Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 162 (2019) 281–289

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Information Technology and Quantitative Management (ITQM 2019)

A comparison between Fuzzy Linguistic RFM Model and traditional RFM model applied to Campaign Management. Case study of retail business.

Rocío González Martínez[a]; Ramón Alberto Carrasco[a,*]; Jesús García-Madariaga[a]; Carlos Porcel Gallego[b]; Enrique Herrera-Viedma[c]

[a] *Dep. of Management and Marketing, UCM 28223 (Spain)*
[b] *Dep. of Computer Science, University of Jaen, Jaen, 23071 (Spain)*
[c] *Dep. of Computer Science and IA, University of Granada, Granada, 18071. (Spain)*

Abstract

Recency Frequency Monetary Value (RFM) is a clear and descriptive way to classify a customer database based on purchasing behavior that direct marketers have used with success since almost the last twenty years. Despite the fashion that exists lately around predictive models and artificial intelligence, direct marketing's RFM, still has a place in modern database marketing. In a real business environment, RFM can still be useful when models are not practical because it is user friendly and the outcome is always interpretable. It can also be used to combine with other models. In this paper, we show a real example about how easy, accurate and explainable can be a customer segmentation based on the traditional RFM model and the 2-tuple RFM model applied to a customer database. It will help us to better understand the benefits of applying the 2-tuple model instead of the traditional one. We will be able to see how, by applying a k-means clustering on top of the 2-tuple model, segments have a great applicability from the business point of view. By using descriptive variables, we will clarify the cluster description and the model will provide us an extremely clear idea about how customers behave. The main goal for developing this example was to define the best target for a direct campaign communication. Data used for this analysis belongs to a worldwide home furniture, Scandinavian Retailer, and are related to its loyalty program which give us the members' historical purchase information.

Keywords: RFM, 2-tuples model; campaign management; k-means clustering; retail customers segmentation

1. Introduction

RFM, (Recency, Frequency, Monetary) model is a very renowned marketing tool that uses transaction history to define behavior-based customer segmentation. In this paper, we follow a previous work from the authors [1] and work around the RFM model which was proposed by Hughes in 1994 [2] and define categories to discover which customers will better react to an offer or to engage successfully it our campaigns. Define segments into a

* Corresponding author.
E-mail address: ramoncar@ucm.es

10.1016/j.procs.2019.11.286

customer database is a well-known method that marketers used to identify the best target group of customers for each campaign.

Nowadays marketers know that, instead of reaching out to 100% of their audience, it is more efficient to focus on a specific customer group that will turn out to be more profitable for their business, also RFM model becomes a great tool to trigger direct marketing actions for retention, activation, regain campaigns when we need to define strategies to create engagement and loyalty in our customer database.[3]-[4].

As almost nothing in life is perfect, the model has weaknesses, mainly related to lack of accuracy specially due to the ordinal scale used to score the RFM variables which normally is based on the set {1,...,5} for each one and helps us to identify customer contributions to revenue. Label 5 indicates the maximum contribution and label 1 would be the minimum contribution to revenue [5].

RFM method connects customer response information with R, F and M values as independent variables and then, define groups of customers into specific RFM categories. Sometimes we are not able to find an equal number of customers under each RFM group due to correlation existing between R, F and M metrics. For example, a customer who is spending more should be also likely to buy more frequency [6].

The fuzzy linguistic approach is a tool developed for modelling qualitative information in a problem. It works around the concept of linguistic variables [7] and it is being used in many situations with very good results. [8]-[9]. The 2-tuple fuzzy linguistic approach is a good solution to represent information related to processes of "computing with words" without detriment of information [10] that it is being used in many business applications [11]-[12].

In this paper, we follow the steps proposed in [1] to link customer segmentation with campaign management applying the 2–tuple model to the RFM computation and to its application based on k-means clustering algorithms. We apply their results to a real business case to better understand how this 2-tuple RFM model yields a greater result interpretability and also enable computing these values with no detriment of information. In consequence, decision makers could easily identify the best customers and launch more effective marketing campaign based on the interpretation of these linguistic results.

We have developed the full example using KNIME 3.7.1 open source software, recognized by Gartner as a Leader in Data Science and Machine Learning Platforms [13], to help other scientist to deploy the full process. This also enables us to be more relevant at the practical level and better understand the pros and cons instead of remaining confined to the theoretical environment.

