

Determination of Consumer Perceptions of Eco-Friendly Food Products Using Unsupervised Machine Learning*

Çevre Dostu Gıda Ürünlerine Yönelik Tüketici Algısının Denetimsiz Makine Öğrenmesi Kullanılarak Belirlenmesi


Yusuf ÇAKMAKÇI^{1*}, Harun HURMA², Cihan ÇAKMAKÇI³


Abstract

This study aims to comparatively determine the consumer perception of food products marketed under ecologically friendly concepts (organic food, good agriculture, and natural production) and food sold directly by farmers, conventional food, and farmer cooperative branded food. For this purpose, a face-to-face survey was conducted with 171 identified consumers. R program was used to perform all of the analyses. Machine learning methods such as Logistic Regression (LR), Correspondence Analysis (CA), and Principal Component Analysis (PCA) are used for determining consumer perception from obtained data. Descriptive statistics results showed that 51.5 percent of those polled were male and 48.5 percent were female. It found that the mean age of the consumers was joined to the survey was 36.4. According to the LR findings, consumer socioeconomic characteristics have a considerable impact on the purchase of various foods (such as organic labeled foods, good agricultural practices foods, producer cooperative branded foods, etc.). It has been discovered as the result of the PCA, people perceived organic branded food and good agricultural practices foods, which are healthier, more flavorful, and more trustworthy than other food. however, it has been discovered that they believe the costs of these types of food are expensive and that they are difficult to obtain. On the other hand, they perceive the pricing of farmer cooperative branded foods and food sold directly by the farmer to be lower. Furthermore, it was observed in CA findings that there was a correlation between these various food groups and purchase locations. While products sold directly by farmers are mostly purchased from public markets, they prefer grocery stores and supermarkets when purchasing foods with good agricultural practices and natural labeled (from the markets). When seen from this perspective, it is possible to conclude that ecologically friendly food and other food products are regarded differently by customers based on product characteristics. The use of marketing techniques that create a positive perspective in terms of affordability and accessibility and the development of policies and production techniques that boost consumers' current perceptions of these items are considered will promote the consumption of these products.

Keywords: Eco-friendly foods, Consumer perceptions, Machine learning, Product appreciation, Purchasing decision

^{1*}**Sorumlu Yazar/Corresponding Author:** Yusuf Çakmakçı, Tekirdağ Namık Kemal University, Faculty of Agriculture, Department of Agricultural Economics, 59000 Tekirdağ, Türkiye. E-mail: cak76yusuf@gmail.com  ORCID: 0000-0002-5136-9102

²Harun Hurma, Tekirdağ Namık Kemal University, Faculty of Agriculture, Department of Agricultural Economics, 59000 Tekirdağ, Türkiye. E-mail: h.hurma@gmail.com  ORCID: 0000-0003-1845-3940

³Cihan Çakmakçı, Van Yüzüncü Yıl University, Faculty of Agriculture, Department of Agricultural Biotechnology, Animal Biotechnology Section, 65080 Van, Türkiye. E-mail: cakmakcicihan@gmail.com  ORCID: 0000-0001-6512-9268.

Atf: Çakmakçı, Y., Hurma, H., Çakmakçı, C. (2024). Çevre dostu gıda ürünlerine yönelik tüketici algısının denetimsiz makine öğrenmesi kullanılarak belirlenmesi. *Tekirdağ Ziraat Fakültesi Dergisi*, 21(3): 634-647.

Citation: Çakmakçı, Y., Hurma, H., Çakmakçı, C. (2024). Determination of consumer perceptions of eco-friendly food products using unsupervised machine learning. *Journal of Tekirdag Agricultural Faculty*, 21(3): 634-647.

*This study was summarized from the Yusuf Çakmakçı MSc. thesis.

©Bu çalışma Tekirdağ Namık Kemal Üniversitesi tarafından Creative Commons Lisansı (<https://creativecommons.org/licenses/by-nc/4.0/>) kapsamında yayımlanmıştır. Tekirdağ 2024

Öz

Bu çalışma, çevre dostu konseptler (organik gıda, iyi tarım ve doğal üretim) altında pazarlanan gıda ürünleri ile doğrudan çiftçiler tarafından satılan gıdalar, konvansiyonel gıdalar ve çiftçi kooperatifi markalı gıdalara yönelik tüketici algısını karşılaştırmalı olarak belirlemeyi amaçlamaktadır. Bu amaçla belirlenen 171 tüketici ile yüz yüze anket yapılmıştır. Tüm analizleri gerçekleştirmek için R programı kullanılmıştır. Elde edilen verilerden tüketici algısının belirlenmesi için Lojistik Regresyon (LR), Yazışma Analizi (CA) ve Temel Bileşenler Analizi (PCA) gibi makine öğrenmesi yöntemleri kullanılmıştır. Tanımlayıcı istatistik sonuçları, ankete katılanların yüzde 51,5'inin erkek ve yüzde 48,5'inin kadın olduğunu gösterdi. Araştırmaya katılan tüketicilerin yaş ortalamasının 36,4 olduğu tespit edildi. LR bulgularına göre, tüketici sosyoekonomik özelliklerinin çeşitli gıdaların (organik etiketli gıdalar, iyi tarım uygulamaları gıdaları, üretici kooperatifi markalı gıdalar vb.) satın alınmasında önemli bir etkisi vardır. PCA sonucunda insanların organik markalı gıdaları ve iyi tarım uygulamaları gıdalarını diğer gıdalardan daha sağlıklı, daha lezzetli ve daha güvenilir olarak algıladıkları saptanmıştır. Ancak bu tür gıdaların maliyetinin pahalı ve elde edilmesinin zor olduğuna inandıkları tespit edilmiştir. Öte yandan, çiftçi kooperatifi markalı gıdaların ve doğrudan çiftçi tarafından satılan gıdaların fiyatını daha düşük olarak algılamaktadırlar. Ayrıca CA bulgularında çeşitli gıda grupları ile satın alma yerleri arasında bağlantı olduğu belirlenmiştir. Çiftçiler tarafından doğrudan satılan ürünler daha çok halk pazarlarından satın alınır iken, iyi tarım uygulamalı gıdalar ve doğal etiketli (marketlerden)gıdalar satın alınırken, daha çok bakkal ve süpermarketleri tercih etmektedirler. Bu açıdan bakıldığında, çevre dostu gıda ve diğer gıda ürünlerinin, ürün özelliklerine göre müşteriler tarafından farklı değerlendirildiği sonucuna varmak mümkündür. Satın alınabilirlik ve erişilebilirlik açısından olumlu bir bakış açısı oluşturan pazarlama tekniklerinin kullanılması ve tüketicilerin bu ürünlere yönelik mevcut algılarını yükselten politikalar ve üretim tekniklerinin geliştirilmesinin bu ürünlerin tüketimini artıracakı düşünülmektedir.

