Expert Systems with Applications 39 (2012) 9359–9366

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Customer data mining for lifestyle segmentation

V.L. Miguéis *, A.S. Camanho, João Falcão e Cunha

Faculdade de Engenharia da Universidade do Porto, Portugal

ARTICLE INFO

Keywords:
Retailing
Clustering
Segmentation
Lifestyle

ABSTRACT

A good relationship between companies and customers is a crucial factor of competitiveness. Market segmentation is a key issue for companies to develop and maintain loyal relationships with customers as well as to promote the increase of company sales. This paper proposes a method for market segmentation in retailing based on customers' lifestyle, supported by information extracted from a large transactional database. A set of typical shopping baskets are mined from the database, using a variable clustering algorithm, and these are used to infer customers lifestyle. Customers are assigned to a lifestyle segment based on their purchases history. This study is done in collaboration with an European retailing company.

1. Introduction

The recent economic and social changes that occurred in Europe transformed the retailing sector. In particular, the relationship between companies and customers changed significantly. In the past, companies focused on selling products and services without searching detailed knowledge concerning the customers who bought the products and services. With the proliferation of competitors, it became more difficult to attract new customers, such that companies had to intensify efforts to keep current consumers. The evolution of social and economic conditions also changed lifestyles, and as a result customers are less inclined to absorb all the information they receive from the companies. This context led companies to evolve from product/service-centered strategies to customer-centered strategies. The establishment of loyalty relationships with customers also became a main strategic goal. Indeed, companies wishing to be at the leading edge have to continually improve the service levels in order to ensure a good business relationship with customers.

Some companies invested in building databases that are able to collect a big amount of customer-related data. For each customer, millions of data objects are collected, allowing the analysis of the complete purchasing history. However, the information obtained is seldom integrated in the design of business functions such as marketing campaigns. In fact, in most companies the information available is not integrated in procedures to aid decision making. The overwhelming amounts of data have often resulted in the problem of information overload but knowledge starvation. Analysts are not being able to keep pace to study the data and turn it into useful knowledge for application purposes.

Data mining (DM) techniques are rising as tools to analyze data resulting from customers' activity, stored in large databases. They can be applied in order to detect significant patterns and rules underlying consumer behavior. However, the use of DM in marketing is still incipient and most companies still use mass strategies to instigate customers loyalty. The marketing segmentation of customers or the identification of customer groups with similar behavior patterns is often done in an ad-hoc way, which constitutes the basis for the definition of customized promotions.

This paper proposes a method for customers segmentation, informed by the nature of the products purchased by customers. This method is based on clustering techniques, which enable segmenting customers according to their lifestyles.

The structure of the remainder of the paper is as follows. Section 2 includes a review of segmentation approaches. Section 3 introduces the company used as case study. Section 4 includes a presentation of the methodology, and Section 5 presents the data and discusses the results. Section 6 suggests marketing actions based on the lifestyle segmentation. The paper finishes with the conclusion.

2. The evolution of segmentation approaches

Segmentation approaches were initially based on geographic criteria, such that companies would cluster customers according to their area of residence or work. This was followed by segmentation based on socioeconomic indicators, such that customers would be grouped according to age, gender, income or occupation. Marketing segmentation research gained momentum in the 1960s. Twedt (1964) suggested the use of segmentation models based on volume of sales, meaning that marketing efforts should focus on customers engaged in a considerable number of transactions. This approached, called "heavy half theory", highlighted that one half of

* Corresponding author. Tel.: +351 225082132.
E-mail address: vera.migueis@fe.up.pt (V.L. Miguéis).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.02.133

the customers can account for up to 80% of total sales. Frank, Massy, and Boyd (1967) criticized this segmentation arguing that this assumes that the heavy purchasers have some socioeconomic characteristics that differentiate them from other purchasers, what was rejected by the regression analysis carried out. Subsequently Haley (1968) introduced a segmentation model based on the perceived value that consumers receive from a good or service over alternatives. Thus, the market would be partitioned in terms of the quality, performance, image, service, special features, or other benefits prospective consumers seek. These models triggered further research that allowed to obtain sophisticated lifestyle-oriented approaches to segment customers. The lifestyle concept, introduced in the marketing field by Lazer (1964), is based upon the fact that individuals have characteristic patterns of living, which may influence their motivation to purchase products and brands. During the 1970s, the validity of the multivariate approaches used to identify the variables that affect deal proneness was criticized (see Green & Wind, 1973), which motivated the development of enhanced theoretical models of consumer behavior (e.g., Blattberg, Buesing, Peacock, & Sen, 1978). One decade later, Mitchell (1983) developed a generalizable psychographic segmentation model that divided the market into groups based on social class, lifestyle and personality characteristics. However, practical implementation difficulties of this complex segmentation model was widely noted during the 1990s in, for example, Piercy and Morgan (1993); Dibb and Simkin (1997).

More recently, the marketing literature raised the concern that customers are abandoning predictable patterns of consumption. The diversity of customer needs and buying behavior, influenced by lifestyle, income levels or age, makes past segmentation approaches less effective. Therefore, current models for marketing segmentation are often based on customer behaviour inferred from transaction records or surveys. The resulting data is then explored with data mining techniques, such as cluster analysis. Examples of applications of data mining for segmentation purposes using survey results include Kiang, Hu, and Fisher (2006). In the context of long-distance communication services, the clients were segmented using psychographic variables, based on data of a survey composed by 68 attitude questions. Min and Han (2005) clustered customers with similar interests in movies based on data containing explicit rating information for several movies provided by each customer. The rating information allowed to infer the perceived value of each movie for each customer. Helsen and Green (1991) also identified market segments for a new computer, system based on the use of cluster analysis with data from a customer survey. The segmentation was supported by the rate of importance given to the product attributes.

