

Team Nine

# Credit Card Users Segmentation

Antonio Moral, Bosoo Kim, Jiajian(Sylar) Guo, Yixuan Wang

3/9/2021

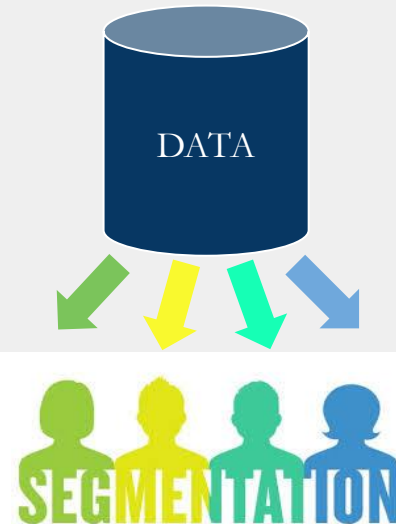


# Business Problem



## Industry Applications:

- Develop customized marketing campaigns
- Design an optimal distribution strategy
- Choose specific product features for deployment
- Prioritize new product development efforts



# Dataset

9000 Credit Card Holders | 18 Variable Columns

- CustomerID
- Balance
- Balance Frequency
- Purchases
- One-Off Purchase
- Installment Purchases
- Cash in Advance
- Purchases Frequency
- One-Off Purchase Frequency
- Tenure
- Purchases Installment Frequency
- Cash Advance Frequency
- No. Transactions Cash in Advance
- No. Purchase Transactions
- Credit Limit
- Payments
- Minimum Payments
- Percent of Full Payment Paid

[www.kaggle.com/arjunbhasin2013/ccdata](https://www.kaggle.com/arjunbhasin2013/ccdata)

# Data Cleaning

```
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
```

**CREDIT\_LIMIT:** 1 Missing

- Drop

**MINIMUM\_PAYMENTS:** 313 Missing

- PAYMENTS ==0 → 0
- PAYMENTS !=0 → Median Value

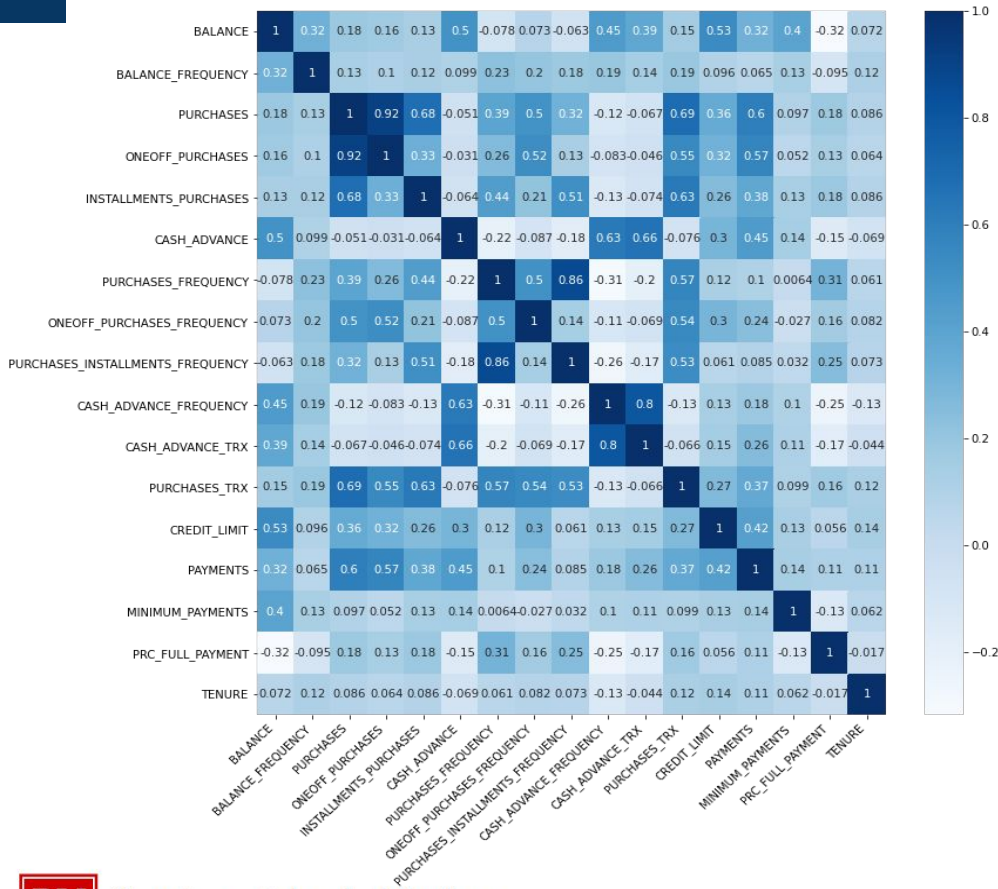
**CUST\_ID:** Meaningless

- Drop All



**8949 Rows & 17 Numeric Variables**

# Exploratory Data Analysis



## Heatmap (Correlation -1~1)

Key findings:

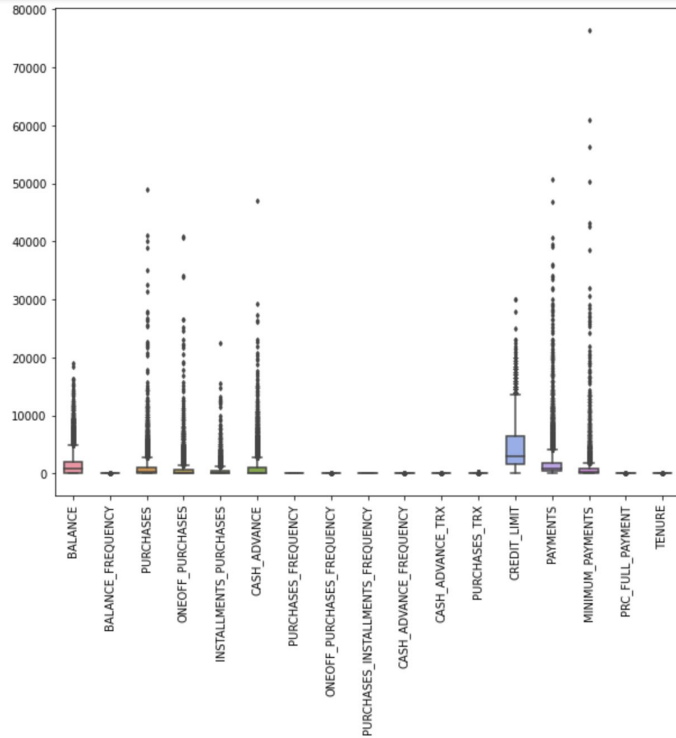
**0.92** : One-off\_Purchase ~ Purchase

**0.86** : Purchase\_Installments\_Freq. ~  
Purchase\_Frequency

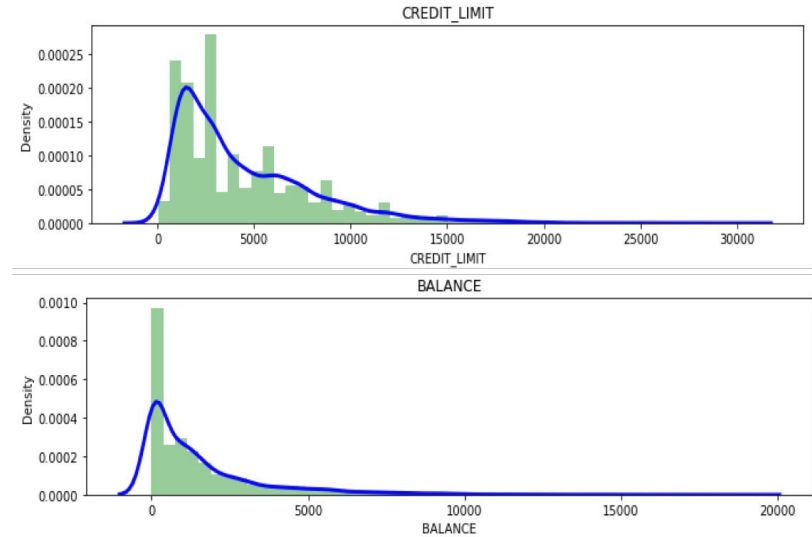
**0.68** : Installments\_Purchase ~ Purchase

