

Application of End-Users Market Segmentation using Statistical Methods

Aleš Žiberna and Vesna Žabkar¹

Abstract

Market segmentation involves identifying homogeneous groups of consumers who behave differently in response to a given marketing strategy than other groups. The purpose of this article is to exercise the possibility of using statistical methods of clustering end-users and to introduce advantages that such methods bring in comparison to traditional methods. We applied statistical methods to group customers into segments according to similarities in usage of health strengthening food supplements, attitudes about health and impulsiveness, based on data from a simple random sample of 241 consumers. Some interesting points and problems encountered are discussed.

1 Introduction

Companies need as much information as possible about client needs, wants and behavioral characteristics, to be able to adjust their marketing activities to achieve successful performance in the served markets. Managers of larger companies usually do not have direct contact with their clients/end-users and therefore do not have enough information at their disposal for making (reliably) effective decisions. Marketing research is usually the source of data about end-users. However, marketing data generally needs some processing to be easily accessible for effective and efficient management decision making. One of the methods of data processing is market segmentation. Results of market segmentation analysis can be applied for gaining a general picture about market conditions, for positioning, and for new product development strategic decisions.

Due to more intensive competition and more demanding end-users the basis for market segmentation is increasingly complex. Researchers are more frequently using statistical methods for clustering individuals in a market, which enables

¹ Faculty of Economics, University of Ljubljana, Kardeljeva pl. 17, 1000 Ljubljana, Slovenia; ales.ziberna@email.si, vesna.zabkar@uni-lj.si

them to consider many more varied and complex characteristics than would be possible without these methods.

2 Market segmentation

Market segmentation is a procedure of sorting the individuals within a market then grouping them into clusters using key strategic variables in order to achieve a clearer picture of the market. In segmentation, the objective is to group consumers into relatively homogeneous groups that respond in a similar manner to marketing activities (Rao and Steckel, 1998: 25). The chosen segmentation characteristics (wishes, needs, knowledge) of consumers should allow the resulting clusters to be homogeneous within and heterogeneous between the segments. This technique also applies to the market of products/services for end-users where there are usually too many users for the company to deal with each user individually (Kotler, 1998: 266). In this article we will concentrate on such end-user markets and use the specific market of health strengthening food supplements.

The segmentation process can be described with the following steps (Dillon, Madden, and Firtle, 1987: 619):

1. Decision about ex-ante or ex-post segmentation
2. Selection and determination of bases for segmentation
3. Selection of variables to describe segments
4. Sample selection
5. Data gathering
6. Formation of segments
7. Description of segments, profile formation
8. Usage of segmentation results for problem solving

Traditionally, market segmentation was performed ex-ante, according to gender, age or some other characteristics of end-users. Ex-post segmentation is done on the basis of data gathered using statistical methods of clustering.

Variables useful for description of the segments could also be used as the basis for segmentation. Different authors define different possible segmentation basis, e.g., Kotler (1998, 271) defines the following segmentation variables:

- Geographical
- Demographical
- Psychographical and
- Behavioral.

Rao and Steckel (1998: 26), cited Frank, Mossy and Wind (1972) and differentiate between general descriptive customer characteristics and characteristics related to consumer behavior.

Segmentation according to behavioral variables, clusters end-users according to the benefits sought, desired application, brand behavior, preferences or sensitivity to marketing mix elements. Furthermore end-users are described with general descriptive customer characteristics. Rao and Steckel (1998: 26) do say that the usage a segmentation basis depends on the circumstances of each environment and firm. However, they claim that the approach of segmenting according to behavioral variables is superior in most cases.

The main advantage of segmentation according to behavioral variables is creation of a relatively homogenous segment solution according to selected behavioral characteristics. Such homogeneity of segments enables better adjustments in marketing activities to end-users. The main weakness of behavioral segmentation lies in the problem of accurately and usefully describing these segments and looking for their “typical” representatives. Accordingly, it is more difficult to determine the size and to target the market segment with the right marketing mix.

Using statistical methods of clustering for selection and determination of bases for segmentation increases the potential for wide use of strategic variables dramatically. The segmentation can be performed using multiple psychographic variables or even a combination of psychographic, behavioral, and demographic (or any other) strategic variables. Such segmentation is not necessarily going to provide better segments, but allows more possibilities to be utilized as clustering criteria.

3 Problem description

The main problems for the producer of health strengthening food supplements related to targeted and effective end-user marketing, are the following:

- What are the general descriptive characteristics of their customers?
- What benefits are sought for health strengthening food supplements?
- Which marketing channel should the company use for these products (stores or pharmacies)?
- Do users have enough information about the products or would they expect more?

Not all of these problems can be solved with segmentation as presented in this article. Here we concentrate on the general descriptive characteristics of the segments such as usage of health strengthening food supplements, segment individuals’ attitudes about health and impulsiveness. In general, the knowledge about end-users is surprisingly limited in the company. It is anticipated that users of the company products are females and older adults.

4 Research methodology

4.1 Bases for segmentation

Several bases for segmentation were applied in the segmentation clustering process, with the anticipation that some variables would not significantly differentiate among respondents in segments. Variables not used for segmentation were applied for the descriptions of different segments.