The rest of the work is organized like following indicated. In section 2 we present some basic aspects about the 2-tuple model. In section 3, the comparison between 2-Tuple RFM Model and Traditional RFM Model is presented. Finally, the obtained conclusions are exposed.

2. The 2-Tuple Model

Let $L = \{l_0,\ldots,l_T\}$ be a linguistic set of terms with odd cardinality. The midterm will mean the neutral value and the rest of the terms will be symmetrical distributed around him. It can be accepted that the semantics of labels are given by means of triangular membership functions and consider the terms on a scale with a total order defined, i.e. $l_i \leqslant l_j \Leftrightarrow i < j$. In this context, if a symbolic process aggregating linguistic information gives us a value $b \in [0,T]$, and $b \notin \{0,\ldots,T\}$, then, a function to calculate is used to express the result in L.

Definition 1 [10]. *Let b be the result of a collection of a set indexes of labels evaluated in the term set L, i.e. the result of a symbolic aggregation operation,* $b \in [0,T]$. *Let* $i = round(b)$ *and* $a = b-i$ *be 2 values, such that* $i \in [0,T]$ *and* $a \in [-0.5,0.5)$, *then a is named as a Symbolic Translation.*

The 2-tuple approach [10] is based on a symbolic translation by representing the linguistic information using 2-tuple (l_i, a_i), $l_i \in L$ and $a_i \in [-0.5, 0.5)$, where l_i represents the linguistic label, and a_i is a numerical value that translate the value from the original result b to the closest index label, i, in the linguistic term set L. The value (l_i,

a_i) also can be named as $l_i \pm a_i$ (+ or - following the sign of a_i).

The model defines the group of transformation functions between numeric values and 2-tuple values.

Definition 2 [10]. *Let $L = \{l_0,\ldots,l_T\}$ be a linguistic term set and $b \in [0,T]$ a value representing the result of a symbolic aggregation operation, then the 2-tuple that represents the corresponding information to b is calculated using the following function:*

$\Delta: [0,T] \rightarrow L \times [-0.5,0.5)$

$$\Delta(b) = (l_i, a), \text{ with } \begin{cases} l_i, & i = round(b) \\ a = b - i, & a \in [-0.5,0.5). \end{cases} \quad (1)$$

where round(·) is the round *operation, l_i has the nearest index label to b and a is the the symbolic translation.*

For all Δ, there exists Δ^{-1}, defined as: $\Delta^{-1}(l_i, a) = i + a$. **(2)**

The negation operator is defined as: $neg((l_i, a)) = \Delta(T - (\Delta^{-1}(l_i, a)))$. **(3)**

Due to the information aggregation gives us a value that resume a set of values, the aggregation of a set of 2-tuples should be also a 2-tuple. Functions Δ and Δ^{-1} translate quantitative values into linguistic 2-tuples and vice versa always with no information detriment. Those functions permit that any aggregation operators could be widespread to able to deal with linguistic 2-tuples. Below we expose the aggregation operators which the authors have applied in their model [1]:

Definition 3. *Let $A = \{(l_1, a_1),\ldots, (l_n, a_n)\}$ be a set of linguistic 2-tuple, the values of their corresponding $b_1,\ldots,b_n$ (see Definition 1) and $W = \{w_1,\ldots, w_n\}$ their corrsponging weights. The 2-tuple weighted average $\bar{A}^w$ is:*

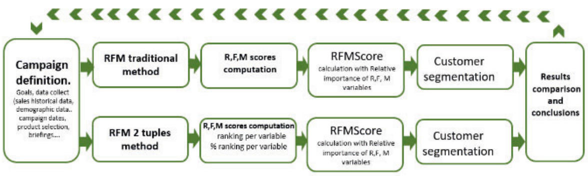
$$\bar{A}^w [(l_1, a_1),\ldots, (l_n, a_n)] = \Delta\frac{\sum_{i=1}^{n}[b_i w_i]}{\sum_{i=1}^{n} w_i}. \quad (4)$$

