Animal Science Papers and Reports vol. 36 (2018), no. 2, 171-184
Institute of Genetics and Animal Breeding, Jastrzębiec, Poland

Using cluster analysis and choice-based conjoint in research on consumers preferences towards animal origin food products. Theoretical review, results and recommendations*

Agnieszka Tekień[1], Krystyna Gutkowska[1]*, Sylwia Żakowska-Biemans[1], Artur Jóźwik[2], Monika Krotki[1]

[1] Department of Organization and Consumption Economics,
Faculty of Human Nutrition and Consumer Sciences Warsaw University of Life Sciences,
Nowoursynowska 159c, 02-776 Warsaw, Poland

[2] Institute of Genetics and Animal Breeding, Polish Academy of Sciences Jastrzębiec,
Postępu 36A, 05-552 Magdalenka, Poland

(Accepted January 17, 2018)

Understanding the consumer requirements is an important issue in food research. Nowadays, constant changes in consumer attitudes towards food products can be observed, therefore recognizing the reasons for consumer decisions can be a challenge. Products of animal origin are no longer considered only in terms of quality (ex. flavour) but also safety, nutritive value, sustainability of production methods and animal welfare standards are becoming increasingly important. In these conditions, cluster analysis is a very useful complex statistical method that allows to investigate consumers behaviour more precisely then using traditional methods. This article presents various consumer segmentation methods used in analysing conjoint choice-based study based on food research exploration. It combines both theoretical aspects and practical recommendations with clear information for researchers how to choose the best segmentation method in order to analyse the food market.

---

*The survey was a part of "BIOFOOD – innovative, functional products of animal origin" project that was co-financed by the European Union from the European Regional Development Fund within the Innovative Economy Operational Programme 2007-2013.
**Corresponding author: krystyna_gutkowska@sggw.pl

KEY WORDS: animal origin products / cluster analysis / conjoint analysis / conjoint choice-based consumer behaviour / segmentation /

Modern world is changing rapidly and so is consumer behaviour. Classical economic perspective [Engel 1968, Fishbein and Ajzen 1975] was focused on defining consumer attitudes, habits and segmentations. Attitudes were clearly defined and fairly constant over time. Nowadays consumer preferences are discussed rather than consumer attitudes [Ariely 2008] which can be changing at any time. Consumers no longer want to maintain a chosen life-style, or to conform to hackneyed attitudes, but also participate in the ongoing decision-making process. They merge various attitudes, outlooks, choices. Consumers do not perceive their everyday choices as elements of a definite lifestyle. They simply see a multitude of possibilities and they "enjoy the moment" by choosing what most appeals to them at a given time. As behavioural economics [Kahneman and Tversky 1979] develops theories about human decision making process, it is worth noting that it involves, among others, uncertainty, automatically generated affect, living in the moment, as well as physiological or emotional states. In view of such unstable consumer behaviour, determining reasons behind particular choices has great importance for consumer science.

Consumer demand in relation to food and especially to animal origin products is increasingly shifting towards products that are safe, nutritious, produced through acceptable methods and of good palatability [Grunert 2006, Poławska *et al*. 2011a]. Consumers are becoming more aware about food of animal origin, especially meat, and are demanding livestock production processes to respect animal welfare, which avoid all forms of environmental contamination and do not misuse synthetic or chemical substances and additives that may pose a potential risk to health [Garcia-Torres *et al.* 2016, Olewnik-Mikołajewska *et al*. 2016].

One of the most important steps in the consumer research design is the identification of appropriate attributes and, subsequently, the specification of feasible attribute levels [Hair *et al*. 1999]. Several studies carried out using conjoint analysis for meat and meat products have used both intrinsic attributes (colour, tenderness, fat content) [Cunhal-Sendim *et al*. 1999] and extrinsic attributes (price, place of purchase, brand, quality label) [Gillespie *et al*. 1998, Ness *et al*. 1994], as well as a mixture of the two [Grunert 1997, Sánchez *et al*. 2000].Segmentation analysis in studies of animal origin products remains a very valuable tool when analysing decisions and declarations, however, it should be borne in mind that the consumers' choices may be changing over time. Nevertheless, their decision making process can certainly be observed using this method. Finding the mechanisms driving consumer behaviours is still an important question both from methodological and practical perspective. It is commonly used in food and beverage studies [Cox *et al.* 2008; García-Torres *et al.* 2016, Lima Filho *et al.* 2015, Mesías *et al.*, 2013, Mesías *et al.* 2009, Romano *et al.* 2015].

