RFM-analysis as a Tool for Segmentation of High-tech Products' Consumers

Marina E. Tsoy, Vladislav Yu. Shchekoldin
Novosibirsk State Technical University, Novosibirsk, Russia

Abstract – **The method for high-tech products' consumer base analysis based on the RFM-analysis application involving classification methods is presented. Instead of the standard partition of the customer base into quintiles the analytical method of ABC-analysis is applied. It has been revealed that the distribution of customer segments is nonuniform. The approbation of the method was carried out on the example of a high-tech and office tools retail company.**

Index Terms – **Segmentation, consumer, RFM-analysis, ABC-analysis.**

I. INTRODUCTION

Last decade the development of high-tech and office tools market in the Russian Federation demonstrates the tendency of essential reduction due to the drastical changes in external economic environment. For instance, the desktop computer sales in Russia reduced on 38% in 2015 and equaled a bit more than 1.88 million items [1]. This situation has arisen due to a number of objective and subjective reasons such as the market development stagnation in the field of IT-technologies; the decline in purchasing capacity of the population and business sector which provokes the irregularity of demand; the absence of own production capacity (impossibility of the total import substitution); the shortage of components; the ageing of equipment; the mutual sanction policy, and other international problems.

In this connection the thorough understanding of consumer behavior could be beneficial in profitability of wholesale and retail trade industry of high-tech and office tools equipment. For this purpose it is necessary to involve the modern methods of client base analysis for customer's segmentation and the implementation of targeted methods and effective sales and promotion programs. It is of great importance especially for domestic markets because of essential distinctions in customer behavior in different regions of the Russian Federation.

RFM-analysis is one of the effective and widely used methods for the customers behavior description based on the clients' database study. It allows building the predictive models of customer behavior. In contrast to the traditional methods it does not require special knowledge of statistics and econometrics [2, 3]. Moreover, the effective application of this method is not necessary to involve the supplementary data. So, every company can perform the client segmentation by using RFM-analysis without any other expenses when conducts its own customer database.

II. PROBLEM STATEMENT

There are numerous studies of RFM-analysis implementation performed by foreign scientists [2-5]. RFM-classification development is based on the investigation of two type's characteristics of consumers: who they are (demographic data) and what they do (behavioral data).

The adequate forecasting is necessary to make effective management decision-making. At the same time the clear awareness of the previous consumer behavior could relieve the determination of his/her behavior in the future. It is important to realize that the RFM-analysis could be applied correctly only for the study of existing customers, but it could not able to predict the behavior of potential customers. The RFM-analysis could carried out for both customers from b2b- and b2c-markets for every area of economic activities that could be considered as a main it's advantage.

In RFM-analysis theory it is assumed that the client behavior (P) is described as a set of three basic characteristics:

$$P = < R, F, M >$$

where R – recency of the last purchasing – it is calculated as a difference between the current date and the date of last purchasing given in days; F – purchasing frequency – the total number of purchasing on observation period; M – monetary – the total amount of outlay money by a client on observation period.

The importance of R-characteristic implies that the purchasing resumption means some consistency in customer behavior. In other words this means a habit or it's acquisition which ensure the stable revenue to the company. When the customer gets rid of a habit his ties to the company will wane and the period between the last purchasing and the next one will rise steeply. As for the F-characteristic it is evident that the more a customer is purchasing the more likely he/she will do another one in the near future. Finally M-characteristic means that the more clients spend the more likely he/she will purchase again.

When using the classical approach of RFM-analysis it is suggested that the possible values for each variable divided into five intervals (quintiles) each of them been encoded with numbers from 1 to 5 according to the following classification of customer behavior: "1" is assigned to quintile with the worst values of the variable, and "5" – for quintile with the best [6].

The analysis of the RFM basics shows that it has a serious drawback: the classes' boundaries are fixed by values of

978-1-5090-4069-8/16/$31.00 ©2016 IEEE

observed variables rather than the homogeneity of the consumer behavior from the same class.

In fact it leads to the misidentification of customers and the development of erroneous sales activities that significantly reduces the company's profit. Therefore it is a necessity in modification and adaptation of the existing method for quintile boundaries determination that meet the real target groups.