# Exploratory Data Analysis

## Boxplot

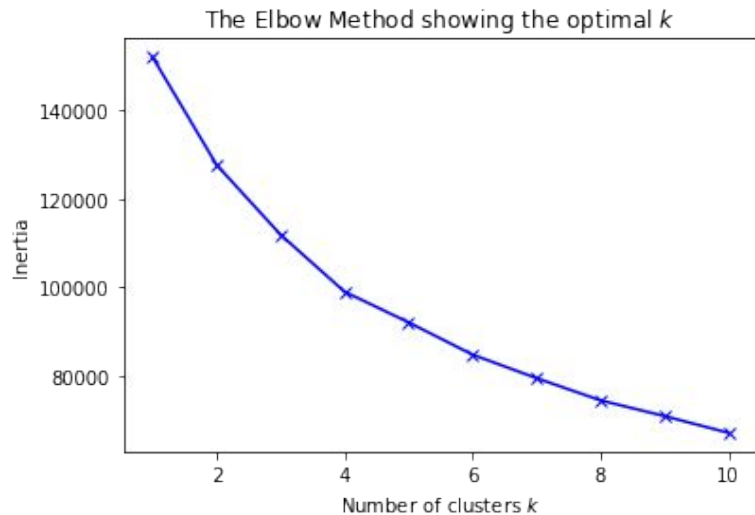


## KDE

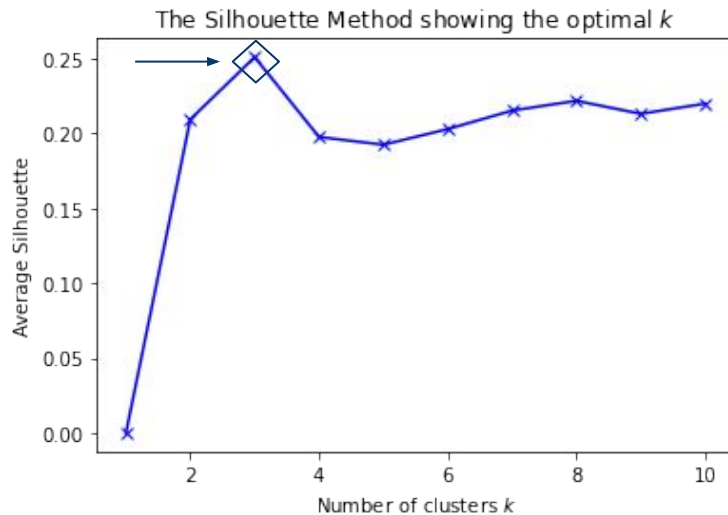


# Analytical Findings

## Elbow Method

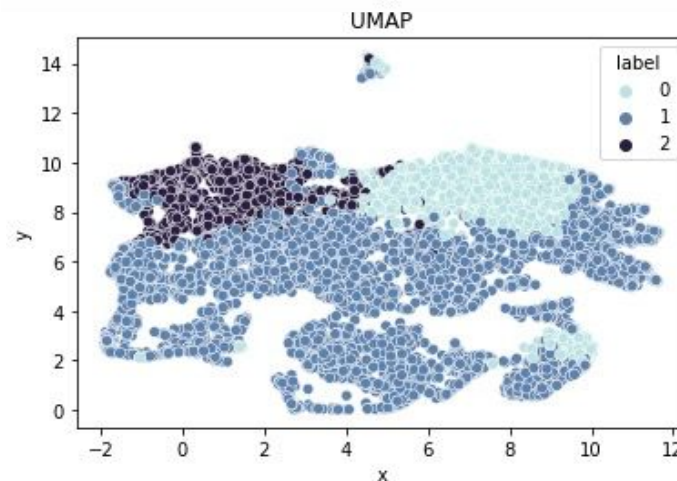
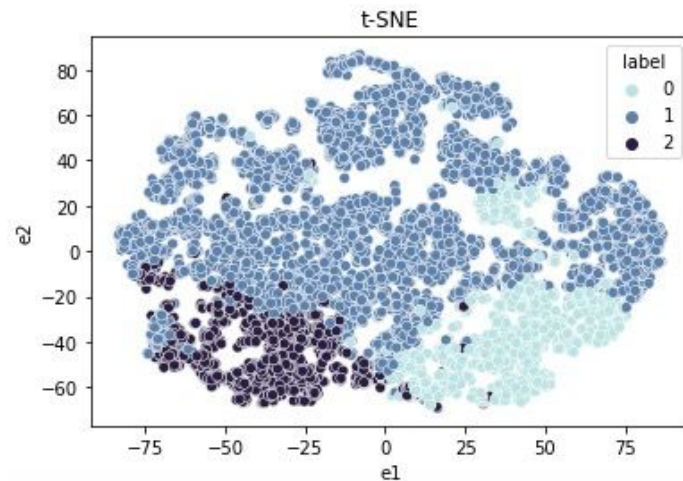
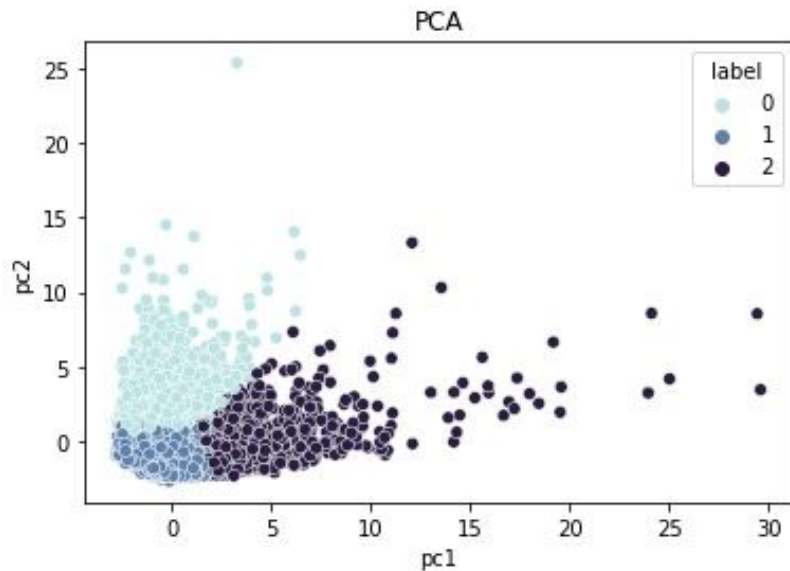


