INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND ANALYSIS

ISSN(print): 2643-9840, ISSN(online): 2643-9875

Volume 05 Issue 09 September 2022

DOI: 10.47191/ijmra/v5-i9-27, Impact Factor: 6.261

Page No. 2521-2530

Market Segment Based on Customer Analytics: An Approach on The S Bank's Big Data (Vietnam)

Tu Van Binh[1], Ngo Giang Thy[2]

[1]University of Economics Ho Chi Minh City

[2]Nguyen Tat Thanh University

ABSTRACT: With the source of big data extracted from the S Bank, the sample of 3,527 customers with 130,000 transactions is employed in the study. Based on the RFM approach plus the mathematic method of K-means analysis taken into account, five market segments are found, such as: spender group (38.8%), shopper group (10.3%), frequent group (14.8%), uncertain group (27.5%), and best group (8.5%). The derived client segmentation is focused on transaction functionality. This finding contributes to the planning, on the basis of its transaction, of a system for segmenting customers to estimate the potential value of the various segments of customers in the bank, in particular for the retail banking sector. In addition, empirical research findings may provide feedback on marketing strategies, develop promotional programs to introduce new products for each category, and stimulate the most profitable consumer group's consumption.

KEYWORDS: Market segment, customer analytics, RFM, banking

1. INTRODUCTION

The exponential growth of information technology in the banking sector has given rise to a major concern that banks need to learn about how to explore internal data in order to serve competitive strategies. Indeed, most banking, insurance and financial services institutions are currently attempting to follow a modern approach to data mining for the advancement and innovation of the services toward customer retention.

Investing in the Big Data infrastructure induces simulation that data mining is of considerable interest to the financial industry because it facilitates the retrieval of useful knowledge from large volumes of data. It particularly helps to recognize customer behavior toward an effective strategy. In fact, big data from businesses is not easy to explore. It generates more profitability once customer analytics on big data are concerned well.

The value of consumer partnerships has now been recognized by banks, which is one of the effective reasons. Nevertheless, problems with how to maintain the most valuable customers and reduce the churn rate are troublesome. Big Data are a valuable resource and very helpful to correctly measure and forecast consumer behavior. The power of data is therefore to extract utility through different domains of its functioning, product through distribution, regulatory compliance management, risk management and management of customer service.

As we know, bank operation specifics produce a large amount of data from unstructured data, such as transaction history, customer to unstructured data, such as the activities of the customer on the website, or the social networking mobile banking application. The most troublesome, though, is how to explore accessible big data. Although there are not a few banks that understand this and want to turn big data into the most powerful tool for market rivalry, new systems, skills, and so on seem to face challenges.

With changes in Vietnam's integration policies through the information technology system, along with fire competition, more and more Vietnamese banks have been paying more attention to investing a massive amount of money into big data system and people's power. S Bank, for example, is one of the banks that is able to spend money very early on in structuring the data structure. To help business plans, it understands data warehouse benefits. In reality, the internal advantages did not improve the effectiveness of internal management, but also helped to increase the competitive advantages and optimize profits.

At present, banks need to be agile in their plans for innovation initiatives in order to capture customer needs and boost customer loyalty and retention. CRM is, by the way, very useful and recognized for customer acquisition and retention. As a result, the bank would be more likely to develop long-lasting and successful customer relationships.

S Bank has offered a capacity-building initiative for workers directly linked to data analytics and business growth to investigate the bank's Big Data infrastructure. In addition, the database department is set up to leverage the internal database that positions an employee's role in the creation of talent in the ecosystem. All, though, seems to be more and more concerned and used for data mining, predictive consumer behavior analytics.

Banks may use data to stand out from competition and improve their position against emerging technology companies. Second, they need a better interpretation of their clients, not one focused on a particular incident or encounter, but one focused on a more nuanced view of actions. By looking at the evolution of financial behavior and how it interacts with non-bank life, banks can better understand what motivates consumers and see evolving patterns in building new goods and services in a radically changed environment.

This paper aims to investigate transaction history of customers who are using bank services and obtain market segments. This study serves as an experiment to help discover customer segments for business strategy.

2. CONCEPTUAL FRAMEWORK

Banking institutions need to be ahead of the curve to stay competitive in this modern era. Banks should work with companies that specialize in data analytics and predictive analysis to get an edge on their competition. To do this, big data system should be concerned. Once data is extensively explored, the bank can get a chance to make a lower cost in business and achieve greater performance. In reality, in the business unit, several forms of costs can be incurred, such as advertisement costs to attract potential clients, costs of a personal sales pitch to new buyers, contact costs to clarify business processes and dealing with new clients. When business units have a strong approach to data mining, these costs can be minimized. Banks in Vietnam, for example, have a significant investment in big data systems to record consumer actions from various channels.

