**Comprehensive Project Report**

**Project Name – Credit Card Segmentation**

**Made by *Moupriya Roy***

**Problem Statement –**

This case requires trainees to develop a customer segmentation to define marketing strategy. The sample dataset summarizes the usage behaviour of about 9000 active credit card holders during the last 6 months. The file is at a customer level with 18 behavioral variables.

**Number of attributes:**

● **CUST\_ID** Credit card holder ID

● **BALANCE** Monthly average balance (based on daily balance averages)

●**BALANCE\_FREQUENCY** Ratio of last 12 months with balance

● **PURCHASES** Total purchase amount spent during last 12 months

● **ONEOFF\_PURCHASES** Total amount of one-off purchases

**● INSTALLMENTS\_PURCHASES** Total amount of installment purchases

● **CASH\_ADVANCE** Total cash-advance amount

● **PURCHASES\_ FREQUENCY**-Frequency of purchases (percentage of months with at least on purchase)

● **ONEOFF\_PURCHASES\_FREQUENCY** Frequency of one-off-purchases

● **PURCHASES\_INSTALLMENTS\_FREQUENCY** Frequency of installment purchases

● **CASH\_ADVANCE\_ FREQUENCY** Cash-Advance frequency

● **AVERAGE\_PURCHASE\_TRX** Average amount per purchase transaction

● **CASH\_ADVANCE\_TRX** Average amount per cash-advance transaction

● **PURCHASES\_TRX** Average amount per purchase transaction

● **CREDIT\_LIMIT** Credit limit

● **PAYMENTS**-Total payments (due amount paid by the customer to decrease their statement balance) in the period

● **MINIMUM\_PAYMENTS** Total minimum payments due in the period.

● **PRC\_FULL\_PAYMENT**- Percentage of months with full payment of the due statement balance

● **TENURE** Number of months as a customer

**Goal of Model:**

***Advanced data preparation:*** Build an ‘enriched’ customer profile by deriving “intelligent” KPIs such as:

1- Monthly average purchase and cash advance amount

2- Purchases by type (one-off, installments)

3- Average amount per purchase and cash advance transaction,

4- Limit usage (balance to credit limit ratio),

5- Payments to minimum payments ratio etc.

6- Advanced reporting: Use the derived KPIs to gain insight on the customer profiles.

7- Identification of the relationships/ affinities between services.

8- Clustering: Apply a data reduction technique factor analysis for variable reduction technique and a clustering algorithm to reveal the behavioral segments of credit card holders

9- Identify cluster characteristics of the cluster using detailed profiling.

10- Provide the strategic insights and implementation of strategies for given set of cluster characteristics

**Project life cycle:**

**Importing the required packages in python:**

Numpy, Matplotlib, Scikit-learn.

**Introduction to data**:

To begin with, dataset has been read in the downloaded CSV file. In 18 variables 14 are of float data type, 3 of integer data type, 1 is object data type which is customer id variable . Float and integer data type variables are numeric values .

**Float Variable :**

**1 BALANCE**

**2 BALANCE\_FREQUENCY**

**3 PURCHASES**

**4 ONEOFF\_PURCHASES**

**5 INSTALLMENTS\_PURCHASES**

**6 CASH\_ADVANCE**

**7 PURCHASES\_FREQUENCY**

**8 ONEOFF\_PURCHASES\_FREQUENCY**

**9 PURCHASES\_INSTALLMENTS\_FREQUENCY**

**10 CASH\_ADVANCE\_FREQUENCY**

**13 CREDIT\_LIMIT**

**14 PAYMENTS**

**15 MINIMUM\_PAYMENTS**

**16 PRC\_FULL\_PAYMENT**

**Integer Variable**

**11 CASH\_ADVANCE\_TRX**

**12 PURCHASES\_TRX**

**17.TENURE**

**Handing Missing Values:**

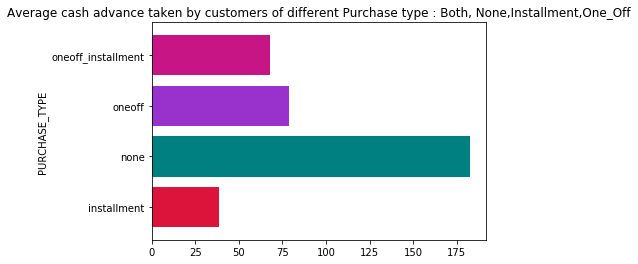
We checked for how many values are missing in our data.

