


# CREDIT CARD CHURN

ISSS621 Data Science for Business

G1 – Group 6

# Table Of Contents



# What is the credit card business?

Credit card business for banks is a lucrative and important revenue stream. Banks earn money by interest, fees and merchant fees

# Declining Credit Card Application Rates

Credit card application has been decreasing overtime. In 2020, credit card application rate fell by ~10%, and rejection rates increased by ~10%

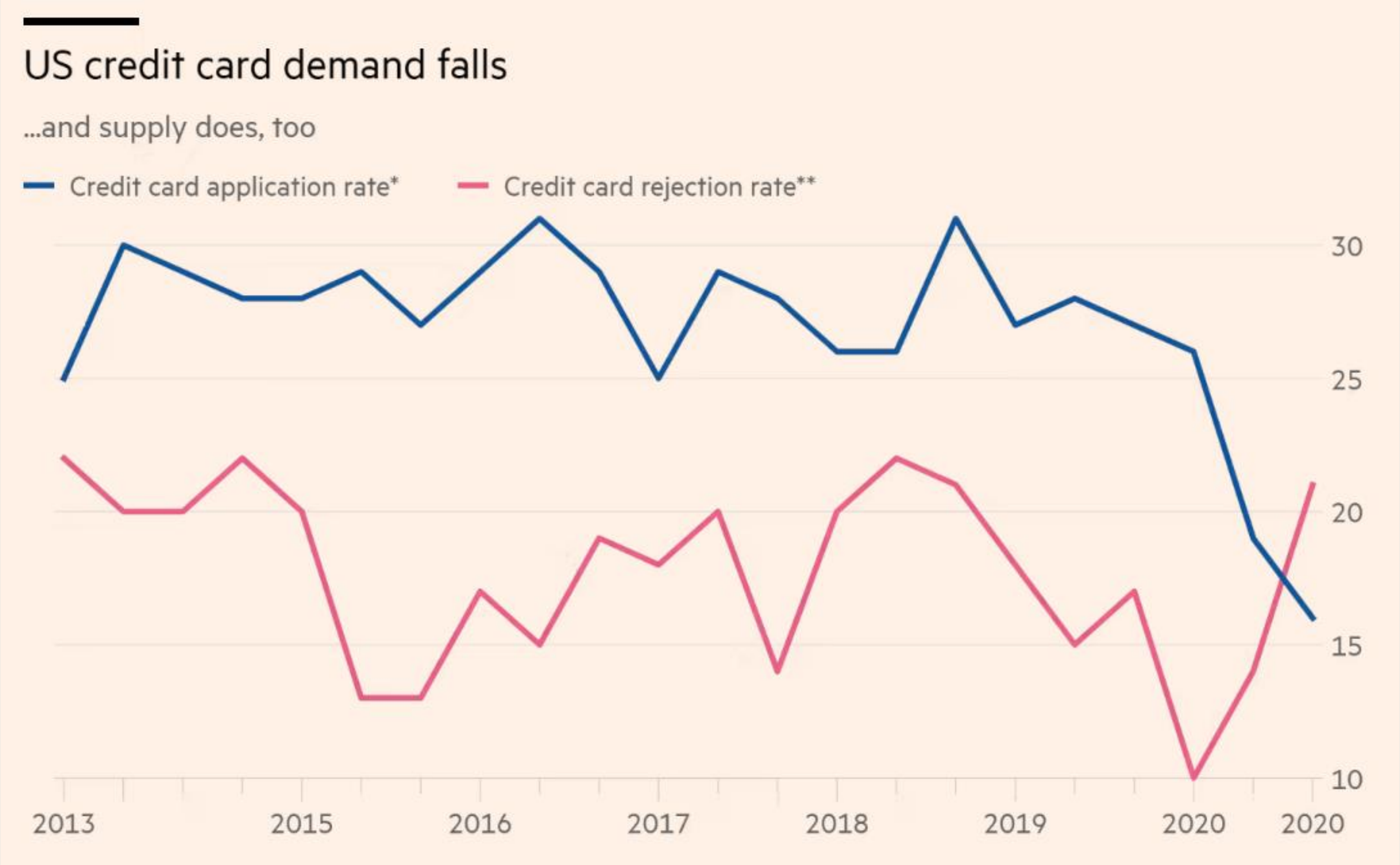



Figure 1: US credit card application rates [www.ft.com]

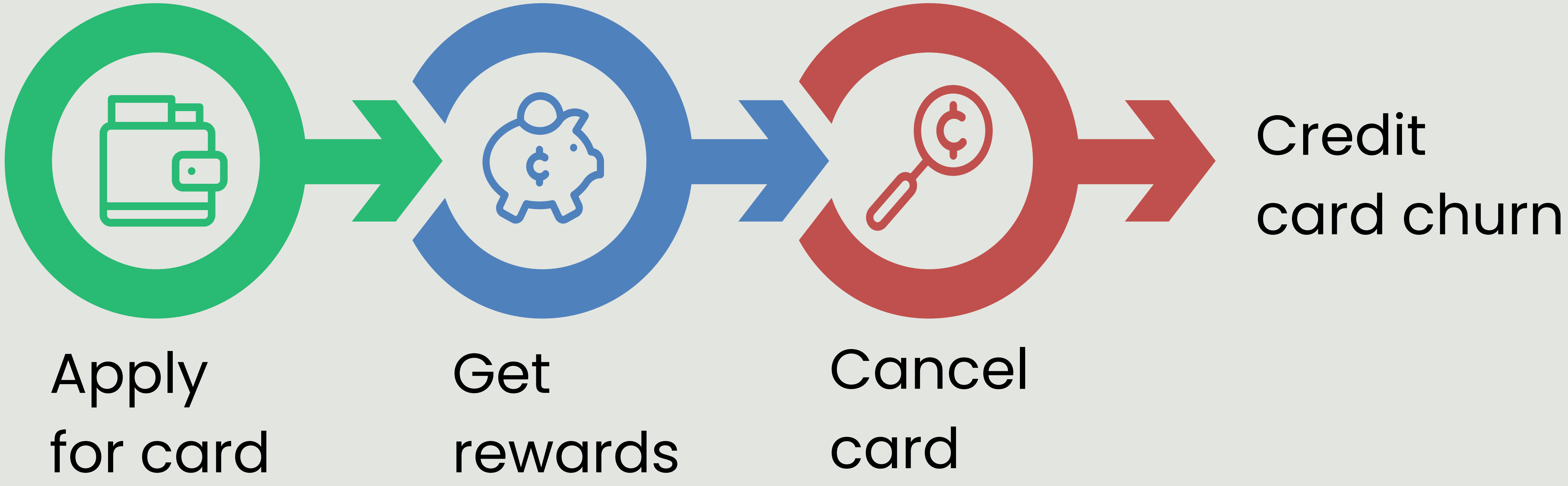
# Credit Card Promotions

Issuers offer increasingly attractive credit card promotions. They are key marketing strategy for banks and credit card issuers which include sign-up bonuses, introductory high interest rates or cash rewards.




# Credit Card Churn

Credit card churn is a strategy used by many savvy customers to take advantage of credit card rewards programs. Customer strategically time their application and cancellations to earn rewards



# Business Problem?


Cost of obtaining a customer is higher than revenue generated from customer leading to lost of profits

# Project Scope

To identify customer groups that are most likely to churn based on demographic and financial data

**Identify who is most likely to churn**

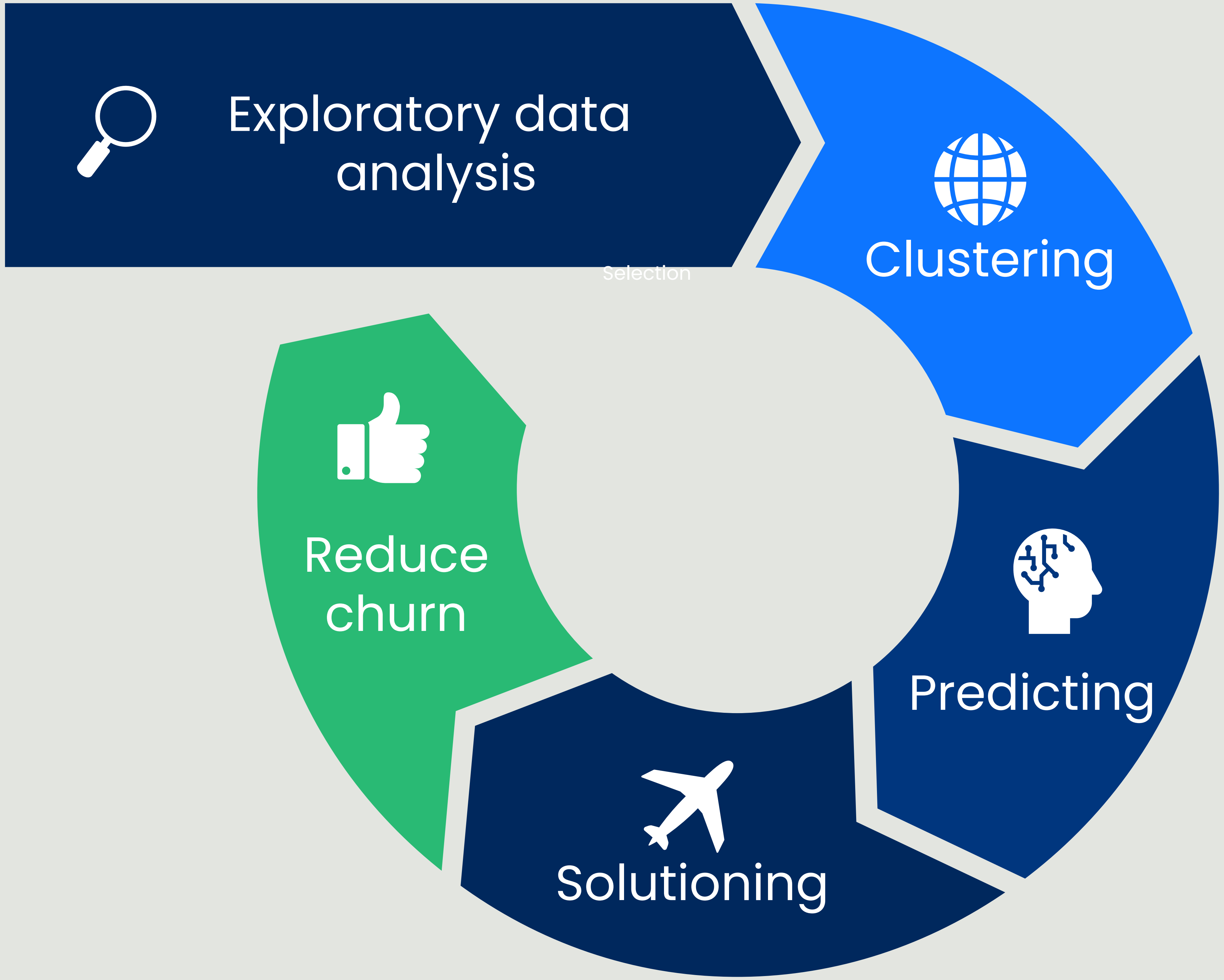
→

**Implement measures to reduce probability of churning**

→

**Reduce operating cost and increase profits**

# Data Science Translation




### Exploratory Data Analysis

To clean data and understand the dataset we are dealing with

### Clustering

To identify customer groups, segmentation and churn ratio within each group using customer data

### Predictive Modelling

To predict which group this customer will belongs to, and whether they are most likely to churn or not

