Analysis of Purchasing Behaviour of Customers in Restaurants Using RFM Model of Data Mining

Oluwafemi Elijah OJO

Department of Business Administration and Management

Rufus Giwa Polytechnic Owo, Ondo State, Nigeria

Tel: +2348035287856   E-mail: adebayoga2016@gmail.com

Abstract

This study is focused on identifying existing customers who are most likely to respond to a new offer especially in service businesses. A restaurant with about 180 customers was used as a case study. A direct marketing model referred to as: recency, frequency, and monetary (RFM) analysis was used to identify customers that score 3 or more points for the RFM score. The analysis shows that 100 customers were qualified and these identified customers would be used by the restaurant management to plan its future advertising and promotion strategies. Entrepreneurs in service businesses are advised to make use of the various direct marketing models to help them as a means to an end in planning their future advertising and promotion strategies for overall profitability of their businesses.

Key words: RFM Analysis, Purchasing Behaviour, Restaurants.

Introduction

Customers have always been and will always be a factor determining the profitability of many businesses. In today's competitive world, no business organization is totally immunes from competitive challenges. Attracting, retaining and securing customers loyalty are onerous task and good marketing strategy must be developed and maintained.(Joshi et al, 2016). There has been much literature also on relationship marketing. A shift is presently on practicalising some direct marketing models in using the past behaviour of customers to analyse future buying pattern especially by businesses that introduce innovation and update with time on their products.

Restaurants are one of such businesses that need to research into customer value analysis and the model usually employed is the RFM analysis- a type of direct marketing tool. In RFM model, three parameters- recency,

frequency and monetary (RFM) are used in clustering classifying and association. Clustering helps to find out customer segment with comparable RFM values, while classification and association are used respectively to assess customer's future buying pattern and product recommendation. Product recommendation guarantees that some customers are most likely to embrace new innovation and product update.

Statement of the problem

The management of a restaurant providing lunch as part of its menu services with provision of tables where customers sit down to eat the lunch, and served by a waiter, is interested in introducing new standard lunch package as different from the existing menus. Specifically, to introduce (1) **Fried Rice with Vegetable + Fish**, (2) **Plantain Porridge with Crayfish** and (3) .The management of the restaurant wants to know those customers that would likely embrace the new innovation. How can these customers be identified?

Objective of the Study

The objective of this paper is:

1., to provide an analytical framework for the restaurant to understand the business customers, segment them into various groups based on their engagement with the business and enable the business to have a better mean to determine which customers should the

restaurant management target with their marketing campaign through the use of RFM Analysis.

Review of related Literature

The development of new business concepts like customers satisfaction, loyalty and relationship marketing have been much focused. These have obliged businesses to segment their markets into smaller homogeneous groups. Some of the advantages as Fauladifar et al, 2016) conceive it was a grasp of marketing, advertising and product development strategies. Realising the attitudes of different categories of customers is a key point to enhance customer relationship within a business.

Relationship Marketing.-

The basis of relationship marketing philosophy is that the attraction of new customers is merely the first step in the marketing process; the important element is the retention of these customers. Kotler and Armstrong (1997) state that relationship marketing involves 'creating, maintaining, and enhancing strong, value laden relationship with customers and other stakeholders" Relationship marketing therefore, is concerned with creating, developing and enhancing relationships with

selected customer segments over times. This involves gaining a detailed understanding of the needs and behaviour of key customer segments and how they change over their life

cycles. Christopher, Payne and Ballantye (1994) compare transaction marketing with relationship marketing in the Table below:

Table 1.Transaction and Relationship Marketing Compared.

Transaction Marketing	Relationship Marketing
A focus on single sales An emphasis upon product features. Short time scales Little emphasis on customer retention. Limited customer commitment Moderate customer contact Quality is essentially the concern of production and no one else	A focus on customer retention and building customer loyalty An emphasis upon product benefits that is meaningful to the customers. Long time scales, recognizing that short term cost may be higher, but so with longterm profits. An emphasis upon high level of service, which are possibly tailored on the individual customer. High customer commitment High customer contact, with each contact being used to gain information and build the relationship. Quality is the concern of all and it is the failure to recognize this that creates minor mistakes, which led to major problems

Source: Christopher, Payne and Ballantye (1994) ;Oloidi (2001)

The RFM Method

The RFM acronym stands for recency, frequency and monetary value. Its goal consists in identifying and targeting the most attractive customers for a specific marketing project. This method is applied in order to select the segment of customers that will better respond to a particular marketing campaign. The application of the method

implies data about the past behavior of customers.