3. Description of the company used as case study

This paper describes an application driven methodology for effectively segmenting customers of an European retailing company. The retailing company used as case study has a chain of food-based stores, i.e. hypermarkets, large supermarkets and small supermarkets. These formats differ essentially by the sales area of the store and by the range and price of products offered. The establishment of loyal relationships with customers is a main strategic goal for this company. The development of the company information system and the implementation of a loyalty program, supported by a loyalty card, have enabled collecting data on each customer profile (e.g. customer name, address, date of birth, gender, number of people in the household, the telephone number and the number of one identification document) and transactions (date, time, store, products and prices). Currently, approximately 80% of the total number of transactions is done by customers using the loyalty card.

The company classifies each product according to the business unit, the category, the subcategory, the product description, the brand and the position of the brand based on its value. For example, a product can be classified in 12 business units (e.g. Drinks, Grocery, Fishery), in 116 categories (e.g. Beers, Desserts, Frozen), in 803 subcategories (e.g. Beers with alcohol, Fruit syrup, Frozen shellfish) and in 5 positions of the brand (i.e. Premium, Sales leader, Secondary, Own brand and Economic). The total number of different products currently commercialized by the company is about 1,363,409.

At present, the company customers are segmented in two ways. One of them consists on grouping customers based on their shopping habits. This segmentation model is a simplified version of the RFM model proposed by Bult and Wansbeek (1995), and is called by the company "frequency and monetary value" (FM) model. According to the values of these two variables, the company specifies 8 groups of customers. Each client integrates one of these groups, according to the average number of purchases done in a 8 week period and the average amount of money spent per purchase. The changes in the percentage of customers belonging to each group are used to guide the marketing actions required to meet the company's objectives. For example, if the number of customers in the clusters with more visits to the store decreases, the company is alerted to launch marketing campaigns in order to motivate customers to go to the stores more often (see Miguéis, Camanho, & Cunha, 2011). The other method of segmentation is based on customer necessities and preferences. In this case, customers are grouped into 7 segments according to the mix of categories of products they purchase. Each segment of clients is defined by using a clustering algorithm, based on the similarity between the products purchased by the client and the categories of products included in pre-defined baskets, evaluated in percentual terms. The insights provided by this segmentation method are currently not fully explored by the company. However, it is expected that in the near future, the information provided by this segmentation method can be used to guide decisions concerning the variety of products available in each store, as well as their prices.

4. Methodology

The methodology followed in this paper aims to segment customers from the retailing company according to their lifestyle. To achieve this purpose we first identify typical shopping baskets, by considering the products more frequently purchased together. In the context of this analysis, a shopping basket is defined as the set of distinct products bought by a customer over the period considered. Customers' lifestyle is inferred by analyzing the products included in the typical shopping baskets. Customers are then assigned to the lifestyle segments by considering the history of their purchases.

Clustering analysis is a widely used data mining technique that maps data items into unknown groups of items with high similarity (i.e., clusters). There is a large variety of clustering algorithms available (see Jain, Murty, & Flynn, 1999 for an overview). Most clustering algorithms can be classified in partitional or hierarchical. A partitional clustering is a division of the data items into non-overlapping groups, such that each item belongs to exactly one cluster. Partitional techniques require the prior specification of the number of clusters. Despite this limitation, partitional techniques have the advantage of allowing the optimization of a criterion related to similarity of objects within clusters or dissimilarity between clusters. Hierarchical algorithms can be classified as agglomerative or divisive. An agglomerative hierarchical clustering starts with clusters containing single items and then merges them until all items are in the same cluster. In each iteration the two most similar clusters are merged. Divisive hierarchical clustering starts with one cluster and iteratively divides it into smaller clusters. Both agglomerative and divisive hierarchical algorithms produce a nested sequence of clusters, with a single all-inclusive cluster at the top, and single-item clusters at the bottom. The resulting structure of the nested clusters enables an easy visualization of the appropriate number of clusters by means of a graphical representation of the variance explained at each clustering level, which is called dendrogram. However, hierarchical techniques do not allow the relocation of objects that may have been "incorrectly" grouped or separated at an early stage.

Whereas the clustering algorithms previously described group items into a set of clusters, there are clustering algorithms whose aim is to group variables into sets of factors, i.e. variable clustering algorithms. Variable clustering can be seen as a clustering of the items where the dataset is transposed. These methods use a measure of correlation between the variables instead of using a distance measure to compute the similarities between the variables. Several procedures for clustering variables have been proposed in the literature, such as VARCLUS (SAS Institute Inc, 2008), Clustering around latent variables (CLV) (Vigneau & Qannari, 2003), and Likelihood linkage analysis (Lerman, 1991). Since in this paper we were mainly interested in discovering lifestyle profiles, which we believe to be more stable than customers' groups, the procedure used integrates a variable clustering algorithm instead of a clustering algorithm for customers. We used the VARCLUS algorithm, integrated in SAS software, to cluster the products. It is based on a divisive algorithm and consequently in the first iteration the set of items to cluster is split into two groups. In each successive iteration, all items in each group are examined; a group will be split as long as there is more than a specified percentage of variation to be explained by splitting. Each item is assigned to a unique cluster, although the item may be reassigned to a different cluster in subsequent iterations, unlike what happens in standard divisive algorithms.

