

CONSUMER BEHAVIOR TESTING IN THE MINERAL WATER MARKET THROUGH FACTOR AND CLUSTER ANALYSIS

PARASZT MARTA ¹

¹*Szent István University, Doctoral School of Management and Organizational Sciences, Godollo, Hungary*

Corresponding author's e-mail: paraszt.marta@gk.szie.hu

Abstract: *Over the last two decades, the consumption of bottled mineral water has become a fashion in Hungary. The purpose of my study was to form homogeneous groups for attitudes towards mineral water selection. The data on the research topic was collected using an online self-filled questionnaire. During the sampling I used an arbitrary method, therefore the results obtained can only be interpreted to the examined population. During factor analysis, I have reduced the answers to the questions as variables to 4 factors. Then I have created 4 clusters along the factors.*

Key words: *consumer behavior, mineral water, questionnaire, attitude*

INTRODUCTION

In the last 2 decades it has become the fashion the consumption of the bottled mineralwater in Hungary too. It can be read on the side of the The Hungarian Mineralwater, Juice and Soft Drinking Federation [4], that in the 80's 90's years the low consumption was characteristic, however from 1993 it has begun to grow dynamically and for today we are in the first 5 country on the European mineralwater-consumption list.

In my thesis I examine the themetopic of the mineralwater consumption, especially, the effect of the applied marketingcommunication tools on the consumers. My goal was to make homogeneous groups/clusters according to the concerning mineralwater-choice attitudes.

MATERIALS AND METHODS

The collection of the examined theme from the quantitative methods was made through self-completed questionnaire, which I updated on online sites. Totally 223 fillings were produced. Through the sampling I applied arbitrary sampling without filter condition, that is why according to Gyenge [1] it can lead to malformed results, when we would like to make general findings. So these results and the findings made from them are interpreted only for the examined crowd. The questions concern on the influencing elements of the mineralwater choose.

The details of the questionnaire were processed by the IBM SPSS Statistics 23 program.

The received answers, as variables I enumerated into factors. The factor-analysis [2] is a structure exploratory method, which means, that we pursue to explore the relations between the variables, what makes the interpretation and the analysis of the details by the reduction of the available numbers of the variables. I chose the main component – analysis as a compression method. The most important feature of the method is that every components explain the variance of the examined variables in order to the own value.

Regarding that, whether the variables are appropriate for factor-analysis I controlled by the anti-image matrix¹, the Bartlett – test (The test examines that whether the variables in the basic crowd uncorrelated are (nullhypothesis). Throug the examination we try to refuse the nullhypothesis, because the basic condition of the factoranalysis that the

¹ The matrix was not published because of the extent limits.

variables must correlate with each other.) and by one of the most important measures, the KMO value.

The KMO value is the average of the MSA values. The MSA value shows, how close relationship can be between a variable and all the other variables. Furthermore, I examined the communality of the factors too, which informs us about that a variance of variables how much part of all the other factors is explained.

On the one hand, for entering the numbers I used the Kaiser criterion, which says, that we have to take into consideration only those factors, which has the value 1. I took into consideration the varianceproportion as well, and according to this the numbers of the factors can be specified by the cumulated percent of the variance (the 60 % is accepted in the social sciences). The figure Scree-plot – which figures its own value in order of the factors - can also help in the defining of the numers of the factors.

According to the methods I received different solutions, the Maximum – likelihood method helped me in the definition of the adequate factor-number. The ML method calculates a fit indicator by the given factor-number and in this case the 0,1 and the lower significant level means that the factor model doesn't fit well to the examined details.

After the definition of the suitability of the variables and the factor numbers I got a factorweight-matrix, and on this matrix I implemented a rotation, which had the goal to receive simpler and more meaningful factorsolutions. From the orthogonal rotation methods I applied the most frequently the Varimax method, which comparing to the others is more stabil and separates the factors better.

With the help of the clusteranalysis, I tabulated the interviewed persons through the received factors. The purpose of the clusteranalysis is to sort the observed units into homogeneous groups according to the coated variables into the analysis. The analysis, different from the factoranalysis try to reduce the number of the observed units instead of the variables [2].