There was one missing value credit limit variable and 313 missing value in minimum payment variable.

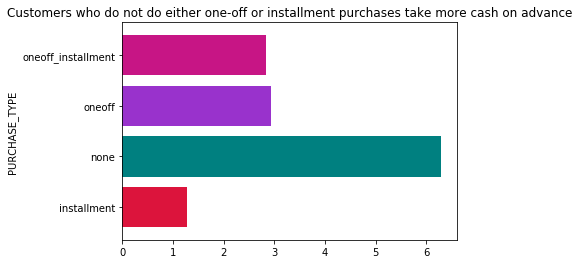
# Insights from ‘intelligent’ KPI’s derived from dataset through visualizations:

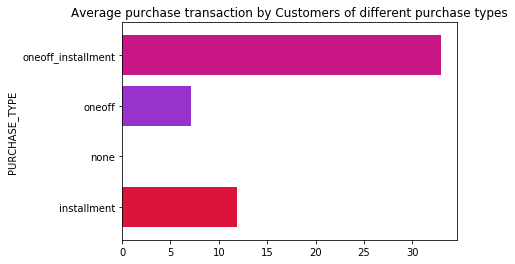


* Customers of one-off purchases has minimum payment ratio that means they are not paying any dues only using their credits for one-off purchases only.
* Customers of installment purchases has maximum payment ratio that means they are paying dues as well as they had been probably using their credits for taking loan.
* Customers who does both type of purchases are comparatively has good Payment ratio.



* Customers of installment purchases are taking least average cash advance
* Customers who doesn’t do either of the purchases is not making enough use of their credits instead taking more cash in advance

****

****

* **One-off and installment purchases Customers are doing the most purchase transactions ended up using good amount of credit limit. They might be considered as potential customers from this perspective.**
* **Customers who don’t do any kind of purchases are more reluctant to use their credit card for purchases.**