### Solutioning

To propose solutions to target high churn group

# Exploratory data analysis

# DATASET

Consumer credit card portfolio

**DEMOGRAPHICS**

**SPENDING BEHAVIOUR**

**RELATIONSHIP WITH CREDIT CARD PROVIDER**

**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# DATASET

Consumer credit card portfolio

**DEMOGRAPHICS**

**AGE**
**GENDER**
**DEPENDENT COUNT**
**EDUCATION LEVEL**
**MARITAL STATUS**
**INCOME CATEGORY**

**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# DATASET

Consumer credit card portfolio

**DEMOGRAPHICS**

**SPENDING BEHAVIOUR**

**RELATIONSHIP WITH CREDIT CARD PROVIDER**

**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# DATASET

Consumer credit card portfolio

**SPENDING BEHAVIOUR**

**TRANSACTION COUNT**
**TRANSACTION AMOUNT**
**MONTHS INACTIVE**
**CARD UTILIZATION RATIO**

**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# DATASET

Consumer credit card portfolio

**DEMOGRAPHICS**

**SPENDING BEHAVIOUR**

**RELATIONSHIP WITH CREDIT CARD PROVIDER**

**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# DATASET

Consumer credit card portfolio

**RELATIONSHIP WITH CREDIT CARD PROVIDER**

**MONTHS ON BOOK**

**NUMBER OF CONTACTS IN LAST 12 MONTHS**

**TOTAL RELATIONSHIPS COUNT**

**CREDIT LIMIT**

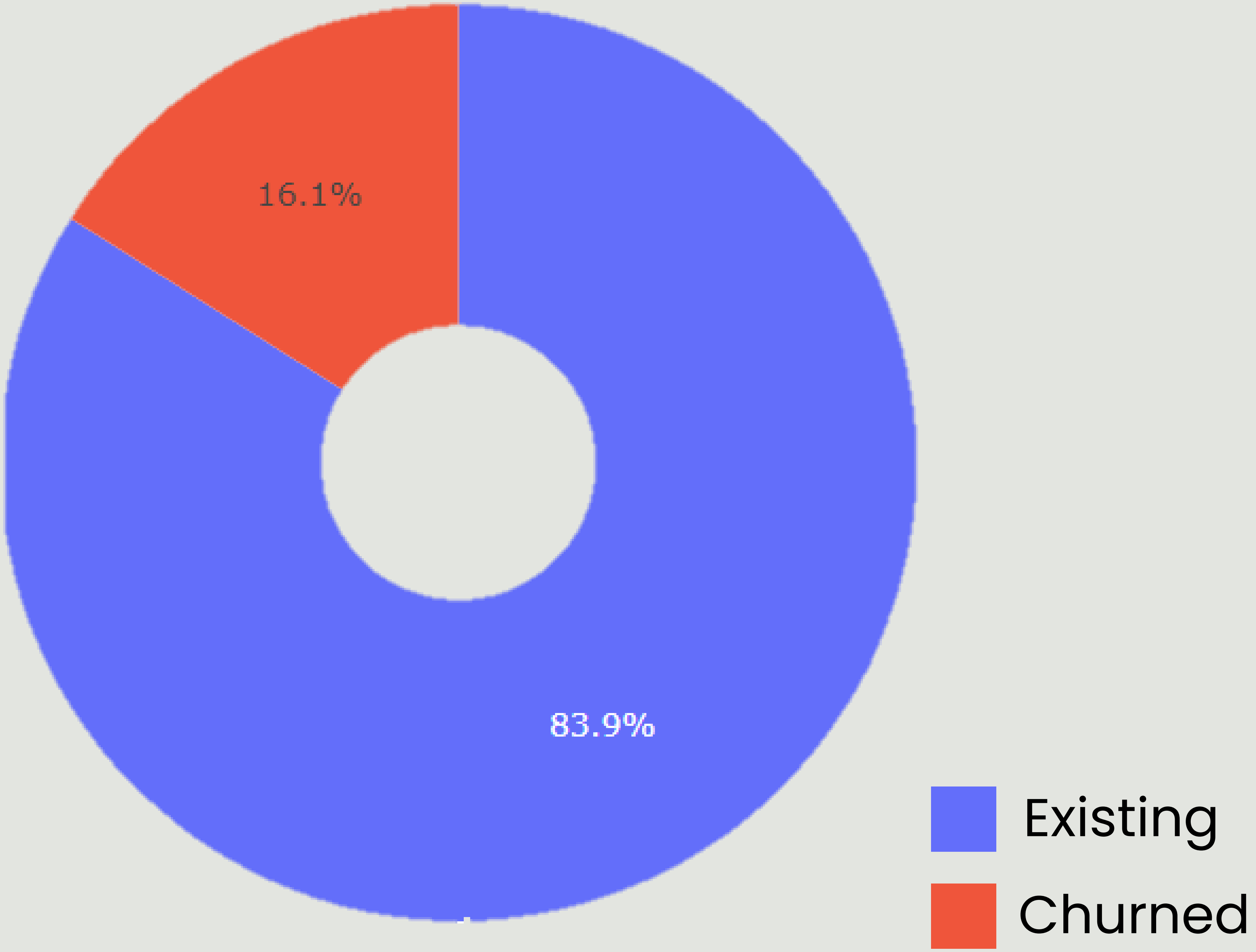
**Source:**
https://www.kaggle.com/datasets/thedevastator/predicting-credit-card-customer-attrition-with-m?resource=download

# Exploratory Data Analysis

Exploratory data analysis

## Churn vs Non-Churn

- 16.1% of customer population have attrited

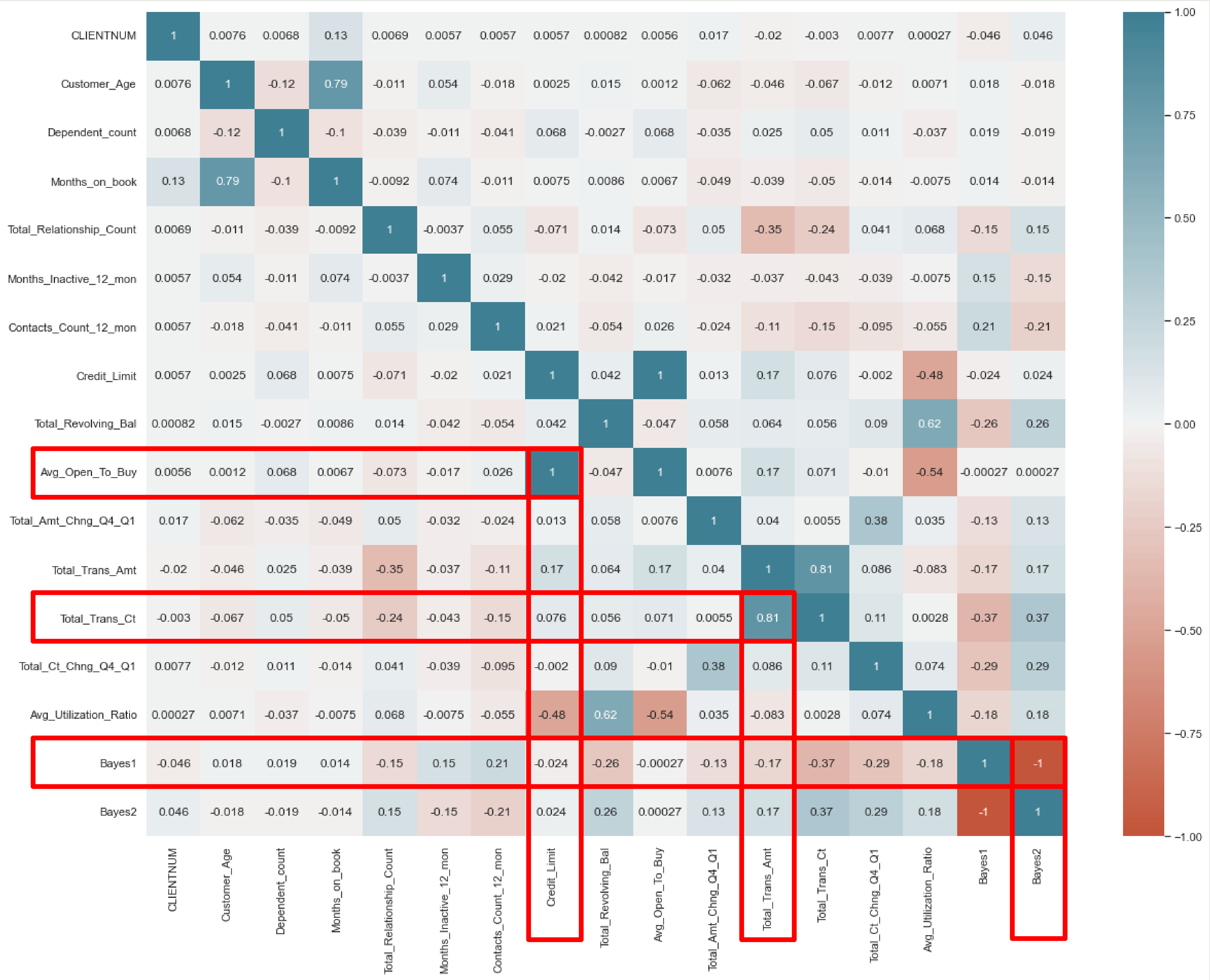


# Exploratory Data Analysis – Multivariate

Exploratory data analysis

## High Correlation

- Credit limit and average Open to buy ratio
- Total Transaction Count and Total Transaction Amount

