Credit Card Customer Segmentation based on Customer Purchase Behaviour

S.A.M.V. SWETHA

Abstract

- Segmentation in business is used to divide customers based on attributes such as behaviour or demographics, in order to provide improved consumer experiences

- The credit card dataset used for the analysis summarizes the usage behaviour of about 9000 active credit card holders during the last 6 months. As per my research question, in order to identify and segment customers - I chose to apply K-Means clustering based on their purchase history, credit limit and balance

- As part of the analysis, optimum number of clusters were calculated using elbow plot, various purchase behaviours were identified using purchase history, clustering was performed using K-Means where-in customers were segmented into Max Payers, Money borrowers, Traders and potential riskers, which can be used to define bank’s marketing strategy

Motivation

- In today’s competitive financial markets, it is essential to have good customer understanding to increase customer lifetime value. The goal of customer segmentation is to group customers by common characteristics to identify profitable, risky segments and enable companies to target each segment with specific offerings

- The focus of this analysis is on business strategy driven customer segmentation, to identify various payment behaviours of customers. Present analysis segments bank customers into Max Payers, ordinary users, money borrowers, Revolvers, Traders, potential riskers

- The information is not only helpful for the bank to understand related characteristics of different customers, but also for the marketing representatives to find potential customers and implement target marketing

Dataset(s)

- CreditCard Dataset from ‘kaggle’ is used for the analysis which summarizes the usage behaviour of about 9000 active credit card holders during the last 6 months, constitutes around 900 KB

- The dataset consists of 18 features about the behaviour of credit card customers. These include variables such as the balance currently on the card, the number of purchases that have been made on the account, the credit limit, and many other attributes which define customer purchase and payment behaviour

- A complete data dictionary can be found on the data download page

Data Preparation and Cleaning

- Data Cleaning was performed to check the missing/corrupted values
- MINIMUM_PAYMENTS and CREDIT_LIMIT columns were identified with (313,1) null values respectively, which were then imputed with zero and median values
- Columns such as CUST_ID are dropped and the input dataset is normalised using Standard Scaler function

Research Question(s)

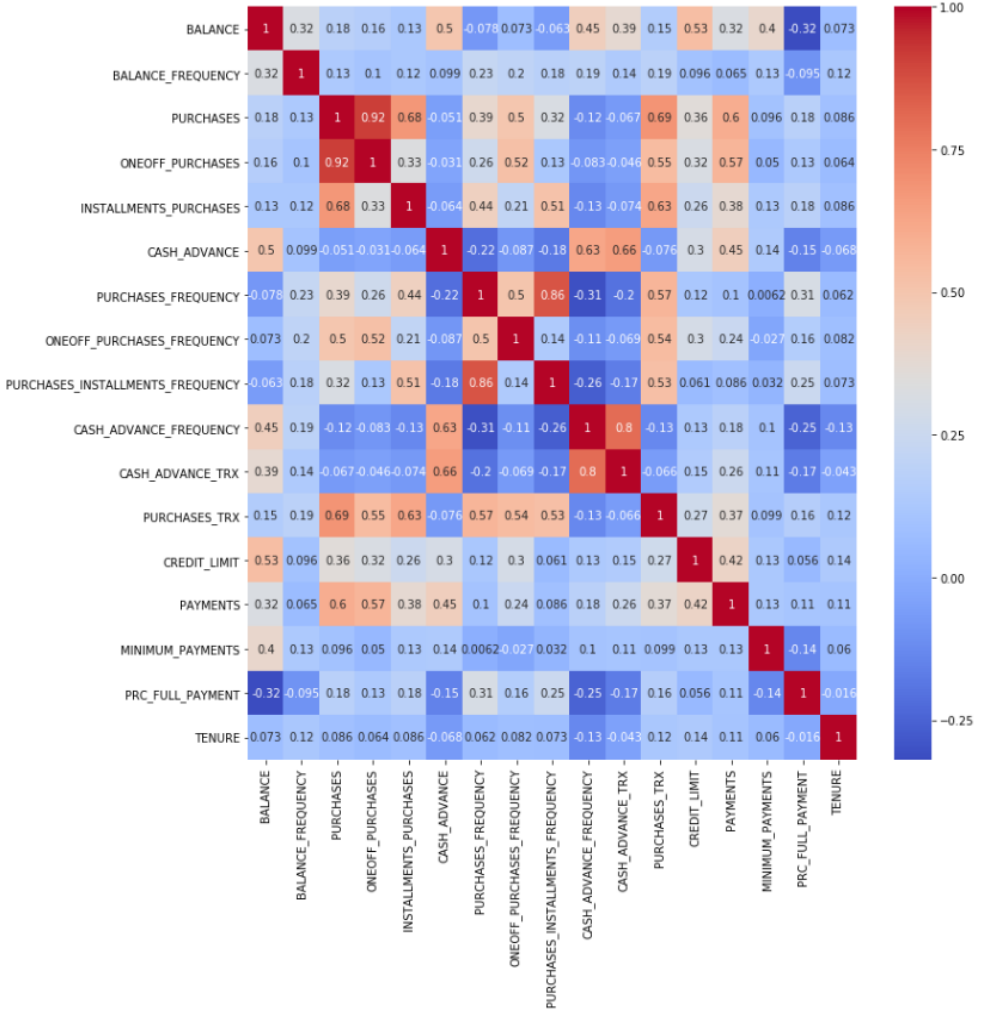
- Which segments of customers are most likely to make payments on time and delay the payments?
- Which segment of customers can be identified as Max Payers, Potential Riskers, Revolvers, Money Borrowers and Traders?
- What are the distinct purchasing behaviours of customers that can be identified from the dataset?
- Based on the purchasing behaviour, Which of these categories do the customers fall into – early payments or late payments?