Anahtar Kelimeler: Çevre dostu gıdalar, Tüketici algıları, Makine öğrenimi, Ürün takdiri, Satın alma kararı

1. Introduction

Consumer concerns about healthy foods and the environment have risen dramatically in recent decades. These concerns have encouraged them to be more conscious of the environment while purchasing products such as technological devices and foods in particular (Petrescu et al., 2020). Consumer attitudes are rapidly integrated into the production of all goods and foods. This rapid integration towards eco-friendly food, organic food, good agricultural practices etc. It has brought along various new production applications as well as new marketing concepts, such as what's driving consumers' preferences while purchasing foods (Kashif et al., 2020).

Consumers' preferences for food products are affected by socioeconomic factors such as gender, education, income as well as, a number of product-related factors such as pricing of products (Echeverría et al., 2022), package of product (Nascimento et al., 2022), location of their residence (Yılmaz et al., 2009), taste, safety, freshness (Oraman et al., 2011), awareness of consumers, availability and accessibility of the product, product appreciation, product quality perception and sensory characteristics, production method (organic, conventional), food safety (antibiotics and hormone usage) and its effects on human beings (Ramya and Ali, 2016; Argemí-Armengol et al., 2019; Çakmakçı and Hurma, 2021). As well as consumer-related factors such as geographical region of production, socioeconomic structure, age, education level, gender, culture, beliefs, and consumption habits (De-Groote and Kimenju, 2008; Ramya and Ali, 2016). Food consumption patterns of consumers, on the other hand, are constantly changing. These changes, however, are linked not only to socioeconomic and cultural developments that influence the whole society, but also to more diverse consumer groups with distinct lifestyles (Ripoll et al., 2018). Changes in consumer expectations for agricultural goods, as well as ethical concerns regarding production procedures, are driving a rise in demand for information about these products' production processes. For this reason, product labels are used to provide information about food products when deciding what to purchase and play a role in purchasing decision-making processes (Klopčič et al., 2012). In recent years, local food production systems have received more attention as a potential alternative to the production patterns of the industrialized global food system. This is due to environmental concerns and the idea that they may make the food production chain more sustainable (Aprile et al., 2016). Local products are often associated with qualitative characteristics such as freshness, flavor, nutritional value, health, and safety (Petrescu et al., 2020). Aside from flavor or safety, some consumer groups may be motivated to purchase products with organic or ecological labels because they believe they are healthier, safer, and free of treatments and nutritional supplement residues such as pesticides, hormones, antibiotics, and chemical fertilizers (Bougherara and Combris, 2009). On the other hand, despite consumers' positive attitudes toward products bearing the eco-friendly label and characterized by their sustainability features, issues such as a lack of market knowledge about the products, the lack of distinctiveness of such products in comparison to conventional products, and the limited availability of such products prevent this mindset from being translated into actual purchase behaviors (Eldesouky et al., 2020). Previous studies have attempted to evaluate consumer perceptions of food products such as preference, trust, quality, and flavor by focusing on a single product or one or two products' characteristics (Nam et al., 2020; Petrescu et al., 2020). The categorization of potential consumers and the acquisition of accurate information about their characteristics are important for predicting their behavior toward food products (Vindigni et al., 2022). Machine learning methods have become widely used approach in such studies (Shiokawa et al., 2018). For cross-selling strategies, consumer segmentation, and qualitative data analysis, this method works perfectly. Principal component analysis (PCA) is one of widely used unsupervised machine learning technique for exploratory data analysis, dimensionality reduction, and data de-noising (IBM, 2022). Besides, Correspondence Analysis (CA) is a statistical method that provides visualize the relation between attributes (Greenacre, 2015), such as the relationship of purchased products and places. On the other hand, Logistic Regression (LR) aims to measure the relationship between a categorical dependent variable and one or more independent variables (continuous or categorical) by drawing the probability scores of dependent variables (Ingwersen et al., 2023).

This study focuses on multiple food products and examines their relative positions based on different attributes. In that respect, the purpose of this study was to identify how consumers perceive food products with different labels, such as conventional, organic, natural, directly from the farmers, good agricultural products, and cooperative labeled, taking into consideration factors such as price, taste, availability, advertising, health, and general trust. Furthermore, determine the impact of consumers' socioeconomic characteristics on their purchasing of foods with different labels.

