

Loyalty Card Membership Challenge: A Study on Membership Churn and their Spending Behaviour

Tong-Ming Lim

Sunway University, 5, Jalan Universiti, Bandar Sunway
47500, Selangor, Malaysia

Angela Siew-Hoong Lee

Sunway University, 5, Jalan Universiti, Bandar Sunway
47500, Selangor, Malaysia

Mark P. C. Chia

SAS Malaysia
Suite 6-3A, Level 6, Menara CIMB
Jalan Stesen Sentral 2, Kuala Lumpur Sentral
50470 Kuala Lumpur, Malaysia.

Wei-Jie Soam

Sunway University, 5, Jalan Universiti, Bandar Sunway
47500, Selangor, Malaysia

Kasper Yuet-Hon Liew

Sunway University, 5, Jalan Universiti, Bandar Sunway
47500, Selangor, Malaysia

ABSTRACT

Understand member spending behaviour and their loyalty is important in all industries. By gaining loyalty from customers and understand how they spend, companies are able to retain their customers, increase their revenue and plan their marketing strategy to continue grow their business in a competitive business ecosystem. This research investigates member spending behaviour and membership churn for a loyalty card company in Malaysia. This research conducts exploratory analysis on three key partners registered with the company to understand their outlets' spending activities and patterns. Meanwhile, this research also model membership churn based on the last 24 months' membership data to identify factors that influence membership churn so that effective strategy can be formulated to retain active members in the company.

Keywords: Customer loyalty, member spending behaviour, loyalty card program, churn model, prediction

INTRODUCTION

Information technology has contributed a lot for marketing managers in terms of creating a new type of customer relationship management strategies (Uncles, et al. 2012). One of the tactics that most of the firms have considered was to establish a customer loyalty card program. By establishing customer loyalty card program to their customers, many attractive incentives and promotions can be planned based on the profiles and purchasing patterns of these customers. Plenty of loyalty program with different sizes and choices have been developed in the past decades. One of the successful stories was the increase airline frequent

flyer programs in the 1980s. Since then, loyalty programs have intensely interlinked with marketing strategies in a new modern way (Gault et al. 2012). Loyalty cards were initially implemented to reward regular customers and to urge them to continue shop in the same store (Arnett, 2013). Companies that provide loyalty cards holds more advantages compare with companies that don't (Acatrinei and Puiu, 2012). The reasons companies offer loyalty cards are because it is easier for them to identify their key customers and to create a deeper relationship with them by understanding their profile and behaviour. By doing this, companies can track their customer shopping patterns.

Organizations always reward loyal customers by providing top-up services or prices with discounts to retain them. In recent years, loyalty rewards programs have become one of the key factors in businesses of many segments. These programs have served as marketing tools to boost up the companies' revenue and sales by encouraging the loyal purchasing behaviour of their loyal customers. Card holders of loyalty card, point's card, rewards card or club card help to profile members' purchasing patterns hence reduce churn and allow ease of segmenting customer based on their needs and buying capability especially in the retail market (Singh & Khan 2012).

Loyalty programs can be used as an incentive by offering perks based on snowballing purchasing over time. Loyalty programs inspire customers to move from inactive or single-period decision making to active or multiple-period decision making. These programs stimulate repeat buying and increase retention rates by offering incentives for customers to buy more often and in higher volumes (Ghaleb Magatef Elham, et al 2015).

Hence, it is important to understand the customers' spending behaviour and factors that predict churn in the loyalty card industry. This research intends to study the factors that result in loyalty cardholders who churn in the data that have been collected for the last 20 years in the company. More than 100 attributes have been provided and churn will be investigated by using 2 data modelling techniques: decision tree and regression model. Several research questions have been defined. These questions include what are the factors that result in churn and what are the modelling techniques that are able to predict churn more reliably.

RELATED WORKS

Loyalty card and Spending Behavior

A loyalty card program is defined as an integrated system of marketing actions that aims to retain customers and increase customer loyalty by developing personalized relationships with them (Sharp & Sharp 1997; Yi & Jeon 2003; Meyer-Waarden 2008). It also acts as popular marketing tools in developing relationships, stimulating product usage and retaining existing prospects and customers (Meyer-Waarden & Benavent 2009). Liu (2007) further explained that loyalty programs is one of the important components of firms' customer relationship management (CRM) strategy and it aims to enhance customer loyalty by rewarding customers for their continuous support and doing business with the firm. By introducing such programs, it enables firms to gain more repeat business and at the same time, attain rich customer data that would be helpful for future CRM effort. For example, Tesco's loyalty program features a business model of program efficiency and a customer model of data-driven knowledge, which allows the firm to undertake tailored strategies and incentives to target on different customer segments (Humby *et al.* 2004).

Buying behavior is the decision processes and acts of people that involved in buying and using towards any products (Sharma 2014). Customer buying behavior refers to the buying behavior of the ultimate customer. There are many factors and characteristics that will influence or encourage the individual in what she is and the customer on her decision-making process,

shopping routines, buying behavior, the brands she purchases or the retailers that she frequently goes. Customer behavior refers to the selection, purchase and consumption of merchandises and services for the fulfillment of their needs and wants (Stavkova et al. 2008). Engel et al, (1986) define customer behavior as “those acts of individuals directly involved in obtaining, using and disposing of economic goods and services, including the decision processes that precede and determine these acts”. The nature of customer choice is too complex and it involves more sophisticated concepts and methods to fully understand, effectively predict and possibly control customer behavior. Instead, psychology, social psychology and sociology are the widely-employed disciplines in this endeavor (Pachauri, 2002).

Given a wide variety of choices and a low switching barrier, it is relatively easy for customers to switch among different firms and this poses significant threats to customer relationships whereby customers are not likely to commit to a single brand or firm (Liu 2007). Loyalty program in this case helps to alleviate this lack of commitment by raising switching costs. Customers tend to focus their purchases in one program in order to maximize the benefits they receive as loyalty programs rewards customers for their repeated patronage (Sharp & Sharp 1997).

Data Quality

Many databases are not error-free and some of them contains a surprisingly large number of error. Arnold (1992) reported that more than 60% of the surveyed organizations have issue regarding data quality. A bank in New York found out that the data within their credit-risk management database were only 60% complete and it requires checking/validation before anyone can use it (Bailey,1983). Meanwhile Wang and Strong (1996) found that a company that has problem accessing all its sales data for a single customer because many different customer numbers were assigned to represent the same customer. To increase data quality, understanding what data quality means to data consumers is vital.

The purpose of improving data quality is to make sure that data can be reused at any time without worrying poor data quality. Motivations for reuse can be diversified, including data verification, reanalysis, meta-analysis, new analysis and reproducing original analysis and results (Wang and Strong, 1996). In order to carry out the data reuse process, the data need to be shared, processed and preserved in a way that ensures they are understandable by every single business unit in an organization and that there is enough information to be understood without the assistance of a subject matter expert (CCSDS,2012).

Customer Churn

Over the past few years, churn prediction has received escalating consideration in the consumer marketing and management research literature. Decision tree is one of the technique used to predict customer churn. Decision tree can be split into classification and regression decision tree. Classification trees are used to predict categorical outcome while regression trees are used to predict continuous outcome. For churn analysis, the outcome of the prediction will be binary. Therefore, a classification tree is used. One of the advantages of decision tree is, it can be easily interpreted because decision tree produce a set of understandable rules (Shaaban et al., 2012). When decision trees do not incorporate a stopping criterion, it is very likely to overfit the training data set. Overfitting occurs when the learning algorithm continues to develop hypotheses that reduce training set error but at the cost of increasing the test set error. There are (2) two ways to avoid overfitting when building a decision tree. Pre-pruning is the first way that stop the tree from growing earlier, before it

perfectly classifies the training data set. Post pruning is the second way where it allows the tree to grow fully first before pruning the tree. The second approach is the more logical approach because it is not an easy task to precisely estimate when to stop growing the tree. In a telecommunication industry, the prominent definition of prepaid churn is based on several successive months with zero usage. For example, a churned customer could be defined as not having any outgoing or incoming calls for the past two months. The definition of zero usage has to be redefined, when a prepaid customer who has already churned, he/she can still have incoming calls from uninformed outsiders. These incoming calls will then be transferred into voicemail that will also be recorded. Therefore, defining zero outgoing calls and zero incoming calls cannot be taken as the optimal decision to decide whether a customer does or does not churn. A new threshold is set to the value of 30 seconds of voicemail per month, which turned out to be a practical value. Therefore, a customer will be confirmed to be churned when he/she has zero usage with the combination of 30 seconds of voicemail per month.