Unlike customer satisfaction and customer loyalty, customer retention is an ability to remain customer over time and measures a relationship between the customer and the firm. Once a customer is retained longer, expectation of loyalty is possible (Jaiswal, Niraj, Park, & Agarwal, 2018). However, according to Kumar et al. (2013), indicating customer satisfaction offers only a small portion of variance in loyalty, but not convincing much to enhance customer retention. Big data is a term used to describe the large volume of data that is being generated by businesses and people. The process of collecting and analyzing database from big data can help companies make better decisions about their business. Once the right strategy based on customer analytics is derived, the churn of bank can be reduced.

Because of competition, retail banks have paid more attention to revamping loyalty programs toward customers. This is strongly considered toward incorporating a reward system into the bank's plans. Currently, many retail banks offer a significant number of potential rewards to promote and solidify customer loyalty. The banks develop a relationship based on two-way street, in which the customer will remain in the relationship only once there is value in doing so. However, once the customers' purchase volume is increased and beneficial relationship gradually takes shape, the customers' relationship with banks is remained (Jaiswal, Niraj, Park, & Agarwal, 2018)

There are many studies using big data to analyze customer consumption (Khajvand & Tarokh, 2011), these studies mention that big data as the tool allow a business unit to manipulate, create and manage huge data sets, also it is stored and required to support the volume of data, characterized by variety, volume and velocity[1].

Internationally, banks are starting to leverage the power of data in order to create value across multiple aspects of their business, ranging from sentiment analysis, product cross-selling, regulatory compliance management, reputational risk management, financial crime management, and so on (Srivastava & Gopalkrishnan, 2015).

Big Data in banking and its contribution

Once the number of customers grows, the bank has to care about delivering quality to a certain degree. Practice demonstrates that the method of tracking and analyzing consumer credit banks and financial institutions, based on vast quantities of data such as records, personal and other sensitive information, has been streamlined by review of existing data. But with the available big data, banks can take advantage of constantly tracking the customer's actions in real time, identifying data sources needed to collect justice solution service. Customer records of assessment processes in real time would gradually improve operational efficiency and profitability, fostering more organizational development.

[1] *Meta Group. Application Delivery Strategies; February 2001*

There are many sources of big data in most industries and in various fields, not just in the banking and financial services sectors. Every customer transaction at banks creates electronic records, the copies are kept in compliance with legal legislation, and transactions at various locations are kept in the office ATMs as well as information stored in the bank. Thanks to big data analysis, the financial services of companies no longer store the data as needed as in the past, but now they are more involved, more involved in extraction to obtain the results that are focused on that provide solutions to enhance operations , increase the organization's profitability.

Segmentation of consumers and consumer records analysis

Once an initial analysis of the customer's spending habits is carried out to identify the type of service, channel transactions are priority customers (i.e. customers who want to save or invest in credit), a database serving the segment will be compiled, the customers will be classified on the basis of the information and documents provided by the clients. Customers who spend comfortably, customers quickly pay the debts and begin to repay the maturing, big data would offer banks customers' information, experience and deep spending patterns, simplifying the task of defining their needs and desires. The bank will be able to identify customers on the basis of different criteria by being able to monitor each transaction by customers, including services frequently used by customers, time to use the service, spending patterns while using a credit card or even a net asset value (net worth plus the value of customer assets minus liabilities).

Benefits that bring customer segmentation is that it allows banks to target customers with better marketing campaigns relevant and designed to meet the exact needs of the customer. A segmented approach to marketing will help banks better serve their specific niche. For example, some banks cater to young professionals, while others offer financing for new parents. To reduce the risk of taking on more risk than they should, banks segment high net worth and low net worth clients. Why not offer a service that helps bridge this gap? Researching the market and securing capital is key to success.

RFM method

As known, RFM is a good tool to measure customer behavior, it seems as an incredible and perceived technique to evaluate the customers' commitment to business units and differentiates important customers from large data by three attributes of recency, frequency and monetary. Based on customers' purchasing history, RFM is quite supported to address customers' classification and ranking.