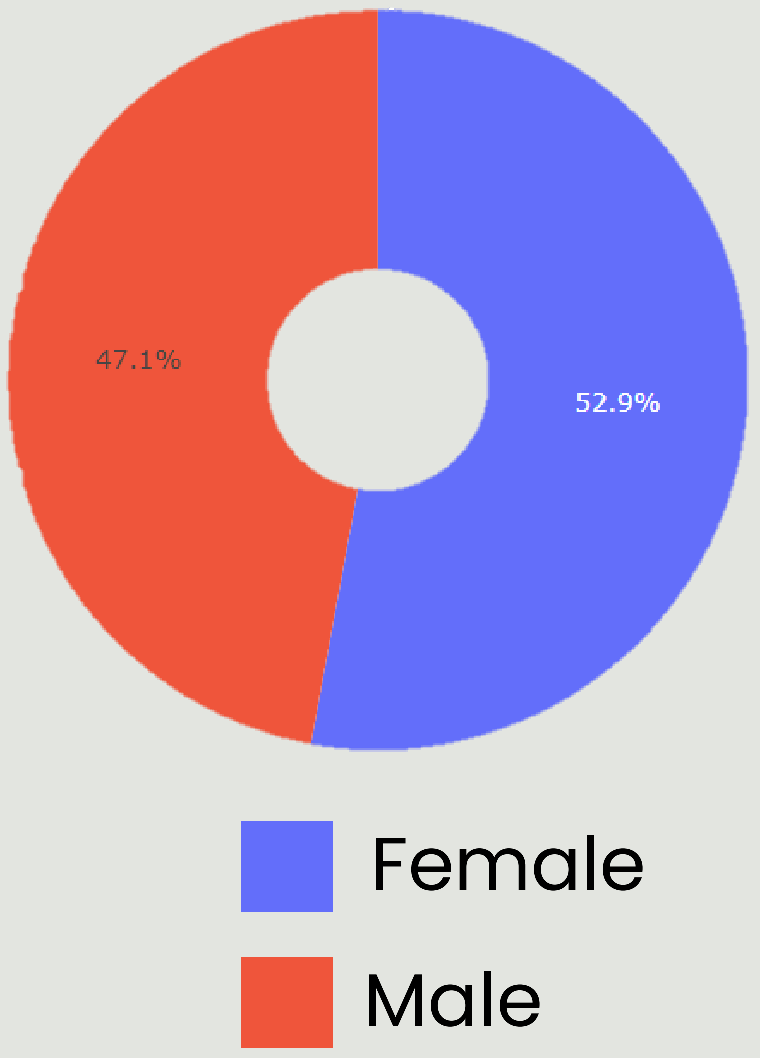


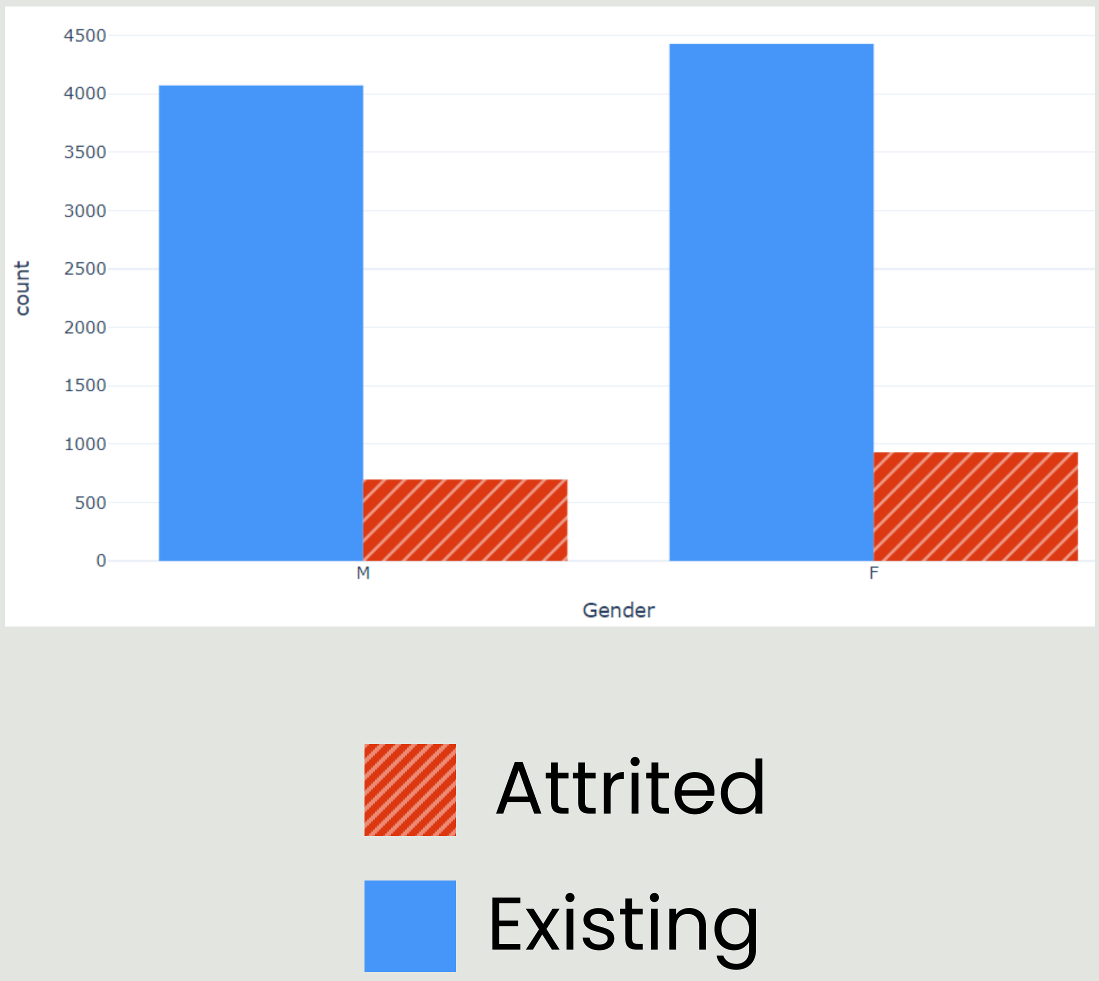
# Exploratory Data Analysis - Demographics

Exploratory data analysis

## Gender

- Higher proportion of customers are Female
- Higher number of Females who have churned



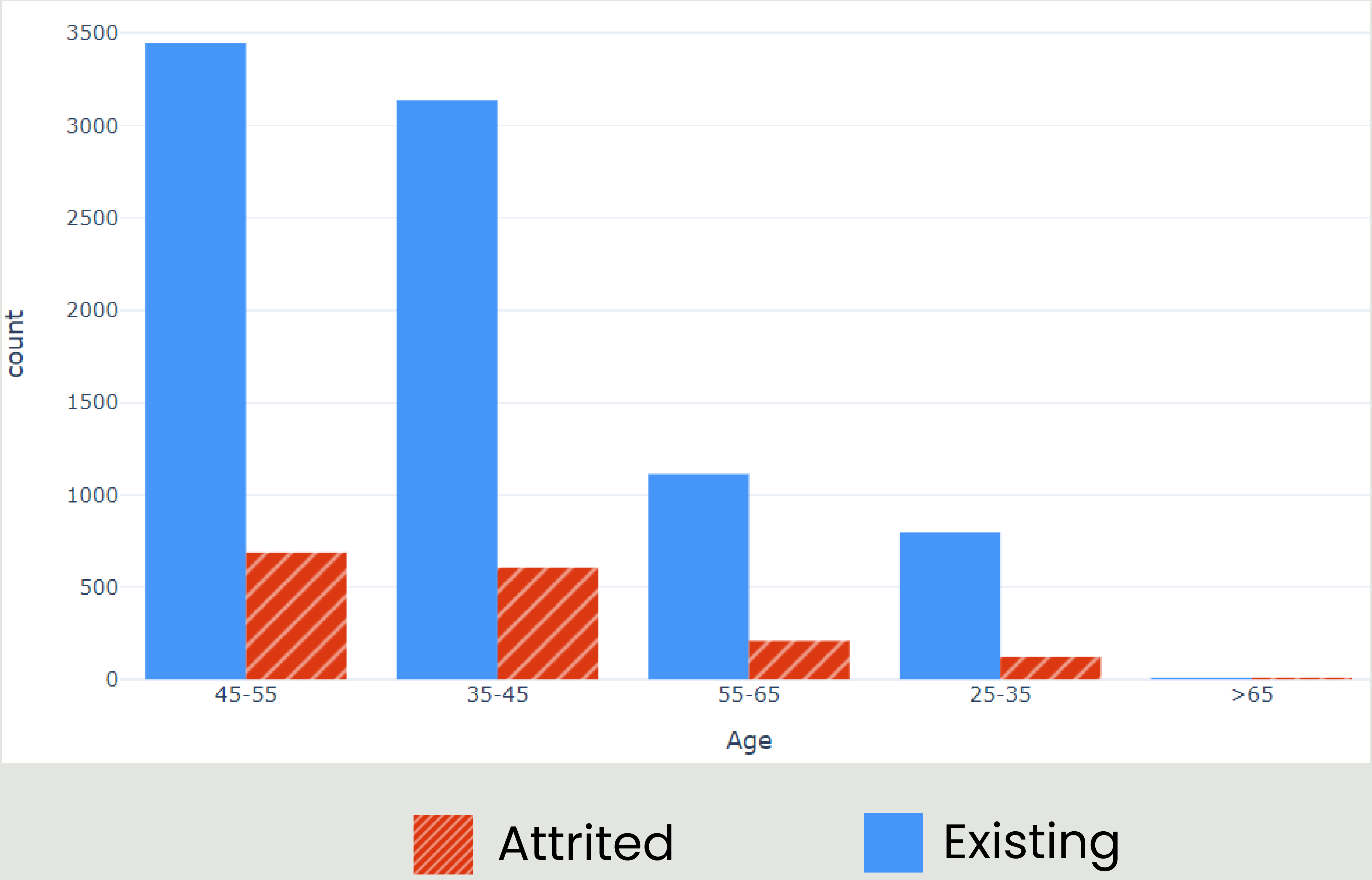



# Exploratory Data Analysis - Demographics

Exploratory data analysis

## Customer Age

- Majority of customers between ages 35 - 55
- Higher proportion of churn in age groups 35-55