During the clusteranalysis, firstly with hierarchical process – with Ward method² – I specified the ideal number of the clusters, clustercentres. The essence of the method, that it calculates for every cluster the average of all the variables, than for every observation units the squar Euclidean distance, it summerizes these distances for all the observation units. It aggregates the 2 clusters at every step, in which scattersquare has the smallest increase in the cluster. During the determination of the cluster numbers it's woth to stop, where the value of the coefficients suddenly increases, so we aggregate distances, which are far from each other. The result was controlled by two-step algorith. Its feature is that we have the opportunity to give a clusternumber, which is graded by the program. During the clusteranalysis I applied the log-likelihood method.

As a next step I tabulated the observation units with non-hierarchical K-middle method, - originated from the hierarchical method – based on center points. The observation unit will be in the cluster, which is the closest to the clustercenter. When every elements are ranked by the program, the central point of the individual groups are recalculated and based on this repeats the group training. The process lasts for that, til we find a stabel center point, after that we don't need to reclassify any more elements [3].

RESEARCH RESULTS

During the factoranalysis I involved 14 pieces 5 graded variables into the research measured on Likert-skale - concerning the selection attitude of the mineralwater, depending on the nature of the questions, the 1 „in the smallest measure influences” and „in the less measure important” – has the meaning.

² The concentration table of the Ward process was not published because of the extent limits.

The first examination locked out from the result the price and the packed formulation. The examination of the remaining 12 variables (figure 2 column 1) gave the following results: KMO value 0,726 (figure 1), according to this value it is right for the analysis of the variables. The null-hypothesis of the Bartlett-test was discarded, because the significant level is lower than 0,05 (figure 1).

KMO and Bartlett's Test			
Kaiser-Meyer-Olkin Measure of Sampling			,726
Bartlett's Test of Sphericity	Approx. Chi-Square		539,718
	df		66
	Sig.		,000

Figure 1. KMO and Bartlett-test

Source: own editing, 2017

During the examination of the communality of the variables (figure 2) I can determine, that in every case it reaches the minimum value 0,25, so every variables has a minimal explaining power.

Communalities		
	Initial	Extraction
Advertisement of the tapwaters have influence	1,000	,377
The quality of the tapwaters in Hungary is adequate	1,000	,409
The Mineralwater is healthier then the tapwater	1,000	,592
Instead of the tapwater when I can do I drink mineralwater	1,000	,705
When I buy mineralwater usually I read the text of the bottle label.	1,000	,416
Brand	1,000	,357
Certificate Trademark (Excellent Hungarian Nutrition, Heart-healthy)	1,000	,671
Prizes in the orginazations	1,000	,679
Prizegames	1,000	,421
Packaging	1,000	,476
The origin of the mineralwater when it was bottled	1,000	,491
Advertisements	1,000	,655

Extraction Method: Principal Component Analysis

Figure 2. The examination of the communality of the variables

Source: own editing, 2017

The determination of the number of the factors: due to the Kaiser criterion the own value of 3 components exceed the 1, however, according to the the Variance-ratio method the 4 factornumber is that, which keeps 60 % of the informations (figure 3). The figure Scree-plot confirms also a 4-factorial solution - the elbow rule says that the numbers of the factors must be maximize there, where the slope of the curve suddenly changes and turns into straight (figure 4).

Component	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,205	26,712	26,712	3,205	26,712	26,712	2,478	20,649	20,649
2	1,701	14,178	40,890	1,701	14,178	40,890	2,024	16,870	37,519
3	1,341	11,178	52,068	1,341	11,178	52,068	1,746	14,548	52,068
4	,994	8,282	60,350						
5	,869	7,245	67,595						
6	,784	6,530	74,125						
7	,718	5,987	80,112						
8	,587	4,893	85,005						
9	,571	4,756	89,762						
10	,505	4,205	93,968						
11	,384	3,199	97,167						
12	,340	2,833	100,000						

Extraction Method: Principal Component Analysis.