Methods

- Firstly, the optimal number of clusters for the K-Means algorithm is found using the ’Elbow Method’ which computes the sum of squared distances for k clusters
- The ‘K-Means Algorithm’ is used which divides the dataset into 6 clusters by segregating customers with similar payment behaviour into the same cluster
- Based on the trends of purchase history and credit limit, customer payment behaviour is predicted for each cluster, and can be used for target marketing
- The purchase history data from the dataset is used to define distinct purchase behaviours, using which the customer payment category were predicted

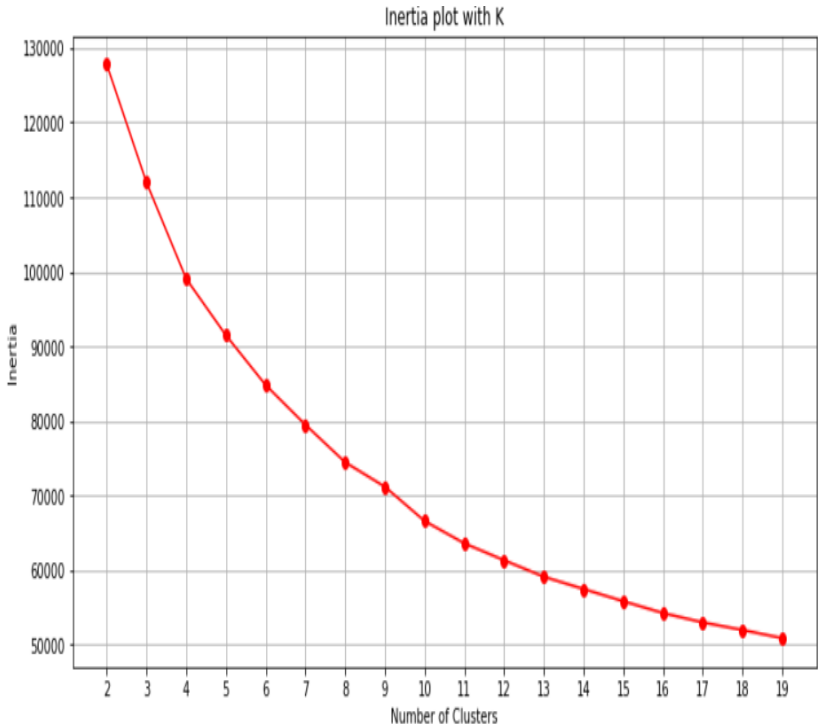
Identifying the attributes which define the customer purchase behaviour

- By using the correlation matrix plot, the multi – collinearity check is performed
- From the matrix: purchases, payments, credit_limit and balance are identified to be correlated and can be further used in the clustering analysis



