ABSTRACT

When Taiwan joins World Trade Organization and National Health Insurance (NHI) program has been implemented, dental services marketing has become crucial and competitive in Taiwan. This study uses RFM (recency, frequency, and monetary) model along with self-organizing maps to segment dental patients of a children’s dental clinic in Taiwan. Four clusters are recommended for the overall 1,532 patients. The average values of R and F variables are computed for each cluster and the overall patients, excluding monetary which is covered by NHI program. The results show that one cluster with both R and F values (506 patients) greater than the overall average R and F values can be viewed as loyal patients, and the dental clinic needs to pay much attention to this group. One group with 238 patients have larger recency value but the other two clusters with 788 patients have smaller average R and F values.

Keywords: dental marketing, RFM model, self-organizing maps, market segmentation, loyal customer.

INTRODUCTION

Capon [6], Wang et al. [48], and Rutsohn and Ibrahim [43] have addressed that marketing strategies have been an essential issue for medical care industry particularly in dental care. There is no exception to Taiwan’s medical care industry because this industry has rapid growth when Taiwan joins World Trade Organization. In addition, with the introduction of the new total payment rule of National Health Insurance (NHI) program executed by Bureau of National Health Insurance of the government in Taiwan since March 1995, Lee et al. [33] and Shieh et al. [44] summarized that the medical care industry has become very competitive, including dental services.

Lee and Shih [34] pointed out that this cost-containment mechanism and global budgeting systems for dental care covered as part of the benefit package in NHI program in Taiwan are unique. A dentist’s income is limited when serving a dental cares covered by NHI program as NHI program will cover most of cost except for the co-payment and registration fee per visit. Under such system and competitive environment, it is critically important for dent clinics to actively identify profitable customers and to retain important customers so as to market dental services.

RFM (Recency, Frequency, and Monetary) model has been widely applied for market segmentation to identify valuable customers [8] [11] [50] [52]. Before the use of RFM model, the number of clusters must be determined. In this study, self-organizing maps (SOM) will be used to segment customers since SOM can help the market managers easily recognize the market segments precisely and compare market maps over time and monitor market responses of every segment [19]. In this case study, the focus is on a children’s dental clinic with 1,532 patients who visited this children’s dental clinic from July 16, 2009 to July 15, 2010 for at least once. The profile for each patient includes the membership number, gender, birth date, the last visit date, and visit frequency. Because the dental care is included in NHI program such that monetary value is proposed to be fixed in this study.

This paper is organized as follows. Section 2 briefly reviews self-organizing maps and RFM model. A case study in a children’s dental clinic in Taiwan is presented in Section 3. Finally, conclusions are drawn in Section 4.

LITERATURE REVIEW

Self-organizing Maps

Clustering techniques are one of the data mining techniques and very practical to divide all customers into appropriate number of clusters based on some similarities among these customers [10] [18] [22] [32] [51]. The philosophy is to both minimize within-group variation and maximize between-group variation in accordance with a distance or dissimilarity function to identify a set of groups. Self-organizing maps is one of the very
common clustering techniques applying unsupervised neural network method to clustering, visualization and abstraction, and market screening [16] [19] [21] [24] [28] [49]. Zhang et al. [55] pointed out that SOM is a competitive and cooperative neural network that the theory is motivated by observing the operation of the brain [28].

Self-organizing maps is trained by an unsupervised competitive learning algorithm and can automatically detect strong features in large data sets. While self-organizing maps maximizes the degree of similarity of patterns within a cluster and minimize the similarity of patterns belonging to different clusters, SOM can produce two-dimensional arrangement of neurons from the multi-dimensional space [1] [28]. Thus, the complexity of the information from multi-dimensional space can be simplified and understood much easier. In addition, SOM network includes an input layer and an output layer [21]. Originally, the patterns of self-organizing maps in a high-dimensional input space are very complicated. After SOM maps similar input vectors onto similar output units on a two-dimensional map, its structure on a projected graphical map display becomes more transparent and more understandable. Therefore, self-organizing maps provides the recognizable technique of the inherent relation between the input and output [12] [13] [24] [55]. More importantly, this unsupervised method can be applied to cluster data without prior knowing the class memberships of the input data [19].