Having obtained the clusters of products using the VARCLUS algorithm, we analyzed the features of the products belonging to each cluster, in order to infer the lifestyle of customers who usually buy those products. In order to ensure the generality of the results, we only focused our analysis on the business units, the categories, and the position of the products' brand concerning the value After having identified the lifestyle segments, we assigned each customer to the lifestyle segment whose typical shopping basket presents more products in common with the customer shopping basket corresponding to past transactions.

5. Lifestyle segments

The analysis reported in this paper is based on transactional data from customers with a loyalty card. The database includes the records from October and November 2009. The preparation of the database for the exploratory analysis involved the integration of data from different sources, and the elimination of outliers. From a total of 1,904,637 customers, we selected only those who had bought at least 10 different products in the two months analysed. From the resulting sample, we randomly selected 100,000 customers to be used for the variable clustering procedure. Since we were interested in analysing only the representative products in terms of sales, we selected from the 105,160 different products transacted only those bought by at least 10,000 customers, i.e. 1831 products. From the transaction records, we then created a binary matrix, which included as items the shopping baskets of the 100,000 customers and as variables the 1831 products. The binary structure of the matrix indicates if a customer did or did not buy each of the products. Since our aim was to investigate the relationship between products, regardless of the quantities bought, we did not prepare the matrix with information on the quantities bought of each product by each customer.

In order to evaluate the number of clusters that would be appropriate to use, we first run the VARCLUS algorithm to obtain the dendrogram, depicted in Fig. 1. From the dendrogram it was possible to define the appropriate number of clusters. This number should ensure that the segments obtained could be substantial

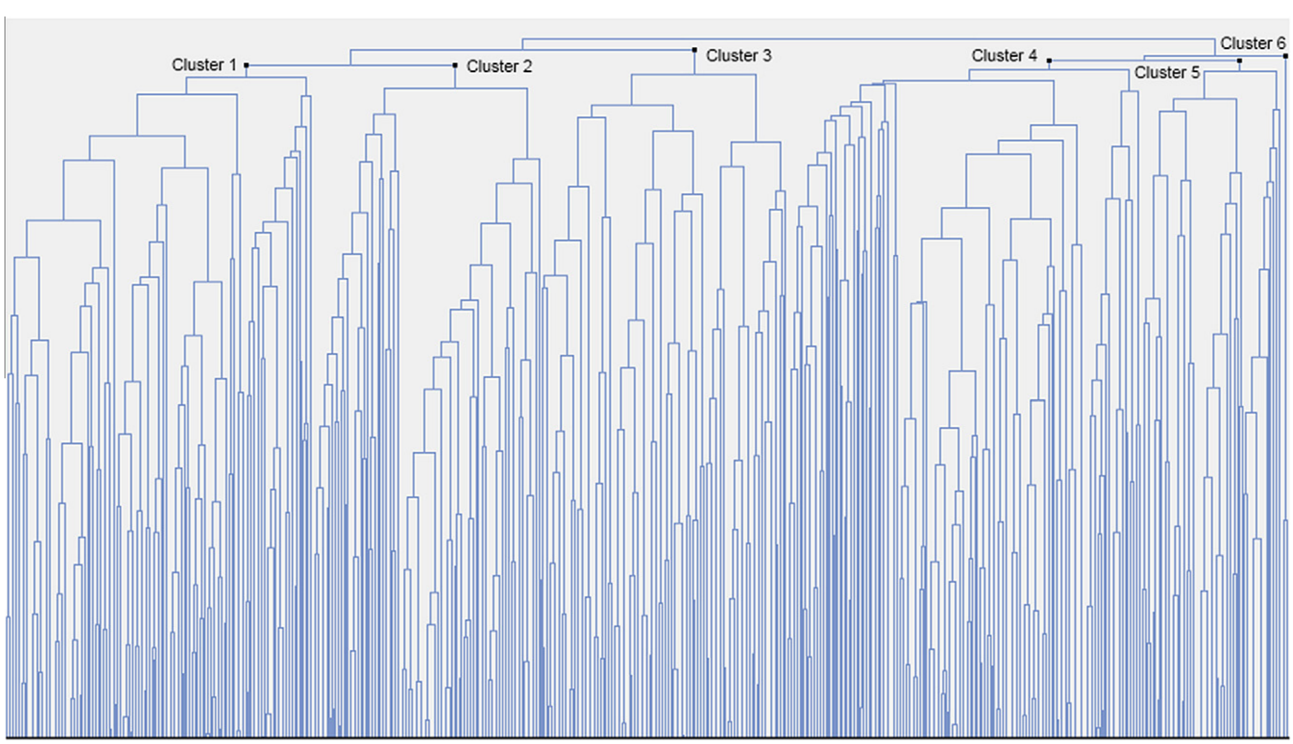

Fig. 1. Dendrogram for products resulting from the VARCLUS algorithm.

Table 1
Number of products in each cluster.