According to Cheng & Chen (2009), the detail definitions of Recency, Frequency and Monetary method are described as the following: (i) Recency (R) is defined as an interval between the time when the latest purchasing order presents and happens such as one week, one month or one quarter. A lower recency value means that customers frequently visit the company. Likewise, the higher value implies that in the near future, customer sometimes or rarely visits the company; (ii) Frequency (F) is shown as the number of purchasing transactions made in a given period of time by customer, for example, one time per year, two times per quarter or three times per month. The higher frequency value, the more loyal customers regarding the company; (iii) Monetary (M) is identified as the total money amount that customers spent during one specified time. So, the more money consumed, the more earnings customers bring to the company.

Accordingly, Wu & Lin (2005) demonstrated that the higher the value of R and F is, the more comparable customers make a new transaction with companies. In addition, the higher M value is, the more comparable customers purchase products or services of companies several times. All customers are analyzed by recency, frequency, and monetary scores which take place in the scale from 1 to 5 as a quintile based on its original records, in which1 being unlikely and 5 being likely. Scores of combinations of RFM are assumed to get remarkable attributes as shown in Table 1. Once all R, F and M are the most recent, the most frequent, and the highest spend, respectively, the score of customers belonging to this group is called "champions" with the highest score of 5. Conversely, once indicators of R, F, M are the least recent, only one transaction, and the lowest spent, respectively, customers belonging to this group are called "lost", due to its score is the smallest score of 1.

Table 1: Recency, Frequency and Monetary Score Description

Score	Classification	Recency	Frequency	Monetary
5	Champions	Most recent	Most frequent	Highest spend
4	Promising	Much recent	Much frequent	Much spend
3	Can't-lose them	Recent	Frequent	Average spend
2	At risk	Less recent	Less frequent	Less spend
1	Lost	Least recent	Only one transaction	Lowest spend

Source: Wu & Lin (2005)

3. Research Methodology

Data collection: The data used is derived from the S Bank's data warehouse. This study is seriously needed to be confidential, so it is not persuaded that the use of data as well as the results of this study must be responsible for who can use the information.

In the data mining process, data preprocessing is taken into account, since it increases the accuracy and efficiency of subsequent modeling (Han & Kamber, 2006). Data pre-processing technology tasks include data cleaning, data transformation, data integration and data reduction.

The database used in the analysis is derived from the bank's data warehouse, with daily data of 12 months from 1 January 2019 to 31 December 2019. Accordingly, there are 3,527 clients employed in the study with 130,000 rows. It means each row is a transaction of customer at the S Bank and one customer has more transaction. As depicted in table 2, the current paper takes into account of 25 variables. Each has a given measure, such as nominal, continuous, date or categorical measurements.

Table 2: Information of variables concerned

Name of variable	**Definition**	**Measurement**
1. CIF	ID of customer	Nominal
2. Branch	ID of branch	Nominal
3. No.document	ID of customer document	
4. Type of customer	Type of customer + Credit customer + Individual customer + family business customer	Nominal variable
5. Access date	Date to register service of customer	date
6. Approved date	Date of documents approved	date
7. Document score	Score on initial documents of customer that the staff evaluates	Continuous
8. Gender	Gender of customer	Nominal
9. Age	Age in years	Continuous
10. Education	Education of customer	Nominal
11. Community	Behavior community relations + Prestige + Good enough + Unknown	Nominal
12. Marital status	Marital status: Single and married	Nominal
13. Job position	Positions + Manager + Worker + Director + Office staff	Nominal
14. House ownership	Status of house property of customers + Owned house + Rented house + Stayed with parents' house	Nominal
15. Usage	Time using loan services from banks + only loan services from 'S' bank + loan services from 'S' bank > another bank + loan services from 'S' bank = another bank + loan services from 'S' < another bank + No loan services from 'S'	Nominal
16. Type of product	Type of loans	Nominal
17. Type of card	Type of card: Main and non-main	Nominal
18. Level of card	Value of card level	Continuous
19. Balance account	Balance account measured in Vietnamese currency (VND)	Continuous

20.	Fee	Annual fee (Yes/No)	Nominal
21.	Card_date	Date to issue the bank card	Date
22.	No.Card	Account number of card	Series number
23.	Ncard_Date	Date to approve the new card	Date
24.	Workingyear	Number of years of customers working	Continuous
25.	Income	Income of customer measured VND/month	Continuous

Methodologies to analyze Data

One of the instruments used in descriptive statistics is data mining. In addition, clustering techniques are employed to cluster customers into groups, which satisfy two main criteria: (i) each group or cluster is homogeneous; (ii) each group or cluster should be different from other clusters. This method mainly gains customer segmentation. In terms of classification method, the cluster of K-means is employed to segment customers (Khajvand & Tarokh, 2011).