This article focuses on methodological aspects of cluster analysis, both theoretical and practical, in relation to food research, especially to conjoint analysis (CA). In the literature there are articles on statistical methods [Desarbo *et al.* 1995, Hartigan 1975, Bryan Orme 2001, Wagstaff *et al.* 2001], there is also a number of studies which use cluster analysis [Mesías *et al.* 2013, 2009, Miklavec *et al.* 2015], but the research does not include a study on methods of segmentation in food research with clear recommendations. The aim of this article is to fill this gap.

Material and methods

This research is based on a conjoint choice-based analysis (CBA), in which near 1000 respondents aged 21+ participated. The sample was selected to be representative of the polish population for region, age and gender. Intrinsic and extrinsic attributes and levels of animal origin products were selected after reviewing previous studies on consumer preferences for meat as well as by direct assessment of their importance by the research team. Special importance was given to colour evaluation as colour preference and purchasing decisions are closely related [Carpenter *et al.* 2001].

Consumer study was complimented by a survey in order to gather more detailed information about declared nutritional beliefs, food selection habits and preferences. The study was conducted in January 2013 using Computer Assisted Personal Interviewing methodology (CAPI - Sawtooth SSI Web CAPI). Consumer survey was linked with questionnaires to increase the complexity of data in the study.

The entire interview contained conjoint choice-based modules, socio-demographic questions and a questionnaire with possible responses on a scale 1 to 7 (where 1 means “I strongly disagree...” and 7 means “I strongly agree...”). Conjoint choice-based script was prepared using “full profile” option in Sawtooth SMRT software v. 4.22.

The product that had been selected for the conjoint choice-based study were eggs. In order to determine the importance of each product’s characteristics six attributes were selected: farming method, hen breed, nutrition claims, egg size, package size and price. Each attribute was assigned different levels – from 2 to 7, as shown in Table 1.

Table 1. Attributes and levels used in the conjoint survey design

Attribute	Levels
Farming method	organic / free range / barn eggs / from caged hens
Breed	no claim / traditional local hen breed
Nutrition claims	no health benefits claim / contains omega3 fatty acids / contains omega3 fatty acids with a positive influence on the cardiovascular system / higher level of A+E vitamins / higher level of A+E vitamins with a positive influence on the cardiovascular system
Size	S / M / L / XL
Package	6 eggs / 10 eggs
Price (per unit)	0.50PLN / 0.60PLN / 0.73PLN / 0.89PLN / 1.07PLN / 1.29PLN / 1.60PLN

In the conjoint study respondents had to decide "Which product would you choose?" – the intention to purchase the product was the most important fact. Each respondent saw similar screens (with 3 different products at a time) with all the attributes defined in accordance with the established levels (presented in Tab. 1) and had to choose one of them. He or she had to choose one of 3 different products a dozen times. It was a forced choice situation – answer "none of those" was not possible. A similar situation takes place in reality, when a person goes shopping, wants to buy eggs and has to choose from what is available on the shelf (not from the entire egg market, which is only a theoretical possibility).

Statistical analysis was performed in SPSS ver. 23.

Theoretical aspects

Conjoint analysis (CA) is one of the methods for examining consumer behaviour. It is popular in food research [Asioli, Næs *et al.* 2016, Cox *et al.* 2008, Garcia-Torres *et al.* 2016, Romano *et al*. 2015] stimuli were presented in the form of mock-up pictures of iced coffees varying in coffee type, production origin, calorie content and price, following an orthogonal design. One group of participants (n=101. CA can be used to identify the most important characteristics of the product in relation to the purchase process. It allows the researcher to obtain information about the product configuration features most desired by consumers, determines the impact of each of the product attributes on its overall usability to respondents and it also gives the researcher the ability to analyse a group of product attributes together.

Conjoint analysis method was created in the 1960s. and 1970s. [Green and Srinivasan 1978] and nowadays it is known as "full profile" version. This method is based on the evaluation and ranking of individual cards describing the offer. Then adaptive conjoint analysis (ACA) was developed [Johnson 2001], a method which uses rank levels, and the respondent indicates how important each of those attributes is. The most recent version of conjoint analysis is the conjoint choice-based (CBA) [Orme 2006. This method simulates real life situations, in which consumers face a limited choice. It is important that CBA can measure the main effects and interactions between them.