4.2 Questionnaire

The questionnaire was developed from the variables/bases of segmentation. There are four parts of the questionnaire:

1. Variables that measure the relationship toward different characteristics of health strengthening food supplements.
2. Variables that measure lifestyle and values.
3. Variables related directly to the company's products.
4. Demographic variables.

4.3 Variables that measure relationship toward different characteristics of health strengthening food supplements

Variables that measure relationship toward different characteristics of health strengthening food supplements were formed according to selected characteristics of these products. The selected characteristics of these products are:

- The strength of the effects
- Made of natural ingredients
- Clear specification of their purpose and functioning
- Location of purchase
- Doctors or pharmacists recommendation
- Price
- Side effects
- Rapid working

Two item statements were formed for each measured characteristic. In the process of testing the questionnaire in most cases both statements gave very similar answers. In those cases, one item was selected. For other characteristics more item statements were added because of the complexity of the measured concept. The variables are listed in Table 1 (those, which were used as a basis of

segmentation) and in Table 5 (the remaining variables, used only for description of the segments).

4.4 Variables of life-style and values

The variables were selected from several sources: from Multi-item Measure of Values (MILOV) from Hersche (1994), Impulse buying tendency scale (Weun, Jones and Beatty, 1997), health consciousness scale (Gould, 1988), the values and lifestyle typology (Mitchell, 1983) and from the VALS survey (VALS survey, 2001). The criteria for selecting items from the scales were the following:

- To cover all areas of segmentation bases (demographic, psychographic, behavioral, and attitudes toward products);
- Only few items were selected from each area to shorten the questionnaire;
- Comprehensibility of the items in the Slovene language;
- Acceptability of items for this cultural environment;
- Subjective judgment about suitability of items.

4.5 Variables related directly to the company's products

The questionnaire also included four questions related directly to the company's products. In the first two questions, all the company's products (15) from the category "health strengthening food supplements" were listed. In the first question, the respondents were asked to mark the products they use and in the second question, they were asked to mark the products they know. From these two questions, variables "The number of the company's products used", "Usage of company' products (yes/no)" and "The number of known company's products". The respondents were also asked if they think they have enough information about the company's products and if they would like to have more information regarding these products.

4.6 Demographic variables

Variables including education (coded as years of attending school), gender, age, size of the household, and household income were recorded. Each household was assigned the median of the class of household income they selected. Using the number of household members, the average income per household member was computed.

4.7 Questionnaire testing

The questionnaire was tested for comprehensibility, length and reliability. Low reliability when using Cronbachs' α (Churchill, 1996) required further adaptation of the questionnaire. Retesting the adapted questionnaire yielded sufficiently good results (see Table 1) for further research using the instrument.

Table 1: Results of testing reliability of questionnaire and univariate statistics of the variables used as basis of segmentation.

Variables related to health strengthening food supplements

Variable description	N	Mean	Standard error	Standard deviation
V1 I am looking for as strong as possible effect in health strengthening food supplements.	241	4,74	0,12	1,93
V2 I prefer health-strengthening food supplements which only contain natural ingredients.	241	6,46	0,09	1,33
V3 I would hardly decide for health strengthening food supplements, without clear specification of their purpose and functioning.	241	5,93	0,11	1,73
V4 I would rather buy health-strengthening food supplements in a nearest store (not in a pharmacy), although I know they could not advise me there.	241	2,82	0,14	2,21
V5 I would buy health strengthening food supplements in a nearest store only under the condition that I know exactly what I want to buy.	241	5,38	0,14	2,13
V6 I rather buy health-strengthening food supplements in pharmacies because they can give me advice and additional information there.	241	5,98	0,10	1,58
V7 I prefer health strengthening food supplements recommended my doctor or pharmacist.	241	6,09	0,09	1,43
V8 Price is for me an important factor when choosing between strengthening food supplements.	241	4,37	0,13	2,04
V9 It does not disturb me too much if health strengthening food supplements have any side effects (unless they are frequent and strong).	241	2,47	0,12	1,91
V10 With health strengthening food supplements I am ready to accept side effects (infrequent, weak) if it means they function better (stronger, faster effect.)	241	2,91	0,13	2,01
V11 Good health strengthening food supplements has to start working rapidly.	241	5,05	0,12	1,83

The following groups of questions were created with the sole purpose of testing the reliability of the questionnaire. The groups were made using personal judgment of the similarity of the information, carried by variables.

Questions 3 – 7 (4. reversed)(information, instructions, advice): $\alpha = 0,4556$

Questions 3, 4 (reversed), 6, 7 (information, instructions, advice, version 2): $\alpha = 0,6388$

Questions 9, 10 (side effects): $\alpha = 0,6241$, $r = 0,454$ ($P < 0,001$)

Questions 1, 10, 11 (functioning): $\alpha = 0,5568$

Questions 1, 9, 10, 11 (functioning2): $\alpha = 0,5929$

Variables related to Health

Variable	N	Mean	Standard Error	Standard Deviation
V15 I am regularly examining my health.	241	4,64	0,11	1,71
V16 I reflect about my health a lot.	241	4,58	0,11	1,76

$\alpha = 0,6271$, $r = 0,457$ ($P < 0,001$)

Variables related to Impulsiveness

Variable	N	Mean	Standard Error	Standard Deviation
V39 I avoid buying things that are not on my shopping list.	241	4,00	0,14	2,12
V40 When I go shopping, I buy things that I had not intended to purchase.	241	4,35	0,13	2,01

α (39 reversed) = 0,4830, $r = - 0,319$ ($P < 0,001$)

Variable related to usage of the company's products

Variable	N	Mean	Standard Error	Standard Deviation
V48NO The number of the company's products used	241	1,50	0,10	1,59

Source: Mail survey, March/April 2002, $n = 241$.