## Silhouette Score



# Analytical Findings

## 1. Clustering on Original Data

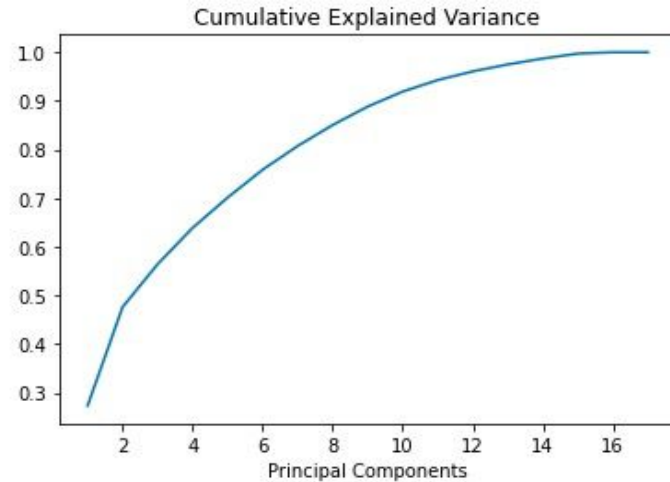
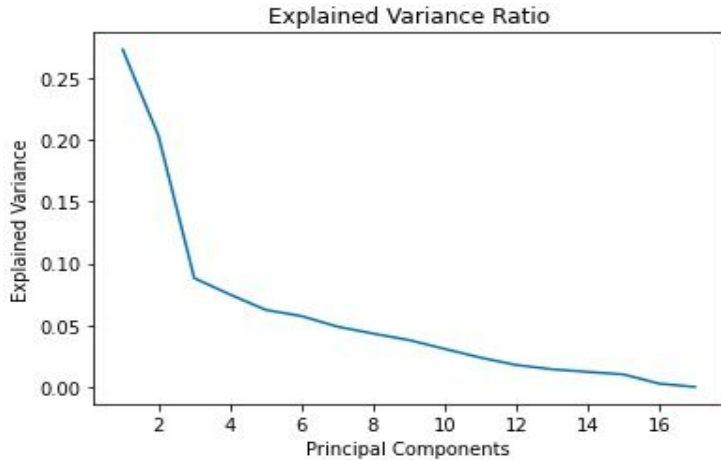