The overall objective of this study was to show that unsupervised machine learning approaches provide a novel strategy for understanding consumers' positioning, purchase locations, and purchasing intentions of foods with different labels. To achieve this, the primary purpose of this research was to determine consumers' attitudes using unsupervised machine learning approach, Principal Component Analysis (PCA), Logistic Regression and Correspondence Analysis (CA).

2. Materials and Methods

2.1. Material and data collection

The data for the study was obtained from a face-to-face survey of 171 consumers that was conducted in 16 major neighborhoods of the Süleymanpaşa district in Tekirdağ province between December 2019 and January 2020.

Size of the sample was estimated using the proportional sampling approach (Eq. 1) and the error margin (7.5 percent) from the population of which the main population is known in order to better reflect the main population of the data (number of surveys) gathered in the study (Newbold, 1995) n : Number of samples, N : Main mass, p : Number of people purchasing specified food products q : Number of people who did not purchase the specified food products, d : margin of error for p (0.075), t : significance level (1.96 for t table 0.5).

$$n = N * t^2 * p * q / (N - 1) * d^2 + t^2 * p * q \quad \text{Eq (1)}$$

2.2. Method

In this study, six different products, including conventional, organic, natural, directly from the farmers, good agricultural products, and cooperative-labeled, were evaluated. The consumers were asked to rate these products based on seven characteristics, including price, taste, availability, advertising, health, and general trust, on a scale ranging from 1 to 7 (1: the product is not good at all, 7: the product is very good).

2.2.1. Consumer perception of food products based on product characteristics

A technique of unsupervised machine learning PCA (principal component analysis) was employed to determine consumer perceptions of food products positioning based on product characteristics. PCA is one of the widely used multivariate statistical approach, which can be applied to almost every scientific field (Abdi and Williams, 2010). PCA, which benefits from covariance or correlation seeks to identify important characteristics of variability and derive inferences from linear relationships between variables matrices (Jolliffe and Cadima, 2016). PCA simplifies identification by reducing the data to a small number of synthetic variables rather than all the original variables (Husson et al., 2011). The aim of this study was to determine the relationships between products and product attributes in the same data table by assessing them together using PCA. Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) sample adequacy measure were calculated using the "psych" R package (Revelle, 2019) to assess the suitability of the data set for PCA. PCA was implemented using "FactoMineR" (Husson et al., 2011). The results of PCA were visualized using "Factoextra" R package (Kassambara and Mundt, 2020).

2.2.2. Relationship between different labeled foods and purchasing places

Following the PCA, Correspondence Analysis (CA) was used to determine the relationship between different labeled foods and purchasing locations. CA is a special case of a generalized principal component analysis used to analyze qualitative data by using the contingency of two nominal variables (Abdi and Williams, 2010). CA is a multivariate analysis method that uses a two-dimensional diagram to explain the associations between the categories of two nominal variables (Di Franco, 2016) and does not require any assumptions about the data distribution (Sourial et al., 2010). The number of dimensions to use to explain the variability in a dataset is determined using similar principles as in principal component analysis. Common criteria include keeping the dimensions that account for more than 70% of the variation or have an eigenvalue greater than 1 (Sourial et al., 2010). Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) sample adequacy measure were calculated using the "psych" R package (Revelle, 2019) to assess the suitability of the data set for PCA. PCA was implemented using "FactoMineR" (Husson et al., 2011). The results of PCA were visualized using "Factoextra" R package (Kassambara and Mundt, 2020).

2.2.3. The impact of consumer personal characteristics purchasing intentions for foods with different labels

Another way to assess consumers' (response) purchasing intentions for products with different labels is to understand the consumers characteristics that may influence them (De-Groote and Kimenju, 2008). Logistic Regression (LR) is the most widely used machine learning method for analyzing such data. LR seeks to determine whether there is a relationship between

a dependent variable (Y) associated with the occurrence or non-occurrence of an event binary and one or more categorical or continuous independent variables (Díaz-Pérez et al., 2019). The response variables are binary type data that represent consumers' purchasing intentions for products with different labels. The response variables Natural, Organic, Good Agriculture, Local and Cooperatives are factors with two levels, and the first level, "no", corresponds to not buying (zero) and the second level, "yes", to buying (1). Stepwise logistic regression models for binary response data were implemented using the logit link function (Fox and Weisberg, 2019) to determine the variables that should be used in the final model. The best subset of independent variables among all subsets was determined by applying stepwise regression to all independent variables. For model selection, the Akaike Information Criterion (AIC) was used as measure of fit. Likelihood ratio (LR) chi-square test was used to test the hypothesis.

Collinearity, which occurs when the data contains highly associated predictor variables, is a critical issue in regression analysis. The appropriateness of the estimated regression coefficients is determined by strong linear correlations among the variables in a linear model (Fox and Weisberg, 2019). The variance inflation factor (VIF) is a useful indicator for collinearity diagnostics (Fox and Monette, 1992). However, because the variance inflation factor cannot be directly applied to sets of related predictors with the degree of freedom higher than 1, Fox and Monette (1992) suggest to use the generalized variance inflation factor ($\sqrt{GVIF^{(1/(2 \cdot Df))}}$), which can be applicable to sets of related predictors. The $\sqrt{GVIF^{(1/(2 \cdot Df))}}$ is equivalent to reporting the square root of the VIF for an individual coefficient (Fox and Monette, 1992). As a general rule, the $\sqrt{GVIF^{(1/(2 \cdot Df))}}$ values greater than 2 indicate strong collinearity. Collinearities were checked by using the "car" R package (Fox and Weisberg, 2019). All analysis were performed using R programming language, R version 4.0.2 (R Core Team, 2020).