CBA used in this study uses Hierarchical Bayesian networks. As it is a discrete choice model, some scientists treat CBA as separate from classic CA analysis [Louviere *et al.* 2010]. Considering a similar purpose of classical conjoint and conjoint choice-based, in this article we treat them it just as different types of analysis that is designed to measure consumer behaviour.

K-means clustering algorithm. The aim of the K-means algorithm is to divide M-points in N-dimensions into K-clusters in order to minimize the within-cluster sum of squares. We seek "local" optima solutions so that no movement of a point from one cluster to another will reduce the within-cluster sum of squares. The algorithm requires as input a matrix of M-points in N-dimensions and a matrix of K-initial cluster centres in N-dimensions. The number of points in cluster L is denoted by NC(L). D(I, L) is the

Euclidean distance between point I and cluster L. The general procedure is to search for a K-partition with locally optimal within-cluster sum of squares by moving points from one cluster to another. Detailed statistical algorithm is described in literature [Hartigan 1975, Telgarsky and Vattani 2010].

Two-step cluster analysis. Two-step cluster analysis [Mooi and Sarstedt 2010] requires only one pass of data (which is important for very large data files) and it can produce solutions based on mixtures of continuous and categorical variables and for varying numbers of clusters. The clustering algorithm is based on a distance measure that gives the best results if all variables are independent, continuous variables have a normal distribution, and categorical variables have a multinomial distribution. This is seldom the case in practice, but the algorithm is thought to behave reasonably well when the assumptions are not met. Since cluster analysis does not involve hypothesis testing and calculation of observed significance levels other than for descriptive follow-up, it's perfectly acceptable to cluster data that may not meet the assumptions for best performance. There is only one drawback: the final solution may depend on the order of cases in the file. To minimize this effect, the cases should be arranged in random order.

Hierarchical clustering: Ward's method. In Ward's method [Anderberg 1973, Punj and Stewart 1983] the means for all variables are calculated for each cluster. Then, for each case, the squared Euclidean distance to the cluster means is calculated. These distances are summed for all of the cases. At each step, the two clusters that merge are those that result in the smallest increase in the overall sum of the squared within-cluster distances. The coefficient in the agglomeration schedule is the within-cluster sum of squares at that step, not the distance at which clusters are joined.

Results and discussin

The aim of this research was to identify the most optimal method of cluster analysis for analysing consumer decisions on food market. Among defined product attributes, price perception (attribute "price") and rearing system importance (as "farming method" attribute can be also understood as consumer awareness of animal welfare) were recognized as the most important aspects. Next, other product attributes and several socio-demographic factors (gender, education, financial status) were considered. Each time a one-way ANOVA was carried out to check if there are significant differences between clusters.

To maintain consistency between the survey declarative question and the conjoint method three questions from the survey were used as the basis for segmentation:

– how often do you buy organic eggs? (scale 1-7);
– information on the packaging is (/is not) very important to me; I need (/do not need) to know what the product contains. (scale 1-7);
– organic eggs are worse/better than non-organic eggs. (scale 1-7).

These questions were used in each analysis performed in this study.

K-means clustering algorithm

K-means clustering algorithm was used at first. Detailed results are shown in Tables 2, 3 and 4.

"Price" attribute, which was defined as most important, significantly varies in 6 from 7 options. Regarding mean relative importance, there are two clusters focused on price (Cluster 1 – RI – 59% and Cluster 3 – RI – 53%) whereas Cluster 4 does not perceive price as the only important egg attribute (RI – 39%). Last cluster (Cluster 2) has an intermediate result in this parameter (RI – 46%).

Differences in "farming method" significance are statistically important in all levels. It can be seen that segments that consider "price" as extremely important pay less attention to attributes related to animal welfare.

Table 2. K-means clustering algorithm: the part-worth utilities with relative importance of attributes for 4 identified clusters