Logistic regression is another technique used to predict customer churn. Logistic regression is a statistical model where the outcome variable is categorical instead of continuous. Logistic regression is widely used in a lot of applications such as agriculture, overweight, obesity, credit scoring, prevention, accident analysis and churn analysis (Nie et al., 2011). According to Nie et al. (2011), there are 6 regression models built to predict credit card churn. The first model built is based on customer information. The second model is built based on customer and basic card information. The third model is built based on the combination of variables from model 1 (customer information) and model 2 (basic card information). Model 4 only takes the transaction information into consideration. The fifth model is built on variables from customer information, card basic information and also risk-related information. The sixth model consists of variables from customer information, card basic information, transaction information and also risk-related information. After evaluating the 6 regression models, the sixth model was determined to be the best model as it best fits the data. The 6 models were evaluated using percentage of Type I error, Type II error, Average error and also misclassification cost. Model 6 appeared to have a 12.43% of Type I error, 18.22% of Type II error, 12.69% of Average error and 80377.25 of misclassification cost which is the lowest compared to the other 5 models (Nie et al., 2011).

Analytical Life Cycle

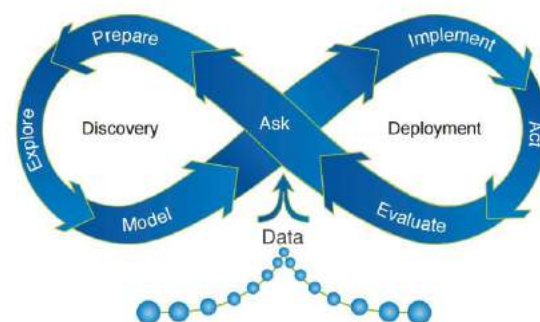


Figure 1: SAS Analytical Life Cycle
Source: SAS (2016)

Figure 1 illustrates the analytical life cycle developed by SAS. Successful organizations identify that analytic models are essential corporate assets that produce and deliver answers to production systems for improved customer relationships, improved operations, increased revenue and reduced risk. SAS developed an iterative analytical life cycle to guide users through the process of going from data to decision. There are eight phases to produce an effective

predictive modelling results.

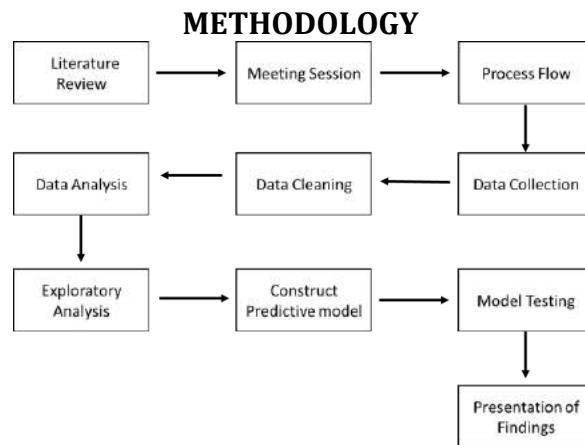


Figure 2: Methodology

Figure 2 illustrates the methodology used in this research. The following sections explain each stage of activities in detail.

LITERATURE REVIEW

A literature review is used to identify various kinds of problems and techniques that have been studied in this research area by previous researchers. The literature review includes data mining, customer buying behaviour, loyalty card, analytic lift cycle and techniques & methods.

Meeting Session

The meeting sessions with the loyalty card company based in Malaysia will be carried out in this research project. In these meeting sessions, domain knowledge is acquired, problems of the industry will be shared and discussed, understanding of databases and tables will take place, updates on the project will be shared and discussed with all members of the project which include supervisors, lecturers from business school and employees from the loyalty card company. The purpose of these meetings is to make sure that the research project is on the right track. A preliminary investigation was conducted to understand the problem scenario faced by the company and the data from the company.

Process Flow

Figure 3 shows the overall chained of research activities needed to achieve the goal of this project. Understanding this process flow is crucial to ensure that everyone is always on the right path with same level of understanding. The DBMS indicates the database to be accessed from the loyalty card company. SQL statement will be written to select important tables and variables from the database. Selected tables will be imported to SAS Enterprise Guide for data quality processing purposes.

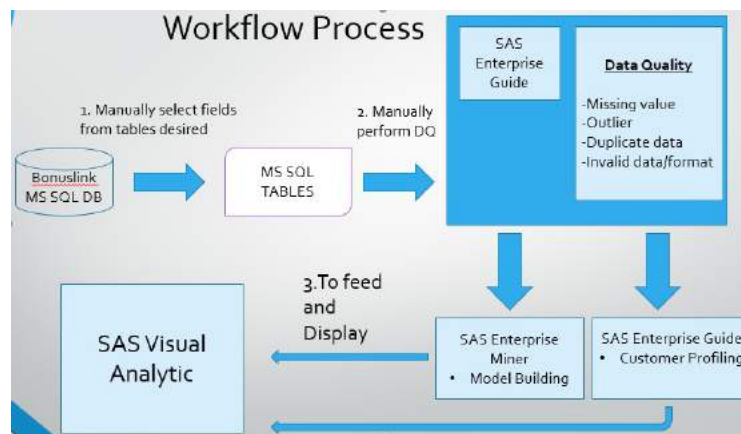


Figure 3: Process Flow

For this part, missing value, outliers, data duplication, invalid data, invalid input format will be identified and removed. Using the dataset provided, predictors and a target variable will be identified to build a predictive model using SAS Enterprise Miner (e-Miner). On the other hand, SAS Enterprise Guide (EG) will be used to perform exploratory analysis. Once this completes, the results will be presented in SAS Visual Analytic for visualisation purposes.

Data Collection

Before data is acquired, an NDA (Non-Disclosure Agreement) form will be signed to make sure all the data given to us will not be disclosed to unauthorized parties. The data contain 9 tables and they contain data from year 2014 and 2015. The data will be used to perform exploratory analysis.

Data Cleaning

SAS EG (Enterprise Guide) are used to clean up the given data. Cleaning up data meaning to detect and remove unnecessary or inaccurate records from the data set. Data cleaning include remove outlier, identify invalid format and variables and others.

Data Analysis

SAS Visual Analytics and SAS Enterprise Guild were the main tools to do analysis, important key matrix will be recognised. For example, the useful information such as customer's age will allow us to know what they usually spend on.

Exploratory analysis

The well-prepared set of variables will be put into SAS Enterprise Guide and SAS Visual Analytic to be analysed in the forms of bar charts and pie charts for better understanding of trends and patterns. This also helps to profile the customers that spend and use the loyalty programme. All the bar charts and pie charts are produced by using data give by the loyalty card company. From the charts, results for each category from different selected loyalty partners are easily shown clearly.

Construct Predictive model

SAS Enterprise Miner is used to build predictive models to forecast and predict customer churn. With this prediction, we will be able to estimate which customers that are about to churn and what appropriate actions should be taken. Factors will be identified at this stage. These independent variables can be used to construct effective marketing activities.

Model Testing

After building the predictive model, tests will be carried out using sample from the dataset to

examine the reliability of the model. If the model is not reliable, corrections will be made to improve the reliability of the model. Several cycles of model building activities will take place until desired objective is achieved.

Presentation of Findings

Results obtained from the SAS Enterprise Guide will then be presented in the form of charts and tables using SAS Visual Analytic for ease of multi-level drill down. A final report comprising all the results and future possible works will be documented at the end of this process.

FINDINGS AND DISCUSSIONS

In this section, findings are presented. Three key loyalty programme partners will be illustrated and discussed. They are from a departmental store, pharmacy chained store and a petrol company. They are chosen because these partners contribute the largest percentage of transactions from the total number of transaction per year. Membership profiling for these partners are illustrated in the first part of the discussion. In Section 4.2, churn model is presented. This is followed by customer spending behaviour in Section 4.3.

Member Spending Outlets Profile

This section briefly highlights the top customers' spending outlets for the three (3) key loyalty programme partners. The exploratory spending analysis provides a good understanding of when and where these loyalty card holders spend in the partners' outlets.