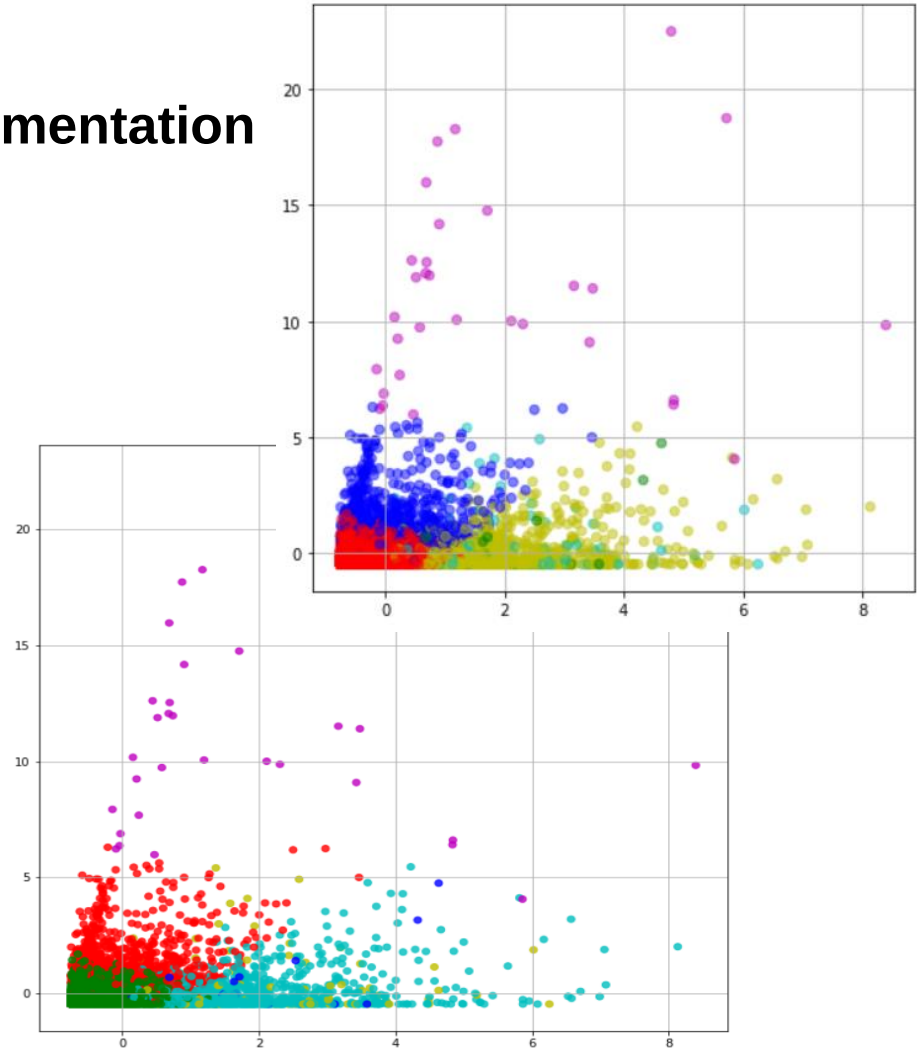
Finding the optimum number of clusters for K-Means Clustering

- The optimal number of clusters for the K-Means algorithm is found using the ’Elbow Method’ which computes the sum of squared distances for k clusters
- Using inertia_plot function, the elbow curve was plotted for the scaled dataset
- From the figure, the optimum number of clusters required are found to be ‘6’