4.8 Sampling

The population represents all inhabitants of Slovenian cities with more than 20.000 inhabitants (Ljubljana, Maribor, Celje, Kranj, Velenje, Koper and Novo Mesto). The sampling frame (the telephone directory of Slovenia, 2002) was not ideal, as all units of the population are not listed. It contains more or less households of Slovenia with the following exceptions:

- Some households are not listed because they do not possess a telephone subscription (91,7 % households according to Statistical Yearbook of Slovenia 2001 should hold a subscription) or do not want their number to be listed in the directory (according to unofficial data there should be about 3% of such subscribers). Such households could not be included in the sample and we face noncoverage error (Churchill, 1996).
- Some households have more numbers listed, which gives them higher likelihood to be selected (overcoverage error, Churchill, 1996).

A probabilistic sampling method, a combination of simple random and systematic sample of households was selected. The “last birthday” method of selecting a member within the household was applied (for family members above 15 years). With the combination of the systematic and simple random sample, the final sample contained the number of households in each city proportional to the number of city dwellers in the total population.

For probability samples (e.g. simple random sample) it is possible to determine the size of the sample according to anticipations, required precision of the estimate, standard error of the estimate and the degree of confidence required. Through calculation we derived that we required at least 100 completely filled and returned questionnaires. Based on the required sample and anticipated response rate of 20% or better, 500 numbers were selected from the directory for the first sample flight

Data was collected using a mail questionnaire to reach geographically dispersed units, not to be subject to interviewer bias, and to ensure anonymity of responses (some questions were personal, sensitive).

Errors arose in the research that probably would not have a major impact on results:

1. Noncoverage errors, overcoverage errors which we already mentioned;
2. Nonresponse errors, frequent in mail surveys. In our research, the problem was not a factor as 57 % of contacted respondents replied (48 % completed the whole questionnaire). Nevertheless the question remains whether the nonrespondents differed in relevant characteristics from those that responded. The characteristics of those respondents that replied early were compared to those that replied after a reminder postcard. A T-test on weighted cases addressing this question showed significant differences only on some variables that could be contributed to random factors. Using a χ^2 – test the distribution of responses was compared to the distribution of contacted households according to the city in which the respondents were living. Differences between the distributions were not significant ($P = 0,2678$).

Also, observation errors arose in the research that could have a major impact on results and could not be avoided, e.g. errors of commission as several respondents replied to several consequential questions with the same value (although some of the questions were of the opposite meaning). It seems, as though not all of the respondents engaged in a reasoning process to arrive at an accurate answer. These respondents were not only a few and could not be treated as outliers. Due to similarity of their responses they were kept in the analysis.

4.9 Sample description

The sample was described according to demographic characteristics and was then compared to the population. The final sample size was 241. Several questionnaires have not been included, as they were not completed. Females appear to be over represented in the sample compared to the number of females in the population.

With a χ^2 -test, the distribution of the sample according to gender and age was compared to the population (10-year classes, last class above 65 years). The distribution in the sample differs significantly from the distribution in the population only for males ($P = 0,0180$), but not for both sexes ($P = 0,1668$) and females ($P = 0,0875$).

The average number of household members in the sample did not significantly differ from the average in the population (although we could expect significant differences because of including only city dwellers in the sample). Average income per household member and average income per household were significantly larger ($p < 0,001$) in the sample compared to the population (Statistical Yearbook 2001, 2001), which could be due to the scale (the lowest income level might have covered too large an interval, especially for households with one or two members only).

4.10 Data preparation for statistical clustering

One of the basic decisions in the clustering procedure was about preparation of data for statistical analysis. We standardized all variables and used the principal components method for determining the importance of selected variables and within variables standardization (due to the fact that not all variables were measured in the same units of measurement). Within variables standardization was used because we applied variables with different variances for clustering. However, we did not want the difference in variances to imply the difference in importance.

Also, methods for identification of disturbing variables could be used here (one of the methods is described in Carmone, Kara, and Maxwell, 1999). In our analysis within case standardization was not relevant. The within cases standardization could be used only on variables that were measured on the same scale (or were previously standardized). This standardization would eliminate differences between respondents with the constant tendency to choose higher, lower, or medium values.

4.11 Methods for setting the importance (usually equal) of dimensions and new variables creation

Three methods of setting the importance (usually equal) of dimensions could be used:

1. Principal components
2. Factor analysis
3. Using one variable as representative for the whole range of variables

Principal components analysis is used to set the importance of dimensions by generating one (as in this case) or only a few components that can explain the highest percentage of variance possible by one or other selected number of variables. Factor analysis is used when we need to determine a factor, or surrogate variable behind the measured variables that constitute a dimension. Representative or surrogate variables are used when one variable can sufficiently represent the whole range of variables that constitute a dimension and when we are trying to avoid FA or PCA.