For more information about self-organizing maps, please refer to Kohonen [30] [31] and Yin and Allinson [54].

**RFM model**

RFM model is a well-known customer value analysis method widely applied for market segmentation [8] [9] [11] [52]. It is a behavior-based model to analyze the behavior of a customer and then make predictions in accordance with the behavior in the database [23] [53]. RFM models is composed of three measures, namely recency, frequency, and monetary, and the definitions are described below [50]. Recency is defined as the number of periods such as days, months, or years since the last purchase. Frequency measures the number of purchase made in a given time period, while monetary measures the total amount of money spent during a given period of time or the average dollar amount per purchase or all purchases to date.

The general way to use RFM model in customer behavior analysis is to sort the customer data by each dimension of RFM variables and then divide the data into five equal quintiles. For recency, the customer database is sorted by purchase dates by descending order. Thus, the top segment is given a value of 5 and the others are descendedly assigned of 4, 3, 2, and 1. For frequency and monetary, sorting customer visiting frequency data and the customer data related to the amount of the money spent in descending order, respectively. The top 20% is assigned the value of 5. The value of 4 is given to the next 20% and so on [23] [26] [47]. These three variables belong to behavioral variables and can be acted as the segmenting variables by observing customers’ attitudes toward the product, brand, benefit, or even loyalty from the database. Marcus [40] suggested that using average purchase amount instead of total accumulated purchase amount is better in order to reduce co-linearity of frequency and monetary. Finally, all customers are presented by 555, 554, 553, …, 111, which thus creates 125 (5×5×5) RFM cells. Moreover, the best customer segment is 555, while the worst customer segment is 111. Based on the assigned RFM behavior scores, customers can be classified into segments and their profitability can be further analyzed [4] [5] [9] [42].

Miglautsch [42] pointed out that the above RFM scoring method is called customer quintile method, which is to sort customers in descending order, i.e., from the best to the worst. The advantage is to yield equal number of customers in each segment. This method, however, has a major disadvantage. It encounters several scoring challenges in the measure of frequency and is relatively sensitive, which leads to break apart customers who have identical behavior at the lower quintiles but group customers together whose buying behaviors have significant differences [2].

Behavior quintile scoring method developed by John Wirth is another approach to sort customers based such that each quintile may have different number of customers. For instance, frequency can be divided into five intervals, including 0-3 months, 4-6 months, 7-12 months, 13-24 months and 25+ months, which are coded as 5, 4, 3, 2, and 1, respectively [41]. Miglautsch [42] proposed an approach and the philosophy is below. The score of frequency is defined that the single purchasers are assigned as a score of 1. Then, average the remaining frequency values to determine the mean. Once a customer’s total frequency value is lower than the mean, a score of 2 is given to this customer. The process may be repeated more than two times. For monetary, five quintiles are still created and each has equal amounts of sales [50].
In addition to use the value of each cell to judge whether the customer is valuable, Liu and Shih [37] [38] and Sohrabi and Khanlari [45] suggested that the combinations of RFM can be used by assigning ↓ or ↑ based on the average R (F, M) value of a cluster being less than or greater than the overall average R (F, M) value. Under such circumstances, eight segments can be created. The composite value of RFM is obtained via multiplying normalized RFM values of each customer and the weight of RFM variables.

RFM model measures when people buy, how often they buy, and how much they buy. By using the past purchase data, firms can identify which customers might be worthy to be contacted and predict their future purchase behaviors. Sohrabi and Khanlari [45] and Colombo and Jiang [15] concluded that RFM models are often developed to target marketing programs such as direct mail for particular customers in order to improve response rates, revealing that RFM facilitates to choose which customers to target with an offer. Hughes [23] and Kahan [26] stated that firms can receive much benefit from RFM model including encompassing increased response rates, lowered order cost, and greater profit.