# Analytical Findings

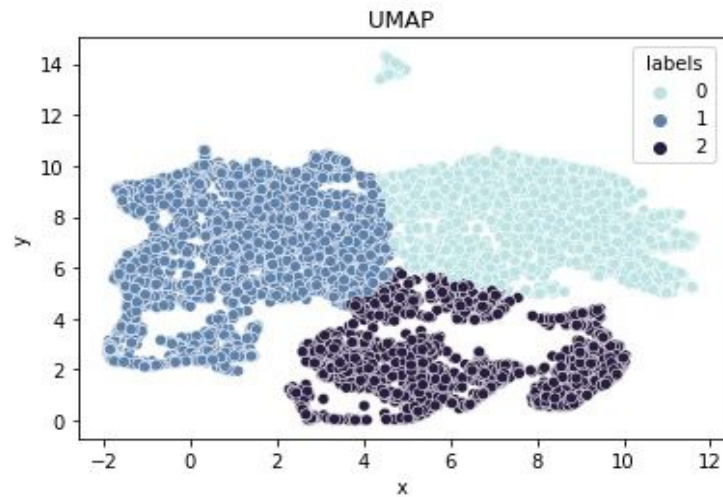
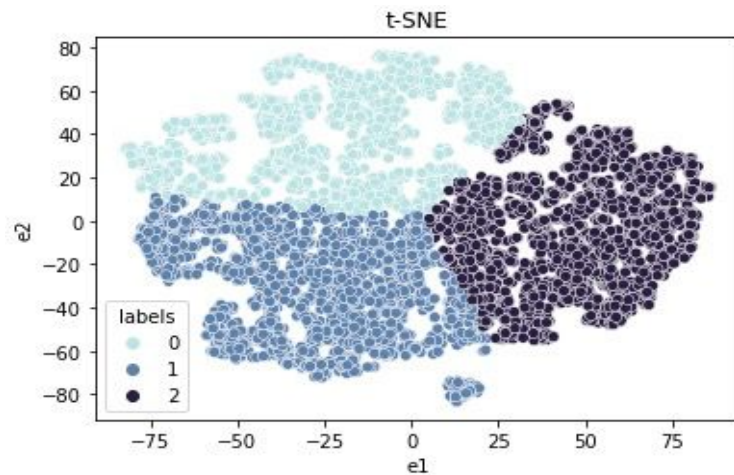
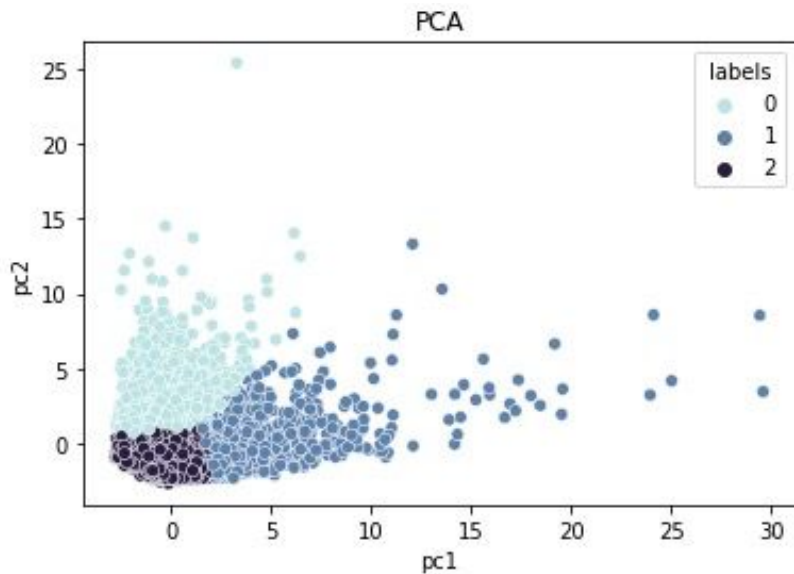
## 2. Clustering on Reduced Data