Schaffer and Green (1998) verified usage of factor analysis for "data preparation" and found several problems. The following statements are most important for our application:

1. In most cases segmentation on original data gave better results than segmentation on factors gained from factor analysis of original data.
2. The problem stems from standardization of values after factor analysis that causes equal importance of all factors after standardization (contrary to differences in explained variance by each factor). The problem arises only when we generate more than one factor from each group of variables.
3. Using principal components analysis (without standardization of principal components) results in a better solution (unfortunately SPSS 10.0 only gives standardized solutions).

It is not always necessary to set the importance of dimensions. There are cases when original or only standardized data are sufficient.

5 Segmentation according to variables of usage of health strengthening food supplements, attitudes about health and impulsiveness

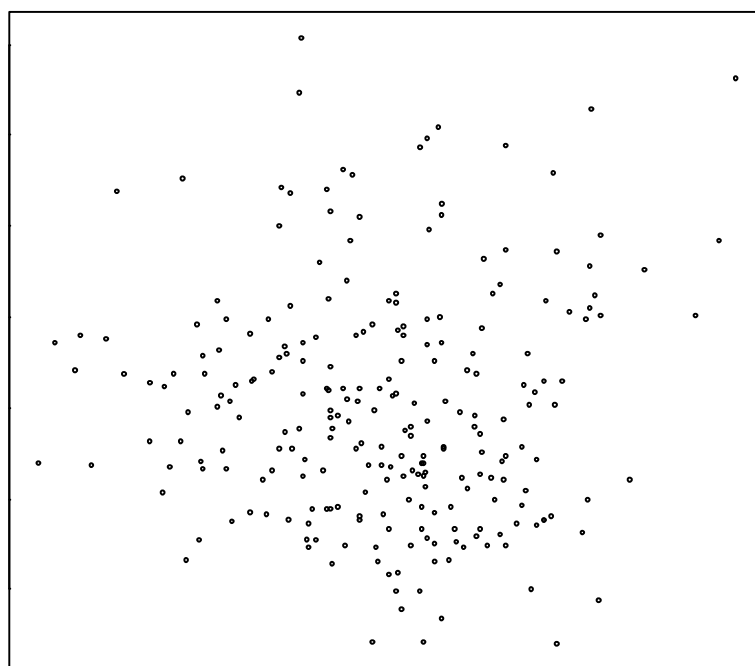
The purpose of segmentation is to determine the existence of segments of users with similar expectations, demands about health strengthening food supplements and to choose the most attractive segments for the company. Additionally it is important to determine if these segments differ from each other according to other

variables, measured by the questionnaire, especially if there are segments that are more inclined to use the company's products.

The segmentation we used here combined variables from different groups, from variables about the usage of health strengthening food supplements, psychographic variables and variables of usage of the company's products. Usage of variables from several groups is recommended if several factors best determine the response of users to marketing mix activities (Neal and Wurst, 2001).

5.1 Graphical presentation using multidimensional scaling

For better presentation of the conditions in the population it is worthwhile to present the distribution of variables according to selected variables. One of the ways to present relative position of units in multidimensional space offers the method of multidimensional scaling (more about the method in Cox and Cox, 1994, and Davison, 1983). The method makes it possible to overview the units in a (most often) two-dimensional space where distances between the units in this space are similar to distances in original (multidimensional) space (Kruskal and Wish, 1986).



Source: Mail survey, March/April 2002, n = 241

Figure 1: Graphical presentation of the data using multidimensional scaling.

It can be seen in the graph presented in Figure 1, that there is a possibility that there are no real segments present in the sample (and therefore presumably in the population). This may be a real situation or may be due to the fact that we showed

units in a two-dimensional space, although each unit has 10 dimensions after the use of principal components method.

5.2 Data preparation and variables used

As the basis of segmentation we used variables listed in Table 1, but only after they were properly prepared.

Units of measurement differed between the variables used. Therefore, usage of the original data was not admissible. Several variables were grouped together using the principal component method. All the principal components were created using a covariance matrix and selecting only the component with the largest percentage of explained variance. The components were standardized. The principal components method was used on selected sets of variables in an attempt to give all dimensions used in the clustering procedure equal importance. The following principal components were created (see Table 1. The variables from Table 1, not mentioned in the following items, were only standardized):

- Component "point of purchase": The component was created from variables V4, V5 and V6.
- Component "side effects": Variables V9 and V10.
- Component "health": Variables V15 and V16.
- Component "impulsive buying": The component was created from variables V39 and V40.

Variables connected to health were included to differentiate between segments according to the characteristics of respondents. Attitudes towards health determine interests about health strengthening food supplements as well as the expected characteristics of the health strengthening food supplements of the company.

Variables related to impulsiveness were included because of the influence of buying behaviour for the company's products (contrary to the majority of the competitors of the company, these products are also sold in retail stores where more shopping occurrences are a result of impulsive buying).

Variables related to the usage of the company's products were included for the purpose of identification of the most attractive segments for the company marketing efforts. Segmentation only on the basis of variables related to health strengthening food supplements resulted in similar segments with regard to the usage of the company's products. Therefore it was difficult to determine more attractive segments (we used the variable *V48NO* – "Number of used company's products", see Table 2).