Input variables recruited in the cluster method are generated by the RFM indicator, which R as Recency, F as Frequency, M as Monetary are generated. Since segmentation is on the basis of Recency – Frequency – Monetary (RFM), the selected features of data to meet RFM is included last transaction date (purchased), count transaction (purchase) and total monetary that customer took loan during one year and count item which refers to variety of customer taken transaction. To count transactions, it is the frequency of customer transactions. In data transformation, the data is transformed in a way that can be exploited by data mining tools.

C&R Tree as a "Classification and Regression Tree" is also employed to investigate customer behavior in more detail. It is supported to develop classification systems toward prediction based on a set of decision rules. Application of this method known as rule induction brings several advantages. Based on this the marketer can think and develop campaigns to retain customers and stimulate their consumption.

4. Empirical analysis

According to the database, the sample size of 3,527 active customers with 130,000 transactions is the main data toward data mining. The research period of data is collected from 1 January to 31 December 2019.

Based on descriptive statistics, four main groups of customers are buying services at the bank, in which customers with their position as staffs account for 26.5 percent, next as workers accounting for 26.1 percent, directors with 24.7 percent, and managers with 22.8 percent.

The bank provides 18 forms of credit facilities at the time of analysis. These credits are broken into 8 main groups to be handy, e.g. VG (Visa Gold), VC (Visa Classic), MG (Master Gold), MS (Master Standard), VP (Visa Platinum), VV (Viva Violet Card), VA (Visa Auto Card), and others. Accordingly, VG is the best concern of customers (24 percent), next as VC with 22 percent, MG with 12 percent, MS with 11 percent, VP with 9 percent, VV 7 percent (Figure 1).

As a consequence, both VG and VC have a share of 46 percent and are common with the citizens on average between 64.8 million VND and 87.9 million VND per month, whilst the VP is associated with customers with the highest income, for example, 347.5 million VND per month on average.

Type of card	Income (VND/month)	Customer amount	Percent(%)
VP	347,487,790	327	9%
VV	215,864,954	243	7%
OTHER	171,826,314	370	10%
MG	131,383,244	426	12%
VG	87,944,515	855	24%
VC	64,843,408	787	22%
VA	33,580,597	120	3%
MS	19,629,088	399	11%
TOTAL		**3,527**	**100%**

Figure 1: Type of products that consumers are concerned

Source: Big data of the S Bank

Note of product types:

VG:	*VISA GOLD*
VC:	*VISA CLASSIC*
MG:	*MASTER GOLD*
MS:	*MASTER STANDARD*
VP:	*VISA PLATINUM*
VV:	*VISA VIOLET CARD*
VA:	*VISA AUTO CARD*
OTHER:	*JP = JCB Platinum; JC = JCB Standard; JETSTAR = JCB Credit Jetstar; JG = JCB Gold; VB = Visa Business; MW = Master One World; PB = Master Passbook Card EEFA = UAFA Card.*

Active customers who have the length of stay (LOS) in the bank service must be more than 12 months. This requirement is to prevent the prejudice of customers with LOS of less than 12 months, as these customers cannot value the services of the bank if their services are only a few months like two or three or so on. Accordingly, the data result that the customers of LOS with the range of >12-31 months account for the highest percent of 37.55, next as the range of >31-45 months occupies 20.7 percent, next as the range of >45-52 months with 16.9%, the range of >52-65 months with 13.4%, and the range of >65-81 months with 11.4%. According to descriptive statistics, three types of products approached by customers, with customers borrowing for personal use accounting for 54 per cent, customers using a credit card accounting for 43 per cent, and the last offering a family business loan for 3 per cent.

Customers are allowed to request one main card from the bank, including both main and extra cards for their demand. The main card is the customer's primary use for regular purchases, e.g. cash withdrawal, Visa card, etc., while the extra cards are additional use for secondary demand, e.g. payment card for fuel, travel, additional services, etc. According to bank regulations, one customer can use one or more cards issued by the bank due to transaction demand. Accordingly, the customers who use both main and extra card accounting for 22%, while 78 percent is occupied by customers with only main cards. Customers who are using both main card and extra cards at the same time are popular for directors with 35 percent of them concerned, less as workers (27 percent), manager (25 percent), and staff (13 percent).

At the same time, the use of two cards is an expression of users ' needs. The growing trend of cross-selling strategies that leverage all consumer demands in the future, in particular for high-income groups and high positions, should also have been noted in this result.