RFM model has been widely applied in many practical areas such as nonprofits, financial organizations, government agencies, on-line industry, travel industry, telecommunication industry, and marketing industry [17] [20] [25] [29] [35] [36] [39] [45] [46]. Moreover, RFM model can be used to segment customers, calculate customer value, compute customer lifetime value, observe customer behavior, estimate the response probability for each offer type, and evaluate on-line reviewers.

Several studies have been successful reported about RFM model in identifying customers and analyze customer profitability. Kaymak [27] used RFM variables as features for characterizing the customers when examining how fuzzy clustering can be used to obtain target selection models. Tsai and Chiu [47] introduced a novel purchased-based market segmentation methodology in accordance with product specific variables such as the purchased items and associated monetary expenses from transactional customer histories to improve unreliable segmentation result due to the traditional adoption of general variables such as customer demographics and lifestyle to segment a market. Jonker et al. [25] provided a decision support system to determine mailing frequency for active customers. The system observes the mailing pattern of customers in terms of RFM variables and provides mailing policies for multiple time periods. The mailing decision process is modeled through a Markov decision chain. Lumsden et al. [39] applied RFM model to distinguish customer value in accordance with pre-purchase motivations of membership initiation in and all-inclusive travel vacation club. Chan [7] proposed an approach that combines customer targeting and customer segmentation for campaign strategies using RFM to identify customer behavior and a customer lifetime value model to evaluate proposed segmented customers through examining Nissan automobile retailer.

CASE STUDY

This case study is based on a children’s dental clinic in Taiwan to identify profitable customers. This dental clinic begins its operation since September 17, 1995. In addition, the patients must be less than 18 years old in order to be classified as children by definition. When the patients become 18 years old, they must go to an adult dental clinic instead. Thus, the “age” factor in this children’s dental clinic should be taken into account. This study collects the data set with 1,532 patients who visited this clinic from July 16, 2009 to July 15, 2010. The profile for each patient consists of the membership number, gender, birth date, the last visit date, and visit frequency. Monetary value for each patient is excluded in the profile as the majority of a dentist’s income is from the co-payment and registration fee per visit by NHI program in Taiwan [34]. Therefore, monetary value is proposed to be fixed. Under such circumstance, only R and F variables are taken into account. The definitions of L and R variables are as follows. Recency is defined as the number of days since the last visit from July 16, 2009 to July 15, 2010. In this study, the assigned value on July 16, 2009 is one, and the value on July 15, 2010 is 365. Therefore, the smallest number of recency is one, while the maximum value is 365. Frequency refers to the number of visit in the same period of time. Thus, the frequency can be measured by counting the number of times a particular patient visits the dental clinic from July 16, 2009 to July 15, 2010.

Among 1,532 patients, 779 patients are male, while 753 patients are female. The majority of patients fall in the ages of 6-10, where the maximum and minimum ages are 17 and 1, respectively. The descriptive statistics for R and F variables of the entire group are presented in Table 1, where max, min, average, and STD represent the maximum value, minimum value, average value of the entire group, and the standard deviation of the entire group, respectively. For the recency, the larger the recency value is, the more recent the patient visits. The maximum and minimum recency values are 365 and 1,
respectively. For the frequency, the maximum and minimum values are calculated to be 24 and 1, respectively.

Table 1 The descriptions of recency and frequency

<table>
<thead>
<tr>
<th></th>
<th>Max</th>
<th>Min</th>
<th>Average</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recency</td>
<td>365</td>
<td>1</td>
<td>239.77</td>
<td>104.74</td>
</tr>
<tr>
<td>Frequency</td>
<td>24</td>
<td>1</td>
<td>3.14</td>
<td>2.54</td>
</tr>
</tbody>
</table>

In order to segment patients into appropriate number of clusters, SOM is applied to determine the number of clusters. Based on Figure 1, four clusters are recommended among 1,532 patients when recency and frequency are the two input variables. Table 2 provides descriptive statistics of these four clusters, including the sample size, gender distribution, average age, average recency, and average frequency. Cluster 1 has the largest number of patients with 506, while Cluster 4 has the smallest number of 238. In addition, the patients in Cluster 1 is the youngest patients with the average age of 7.28 among four clusters.