Table 2: Means and share of users by segments for the company products' users.

Variable	Segments							
	1	2	3	4	5	6	7	Total
Mean for number of the company products used (V48NO)	1,4104	0,8408	4,0033	1,0948	1,0944	1,7322	1,7554	1,5000
Share of users of the company's products (USAGE)	0,6968	0,5505	1,0000	0,5998	0,5906	0,8290	0,9384	0,6918

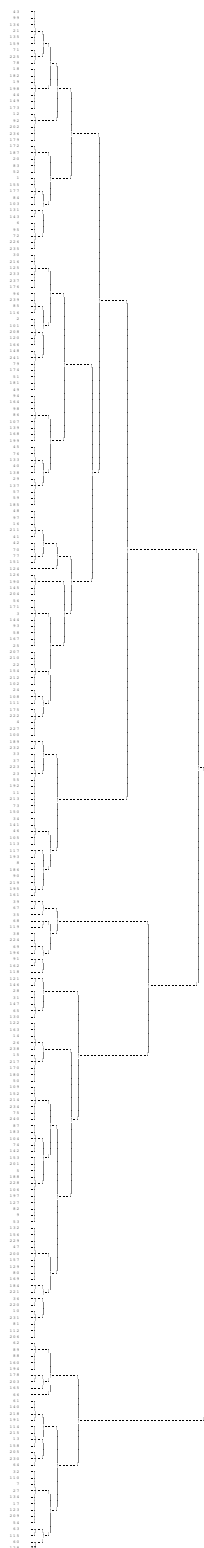
Source: Mail survey, March/April 2002, n = 241.

5.3 Hierarchical clustering procedure

The Ward's method was selected for the hierarchical clustering procedure. This method was selected due to its ability to create compact clusters of similar sizes and shapes, which is also one of its main advantages. The method forms clusters in such a way, that it maximizes the homogeneity of the clusters, which is measured with within clusters sum of squares, which is of course minimized. The method is also often regarded as the most suitable, especially in the field of market segmentation (Sharma, 1996: 193).

The main disadvantages of this method can be "read" from its advantages. It is inclined to forming clusters of approximately similar sizes, even if the data do not support such division. Of course it is also shares the disadvantages and advantages common to all hierarchical clustering procedures.

Based on the dendrogram (Figure 2), agglomeration schedule and examination of solutions with different number of segments the solution with seven segments was selected. Deciding on the number of clusters to accept as practical and most useful is usually one of the toughest decisions in the process. In the case of this analysis, the decision about the number of segments was based on theoretical foundations and practical judgment. The distances between clusters at successive steps served as a useful guideline. Less than seven clusters would in line with dendrogram mean five clusters. In such a case, the cluster with the highest and the cluster with the lowest proportion of the company users would be joined. More than nine clusters would not be very applicable to the development of marketing strategies and activities. Decision for seven segments should therefore not be interpreted as giving all weight to proportion of company users. When determining the most suitable segmentation it should be taken into consideration the content of the segments as well as the practicability of the solution.



Source: Mail survey, March/April 2002, n = 241

Figure 2: Dendrogram for hierarchical cluster and analysis using Ward's method (variables of health strengthening food supplements, usage of the company's products, health and impulsiveness).

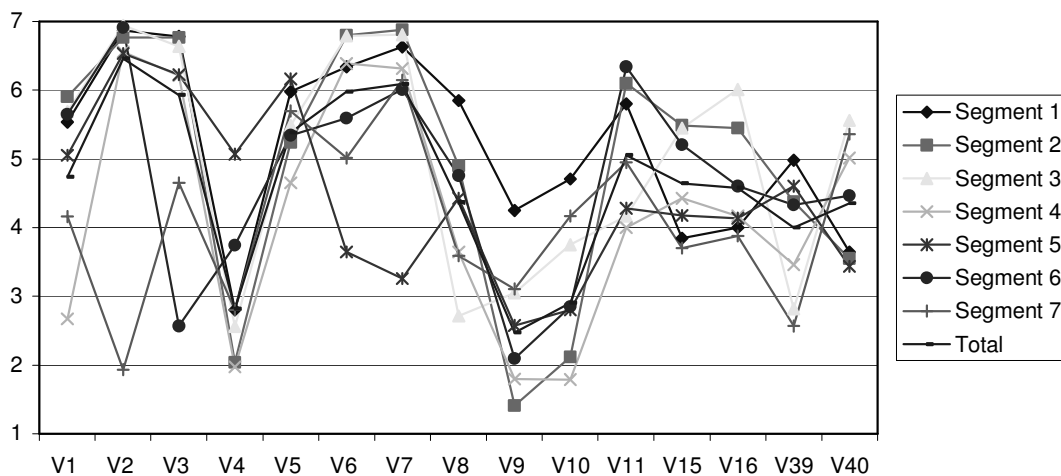
5.4 Non-hierarchical clustering procedures

The non-hierarchical cluster analysis (K-means) with initial seed points from Ward's method of hierarchical analysis was used in an attempt to improve the solution and revealed similar results (81,5% of cases remained in the same cluster). The final solution (using k-means) can be described with the segment means of the variables that were used as basis of segmentation (see Table 3).

5.5 Results

Description of the segments is given in Table 3. For the company, all segments with above average inclination toward their products are interesting. It is also interesting to examine segments with the largest users and the largest number of users.