RMF and Market Segment

As stated previously, applying the model of RFM is to analyze customer behavior, its application is based on Wu & Lin (2005) and Khajvand, Zolfaghar, Ashoori, & Alizadeh (2011) to develop the matrix by a combination between LOS and Recency.

Concerning the matrix of combination between LOS and Recency of WU & Lin (2005), the result is derived and presented in Figure 2. There are four cells with paired interaction, each one has a number of customers with card usage as follow

(1) CLOSE RELATIONSHIP (CR): it is an integration between High of LOS and LOW of Recency, so called "potential relationship", accounting for 1,777 customers, and equivalent to 50.38 percent of total customers. The card is the best consideration of this group as Visa Gold (VG), accounting for 26.1 percent, next as Visa Class (VC) with 16.6 percent, Master Gold (MG) with 14.1 percent, Visa Platinum (VP) with 12.2 percent, and Master Standard (MS)

(2) LOST RELATIONSHIP (LR): it is integration between LOW of LOS and LOW of Recency, so called "lost relationship", accounting for 49.22 percent of total customers. The card of Visa Class (VC) is the best consideration of customers with the share of 28.2 percent, next as Visa Gold (VG) card with 22.4 percent, Master Standard (MS) with 10.5 percent and Master Gold (MG) with 105.

(3) POTENTIAL RELATIONSHIP (PR): it is an integration between LOW of Recency and HIGH of LOS. However, this PR has only 6 customers. As a result, the following analysis will not focus on this group, due to very few customers

(4) ESTABLISHING RELATIONSHIP (ER): It is an integration between HIGH of Recency and LOW of LOS. However, there are only 8 customers in this cell. As a result, the following analysis will ignore this group.

As found in the matrix, we can see a good picture and bad picture of the bank, in which the good picture that the bank has 50.38 percent customers belong to "Close Relationship", the bad picture that the bank has 49.22 percent customers belong to "Lost

Relationship". This finding brings an important message of alarming churn of customers if the bank does have an urgent strategy to retain the group of "Lost Relationship".