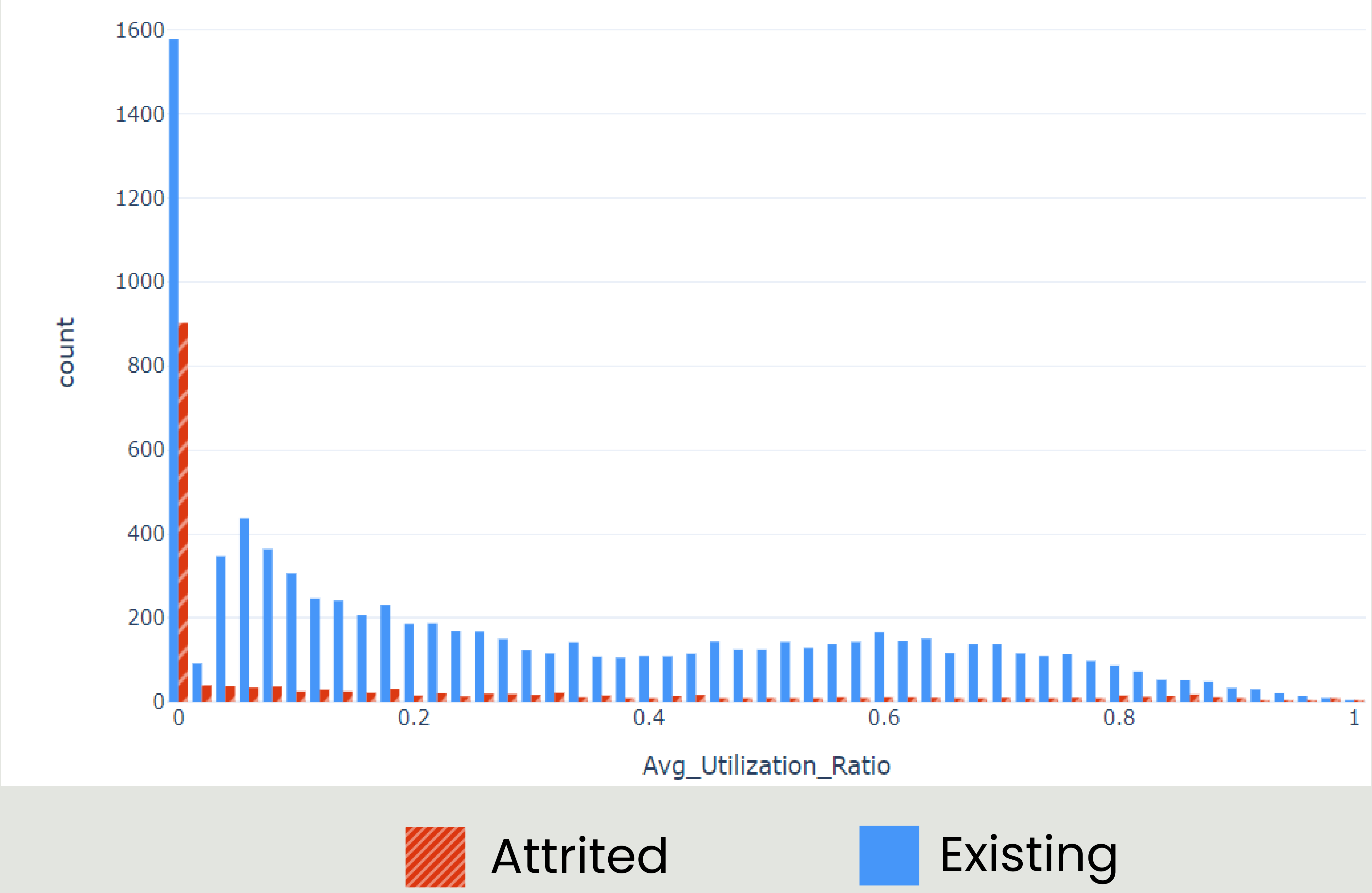


# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

## Utilization Ratio

- High churn at low utilization ratio

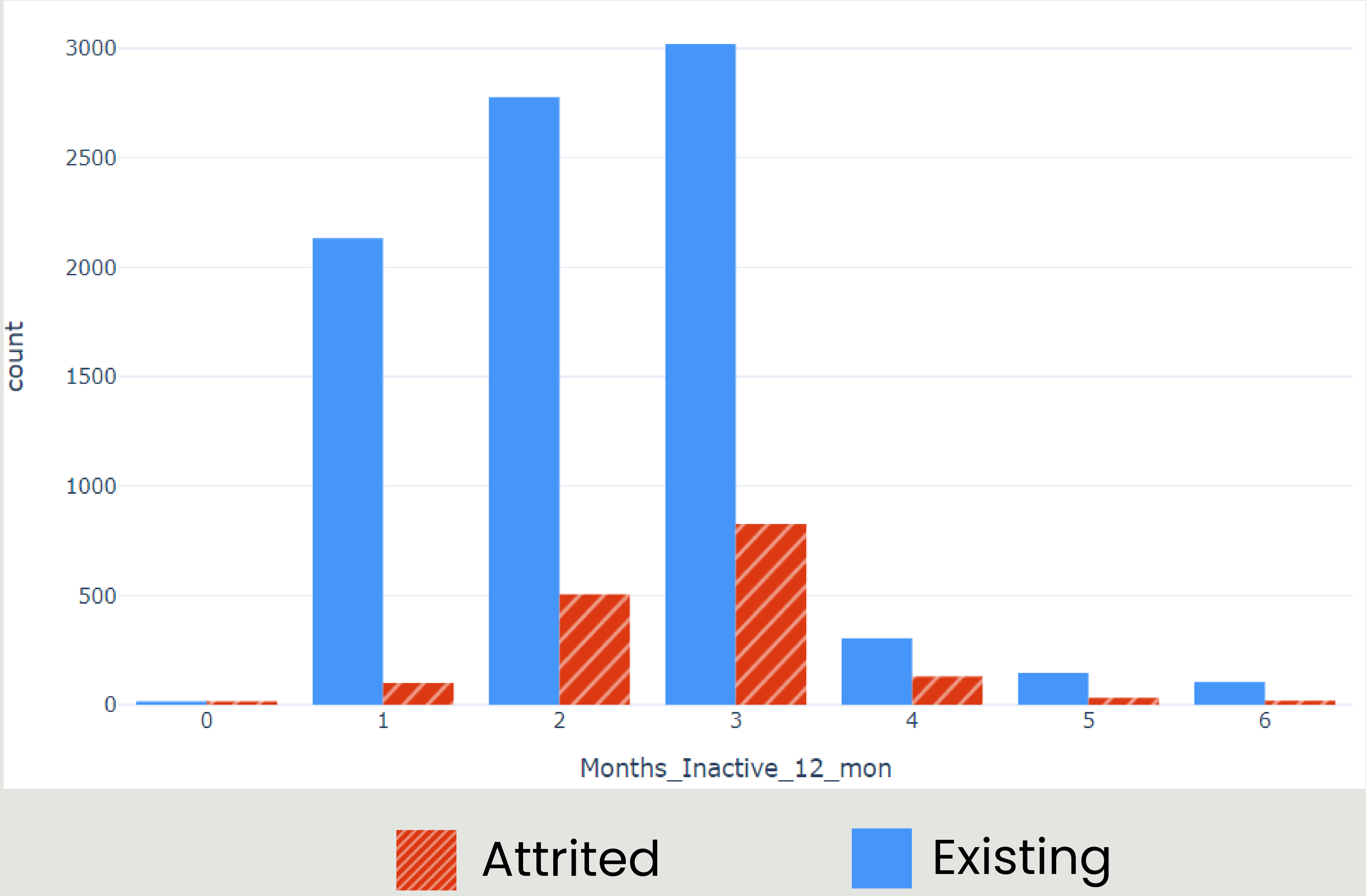


# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

## Months inactive

- High churn at 3 months
- Period of promotion

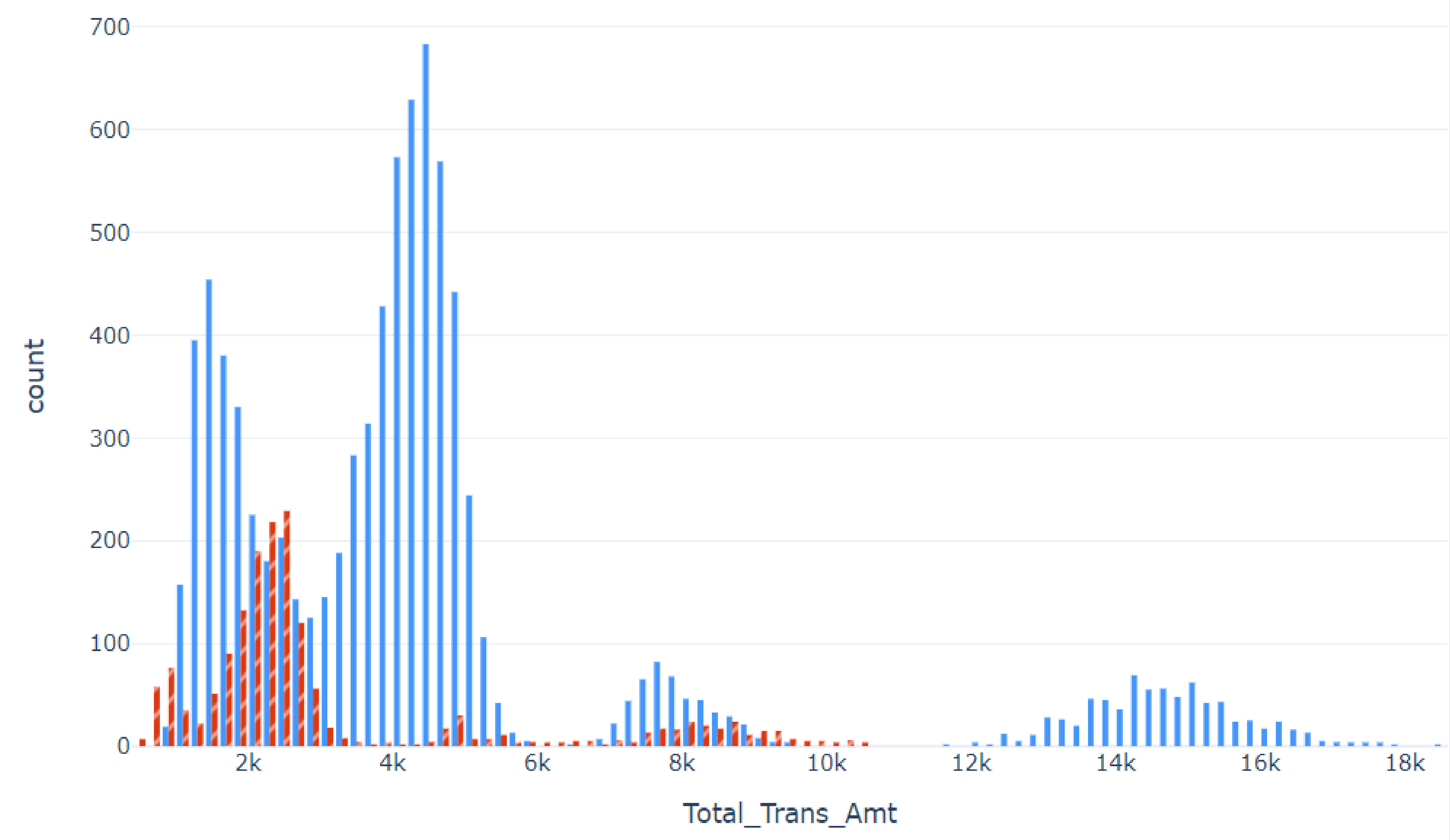


# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

## Transaction Amount

- Higher churn at lower end of transaction amounts from 0k to 3k
- Low churn at high transaction amounts