Definition 4. *Let $A = \{(l_1, a_1),\ldots, (l_n, a_n)\}$ be a set of linguistic 2-tuple and the values of their corresponding $b_1,\ldots,b_n$ (see Definition 1). The 2-tuple average $\bar{A}$ is:*

$$\bar{A} [(l_1, a_1),\ldots, (l_n, a_n)] = \Delta\left[\frac{\sum_{i=1}^{n} b_i}{n}\right]. \quad (5)$$

3. A comparison between 2-Tuple RFM Model and Traditional RFM Model

The basic idea presented in [1] has been to use the 2-tuple representation instead of the traditional scores (usually quintiles) used in classical RFM in order to achieve greater precision without losing linguistic interpretability of these scores. In this section we are going to compare this proposal with the classic one applied to a real retail business to really verify the improvements of the 2-tuple version. For this we will follow the scheme presented in Fig. 1. Results from this comparison will be applied to a real business case in Retail. The marketing department of the retailer is continuously launching campaigns on the customer database and they



need to be able to prioritize and define different actions for different customers' profiles. We will provide them a good description of their customer database based on these RFM analysis and subsequent segmentations.

Fig 1. Process of customer segmentation with retail campaign activities using both RFM methods

3.1. The RFM scores computation with traditional model

In the traditional model, customers are sorted separately by each of the RFM variables and then divided into five groups, typically quintiles. This generates a table: *CustomerRFM (with variables like PersonalCard, Recency, Frequency and Monetary Score)* with all score fields ∈ {1,…, 5} (an extract is shown in Fig 2). Consequently, the RFM measures will be ordinal scores were level 1 has the 20% of customers with the lowest values and level 5 will include the 20% with the highest values in the corresponding variable. In the case of the *Recency* attribute, the scale of the calculated ordinal score, *Recency Score*, should be inverted so larger scores will mean the most recent buyers.

Renamed/Retyped table - 2:44 - Column Rename

File Hilite Navigation View

Table "default" - Rows: 223781 | Spec - Columns: 10 | Properties | Flow Variables

...	S Personal_Card	I Freque...	D Monetary	L Recenc...	S RECEN...	S FREQU...	S MONETARY INTERVAL	S RECENCY...	S FREQUENCY...	S MONETARY...
..	6275980464000000108	24	112.293	224	(161,337]	(13,33]	(93.33,280.42]	3	2	2
..	6275980464000000113	319	6,424.611	54	[1,58]	(145,34,355]	(1,512.437,954,380.315]	5	5	5
..	6275980464000000114	24	322.331	301	(161,337]	(13,33]	(280.42,649.383]	3	2	3
..	6275980464000000115	73	550.711	261	(161,337]	(69,145]	(280.42,649.383]	3	4	3
..	6275980464000000116	82	287.365	89	(58,161]	(69,145]	(280.42,649.383]	4	4	3

Fig 2. Extract of Customer RFM table

We also consider useful to have a unique measure instead of three different classifications per customers, so the R, F, and M values are summed to define the *RFM Score*, which characterizes together the RFM scores. In the simplest case, the RFM scores will be calculated for each customer just using a sequential sorting based on RFM metrics, but we can use an alternative method which applies regression techniques to calculate the weights of the R, F and M metrics and this relative weight can be apply in the calculation of cumulative points for each customer [6]. These weights per each variable could be named as w_R, w_F, w_M, and they would be use to prepare the final unique *RFMScore*. The RFM score is the weighted average of its individual components and is calculated as follows:

$$RFMScore = Recency\ Score \times w_R + Frequency\ Score \times w_F + Monetary\ Score \times w_M.$$

3.2. The RFM scores computation with the 2 Tuple model

In this step the goal is to calculate and store the scores that conform the output table (CustomerRFM 2-tuples), i.e., Recency Score, Frequency Score, Monetary Score and RFM Score using the 2-tuple model [12].

As a first step we will define the symmetric and uniformly distributed domain L using five linguistic labels. The labels have a semantic definition for the variables of the RFM model that can be translated as the degree of agreement on the goodness of the variable:

Let L = {l0,…, lT}, T = 4: s0 = Very Bad = VB, s1 = Bad = B, s2 = Neutral = N, s3 = Goof = G, and s4 = Very Good = VG, with the definition showed in Fig. 3.