Attribute	Attribute level	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Sig.
Farming method	mean relative importance (%)	15	27	19	35	
	from caged hens	-1.128	-1.792	-1.466	-2.162	0.000
	organic	0.240	0.751	0.519	1.125	0.000
	free range	0.905	1.422	1.126	1.770	0.000
	barn eggs	-0.017	-0.381	-0.178	-0.732	0.000
Breed	mean relative importance (%)	2	2	1	2	
	no claim	-0.124	-0.094	-0.055	-0.107	0.214*
	traditional local hen breed	0.124	0.094	0.055	0.107	0.214*
Nutrition and health claims	mean relative importance (%)	3	8	5	6	
	no health benefits claim	-0.257	-0.523	-0.422	-0.430	0.019
	contains omega 3 fatty acids	-0.209	-0.200	-0.096	-0.190	0.261*
	contains omega 3 fatty acids with a positive influence on the cardiovascular system	0.120	0.355	0.233	0.264	0.022
	higher level of A +E vitamins	0.130	-0.005	0.015	0.064	0.075*
	higher level of A + E vitamins with a positive influence on the cardiovascular system	0.216	0.374	0.269	0.292	0.296*
Size	mean relative importance (%)	15	12	15	15	
	S	-1.080	-0.798	-1.061	-0.968	0.069*
	M	-0.212	-0.126	-0.214	-0.034	0.005
	L	0.352	0.257	0.383	0.283	0.15*
	XL	0.939	0.666	0.892	0.718	0.014
Package	mean relative importance (%)	6	5	7	3	
	6 eggs	0.431	0.312	0.488	0.197	0.000
	10 eggs	-0.431	-0.312	-0.488	-0.197	0.000
Price (per unit)	mean relative importance (%)	59	46	53	39	
	0.50 PLN	3.670	2.345	3.152	1.883	0.000
	0.60 PLN	2.803	1.709	2.433	1.447	0.000
	0.73 PLN	1.465	1.047	1.368	0.904	0.000
	0.89 PLN	0.222	0.263	0.361	0.343	0.135*
	1.07 PLN	-1.201	-0.628	-1.064	-0.537	0.000
	1.29 PLN	-2.656	-1.726	-2.402	-1.475	0.000
	1.60 PLN	-4.303	-3.010	-3.847	-2.565	0.000

*Differences between groups not significant (p-value>0.05).

Table 3. K-means clustering algorithm: socio-demographic profile of the participants (%)

Item	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Sig.
Gender					0.007
male	55	50	45	43	
female	45	50	55	57	
Education					0.000
primary school & gymnasium	14	15	10	6	
basic vocational school	43	35	38	26	
high school	33	35	37	44	
graduate degree	9	13	13	23	
no response	1	2	2	1	
Financial status					0.011
we can afford all we need	6	11	8	13	
we live frugally and we satisfy all our needs	32	34	37	39	
we live very frugally to save for more important needs	30	32	33	32	
we have money only for the cheapest food and clothes	22	15	13	8	
we do not have enough money, even for the cheapest food and clothes	1	0	1	0	
no response	9	8	8	8	

*Differences between groups not significant (p-value>0.05).

Table 4. K-means clustering algorithm: number of respondents in clusters

Item	No of respondents	%
Cluster 1	301	30
Cluster 2	155	15
Cluster 3	305	30
Cluster 4	242	24
Total sample	1 003	100

Less important attributes gave mixed results: “packaging” has statistically important differences, “size” – 2 out of 4 levels differ, “nutrition and health claim” – 2 out of 5, while in “breed” differences between groups are not significant (p-value >0,05).

Upon considering selected socio-demographic data, clusters differ significantly in terms of gender and education as well as in financial status. It is worth noting that number of respondents in clusters established by K-mean clustering algorithm is rather balanced.

Two-step cluster analysis

The next method taken into consideration was two-step cluster analysis. As a result, three clusters were established.

The detailed results (Tab. 5, 6 and 7) show that “price” attribute has statistically significant differences among nearly all levels (6 out of 7). “Farming method” attribute is divided into two segments: in one of them this attribute does not play an important

Table 5. Two-step cluster analysis: the part-worth utilities with relative importance of attributes for four identified clusters