Figure 3. Kaiser-criterion and Variance – ratio method
 Source: own editing, 2017

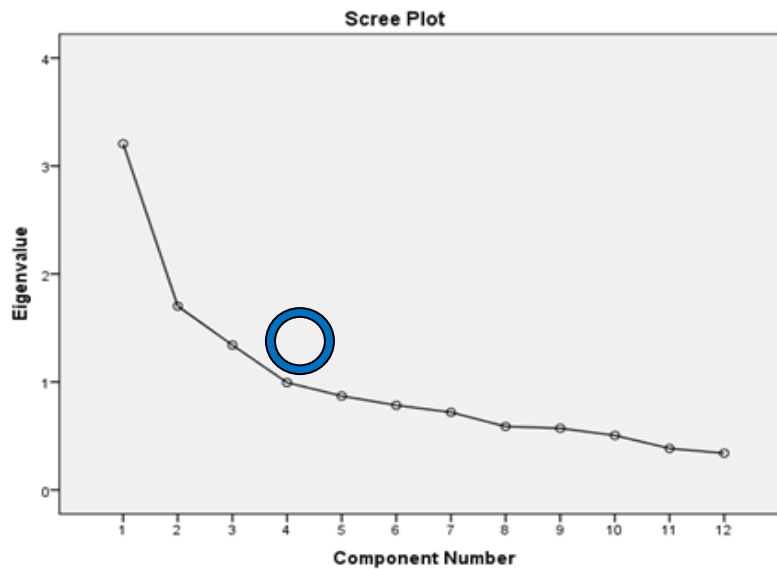


Figure 4. Scree-plot figure
 Source: own editing, 2017

The examination of the fit of the model confirms (figure 5) that the 4 - factorial model fits better the details.

Factor Analysis		
Goodness-of-fit Test		
3 factorial	Chi-Square	Sig.
	49,280	,034

Factor Analysis		
Goodness-of-fit Test		
4 factorial	Chi-Square	Sig.
	31,069	,152

Figure 5. Examination of the fit of the 3- and 4-factoral model
 Source: own editing, 2017

The figure 6 shows the 4 main components according to the sequences of the growth of the explained variance. After the rotation it can be identified, that in the case of the 4 main component we reached the minimum 60 % information retention level, what is more, in case of 4 factors the KMO value exceeds the expected 1 value.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared			Rotation Sums of Squared		
	Total	Variance	e %	Total	Variance	e %	Total	Variance	e %
1	3,205	26,712	26,712	3,205	26,712	26,712	2,427	20,223	20,223
2	1,701	14,178	40,890	1,701	14,178	40,890	1,737	14,474	34,696
3	1,341	11,178	52,068	1,341	11,178	52,068	1,554	12,947	47,644
4	,994	8,282	60,350	,994	8,282	60,350	1,525	12,706	60,350
5	,869	7,245	67,595						
6	,784	6,530	74,125						
7	,718	5,987	80,112						
8	,587	4,893	85,005						
9	,571	4,756	89,762						
10	,505	4,206	93,968						
11	,384	3,199	97,167						
12	,340	2,833	100,000						

Extraction Method: Principal Component Analysis.

Figure 6. The fully explained variance
 Source: own editing, 2017

The variables having the factorweight over 0,5 were classified in the rotated factorweight-matrix (figure 7). The factorweight shows the correlation between the origin variable and the factor, and its value can change from -1 to 1. The number of the elements must be taken into consideration to get significant factorweight. In this case, according to the number of the elements the value of the factorweight must reach the 0,4.

Rotated Component Matrix				
	Component			
	1	2	3	4
Certificate Trademark (Excellent Hungarian Nutrition, Heart-healthy)	,817	-,017	-,047	,091
Prizes in the organizations	,779	-,041	,152	,225
The origin of the mineralwater when it was bottled	,659	,148	,143	,124
When I buy mineralwater usually I read the text of the bottle label.	,637	,133	,200	-,137
Instead of the tapwater when I can do I drink mineralwater	,134	,815	-,062	,144
The Mineralwater is healthier then the tapwater	-,053	,755	-,113	,106
The quality of the tapwaters in Hungary is adequate	-,120	-,664	-,187	,191
Prizegames	,051	-,082	,831	,031
Packaging	,236	,052	,679	,222
Advertisement of the tapwaters have influence	-,011	,002	,118	,772
Advertisements	,143	,075	-,500	,625
Brand	,444	,067	-,016	,571

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 5 iterations.