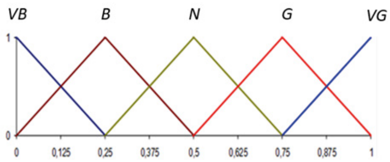


Fig 3. Definition of the set S.

As a consequence, we will have the variables to compute: *Recency Score, Frequency Score, Monetary Score, RFM Score* $\in L \times [-0.5, 0.5)$.

For each customer $i = 1,..., n$, we obtain $A_i = (A_{i1}, A_{i2}, A_{i3})$ with $A_{i1} = RecencyScore_i$, $A_{i2} = FrequencyScore_i$ and $A_{i3} = MonetaryScore_i$. Firstly, customers are organised in ascending order followign each of the individual RFM components $B_i = (B_{i1}, B_{i2}, B_{i3})$, with $B_{i1} = Recency_i$, $B_{i2} = Frequency_i$ and $B_{i3} = Monetary_i$, stored in the customer transactions database. After this, we define $rank_{ij} \in \{1,\ldots, n\}$ as the ranking of each client respect to each of these variables:

$$percent\_rank_{ij} = (rank_{ij}-1) / (n-1)$$

with $percent\_rank_{ij} \in [0, 1]$, $i = 1,..., n$, $j = 1,..., 3$ and $n > 1$. The final 2-tuple score A_{ij} is obtained as following:

$$A_{ij} = \begin{cases} \Delta((percent\_rank_{ij}), \text{ if } j \neq 1 \\ neg(\Delta(percent\_rank_{ij})), \text{ if } j = 1 \end{cases} \qquad \textbf{(6)}$$

where $\Delta(\cdot)$ and $neg(\cdot)$ were defined in Eq. 2 and 4 respectively in the paragraph related with 2 Tuple model. We use the negation function on *Recency* in case of the higher scores means the most recent buyers.

We compute the 2-tuple $RFMScore_i$, which characterizes together the RFM scores for each *i*-customer using the Eq. (4) as follows:

$$RFMScore_i = \bar{A}^w [A_{ij}] \qquad \textbf{(7)}$$

where, $j = 1,...,3$ with the user-defined weights $W = \{w_R, w_F, w_M\}$ same that previously defined for the traditional RFM model.

Fig 4 show us the results of this process.

Input with appended column - 2:127:124 - Column Combiner

File Hilite Navigation View

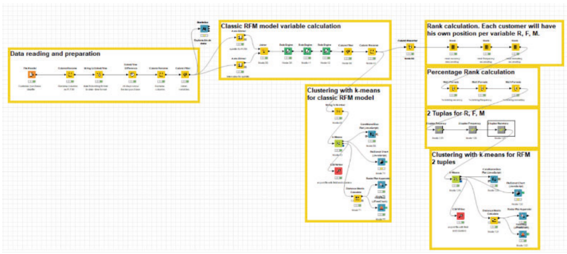
Table "default" - Rows: 223781 | Spec - Columns: 28 | Properties | Flow Variables

...	S Personal_Card	L Recency...	I Freque...	D Monetary	I RANK R...	I RANK F...	I RANK M...	D percentag...	D percent...	D percent...	S 2 TUPLAS RECENCY	S 2 TUPLAS FREQUENCY	S 2 TUPLAS MONETARY
..	6275980464000000108	224	24	112.293	113438	70275	51002	0.507	0.314	0.228	"N","0.007"	"B","0.064"	"B","-0.022"
..	6275980464000000113	54	319	6,424.611	181603	211387	219563	0.812	0.945	0.981	"G","0.062"	"VG","-0.055"	"VG","-0.019"
..	6275980464000000114	301	24	322.331	96491	70275	96429	0.431	0.314	0.431	"N","-0.069"	"B","0.064"	"N","-0.069"
..	6275980464000000115	261	73	550.711	104435	136439	125150	0.467	0.61	0.559	"N","-0.033"	"N","0.11"	"N","0.059"
..	6275980464000000116	89	82	287.365	162458	143816	90740	0.726	0.643	0.405	"G","-0.024"	"G","-0.107"	"N","-0.095"

Fig 4. Extract of Customer RFM 2 tuples table with full group of fields.