Attrited

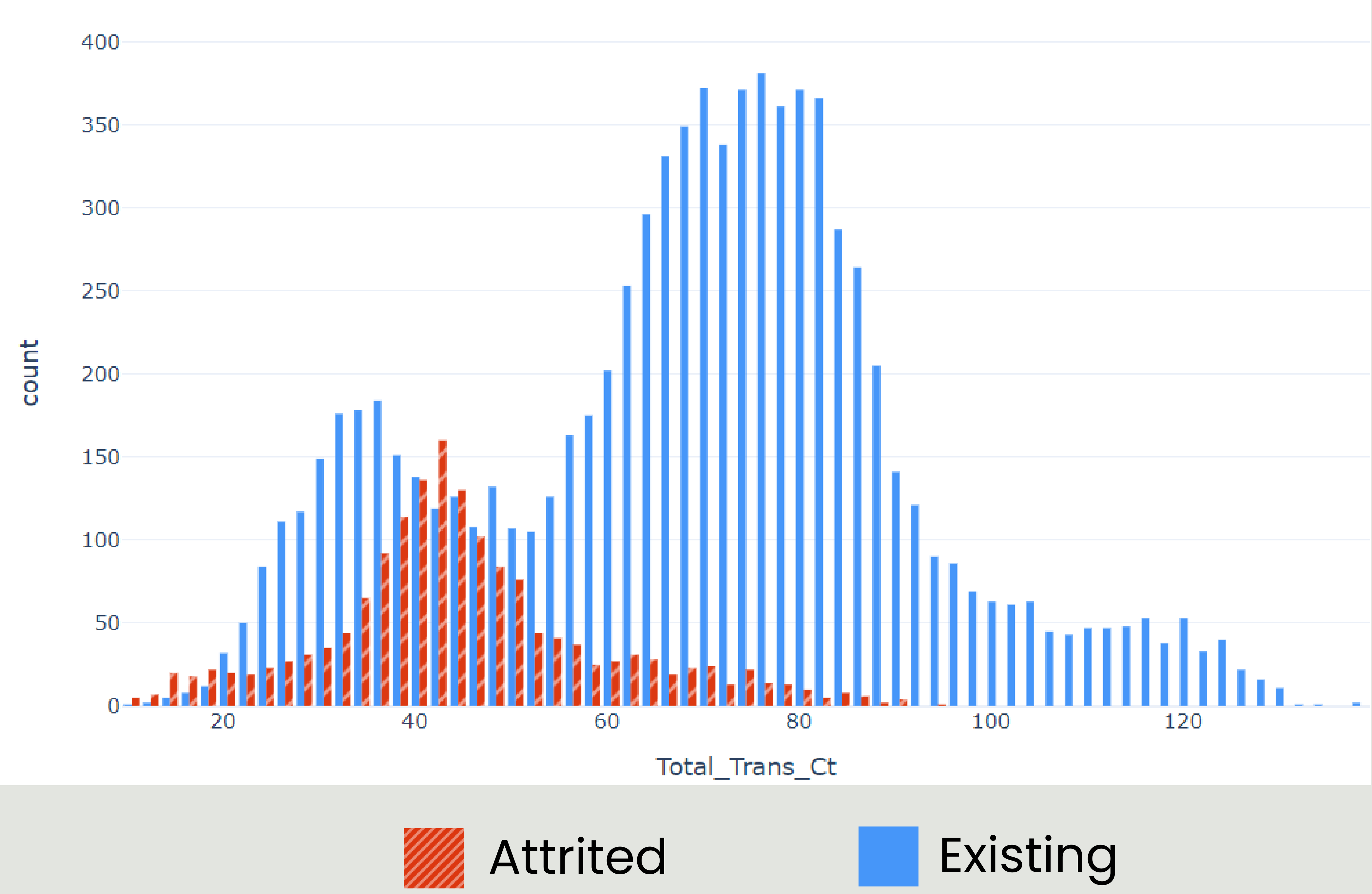
Existing

# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

## Transaction Count

- Higher churn with peak at transaction count of 40-45



# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

## Transaction Count vs Amount

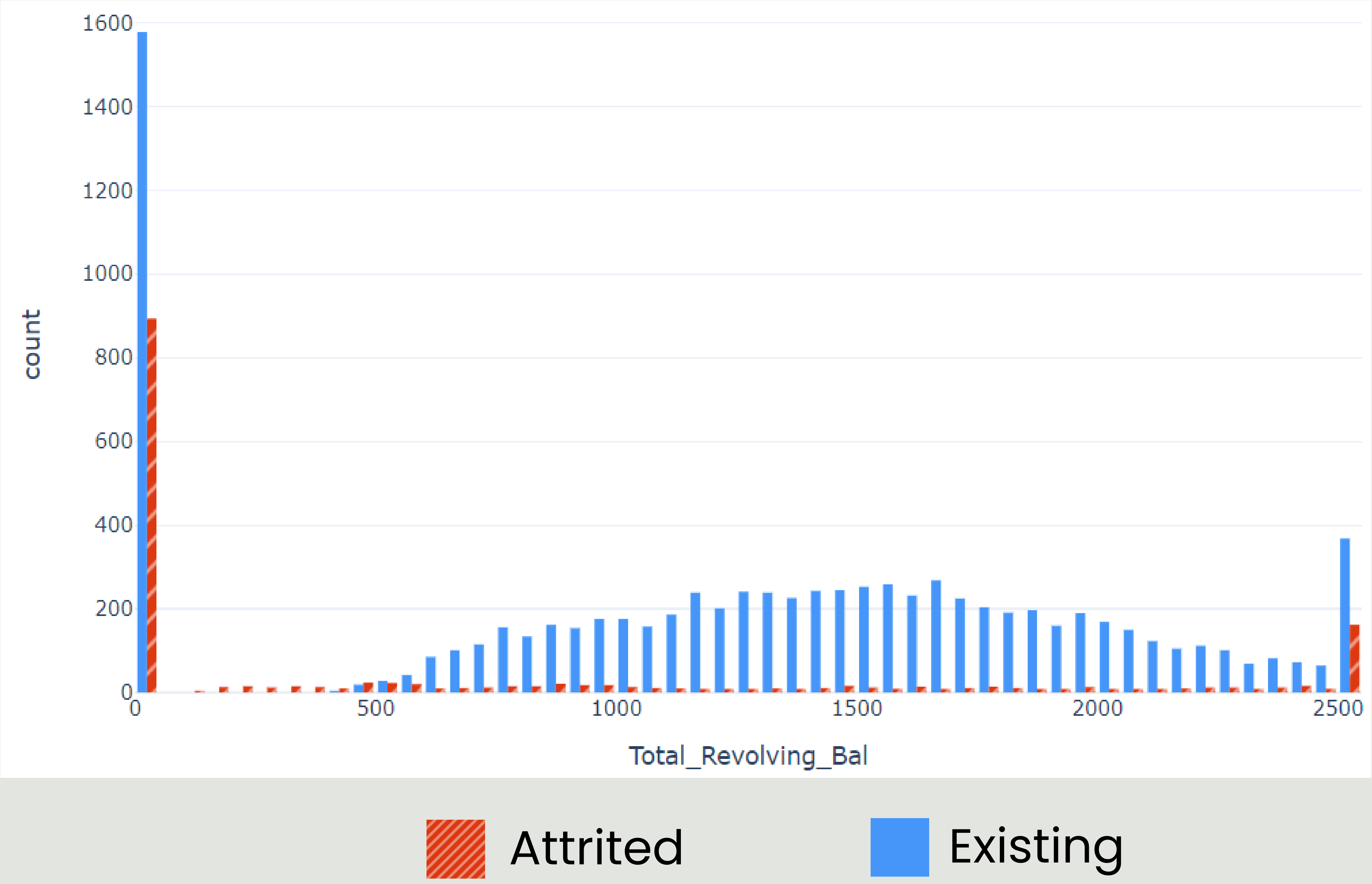
- Positive correlation
- 3 distinct segments
- Low count, low amount
- Moderate count, moderate amount
- High count, high amount

# Exploratory Data Analysis – Spending Behaviour

Exploratory data analysis

**Total Revolving Balance**

- Higher churn at low revolving balance
- Relatively high churn at high revolving balance

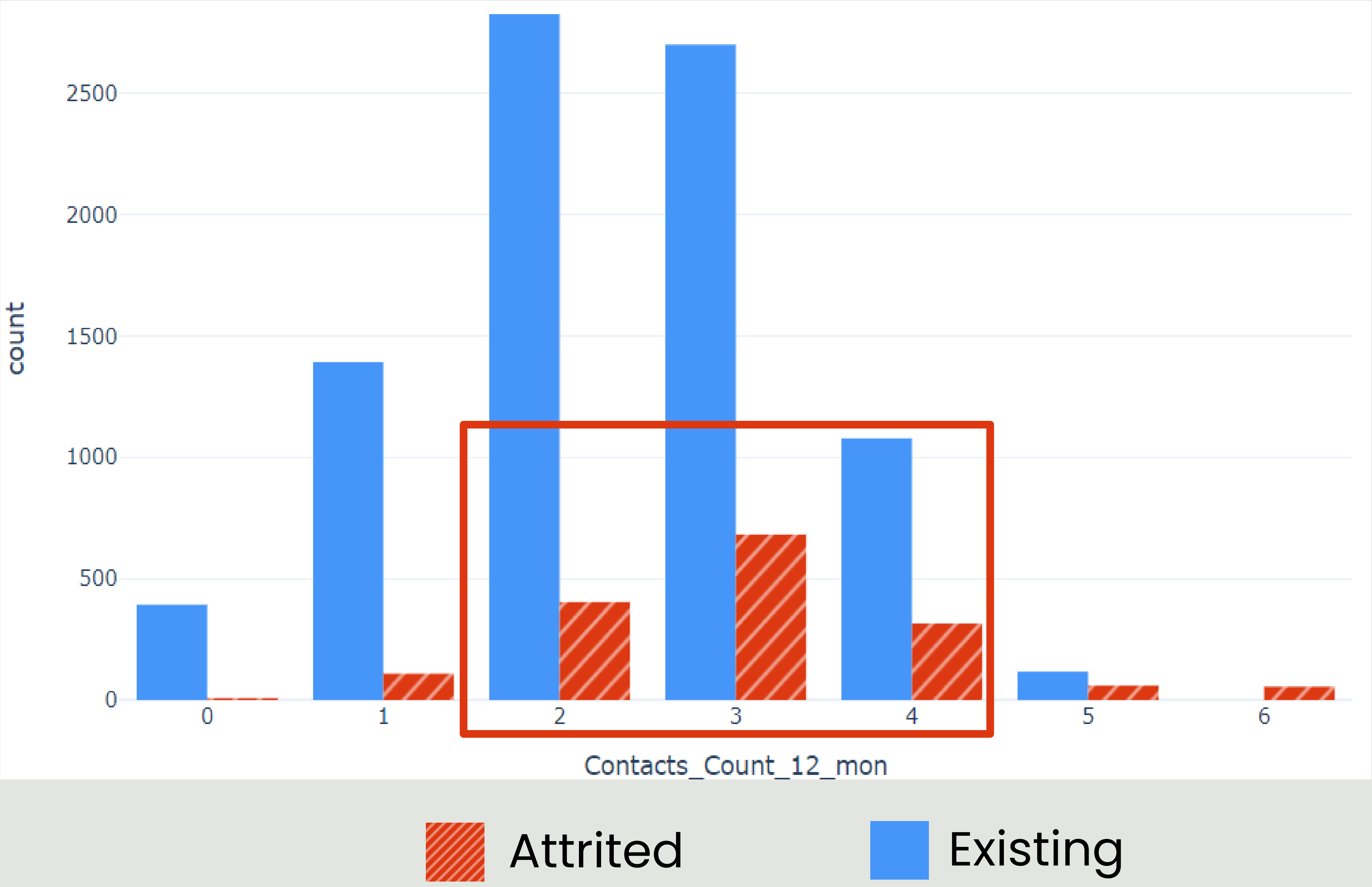


# Exploratory Data Analysis – Relationship

Exploratory data analysis

## Contacts Count

- Churn concentrated with customers who between 2 & 4 contacts



# Clustering

# Clustering


- Key characteristics of high/low churn customers were identified in Exploratory Data Analysis
- Clustering was utilized to identify customer groups with sets of these characteristics

# Selecting Clustering Method

| Clustering Algorithm | Advantages | Disadvantages |
|---|---|---|
| **K-Means Clustering** | • Efficiency and ease of use<br>• Well-suited for large datasets | • Need to specify number of clusters<br>• Sensitive to where centroids are initialised |
| **Hierarchical Clustering** | • Able to visualise using dendrograms<br>• Easy to interpret results and identify meaningful clusters in dendrograms | • Computationally expensive<br>• Not well suited for large datasets |
| **Density-Based Clustering** | • Able to handle clusters of arbitrary shapes | • Not suitable for datasets with varying densities or for datasets with clusters of similar densities.<br>• Choice of distance threshold and minimum number of points required to form a cluster can have a significant impact on the results |

# Process to segment customers

**K-means Clustering**

- Data Preparation e.g. standardize numerical variables
- Run the k-means clustering from 1 to 9