Cluster	# Products	% Products
Cluster 1	367	20.0
Cluster 2	224	12.2
Cluster 3	93	5.1
Cluster 4	226	12.3
Cluster 5	501	27.4
Cluster 6	420	23.0
Total	1831	100

(in the sense they should be large enough), and differentiable (i.e., truly distinct from other segments). These criteria are recognizable requirements for segments to be considered effective (Kotler, Brown, Stewart, & Armstrong, 2003). By analyzing the dendrogram we decided to group products in six clusters. The distribution of the 1831 products considered in the analysis by the clusters is presented in Table 1. Note that each product is assigned to one of the six clusters specified.

In order to infer the purchase patterns that may underlie each cluster of products, we analyzed the products' business units more relevant in each cluster. For that purpose, we computed the ratio between the proportion of products within a cluster belonging to each business unit and the average proportion of products from that business unit in the sample analysed. The results obtained are shown in Fig. 2.

We also analyzed the product categories more relevant in each cluster. For this purpose we identified, within each cluster, the categories whose proportion of products within the cluster was at least two times higher than the average proportion in the sample. Table 2 presents the categories found to be relevant for at least one cluster, indicating the cluster(s) for which the number of products from that category is higher than the average in the sample. Fig. 3 illustrated the results obtained for each cluster.

Finally, for each cluster we explored the most prominent positions of the products' brand concerning the value (see Fig. 4), in order to deduce the economic power of the potential buyers.

The conjunction of these analysis allowed to infer the lifestyle of the buyers of each group of products.

Cluster 1 is characterized by a relative high proportion of products from delicatessen, bakery and takeaway. Concerning the categories, baby food, beers, special fruits, bread, cheese on shelf,

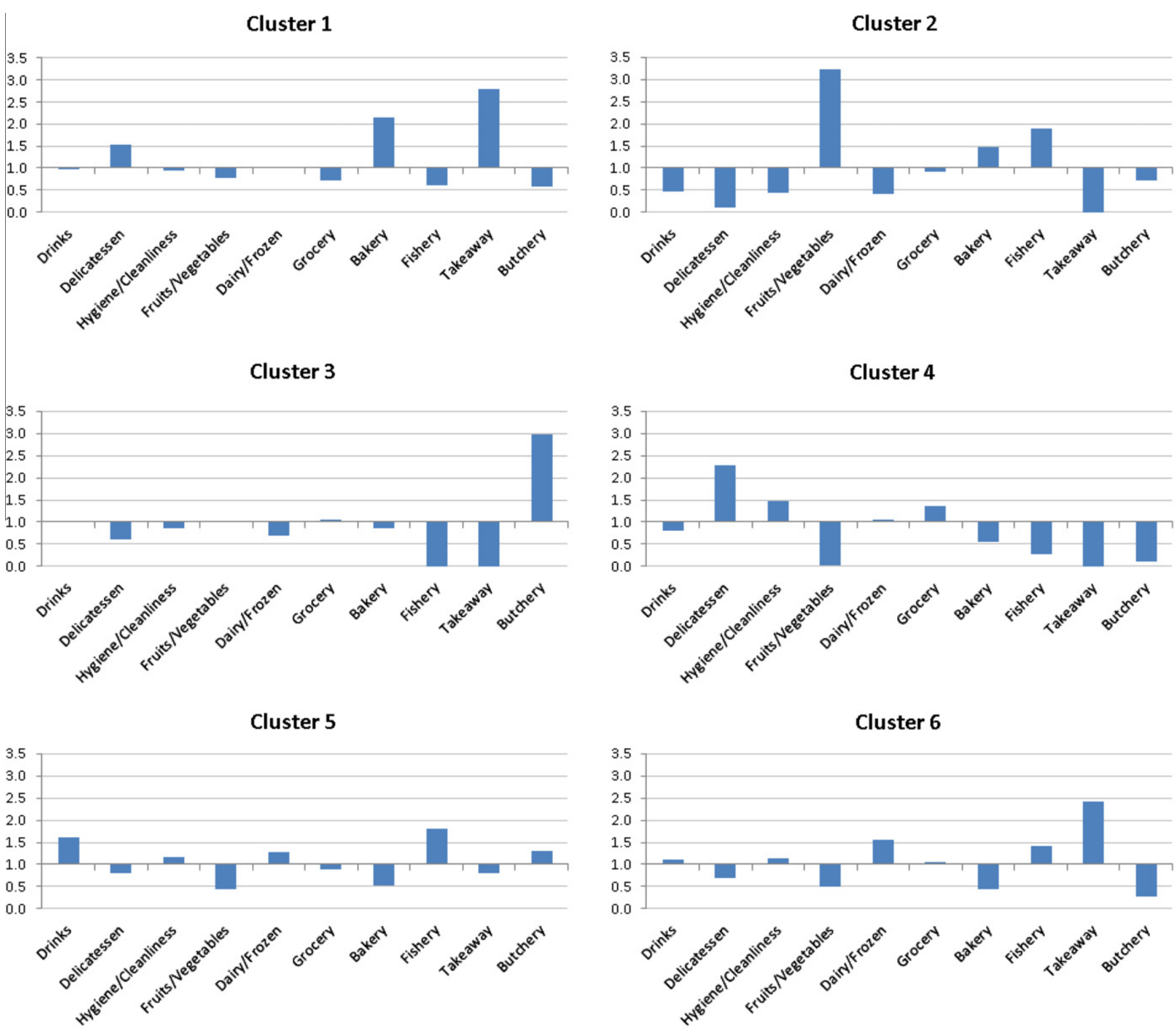

Fig. 2. Proportion of products in each business unit.

