Credit Card Customer Attrition

Gokul Govindasamy Sutharsan

2022-09-22

BUSINESS TASK

Credit card attrition usually hurts the cards business's financial statements, resulting in revenue and profit loss and asset/balance deterioration. Numerous factors have contributed to account attrition, such as dissatisfaction with customer support, the card's pricing structure falling short of the customers' expectations, and more lucrative offers. These situation demands a need for analysis of customers data to analyse the various features that affects customer attrition.

PREPARING THE DATA

```
install.packages('tidyverse')
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.0      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Importing Data

```
credit_data = read.csv("BankChurners.csv")
```

```
head(credit_data)
```

```
##   CLIENTNUM    Attrition_Flag Customer_Age Gender Dependent_count
## 1 768805383 Existing Customer           45      M               3
## 2 818770008 Existing Customer           49      F               5
## 3 713982108 Existing Customer           51      M               3
## 4 769911858 Existing Customer           40      F               4
## 5 709106358 Existing Customer           40      M               3
## 6 713061558 Existing Customer           44      M               2
##   Education_Level Marital_Status Income_Category Card_Category Months_on_book
## 1     High School        Married     $60K - $80K          Blue             39
## 2        Graduate         Single  Less than $40K          Blue             44
## 3        Graduate        Married    $80K - $120K          Blue             36
## 4     High School        Unknown  Less than $40K          Blue             34
## 5      Uneducated        Married     $60K - $80K          Blue             21
```

```
## 6       Graduate        Married     $40K - $60K          Blue             36
##   Total_Relationship_Count Months_Inactive_12_mon Contacts_Count_12_mon
## 1                        5                      1                     3
## 2                        6                      1                     2
## 3                        4                      1                     0
## 4                        3                      4                     1
## 5                        5                      1                     0
## 6                        3                      1                     2
##   Credit_Limit Total_Revolving_Bal Avg_Open_To_Buy Total_Amt_Chng_Q4_Q1
## 1        12691                 777           11914                1.335
## 2         8256                 864            7392                1.541
## 3         3418                   0            3418                2.594
## 4         3313                2517             796                1.405
## 5         4716                   0            4716                2.175
## 6         4010                1247            2763                1.376
##   Total_Trans_Amt Total_Trans_Ct Total_Ct_Chng_Q4_Q1 Avg_Utilization_Ratio
## 1            1144             42               1.625                 0.061
## 2            1291             33               3.714                 0.105
## 3            1887             20               2.333                 0.000
## 4            1171             20               2.333                 0.760
## 5             816             28               2.500                 0.000
## 6            1088             24               0.846                 0.311
##   Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educatio
## 1
## 2
## 3
## 4
## 5
## 6
##   Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educatio
## 1
## 2
## 3
## 4
## 5
## 6
```

```
colnames(credit_data)
```

```
##  [1] "CLIENTNUM"
##  [2] "Attrition_Flag"
##  [3] "Customer_Age"
##  [4] "Gender"
##  [5] "Dependent_count"
##  [6] "Education_Level"
##  [7] "Marital_Status"
##  [8] "Income_Category"
##  [9] "Card_Category"
## [10] "Months_on_book"
## [11] "Total_Relationship_Count"
## [12] "Months_Inactive_12_mon"
## [13] "Contacts_Count_12_mon"
## [14] "Credit_Limit"
## [15] "Total_Revolving_Bal"
## [16] "Avg_Open_To_Buy"
```

```
## [17] "Total_Amt_Chng_Q4_Q1"
## [18] "Total_Trans_Amt"
## [19] "Total_Trans_Ct"
## [20] "Total_Ct_Chng_Q4_Q1"
## [21] "Avg_Utilization_Ratio"
## [22] "Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educa
## [23] "Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educa
```

PROCESSING THE DATA

```
colnames(credit_data)
```

```
##  [1] "CLIENTNUM"
##  [2] "Attrition_Flag"
##  [3] "Customer_Age"
##  [4] "Gender"
##  [5] "Dependent_count"
##  [6] "Education_Level"
##  [7] "Marital_Status"
##  [8] "Income_Category"
##  [9] "Card_Category"
## [10] "Months_on_book"
## [11] "Total_Relationship_Count"
## [12] "Months_Inactive_12_mon"
## [13] "Contacts_Count_12_mon"
## [14] "Credit_Limit"
## [15] "Total_Revolving_Bal"
## [16] "Avg_Open_To_Buy"
## [17] "Total_Amt_Chng_Q4_Q1"
## [18] "Total_Trans_Amt"
## [19] "Total_Trans_Ct"
## [20] "Total_Ct_Chng_Q4_Q1"
## [21] "Avg_Utilization_Ratio"
## [22] "Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educa
## [23] "Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Educa
```

Checking Duplicate entries in CLIENTNUM