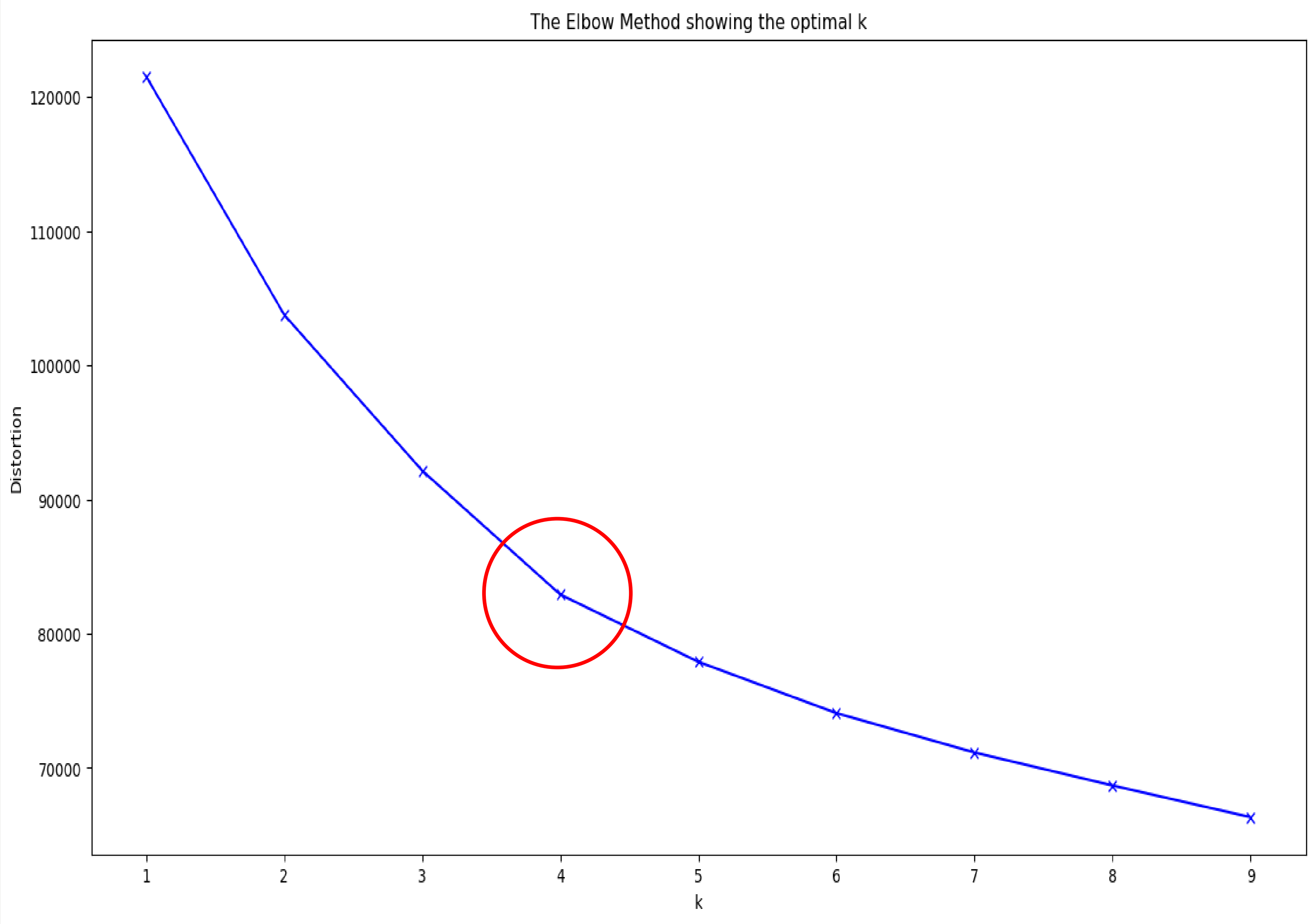
**Elbow Method**

- Leverage Elbow Method to find out the optimal number of clusters for customer segmentation

**Find percentage churn**

- Percentage churn is defined as number of churners / number of customers in the group
- Groups with percentage churn rates will be targeted for solutioning

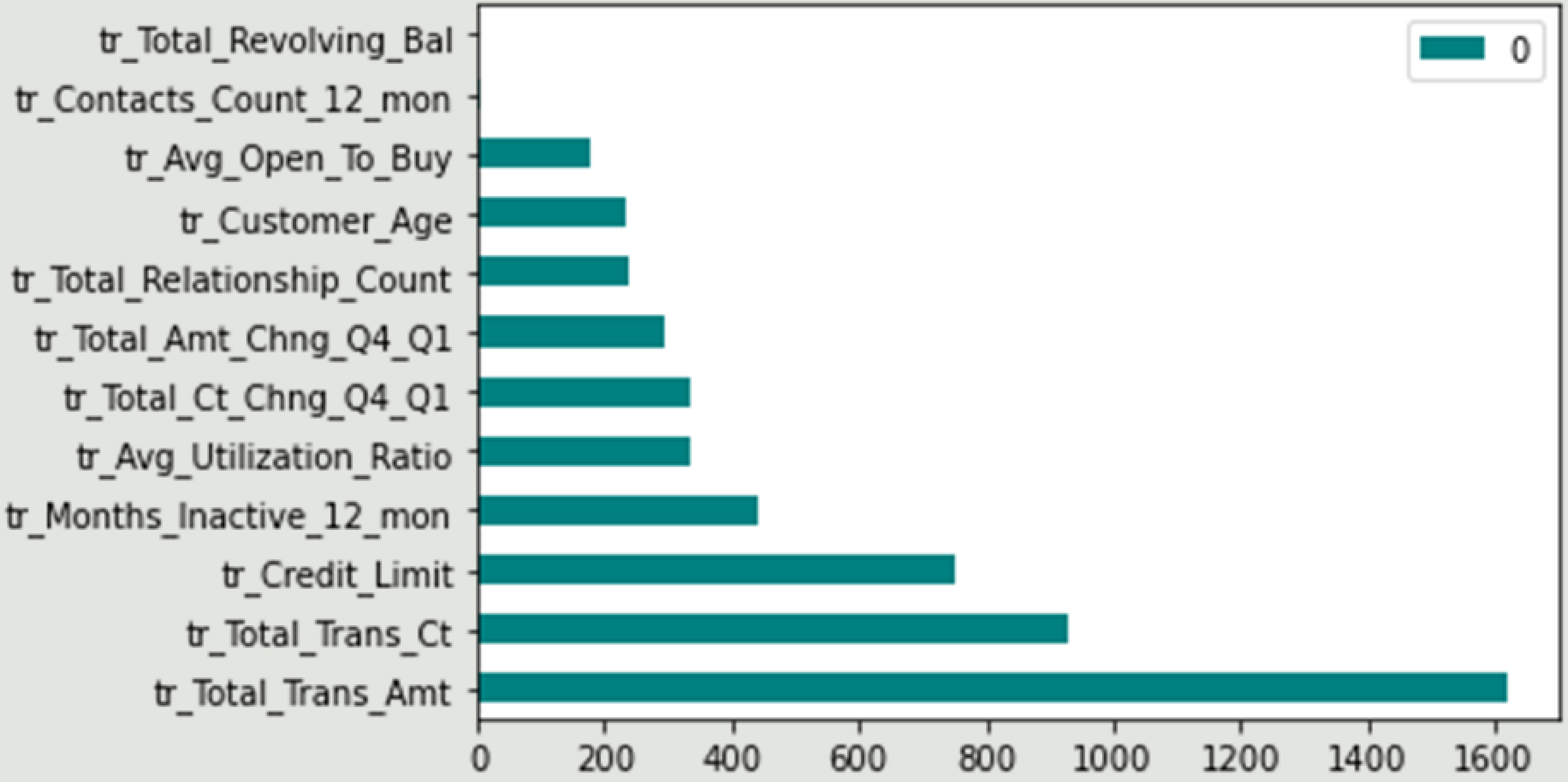
# Findings from Elbow Method




- A technique to find optimal number of clusters in a dataset by plotting SSE against the number of clusters

- Optimal number of clusters is identified by finding the point of maximum curvature in a plot of SSE versus the number of clusters.

# Feature Importance




- Feature importance used to evaluate variables using f-test
- tr_Contacts_Count_12mon and tr_Total_Revolving_Bal have low importance as their scores are negligible, and are removed

# Defining Churn Rates

$$Percentage\ Churn_i = \frac{(Total\ number\ of\ churners)_i}{(Total\ number\ of\ customers)_i}$$

Overall Percentage Churn = 0.16

High churn rate = Percentage Churn of cluster > Overall Percentage Churn

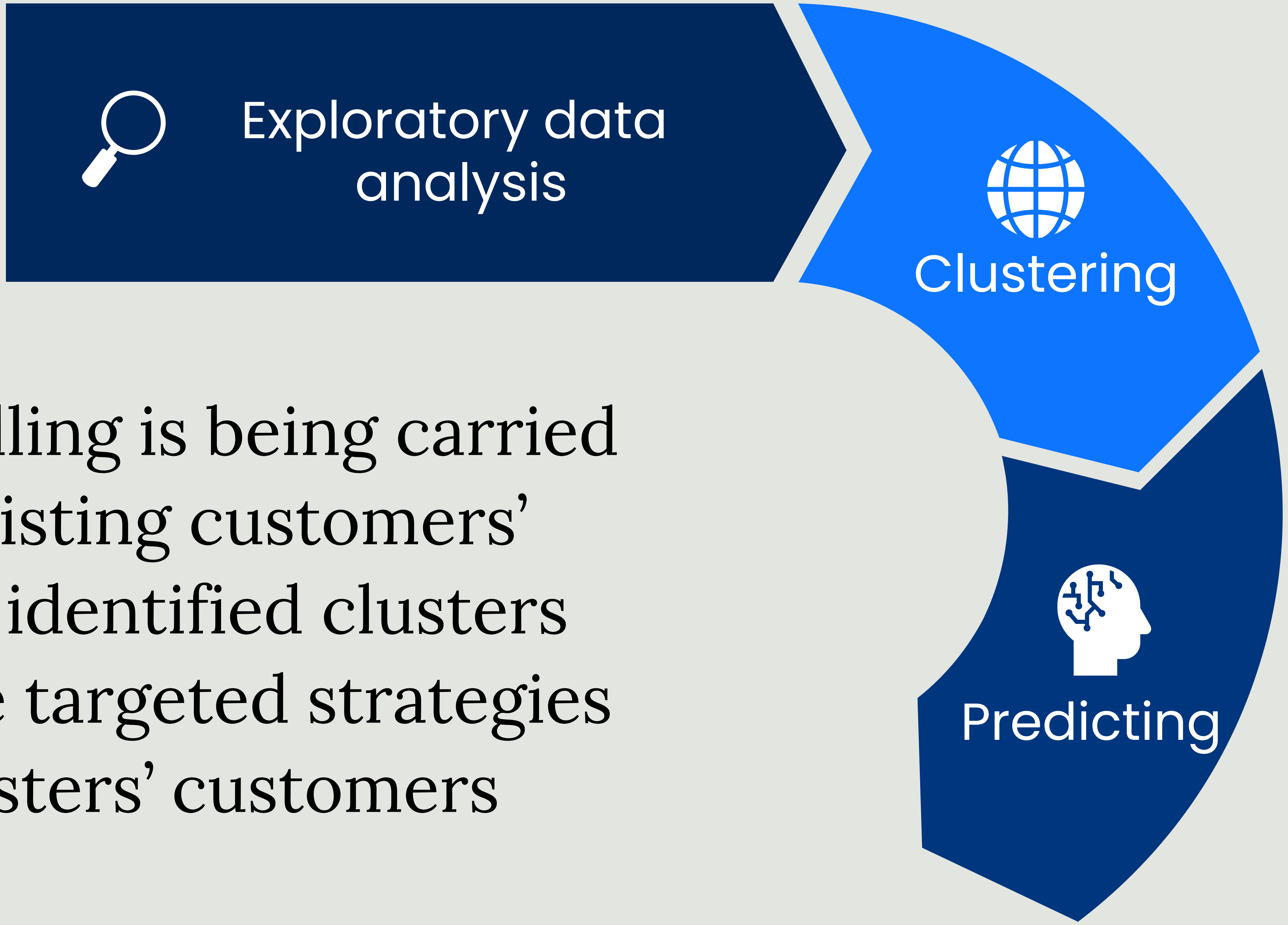
Low churn rate = Percentage Churn of cluster < Overall Percentage Churn