The RFM method may be implemented for the customer's existing in the portfolio of the company, not in the case of prospects. It helps an organization to predict the reaction of customers to an offer made by the organisation, based on dynamic data about their purchase history. Consequently, RFM method has as result a behavioral

segmentation of customers and the selection of those customers that will respond more favorably to the marketing proposition. The main idea on which the method relies is that past behavior is a better predictor of future behavior than customer declarations relative to their purchasing intentions and attitudes. (Spiller et al, 2005) The likelihood of response to a specific marketing campaign is more reliably estimated by data on historical behavior than by the reasoning and plans stated by customers.'

Basic RFM Concepts

Data mining: Data Mining (DM) is the procedure of dedicated searching of customers' transactions records and it is synonymous with extraction of complex and sometimes numerous range of raw data from a cumulative database. to in order to explore new patterns or rules. Data mining usually is conducted to reach two main objectives: first, predictions upon hidden patterns or relations between the quantities within the database and second, description which is actually the interpretation of database (Bham bri, 2011; Garg et al., 2008). Data mining involves searching, extracting, sieving using customers' past records.

Clustering: One of the most common DM methods is clustering upon which, the similar quantities are collected into same clusters

(Garg et al., 2008). In fact, it is about finding the sets in heterogeneous data by minimizing some measure of dissimilarity (Amiri and Fathian, 2007). Clustering is based on measuring Euclidean distance between two data observations and performs disjoint analysis computed from one or more variables which market researchers define.

Classification using RFM-Recently, integration of classification techniques and RFM was studied by Olson et al. (2009) to analyze customers' response possibilities to a specific product promotion.. Cheng and Chen (2009) also combined RFM attributes and rough set theory (the LEM2 algorithm) to mine classification rules that help enterprises finding out the characteristics of customers in order to strengthen customers relationship marketing (CRM). Ha (2007) used decision tree technique to track changes in RFM values of customers over time, to discover classification rules related to transition paths and thus to predict the next customers' RFM values from the current customers' RFM values. **Data Mining-**In data mining, association rules help in product recommendation to identify those customers that are most likely to embrace new innovation and product update.

How RFM Analysis Work - *Recency Score*-Customers are assigned a *recency* score based

on date of most recent purchase or time interval since most recent purchase. This score is based on a simple ranking of *recency* values into a small number of categories. For example, if you use five categories, the customers with the most recent purchase dates receive a *recency* ranking of 5, and those with purchase dates furthest in the past receive a *recency* ranking of 1. **Frequency Score**- In a similar fashion, customers are then assigned a **frequency ranking** with higher values representing a higher frequency of purchases. For example, in a five category ranking scheme, customers who purchase most often receive a *frequency* ranking of 5. **Monetary Score-** Finally, customers are ranked by monetary value, with the highest monetary values receiving the highest ranking. Continuing the five-category example, customers who have spent the most would receive a monetary ranking of 5. The maximum a customer can score is 555, that is 5 for recency, 5 for frequency and 5 for monetary.

The Restaurant Business

The objective of this paper does not include going into the management of restaurants and has been limited to managing customers for profit

A restaurant is a business which prepares and serves food and drinks to customers in exchange for money.(Wikipedia, n.d.) Meals are generally served and eaten on the premises, but many restaurants also offer take-out and food delivery services, and some only offer take-out and delivery. Restaurants vary greatly in appearance and offerings, including a wide variety of cuisines and service models ranging from inexpensive fast food restaurants and cafeterias to mid-priced family restaurants, to high-priced luxury establishments.. Some restaurants serve all the major meals, such as breakfast, lunch, and dinner (e.g., major fast food chains, diners, hotel restaurants, and airport restaurants). Other restaurants may only serve a single meal (e.g., a pancake house may only serve breakfast) or they may serve two meals (e.g., lunch and dinner) or even a kids' meal. Various **types of restaurant** fall into several industry classifications based upon menu style, preparation methods and pricing. Additionally, how the food is served to the customer helps to determine the classification.

Historically, *restaurant* referred only to places that provided tables where one sat down to eat the meal, typically served by a waiter. Following the rise of fast food and take-out restaurants, a "standard" restaurant was created coined as '**sit-down restaurant'** (Wikipedia, n. d.) Most commonly, "sit-down

restaurant" refers to a casual dining restaurant with table service, rather than a fast food restaurant or a diner, where one orders food at a counter. Sit-down restaurants are often further categorized, in North America, as "family-style" or "formal". The focus of this paper is on the sit- down restaurant group.