3. Results and Discussion

3.1. Participants' socioeconomic characteristics

Socio-economic and demographic data of the consumers are given in *Table 1*. A 51.5% of the consumers are male and 48.5% are female. 56.1% of consumers are "36 years old or younger," while 43.9% are "over the age of 36". 36.2% have "graduate and postgraduate" education, while 46.8% have "high school and undergraduate" education, and 17.0% have "pre-high school" education.

Table 1. Consumer social and demographic characteristics (%)

Gender	%	Age	%
Female	48.5	36 years and below	56.1
Male	51.5	Over 36	43.9
Education		Income*	
Graduate and Post Graduate	36.2	0-2900TL	36.3
High School and Under Graduate	46.8	2901-4500TL	35.1
Pre-High School	17.0	4501TL and Higher	28.5

3.2. Logistic Regression (LR)

Stepwise regression was applied for all independent variables to obtain the optimal subset of independent variables that explain the variations in dependent variables. The results indicated that the final models that explain dependent variables were significantly different from null models (*Table 2*).

House ownership was the only factor that had a significant effect on the purchasing of natural labeled products ($P < 0.05$). Gender, habits, and age of consumers all had a significant effect on their purchase intentions for organic labeled products ($P < 0.05$). Additionally, habits significantly effect the purchase of products with good agriculture labels, while age was the only factor that affected the purchase of Direct sells by farmers ($P < 0.05$). On the other hand, both marital status ($P < 0.05$) and family size ($P < 0.001$) had a significant effect on the purchase of Farmer Cooperatives branded foods (*Table 3*).

The Variance Inflation Factor (VIF) for each independent variable was used to check for multi-collinearity (*Table 4*). It is a measure of multicollinearity among several regression variables. The greater the value of VIF, the greater the connection between this variable and the others. When the value of VIF is $1 \leq VIF \leq 5$, the variables are moderately related. The VIF difficulty value runs from 5 to 10, suggesting that the variables are highly related. If VIF is between 5 and 10, the predictors in the regression model will be multicollinear, and VIF larger than 10

indicates that the regression coefficients are inadequately calculated due to multicollinearity. Table 4 shows that there are no issues with multicollinearity between the independent variables ($VIF < 5$).

Table 2. Hypothesis test for final model vs null model

Dependent Variables	Model	Resid. Df	Resid. Dev	Pr(>Chi)
Natural labeled (from market)	Null Model	167	232.87	0.024*
	Final Model	165	225.44	
Organic labeled	Null Model	167	230.01	0.002***
	Final Model	159	206.01	
Good Agriculture	Null Model	167	226.76	0.010***
	Final Model	165	217.62	
Direct sells by farmers	Null Model	167	222.28	0.048*
	Final Model	166	218.39	
Farmer Cooperatives branded foods	Null Model	167	210.99	0.000***
	Final Model	164	190.37	

Table 3. Test of individual parameters and pseudo-R squared for final model

Dependent Variables	Exploratory variables	LR Chisq	Df	Pr(>Chisq)	AIC	McFadden's Pseudo R2
Natural labeled (from market)	House ownership	4.97	1	0.026*	2231	0.032
	Age	3.77	1	0.052*		
	Gender	4.74	1	0.030**		
Organic labeled	Environmental consumption level	7.74	2	0.021**	2224	0.104
	Children	2.50	1	0.114		
	Age	4.63	1	0.031**		
	Occupation	6.04	3	0.110		
Good Agriculture	Environmental consumption level	9.15	2	0.010***	2224	0.040
Direct sells by farmers	Age	3.90	1	0.048**	2222	0.018
Farmer Cooperatives branded foods	Marital status	5.19	1	0.023**	1198	0.098
	Family size	18.37	2	0.000***		

* $P < 0,10$, ** $P < 0,05$, *** $P < 0,01$

Table 4. Variance inflation factors to check multicollinearity

Products	Exploratory variables	GVIF	Df	GVIF ^{^(1/(2*Df))}
Natural labeled (from market)	House ownership	1.100	1	1.049
	Age	1.100	1	1.049
	Gender	1.126	1	1.061
Organic labeled	Environmental consumption level	1.141	2	1.033
	Children	1.096	1	1.047
	Age	1.100	1	1.049
	Occupation	1.198	3	1.031
Farmer Cooperatives branded foods	Marital status	1.073	1	1.036
	Family size	1.073	2	1.018

The fact that the odds are 52% higher for house owners than for renters indicates that renters are less likely to purchase products with natural labels compared to homeowners (OR = 0.48, 95% CI [0.25, 0.92]), P=0.028) (Table 5). In terms of purchasing habits, consumers in the low-eco-friendly category are less likely to buy organic-labeled products than consumers in the high-eco-friendly category (OR = 0.24, 95% CI [0.08, 0.67]), P=0.008). Similarly, male consumers were 54% less likely to buy organic-labeled products than female consumers (OR = 0.46, 95% CI [0.23, 0.93]), P=0.032). Those under the age of 37 were more likely to purchase organic-labeled products than those older than 36 (OR = 2.11, 95% CI [1.07, 4.28]), P=0.034). Large families are less likely to purchase Farmer Cooperatives branded foods than those small families (P<0.05). Marital status also significantly influenced consumers' preferences for products with Farmer Cooperatives branded foods. Married consumers were more likely to purchase products Farmer Cooperatives branded foods than single consumers (OR = 0.44, 95% CI [0.21, 0.89]), P=0.024). In terms of habits, consumers in the high-eco-friendly category were more likely to purchase products with good agriculture labels(P<0.05) than those in the other categories (OR = 0.2, 95% CI [0.05, 0.58]), P=0.006). In addition, younger consumers tend to buy direct sells by farmers more than other consumers (OR = 0.53, 95% CI [0.27, 1.00]), P=0.051).