Top 5 outlets in 2014 & 2015 For Departmental Store by-quarterly

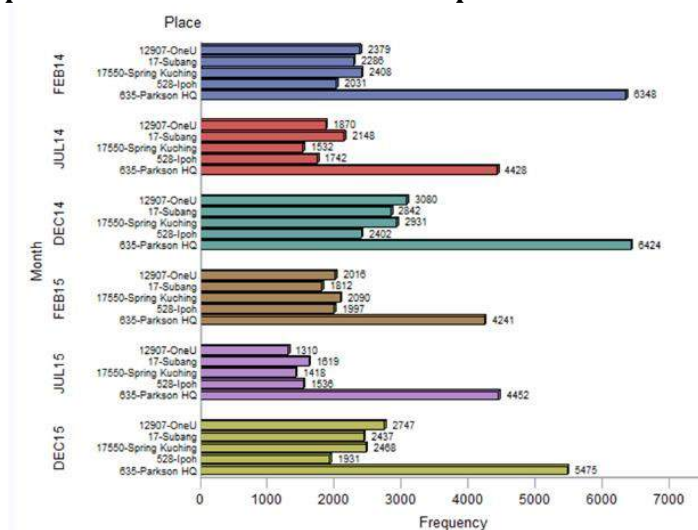


Figure 4: Top 5 transaction frequency outlets.

Figure 4.0 illustrates the top 5 outlets for departmental store in 2014 & 2015 by quarterly. In Figure 4.0, every five outlets is grouped by the high transactional frequency quarterly. Based on 2014 and 2015 data, people tend to shop more during these three months. From the analysis, it shows that the departmental store HQ have the highest frequency of transaction for all the months. For other outlet, the month that has the highest frequency of transaction is on December 2014. This indicates that people shop more during year end promotions but people will also shop more during festive season such as Chinese New Year or Hari Raya.

Top 5 outlets of 2014 & 2015 For Pharmacy chain store by quarterly

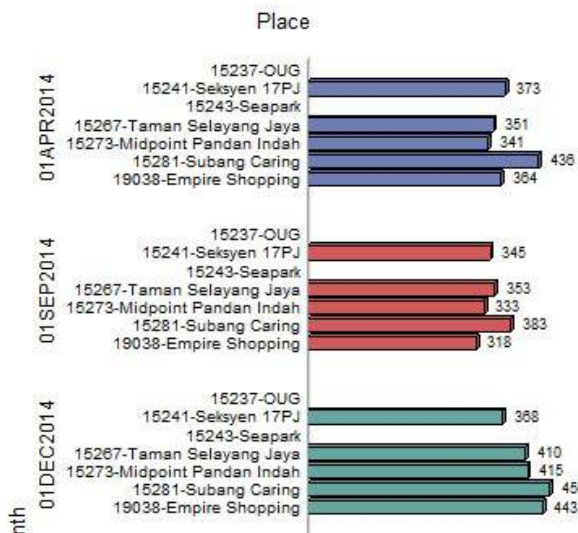


Figure 5: Frequency of top 5 outlet on 2014

According to Figure 5, the five outlets that were grouped by the highest frequency (by transaction) month is shown. Based on 2014 data, people tend to shop more during these three months. From the bar chart, it shows that Subang Pharmacy chain store have the highest frequency of transaction for all the months. For the other outlets, it fluctuates over the three months. This indicates that, people will shop in the pharmacy chain store mostly during end of the year.

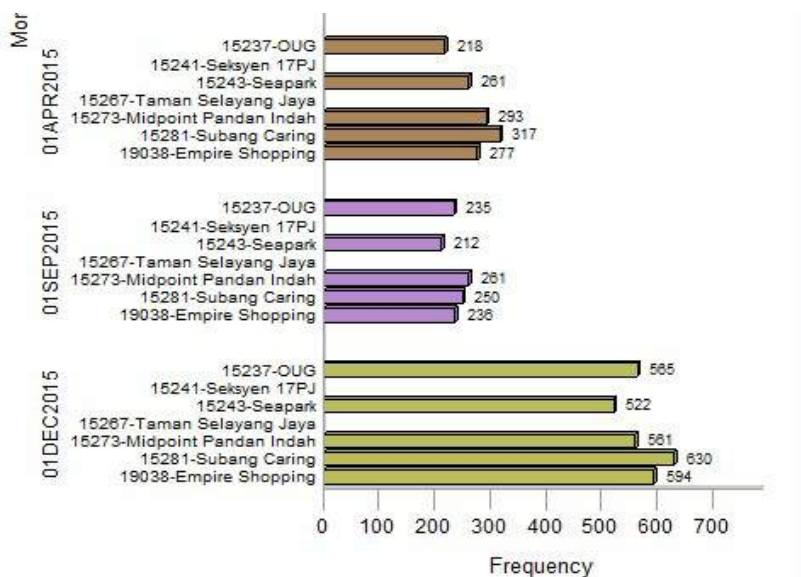


Figure 6: Frequency of top 5 outlet on 2015

In Figure 6, it was found that the five outlets are grouped by the highest transaction frequency month. Based on the 2015 data, card holders tend to shop more during these three months. From the chart, it shows that Subang Pharmacy chain store has the highest transaction frequency for all the months except September 2015. For the other outlets, the highest transaction frequency month is December 2015. This indicates that people will shop in The pharmacy chain store mostly at end of the year.

Top 5 outlets for the petrol company in 2014 & 2015 by-yearly

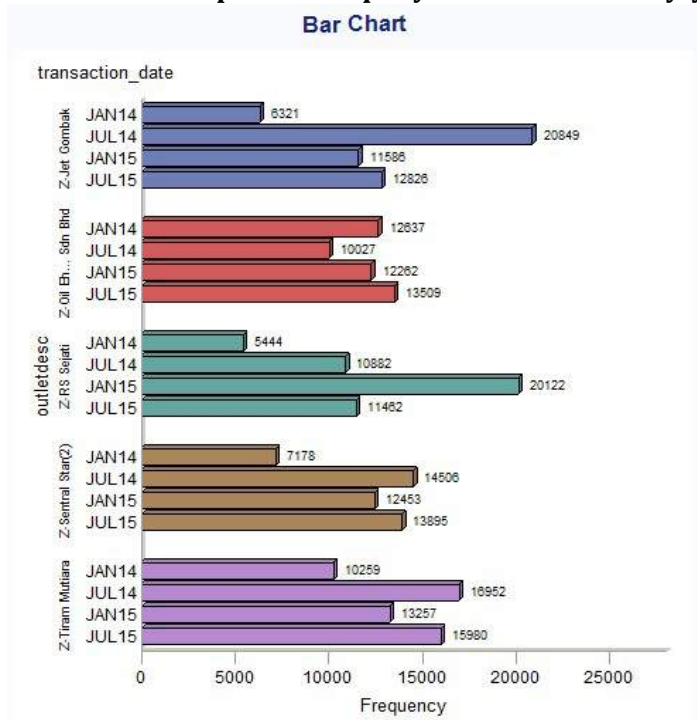


Figure 7: Frequency of top 5 outlet by 6months

From Figure 7, it was found that every five outlets are grouped by-yearly for year 2014 and 2015. Each outlet has different transaction frequency for the first and second half of the year. From the figure, in 2014 data, outlet Jet Gombak shows that the second half of the year is higher than the first half of the year. However, in 2014 data, outlet Oil Ehsan shows that the first half of the year is higher than the second half. This indicates that different The petrol company outlets have different transaction frequency. This might be due to the location of the outlets.

CUSTOMER CHURN MODEL

Churn Model based on Customer Transaction data

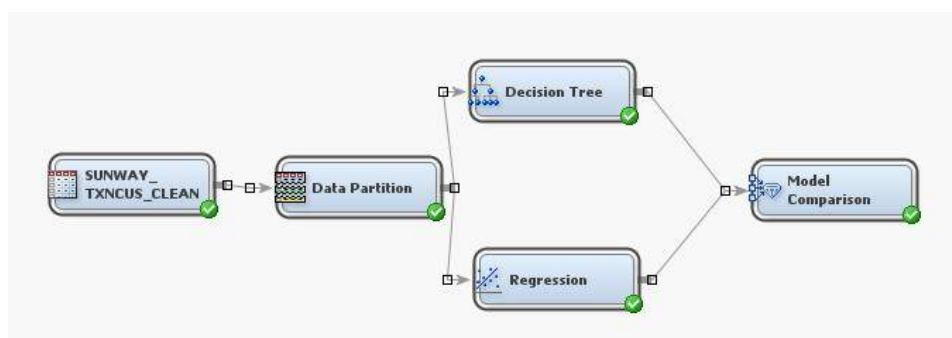


Figure 8: Predictive model for Customer Transaction table.

Figure 8 shows the flow of the predictive model constructed using SAS Enterprise Miner. The first node on the right represents the dataset used. In this case, it is the Customer Transaction table. The next node is the data partition node. This node separates the dataset into 3 parts. The first part is validation; it is used for performance tuning for the model built. The tuning process usually involves selecting among models of different types and complexities. The second part is testing; it is used for model assessment. The third part is training; it is used to

build the models. The data is partition into 60% training, 30% validation and 10% test. The next two nodes are the deductive models developed.