RFM and Restaurants

The RFM analysis will help the restaurants to locate the most valuable customers and we can go that extra mile to encourage them to spend with the restaurants. A particular restaurant was chosen as case study. The renown restaurant is having average customers of about 180 on records. A random sample of 15 customers were made using Random Numbers, to demonstration only since the expected characteristics of any

Table 2 Recency ranking would be given as follows: How recent within 60 service days.

Recency Ranking	Recency (Days since last purchase)
5	Less than 12 days
4	12-24 days
3	24-36 days
2	36-48 days
1	48-60 days

customer are all exhibited in this template. RFM Analysis was limited to customers who made purchases between March and May 2016. That is a total of 60 days for a 5 day per week services. Saturdays are not predictable because of socio-domestic challenges since service type is the **sit-down restaurants**. Services are limited to Monday - Friday every week. Customers' lunch patterns were studied for a period of three months based on the scoring and ranking in Tables 2-4.

Scoring the RFM Model

There is no hard and fast rule on scoring. Experts helping the particular organization would study how best to model scoring. Whichever criteria used, what is most important is that it should be uniformly applied.

Table 3 Frequency ranking would be given as follows: How frequent with 60 service days

Frequency Ranking	Frequency (Number of Lunch Taken)
5	48-60
4	36-48
3	24-36
2	12-24
1	< 12

Table 4 Monetary ranking would be given as follows. Cumulative amount spent for 60 service days.

Monetary Ranking	Monetary (Amount spent on lunch) N'00
5	9600-12000
4	7200-9600
3	4800-7200
2	2400-4800
1	< 2400

Table 5 is the outcome of some processes from the transaction records of the restaurant. The following steps were taken before arriving at the RFM scores in Column 6.

The following steps are taken to ensure eventual RFM analysis that would help to identify customers that would likely respond to standard lunch package and other innovation to restaurant services.

We compute RFM scores from transaction data. A transaction data is the records of all customers as they patronize the restaurants on daily basis, together with amount of lunch taken each time and the number of lunch between March and May 2016. A new dataset is created that includes the new RFM scores.

By default, the dataset includes the following information for each customer:

- Customer ID variable(s)

- Date of most recent lunch
- Total number of lunch
- Summary of total value of lunch for the three months in N'00 amount (the default is total)
- Recency, Frequency, Monetary, and combined RFM scores

The new dataset contains only one row (record) for each customer. The original transaction data been aggregated by values of the customer identifier variables. The identifier variables are always included in the new dataset; otherwise one would have no way of matching the RFM scores to the customers.

The dataset must contain variables that contain the following information:

- A variable or combination of variables that identify each case (customer).
- A variable with the date of each lunch
- A variable with the monetary value of each lunch.

Methodology

Data collection

A renowned restaurant is having about 160 customers that usually take lunch with them as the transaction records show. A table of random number was used to select 15 customers to demonstrate RFM Analysis out of the 180 customers that scored 3 point or more.

Model Specification

The RFM Analysis Model was used and had been explained.

Results and Discussion

The outcome of the RFM Analysis is in Table A-1. Column 1 is the customer's identification or simply account number. How the ID numbers are arranged in of little importance. In most cases, ID number follows chronological order of patronage. Column 2 follows the distribution of the recency ranking in Table 2 with the accompanied months. Likewise is Column 3 where the days recency in Column 3a is ranked in Column 3b depending on the lunch pattern of customers. Table 3 is also applied to frequency Column 4. Table 4 is for monetary Column 5. The result of the ranking in columns 3a, 4a, and 5a make the RFM score in Column 6. For example, the combined RFM score for each customer is simply the concatenation of the three individual scores, computed as: (**recency** x 100) + (**frequency** x 10) + **monetary.** The ranking column under

recency, frequency and monetary constitute the RFM scores

For example, the RFM score for ID-16 is (3x100 + 5x10 + 5) = 300 + 50 + 5 = 355. The same process is applicable to all other customers. To qualify, a customer must score at least 3 points for **recency.** Eight customers were qualified and such customers were ticked under column 7**.** The process of data mining, clustering, classification and association were all completed in Table 5. Therefore the marked customers are those that are likely to respond to new standard lunch package. This result would help the restaurants of the future advertising strategies since the customers purchasing behaviours have been analysed using the RFM analysis. Eventually, 100 customers out of the 180 customers that patronize the restaurant for lunch were qualified having been scored 3 points and above on the RFM Analysis.