17 Features  $\rightarrow$  10 Principal Components



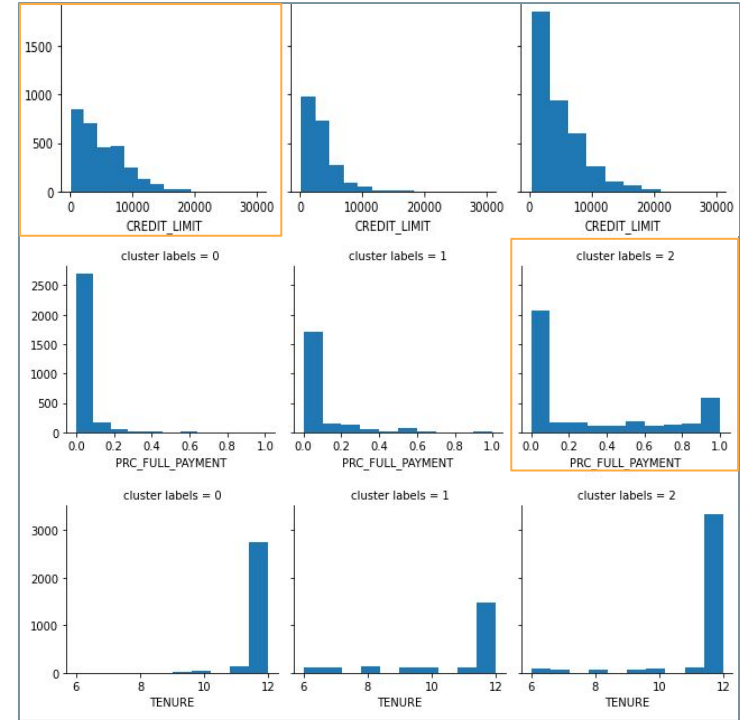
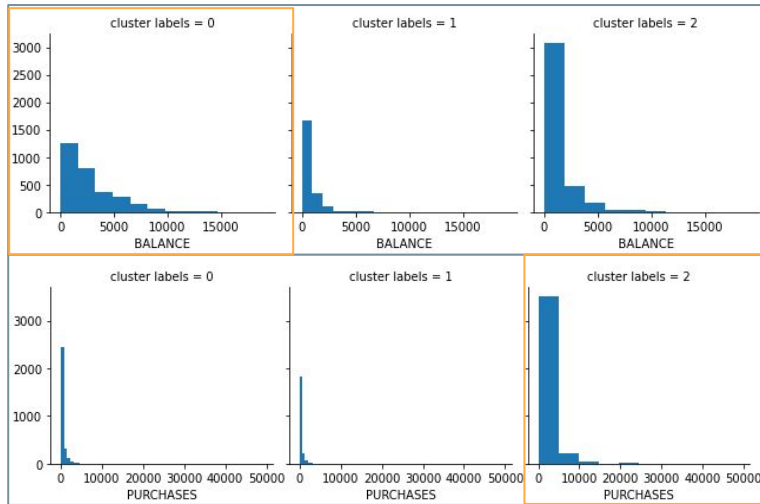
# Analytical Findings

## 2. Clustering on Reduced Data

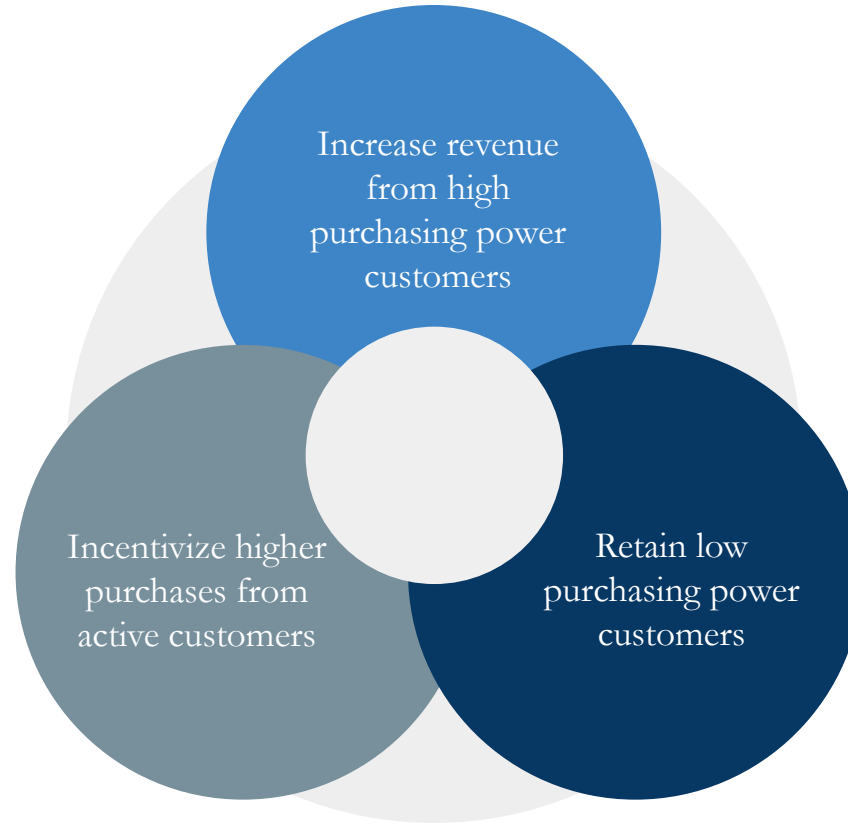


# Conclusions

## 3 Separate Clusters with Defined Characteristics.



# Recommendations



Q&A

# Thank You

---

Antonio Moral, Bosoo Kim, Jiajian(Sylar) Guo, Yixuan Wang

3/9/2021