Table 2
Number of products in each cluster.

Category	Business unit	Relevant clusters
Bread	Bakery	1
Fowl meat	Butchery	3
Frozen	Butchery	5
Pork meat	Butchery	5
Veal meat	Butchery	3
Frozen desserts	Dairy/frozen	5
Frozen meals	Dairy/frozen	6
Frozen vegetables/fruits	Dairy/frozen	3
Cheese on conter	Delicatessen	3
Cheese on shelf	Delicatessen	1
Meat on counter	Delicatessen	4
Meat on shelf	Delicatessen	4
Beers	Drinks	1
Current wines	Drinks	1,5
Fortified wines/champagne	Drinks	2
Spirit drinks	Drinks	3
Codfish	Fishery	6
Fresh fish	Fishery	2
Fruits	Fruits/vegetables	2
Special fruits	Fruits/vegetables	1
Vegetables	Fruits/vegetables	2
Appetizers	Grocery	3
Baby food	Grocery	1,6
Basic ingredients	Grocery	4
Canned food	Grocery	4
Cereals	Grocery	6
Cookies	Grocery	2
Eggs	Grocery	4
Fats	Grocery	6
Honey/jams	Grocery	4
Liquid fats	Grocery	3
Pet care	Grocery	4
Powdered drinks/mixes	Grocery	6
Soups	Grocery	5,4
Spices	Grocery	5
Baby hygiene/protection	Hygiene/cleanliness	5
Body hygiene	Hygiene/cleanliness	4
Consumables	Hygiene/cleanliness	5
Dishwashing products	Hygiene/cleanliness	3
Health care	Hygiene/cleanliness	2,5
Laundry	Hygiene/cleanliness	4
Men products	Hygiene/cleanliness	5
Oral hygiene	Hygiene/cleanliness	5
Perfumes/Cosmetics	Hygiene/cleanliness	3,5
Barbecue chicken	Takeaway	5
Pre-cooked meals	Takeaway	1,6

pre-cooked meals and current wines are the most representative. Regarding the brands, most products included in this cluster are from a secondary brand or own-brand. Summing up, Cluster 1 may respect to customers with medium purchasing power that are mainly focused on practical meal solutions, preferring to buy takeaway food, bread and delicatessen products (eventually to prepare sandwiches). The potential buyers of these products have babies and seem to appreciate wine.

Cluster 2 is characterized by a relative high proportion of fruits/vegetables, bakery products and fishery products. Concerning the categories, health care, cookies, fruits, vegetables, fresh fish and fortified wines/champagne are the most representative in this cluster. Similarity to Cluster 1, this cluster includes mainly products from a secondary brand or own-brand. To conclude, the potential buyers of this group of products may have a medium purchasing power and seem to follow a balanced diet, evidenced for example by the purchase of vegetables, fruits and fish. The potential buyers of these products seem to enjoy socializing, given the diversity of fortified wines or champagnes purchased.

Cluster 3 presents a relative high proportion of products included in drinks, fruits/vegetables, grocery and butchery business units. The categories most relevant in this cluster are: appetizers, fowl meat, spirit drinks, veal meat, liquid fats, perfumes/cosmetics, dishwashing products, cheese on counter and frozen vegetables/fruits. Most products included in this cluster are from a secondary brand or own-brand. Concluding, the potential buyers of such products are people with medium purchasing power who appreciate meat. They seem to care about the personal appearance, reflected by the purchase of perfumes/cosmetics. These customers also seem to enjoy socializing, given the diversity of spirit drinks and appetizers bought.

Cluster 4 is characterized by a relative high proportion of delicatessen, hygiene/cleanliness, grocery and butchery products. The categories most relevant in this cluster are: meat on counter, meat on shelf, canned food, body hygiene, basic ingredients, honey/jams, eggs, pet care, laundry and soups. Most products included in this cluster are from a premium and economic brand. Summing up, the potential customers who buy the products included in Cluster 4 have low purchasing power, although for certain products they appreciate premium brands. These customers are likely to have a pet and often prepare dishes with basic ingredients.

Cluster 5 has a relative high proportion of products from the following business units: drinks, hygiene/cleanliness, dairy/frozen, fishery and butchery. The categories more relevant within this cluster are: health care, frozen fish, baby hygiene/protection, oral hygiene, perfumes/cosmetics, consumables, men products, soups, barbecue chicken, frozen desserts, pork meat, spices and current wines. Most products included in this cluster are from a premium or a leader brand. Therefore, this cluster may include customers with a relative high purchasing power and with babies. These customers seem to prefer frozen products in general and to like chicken barbecue and wine. These customers also appear to be particularly interested in health, hygiene and cosmetic products.

Cluster 6 presents a high relative proportion of products from drinks, hygiene/cleanliness, dairy/frozen, fishery and takeaway business units. The most distinctive categories are baby food, cod-fish, powdered drinks/mixes, cereals, fats, frozen meals and pre-cooked meals. Most products included in this cluster are from a premium, leader and own-brand. Summarizing, the potential buyers of these products seem to have a high economic power (despite being lower than the economic power of customers corresponding to Cluster 5). These customers may have babies, and appreciate practical meal solutions, such as cod-fish meals.