To get a solution to the problem mentioned above for characteristics *F* and *M* in [5] it has been suggested the application of the so-called Wirth method of the mean scoring. According to this approach the class labeled "1" includes only "sporadic" consumers, i.e. those who committed only one purchase for the period under review, while for the remaining part of customers the average value of the indicator (*F* or *M*) is calculated. The consumers whose own value of such an index is lower than the average will include in the class labeled "2". The process is repeating for the rest of non-numbered consumers. This procedure allows evaluating the quintile boundaries with respect to certain tendencies contained in the initial data. However, it is well known that the arithmetic mean value is rather unstable measure of location, especially in the presence of outliers in the sample [7]. It will be shown below that the distribution of characteristics *F* and *M* is going to be close to exponential, and therefore, the appearance of such values in the sample has a nonzero probability. So the average value will be distorted, and the results of RFM-analysis will become incorrect.

III. RFM-ANALYSIS ENHANCEMENT

The segmentation method suggested by authors is based on customer database partitioning into quintiles by applying a modified ABC-analysis has been considered in [8] and developed in [9, 10]. It allows estimating the flexible quintile boundaries. In doing so their values are going to be optimized to provide the great likelihood for identical response for every observation within the allocated segments. It gives an opportunity to determine the most numerous groups of consumers and to develop appropriate marketing strategies for each target group individually.

In general the enhanced procedure of RFM-analysis could be expounded as follows:

1. Fulfilling the client database and freezing the current date.
2. Calculating the values of *R*, *F*, *M* characteristics. Conducting the primary statistical analysis.
3. Calculating the pivotal points of the cumulative curves for each of the characteristics.
4. Assessing the cumulative curves by the ordinary least squares method (OLS). Selecting the best approximation for each characteristic.
5. Constructing the optimal partition into quintiles based on the analytical method ABC-analysis.
6. Calculating the quintile boundaries, counting the number of customers in every segment.
7. Setting up the RFM-encoding for customers from quintiles.
8. Analyzing the selected customer segments.

The implementation of the supposed method will be demonstrated on the example of enterprise engaged in retail sales of high-tech and office tools.

IV. IMPLEMENTATION OF THE METHOD SUPPOSED

At the commencement of research the client database contained 5818 records on purchases made by company's loyalty card holders during 2015. It allows correctly implementing the RFM-analysis. The loyalty cards were issued to the clients who one-time spent more than 3000 rubles per purchase.

The primary statistical analysis results for *R*, *F*, *M*-characteristics are shown in Figure 1.

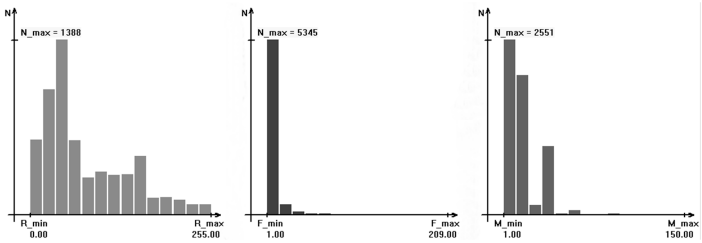

Fig. 1. Sample distributions for R-F-M values

Comparing the histograms one could conclude that the largest number of customers (1388) purchased goods recently; the vast majority of them did one or two purchases in 2015. The plot for *M*-characteristic shows that there are two sizeable segments: the first one contained 2551 clients who spent from 1 to 10 thousand rubles per purchase, and the other one – 994 clients who spent from 25 to 50 thousand rubles. The histograms represent the estimation for the density distribution of the analyzing characteristics. Therefore one can assume that for *F* and *M*-characteristics the most suitable distribution will be of exponential type. Since in average with an increasing in the values of these characteristics the corresponding number of clients is reduced [11].

The next step is the determination of pivotal points for cumulative curves for every characteristics and the evaluation of appropriate relations. The parameterization patterns of cumulative curves were chosen according to the [9, 11]. Table 1 contains the appropriate OLS estimators. The best approximating functions (corresponding to the least values of residual sum of squares ESS) are in bold.

TABLE 1.
DETERMINATION OF BEST PARAMETRIZATION FOR CUMULATIVE CURVES

PARAMETRIZATION	**ESS (R)**	**ESS (F)**	**ESS (M)**
Power function	20.183	6.288	9.023
Ballou function	1.547	4.726	**0.268**
Fourier function	1.061	4.037	3.992
Chotikapanich function	**0.149**	15.502	1.047
Pearl-Reed function	9.741	**1.919**	2.518

When analyzing the Table 1 data, it should be noted that the best parameterization for *R*-characteristic is the Chotikapanich function, for *F* – the Pearl-Reed function, and for *M* – the Ballou function. In this case the power function which is the simplest and most commonly used parameterization in similar studies became the worst one for all three characteristics, especially for *R*, that confirms its theoretical properties (the weak approximating ability for different initial data).