The first node on the top is decision tree. Since the predictive type is estimate, “average square error” method is used in the assessment measure. In order to ensure all the data is used to run this model, a default sample size is changed to “none” representing all data will be used for modelling. As for the second modelling node, regression model is used. Since the target variable is a categorical variable, the regression type used is “logistic regression” instead of linear regression. The selection model for this model is “stepwise” approach.

Decision Tree Churn Model

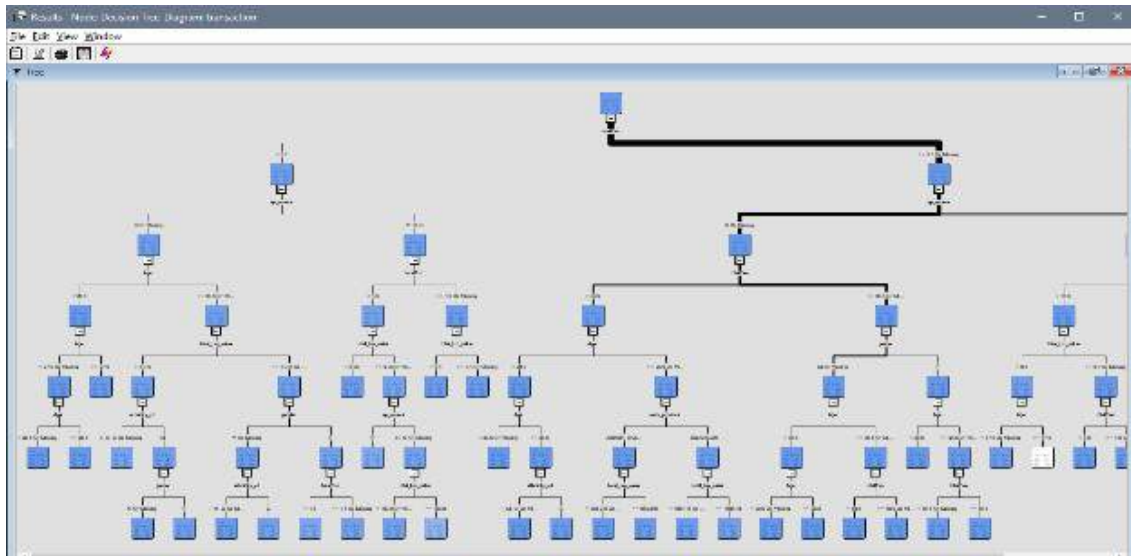


Figure 9: Maximum tree plot by decision tree model

Figure 9 shows the maximum tree plot generated by SAS Enterprise Miner. It has a total of 37 leaf nodes produced by the tree.

English Rule on Decision Tree for Transaction Table

<p>if totalTxn < 5.5 AND ap_source IS ONE OF: S or MISSING AND Age < 27.5 AND Age >= 25.5 then Tree Node Identifier = 33 Number of Observations = 16706 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.00 Predicted: card_status=A = 1.00</p>	<p>if total_txn_value < 4.015 AND totalTxn < 5.5 AND ethnicity_pd IS ONE OF: I, C, 0 or MISSING AND ap_source IS ONE OF: S or MISSING AND Age >= 28.5 or MISSING then Tree Node Identifier = 36 Number of Observations = 7675 Predicted: card_status=P = 0.01 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.01 Predicted: card_status=A = 0.99</p>	<p>if total_txn_value >= 3.95 or MISSING AND totalTxn < 1.5 AND ap_source IS ONE OF: P then Tree Node Identifier = 40 Number of Observations = 6467 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.02 Predicted: card_status=C = 0.00 Predicted: card_status=A = 0.98</p>
--	--	---

Figure 10: English rule for decision tree churn model.

Figure 10 shows the English rules generated by the decision tree model. From Figure 5.20, we

observed that if totalTxn < 5.5 AND ap_source is S and age < 27.5 and age >= 25.5 then the predicted card_status is A. A represents 'card aborted' or 'card withdrew'. If total total_txn_value < 4.015 AND totalTXN < 5.5 AND ethnicity_cd is I, C, O AND ap_source is S AND age >= 28.5 then it is predicted that 1% will have the card_status equal to P and C while 99% of the card_status is equal to A. P represents card is phantom where card is new and active but transactions have not been registered with the server system yet. C represents card has been cancelled. Finally, if total_txn_value >= 3.95 and totalTxn < 1.5 AND ap_source is P then it is predicted that 2% will have the card_status equal to F while 98% of the card_status is predicted as A. F represents card has been force closed or permanently terminated.

<pre> if total_txn_value < 0.6 AND totalTxn < 15.5 AND totalTxn >= 5.5 AND ap_source IS ONE OF: P, N, O AND Age >= 77.5 then Tree Node Identifier = 55 Number of Observations = 34 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.12 Predicted: card_status=A = 0.88 </pre>	<pre> if total_txn_value >= 0.6 or MISSING AND totalTxn < 7.5 AND totalTxn >= 5.5 AND ap_source IS ONE OF: P, N, O then Tree Node Identifier = 56 Number of Observations = 22385 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.00 Predicted: card_status=A = 1.00 </pre>	<pre> if total_txn_value >= 0.6 or MISSING AND totalTxn < 15.5 AND totalTxn >= 7.5 or MISSING AND ap_source IS ONE OF: P, N, O then Tree Node Identifier = 57 Number of Observations = 75153 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.00 Predicted: card_status=A = 1.00 </pre>
---	--	--

Figure 11: English rule for decision tree churn model.

Figure 11 shows the English rule for decision tree churn model developed. From Figure 11, we can observe that if total_txn_value < 0.6 AND totalTxn <15.5 AND totalTxn >= 5.5 AND ap_source is P, N or O AND age > 77.5 then it is predicted that 12% of the card_status is C (or cancel) while 88% it is predicted that the card_status A (or abort). On the other hand, if total_txn_value >= 0.6 AND totalTxn < 7.5 AND totalTxn >5.5 AND ap_source is P, N or O then the predicted card_status is A (or abort). Finally, if total_txn_value >= 0.6 AND totalTxn < 15.5 AND totalTxn >= 7.5 AND ap_source is P, N or O then the predicted card_status is equal to A (or abort).

<pre> if total_txn_value >= 4.015 or MISSING AND totalTxn < 1.5 AND gender IS ONE OF: F AND ap_source IS ONE OF: S or MISSING AND Age >= 28.5 or MISSING then Tree Node Identifier = 70 Number of Observations = 27710 Predicted: card_status=P = 0.01 Predicted: card_status=F = 0.01 Predicted: card_status=C = 0.00 Predicted: card_status=A = 0.98 </pre>	<pre> if total_txn_value >= 4.015 or MISSING AND totalTxn < 5.5 AND totalTxn >= 1.5 or MISSING AND gender IS ONE OF: F AND ap_source IS ONE OF: S or MISSING AND Age >= 28.5 or MISSING then Tree Node Identifier = 71 Number of Observations = 66189 Predicted: card_status=P = 0.01 Predicted: card_status=F = 0.00 Predicted: card_status=C = 0.00 Predicted: card_status=A = 0.99 </pre>	<pre> if total_txn_value < 469.9 AND total_txn_value >= 3.95 or MISSING AND totalTxn < 1.5 AND ap_source IS ONE OF: N, O or MISSING then Tree Node Identifier = 72 Number of Observations = 19122 Predicted: card_status=P = 0.00 Predicted: card_status=F = 0.01 Predicted: card_status=C = 0.00 Predicted: card_status=A = 0.99 </pre>
--	--	---

Figure 12: English rule for decision tree churn model.

Figure 12 shows the English rule for decision tree generated. From Figure 12, we have observed that if total_txn_value >= 4.015 or Missing AND totalTxn < 1.5 AND gender is F AND ap_source is S or Missing AND age >=28.5 or Missing then it is predicted that 1% of the card_status is equal to P (or phantom) and F (or force closed) while 98% is predicted to have card_status equal to A (or abort). On the other hand, if total_txn_value >= 4.015 or Missing AND totalTxn < 5.5 AND totalTxn >= 1.5 or Missing AND gender is F AND ap_source is S or Missing

AND Age \geq 28.5 or missing the it is predicted 1% of the card_status is equal to P (or phantom) while 99% of the card_status is predicted to be A (or abort). Finally, if total_txn_value $<$ 469.9 and total_txn_value $>$ 3.95 or Missing AND totalTxn $<$ 1.5 AND ap_source is N, O or Missing then it is predicted that 1% of the card_status equal to F (or force closed) while 99% of the card_staus is predicted to be A (or abort).