# Customer groups identified after K-means

| Cluster No. | Total | Churners | Percentage Churn | |
|---|---|---|---|---|
| 0 | 1386 | 250 | 0.18 | High Churn |
| 1 | 3385 | 1093 | 0.32 | |
| 2 | 4326 | 215 | 0.050 | Low Churn |
| 3 | 1030 | 69 | 0.067 | |

# Predictive Modelling

# Predictive Modelling




- Predictive Modelling is being carried out to predict existing customers' belonging to the identified clusters
- To be able to use targeted strategies on predicted clusters' customers

# Selected Predictive Model

| Predictive Model | Advantages | Disadvantages |
|---|---|---|
| **Logistic Regression** | • Easy to interpret and explain the model<br>• Can handle binary or multi-class classification problems<br>• Computationally efficient | • Assumes linear relationships between dependent and independent variables<br>• May not work well for complex datasets with non-linear relationships<br>• Limited ability to capture interactions between variables |
| **Random Forest** | • Can handle high-dimensional and complex datasets<br>• Low risk of overfitting (the use of multiple decision trees)<br>• Non-parametric model, ability to capture complex relationships and interactions | • Computationally expensive for large datasets<br>• Difficult to interpret and explain the model compared to single decision tree |
| **Support Vector Machine** | • Effective in high-dimensional spaces<br>• Good at capturing non-linear relationships and interactions<br>• Low risk of overfitting | • Computationally expensive for large datasets<br>• Prone to overfitting with noisy datasets<br>• Difficult to interpret and explain the model |

# Predictive Modelling

**“Confusion Matrix”**

| Predicted Cluster \ Actual Cluster | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 805 | 22 | 0 | 3 |
| 1 | 75 | 1046 | 14 | 12 |
| 2 | 0 | 3 | 311 | 4 |
| 3 | 1 | 2 | 0 | 234 |

# Model Evaluation

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP+TN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision+Recall}}$$

| | Cluster 0 | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|---|
| **Accuracy** | 0.960 | 0.949 | 0.992 | 0.991 |
| **AUC** | 0.949 | 0.953 | 0.977 | 0.962 |
| **Specificity** | 0.985 | 0.931 | 0.997 | 0.999 |
| **Precision** | 0.970 | 0.912 | 0.978 | 0.987 |
| **Recall** | 0.914 | 0.975 | 0.957 | 0.925 |
| **F1-Score** | 0.941 | 0.942 | 0.967 | 0.955 |

# Feature Analysis within Clusters

| Cluster number | Percentage Churn |
|---|---|
| 0 | 0.18 |
| 1 | 0.32 |
| 2 | 0.050 |
| 3 | 0.067 |

## Cluster 0

- Average Utilisation Ratio >= 0.5
- Female
- Credit Limit < \$3,000
- Low Relation Count (3-4)

## Cluster 1

- Average Utilisation Ratio < 0.2
- Equally distributed between 2 genders
- Income Category < \$40,000

## Cluster 2

- Credit Limit > \$7,000
- Average Utilisation Ratio < 0.2
- Male
- Income Category \$80k – \$120k

## Cluster 3

- Transaction Count > 70
- Relationship Count < 3
- Average Utilisation Ratio < 0.2
- Credit Limit > \$7,000
- Male

# Proposed Solutions

# 3 Step Process to Reduce Churn

Card application

1 Customer applies for card

- Collect demographic data
- Monitor spending behavior

Analysing customer data

2 Identify customers profiles most likely to churn

Retention strategy

3 Apply measures to reduce probability that customers would churn

Banks and credit card issuers strategy to reduce churn rates

# Customer Profiling

To retain customer with higher churn rates, or acquire more customers most likely not to churn, we analysed their characteristics within each groups, and tailor a solution specifically to target these groups

0

**Convert**

**Female customers with high churn**

1

**Convert**

**Not gender specific high churn**

2

**Attract**

**High earning male customers with low churn**

# Group 0 : Female with high churn rate

**High churn rate**

### Female dominant

73% of this high churn group comprise of females

### Low credit limit

Credit limit is below $3k

### Low relationship count with issuer

Have relatively lower relationship (3-4) with issuer (i.e other accounts, investments, insurance or other products with the same issuer)

# Group 0 : Retention strategy

## Retention Strategy

### 1 Targeted marketing

Marketing strategy should be tailored for the female gender

### 2 Increase credit limit

To encourage spending, offer flexibility to increase credit limit

### 3 Increase relations

Market other banking products to entrench customer into the issuer's ecosystem

# Group 1 : Not gender specific

**High churn rate**

### Not gender specific

This group could either be male or female

### Low utilisation rate

Has low utilization rates with lower than 50 transactions per month

### Relatively lower spending

Spends less than $500 dollars per month on the credit card

### Relatively lower income

Earning less than $40k per year

# Group 1 retention strategy

**Value**

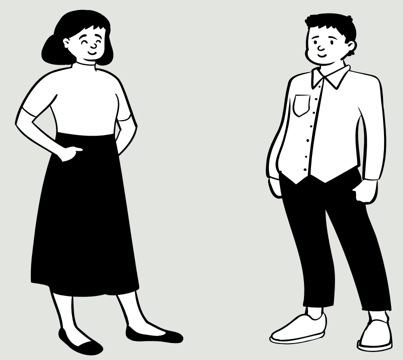
As this group have lower income, provide greater value by lowering annual fees

**Incentivize**

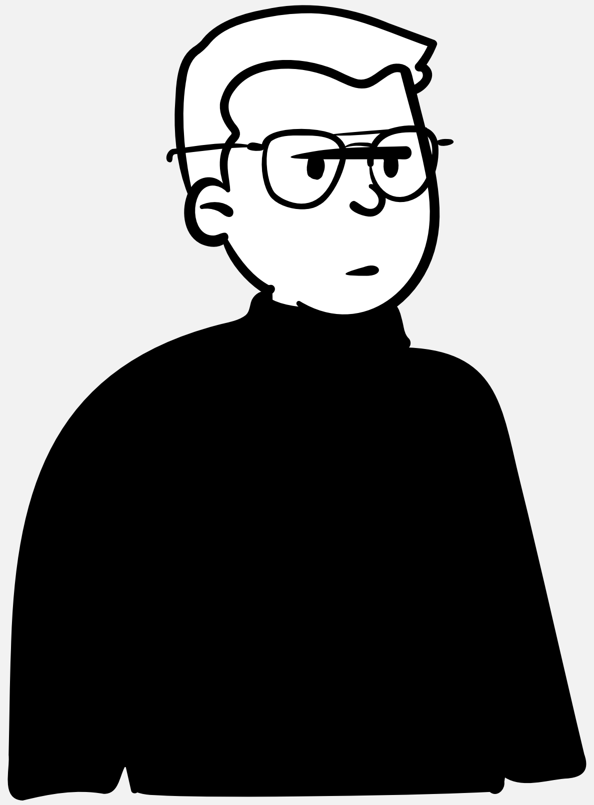
To increase transaction rates, provide additional incentive to spend, such as increasing transaction count (i.e higher miles per dollar, vouchers, higher cashback)

**Target**

As this group could either be male or female, targeted marketing strategy should be gender neutral



# Group 2 : Male with low churn rate



**Low churn rate**

### Male dominant

83% of this low churn group comprise of males

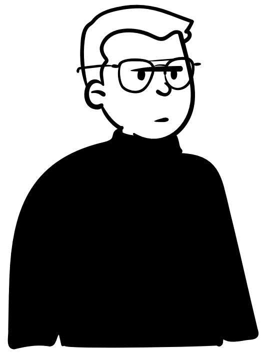
### High income

Has >$80k annual income

### Low existing relations with issuer

This group has 3-4 existing relations with the issuer

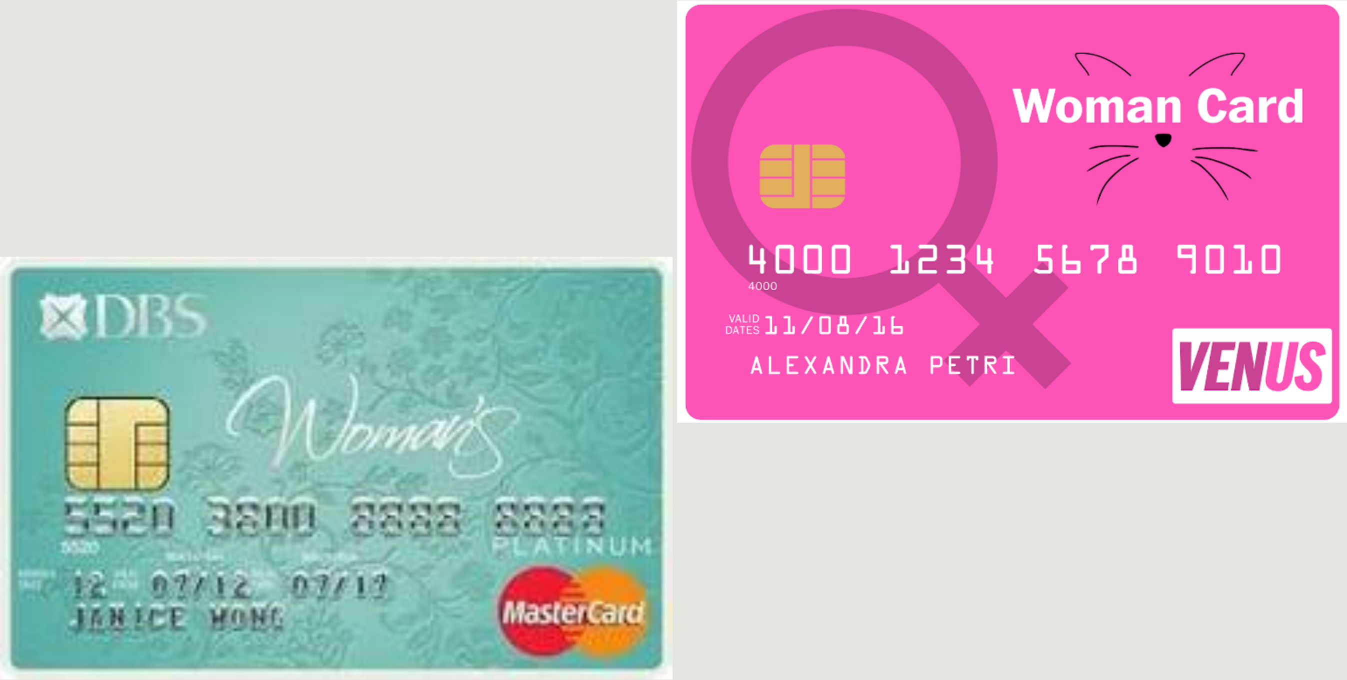
# Group 2 : Acquisition strategy



As an issuer you know...