Figure 7. Rotated factorweight matrix
 Source: own editing, 2017

The certification trademarks, prizes awarded at events, the importance of the origin and the perusal of the bottle label belong to the factor 1, which have positive sign, so they have one way movement. During the election it is an essential aspect, that the product must be qualified excellent, there happen with bigger probability the perusal of the bottle label.(for example: with respect of the origin or the composition) The factor 1 took up the name of „*Consciousness factor*”.

The second factor includes 3 variables. The factorweight is a variable – the quality of the tap water in Hungary is adequate – in this case negativ, which is so interpratable, that those respondents, who „can do” consume mineralwater, because/and they think, that the mineralwater is healthier than the tap water, they think the quality of the tap water in our country is not appropriate for drinking. The called the second factor „*Anti -Tap water factor*”.

The 3rd factor contains the prize games and the role of the wrapping ant its importance. Both of the variables have pozitiv factorweight, so the variables are moving together. When the prize game is important to choose the bottle, than to overview the packaging has a major role – in order to recognize the prize game. The role of the advertisements appears in a little measure (0,5) in this factor too (and in the 4th too). It often occurs that we meet prize games in the advertisements, therefore it has an indirect influence on the election. The third factor got the name „*Game factor*”.

The 4th component – „*Advertisement consumer factor*” has the following content: the influencing role of the ads appears during the mineral water election, the ads affect on the consumers in every area and the brand loyalty is here the strongest. According to the pozitiv sign of the factorweights, the loyalty arisens in the case of that brand, which occure the most frequently in the advertisements.

According to the received 4 factors with the usage of hierarchical clusteranalysis (Ward proceeding) I strived for the formation of the optimum groupnumber. Due to the elbowrule, the 4 cluster solution is appropriate. According to the linediagram (figure 8) I can state, that a sharper fracture can be found in the 220th step.

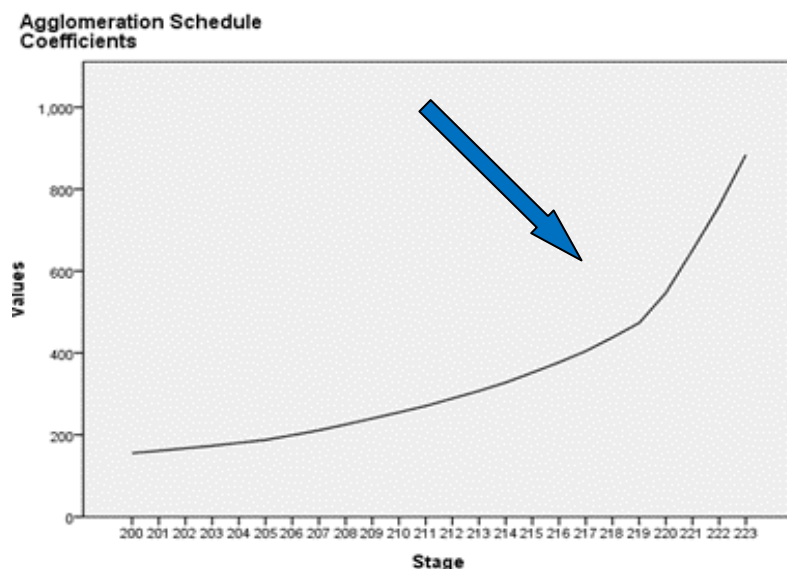


Figure 8. Figure of the coefficients depending on the steps of the concentration
 Source: own editing, 2017

The clusteranalysis in 2 steps determined the 4 clusters solution too (figure 9).