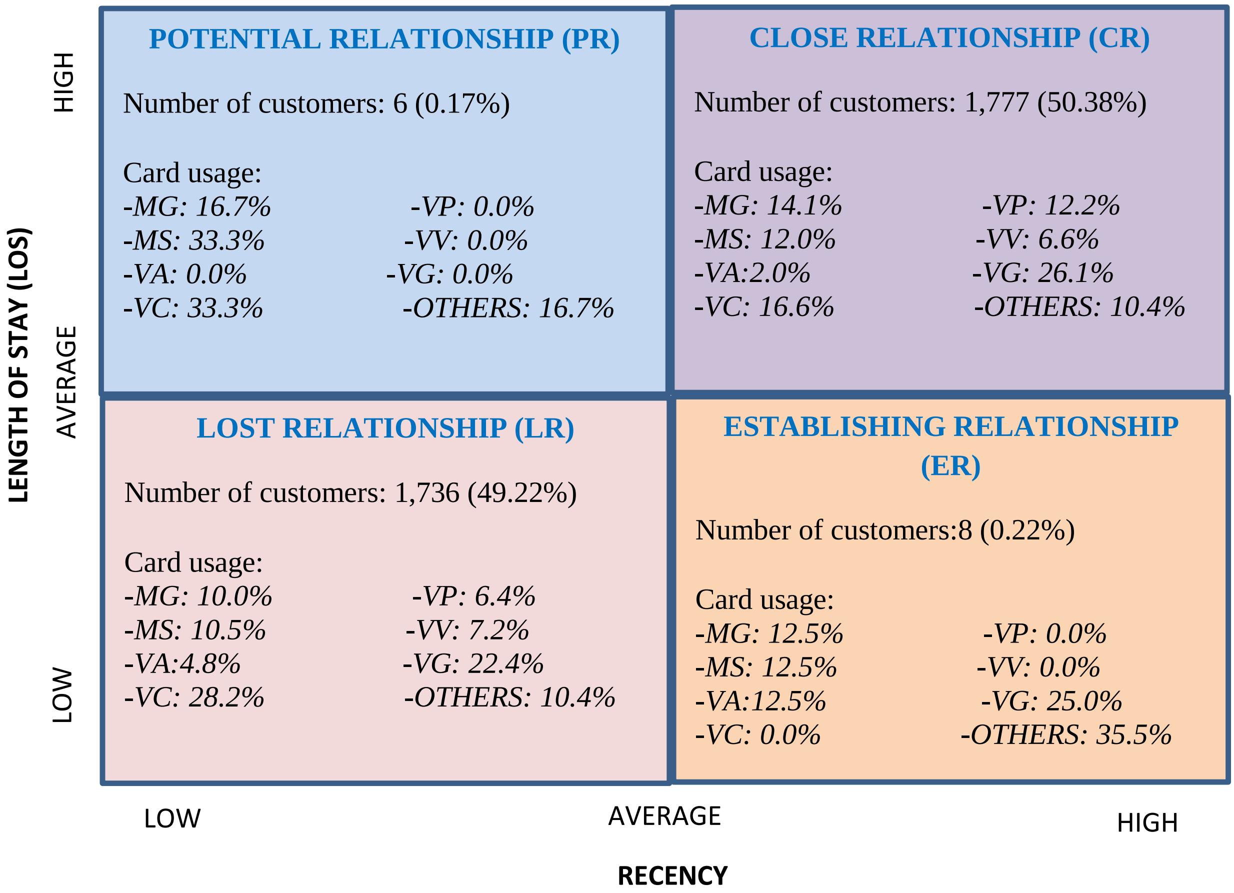

Figure 2: Matrix of LOS and recency
Source: Big data of the S Bank

Note of product types:

VG:	VISA GOLD
VC:	VISA CLASSIC
MG:	MASTER GOLD
MS:	MASTER STANDARD
VP:	VISA PLATINUM
VV:	VISA VIOLET CARD
VA:	VISA AUTO CARD
OTHER:	JP = JCB Platinum; JC = JCB Standard; JETSTAR = JCB Credit Jetstar; JG = JCB Gold; VB = Visa Business; MW = Master One World; PB = Master Passbook Card EEFA = UAFA Card.

RFM and its market segment

As defined previously, indicators of RFM are employed as inputs of the cluster method, which we can find consumers grouped into market segments.

Recency: one of the most important elements identifies the potential customers with the highest transaction value at the bank not far from the present. The most current transaction of a customer can confirm that she or he may be using services from the bank again, or they are new customers.

Frequency: it can play as a second element to measure frequency of customers who visited and took transactions at the bank during the period of study. The higher frequency, the higher chance the customer can buy more products or use more services from the bank.

Monetary: it is measured by the total value of balance account that customers have transactions at the bank during the period of the study. Once the higher value of the balance account is presented, it is a signal that customers have spent to buy more service or product. The more they spend, the more they are much attracted by the services of the bank.

In fact, RFM is popularly used to segment market, where the method of cluster analysis is applied, directly K-Means. As resulted in Figure 3, there are 3,527 customers grouped to 5 clusters, each cluster displayed as a market segment.

According to the method of RFM, the value of each element in RFM is ranked by a five-point scale, with 1 being extremely poor, 5 being extremely high. But R (recency) is opposite, which 1 being the time of the last buying is far away from the current time of study, 5 being the time of last buying is the closest to the current time of study. As a result, values of each element of RFM are mean values, the higher value, the higher likely.

Cluster	Label	Description	Size	Inputs		
Cluster-1	R=LOW; F= LOW; M= HIGH	Spender group	38.8%	Recency (3.01)	Frequency (1.00)	Monetary (3.99)
Cluster-2	R=HIGH; F= HIGH; M= LOW	Shopper group	10.3%	Recency (4.15)	Frequency (3.90)	Monetary (1.93)
Cluster-3	R=LOW; F= HIGH; M= LOW	Frequent group	14.8%	Recency (1.42)	Frequency (3.98)	Monetary (2.77)
Cluster-4	R=LOW; F= LOW; M= LOW	Uncertain group	27.5%	Recency (1.22)	Frequency (1.00)	Monetary (1.16)
Cluster-5	R=HIGH; F= HIGH; M= HIG	Best group	8.5%	Recency (4.01)	Frequency (4.32)	Monetary (4.65)

Figure 3: Marketing segment based on RFM

Source: Source: Big data of the S Bank

To cluster 1: it is defined as a spender group, in which customers have the low recency (R=LOW) and low frequency (F=LOW), while the monetary is high (M=HIGH). This means that the customers have a high balance account, due to the large loan during the study period during 12 months. As a result in Figure 3, the customers belong to this spender group accounting for 38.8 percent. The bank should consider these customers to avoid a risk of payment capacity.