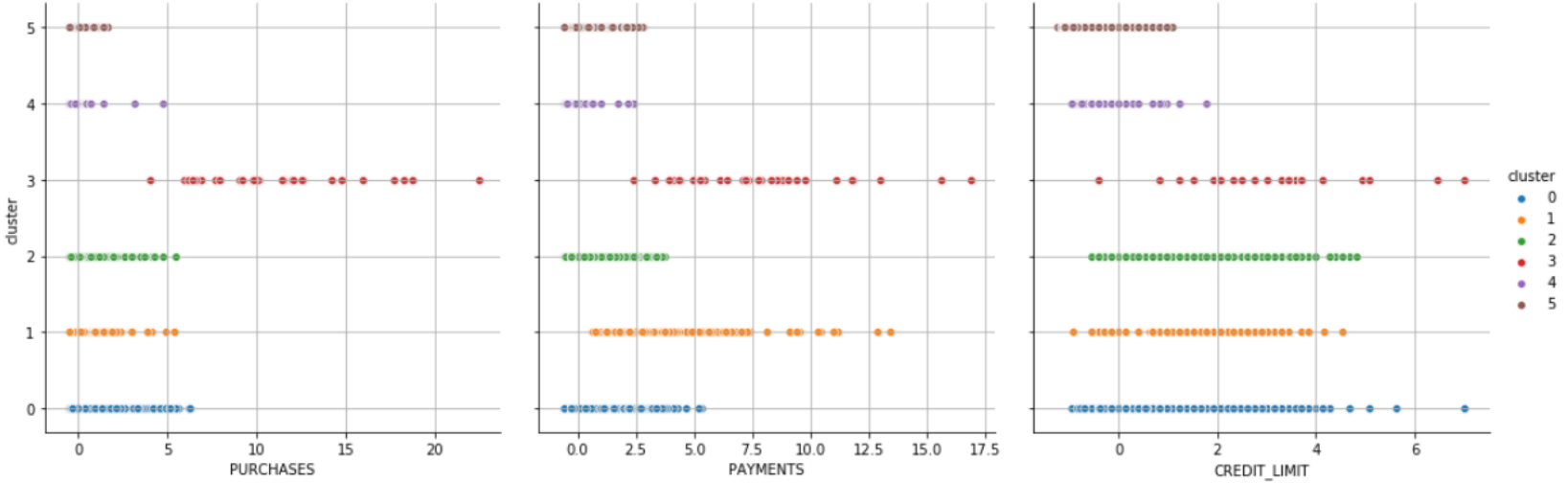


K-Means Clustering for Customer Segmentation

- Using scikit-learn machine learning library, K-Means clustering is implemented on the scaled dataset
- As per the elbow method, the customers with similar purchasing and payment behaviour are segmented into 6 different clusters
- The adjacent scatter plots show 6 different clusters, customer segmentations



Identifying Max Payers, Ordinary Users and Revolvers



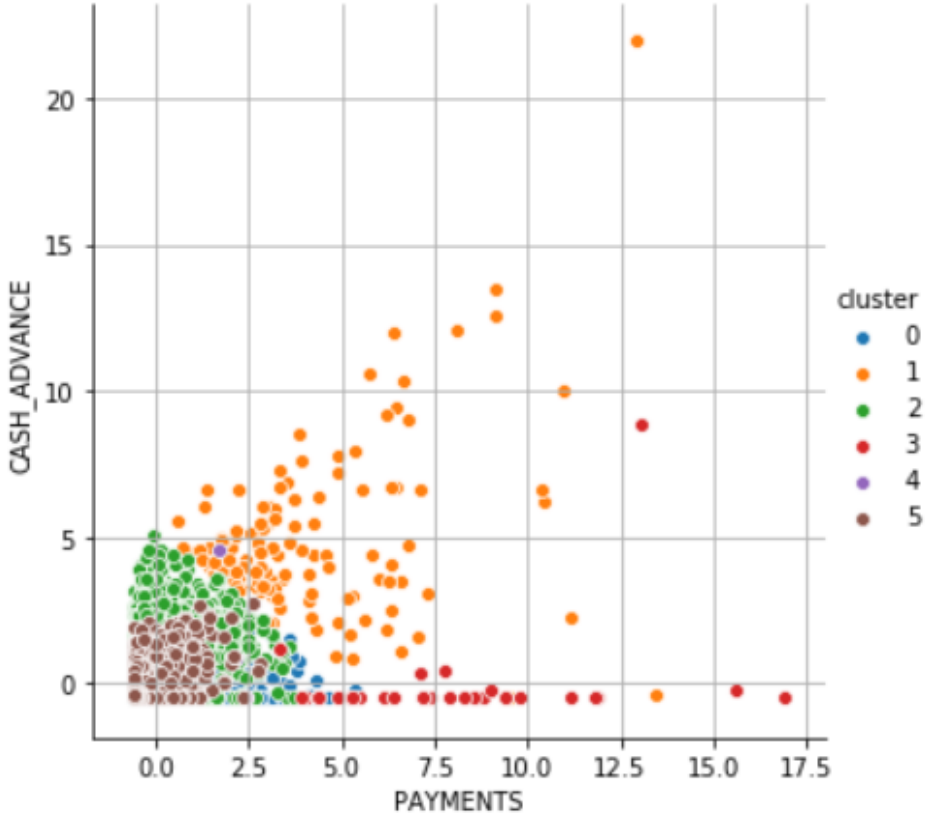
- **Cluster 0 (Blue):Ordinary Users** - This group of users have average purchases and payments, and lowest credit score (*Credit Score - Lower implies customers are maintaining their balance properly)

- **Cluster 3 (Red):Max Payers** - This group of users have highest purchases, highest payments, and low credit score values. They make payments on time