Attribute	Attribute level	Cluster 1	Cluster 2	Cluster 3	Sig.
Farming method	mean relative importance (%)	19	17	29	
	from caged hens	-1.411	-1.219	-1.937	0.000
	organic	0.325	0.315	0.966	0.000
	free range	1.197	1.009	1.487	0.001
	barn eggs	-0.110	-0.105	-0.516	0.000
Breed	mean relative importance (%)	2	1	2	
	no claim	-0.158	-0.064	-0.095	0.051*
	traditional local hen breed	0.158	0.064	0.095	0.051*
Nutrition and health claims	mean relative importance (%)	4	5	6	
	no health benefits claim	-0.357	-0.397	-0.398	0.873*
	contains omega 3 fatty acids	-0.209	-0.097	-0.209	0.084*
	contains omega 3 fatty acids with a positive influence on the cardiovascular system	0.131	0.227	0.260	0.198*
	higher level of A+E vitamins	0.184	-0.002	0.059	0.006
	higher level of A+E vitamins with a positive influence on the cardiovascular system	0.250	0.270	0.289	0.865*
Size	mean relative importance (%)	14	14	15	
	S	-1.037	-1.040	-0.962	0.577*
	M	-0.191	-0.226	-0.088	0.011
	L	0.338	0.358	0.305	0.516*
	XL	0.890	0.908	0.745	0.064*
Package	mean relative importance (%)	6	7	5	
	6 eggs	0.395	0.473	0.287	0.002
	10 eggs	-0.395	-0.473	-0.287	0.002
Price (per unit)	mean relative importance (%)	54	56	44	
	0.50	3.367	3.481	2.215	0.000
	0.60	2.518	2.599	1.752	0.000
	0.73	1.307	1.404	1.075	0.000
	0.89	0.183	0.297	0.346	0.079*
	1.07	-0.936	-1.173	-0.694	0.000
	1.29	-2.438	-2.554	-1.723	0.000
	1.60	-4.002	-4.054	-2.970	0.000

*Differences between groups not significant (p-value>0.05).

role (Cluster 1 – 19%, Cluster 2 – 17%), and in the other it is an important aspect (Cluster 3 – 29%). Regarding less important product attributes, it can be observed that “nutrition and health claims” are not varied between clusters (only 1 out of 5) the same as “size” (1 out of 4).

It is worth noting that the discussed algorithm generated segments that are not varied with respect to gender. Other socio-demographic factors have statistically significant differences, but when looking at segment sizes it can be stated that they are not balanced: Cluster 3 is more than twice the size of Cluster 1.

Table 6. Two-step cluster analysis: socio-demographic profile of the participants (%)

Item	Cluster 1	Cluster 2	Cluster 3	Sig.
Gender				0.517*
male	49	51	47	
female	51	49	53	
Education				0.000
primary school & gymnasium	13	13	8	
basic vocational school	41	43	30	
high school	34	33	41	
graduate degree	10	10	19	
no response	2	1	2	
Financial status				0.000
we can afford all we need	4	8	11	
we live frugally and we satisfy all our needs	29	33	40	
we live very frugally to save for more important needs	34	31	31	
we have money only for the cheapest food and clothes	20	18	11	
we do not have enough money, even for the cheapest food and clothes	2	1	0	
no response	11	9	7	

*Differences between groups not significant (p-value>0.05).

Table 7. 2-step cluster method: number of respondents in clusters

Item	No of respondents	%
Cluster 1	178	18
Cluster 2	361	36
Cluster 3	464	46
Total sample	1 003	100

Interpretation of segments that are not adequately differentiated (eg. "farming method" attribute, gender or segment sizes) may be impeded.

Hierarchical cluster analysis – Ward's method

Another segmentation algorithm that was considered was hierarchical cluster analysis – Ward's method. The analysis resulted in two versions: with three clusters and with four clusters. Due to editorial constraints Table 8 and Table 9 present detailed results for three established clusters. Table 10 shows the numbers of respondents in clusters in both analyses. Detailed version with four segments has not been posted here.

Hierarchical cluster analysis using Ward's method gives satisfying results in both versions: with three and four clusters. More specifically, analysis with three clusters gives statistically differentiated results in main product attributes: "price" (Cluster 1 – 54%, Cluster 2 – 57%, Cluster 3 – 40%) and "farming method" (Cluster 1 – 18%,

Table 8. Hierarchical cluster analysis: the part-worth utilities with relative importance of attributes for three identified clusters