To cluster 2: it is defined as a shopper group, which customers have the high recency (R=HIGH) and the high frequency (F=HIGH), while the monetary is low (M=LOW). This means that these customers often have visiting frequency for transactions, particularly for current time. However, transaction times have low shopping as low loans. According to Figure 3, the customers belonging to the shopper group account for 10.3 percent. The bank can check transaction behavior as a profile of customers to select good ones to attract them more shopping at the bank.

To cluster 3: it is defined as a frequent group, in which customers have the low recency (R=LOW) and the low monetary (M=LOW), while the frequency is high (F=HIGH). This means that the customers take transactions for payment or doing something else. May be said that this group has a good responsibility for loan payment. The bank should take care of these customers to attract them for new services, because they are quite potential and not risky. As depicted in Figure 3, the cluster of the frequent group accounts for 14.8 percent.

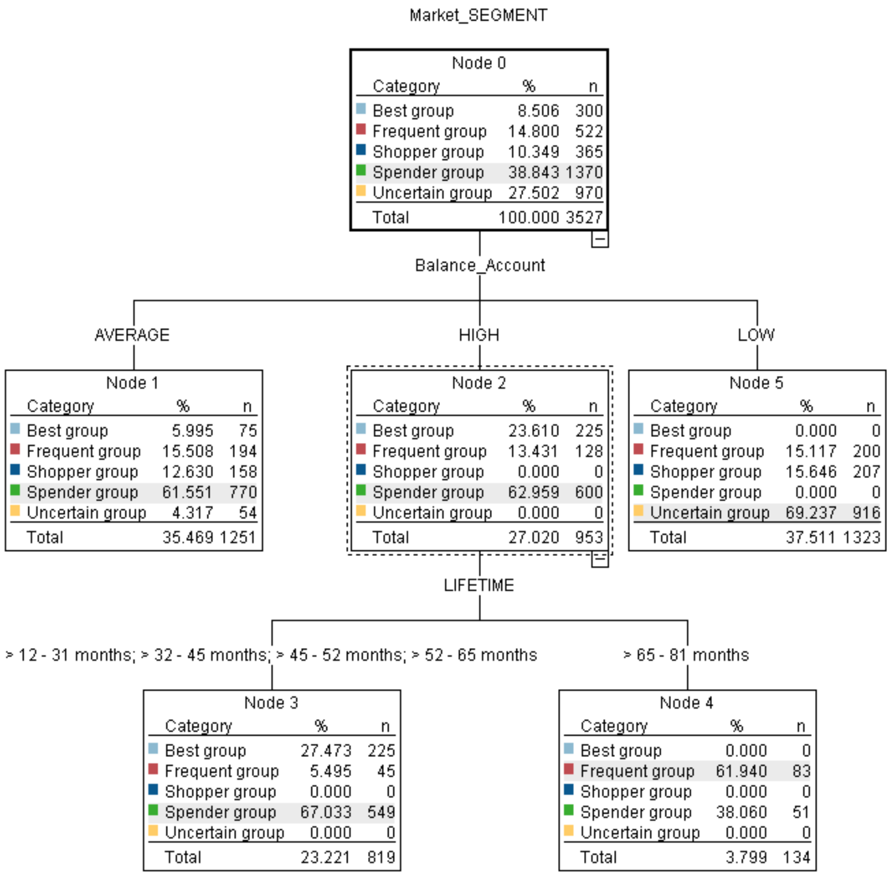

Figure 4: Result of Tree Decision of five market segments
Source: Source: Big data of the S Bank

To cluster 4: it is defined as an uncertain group, in which customers have the low recency (R=LOW), the low frequency (F=LOW), and the low monetary (M=LOW). As a result, the customers of this cluster spent not much on buying during the research study, and they did come back to visit for transactions very few. This group accounts for 27.5 percent (Figure 3).
To cluster 5: it is defined as a best group, which the recency, frequency, and monetary of customers all are high (R=HIGH, F=HIGH, M=HIGH). This means that customers have highly current transactions, more visiting frequency to the bank and high transaction values. The contribution of this best group is valuable, so the bank should have premium programs for these customers to retain them and make them to be loyal in the future.
Applying the method of Tree Decision as mentioned derives the result shown in Figure 4 Accordingly, there are five nodes, each node presents descriptive statistics. Node 1 presents the average balance account, which the customers are spenders accounting for the highest percent of 61.55 percent, equivalent to 770 customers, while the best group occupies almost 6 percent.
Similarly, node 5 is the low balance account, which the customers are uncertain for their transaction, accounting for the highest percent of 69.24 percent, equivalent to 916 customers. Reasons of customers with low balance accounts happened in the market segment of the uncertain group may be that they are workers or staffs often have a low loan.