Having identified the lifestyle segments corresponding to the clusters of products, we assigned the customers of the company to these six segments. From the total number of customers (1,904,637) whose purchases with a loyalty card were recorded in the transactions database in the months analysed (October and November), we classified only those who bought at least 10 distinct products from those included in the segmentation analysis (i.e., from the 1831 products considered to be the most representative for the construction of the clusters). This resulted in the segmentation of 1,712,307 customers, as the remaining customers would have classifications with little support. For this purpose, we used a binary matrix revealing if a customer did or did not buy a product, and calculated the total number of products that each customer bought from each cluster. Note that since our aim was to investigate lifestyles, which we believe that it revealed by the type of products bought, rather than the quantities bought, we did not prepare the matrix with information on the value of the purchases. The customers were then assigned to the cluster whose products are most similar to those included in the shopping corresponding to the customer past purchases.

The distribution of customers by the clusters is shown in Table 3. By analyzing this table, it can be concluded that most customers belong to Cluster 2 (i.e., customers with medium purchasing power who follow a balanced diet and who enjoy socializing) and Cluster 1 (i.e., customers with medium purchasing power who have babies

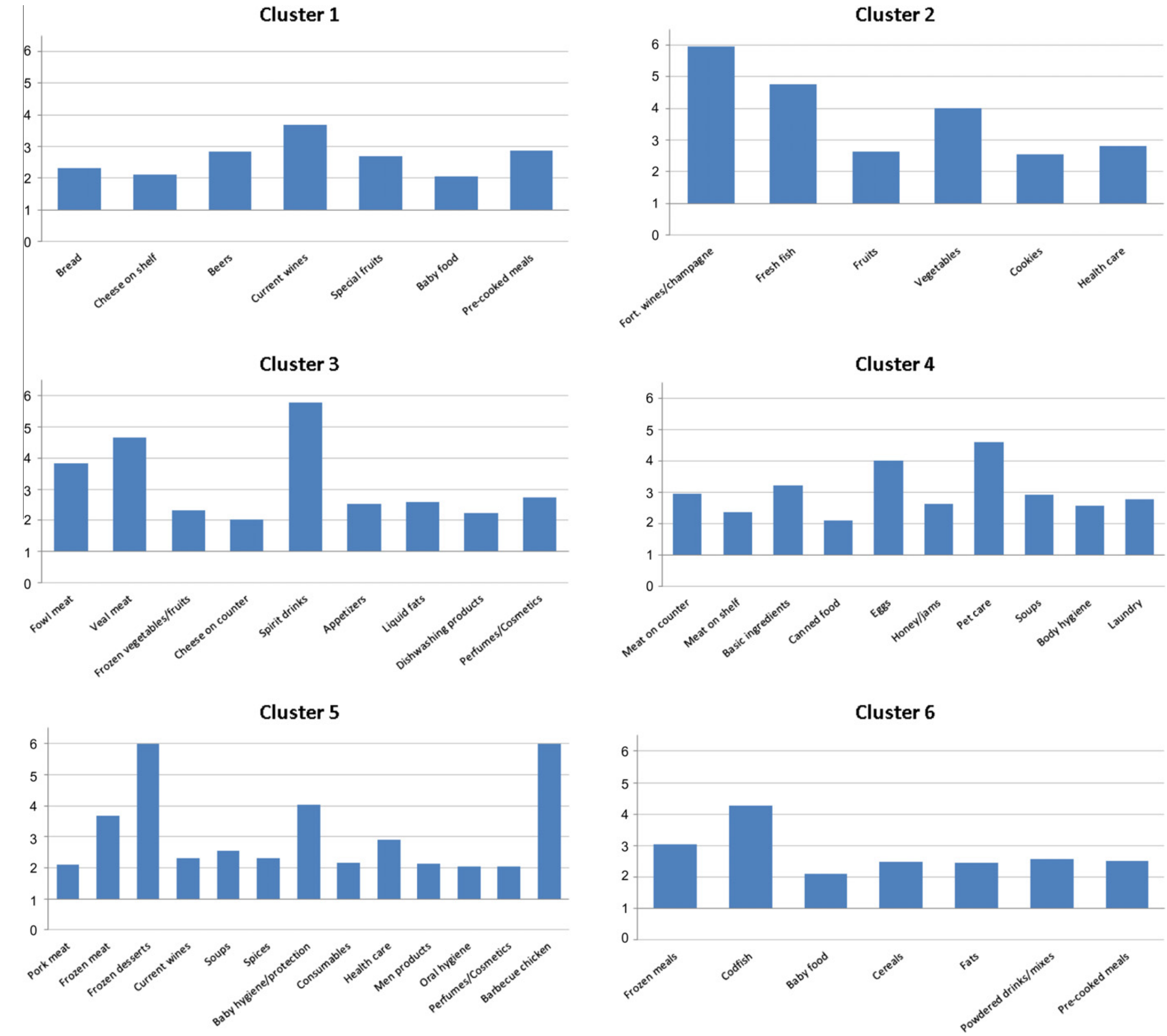

Fig. 3. Proportion of products in the main category.

and appreciate practical meal solutions and wine). In contrast, Cluster 5 (i.e., customers with high purchasing power who have babies, and often buy frozen products, chicken barbecue, wine and health, hygiene and cosmetic products) and Cluster 6 (i.e., customers with high purchasing power who have babies and appreciate practical meal solutions) are the least typical. Therefore, we can conclude that only a minority of company customers has high purchasing power.