```
sum(duplicated(credit_data['CLIENTNUM']))
```

```
## [1] 0
```

ANALYSING DATA

Getting the summary of data

```
summary(credit_data)
```

```
##    CLIENTNUM         Attrition_Flag      Customer_Age      Gender
##  Min.   :708082083   Length:10127       Min.   :26.00   Length:10127
##  1st Qu.:713036770   Class :character   1st Qu.:41.00   Class :character
##  Median :717926358   Mode  :character   Median :46.00   Mode  :character
##  Mean   :739177606                      Mean   :46.33
##  3rd Qu.:773143533                      3rd Qu.:52.00
##  Max.   :828343083                      Max.   :73.00
##  Dependent_count Education_Level    Marital_Status     Income_Category
```

```
##  Min.   :0.000   Length:10127       Length:10127       Length:10127
##  1st Qu.:1.000   Class :character   Class :character   Class :character
##  Median :2.000   Mode  :character   Mode  :character   Mode  :character
##  Mean   :2.346
##  3rd Qu.:3.000
##  Max.   :5.000
##  Card_Category      Months_on_book  Total_Relationship_Count
##  Length:10127       Min.   :13.00   Min.   :1.000
##  Class :character   1st Qu.:31.00   1st Qu.:3.000
##  Mode  :character   Median :36.00   Median :4.000
##                     Mean   :35.93   Mean   :3.813
##                     3rd Qu.:40.00   3rd Qu.:5.000
##                     Max.   :56.00   Max.   :6.000
##  Months_Inactive_12_mon Contacts_Count_12_mon  Credit_Limit
##  Min.   :0.000          Min.   :0.000         Min.   : 1438
##  1st Qu.:2.000          1st Qu.:2.000         1st Qu.: 2555
##  Median :2.000          Median :2.000         Median : 4549
##  Mean   :2.341          Mean   :2.455         Mean   : 8632
##  3rd Qu.:3.000          3rd Qu.:3.000         3rd Qu.:11068
##  Max.   :6.000          Max.   :6.000         Max.   :34516
##  Total_Revolving_Bal Avg_Open_To_Buy Total_Amt_Chng_Q4_Q1 Total_Trans_Amt
##  Min.   :   0        Min.   :    3   Min.   :0.0000       Min.   :  510
##  1st Qu.: 359        1st Qu.: 1324   1st Qu.:0.6310       1st Qu.: 2156
##  Median :1276        Median : 3474   Median :0.7360       Median : 3899
##  Mean   :1163        Mean   : 7469   Mean   :0.7599       Mean   : 4404
##  3rd Qu.:1784        3rd Qu.: 9859   3rd Qu.:0.8590       3rd Qu.: 4741
##  Max.   :2517        Max.   :34516   Max.   :3.3970       Max.   :18484
##  Total_Trans_Ct   Total_Ct_Chng_Q4_Q1 Avg_Utilization_Ratio
##  Min.   : 10.00   Min.   :0.0000      Min.   :0.0000
##  1st Qu.: 45.00   1st Qu.:0.5820      1st Qu.:0.0230
##  Median : 67.00   Median :0.7020      Median :0.1760
##  Mean   : 64.86   Mean   :0.7122      Mean   :0.2749
##  3rd Qu.: 81.00   3rd Qu.:0.8180      3rd Qu.:0.5030
##  Max.   :139.00   Max.   :3.7140      Max.   :0.9990
##  Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Education_
##  Min.   :0.0000077
##  1st Qu.:0.0000990
##  Median :0.0001815
##  Mean   :0.1599975
##  3rd Qu.:0.0003373
##  Max.   :0.9995800
##  Naive_Bayes_Classifier_Attrition_Flag_Card_Category_Contacts_Count_12_mon_Dependent_count_Education_
##  Min.   :0.00042
##  1st Qu.:0.99966
##  Median :0.99982
##  Mean   :0.84000
##  3rd Qu.:0.99990
##  Max.   :0.99999
```

Average age of Attrited Customer = 46
Average age of Existing Customer = 46.
Average Months on Book = 36 and min = 13.
Average Months inactive = 2 months and max = 6 months.
Average number of time the customer has contacted customer service = 2 and max = 6. Av-

erage revolving balance = 1163 and Max 2517.

```
c =credit_data%>%
  count(Attrition_Flag)
```

Total cutomers = 10127.
Existing customers = 8500.
Customer attrited in the past 12 months = 1627.
% of customers attrited in the past 12 months = 16.07

Finding number of male and female attriters

```
female_attrited = credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer', Gender =="F")
count(female_attrited)
```

```
##     n
## 1 930
```

Number of Female Attriter = 930
Number of Male Attriter = 697

ANALYSING THE CARD TYPE

```
credit_data%>%
  count(Card_Category)
```

```
##   Card_Category    n
## 1          Blue 9436
## 2          Gold  116
## 3      Platinum   20
## 4        Silver  555
```

Blue card holders = 9436
Silver card holders = 555
Gold card holders = 116
Platinum card holders = 20

Rate of Attrition by card type

```
attrition_by_card = credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer')%>%
  group_by(Card_Category)%>%
  count(Card_Category)
names(attrition_by_card)[2] = 'Number of customers'
attrition_by_card
```