Attribute	Attribute level	Cluster 1	Cluster 2	Cluster 3	Sig.
Farming method	mean relative importance (%)	18	17	33	
	from caged hens	-1.393	-1.182	-2.084	0.000
	organic	0.437	0.297	1.056	0.000
	free range	1.048	0.973	1.677	0.000
	barn eggs	-0.091	-0.087	-0.649	0.000
Breed	mean relative importance (%)	2	1	2	
	no claim	-0.108	-0.067	-0.107	0.396*
	traditional local hen breed	0.108	0.067	0.107	0.396*
Nutrition and health claims	mean relative importance (%)	5	4	7	
	no health benefits claim	-0.365	-0.298	-0.482	0.041
	contains omega 3 fatty acids	-0.145	-0.163	-0.195	0.678
	contains omega 3 fatty acids with a positive influence on the cardiovascular system	0.151	0.206	0.301	0.042
	higher level of A+E vitamins	0.096	0.045	0.026	0.323
	higher level of A+E vitamins with a positive influence on the cardiovascular system	0.263	0.211	0.349	0.110
Size	mean relative importance (%)	15	14	14	
	S	-1.123	-1.001	-0.901	0.041
	M	-0.234	-0.186	-0.076	0.006
	L	0.407	0.321	0.268	0.022
	XL	0.950	0.866	0.709	0.009
Package	mean relative importance (%)	7	7	4	
	6 eggs	0.477	0.442	0.233	0.000
	10 eggs	-0.477	-0.442	-0.233	0.000
Price (per unit)	mean relative importance (%)	54	57	40	
	0.50	3.414	3.426	1.946	0.000
	0.60	2.674	2.569	1.451	0.000
	0.73	1.444	1.399	0.909	0.000
	0.89	0.315	0.251	0.329	0.467*
	1.07	-1.137	-1.135	-0.527	0.000
	1.29	-2.589	-2.463	-1.495	0.000
	1.60	-4.122	-4.048	-2.612	0.000

*Differences between groups not significant (p-value>0.05).

Cluster 2 – 17%, Cluster 3 – 33%). Differences between segments are also statistically significant in less important characteristics such as "size" and "package" (Tab. 8). As for "nutrition claims" the version of analysis with three clusters gives satisfying results (Cluster 1 – 5%, Cluster 2 – 4%, Cluster 3 – 7%). Everything considered, socio-demographic data display statistically significant differences (Tab. 9).

The results are quite satisfactory for both analysis: with three and four generated clusters.

When comparing cluster sizes, it can be noticed that although they are quite balanced in four clusters (33%, 26%, 25%, 12%), they are far more steady (i.e. have more similar size) when there are three of them (33%, 26%, 37%).

Table 9. Hierarchical cluster analysis: socio-demographic profile of the participants (%) in three established clusters

Item	Cluster 1	Cluster 2	Cluster 3	Sig.
Gender				0.000
male	45	57	48	
female	55	43	52	
Education				0.010
primary school & gymnasium	11	14	9	
basic vocational school	36	47	28	
high school	38	30	42	
graduate degree	13	9	20	
no response	2	0	1	
Financial status				0.017
we can afford all we need	6	8	13	
we live frugally and we satisfy all our needs	38	30	37	
we live very frugally to save for more important needs	32	31	31	
we have money only for the cheapest food and clothes	16	19	11	
we do not have enough money, even for the cheapest food and clothes	1	2	0	
no response	7	10	8	

*Differences between groups not significant (p-value>0.05).

Table 10. Hierarchical cluster analysis: number of respondents in clusters

Item	3 cluster		4 cluster	
	no of respondents	%	no of respondents	%
Cluster 1	335	33	335	33
Cluster 2	263	26	263	26
Cluster 3	369	37	247	25
Cluster 4	-	-	122	12
Total sample	967	100	967	100

The aim of this study was to compare clustering methods and to select the most optimal one. At first glance one might conclude that all methods give equally good results. This is the moment to look both at the theoretical details of each method and the results obtained in defined key areas as well as in supporting ones: attributes defined as less important, socio-demographic data and also the size of individual segments. The criteria to be considered when comparing methods were the theoretical details of each method and the results obtained in defined key areas as well as in supporting ones: attributes defined as less important, socio-demographic data and also the size of individual segments.

Segmentation is not a purely scientific pursuit [Horn and Huang 2009] but the choice of segmentation method should not be random because the method itself has a significant impact on the results obtained.

In this paper, a conjoint choice-based study on eggs was examined. As main attributes “price” and “farming method” were chosen, other (“breed”, “nutrition claims”, “size”, and “package”) were defined as less important but were taken into consideration later on. Socio-demographic data (“gender”, “education”, “financial status”) and the size of each segment were of secondary importance.

Considering the results, it can be stated that the most optimal clustering method in this presented article was Hierarchical Cluster analysis - Ward’s method. It gives good, differentiated results in main goals, other attributes, socio-demographic data and segment’s sizes alike. It is also has strong theoretical background.