ANOVA was run on the variables that served as a basis for segmentation and significant differences for all variables ($p < 0,001$) were found. The means for the variables that were measured with seven point Likert scale are presented graphically in Figure 3 and the means for the remaining variables are presented in Table 2. Overview of the segment characteristics regarding those variables not used as basis of segmentation is presented in Table 4.



Source: Mail survey, March/April 2002, n = 241.

Figure 3: The means for the variables that were used as basis of segmentation.

Table 3: The descriptions of all segments using the basis of segmentation.

Nr.	Name	Description	Size
1	Effect, accepts side effects	The segment has an average proportion of users (USAGE) and uses an average number of the company products (V48NO). Fast (V11) and strong (V1) effects, natural substances (V2), description of functioning (V3) and shopping in pharmacies (V4, V5, V6) are desired. The segment is price sensitive (V8), but it accepts side effects (V9, V10). Health (V15, V16) is not seen as very important. The segment is also the least impulsive (V39, V40).	16,5 %
2	Little users, effect	The segment has the smallest proportion of users (USAGE) and uses the fewest (V48NO) company products. Fast (V11) and strong (V1) effects, natural substances (V2), description of functioning (V3), shopping in pharmacies (V4, V5, V6) and doctor advices (V7) are desired. The members of the segment do not want to shop in regular stores (V4, V5, V6). They are not prepared to accept side effects (V9, V10) and give great importance to health (V15, V16). They also are not impulsive (V39, V40).	17,5 %
3	Users only, impulsive	The only users (USAGE), the heaviest (V48NO) users. They desire strong (V1) effects, which do not have to be fast (V11). Natural substances (V2), description of functioning (V3), shopping in pharmacies (V4, V5, V6) and doctor advices (V7) are also desired. They are the least price sensitive (V8) and give great importance to health (V15, V16). They are also very impulsive (V39, V40).	9,8 %
4	Weak effects	The segment has a small proportion of users (USAGE) and uses little (V48NO) of the company products. Strong (V1) and fast (V11) effects and shopping in regular stores (V4, V5, and V6) are not desired, but shopping in pharmacies (V4, V5, and V6) and doctor advices (V7) are. They are not prepared to accept side effects (V9, V10) but are not very concerned about their health (V15, V16). They are not price sensitive (V8) and are quite impulsive (V39, V40).	26,2 %
5	Regular stores	The segment has a small proportion of users (USAGE) and uses little (V48NO) of the company products. Fast effects (V11) are not required. This segment is most prepared to shop in regular stores (V4, V5, V6) and least prepared to shop in pharmacies (V4, V5, V6). Doctor advices (V7) and health (V15, V16) are not perceived as very important. The segment is also not impulsive (V39, V40).	12,3 %
6	Description of functioning is not relevant	Mainly users (USAGE). They desire relatively strong (V1) effects, but especially fast effects (V11) and natural substances (V2). Description of functioning (V3) is not perceived as important. They are quite prepared to shop in regular stores (V4, V5, V6). They are regularly examining their health (V15, V16).	11,9 %
7	Many users, natural substances are not preferred, impulsive	Mainly users (USAGE). Strong (V1) and fast (V11) effects and natural substances (V2) are not desired. The segment is not sensitive to either price (V8) or side effects (V9, V10). They do not have a strong preference to shopping in pharmacies (V4, V5, V6). They are not worried about their health (V15, V16) and are very impulsive (V39, V40).	5,7 %

Remarks: The brackets contain the variables. See Table 1 and 6 for interpretation.

Source: Mail survey, March/April 2002, n = 241.

Table 4: The descriptions of all segments using variables not used as basis of segmentation.

Nr.	Name	Description	Size
1	Effect, accepts side effects	Members of this segment are not very self-confident (V23, V24). They are quite conservative (V44, V45) and like recreation (V33). They know little more company products (V49TOT) than an average person. The segment has the highest share (V52) of men and the lowest average education (V53).	16,5 %
2	Little users, effect	The segment perceives security (V12, V13, V14), human relationships and a sense of belonging (V17 – V22), morals (V26) and being well respected (V27) as quite important. Their members think that there is too much sex on television today (V45). It has the highest share of women (V52) and the lowest average number of household members (V55).	17,5 %
3	Users only, impulsive	This segment values human relationship (V17, V18, V19). It is quite self-confident (V23, V24) and likes variety (V29, V30). Its members often buy the best of everything (V37) and are prepared to pay more for (food) products that do not contain unnatural substances (V41). They are more interested in social than in natural sciences (V46). They are not conservative (V44, V45). They more or less (the highest average score among segments) think they have enough information about the company products (V50), but they would like to have even more information (V51). On average they know the largest number of the company products (V49TOT). They are quite well educated (V53), young (V54) and have larger households (V55).	9,8 %
4	Weak effects	This segment does not give great importance to security (V12, V13, V14), human relationships (V17, V18, V19) and other people's opinions (V28). It's not conservative (V44, V45) and the number of the company products known (V49TOT) by this segment is lower than the average.	26,2 %
5	Regular stores	The members of this segment do not find security (V12, V13, V14) very important. They are more confident (V23, V24) than an average person. They know the fewest company products (V49TOT) and have the highest education (V53).	12,3 %
6	Description of functioning is not relevant	Members of this segment find security (V12, V13, V14) and human relationships (V17, V18, V19) very important. They know few company products (V49TOT). They represent the most conservative (V44, V45) and the oldest segment (V54). They have the lowest income (V56, AVG_INC) and under average education (V53).	11,9 %
7	Many users, natural substances are not preferred, impulsive	Security (V12, V13, V14) is not a very important issue for this segment. Its members have fairly low self-confidence (V23, V24) and do not have preferences to variety (V29, V30) and recreation (V33). They are not as prepared to pay more for products without unnatural substances (V41) as others. They are the least conservative (V44, V45) and the most oriented towards social sciences (V46). Their education level (V53) and age (V54) are under average, but they have the largest households (V55) and income pre household (V56).	5,7 %