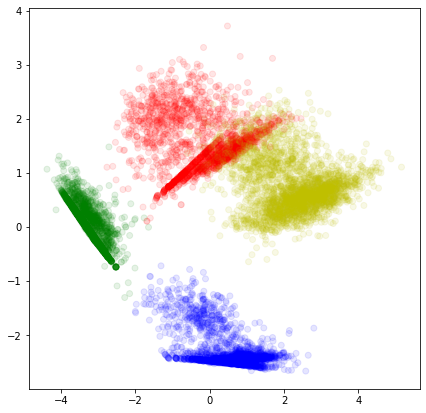


# Beneficial Information’s from Heat map

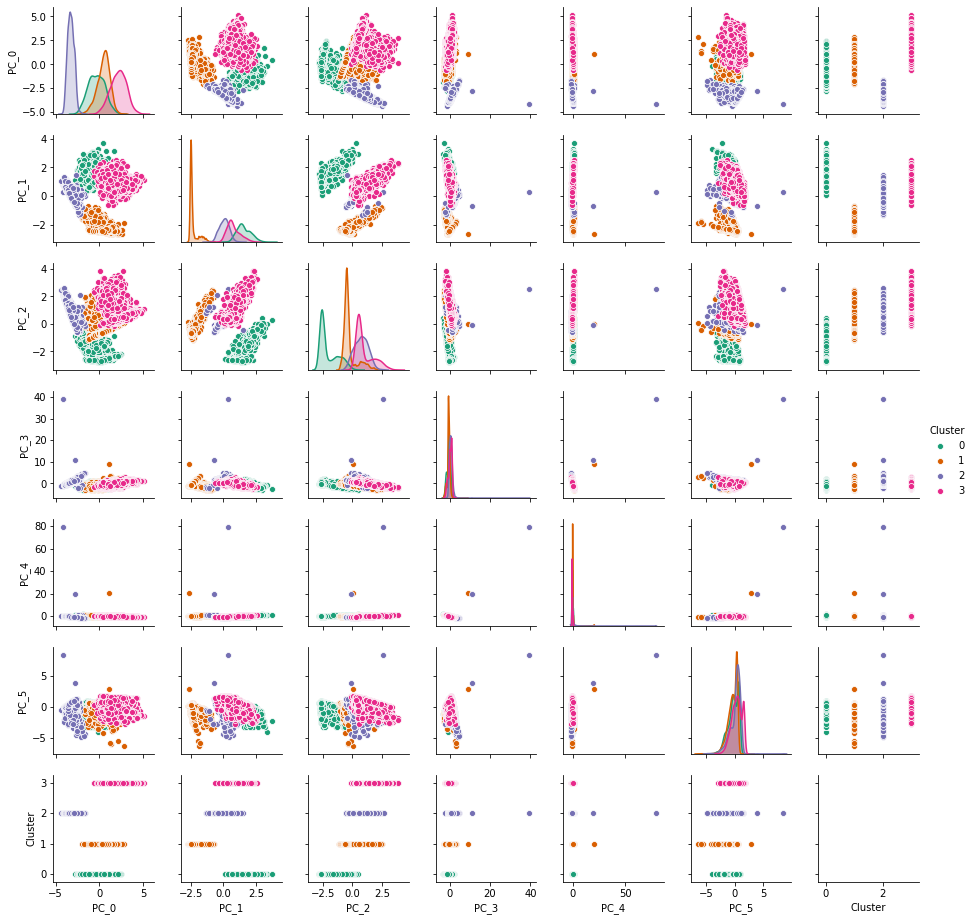
* One-off purchases and installment purchases are highly correlated with purchase transaction and negatively correlated with cash-advance=transaction.
* Purchase Frequency is more correlated with Installment purchases than One-off purchases.
* Purchase Transaction is highly correlated with Monthly average transaction
* Customers who does both one-off and installmentment purchases tend to have higher purchase frequency. Their Monthly average purchases are also higher than other customers
* Customer who doesn't do either of both the purchases tend to transact more during purchase rather than advance transactions

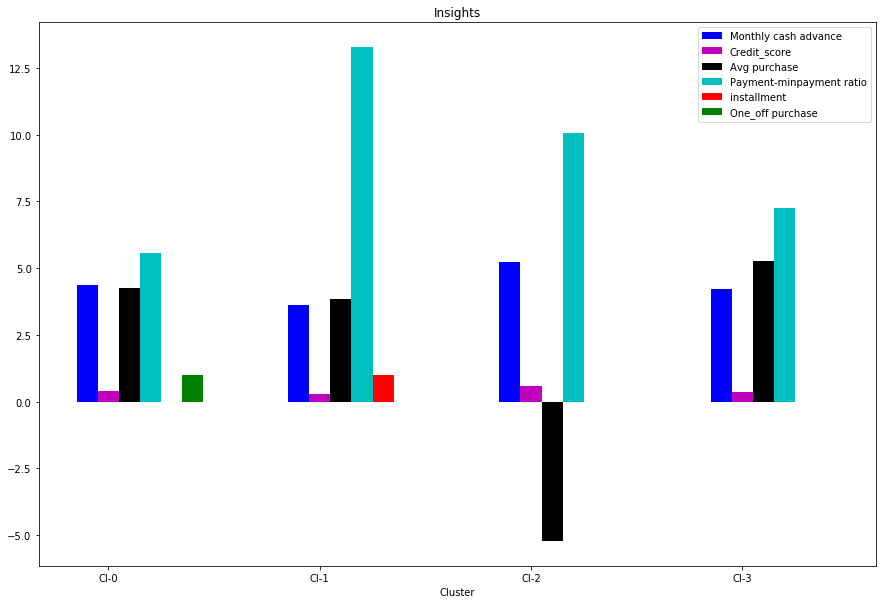
### Clustering

**Based on the intuition on type of purchases made by customers and their distinctive behavior exhibited based on the purchase\_type (as visualized in Insights from KPI) , we can start with 4 clusters.**