Considering results cannot be separated from analyzing p-value. It is worth to mention that American Statistical Association issued a statement about p-value interpretation and its influence on decision making process [ASA 2016]. P-value should always be perceived as one of parameters, but the researcher should not focus only on its result. This is not to say that p-value is not helpful in the data evaluation, but that common-sense analysis is a crucial element in choosing the method.

Conclussion

Regarding cluster analysis and summarizing carried out analyzes, the following steps are recommended:

– clear definition of the goals;
– employing varied methods;
– considering different numbers of clusters;
– analysis of the results with respect to main goals and subsequent characteristics;
– common-sense analysis of the results and final choice of the method.

When looking at data presented in this article, it can be stated that both segmentation versions (three and four clusters) of hierarchical cluster analysis with Ward’s method gave most satisfying results and are recommended for consideration when analysing consumers on food market.

Results indicate that hierarchical cluster analysis with Ward’s method is one of the methods which give good and stable results. It is recommended in consumer segmentation in food research.

It can be stated that cluster analysis may be a part of an in-depth study of consumer behaviour towards animal origin food products. Focusing not only on theoretical aspects of segmentation methods but also on defined goals can help choosing the optimal method of segmentation. Consumers are unstable in their choices and opinions, therefore analysing their behaviour is a difficult task but still worth consideration. Cluster analysis makes those consumer preferences and decisions more transparent and easier to explore.

REFERENCES

1. ANDERBERG M.R., 1973 – Cluster Analysis for Applications: Probability and Mathematical Statistics. Academic Press (19/09/2014).
2. Ariely D., 2008 – Predictably irrational: the hidden forces that shape our decisions. HarperCollins Publishing, USA.
3. ASA., 2016 – STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES. https://doi.org/10.1080/00031305.2016.1154108#.Vt2XIOaE2MN
4. ASIOLI D., NÆS T., ØVRUM A., ALMLI V.L., 2016 – Comparison of rating-based and choice-based conjoint analysis models. A case study based on preferences for iced coffee in Norway. ***Food Quality and Preference*** 48, 174-184.
5. CARPENTER C.E., CORNFORTH D.P., WHITTIER D., 2001 – Consumer preferences for beef color and packaging did not affect eating satisfaction. ***Meat Science*** 57, 359-363.
6. COX D.N., EVANS G., LEASE H.J., 2008 – Australian consumers' preferences for conventional and novel sources of long chain omega-3 fatty acids: A conjoint study. ***Food Quality and Preference*** 19(3), 306-314.
7. CUNHAL-SENDIM A., ALBIAC MURILLO J., DELFA BELENGUER R., LAHOZ CASTELLÓ F., 1999 – Quality Perception of Light Lamb. ***Carcass. Archive Zootechnic*** 48, 187-196.
8. DESARBO W.S., RAMASWAMY V., COHEN S.H., 1995 – Market segmentation with choice-based conjoint analysis. ***Marketing Letters*** 6(2), 137-147.
9. Engel J., 1968 – Consumer behavior. New York: Holt Rinehart and Winston.
10. FISHBEIN, AJZEN., 1975 – Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Retrieved January 24, 2017, from http://people.umass.edu/aizen/f&a1975.html
11. GARCIA-TORRES S., LÓPEZ-GAJARDO A., MESÍAS F.J., 2016 – Intensive vs. free-range organic beef. A preference study through consumer liking and conjoint analysis. ***Meat Science*** 114, 114-120.
12. GILLESPIE J., TAYLOR G., SCHUPP A., WIRTH F., 1998 – Opinions of professional buyers to-ward a new, alternative red meat: Ostrich. ***Agribusiness*** 14, 247-256.
13. Green P., Srinivasan V., 1978 – Conjoint analysis in consumer research: Issues and outlook. ***Journal of Consumer Research*** 5, 103-123.
14. GRUNERT K.G., 1997 – What's in a steak? A cross-cultural study on the quality perception of beef. ***Food Quality and Preference*** 8(3), 157-173.
15. GRUNERT K.G., 2006 – Future trends and consumer lifestyles with regard to meat consumption. ***Meat Science*** 74(1), 149-160.
16. HAIR J.F., ANDERSON R.E., TATHAM R.L., BLACK W.C., 1999 – Multivariate Data Analysis. NJ: Prentice Hall. 5th Edition.
17. HARTIGAN J.A., 1975 – Clustering Algorithms. New York: Wiley.
18. HORN B., HUANG W., 2009 – Comparison of Segmentation Approaches. Retrieved from http://hbanaszak.mjr.uw.edu.pl/TempTxt/HornHuang_2009_Comparison of Segmentation Approaches.pdf
19. JOHNSON R.M., 2001 – History of ACA. Proceedings of the Sawtooth Software Conference, 205–212. Victoria, BC, Canada.
20. KAHNEMAN D., TVERSKY A., 1979 – Prospect Theory: An Analysis of Decision under Risk. ***Econometrica*** 47(2), 263-291.
21. LIMA FILHO, T., DELLA LUCIA, S. M., LIMA, R. M., & MINIM, V. P. R. (2015). Conjoint analysis as a tool to identify improvements in the packaging for irradiated strawberries. ***Food Research Internationa**l* 72, 126-132.