Conclusion and Recommendation.

The RFM Analysis has been used to segment and cluster about180 customers patronizing a restaurant for lunch. Out of the 180customer, 100 of them were able to score 3 points or more. This will help the restaurant

management to effectively plan its future advertising strategies.

Entrepreneurs in service businesses are advised to make use of the various direct marketing models to help them as a means to an end in future advertising and promotion strategies for overall profitability of their businesses.

References

Amiri, B. and M. Fathian, 2007. Integration of Self-organizing feature maps and Honey Bee

Management. Optimization Algorithm for Market Segmentation. J. Theoretical. Application

of Information Technology.

Bhambri, V., 2011. Application of Data Mining in Banking Sector. IJCST, Mandi Gobindgarh,

Punjab, India, Vol. 2, Issue 2.

Chuang, H. & Shen, C. (2008). A study on the applications of data mining techniques to

enhance customer lifetime value - based on the department store industry,

Proceedings of the 7th International Conference on Machine Learning and Cybernetics, pp.

168-173, ISBN: 978-1424420964, Kunming, China, July 2008, IEEE.

Garg, K., D. Kumar and M.C. Garg, 2008. Data Mining Techniques for Identifying the Customer

Behavior of Investment in Life Insurance Sector in India. Intemat. J. Informat. Technolo.

Knowledge. Management., 1 (1): 51-56

Ha, S.H. (2007). Applying knowledge engineering techniques to customer analysis in the

service industry, *Advanced Engineering Informatics,* Vol. 21, No. 3, (July 2007) 293-

301, ISSN:1474-0346.

Khajvand, M. and M.J. Tarokh, 2011. Estimating Customer Future Value of Different Customer

Segments Based on Adapted RFM Model in Retail Banking Context. Elsevier Ltd., Procedia

Computer Sci., 3: 1327-1332.

Liu, D-R. & Shih, Y-Y. (2005). Integrating AHP and data mining for product

recommendation based on customer lifetime value, *Information & Management,* Vol.

42, No. 3, (March 2005) 387-400, lSSN:0378-7206.

Oloidi, G. A. (2001) Re-Engineering Relationship Marketing as Strategy to Face the Challenges of

Globalisation, Privatisation and Matured Markets in Nigeria. Nigerian Journal of

Development Issue: Education, Socio-Political and Economic Development, 5 (12), 1-22.

Olson, D.L.; Cao, Q.; Gu, C. & Lee, D. (2009). Comparison of customer response models,

Service Business, Vol. 3,(2), 117-130

Restaurants and Types (n.d.) in Wikipedia. Reviewed August 22, 2016 from http://en.wikipedia.org/wiki/restaurants

Spiller L., Baier M. - "Contemporary Direct Marketing". Pearson Education, Upper Saddle River,

New Jersey. 2005.

APPENDIX

Table A-1 RFM Analysis from Transaction Data in Records March to May 2016.

ID (1)	Days Recent (2) Most		Recency (in days since last lunch) (3)		Frequency(time lunch was taken) (4)		Monetary Score (5)		RFM Score (6)	Customer that would respond to new lunch packagewith (7)
	Days	Month	Least days (a)	Ranking (b)	Lunch Frequency(no of times) (a)	Ranking (b)	Amt (a)	Ranking (b)		
01	24-36	May/April	28	3	30	3	5000	3	333	✓
02	36-48	April/March	45	2	20	2	9800	5	225	
08	< 12	May	5	5	25	3	1100	1	531	✓
16	24-36	May/April	24	3	55	5	11000	5	355	✓
24	48-60	March	50	1	15	2	800	1	121	
28	48-60	March	58	1	34	3	6000	3	133	
30	36-48	April/March	40	2	30	3	7000	3	233	
35	24-36	May/April	30	3	20	2	2600	2	322	✓
49	24-36	May/A pr	25	3	10	1	12000	5	315	✓
72	24-36	May/April	36	3	27	3	4800	3	333	✓
75	48-60	March	55	1	36	3	900	1	131	
77	36-48	April/March	38	2	22	2	2000	1	221	
81	12-24	May	14	4	8	1	2200	1	411	✓
83	36-48	April/March	42	2	5	1	9000	4	214	
89	< 12	May	8	5	24	2	3200	2	522	✓