Table 5. Coefficients, and exponentiated coefficients as odds-ratios (OR) and confidence intervals of OR

	Variables	Estimate	Std. Error	z value	Pr (> z)	OR Estimate	OR 2.5%	OR 97.5%
Natural labeled (from market)	(Intercept)	-0.074	0.255	-0.292	0.770	0.928	0.561	1.530
	House_ownershipRental	-0.724	0.329	-2.201	0.028	0.485	0.252	0.91
	AgeAge_Under_37	0.625	0.325	1.923	0.055	1.869	0.994	3.568
Organic labeled	(Intercept)	0.552	0.415	1.329	0.184	1.736	0.776	3.989
	GenderMale	-0.766	0.357	-2.148	0.032	0.465	0.228	0.927
	low environmental consumption	-1.427	0.537	-2.657	0.008	0.240	0.080	0.670
	medium environmental consumption	-0.548	0.373	-1.469	0.142	0.578	0.275	1.194
	ChildrenYes	0.622	0.400	1.553	0.120	1.862	0.863	4.186
	AgeAge_Under_37	0.749	0.352	2.125	0.034	2.114	1.069	4.275
	OccupationOwn_busines	-0.722	0.616	-1.172	0.241	0.486	0.140	1.602
Good Agriculture	OccupationPrivate	-0.120	0.419	-0.285	0.776	0.887	0.388	2.024
	OccupationPublic	0.777	0.540	1.438	0.150	2.175	0.771	6.507
	(Intercept)	-0.030	0.244	-0.122	0.903	0.971	0.599	1.570
Direct sells by farmers	low environmental consumption	-1.628	0.598	-2.724	0.006	0.196	0.053	0.582
	medium environmental consumption	-0.343	0.338	-1.014	0.310	0.710	0.364	1.375
Farmer Cooperatives branded foods	(Intercept)	0.879	0.254	3.467	0.001	2.409	1.487	4.040
	AgeAge_Under_37	-0.642	0.329	-1.953	0.051	0.526	0.274	0.995
	(Intercept)	2.225	0.479	4.650	0.000	9.255	3.870	25.686
Farmer Cooperatives branded foods	Marital_statusSingle	-0.828	0.368	-2.252	0.024	0.437	0.210	0.891
	Family_size3-4	-1.210	0.475	-2.551	0.011	0.298	0.110	0.720
	Family_size4<	-2.469	0.621	-3.979	0.000	0.085	0.024	0.273

3.3. Principal Component analysis

Bartlett's test of sphericity (X²=40.987, P<0.01) and Kaiser-Meyer-Olkin (KMO) measure of sample adequacy (0.72) were calculated to test the validity of PCA in the dataset. The first two PCA components account for 98.9% of the variation in the dataset. This percentage is particularly high and thus the first plane perfectly represents the dataset variability. The parallel analysis showed that only the first component should be kept (Figure 1a). The first

component with an eigenvalue bigger than one accounted for 83.9% of the total variation (*Figure 1b*). The first two dimensions were taken into account in order to visualize the dataset on a two-dimensional plane.

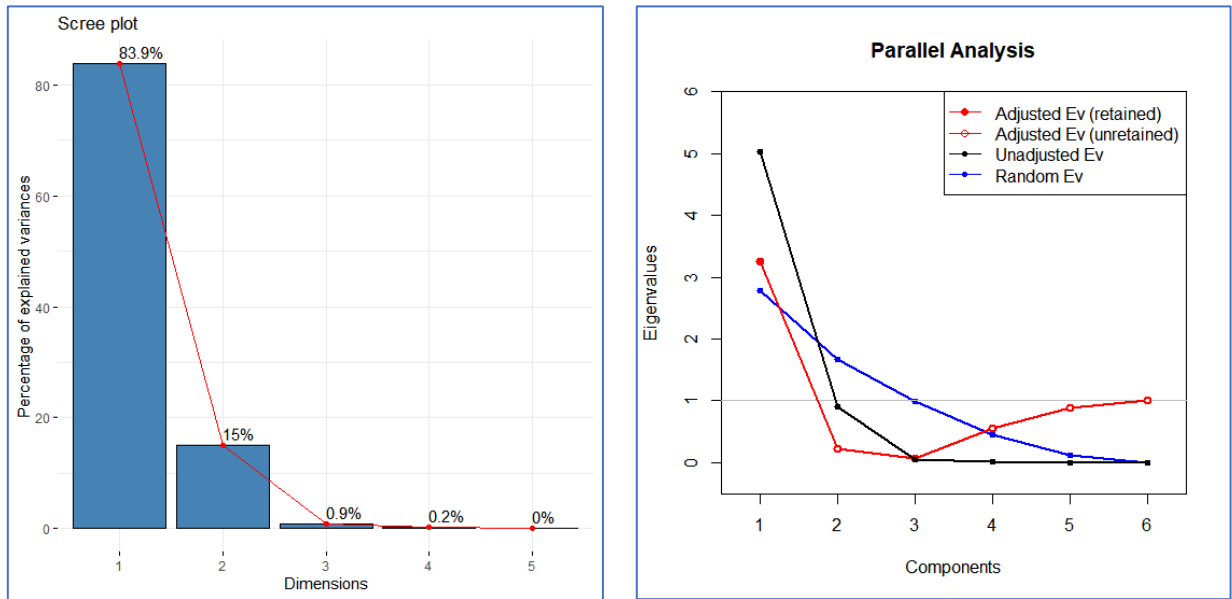


Figure 1. a: The proportion of variations that each principal component explains. b; Plot of parallel analysis