6. Marketing actions

The results of the lifestyle segmentation approach proposed in this paper can contribute to the design of company's strategic actions. A well defined market segmentation enables companies to enhance their relationship with customers, leading to higher sales. In this section we provide examples of managerial policies that can be implemented based on the insights provided by the lifestyle segmentation.

One possibility is to use the segmentation for promotional campaigns. In this case, the advantage of lifestyle segmentation is to enable an easy identification of the customers who may be interested in a given product. A promotion is likely to be more successful if there is affinity between the product and the customer needs, such that he/she will feel that the company understands his/her interests.

The range of products in each store can also be adjusted taking into account the segments more representative for each store. Each store should include a considerable diversity of products belonging to the categories more representative for those segments. This action will allow the company to successfully meet the needs of most customers that go to that store, which may lead to an increase in sales and customers' satisfaction.

The layout of each store can also be defined in order to have the categories more representative for the segments of clients that more often visit the store in areas of greater visibility. The increase in convenience for customers can lead to an increase in sales.

We illustrate next some insights that can be gain from the segmentation, and that can be used at the store level. For illustration

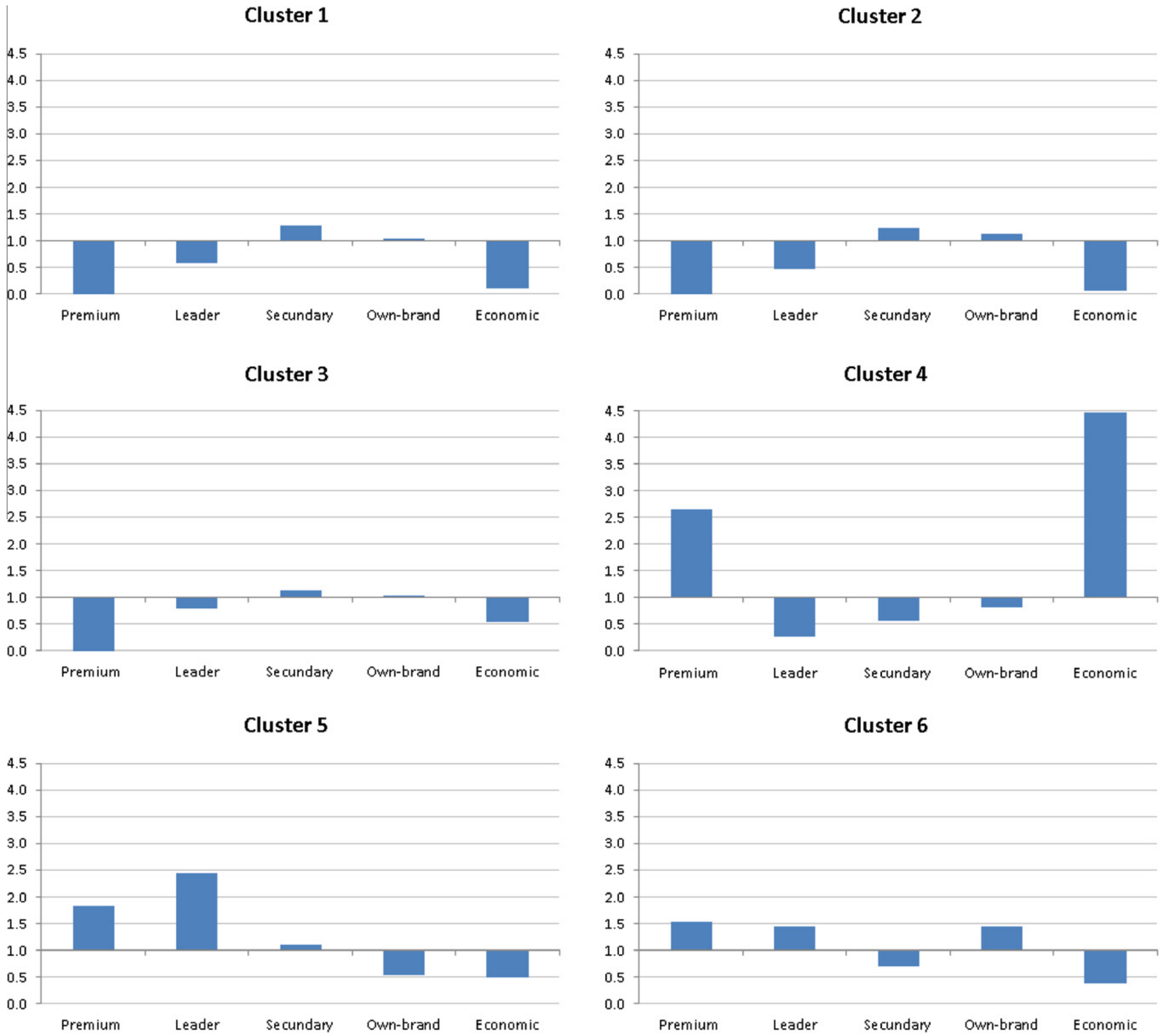

Fig. 4. Proportion of products of each brand position.

Table 3
Distribution of customers by the clusters.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Total
% Customers	22.3	31.7	15.4	17.4	6.4	6.8	100

Table 4
Distribution of customers by the clusters for specific stores.