Fit Statistic for Decision Tree

Fit Statistics	Statistics Label	Train	Validation	Test
NOBS	Sum of Frequencies	2460666	1230333	410115
MISC	Misclassification Rate	0.003041	0.003041	0.003048
MAX	Maximum Absolute Error	0.999997	1	0.999965
SSE	Sum of Squared Errors	14883.24	7444.605	2487.497
ASE	Average Squared Error	0.001512	0.001513	0.001516
RASE	Root Average Squared Error	0.038886	0.038894	0.03894
DIV	Divisor for ASE	9842664	4921332	1640460
DFT	Total Degrees of Freedom	7381998		

Figure 13: Fit statistic for decision tree.

In Figure 13, there are a total of 2460666 observations for train data, 1230333 observations for validation data and 410115 observations for test data. The misclassification rate for train and validation data is 0.003041, as for test data the misclassification rate is 0.003048. The misclassification rate difference between training, validation and test data is very small/not significant. The average squared error for training data is 0.001512, validation data is 0.001513 and test data is 0.001516. The difference in average square error between all validation, test and training data is not significant.

Subtree Assessment Plot for Decision Tree Churn Model

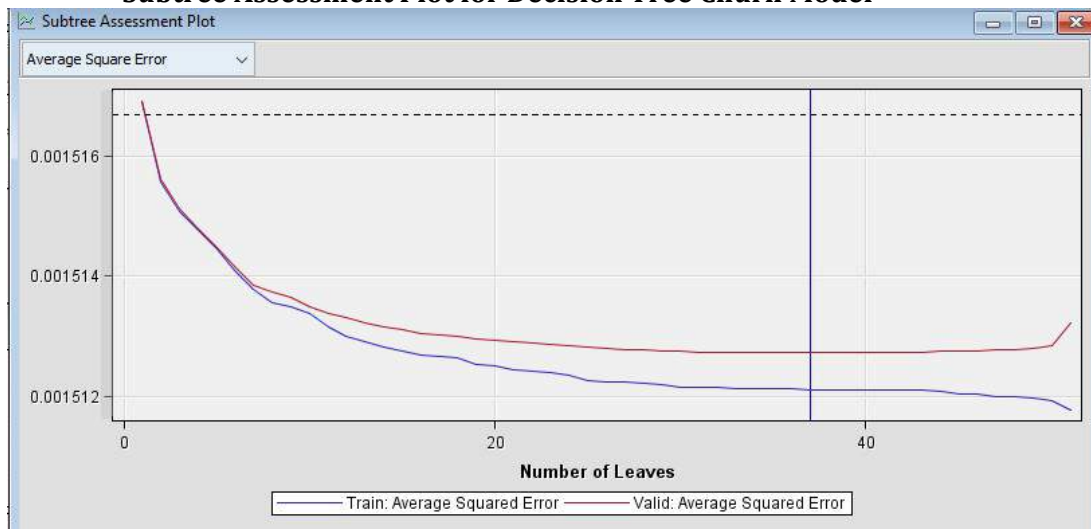


Figure 14: Subtree assessment plot

The plot in Figure 14 shows the Average Square Error corresponding to each subtree as the data is sequentially split. This plot is similar to the one generated with the Interactive Decision Tree, and it confirms suspicions about the optimality of the 37-leaf tree. The performance on the training and validation sample becomes monotonically better as the tree becomes more complex. The average square error decreases as the number of leaves in the tree increases. However, until approximately at leaf 31 the average square error remains the same until the tree matures.

Regression Churn Model

Summary of Stepwise Selection

Step	Effect Entered	DF	Number		Score Chi-Square	Wald	
			In	Chi-Square		Chi-Square	Pr > ChiSq
1	ap_source	9	1	2407.5516			<.0001
2	totalTxn	3	2	1421.4475			<.0001
3	Age	3	3	984.9698			<.0001
4	gender	3	4	574.7671			<.0001
5	work_province	45	5	617.7542			<.0001
6	ethnicity_cd	9	6	176.6191			<.0001

The selected model is the model trained in the last step (Step 6). It consists of the following effects:

Intercept Age ap_source ethnicity_cd gender totalTxn work_province

Figure 15: Summary result of regression model using Customer Transaction table.

Figure 15 shows that there are a total of 6 variables selected by the regression model. The variables are ap_source, totalTxn, Age, gender, work_province and also ethnicity_cd. They are the predictors for the customer churn model using regression technique.

Odd Ratio Analysis

Odds Ratio Estimates

Effect	card_status	Point Estimate
Age	P	1.025
Age	F	1.014
Age	C	1.033
ap_source N vs S	P	<0.001
ap_source N vs S	F	2.089
ap_source N vs S	C	1.305
ap_source 0 vs S	P	<0.001
ap_source 0 vs S	F	2.161
ap_source 0 vs S	C	1.505
ap_source P vs S	P	0.004
ap_source P vs S	F	2.264
ap_source P vs S	C	1.356
ethnicity_cd C vs 0	P	1.491
ethnicity_cd C vs 0	F	1.399
ethnicity_cd C vs 0	C	1.863
ethnicity_cd I vs 0	P	1.032
ethnicity_cd I vs 0	F	1.393
ethnicity_cd I vs 0	C	1.966
ethnicity_cd M vs 0	P	1.189
ethnicity_cd M vs 0	F	1.448
ethnicity_cd M vs 0	C	0.992
gender F vs M	P	1.746
gender F vs M	F	1.420
gender F vs M	C	0.636
totalTxn	P	0.992
totalTxn	F	0.917
totalTxn	C	0.993
work_province JOHOR vs WILAYAH PERSEKUTUAN	P	0.463
work_province JOHOR vs WILAYAH PERSEKUTUAN	F	0.662
work_province JOHOR vs WILAYAH PERSEKUTUAN	C	0.586
work_province KEDAH vs WILAYAH PERSEKUTUAN	P	0.798
work_province KEDAH vs WILAYAH PERSEKUTUAN	F	0.850

Figure 16: Odd ratio estimate for regression churn model

Figure 16 shows the odd ratio estimate for each predictor selected by SAS Enterprise Miner. Odd Ratio expresses the increase in primary outcome odds associated with a unit change in an input. It is obtained by exponentiation of the parameter estimate of the input of interest.

Age

For Age with card_status predicted equals to P (or phantom), the odd ratio estimate equals 1.025. This means that for each addition age, the odds of predicting card_status equals to P change by a factor of 1.025, a 2.5% increase. On the other hand, for Age with card_status predicted equals to F (or force closed), the odd ratio estimate equals 1.014. This means that for each addition age, the odds of predicting card_status equals to F (or force closed) change by a factor of 1.014, a 1.4% increase. Finally, For Age with card_status predicted equals to C (or cancel), the odd ratio estimate equals 1.033. This means that for each addition age, the odds of predicting card_status equals to C(or cancel) change by a factor of 1.033, a 3.3% increase.

Ap_source (P vs S)

For ap_source with card_status predicted equals to P (or phantom), the odd ratio (P vs S) estimate are 0.004. This means that cases with P value for ap_source are 0.004 more likely to predict card_status equals to P than cases with a value of S for ap_source. For ap_source with card_status predicted equals to F, the odd ratio (P vs S) estimate are 2.264. This means that cases with P value for ap_source are 2.264 more likely to predict card_status equals to F (or force closed), than cases with a value of S for ap_source. For ap_source with card_status predicted equals to C (or cancel), the odd ratio (P vs S) estimate are 1.356. This means that cases with P value for ap_source are 1.356 more likely to predict card_status equals to C(or cancel), than cases with a value of S for ap_source.

Ethnicity_cd (C vs O)

For ethnicity_cd with card_status predicted equals to P (or phantom), the odd ratio (C vs O) estimate are 1.491. This means that cases with C value for ethnicity_cd are 1.491 more likely to predict card_status equals to P (or phantom), than cases with a value of O for ap_source. For ethnicity_cd with card_status predicted equals to F (or force closed), the odd ratio (C vs O) estimate are 1.399. This means that cases with C value for ethnicity_cd are 1.399 more likely to predict card_status equals to F (or force closed), then cases with a value of O for ethnicity_cd. For ethnicity_cd with card_status predicted equals to C, the odd ratio (C vs O) estimate are 1.863. This means that cases with C value for ethnicity_cd are 1.863 more likely to predict card_status equals to C (or cancel) than cases with a value of O for ethnicity_cd.