22. LOUVIERE J. J., FLYNN T. N., CARSON R.T., 2010 – Discrete choice experiments are not conjoint analysis. ***Journal of Choice Modelling*** 3(3), 57-72.
23. MESÍAS F. J., GASPAR P., PULIDO Á. F., ESCRIBANO M., PULIDO F., 2009 – Consumers' preferences for Iberian dry-cured ham and the influence of mast feeding: An application of conjoint analysis in Spain. ***Meat Science*** 83(4), 684-690.
24. MESÍAS F.J., PULIDO F., ESCRIBANO M., GASPAR P., PULIDO Á.F., ESCRIBANO A., RODRÍGUEZ-LEDESMA A., 2013 – Evaluation of New Packaging Formats for Dry-Cured Meat Products Using Conjoint Analysis: An Application to Dry-Cured Iberian Ham. ***Journal of Sensory Studies*** 28(3), 238-247.
25. MIKLAVEC K., PRAVST I., GRUNERT K.G., KLOPČIČ M., POHAR J., 2015 – The influence of health claims and nutritional composition on consumers' yoghurt preferences. ***Food Quality and Preference***, 43, 26-33.
26. MOOI E., SARSTEDT M., 2010 – Introduction to Market Research. In: A Concise Guide to Market Research (pp. 1-9). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12541-6_1
27. NESS MITCHELL R., GERHARDY H., 1994 – Consumer Preferences for Quality and Freshness Attributes of Eggs. ***British Food Journal*** 96(3), 26-34.
28. OLEWNIK-MIKOŁAJEWSKA A., GUZEK D., GŁĄBSKA D., SAJDAKOWSKA M., GUTKOWSKA K., 2016 – Fodder enrichment and sustaining animal well-being as methods of improving quality of animal-derived food products, in the aspect of consumer perception and acceptance. ***Animal Science Papers and Reports*** 34(4), 361-372.
29. ORME B., 2001 – Sawtooth Software Research Paper Series: Hierarchical Bayes: Why All the Attention? 98382(360).
30. ORME B., 2006 – Getting started with conjoint analysis. Madison, WI: Research Publishers LLC, USA.
31. POŁAWSKA E., MARCHEWKA J., COOPER R.G., SARTOWSKA K., POMIANOWSKI J., JÓŹWIK A., STRZAŁKOWSKA N., HORBAŃCZUK J.O., 2011 – The ostrich meat – an updated review. II. Nutritive value. ***Animal Science Papers and Reports*** 29(2), 89-98.
32. PUNJ G., STEWART D.W., 1983 – Cluster Analysis in Marketing Research: Review and Suggestions for Application. ***Journal of Marketing Research*** 20(2), 134.
33. ROMANO K.R., ROSENTHAL A., DELIZA R., 2015 – How do Brazilian consumers perceive a non-traditional and innovative fruit juice? An approach looking at the packaging. ***Food Research International*** 74, 123-130.
34. SÁNCHEZ M., GOÑI C., MARAÑÓN I., MARTÍN S., 2000 – Diferencias en las preferenciasentre los consumidores de carne de vacuno etiquetada y no etiquetada (Preference differences between consumers of bovine meat with and without label). In Spanish with English summary. ***ITEA-Informacion Tecnica Economica Agraria*** 96A (1), 40-55.
35. TELGARSKY M., VATTANI A., 2010 – Hartigan's Method: k-means Clustering without Voronoi. Proceedings of the 13th International Conference on Artificial Intelligence and Statisctics (AISTATS), Chia Lagune Resort, Sardinia, Italy. 9, 820-827.
36. WAGSTAFF K., CARDIE C., ROGERS S., SCHROEDL S., 2001 – Constrained K-means Clustering with Background Knowledge, 577-584.