The positions of the consumers along with the product features that consumers consider when purchasing a product are shown on a two-dimensional plane in *Figure 2*. The variables taste, health, and trust are highly correlated with the first dimension (a respective correlation of 0.93, 0.95, and 0.99, $P < 0.01$) whereas the variables accessibility and advertisement were highly correlated with the first dimension ($r = -0.99$ and $r = -0.96$, $P < 0.01$). On the other hand, price was closely linked with the second dimension ($r = 0.88$, $P < 0.05$). Although their link was unclear, consumers who perceived the products as healthy, reliable, and tasty also considered the price more reasonable. Consumers believed that the prices of products labelled as Farmer Cooperatives branded foods or Direct sells by farmers were reasonable, but their promotion was insufficient. Although consumers found conventional food products more accessible, they did not consider them as reliable, tasty, or healthy. Consumers believed that advertisements for products with natural labeled (from markets) were adequate, but that their prices were high. Good agriculture labelled foods were positioned close to the origin of the plane. Consumers regarded these products as tasty, healthy, and reliable, but lacking in terms of accessibility.

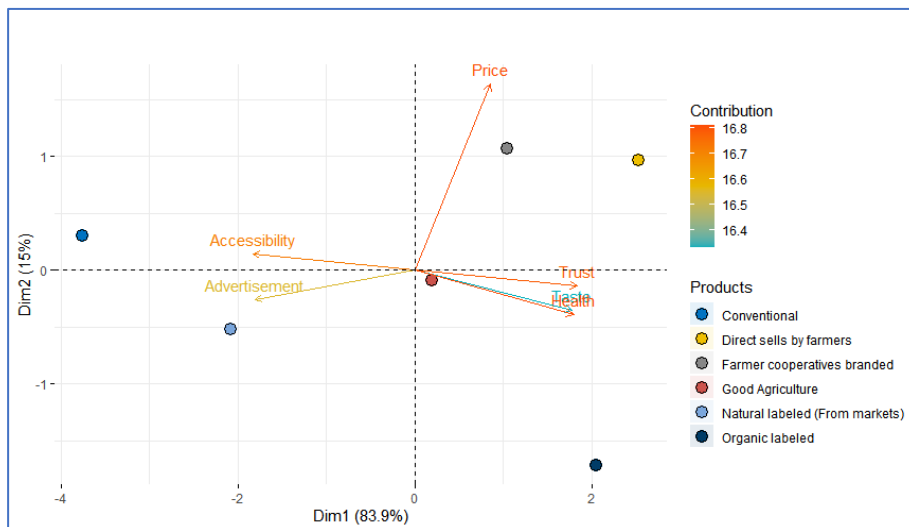


Figure 2. PCA-Biplot for food product perceptions of consumers.

3.4. Correspondence Analysis (CA)

Bartlett’s test for sphericity (Chi square [X2] =28.91591, P <0.01) and the Kaiser-Meyer-Olkin measure of sample adequacy (KMO=0.67) indicated that the dataset was suitable for CA. Parallel analysis revealed that only the first dimension was necessary; however, the second dimension was also maintained so that the result could be displayed on a two-dimensional plane (Figure 3a). The first two components obtained from the CA account for 93.5% of the total variation in dataset (Figure 3b).

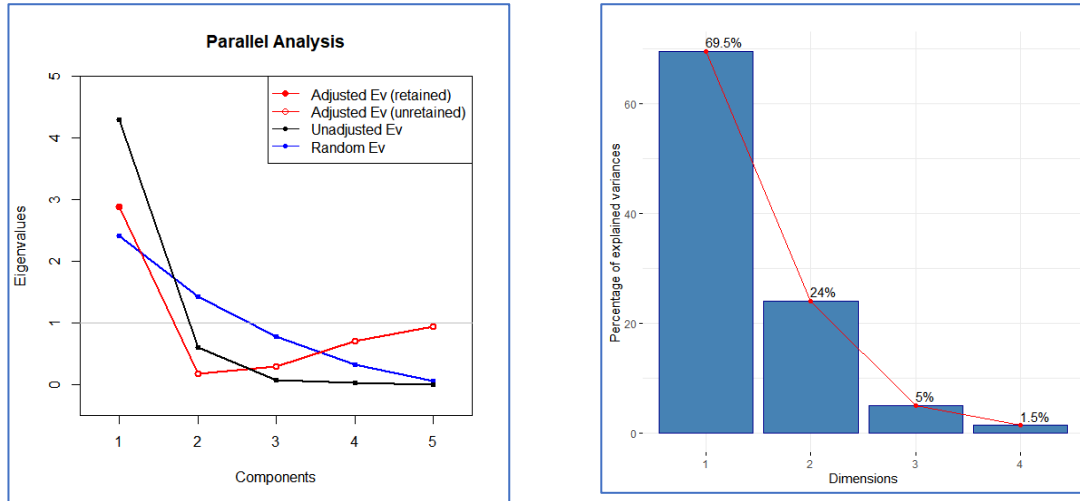


Figure 3. a) Parallel analysis, b) Percentage of explained variance for each dimension

