CRM PROJECT REPORT

CUSTOMER DISTRIBUTION AND DEACTIVATION ANALYSIS

Submitted By:

Sherin Joseph

22 / July / 2021

# INTRODUCTION

* Customer churn is one of the major problems with companies.
* The companies implement several strategies to retain their valuable customers.
* Despite the service provided by the company, there are high chances for customer churn due to several factors such as market competition, financial situations and so on.
* In this project, the wireless company would like to use behavioral analysis for 2 years to categorize customers and design strategies that will do to minimize customer attrition and sustain and maximize the activities of the most valuable customers.

# OBJECTIVE

* The objective of Given today's marketplace demands, it is more important than ever for businesses to reduce customer attrition.
* Gain valuable insight on why your relationships ended and what might have been done to preserve them.
* Evaluate specific strengths, weaknesses and gaps within your product line, your client service delivery structure, and your technology platform.
* Based on the insights what need to be done to keep the existing customers to stay in business with the company.

# DATA SET

Importing data to SAS: Text

Description automatically generated

This following are the details of the data set:

Table

Description automatically generated

Table

Description automatically generated

Result:

A picture containing text

Description automatically generated

Graphical user interface, text

Description automatically generatedTable

Description automatically generated

Table

Description automatically generated with medium confidence

**Analysis Requests:**

**1.1** Explore and describe the dataset briefly. For example, is the acct no unique? What

is the number of accounts activated and deactivated? When is the earliest and

latest activation/deactivation dates available? And so on….

* Is the Account Number Unique?

Graphical user interface

Description automatically generated with medium confidence

Table

Description automatically generated

As the total count of observations and the count of discrete account numbers are the same, we can confirm that the Account numbers are unique.

* What is the number of accounts activated and deactivated?

A picture containing text

Description automatically generated

Table

Description automatically generated

* When is the earliest and latest activation/deactivation dates available?

Graphical user interface, text, application

Description automatically generated

Graphical user interface, table

Description automatically generated with medium confidence

* What is the main reason for deactivation?

Text

Description automatically generated with medium confidence

Table

Description automatically generated

From the table, its clear that the main reason for majority of the customers to leave the company is “NEED”.

* How many customers with good and bad credit?

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generatedChart, pie chart

Description automatically generated

There are 69.44% customers with good credit and 30.56% customers with bad credit.

* Find the distribution of rate plans for customers are see the which plan has more customers.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

The “Low” plan has the maximum number of customers.

* Find distribution of dealer type of customers.

Text

Description automatically generated with medium confidence

Text

Description automatically generated

Table

Description automatically generated Chart, bar chart

Description automatically generated

The A1 dealer had maximum customers.

* Find the distribution of Age of Customers.



Graphical user interface, text, application, email

Description automatically generated Chart, histogram

Description automatically generated

A picture containing table

Description automatically generated

* Which province has the maximum and min number of customers?



Text

Description automatically generated Table

Description automatically generatedChart, pie chart

Description automatically generated

From the summary of provincial data, we see that ONTARIO had maximum customers and Quebec has minimum number of customers.

* Find the distribution of Sales among customers.



Text

Description automatically generated Diagram, histogram

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**1.2** What is the age and province distributions of active and deactivated customers?

* Age distribution of active and deactivated customers

Graphical user interface, text, application, email

Description automatically generated

Chart, line chart

Description automatically generatedChart

Description automatically generated

Chart, box and whisker chart

Description automatically generatedTable

Description automatically generated

The mean age of active customers is 47.57 and deactivated customers is 47.67 years.

* Province distribution of active and deactivated customers

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

Chart, bar chart

Description automatically generated

The chart above gives clear picture on the province wise distribution of active and deactivated customers. ONTARIO has maximum active customers and deactivated customers where Quebec has least active customers and deactivated customers.

**1.3** Segment the customers based on age, province, and sales amount:

Sales segment: < $100, $100---500, $500-$800, $800 and above.

Age segments: < 20, 21-40, 41-60, 60 and above.

Table

Description automatically generated

* Customers by AGE segment



Text

Description automatically generated

Table

Description automatically generatedChart, bar chart

Description automatically generated

The above summary shows that majority of customers are in the age group “41 - 60” and least in the age group “<=20”.

* Customers by SALES segment

Text

Description automatically generated with medium confidence

Text

Description automatically generated

Table

Description automatically generatedChart

Description automatically generated

The Sales group “<=$100” have maximum customers and “>=800$” has the least customers.

* Customers by PROVINCE segment



Text

Description automatically generated

Table

Description automatically generated Chart, bar chart

Description automatically generated

The maximum number of customers are in ONTARIO and least in QUEBEC province.

**1.4.** Statistical Analysis:

1) Calculate the tenure in days for each account and give its simple statistics.

Text

Description automatically generatedText

Description automatically generated

Text

Description automatically generated

Chart, histogram

Description automatically generated

The mean of the tenure days of customers is 282.57 days.

2) Calculate the number of accounts deactivated for each month.

Graphical user interface, text, application

Description automatically generated

Chart, bar chart, histogram

Description automatically generated

The maximum deactivations occurred in the month of December and least were in the month of February.