3.3. Model development and comparison

In a previous paper [1] the authors have proposed both, the traditional model RFM and the representation data type 2-tuple for campaign management using the k-means clustering technique. The basic idea of these authors is to apply the k-means algorithm on 2-tuple values (result of the 2-tuples RFM) calculating the mean of each centroid using the mean $\bar{A}$ (see Definition 4). The practical implementation is done by the authors using IBM SPSS Modeler. In this case, as a second step, we develop the model using KNIME 3.7.1 and we use a real



customer data to define the best customers groups to target on the different campaigns and engagements activities. Thus, Fig. 5 show the scheme of the dataflow to solve the example of the previous sections.

Fig 5. Data flow for model development.

A good tool to analyse a clustering result is the Conditional Box Plot node that KNIME offers to compare the clusters centroids. We can see how detailed is the information that the Box plot gives us for the 2-tuples model (Fig 6 right graph) in comparison with the Traditional model (Fig 6 left graph) because of the nominal values for each variable which decrease the possibility to differentiate customer behaviours.

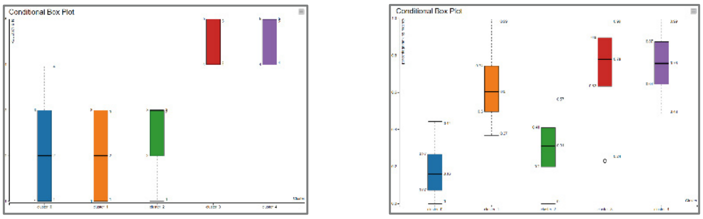
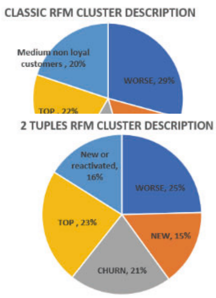


Fig 6. Box plot for RecencyScore, in K-means for Traditional and 2-Tuples RFM model

As a summary, we can describe the five clusters for both models like follows:



		TRADITIONAL RFM CLUSTER DESCRIPTION			Centroids		
cluster	name	description	size	%	Recency	Frequency	Monetary
CLUSTER 0	WORSE	Customers with very low values in Recency, Frequency and Monetary.	64,767	29%	2	1	1
CLUSTER 3	NEW	Very good value for Recency. Still medium for Frequency and Monetary.	36,013	16%	4	3	3
CLUSTER 2	CHURN	Customers with bad Recency but very good Frequency and Monetary. They used to be quite good but are becoming sleeper customers	28,473	13%	2	4	4
CLUSTER 4	TOP	Good customers in the 3 variables	48,122	22%	5	5	5
CLUSTER 1	Medium non loyal customers	Customers with meidum values in R, F and M. They can be customers that are loosing interest but they never were realy engaged with the brand.	44,039	20%	2	3	3

		2 TUPLES RFM CLUSTER DESCRIPTION			Centroids		
cluster	name	description	size	%	2-tuples R	2-tuples F	2-tuples M
CLUSTER 0	WORSE	Customers with very low values in Recency, Frequency and Monetary.	55,028	25%	"B","-0.077"	"B","-0.064"	"B","-0.055"
CLUSTER 1	NEW	Customers with very low values in Frequency and Monetary but good in Recency. Means we have no historical info about them but they are actives now.	34,333	15%	"G","-0.122"	"B","-0.015"	"B","-0.01"
CLUSTER 2	CHURN	Customers with bad Recency but Neutral Frequency and Monetary. They used to be quite good but are becoming sleepers customers.	46,333	21%	"B","0.052"	"N","0.089"	"N","0.119"
CLUSTER 3	TOP	Good customers in the 3 variables Eventhough they can still grow.	51,868	23%	"G","0.006"	"G","0.116"	"G","0.107"
CLUSTER 4	New or reactivated	Better customers that cluster 1. Recency Good and better F and M so they can be old customers reactivated or good new customers	36,013	16%	"G","0.009"	"N","0.069"	"N","0.046"

Fig 7. Cluster description for Traditional RFM model.

Fig 8. Cluster description for 2- tuples RFM model.

In order to really understand how good are these groups defined and also to be able to clarify the doubts about New members distribution that showed up in the Heat map we can introduce new variables to the study, not to take them into account as active variables in the k-means clustering but just to describe the segments and enrich the conclusions.