Further to the results obtained the analytical method of ABC-analysis was applied. It involves the cumulativity property's usage for the partitioning of the data by quintiles A-B-C-D-E [8, 12]. In accordance with the RFM-analysis principles the whole

set of customers and, accordingly, the database records should be divided into five groups. Their boundaries will be determined by the initial point (0, 0), ending point (1, 1), and four inner points, graphically located on x-axis, satisfying the relations

$$0 < x_1 < x_2 < x_3 < x_4 < 1.$$

The partitioning by these points is carried out on the maximization of the sum of the areas of figures formed under the cumulative curve.

V. RESULTS DISCUSSION

When analyzing the partitions obtained above all characteristics it has been found out that the classes structure varies seriously, which proves the incorrectness of application of uniformly precise partitioning into quintiles recommended in the standard approach of RFM-analysis. The Table 2 contains the database proportions corresponded to each of the quintiles.

TABLE 2.
QUINTILES FILLING FOR R-F-M CHARACTERISTICS

Characteristic \ Quintile	5	4	3	2	1
R	13.60%	16.30%	19.31%	23.12%	27.67%
F	4.25%	13.23%	21.92%	29.39%	31.21%
M	10.19%	16.89%	20.61%	24.10%	28.21%

Considering the Table 2 elements, one could note that in quintiles with the best values of the parameters (coded as "5" and "4") contained the fewer number of customers than in the others, which is consistent with the main idea of ABC-approach as «vital few, trivial many» [8, 9].

The next step was the RFM-coding construction and the visual presentation for the results. Figure 2 shows the projections of the customer number combined with the same encoding for M=1,...,5 in the R-F-coordinates. In a total the 54 different sets of customers from the analyzed database were singled out.

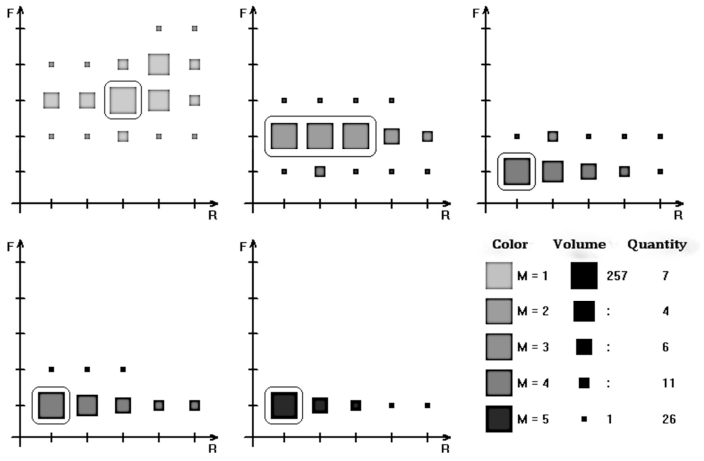

Fig.2. RFM-analysis for company's clients

Thus the study allowed singling out the seven largest groups of customers, forming the homogeneous segments of buyers. On Fig. 2 they are indicated by squares with the largest area, and circled for clarity. The minimum number of elements for each group is 257 clients. Down below one can find the results interpretation.

1) RFM = 331. The target group which did insignificant and inexpensive purchases but regularly and not so long ago. These customers are interested in a variety of accessories for computer equipment, such as routers, connectors, optical peripherals, etc. The marketing strategy for this consumers should be based on the increasing the number of inexpensive product lines and brands, the usage of competitive and psychological pricing to reduce the perception of the price among buyers. The promotion mix should include not only some different events for sales promotion, such as coupons, product samples, contests, sweepstakes, giveaways, discounts, etc., but also actions aimed to support the customer loyalty. For example, for improvement the customer loyalty program and enlargement of its opportunities could be effective to implement the bonus accumulation, progressive scale of discounts, fixed prices to increase purchase frequency, and so on.
2) RFM = 122, 222 и 322. These customers have already spent more money than the customers from the previous group. Some of them have bought recently, and others did their last purchase since 3-4 months. Most likely, these customers have bought monitors, netbooks, tablets, game consoles, and other computer equipment. The most effective marketing strategy for this group should be based on sales promotion by offering all sorts of discounts. For instance, there are discounts with colored price tags, bonus discounts (given to the regular customers when they are purchasing the certain amount of goods for the certain period of time), "happy hours" (morning hours in weekdays, when buyer is awarded with additional 5% discount on any good), and so on. For this target group the provision of lending services and the payment by installments on their purchases is also effective.
3) RFM = 113, 114 и 115. The customers from these groups have spent a lot of money, as a rule, for one-time purchase. Perhaps it has been the acquisition of graphic workstations, servers, computer classes, etc. In this case it would be suggested to apply the marketing strategy providing the image of enterprise engaged in retail sales of high-tech and office tools. It may include a wide range of pre-sales and after-sales services: packaging, delivery, installation, configuration, information support, and various kinds of repairs, including warranty, etc. Pricing strategy could be based on the market skimming: to set up the high initial prices for new prestigious products with further reducing it while the market is filling with such products. The development and implementation of pre-order services at company's website would be the other impactful approach to stimulate clients from this group. One can find in [9] the recommendations for effective service of this type of clients.