Gender (F vs M)

For gender with card_status predicted equals to P (or phantom), the odd ratio (F vs M) estimate are 1.746. This means that cases with F value for gender are 1.746 more likely to predict card_status equals to P (or phantom) than cases with a value of M for gender. For gender with card_status predicted equals to F (or force closed), the odd ratio (F vs M) estimate are 1.420. This means that cases with F value for gender are 1.420 more likely to predict card_status equals to F (or force closed) than cases with a value of M for gender. For gender with card_status predicted equals to C (or cancel), the odd ratio (F vs M) estimate are 0.636. This means that cases with F (or force closed) value for gender are 0.636 more likely to predict card_status equals to C than cases with a value of M for gender.

Work_province (Kedah vs Wilayah Persekutuan)

For work_province with card_status predicted equals to P, the odd ratio (Kedah vs Wilayah Persekutuan) estimate are 0.798. This means that cases with Kedah value for work_province are 0.798 more likely to predict card_status equals to P (or phantom) than cases with a value of

Wilayah Persekutuan for work_province. For work_province with card_status predicted equals to F (or force closed), the odd ratio (Kedah vs Wilayah Persekutuan) estimate are 0.850. This means that cases with Kedah value for work_province are 0.850 more likely to predict card_status equals to F (or force closed) than cases with a value of Wilayah Persekutuan for work_province. For work_province with card_status predicted equals to C (or cancel), the odd ratio (Kedah vs Wilayah Persekutuan) estimate are 0.443. This means that cases with Kedah value for work_province are 0.443 more likely to predict card_status equals to C (or cancel) than cases with a value of Wilayah Persekutuan for work_province.

Work_province (Sarawak vs Wilayah Persekutuan)

For work_province with card_status predicted equals to P (or phantom), the odd ratio (Sarawak vs Wilayah Persekutuan) estimate are 0.873. This means that cases with Sarawak value for work_province are 0.873 more likely to predict card_status equals to P (or phantom) than cases with a value of Wilayah Persekutuan for work_province. For work_province with card_status predicted equals to F (or force closed), the odd ratio (Sarawak vs Wilayah Persekutuan) estimate are 1.057. This means that cases with Sarawak value for work_province are 1.057 more likely to predict card_status equals to F (or force closed) than cases with a value of Wilayah Persekutuan for work_province. For work_province with card_status predicted equals to C (or cancel), the odd ratio (Sarawak vs Wilayah Persekutuan) estimate are 0.870. This means that cases with Sarawak value for work_province are 0.870 more likely to predict card_status equals to C (or cancel) than cases with a value of Wilayah Persekutuan for work_province.

Fit Statistic for Regression Churn Model

Fit Statistics	Statistics Label	Train	Validation	Test
AIC	Akaike's Information Criterion	105122.6		
ASE	Average Squared Error	0.001513	0.001513	0.001516
AVERR	Average Error Function	0.010665	0.010715	0.010685
DFE	Degrees of Freedom for Error	7381923		
DFM	Model Degrees of Freedom	75		
DFT	Total Degrees of Freedom	7381998		
DIV	Divisor for ASE	9842664	4921332	1640460
ERR	Error Function	104972.6	52730.54	17527.68
FPE	Final Prediction Error	0.001513		
MAX	Maximum Absolute Error	1	1	1
MSE	Mean Square Error	0.001513	0.001513	0.001516
NOBS	Sum of Frequencies	2460666	1230333	410115
NW	Number of Estimate Weights	75		
RASE	Root Average Sum of Squares	0.038897	0.038902	0.038941
RFPE	Root Final Prediction Error	0.038897		
RMSE	Root Mean Squared Error	0.038897	0.038902	0.038941
SBC	Schwarz's Bayesian Criterion	106158.7		
SSE	Sum of Squared Errors	14891.62	7447.724	2487.57
SUMW	Sum of Case Weights Times Freq	9842664	4921332	1640460
MISC	Misclassification Rate	0.003041	0.003041	0.003048

Figure 17: Fit statistic for regression model.

From Figure 17, the average square error for training and validation data is 0.001513, while test data is 0.001516. The misclassification rate for training and validation data is 0.003041 while test data is 0.003048. The Sum of Squared Errors for training data is 14891.62, while for validation data is 7447.724 and for test data is 2487.57. The average square error difference between training, validation and test data is not significant because the difference value is only 0.000003. Not only that, the difference in misclassification rate between training, validation and test data is also not significant.

Model Comparison for Decision Tree and Regression

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Misclassification Rate	Train: Akaike's Information Criterion	Train: Average Squared Error
Y	Reg3	Reg3	Regression	card_status	card_status	0.003041	105122.6	0.001513
	Tree3	Tree3	Decision Tr...	card_status	card_status	0.003041	.	0.001517

Figure 18: Fit Statistic for model comparison.

Figure 18 shows the fit statistic for model comparison. Based on the 2 models, the model comparison node chooses regression as the preferred model because the Average Square error for Regression is slightly lower than Decision Tree. The difference is 0.000004 therefore decision tree and regression model can be used for prediction.

CUSTOMER SPENDING BEHAVIOUR

Discussion on the Departmental store Customer's spending behaviour

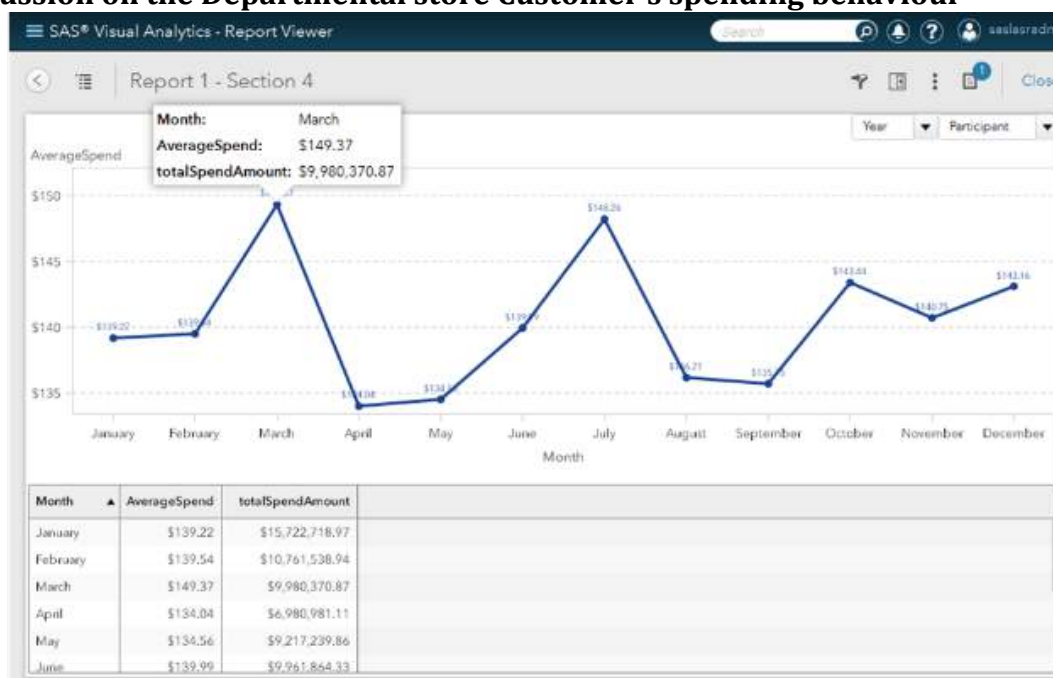


Figure 19: Total and average spending for the chosen departmental store

Figure 19 illustrates shows the total and average spending in two years for card holders at outlets of the departmental store. This shows that customers spend more in March and July. The reason that the departmental store achieves high sales in March is because Malaysians are likely to purchase more before the implementation of goods and service tax (GST) on April 2015 (Eugene Mahalingam, 2014). People tend to spend more again during July due to festive promotion and on Member day held in 2014. Furthermore, the table below the line chart is to show the details for every month's average and total amount spent by the customers.