This Retail Company considers as active customer all those customers with almost one purchase during the las 12 months, so, following this criteria we can have three different types of customers:

- New members: active customers but his first purchase has been done during the last 12 months.

- Active customers: almost one purchase during the last 12 months but not including New members.
- Inactive customers: no purchases in the last 12 months but they use to be actives before.

As we have purchase data since 2014, the company can classify his customers using historical purchase information into four categories, let's call them as A, B, C, D where A are the best customers, having a big number of products and specially some of the items that are more relevant for the retailer and D are the worse customers in term of number and quality of product purchases. Thant means that is a very weak group for the retailer and is the most difficult group to engage.

After these definitions what we have done, Fig 9 show the clusters description with external variables for Traditional RFM model.

TRADITIONAL RFM CLUSTERING		
Trad Cluster RFM	ACTIVITY	% CUSTOMERS
WORSE	ACTIVE	12%
	INACTIVE	71%
	NEW	17%
MEDIUM NON LOYAL CUSTOMERS	ACTIVE	37%
	INACTIVE	57%
	NEW	6%
CHURN	ACTIVE	60%
	INACTIVE	39%
	NEW	2%
NEW	ACTIVE	65%
	NEW	35%
TOP	ACTIVE	96%
	NEW	4%

RFM CLASSIC CLUSTERING		
Trad Cluster RFM	TYPOLOGY	% CUSTOMERS
WORSE	a	0%
	b	12%
	c	30%
	d	58%
MEDIUM NON LOYAL CUSTOMERS	a	3%
	b	49%
	c	40%
	d	7%
CHURN	a	50%
	b	49%
	c	1%
NEW	a	3%
	b	41%
	c	38%
	d	19%
TOP	a	61%
	b	38%
	c	1%

RFM 2-TUPLES CLUSTERING		
Cluster 2 tuples	ACTIVITY	% CUSTOMERS
WORSE	ACTIVE	4%
	INACTIVE	94%
	NEW	2%
NEW	ACTIVE	44%
	NEW	56%
CHURN	ACTIVE	33%
	INACTIVE	64%
	NEW	3%
TOP	ACTIVE	96%
	INACTIVE	1%
	NEW	3%
NEW OR REACTIVATED	ACTIVE	83%
	NEW	17%

RFM 2-TUPLES CLUSTERING		
Cluster 2 tuples	TYPOLOGY	% CUSTOMERS
WORSE	a	0%
	b	14%
	c	31%
	d	55%
NEW	a	0%
	b	19%
	c	35%
	d	46%
CHURN	a	18%
	b	57%
	c	23%
	d	2%
TOP	a	68%
	b	31%
	c	0%
NEW OR REACTIVATED	a	7%
	b	57%
	c	32%
	d	4%

Fig 9. Cluster description with external variables

- The 2 - tuples model detect in the "WORSE" group 94% of inactive customers and 86% of low product penetration while Traditional RFM model gives us just a 71% of inactive customers and include here 17% of NEW customers which can be right in term of activity but not in term of business.
- 2-tuples find 56% of real New members for the group "NEW" while Traditional model give us just 35%.
- "CHURN" group is also very well defined for 2-tuples model with 64 % of inactive customers with high product penetration so they really used to be good customers. Traditional model finds just 39% inactive.
- The 2-tuples model find a very interesting group that we have named as "NEW OR REACTIVATED" that is really well-defined wit 83% of active customers and 17% of New customers. that are purchasing now and have already a quite good penetration in terms of products.
- For the case of Traditional model, we have the "MEDIUM NON-LOYAL CUSTOMERS" which is a group similar to the "WORSE" group because of cluster distance but they are better customers in terms of product penetration. 57% out of them are inactive customers and 37% active customers but all of them with similar RFM. Medium - low product penetration.
- When we check the size of "CHURN" group we can see how in the 2-tuples RFM model this group represents the 21% out of the total customer database, while, Traditional model, gives us just a group with the 13% of customers. Knowing how important this group can be for retention actions, we consider it something to remark.
- Traditional model is not able to isolate NEW customers properly which is a big issue because for this Retailer in terms of business because the first year as a customer normally is the most profitable one for the company.