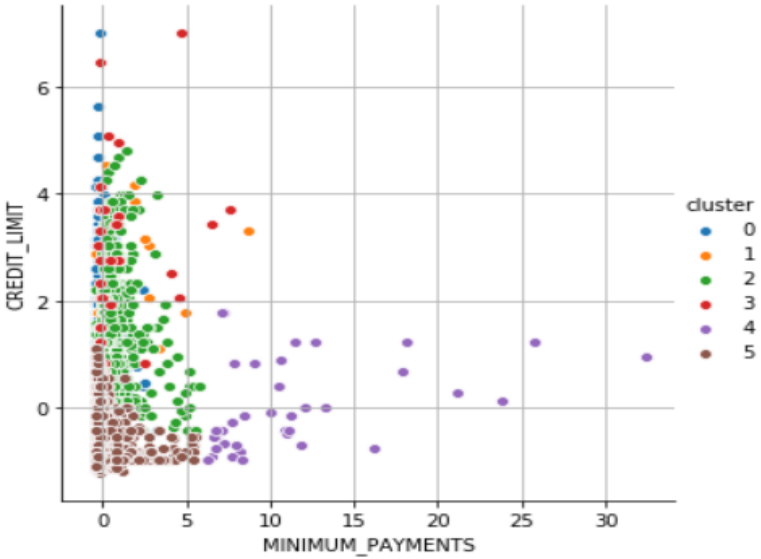
- **Cluster 2 (Green):Revolvers**- This group has the varied purchases, payments, highest minimum payments, but the other features are also wildly varied in values

Identifying the ‘Money Borrowers’

- **Cluster 1 (Orange): Money Borrowers**
- From the plots, it is noted that the customers who belong to this cluster have wildly varied balance, second highest payments, average purchases
- Also, these customers have the highest cash advance by far - there is even one extreme case that has 25 cash advance points

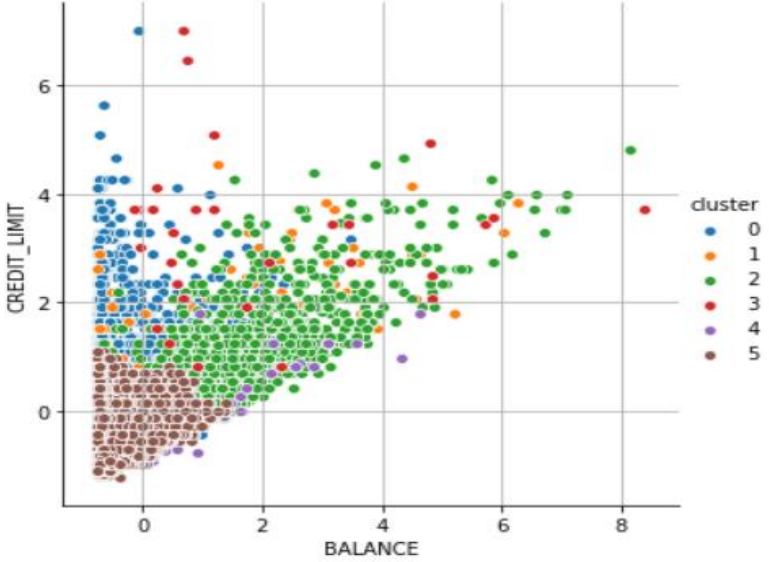


Identifying the High Riskers and Traders



- **Cluster 4 (Purple): Traders**

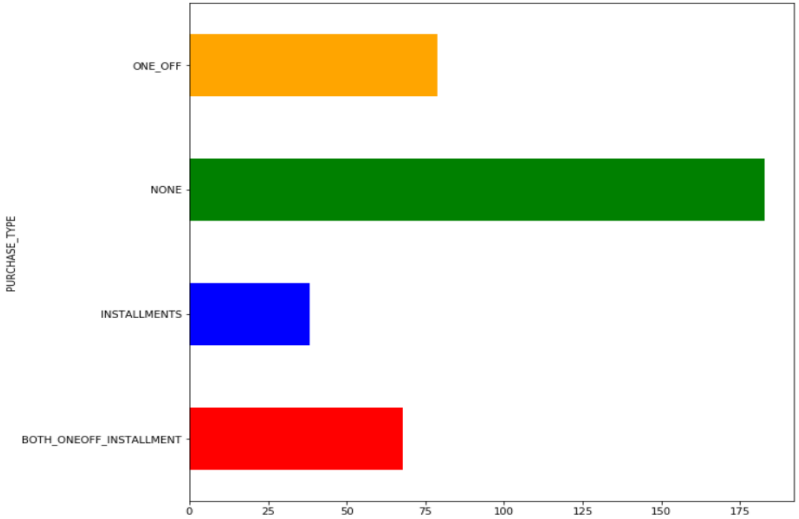
This group has highest minimum payments while having the second lowest credit limit. Therefore, bank can identify them as Traders