Discussion on the pharmacy chain store Customer's spending behaviour

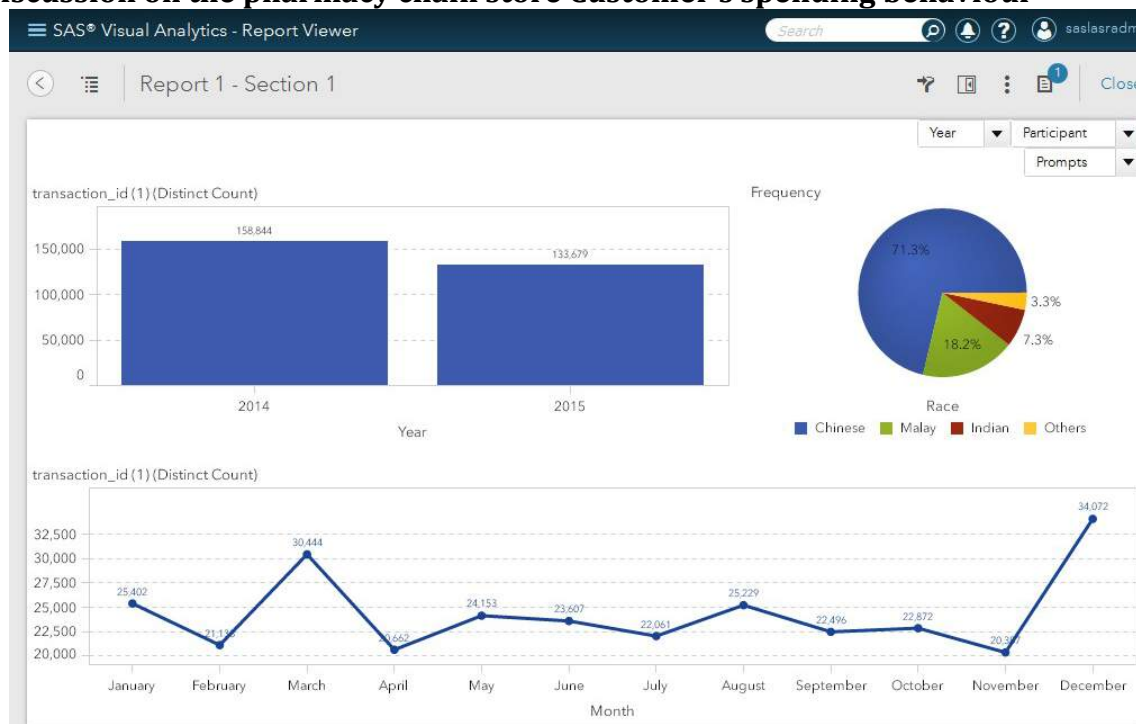


Figure 20: Section 1 for the pharmacy chain store

Figure 20 shows the transaction frequency in two years and the customers' race for the pharmacy chain store. From the result, it can be seen that there is a huge change in customer's buying behaviour in 2015. This change caused a reduction in sales for all the pharmacy chain store outlet resulting in smaller number of transactions. The issue that causes this is the implementation of Goods and Service Tax (GST) in Malaysia at that point of the year. Other than that, the sharp drop of Malaysian currency also introduces price hike for many essential goods. Furthermore, the pharmacy chain store tends to have more Chinese customers as compared to other races. Because of the price hike, all the races tend to spend their money wisely and carefully. From the line graph, it shows that in March, there is a very high frequency of transaction. This is because customers tend to purchase more items before the implementation of GST in April 2015. This has caused the month before April 2015 to have very high number of transaction.

Discussion on the petrol company customer's spending behaviour

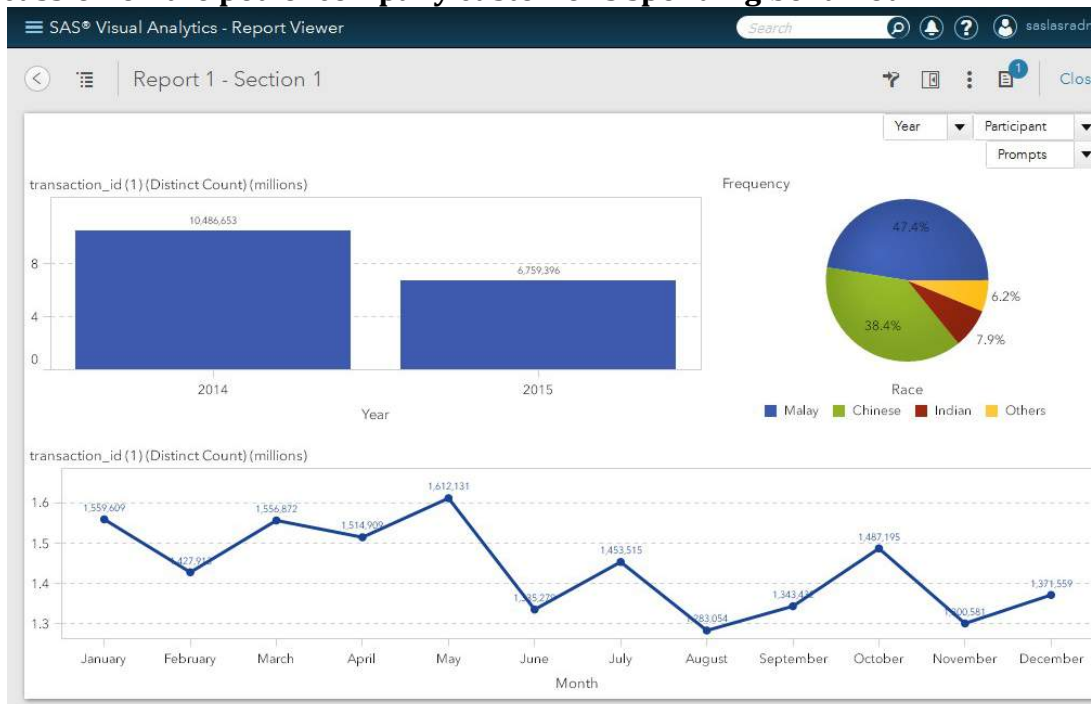


Figure 21: Section 1 for the petrol company

Figure 21 shows the transaction frequency for two years and the races of customers for the petrol company. From the result in Figure 21, it is found that there is a huge change in customer's buying behaviour in 2015. This has caused a significant impact on the petrol company's sales. As illustrated from the line graph, the number of transactions are not stable and fluctuate every month. This may be due to the volatility of oil price which changes every month. With the implementation of GST and the depreciation of Malaysia Ringgit, customers have become very price sensitive and are watchful for low petrol price. When petrol price is low, customers tend to purchase more petrol before the petrol price increases again next month. Thus, resulting more transactions on some months. When petrol price is high, customers will pump less petrol. This has resulted less transactions on that month. Thus, the transactions for the petrol company are volatile due to changes of oil price. In addition, the chart also shows that the customer group for the petrol company consists more of Malay customers.

RECOMMENDATIONS FOR THE DEPARTMENTAL STORE, PHARMACY CHAINED STORE AND PETROL COMPANY.

Recommendation for the departmental store

From the results discussed, the loyalty card company and the departmental store are able to identify when their customers would spend or shop less. In the analysis, it was found that card holders spend less in April and September as compared to March and July. In order to boost up the sales or to have better sales every month, the marketing team of the departmental store can give incentive such as discounts on promotional items. This will provide incentive to the card holders purchase these promotional items. According to Arief Irsyad (2014), consumers will go crazy whenever consumers see the word "sale" or "discount" because this makes them think that the prices are lower. The above figure also shows that the number of transactions in April is the lowest. This might be due to gifts available for redemption are not attractive enough. Therefore, the departmental store could come out with more appealing gifts for their customers to attract them to redeem their earned points. As a result, customers will tend to use the loyalty card more whenever they make their purchases because the customers will want to

collect more points in order to be able to redeem attractive gifts. Instead of having exclusive sales for loyalty card members every month, the departmental store should reduce the frequency of having monthly sales or promotions and should carry out sales or promotions in the month of April and September to make their sales improved. The departmental store can do so by increasing the points per transaction for months April and September to attract customers to use the card when making purchases. Furthermore, the loyalty card company can set up booth in the departmental store during peak season to allow customers to sign up for the card and promote the benefits of the card.

Recommendation for the pharmacy chain store

From the findings, the management for the loyalty card company and the pharmacy chain store can understand the reasons why sales for certain months are not constant. There are so many competitors out there in Malaysia in the line of pharmacy stores. In order to boost up the sales, should consider frequent promotional activities on the healthcare products. This will attract health conscious card holders to shop at the pharmacy chain store. Other than that, the loyalty card company should also provide incentives such as big bonus points or high value free gifts for card holders when they shop at the pharmacy chain store during sales periods. Other than that, the pharmacy chain store can come out with cross selling products or services for their loyal customers. This is because the pharmacy chain store does not only sell healthcare products, but they also sell cosmetic and daily necessities. Furthermore, the loyalty card company can set up a booth in the store once a month to allow their customers to sign up for as new loyalty card members and to promote the benefits of the loyalty card.