4. Concluding remark and future work

We have presented here a real business case of how to apply the RFM model to a retailer company. We have developed the RFM following the Traditional model based on the RFM variable scores expressed by an ordinal scale and also the RFM model based on the 2-tuple fuzzy linguistic approach. Using a graphical representation of the clusters, allows us to identify and define the differences between groups and also, once we already had the cluster definition, we have checked the quality of the segmentation (from a business point of view) by introducing descriptive variables that offered us a very clear idea about how well were clusters identified.

We have applied the implemented models on a real dataset currently in use to try to solve a business problem related to gaining knowledge about the activity of the home-furnishing retailer customer database and find a tool to help the company to focus their marketing campaigns with different purposes as could be, activate the inactive customers, create engaging on the current good and active customers, define actions to retain the Top customers, define triggers to increase loyalty to the brand, etc.

We have compared both models, and the conclusion is clear. Both models are useful but, 2-tuples model has several advantages against the traditional model. After applying the segmentation k-means algorithm on both models, the 2-tuples is able to define more clear clusters. We mention the possibility to apply weights per each variable R, F, M, to modify the effect of each one on the model, but in this case the RFM values are assigned for each customer by sequential sorting based on RFM metrics per each model.

Future works will be focused on the definition of this weight vector by using a test campaign to analyze answers per RFM groups as Kumar mentioned [6].

5. Acknowledgments

This paper has been developed thanks to the financial support of FEDER funds in the Spanish National research project TIN2016-75850-R.

6. References

[1] R. A. Carrasco, M. F. Blasco, J. García-Madariaga, E. Herrera-Viedma. A Fuzzy Linguistic RFM Model Applied to Campaign Management, International Journal of Interactive Multimedia and Artificial Intelligence (2018), http://dx.doi.org/10.9781/ijimai.2018.03.003

[2] A. M. Hughes, “Strategic database marketing,” Chicago: Probus Publishing Company, 1994.

[3] K. Tsiptsis, A. Chorianopoulos, “Data mining techniques in CRM: inside customer segmentation,”, John Wiley & Sons, 2011.

[4] W. Buckinx, D. Van den Poel, “Customer base analysis: partial defection of behaviorally loyal clients in a non-contractual FMCG retail setting,” European Journal of Operational Research, vol. 164, no. 1, pp. 252–268, 2005.

[5] C. H. Cheng, Y. S. Chen, “Classifying the segmentation of customer value via RFM model and RS theory,” Expert systems with applications, vol. 36, no. 3, pp. 4176–4184, 2009.

[6] V. Kumar, W Reinartz – Customer relationship management, pp 102 – 134, 2012 – Springer

[7] L. A. Zadeh, “The concept of a linguistic variable and its applications to approximate reasoning,” Inf Sci, vol. 8, pp. 199–249, 1975.

[8] F. J. Cabrerizo, R. Al-Hmouz, A. Morfeq, A. S. Balamash, M. A. Martínez, E. Herrera-Viedma, “Soft consensus measures in group decision making using unbalanced fuzzy linguistic information,” Soft Computing, vol. 21, no. 11, pp. 3037-3050, 2017.

[9] S. Massanet, J. V. Riera, J. Torrens, E. Herrera-Viedma, “A new linguistic computational model based on discrete fuzzy numbers for computing with words,” Information Sciences, vol. 258, pp. 277-290, 2014.

[10] F. Herrera, L. Martínez, “A 2-tuple fuzzy linguistic representation model for computing with words,” IEEE Transactions on Fuzzy Systems, vol. 8, no. 6, pp. 746–752, 2000.

[11] A. Cid-López, M. J. Hornos, R. A. Carrasco, E. Herrera-Viedma, “Applying a linguistic multi-criteria decision-making model to the analysis of ICT suppliers’ offers,” Expert Systems with Applications, vol. 57, pp. 127-138, 2016.

[12] R. A. Carrasco, M. F. Blasco, E. Herrera-Viedma, "A 2-tuple fuzzy linguistic RFM model and its implementation," Procedia Computer Science, vol. 55, pp. 1340-1347, 2015.
[13] https://www.gartner.com/reviews/market/data-science-machine-learning-platforms/vendor/knime/product/knime-analytics-platform