3.5. Correspondence Analysis (CA)

Correlations between variables are provided in Figure 4a. The sign and magnitude of correlations were used to explain and interpret the relationship between rows and columns. Blue colors show positive correlations, while red colors show negative correlations. The size of the nodes indicates the magnitude of the correlations. There was a highly positive correlation between municipal direct sale points and cooperative labeled products, which indicated that consumers mostly prefer to buy cooperative labeled products from municipal direct sale points (Figure 4b). Similarly, while locally labeled (Direct sells by farmers) products were mostly purchased from public markets, this was followed by purchases directly from the producers. In addition, when purchasing products with good agricultural practices and natural labels (from markets), they mostly prefer grocery stores and supermarkets.

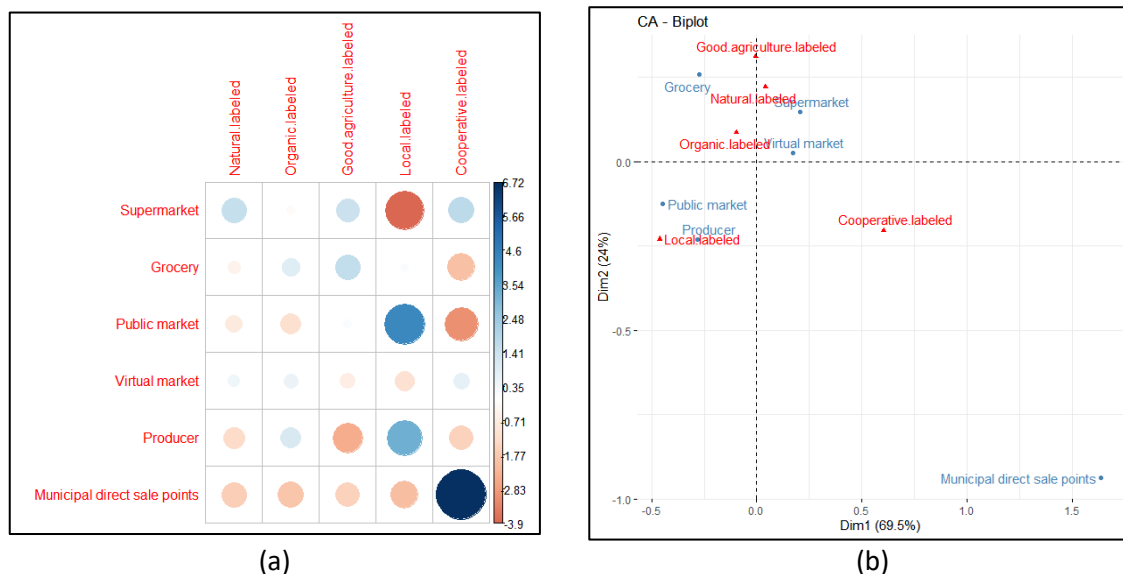


Figure 4. a) Correlations between variables, b) Biplot of correspondence analysis.

This study revealed that householders were more likely to buy organic foods than renters. Similarly, younger individuals (under 37) were shown to be more likely to buy natural labeled products than older ones. In line with these results, (Banytė et al., 2010) reported that young consumers have a higher environmental awareness and desire for a healthier lifestyle than older consumers, and they are more likely to buy eco-friendly food products.

Habit and age had an effect on the probability of purchasing organic labeled products. Male customers were 54% less likely to buy organic labeled foods than female consumers. Similarly, (Curl et al., 2013) reported that female consumers are more likely than male consumers to buy organic foods. Moreover, (Witek and Kuźniar, 2021) reported that females showed more enthusiasm to support environmental causes and green goods than males. In contrast, (Shrestha and Baral, 2019) discovered that men were 2% more likely than women to be willing to pay a higher price for organic goods. Given their purchasing habits, consumers in the low-eco-friendly category were less likely to buy food labeled organic or produced using good agricultural practices than high-eco-friendly consumers. The age of consumers is another factor that affects the purchasing of organically labeled foods. These findings are consistent with those of earlier research. In this respect, (Curl et al., 2013) showed that age was firmly connected with organic consumption preferences, such that the probability of being an organic consumer decreased by 13% for every 10-year increase in age. Similarly, study carried out in Bangladesh observed that younger consumers are more likely to buy eco-friendly food products (Fahmida et al., 2020). According to Polish research (Bryła, 2016), Polish young individuals trust organic food more since it is associated with stricter controls, looks much nicer, and is more environmentally friendly. However, research done in Thailand by (Roitner-Schobesberger et al., 2008) found that older consumers were more likely to buy organic vegetables.

There is evidence that older customers are more to purchase products directly from producers than younger consumers. The odd ratio was %47 higher for older customers, indicating that consumers over the age of 36 were more likely to purchase directly from producers than those under the age of 37. Similarly, (Onianwa et al., 2005) found that while there was no significant correlation between consumer age and buying food directly from farmers (producers), older individuals were more likely to buy food directly from farms. According to the findings of a study conducted in Romania by Polimeni et al. (2018) age has been found as an important factor while shopping at the piata, or farmer's market. It has been stated that older Romanian consumers choose locally produced goods and are less concerned about price, which may be due to the fact that older Romanians have purchased at the piata for years and formed relationships with sellers (Polimeni et al., 2018). In this study, the perception that foods purchased directly from farmers are healthier, of higher quality, less expensive, and produced using traditional techniques may be the underlying reason why older consumers in Turkey prefer to buy directly from farmers.