Remarks: The brackets contain the variables. See Table 1 and 6 for interpretation.

Source: Mail survey, March/April 2002, n=241.

Table 5: ANOVA on variables that were not used as basis of segmentation.

Variable	significance
V12 I am often concerned about my physical safety.	0,012
V13 Knowing that I am physically safe is important to me.	0,014
V14 Financial security is very important to me.	0,022
V17 I try to be as open and genuine as possible with others.	0,120
V18 When those who are close to me are in pain, I hurt too.	0,238
V19 Without my close friends, my life would be much less meaningful.	0,039
V20 I would rather spend a quiet evening at home than go out to a party.	0,275
V21 It's important to me to feel I am a part of a group.	0,029
V22 I need to feel there is a place that I can call "home".	0,606
V23 I think I am more self-confident than most people.	0,044
V24 I have more ability than most people.	0,486
V25 I like being in charge of a group	0,349
V26 Knowing that I am doing the right thing in a given situation is worth any price.	0,005
V27 My social position is very important to me.	0,163
V28 I am easily hurt by what others say about me.	0,206
V29 I like a lot of variety in my life	0,710
V30 I like trying new things.	0,174
V31 My life is pretty much the same from week to week.	0,013
V32 I think I am sometimes unpredictable or spontaneous.	0,595
V33 Recreation is an integral part of my life.	0,192
V34 Recreation is important for a healthy life.	0,428
V35 Having fun is important to me.	0,967
V36 I like to learn about things even if they may never be of any use to me.	0,107
V37 I like to buy the best of everything when I go shopping.	0,474
V38 I treat myself well.	0,555
V41 I prefer buying food products without artificial substances (conservatives, colors, flavors), even if they are more expensive.	0,007
V42 I would rather make something than buy it.	0,305
V43 Air pollution is a major worldwide problem.	0,156
V44 A woman's life is fulfilled only if she can provide a happy home for her family.	0,000
V45 There is too much sex on television today.	0,000
V49TOT The number of known company's products	0,000
V50 I think I have enough information on Company health-strengthening food supplements, its effects and usage.	0,187
V51 I would like to have more information on Company health-strengthening food supplements, its effects and usage.	0,046
V52 Gender	0,285 *
V53 Years of education	0,166
V54 Age	0,002
V55 Number of members of the household	0,035
V56 Household net income	0,098
AVG_INC Household net income per household member	0,260
USAGE Usage of Company products (yes/no)	0,000 *

Remarks:

* By this question we are not testing hypothesis about means, but hypothesis about proportions (which are actually represented by the means).

Source: Mail survey, March/April 2002, n=241.

5.6 Validation

Validation was done by searching for significant differences between segments regarding variables that were not used as basis of segmentation. Table 5 contains significance levels for variables that are related to the discussed problem. Based on results presented we could talk about ability of clusters to show the differences on variables not used to form the clusters. Criterion validity is therefore supported.

6 Conclusions

From the company point of view, segment 3 – *"Users only, impulsive"* seem the most appealing. It contains strictly those users of company products, and those who use several of its products. The segment also does not demand fast effects and wants natural substances, but unfortunately also desires strong effects, shopping in pharmacies and doctors advice. Its favorable characteristics also include impulsiveness, health-consciences and lack of price-sensitivity.

There are two more interesting segments. Segment 6 – *"Description of functioning is not relevant"* seems to present mainly older costumers who probably use primarily older products that they are used to purchasing and using. The other interesting segment is segment 7 – *"Many users, natural substances are not preferred, impulsive"*. Its interesting characteristics are its lack of desire for natural substances and lack of interest in health, characteristics that one would expect from company customers. On the other hand they also do not demand fast and strong effects and are not price sensitive, but impulsive. Their young age suggests that this segment could consist of young people who are mainly consumers, but are not also buyers. They find the company products at home or they more or less buy them in an impulsive gesture, since they experienced them at home.

7 Interesting points and problems

The selection of variables used as basis of segmentation in this segmentation is a rather unusual mix. Usually, all segmentation variables come from the same group (demographic, behavioral, psychographic), but in this case variables come from several different groups of variables. The included variables for segmentation measure benefits sought as well as some psychographic and usage characteristics. The usage of such a basis of segmentation is justified, when all variables included influence consumers response to the marketing mix (Neal, Wurst, 2001, p. 14–18).