	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 4 (%)	Cluster 5 (%)	Cluster 6 (%)
Store 1	21.2	32.9	13.8	19.8	4.9	7.4
Store 2	21.8	23.9	10.2	33.3	4.3	6.5

purposes, we show in Table 4 the percentage of customers included in each segment for two stores. Store 1 is a hypermarket located in a metropolitan area while Store 2 is a large supermarket located in the countryside. According to the national institute of statistics, in 2007, the average purchasing power of an inhabitant living in the county where Store 1 is located is 1.5 times higher than the national average purchasing power per capita. Conversely, for an inhabitant living in the county where Store 2 is located the average purchasing power is only 66% of the average national purchasing power per capita.

Table 4 shows that the cluster more representative in Store 1 is Cluster 2, corresponding to customers with medium purchasing power that favor a balanced diet and enjoy socializing. Cluster 4 is the most representative for Store 2, corresponding to customers with low purchasing power, who are likely to have a pet and prepare basic meals. Note that the high representativeness of Cluster 4 in Store 2 could be expected, given the low value of the purchasing power indicator in the store catchment area. As a result, it could be advisable to have in Store 1 a good diversity of products such as fruits, vegetables and fresh fish, whereas in Store 2 the diversity of products such as pet food, basic ingredients and canned food should be larger. These products can be disposed, for example, close to the entrance and main corridors.

7. Conclusion

Customers segmentation can be used to support companies' strategic actions and promote competitiveness. Recognizing customers' differences can be the key to successful marketing, since it can lead to a more effective satisfaction of customers' needs.

This paper segmented customers of an European retailing company according to their lifestyle and proposed promotional policies tailored to customers from each segment, aiming to reinforce loyal relationships and increase sales.

Data mining tools allowed to identify typical shopping baskets based on transactional records stored in the company loyalty card database. These typical shopping baskets were identified using a divisive cluster analysis technique, which considers the correlation between the products purchased. As a result, the products were grouped into six clusters. The methodology also involved the inference of the lifestyle corresponding to each cluster of products, by analyzing the type of products included in each cluster. In particular, it was analysed the business unit, the category and the position of the product brand concerning the value. Each customer was then assigned to the segment whose shopping basket is more similar to the customer's past purchases. The research described in this paper also identified some marketing actions, such as customized promotional campaigns, adjustment of stores' range of products and adjustment of stores' layout, that can help to reinforce the relationship between companies and customers.

Acknowledgements

The funding of this research through the scholarship SFRH/BD/60970/2009 from thePortuguese Foundation of Science and Technology (FCT) is gratefully acknowledged by the first author.

References

Blattberg, R., Buesing, T., Peacock, P., & Sen, S. (1978). Identifying deal prone segment. *Journal of Marketing Research, 15*(3), 369–377.

Bult, J. R., & Wansbeek, T. (1995). Optimal selection for direct mail. *Marketing Science, 14*(4), 378.

Dibb, S., & Simkin, L. (1997). A program for implementing market segmentation. *Journal of Business and Industrial Marketing, 12*(1), 51–65.

Frank, R. E., Massy, W. F., & Boyd, H. W. (1967). Correlates of grocery product consumption rates. *Journal of Marketing Research, 4*(2), 184–190.

Green, P. E., & Wind, Y. (1973). *Multiattribute decisions in marketing*. Dryden Press.

Haley, R. I. (1968). Benefit segmentation: A decision-oriented research tool. *Journal of Marketing, 32*(3), 30–35.

Helsen, K., & Green, P. E. (1991). A computational study of replicated clustering with an application to market-segmentation. *Decision Sciences, 22*(5), 1124–1141.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*(3), 264–323.

Kiang, M. Y., Hu, M. Y., & Fisher, D. M. (2006). An extended self-organizing map network for market segmentation–a telecommunication example. *Decision Support Systems, 42*(1), 36–47.

Kotler, P., Brown, L., Stewart, A., & Armstrong, G. (2003). Marketing. Pearson Education Australia, 6th edition.

Lazer, W. (1964). Lifestyle concepts and marketing. In S. A. Greyser (Ed.), *Toward scientific marketing*. Chicago: American Marketing Association.

Lerman, I. C. (1991). Foundations of the likelihood linkage analysis (LLA) classification method. *Applied Stochastic Models and Data Analysis, 7*(1), 63–76.

Miguéis, V. L., Camanho, A. S., & Cunha, J. F. (2011). Mining customer loyalty card programs: The improvement of service levels enabled by innovative segmentation and promotions design. *Lecture Notes in Business Information Processing, 82*, 83–97.

Min, S., & Han, I. (2005). Detection of the customer time-variant pattern for improving recommender systems. *Expert Systems with Applications, 28*(2), 189–199.

Mitchell, A. (1983). *The nine American lifestyles: who we are and where we're going*. New York: Warner.

Piercy, N., & Morgan, N. (1993). Strategic and operational market segmentation: A managerial analysis. *Journal of Strategic Marketing, 1*, 123–140.

SAS Institute Inc. (2008). SAS/STAT 9.2 User's Guide. Cary, North Carolina.

Twedt, D. W. (1964). How important to marketing strategy is the Heavy user? *The Journal of Marketing, 28*(1), 71–72.

Vigneau, E., & Qannari, E. M. (2003). Clustering of variables around latent components. *Communications in Statistics Simulation and Computation, 32*(4), 1131.