Therefore one could advise the company to develop three directions of marketing activity each of those should include the customers from determined segments.

The optimization of product portfolio structure based on the modern methods of category management and the RDE-type approaches [13] allowing to identify the SKU's best fitted to the customers from determined segments could also provide the significant contribution to the sales' improvement.

VI. CONCLUSION

Summarizing the whole study it can be noted that the RFM-analysis implementation as a segmentation tool is an effective way to amend the work with client groups and, as a consequence, to increase sales and profits. Adaptation of enhanced RFM-analysis on the example of wholesale and retail trade of high-tech products has identified three target groups of consumers. According to this it is recommended to improve the marketing and promotion policy.

In addition the method proposed in the study is not focused on specific product categories or markets which makes it versatile and allows to significantly expanding its scope.

REFERENCES

[1] Lobova O. The 12-years-old level: 2015-results on Russian PC market. http://idcrussia.com/ru/about-idc/press-center/63463-press-release. (in Russian)

[2] Goodman, J. Leveraging the Customer Database to Your Competitive Advantage. Direct Marketing, 1992. Vol. 8(55). pp. 26–27.

[3] Hughes, A.M. Boosting Response with RFM. New York: Marketing Tools, 1996.

[4] Hughes, A.M. The Complete Database Marketer: Tapping Your Customer Base to Maximize Sales and Increase Profits. New York: Probus Pub Co, 1991.

[5] Miglautsch J.R. Thoughts on RFM Scoring. The Journal of Database Marketing, 2000. Vol. 8. pp. 67-72.

[6] Hughes A.M. Strategic Database Marketing. 2nd edition. New York, McGraw-Hill, 2000.

[7] Rousseeuw P.J. Tutorial to Robust Statistics. Journal of Chemometrics, 1991. Vol 5. pp. 1-20.

[8] Models and Methods of Logistics Theory. Edited by V.S. Lukinsky. St. Peterburg, Piter-Press, 2007. (in Russian)

[9] Shchekoldin V. The Identification of the Internet-shops' Consumers on the Basis of ABC-modification of the Factor Analysis. Materials of the 6th International scientific and practical conference “Logistics – Eurasian bridge”, Krasnoyarsk, 2011. (in Russian)

[10] Shchekoldin V., Tsoi M. The Assessment of the Effectiveness of Sales Promotion Campaign on the Basis of the ABC-analysis. Pacific State University of Economics bulletin, №3, 2010. (in Russian)

[11] Shchekoldin V., Timofeev V., Faddeenkov A. Econometrics. The manual. 3rd ed. URAIT, Moscow, 2015. (in Russian)

[12] Aaberge R. On the Problem of Measuring Inequality. Discussion Paper No. 14, Research Department, Statistics Norway, 1986.

[13] Shchekoldin V., Tsoi M. Development of a New Range of Products on the Basis of RDE-method. Practical Marketing, 2015. Vol. 4 (218). pp.17-22. (in Russian)

Shchekoldin Vladislav Yurievitch, Candidate in Science (PhD) in Computer Science, Associate Professor of Marketing and Service Department Novosibirsk State Technical University. Research interests are experimental design, logistics, econometrics, statistics, and marketing research. More than 50 papers and manuals have been published.

Tsoy Marina Yevgenienva, Candidate of Science (PhD) in Economy Science, Associate Professor, Head of Marketing and Service Department Novosibirsk State Technical University. Research interests are consumer behavior, marketing research, and marketing logistics. More than 50 papers and manuals have been published.