5. CONCLUSION AND IMPLEMENTATION

Based on data mining of 3,527 customers from the source of big data, the mathematics methods of K-Means cluster and Decision Tree are taken into account. Initially, the overall image of the S Bank's real condition is portrayed by descriptive statistics. The S Bank's service is used by four major categories of clients, such as, directors, managers, staff, and some of workers. These customers have a long period of loyalty to the bank's service, with 60 per cent in the length of stay (LOS) of the customer remaining with the bank for more than 3 years, with about 11 per cent in LOS for 5 years and above.
Many consumers pay greater heed when utilizing two cards, the so-called main card and the extra card. At the same time, the use of two cards is an expression of users ' needs. The growing trend of cross-selling strategies that take advantage of all consumer

demands in the future , particularly for high income groups, and have high positions should also have been noted in this overall result.

These cards are grouped into eight main groups, such as VG (Visa Gold), VC (Visa Classic), MG (Master Gold), MS (Master Standard), VP (Visa Platinum), VV (Viva Violet Card), and VA (Visa Auto Card). VG and VC, however, are both popular and concerned with the highest. Clients of the top two products, e.g. VG and VC, which appear to be more interested in directors for VG, and VC are more interested in staff. Moreover, consumers that use VG have higher income than VC.

With the application of K-means analysis, in which cluster analysis involves three components of R, F, and M. Five segments are constructed on that basis. The uncertain group is a bit high at 27.5 percent, while 8.5 percent of the customers are the best overall. Moreover, many clients also come to the bank for small transactions.

The findings carry a significant message, the best group is populated by customers who are directors and managers, while the customers belonging to the unpredictable group are mostly workers, since the income level of employees is limited and unstable, this is consistent with practices. In order to build suitable plans and stimulate more transaction values in the future, the results are good for the bank. The message, on the other hand, also confirms that workers & staff are insecure, not unexpectedly, since these people are low-income, particularly for employees who are often unstable.

Moreover, the results of the business segment show that the high percentage of consumer purchase spending is accounted for by credit card customers and personal use customers. This means that these clients often have an immense amount of transactions. As seen, the demand characteristics of five consumer segments are extremely accurate since these results are measured on the consumer usage experience reported in the framework. To leverage the market efficiently, the bank's marketing plan needs to build typical campaigns to meet the demand of the consumer.

The derived client segmentation is focused on transaction functionality. This finding contributes to the planning, on the basis of its transaction, of a system for segmenting customers to estimate the potential value of the various segments of customers in the bank, in particular for the retail banking sector. In addition, empirical research findings may provide feedback on marketing strategies, develop promotional programs to introduce new products for each category, and stimulate the most profitable consumer group's consumption.

REFERENCES

1) Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications, 36*, pp.4176-4184.
2) Han, J., & Kamber, M. (2006). *Data mining concepts and techniques* (2nd ed.). Burlington, USA: Morgan Kaufimann.
3) Jaiswal, A. K., Niraj, R., Park, C. H., & Agarwal, M. K. (2018). The effect of relationship and transactional characteristics on customer retention in emerging online markets. *Journal of Business Research,* 92, pp.25-35.
4) Khajvand, M., & Tarokh, M. J. (2011). Estimating customer future value of future value of different customer segments based on adapted RFM model in retail banking context. *Procedia Computer Science,* 3, pp.1327-1332.
5) Kumar, V., Pozza, I. D., & Ganesh, J. (2013). Revising the satisfaction-loyalty relationship: Empirical generalizations and directions for future research. *Journal of Retailing,* 89 (3), pp.246-262.
6) Liu, D. R., & Shih, Y. Y. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime. *Information & Management,* 42, pp.387-400.
7) Meta Group. (2001). *Application Delivery Strategies*. East Coast, USA: Meta Group Inc.
8) DataPlus Millenium (2001). *Data-driven analysis tools and techniques.* KC, USA: DataPlus Millenium
9) Srivastava, U., & Gopalkrishnan, S. (2015). Impact of big data analytics on banking sector: Learning for Indian Banks. *Procedia Computer Science,* 50, pp.643-652.
10) Szus, I., & Tóth, Z. (2008). Customer loyalty problems in retail banking. Budapest, Hungary: MEB 2008-6th International Conference on Management, Enterprise and Benchmarking.
11) Wang, X., & Keh, H. T. (2017). Consumer susceptibility to cross-selling persuasion: The roles of self-construal and interpersonal harmony. *Journal of Retailing and Consumer Services,* 34, 177-184