Figure 1 Four clusters recommended by SOM

Table 2 Descriptive statistics of four clusters based on SOM technique

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Patients</td>
<td>506</td>
<td>472</td>
<td>316</td>
<td>238</td>
<td>1,532</td>
</tr>
<tr>
<td>Male Patients</td>
<td>272</td>
<td>228</td>
<td>156</td>
<td>123</td>
<td>779</td>
</tr>
<tr>
<td>Avg. R</td>
<td>296.4</td>
<td>172.8</td>
<td>230.8</td>
<td>264.2</td>
<td>239.8</td>
</tr>
<tr>
<td>Avg. F</td>
<td>5.91</td>
<td>1.00</td>
<td>2.00</td>
<td>3.00</td>
<td>3.14</td>
</tr>
<tr>
<td>Avg. age</td>
<td>7.28</td>
<td>8.99</td>
<td>8.87</td>
<td>8.65</td>
<td>8.35</td>
</tr>
<tr>
<td>Item(s) above average</td>
<td>RF</td>
<td>---</td>
<td>---</td>
<td>R</td>
<td></td>
</tr>
</tbody>
</table>

Based on Table 2, Cluster 1 has the largest R value and F values among four clusters. Cluster 4 has the second largest R value, which is also larger than the average of the entire group, and has the second best F value, which is less than the average. In contrast to Clusters 1 and 4, Cluster 2 has the lowest R and F values among four clusters. When the resources are limited, this dental clinic might put less on these patients since they have not been visited for more than six months. This might indicate that these patients might turn to other dental clinic for services. Cluster 3 has slightly lower than the average R value but the second lowest F value. This shows they have not been to this dental clinic for more than four months. However, compared with Cluster 2, these patients tend to visit the dental clinic more. Therefore, the dental clinic might use some resources to encourage patients to come back.

Cluster 4 has relatively high R value but slightly lower than the average F value. This might indicate that they are new customers. Pertaining to the marketing strategies for customers in Cluster 4, the dental clinic should enhance customer relationship management to keep in touch with the patients and find ways to meet their demand by attracting them to visit more often. For example, special attention and treatment is particularly vital to children patients when receiving dental care [3].

The patients in Cluster 1 have both largest R and F values, far better than the average values. These patients might be the core customers since they have a relatively closer relationship with the dental clinic by visiting more recent and more frequent. They are loyal both in attitude and behavior in the dental clinic. Loyal customers are profitable as they would contribute positively to the success of their dental care experience. Therefore, the dental clinic should focus on the patients in Cluster 1.

CONCLUSIONS

Marketing expertise plays a vital role of dental clinics in remaining viable in the dental market [34]. Customers believe dental advertising can be done according to each unique taste and dentists should market their services [14]. A successful marketer must identify the needs of customers in different segments [6]. This paper uses RFM model along with self-organizing maps to classify patients and target important patients in a children’s dental clinic in Taiwan. Four clusters are formed by SOM. Patients in Cluster 4 are new customers. The strategy for this dental clinic is to strengthen the interaction with the patients in order to maintain a long-term dental care relationship. Besides, the dental clinic can make more dental advertising or adopt some promotional channels such as offering discount for recommending friends and free registration fee to enhance the number of the visit of patients. The patients in Cluster 1 are the most important patients for the clinic. They would be regarded as
the highest loyal customers in the clinic, revealing the necessary for the dental clinic to committing more resources to the patients. To increase the frequency of the patients, the dental clinic can provide better after-medical care activities and small gifts with the provision of the medical services.

Acknowledgement

This study was partially supported by the National Science Council with grant number of NSC 99-2221-E-018-012-MY2.

REFERENCES


The 11th International DSI and the 16th APDSI Joint Meeting, Taipei, Taiwan, July 12 – 16, 2011.