Initially we intended to use only variables measuring benefits sought as the basis of segmentation. The reason for the decision to include other variables is the

failure of the segmentation on the previously mentioned basis to produce satisfactory results. The main problem was that the segments did not differ significantly regarding the usage of the company's products. There may be several reasons for such results. First, one of the problems could be an inappropriate measuring scale. Here, a seven (7) point Likert scale was used to measure benefits sought, but the returned answers imply that the respondents may not have been able to accurately assess the benefits they sought nor assess their psychographic characteristics properly, using this scale. Second, the seven (7) point Likert scale was used as an interval scale although it could be interpreted as ordinal. These two problems present a minor limitation of the study: seven-point scales are usually considered to be sufficient for practical purposes and noncomparative Likert scales with interval scale characteristics are widely used in marketing research (Malhotra and Birks, 2003; Churchill, 1996).

It is interesting to see, that although one variable that measures usage of the company's products was used as one of the basis of segmentation, there are no segments without users in the solution. This is probably the result of two factors. Firstly, users of the company's products present a large portion of the population (approximately 70%), and secondly, many users of the company's products do not differ significantly from nonusers regarding the other variable basis of segmentation explored in this research. Furthermore, when segmentation was done solely on the basis of the status of consumers regarding usage of the company's products (users/nonusers), very few significant differences were found.

References

- [1] Bearden, W.O. and Netemayer, R.G. (1993): *Handbook of Marketing Scales*. Newbury Park [etc.]: Sage, **XII**, 352.
- [2] Bearden, W.O. and Netemayer, R.G. (1999): *Handbook of Marketing Scales*. Second edition. London [etc.]: Sage: Association for Consumer Research. **XIV**, 537.
- [3] Carmone, F.J., Kara, A., and Maxwell, S. (1999): HINoV: A new model to improve market segment definition by identifying noisy variables. *Journal of Marketing Research*, **36**, 501-509.
- [4] Churchill, G.A. (1996): *Basic Marketing Research*. 3rd ed. Forth Worth: The Dryden Press, 863.
- [5] Cox, T.F. and Cox, M.A. (1994): *Multidimensional Scaling*. 1st ed. London: Chapman & Hall, 213.
- [6] Davison, M.L. (1983): *Multidimensional Scaling*. New York: John Wiley & Sons, 242.
- [7] Dilon, W.R., Madden T.J., and Firtle N.H. (1987): *Marketing Research in Marketing Environment*. Second Edition. Homewood: Irwin, 853.

- [8] Frank, R.E., Massy, W.F., and Wind, Y. (1972): *Market Segmentation*. Englewood Cliffs, NJ: Prentice –Hall.
- [9] Gould, S.J. (1988): Consumer attitudes toward health and health care: A differential projective. *Journal of Consumer Affairs*, **22**, 98–118. In W.O. Bearden and R.G. Netemayer: *Handbook of Marketing Scales*. Second edition. London: Sage: Association for Consumer Research, cop., 1999, 132–133.
- [10] Hersche, J. (1994): Measuring social values: Multi-item adaptation of the list of values (MILOV). Working paper report number 94–101, Cambridge, MA: Marketing Science Institute. In W.O. Bearden and R.G. Netemayer: *Handbook of Marketing Scales*. Second edition. London: Sage: Association for Consumer Research, 23–25.
- [11] Kotler, P. (1998): *Marketing Managment*. Ljubljana: Slovenska knjiga, 832.
- [12] Kruskal, J.B. and Wish, M. (1986): *Multidimensional Scaling*. Beverly Hills: Sage, 93.
- [13] Malhotra, N.K. and Birks, D.F. (2003): *Marketing Research: An Applied Research*. Harlow: Pearson Education, 786.
- [14] Mitchell, A.: *The Nine American Lifestyles: Who We Are and Where We're Going*. New York: MacMillan, 1983. In W.O. Bearden and R.G. Netemayer: *Handbook of Marketing Scales*. Newbury Park: Sage. 1993, 89–94.
- [15] Neal, W.D. and Wurst, J. (2001): *Advances in Market Segmentation*. Marketing Research, Chicago, **13**, 14-18.
- [16] Rao, V.R. and Steckel, J.H. (1998): *Analysis for Strategic Marketing*. Reading (Mass.): Addison-Wesley, **XIII**, 514.
- [17] Schaffer, C.M. and Green, P.E. (1998): Cluster-based market segmentation: Some further comparisons of alternative approaches. *Journal of the Market Research Society*; London: Market Research Society, **40**, 155-163.
- [18] Sharma, S. (1996): *Applied Multivariate Techniques*. New York: John Wiley & Sons, 493.
- [19] Statistical Yearbook 2001. Ljubljana: Statistični urad Republike Slovenije, 2001.
- [20] VALS Survey. [<http://future.sri.com/vals/surveynew.shtml>], SRI Consulting Business Intelligence, 2001.
- [21] Weun, S., Jones, M.A., and Beatty, S.E. (1997): A Parsimonious Scale to Measure Impulse Buying Tendency. In W.M. Pride and G.T. Hult: AMA Educator's Proceedings: *Enhancing Knowledge Development in Marketing*. Chicago: American Marketing Association, 306–307. In W.O. Bearden and R.G. Netemayer: *Handbook of Marketing Scales*. Second edition. London: Sage: Association for Consumer Research, cop., 1999, 57.