* Even though we can see 4 clusters arriving but it is little difficult to understand this way as the clusters might overlap one on another and also we're unable to get individual plot for each cluster this way





# Conclusion with 4 Clusters

## Monthly cash advance Transaction

* Highest:Cluster 2
* lowest:Cluster 1
* Similar(aprox):Cluster 0,Cluster 3

## Highest Credit score

* cluster 2

## Average Purchase

* Highest: Cluster 3
* Lowest: Cluster 2

## Payment to min\_payment ratio

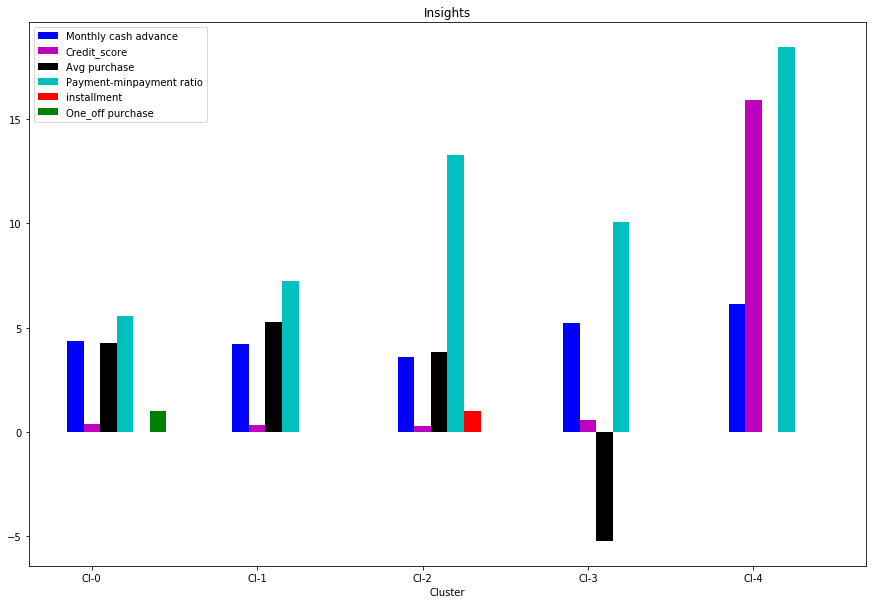
* Highest:Cluster 1
* Lowest: Cluster 0

## Only one\_off purchase done by customer of Cluster

* Cluster 0

## Only installment purchase done by customer of Cluster

* Cluster 1



# Conclusion with 5 Clusters

## Monthly cash advance Transaction

* Highest:Cluster 4
* Next to highest: Cluster 3
* lowest:Cluster 2
* Similar(aprox):Cluster 0,Cluster 1