Recommendation for the departmental store and petrol company

The results found from the discussion allows the management of the loyalty Card Company and the petrol company to identify when their customers spend or pump petrol less. When they anticipate to have less customers spend at their outlets, they can find ways to attract these customers to spend at the petrol company petrol stations. The petrol company can consider giving rewards to their customers for their continuous support with more appealing products and services such as cash voucher when they spend up to certain amount at their petrol stations. Hence these card holders can collect points from their transactions at the petrol company. Besides, can boost up their sales by having more contests or lucky draws to attract their customers to pump petrol at the petrol company. The petrol company can also provide training for their workers and remind their customers to use the loyalty card when filling up petrol. If customers do not hold any of the loyalty card, the petrol company's workers can promote them the benefits of collecting points and direct them to the sign-up booth at the petrol company. Therefore, loyalty Card company can set up the booth at the petrol company every month for non-members to sign up and thus will lead to more transactions when customers spend at The petrol company.

STRATEGIES FOR THE CHURN PREDICTIVE MODEL

One of the strategies proposed in this research is to offer a simple way to pay according to Meagan Rhodes, the Digital Marketing Lead of @Pay, a simple secure email payment technology company. As for loyalty card company, allowing card holders to easily use their loyalty card while making any transactions will be able to significantly increase the number of transaction by existing card user (Stringfellow, 2016). For example, not all customers carry their loyalty card everywhere they go. So, by allowing them to store their card details in their phone and make transaction through their phone will significantly increase the usage of the loyalty card because everyone do carry their phone everywhere they go. This strategy will need a substantial amount of investment and it takes time and effort before seeing the result.

A more economical way is to ask for feedback from customer on ways to improve their service said Chip Bell, a Senior Partner in Chip Bell Group (Stringfellow, 2016). For example, based on the predicted customers from the decision tree and regression model who will potentially churn, the loyalty company can reach out to these potentially churn customers and ask for their feedback, listen to their needs on the loyalty program and then based on their feedbacks and act accordingly.

Therefore, based on the English rule in Decision Tree model from figure 4.6 to figure 4.8 and the odd ratio estimate shown in figure 4.12 from the regression model, we will then determine who are the potential churners and take appropriate action.

Based on the English rules, we observe that if $\text{total_txn_value} < 8.1$ AND $\text{totalTxn} < 5.5$ AND $\text{totalTxn} \geq 1.5$ AND ap_source is P, N, or O then it is predicted that approximately **43** customers will churn. Besides that, if $\text{total_txn_value} \geq 3.95$ AND $\text{totalTxn} < 1.5$ AND ap_source is P then it is predicted that approximately **129** customers will churn. On the other hand, if $\text{total_txn_value} < 0.6$ AND totalTxn is between 5.5 and 15.5 AND ap_source is P, N or O AND $\text{age} > 77.5$ then it is predicted that approximately **29** customers will churn. If $\text{total_txn_value} \geq 4.015$ AND $\text{totalTxn} < 1.5$ AND gender is F AND ap_source is S AND $\text{age} \geq 28.5$ then it is predicted that approximately **277** customers will churn. Furthermore, if total_txn_value is between 3.95 and 469.9 AND $\text{totalTxn} < 1.5$ AND ap_source is N or o then it is predicted that approximately **191** customers will churn.

Actions can be taken on the application source (ap_source) P, N and O. As these 3 out of 4 application sources seems to be one of the cause of member churn. A more useful approach of membership application should be formulated to replace these application sources (with value equals to P, N and O).

Based on the odds ratio analysis in Figure 4.12, strategies can be implemented on card holders with their work province address equals to Perlis (3.334), Labuan (1.275), Negeri Sembilan (1.141), Sarawak (0.870), Selangor (1.073) and Perak (1.052) because card holders in these six (6) states have the highest odds of churning as compared to Wilayah Persekutuan. This indicates that the loyalty card company should focus more on customers from these 6 states which have higher churn probability. As for ethnicity, card holders who are Chinese and Indian have higher odd of churning as compared to Malay. Therefore, it is logical to implement strategies mentioned above on Chinese and Indian card holders more than Malay.

RECOMMENDATION FUTURE RESEARCH

Some of the limitations and challenges faced in this project can be solved to improve future works. In this project, two years of the loyalty membership and transactions data was collected and the data consists of 70 GB. The prediction would be more accurate if more years of data can be provided such as 4 to 8 years in order to see the pattern changes on the member spending patterns and churn factors. Preliminary investigation such as understanding the data with some questions being asked to the Head of IT has been conducted to understand the current situation better. For future enhancement, perhaps an interview and observation session can be conducted to enhance the data collection in the company.

CONCLUSION

The loyalty card company in this case study is a well-known loyalty card company in Malaysia with long history of establishment. In this research, the result generated by SAS Visual Analytic and SAS Enterprise Guide on demographic analysis such as gender, race and age group have shown interesting with different spending behaviours. For example, different races tend to

spend more on different month and different age groups tend to spend more at different partners. Demographics factors can enable loyalty card company to further analyse on customer spending behaviour and produce suitable approaches to target at their card holders. The graphical presentation produced by SAS Visual Analytic has clear and flexible presentation capabilities that can assist the loyalty card company to improve their businesses in the future.

References

Barnard, A. (2002). Feedback Seeking in Customer Service Relationships. [online] Available at: <http://citeseerx.ist.psu.edu/viewdoc/download>

Barlow, J. & Moller, C. (1996) *A Complaint is a Gift: Using Customer Feedback as a Strategic Tool*. [ebook] San Francisco: Berrett-Koehler Publishers. Available from: <https://books.google.com.my/>

Bholat, D., Hansen, S., Santos, P., & Bailey, C.S. (2015) *Text Mining for Central Banks*. Centre for Central Banking Studies.

Blake, C. (2011) *Text Mining*. Annual Review of Information Science and Technology.

Chakraborty, G., Pagolu, M. & Garla, S. (2013) *Text Mining Analysis: Practical Methods, Examples, and Case Studies using SAS*. Cary, NC: SAS Institute Inc.

Chinsha, T.C. & Joseph, S. (2014) *Aspect Based Opinion Mining from Restaurant Reviews*. International Journal of Computer Applications.

Dalmolen, S. (2010). *Defining patterns in unstructured manifests in a volatile cross-domain environment*. [online] Available from: <http://www.cs.rug.nl/~aiellom/tesi/dalmolen>.

Dominici & Guzzo (2010), *Customer Satisfaction in the Hotel Industry: A Case Study from Sicily*, International Journal of Marketing Studies, 2(2).

Fan, W., Wallace, L., Rich, S. & Zhang, Z. (2006) *Tapping the Power of Text Mining*. Communication of the Ach. [Online]. 49(9), 76-82. Available from: <http://dl.acm.org/citation.cfm?id=1151032>

Godden, S. & Pollock, A.M. (2009) *Waiting List and Waiting Time Statistics in Britain: A Critical Review*. Public Health [Online] 123, 47- 51. Available from: www.elsevierhealth.com/journals/pubh.

Gupta, V. and Lehal, G. (2009). *A Survey of Text Mining Techniques and Applications*. JOURNAL OF EMERGING TECHNOLOGIES IN WEB INTELLIGENCE, 1(1).

Hendry, S., & Madeley, A. (2010) *Text Mining and the Information Content of Bank of Canada Communications*. Bank of Canada. [Online] Available from: <https://core.ac.uk/download/pdf/6699593.pdf>

Hotho, A., Nurnberger, A. and Paab, G. (2005). *A Brief Survey of Text Mining*. [online] Available at: <http://www.kde.cs.uni-kassel.de/hotho/pub/2005/hotho05TextMining.pdf>

Lau, K., Lee, K. and Ying, H. (2005). *Text Mining for the Hotel Industry*. 46(3).

Lavrakas, P.J (2008). *Encyclopedia of Survey Research Methods*. 2nd Edition. California. Thousand Oaks, Calif.

Liu, B., (2012). *Sentiment Analysis and Opinion Mining*. Sentiment Analysis and Opinion.

Mattila & Neill n.d, *Relationship between Hotel Room Pricing, Occupancy, And Guest Satisfaction: A Longitudinal Case of a Midscale Hotel in United States*.

McLoughlin, P (2013), *Review of Measure to Reduce Costs in the Private Health Insurance Market*.

Oja, P. (2010). *Significance of Customer Feedback An Analysis Of Customer Feedback Data In A University Hospital Laboratory*. [online] Available at: <http://jultika.oulu.fi/files/isbn9789514262739.pdf>.

Vinodhini, G., Professor, A. & Chandrasekaran, R., (2012). *International Journal of Advanced Research in Computer Science and Software Engineering Sentiment Analysis and Opinion Mining: A Survey*, 2(6).