3) Segment the account, first by account status “Active” and “Deactivated”, then by

Tenure: < 30 days, 31---60 days, 61 days--- one years, over one year. Report the

number of accounts of percent of all for each segment.

Text

Description automatically generated

Chart, bar chart

Description automatically generated

The most of deactivated customers had belonged to tenure group “61 days – 1 year”.

Most of active customers are in the tenure group “Over 1 year”.

4) Test the general association between the tenure segments and “Good Credit”

“Rate Plan ” and “Dealer Type.”

* Tenure segments and “Good Credit” - [Ho: They are independent]

Text

Description automatically generated

Table

Description automatically generatedChart, bar chart

Description automatically generated

A chi-square test is done between the Tenure segment and Good Credit. The P-value = 0.0001 < 0.05 and hence we can reject the null hypothesis that they are independent of each other at 5% significance level.

* Tenure segments and “Rate Plan”- [Ho: They are independent]

Text

Description automatically generated

Table

Description automatically generatedChart, bar chart

Description automatically generated

A chi-square test is done between the Tenure segment and Rate Plan. The P-value = 0.0001 < 0.05 and hence we can reject the null hypothesis that they are independent of each other at 5% significance level.

* Tenure segments and “Dealer Type” – [Ho: They are independent]

Text

Description automatically generated

Table

Description automatically generatedChart, bar chart

Description automatically generated

A chi-square test is done between the Tenure segment and Dealer Type. The P-value = 0.0001 < 0.05 and hence we can reject the null hypothesis that they are independent of each other at 5% significance level.

5) Is there any association between the account status and the tenure segments?

Could you find out a better tenure segmentation strategy that is more associated

with the account status?

Graphical user interface, text, application

Description automatically generated with medium confidence

Table

Description automatically generated

P-value in chi-square test is less then 0.05 and hence we could reject their association.

Text

Description automatically generatedChart, histogram

Description automatically generated

The distribution of tenure days shows that the tenure days is more related quadrant wise and hence a new tenure segment is made with Q1, Q2, Q3 and Q4.

Text

Description automatically generated

The chi-square test is again performed.

Table

Description automatically generated

We could see from the result that the Mantel-Haenszel Chi-square value have increased drastically compared to the previous tenure segmentation and hence we could conclude that the Quadrant wise tenure classification has more association with the Status-Type.

Practically, if a customer falls in first quadrant, then special care should be taken to avoid customer churn.

6) Does Sales amount differ among different account status, Good Credit, and

customer age segments?

* Does sales amount differ among different account status?

Graphical user interface, text, application, email

Description automatically generated

Table

Description automatically generated

Checking the Homogeneity by Levene’s Test.

Table

Description automatically generated with medium confidence

From the Levene’s test, The P-value = 0.0520 > 0.05 and hence take normal (Pooled) ttest result.

Graphical user interface, text, application, email

Description automatically generated

Table

Description automatically generated

Pooled P-Val = 0.3845 > 0.05. Hence at 5% significant level we fail to reject null hypothesis that sales amount across account status are the same.

* Does sales amount differ among different Good Credit?

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

Checking the Homogeneity by Levene’s Test.

Table

Description automatically generated

Form the Levene’s test, The P-value = 0.7540 > 0.05 and hence use normal (Pooled) ttest result.

The Tukey plot explains the variance of sales in each group.

Chart, box and whisker chart

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

Pooled t-test P-Val = 0.8844 > 0.05. Hence at 5% significant level we fail to reject null hypothesis that sales amount for Good Credit are the same.

* Does sales amount differ among different Age Segments?

Graphical user interface, text, application, chat or text message

Description automatically generated

Table

Description automatically generated

Checking the Homogeneity by Levene’s Test.

Table

Description automatically generated

The P-Val for Levene’s test = 0.0686 > 0.05 and hence we use the standard one-way ANOVA results i.e., P-Val = 0.4770 is considered.

The Tukey plot explains the variance of sales in each group.

Chart, box and whisker chart

Description automatically generated

Since 0.4770 > 0.05, we conclude that at 5% significant level, we fail to reject the null hypothesis that sales amount for customer age segments are similar

# CONCLUSION

To conclude, the project gave the following insights:

* The main reasons for customer deactivation were “NEED”.
* Majority of the customers have a good credit.
* Customers are more inclined to have “LOW” Rate plan.
* Majority of customers had “A1” as their dealer type.
* ONTARIO has maximum active customers and deactivated customers where QUEBEC has least active customers and deactivated customers.
* Mean age of Deactivated Customers was 47 years.
* Mean Tenure of the customers is 282 days.
* The maximum deactivations occurred in the month of December and least were in the month of February.
* Tenure Segments is not significantly related Good Credit, Rate Plan and Dealer Type.
* The Sales amount does not differ among the account status, Good Credit and Customer Age Segments.

# Recommendation

* When the tenure of customers was segmented quadrant wise, it had more association with Customer status.
* Hence, if a customer is in Q1, more precautions should be taken to avoid the customer churn. There should be policies adopted to give special care and promotions to these customers and by doing so, there is less possibility that they will be deactivated.