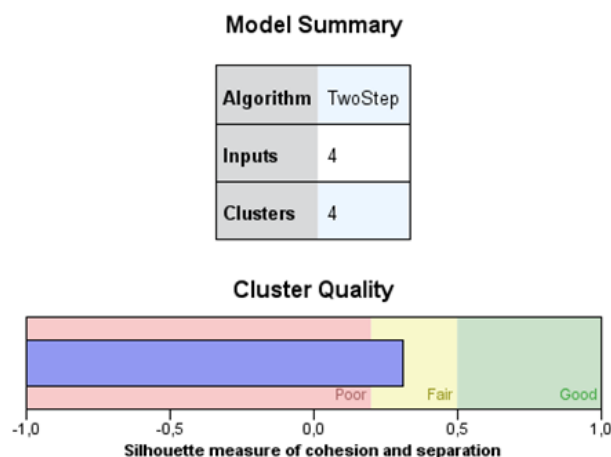


Figure 9. Clusteranalysis – model summary in 2-steps

Source: own editing, 2017

The analysis of the clusters happened according to the K-middle method determined by the clustercentroids (figure 10).

Final Cluster Centers				
	Cluster			
	1	2	3	4
Consciousness factor	-,32217	,21312	,75410	-,73886
Anti-tap water factor	1,18833	,05597	-,50799	-,54301
Game factor	-,39660	1,29748	-,58923	,18915
Advertisement consumer factor	-,02377	,96483	,13294	-,78970

Figure 10. Clustercentroids

Source: own editing, 2017

The feature of the observation units which belong to the first cluster (figure 10, column 1) who think that the mineral water is healthier, than the tap water and therefore they are dissatisfied with the quality of the tap water. The consciousness isn't characteristic for them, it doesn't matter in the election of the mineral water, whether it has a certification trademark or prize or not, so they don't read the label on the bottle. The prize game are neither important for them and the advertisements affect them only at an average level. By and large, we can state, that it's all the same with them, what kinds of mineralwater they drink, just do not tap water. They are the “*at all costs cluster*”.

The second group prefer paramountly the prize games (figure 10, column 2), and they choose the mineralwater depending on the prizes. At every area of their life is characteristic that the ads have influence on them. The mineralwater of those distributors are consumed first of all, who advertise prize game sin the ads. They neither think, may the mineralwater is healthier than the tap water, nor the quality of the tap water is bad in Hungary. There is a mild consciousness, what is characteristic them. They are the “*prize cluster*”.

The group number 3 has the opinion, that is is more important for them to drink a mineralwater, which has a certificate or trademark (figure 10, column 3). We consider it more consciously the product label to get a high quality product from the shelves. The listen to the ads barely and they are not interested in the prize games. They don't have the goal at all costs – of any quality – to drink mineralwater instead of the tap water. Probably,

they don't consider it as a bad quality. Certainly, they have bottled water at home only for the hospitality. They are the “*self-exhibitionist cluster*”.

The 4th cluster (figure 10, column 4) is the “*not caring group*”. They can't be influenced by the advertisements – mentioned any kinds of the areas, the quality certifications, they are really not interested in the brand compared to the other groups. As a matter of fact they think the less that the tap water have bad quality, that is why they don't want to drink mineralwater. If they would drink bottled water, than they choose it according to the packaging, maybe it can happen because of the chance of the prize.

CONCLUSIONS

Analysing the mineralwater-consumption habits of the the examined 223 persons, I managed to classify the decision criterions and the attitudes into 4 factors, what is more, along the 4 factors I could make 4 really well separated groups regarding the mineralwater-consumption attitude.

REFERENCES

- [1]. **GYENGE B.**, 2009, Marketingkutatás jegyzet, Szent István Egyetemi Kiadó, Gödöllő, p.154
- [2]. **SAJTOS L.-MITEV A.**, 2007, SPSS kutatási és adatelemzési kézikönyv, Alinea Kiadó, Budapest, p. 245-328
- [3]. **SZÚCS I.**, 2004, Alkalmazott statisztika, AGROINFORM Kiadó és Nyomda Kft., Budapest, p. 496-510
- [4]. <http://www.asvanyvizek.hu/index.php/erdemes-tudni/asvanyviz-fogyasztasi-adatok>