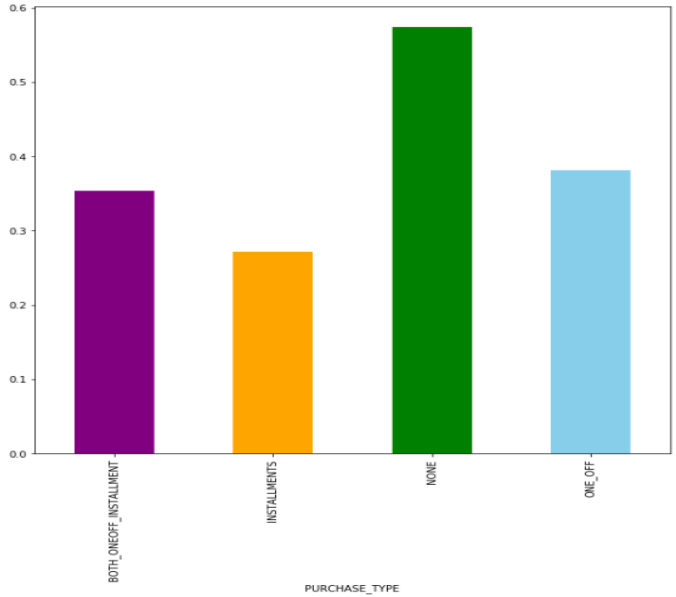


- **Cluster 5 (Brown): High Riskers**

This group is troublesome to analyse as both their credit limit and balance values are low, they have least purchases and payments. They usually delay their payments

Purchasing Behaviours of Customers





- Based on the monthly average cash advance calculated, the customers who do not do ONEOFF or INSTALLMENTS take more cash advance and can be categorized into late payments
- Based on the Credit Limit Usage calculated, Customers with installment purchases have good credit score and are paying dues, they make early payments

Limitations

- The present analysis is confined to a smaller dataset with less scope on the ‘time of purchases’ made. This attribute will be helpful to detect much more insights on the customer behaviour

- Based on the time and demographics of the purchase made, there is a scope to identify if the purchase was made by a trustworthy customer or a fraudulent customer

- In order to identify the delinquent customers, the dataset with near 6 months information may not suffice to derive into conclusions

Conclusions

- As per the analysis, **Max Payers** are important to the credit card companies because they make payments on time which allows them to lend out more money

- The **Ordinary Users** do not use credit card much in their daily life. They have healthy finances and low debts. Encouraging these people to use credit cards more is necessary for the company's profit

- **Revolvers** are a type of customers who pay only minimum amount due, and go about making purchases on their cards as usual. The interest rates and late fees these users pay are the bread and butter of the credit card industry

- **Money Borrowers** are the customers who take more cash advances. Companies must look after these users more diligently

Cont’d…

Conclusions

- A good way to manage **High Riskers** is by giving them low credit limits, cash advances with regular follow-up, else they might delay their payments

- **Traders** are the smartest credit card users. They always lookout for cards with maximum rewards, cash back, longest payback tenure etc. Companies can provide offers, cashbacks to attract there customers

- Customers with installment payments are paying dues and have good credit score, customers who do not do ONOFF or INSTALLMENTS take more cash advance

Acknowledgements

- The dataset was obtained from kaggle website which defines credit customer behaviour with 18 various attributes
- No other informal analysis has been used in this work
- The data was collected by myself from kaggle website and no feedback has been received
- Thanks to the Course Instructors Leo Porter and Ilkay Altinas for their amazing work

References

- I have completed the analysis on my own with the help of weekly notebooks provided in the course

- The information about different types credit card customers in the banking sector was studied from the following websites:
  - https://in.finance.yahoo.com/news/5-types-credit-card-users-143051341.html
  - https://cashmoneylife.com/5-types-of-credit-card-users/

- I’m attaching a PDF of my final-project notebook https://github.com/SwethaSripada/PythonNotebooks/blob/master/Credit%20Card%20Customer%20Segmentation.pdf