2	1095.8
PRC_FULL_PAYMENT	0.2	0.0	0.2
PURCHASES	1082.9	972.5	378.5
PURCHASES_FREQUENCY	0.6	0.4	0.3
PURCHASES_INSTALLMENTS_FREQUENCY	0.4	0.3	0.2
PURCHASES_TRX	16.6	14.1	5.0
TENURE	11.6	11.5	11.3
Monthly_avg_purchase	92.8	84.2	33.7
Monthly_cash_advance	29.5	324.2	32.1
limit_usage	0.4	0.6	0.0
installment	0.3	0.1	0.4
none	0.2	0.4	0.2
one_off	0.2	0.2	0.3
Balance	Med	High	Low
Credit Limit	Med	High	Med
Installment Purchase	High	Med	Low
Purchase / Installment Frequency	Med	Med	Low
Monthly_avg_purchase	Med	Med	Low
Monthly_cash_advance	Low	High	Low
limit_usage	Med	High	Low

Target - marketing strategy required

Good - may be different schemes increase purchase even more

Risky - less marketing required

Insights and type of marketing strategy

Cluster 1: Spender - Good

Cluster 2: Passive - Target

Cluster 3: Spends more with low balance – Good

4 #Clusters

	KM4_1	KM4_2	KM4_3	KM4_4
Seg_size	448.0	3879.0	977.0	961.0
Seg_Pct	0.1	0.6	0.2	0.2
BALANCE	4541.1	988.8	3698.4	120.3
BALANCE_FREQUENCY	1.0	1.0	1.0	0.4
CASH_ADVANCE	795.4	327.8	3918.3	315.2
CASH_ADVANCE_FREQUENCY	0.1	0.1	0.5	0.0
CASH_ADVANCE_TRX	2.0	1.2	12.9	0.7
CREDIT_LIMIT	8556.5	3688.6	6532.7	3672.8
INSTALLMENTS_PURCHASES	1581.0	354.8	180.3	163.3
MINIMUM_PAYMENTS	3348.5	454.4	1278.5	181.5
ONEOFF_PURCHASES	1519.8	563.7	370.2	223.0
ONEOFF_PURCHASES_FREQUENCY	0.4	0.2	0.2	0.1
PAYMENTS	3747.2	1204.3	2963.0	1104.1
PRC_FULL_PAYMENT	0.1	0.2	0.0	0.2
PURCHASES	3275.6	926.5	552.5	385.6
PURCHASES_FREQUENCY	0.8	0.6	0.3	0.3
PURCHASES_INSTALLMENTS_FREQUENCY	0.7	0.4	0.2	0.2
PURCHASES_TRX	42.3	14.9	8.6	5.1
TENURE	11.9	11.6	11.3	11.3
Monthly_avg_purchase	275.6	79.9	49.0	34.1
Monthly_cash_advance	68.6	30.1	357.5	29.3
limit_usage	0.6	0.4	0.6	0.0
installment	0.2	0.3	0.1	0.4
none	0.1	0.2	0.5	0.2
one_off	0.1	0.2	0.2	0.3
Balance	High	Low	High	Low
Credit Limit	High	Med	High	Med
Installment Purchase	High	Med	Low	Low
Purchase / Installment Frequency	High	Med	Low	Low
Monthly_avg_purchase	High	Med	Low	Low
Monthly_cash_advance	Med	Low	High	Low
limit_usage	High	Med	High	Low

Target - marketing strategy required

Good - may be different schemes increase purchase even more

Risky - less marketing required

Insights and type of marketing strategy

Cluster 1: Spender - Good

Cluster 2: Spender - Good

Cluster 3: Passive - Target

Cluster 4: Retiring - Risky

7 #Clusters

	KM7_1	KM7_2	KM7_3	KM7_4	KM7_5	KM7_6	KM7_7
Seg_size	108.0	941.0	2974.0	944.0	361.0	498.0	439.0
Seg_Pct	0.0	0.2	0.5	0.2	0.1	0.1	0.1
BALANCE	3916.7	117.8	821.3	1762.0	4885.9	1523.8	5550.3
BALANCE_FREQUENCY	1.0	0.4	1.0	1.0	1.0	1.0	1.0
CASH_ADVANCE	819.4	300.3	182.8	1905.0	5898.1	261.1	1791.4
CASH_ADVANCE_FREQUENCY	0.1	0.0	0.0	0.4	0.6	0.0	0.2
CASH_ADVANCE_TRX	2.9	0.6	0.6	8.3	18.0	0.7	3.4
CREDIT_LIMIT	4852.3	3641.2	3373.3	3313.8	8438.9	7399.7	9592.5
INSTALLMENTS_PURCHASES	835.2	148.1	272.5	111.6	290.3	2059.2	323.8
MINIMUM_PAYMENTS	7820.3	179.6	411.7	662.5	1763.7	771.6	1826.7
ONEOFF_PURCHASES	787.5	221.9	523.6	248.7	536.5	1504.8	944.1
ONEOFF_PURCHASES_FREQUENCY	0.1	0.1	0.2	0.1	0.2	0.4	0.3
PAYMENTS	2722.5	1076.2	1026.2	1538.5	4482.9	3506.8	2575.3
PRC_FULL_PAYMENT	0.0	0.2	0.2	0.0	0.0	0.3	0.0
PURCHASES	1858.3	369.3	802.3	362.1	837.6	3651.2	1309.3
PURCHASES_FREQUENCY	0.5	0.3	0.6	0.3	0.4	0.9	0.5
PURCHASES_INSTALLMENTS_FREQUENCY	0.5	0.2	0.4	0.2	0.2	0.9	0.3
PURCHASES_TRX	24.3	5.0	13.5	6.1	12.3	48.8	18.0
TENURE	11.8	11.3	11.6	11.1	11.4	11.9	11.8
Monthly_avg_purchase	161.8	32.8	69.5	33.2	73.1	307.0	109.6
Monthly_cash_advance	76.9	28.1	16.5	179.6	532.4	21.9	156.3
limit_usage	0.9	0.0	0.4	0.6	0.6	0.3	0.6
installment	0.3	0.4	0.3	0.1	0.1	0.2	0.1
none	0.3	0.2	0.1	0.5	0.4	0.0	0.3
one_off	0.2	0.3	0.2	0.2	0.2	0.0	0.2
Balance	High	Low	Low	Med	High	Med	High
Credit Limit	Med	Med	Low	Low	High	High	High
Installment Purchase	High	Low	Low	Low	Low	High	Med
Purchase / Installment Frequency	High	Low	Med	Low	Low	High	Med
Monthly_avg_purchase	High	Low	Med	Low	Med	High	High
Monthly_cash_advance	Med	Low	Low	High	High	Low	High
limit_usage	High	Low	Med	High	High	Low	High