## Credit score

* Highest:cluster 4
* Lowest Cluster 3

## Average Purchase

* Highest: Cluster 1
* Lowest: Cluster 3

## Payment to min\_payment ratio

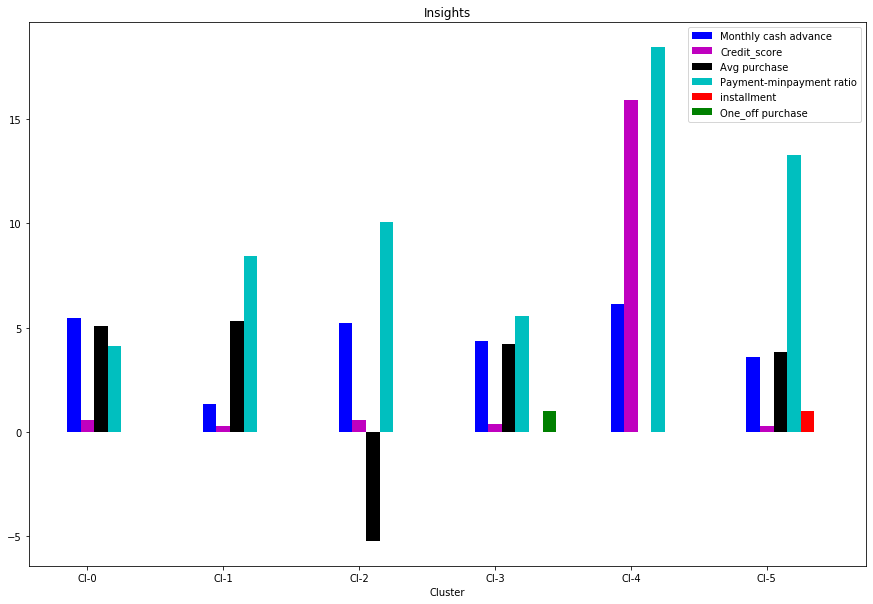
* Highest:Cluster 4
* Lowest: Cluster 0

## Only one\_off purchase done by customer of Cluster

* Cluster 0

## Only installment purchase done by customer of Cluster

* Cluster 2



# Conclusion with 6 Clusters

## Monthly cash advance Transaction

* Highest:Cluster 4
* Next to highest: Cluster 0
* lowest:Cluster 1
* Decrement Order of clusters: 4,0,2,3,5,1

## Credit score

* Highest:cluster 4
* Lowest Cluster 1

## Average Purchase

* Highest: Cluster 1 ,Cluster 0
* Lowest: Cluster 2

## Payment to min\_payment ratio

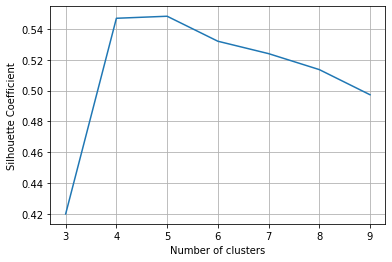
* Highest:Cluster 4
* Lowest: Cluster 0

## Only one\_off purchase done by customer of Cluster

* Cluster 3

## Only installment purchase done by customer of Cluster[¶](http://localhost:8892/notebooks/Downloads/EDWISER/Project%201/Credit%20Card%20Customer%20Segmentation.ipynb#Only-installment-purchase-done-by-customer-of-Cluster)

* Cluster 5



## Using Silhouette Coefficient we can see that K-means with 4 cluster is able to show distinguished characteristics of each cluster.

## Insights with 4 Clusters

* Cluster 0 customers are doing maximum One\_Off transactions and least payment ratio and credit\_score on lower side though higher than cluster 1;This group is about 21% of the total customer base
* Cluster 1 customers have maximum Payment ratio and are paying dues and are doing maximum installment purchases and has a credit\_score on lower side. This group is about 25% of the total customer base
* Cluster 2 is taking maximum advance\_cash and have lower payment ratio i.e is paying comparatively less minimum payment and even though maintaining highest credit\_score & doing no purchase transaction. This group is about 23% of the total customer base
* Cluster 3 is the group of customers who have highest Monthly\_avg purchases and doing both installment as well as one\_off purchases, have comparatively better credit score than cluster 1 customers. This group is about 31% of the total customer base

## Marketing Strategy Suggested:

* A. Group 0
  + This group is has minimum paying ratio and using card for just oneoff transactions (may be for utility bills only). This group seems to be risky group.
* B. Group 1:
  + They are potential target customers who are only doing installment purchases and paying dues but with a lower credit score So we can increase credit limit or can lower down interest rate or they can be given premium card /loyality cards to increase transactions
* C. Group 1
  + Even though they have highest credit score i.e not using their credit limit very much and not doing any kind of purchase very much even if they do then they're only taking only cash on advance. We can target them by providing less interest rate on purchase transaction.
* d. Group 3
  + This group is doing both one-off and installment purchase and performing best among all as customers ,are maintaining good credit score and paying dues on time. Good point is they falls under the highest percentage of customer base. Giving rewards point will make them perform more purchases.

**Instructions to run and deploy code**

As the code file is written in ipynb file it should be opened by Jupyter Notebook installed in the users system. If the user don’t have Jupyter notebook installed he or she should be installing it to open the notebook/code file.