- Existing customer data as they have established links with the issuer from other products sold
- Bank issuers should identify this profile from existing database and apply targeted marketing to acquire them
- Offer referral / sign up bonus to attract new customers in the same network as the target customer profile

# Examples of gender marketing




Women’s credit card


Advertisements targeting men

# Closed Loop Solution

1. Update customer groups for existing customers

    - Rerun the prediction model to observe if any customers have switched groups after implementation of recommendations

2. Compare ‘before and after’ data for existing customers

    - Collect spending data, before and after implementation of recommendations
    - Should be regularly monitored, and adjust incentives when necessary

# Future Work

1. Improving the clustering model

   - Explore other clustering models such as Neural Networks, Agglomerative Clustering, Gaussian Mixture Models that are able to handle categorical data

2. Improving the predictive model

   - Reducing the number of variables sampled at each split
   - Rank features based on importance before forwarding to the machine learning model

# Future Work

## 3 Using a larger dataset

- Low attrition count presented in dataset – 16%
- Customer transaction history and communication preference
- That contains geolocation data (e.g. residential location, work location)
- That contains demographic data (e.g. employment information)

# THANK YOU

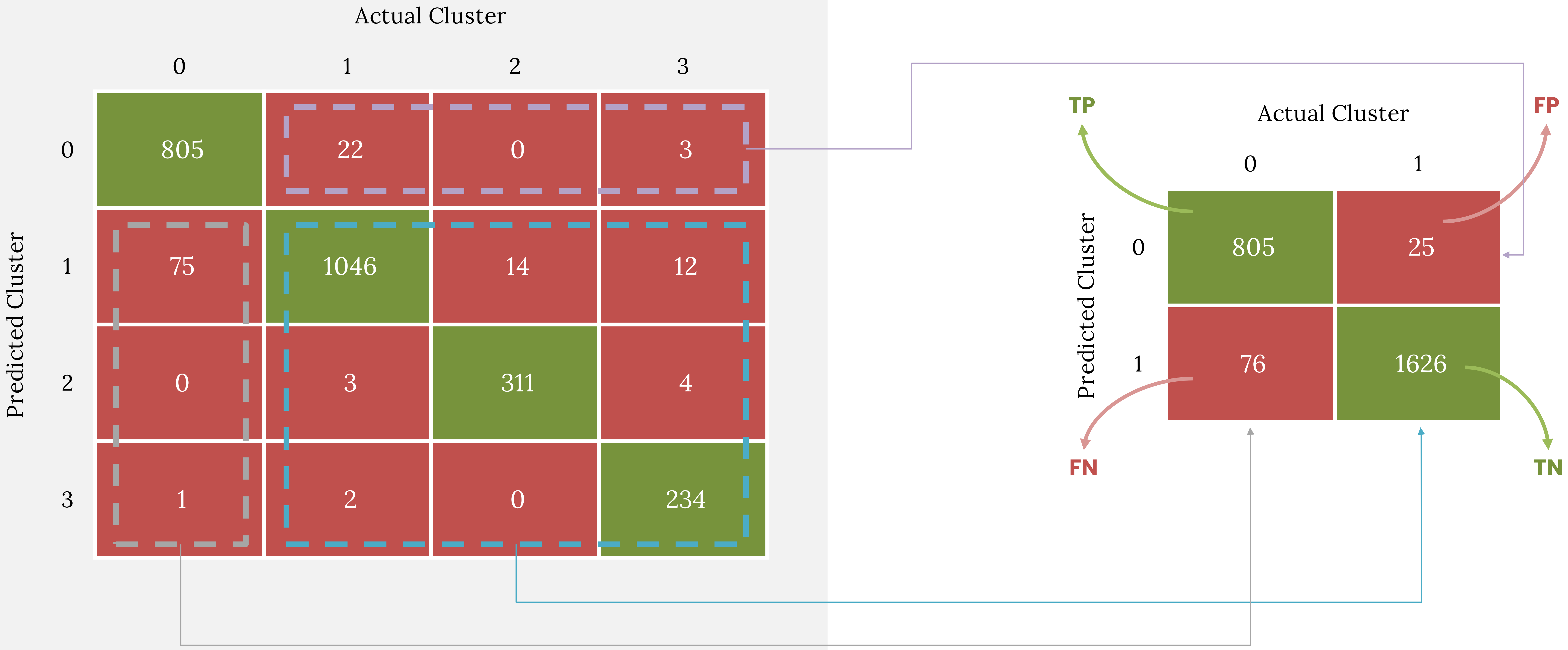
# ANNEXES



1. Conversion of confusion matrix

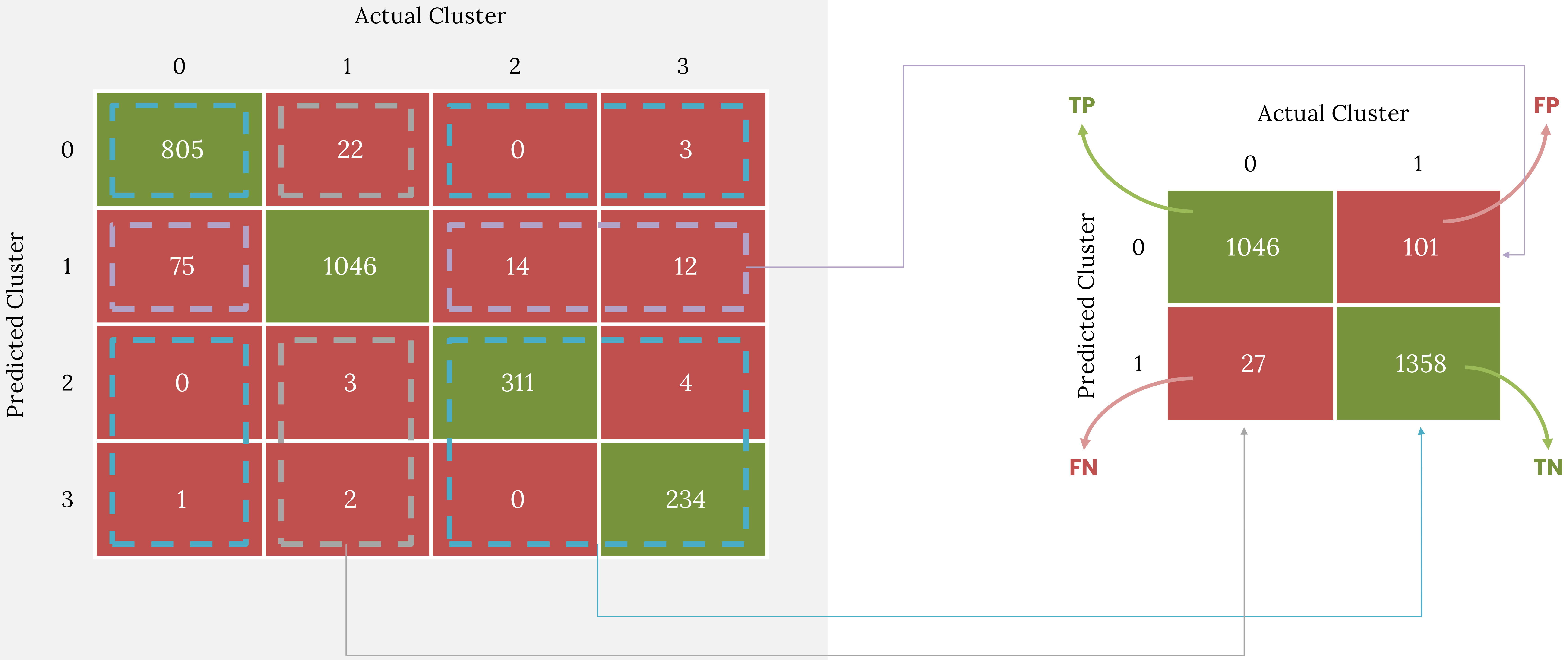
# Converting to One-vs-All Matrix (Cluster 0)

| | |
|---|---|
| **Specificity** | 0.958 |
| **Precision** | 0.970 |
| **Recall** | 0.914 |
| **F1-Score** | 0.941 |



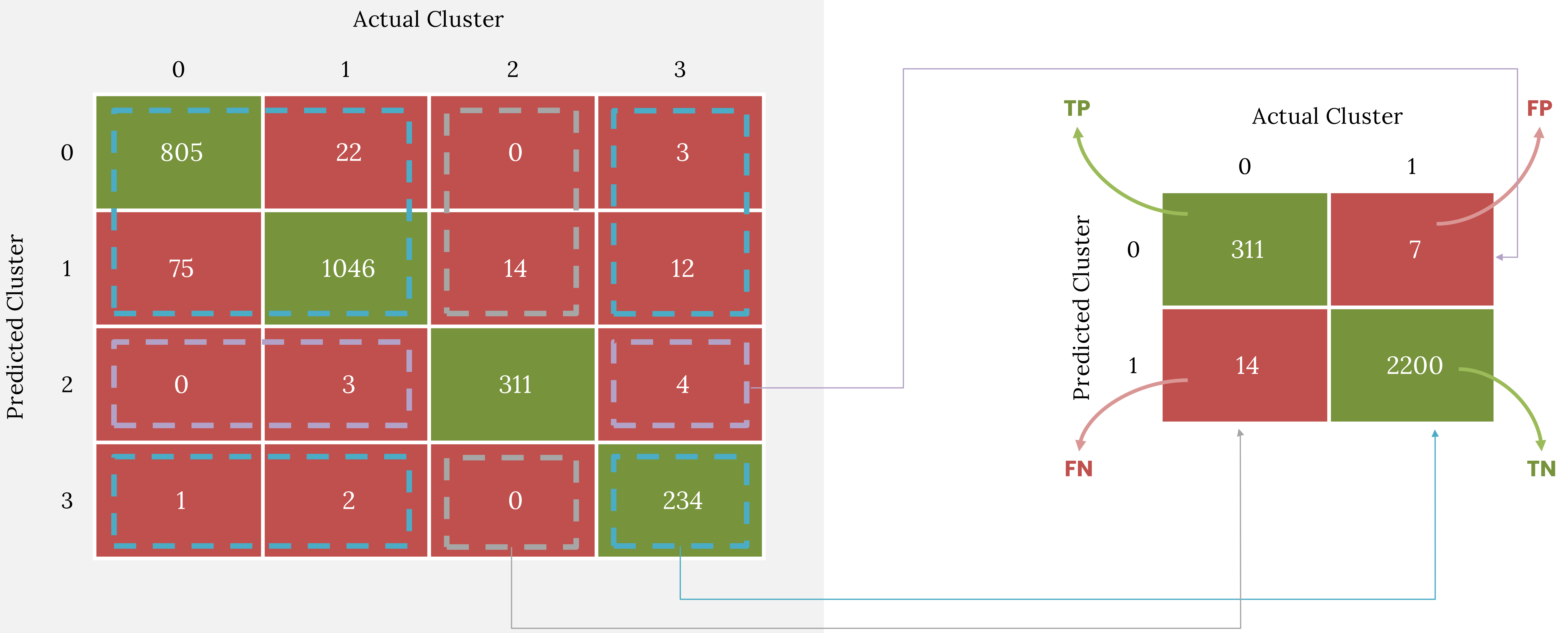
# Converting to One-vs-All Matrix (Cluster 1)

| | |
|---|---|
| **Specificity** | 0.931 |
| **Precision** | 0.912 |
| **Recall** | 0.975 |
| **F1-Score** | 0.942 |



# Converting to One-vs-All Matrix (Cluster 2)

| | |
|---|---|
| **Specificity** | 0.997 |
| **Precision** | 0.978 |
| **Recall** | 0.957 |
| **F1-Score** | 0.967 |



# Converting to One-vs-All Matrix (Cluster 3)

| | |
|---|---|
| **Specificity** | 0.999 |
| **Precision** | 0.987 |
| **Recall** | 0.925 |
| **F1-Score** | 0.955 |