```
## # A tibble: 4 x 2
## # Groups:   Card_Category [4]
##   Card_Category `Number of customers`
##   <chr>                         <int>
## 1 Blue                           1519
## 2 Gold                             21
## 3 Platinum                          5
## 4 Silver                           82
```

Blue attrition rate = 1519/9436 = 16.1% attrited, 58.6% of attrited customers are female.
Silver attrition rate = 82/116 = 14.7% attrited, 34.14% of attrited customers are female.
Gold attrition rate = 21/116 = 18.10% attrited, 38.1% of attrited customers are female.
Platinum attrition rate = 5/20 = 25% attrited, 80% of attrited customers are female.

Rate of Attrition by Income range

Number of attrited customer in each income category

```
income_category = credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer')%>%
  count(Income_Category)
```

Total number of customers in each income category

```
total_income_category = credit_data%>%
  count(Income_Category)
```

Renaming Column names

```
names(total_income_category)[2] <- 'Total customers'
```

```
names(income_category)[2] <- 'Attrited cusotmer'
```

Merging income category and total income category

```
attrition_by_income = merge(income_category,total_income_category, by = 'Income_Category')
```

Identifying number of men and women attrited in each income range

```
attrition_by_gender =  credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer', Gender == 'F')%>%
  count(Income_Category)
attrition_by_gender
```

```
##   Income_Category   n
## 1     $40K - $60K 166
## 2  Less than $40K 582
## 3         Unknown 182
```

```
names(attrition_by_gender)[2] <- 'Attrited Females'
```

Merging attrition_by_gender with attrition_by_income

```
attrition_by_income = merge(attrition_by_income, attrition_by_gender, by = 'Income_Category', all.x = Tl
```

```
attrition_by_income[is.na(attrition_by_income)] <-0
```

```
attrition_by_income
```

```
##   Income_Category Attrited cusotmer Total customers Attrited Females
## 1         $120K +               126             727                0
## 2     $40K - $60K               271            1790              166
## 3     $60K - $80K               189            1402                0
## 4    $80K - $120K               242            1535                0
## 5  Less than $40K               612            3561              582
## 6         Unknown               187            1112              182
```

Finding the number of men in each category

```
attrition_by_income['Attrited Males'] = attrition_by_income["Attrited cusotmer" ]-attrition_by_income['
```

Identifying the percentage of attrition in each income category

```
attrition_by_income['Percetage of Attrition'] = attrition_by_income["Attrited cusotmer" ]/attrition_by_i
```

Finding the the average numeber of times customers in each income category contacted customer care

```
call_mean_income = credit_data%>%
  group_by(Income_Category)%>%
  summarise(call_mean = mean(Contacts_Count_12_mon))
```

ANALYSIS OF RELATIONSHIP COUNT

```
attrited_relation = credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer')%>%
  count(Total_Relationship_Count)
```

```
existing_relation = credit_data%>%
  filter(Attrition_Flag == 'Existing Customer')%>%
  count(Total_Relationship_Count)
```

```
names(attrited_relation)[2] = 'Attrited Customer'
```

```
names(existing_relation)[2] = 'Existing Customer'
```

```
relationship_count = merge(attrited_relation,existing_relation, by = 'Total_Relationship_Count')
```

```
relationship_count['Percentage of Attrition'] = relationship_count['Attrited Customer']/relationship_cou
relationship_count
```

```
##   Total_Relationship_Count Attrited Customer Existing Customer
## 1                        1               233               677
## 2                        2               346               897
## 3                        3               400              1905
## 4                        4               225              1687
## 5                        5               227              1664
## 6                        6               196              1670
##   Percentage of Attrition
## 1                34.41654
## 2                38.57302
## 3                20.99738
## 4                13.33729
## 5                13.64183
## 6                11.73653
```

ANALYSING INACTIVE MONTHS

Average number of months attrited customers are inactive = 2.69
Average number of months existing customers are inactive = 2.27

ANALYSING CUSTOMER SUPPORT CONTACT DATA

```
credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer')%>%
```

```
  select(CLIENTNUM, Contacts_Count_12_mon)%>%
  summarise(mean_contace = mean(Contacts_Count_12_mon))
```

```
##   mean_contace
## 1     2.972342
```

```
credit_data%>%
  filter(Attrition_Flag == 'Existing Customer')%>%
  select(CLIENTNUM, Contacts_Count_12_mon)%>%
  summarise(mean_contace = mean(Contacts_Count_12_mon))
```

```
##   mean_contace
## 1     2.356353
```

Average number of times attrited customers call customer support = 2.97
Average number of months existing customers call customer support = 2.36

ANALYSING REVOLVING BALANCE

```
credit_data%>%
  filter(Attrition_Flag == 'Attrited Customer')%>%
  summarise(mean_revolving_balance = mean(Total_Revolving_Bal))
```

```
##   mean_revolving_balance
## 1                672.823
```

```
credit_data%>%
  filter(Attrition_Flag == 'Existing Customer')%>%
  summarise(mean_revolving_balance = mean(Total_Revolving_Bal))
```

```
##   mean_revolving_balance
## 1               1256.604
```

Average Revolving Balance of Attrited Customer = $673
Average Revolving Balance of Existing Customer = $1257

Visualisation done using **Tableau**
Click here to check the presentation