The findings of this study suggest that consumers perceive organic and good agricultural practices products, which are ecologically friendly food, to be distinct from other food products (conventional food, Farmers Cooperatives branded and market products). They are confident in these products and believe they are safe for consumption. Consumers, on the other hand, believe that the price and accessibility of these products are limited. Comparable research indicates that consumers see environmentally friendly foods as healthy, natural, fresh, tasty, trustworthy, nutritious, and safe for human and animal health (Dardak et al., 2009; Ağır et al., 2014). According to other surveys, consumers regard organic foods to be of better quality than conventional foods, despite their higher pricing (Bilgen, 2017; Rodríguez-Bermúdez et al., 2020). Similarly, consumers prefer organic foods because they perceive them to be nutritious, natural, fresh, tasty, pesticide-free, and caring for animal welfare (Dardak et al., 2009; Bahsi and Akça, 2019; Rodríguez-Bermúdez et al., 2020; İnan et al., 2021). On the other hand, the most significant barriers to purchasing organic food are their high price, limited availability (Ağır et al., 2014; Bahsi and Akça, 2019), and lack of trust in these products (Doğan and Gürel, 2016). The findings of this study are consistent with earlier studies. This study differs from the others in that it assesses consumer perception of a variety of products and product attributes presented under different concepts.

4. Conclusions

The findings of this study suggest that socioeconomic aspects of consumers affect product positioning. In addition, consumers are familiar with food products marketed under various concepts and can evaluate the products based on their attributes. In addition, it was established that there is a perception that the specified food products have distinct product qualities. Consumers perceive ecologically friendly products to be tasty, healthy, and safe, but they find their prices to be high and their availability to be limited. Based on these results, extending the availability and

advertising/promotion efforts for commercially available environmentally friendly foods, as well as implementing steps to reduce their prices, may result in an increase in the purchase of these products. The direction of future studies may be to identify consumers' perceptions of different products by segmenting them according to their characteristics using different machine learning techniques.

Ethical Statement

There is no need to obtain permission from the ethics committee for this study.

Conflicts of Interest

We declare that there is no conflict of interest between us as the article authors.

Authorship Contribution Statement

Concept: Hurma, H, Çakmakçı, Y.; Design: Hurma, H, Çakmakçı, Y.; Data Collection or Processing: Hurma, H, Çakmakçı, Y., Çakmakçı, C.; Statistical Analyses: Hurma, H, Çakmakçı, Y., Çakmakçı, C.; Literature Search: Çakmakçı, Y.; Writing, Review and Editing: Hurma, H, Çakmakçı, Y., Çakmakçı, C.

References

- Abdi, H. and Williams, L. J. (2010). Principal Component Analysis (Issue 2). www.utdallas.edu/.
- Ağır, B. H., Poyraz, N., Yılmaz, H. İ. and Boz, İ. (2014). Organic product perception of consumers: Sample of Kayseri Province. *XI. Congress of Agricultural Economics*, 3-5 September, Samsun, Türkiye.
- Aprile, M. C., Caputo, V. and Nayga, R. M. (2016). Consumers' Preferences and Attitudes Toward Local Food products. *Journal of Food Products Marketing*, 22(1): 19–42. <https://doi.org/10.1080/10454446.2014.949990>
- Argemí-Armengol, I., Villalba, D., Ripoll, G., Teixeira, A. and Álvarez-Rodríguez, J. (2019). Credence cues of pork are more important than consumers' culinary skills to boost their purchasing intention. *Meat Science*, 154: 11–21. <https://doi.org/10.1016/j.meatsci.2019.04.001>
- Bahsi, N. and Akça, A. (2019). A research on the determination of consumers' perspectives on organic agricultural products: Case study in Osmaniye and Şanlıurfa Provinces. *Journal of Agriculture and Nature*, 22(1): 26–34. <https://doi.org/10.18016/ksutarimdogavi.443228>
- Banytė, J., Brazionienė, L. and Gadeikienė, A. (2010). Investigation of green consumer profile: A case of Lithuanian market of eco-friendly food products. *Economics and Management*, 15: 374–383.
- Bilgen, İ. (2017). Perceived quality in organic agriculture products: A research on the consumers in Istanbul. *The Turkish Online Journal of Design, Art and Communication*, 7(4): 678–685. <https://doi.org/10.7456/10704100/013>
- Bougherara, D. and Combris, P. (2009). Eco-labelled food products: What are consumers paying for? *European Review of Agricultural Economics*, 36(3): 321–341. <https://doi.org/10.1093/erae/jbp023>
- Bryła, P. (2016). Organic food consumption in Poland: Motives and barriers. *Appetite*, 105: 737–746. <https://doi.org/10.1016/J.APPET.2016.07.012>
- Çakmakçı, Y. and Hurma, H. (2021). The relationship between socio-economic characteristics and environmental awareness levels of consumers and the factors effective in purchasing environmentally friendly food products. *Turkish Journal of Agriculture - Food Science and Technology*, 9(7): 1297–1303. <https://doi.org/10.24925/turjaf.v9i7.1297-1303.4439>
- Curl, C. L., Beresford, S. A. A., Hajat, A., Kaufman, J. D., Moore, K., Nettleton, J. A. and Diez-Roux, A. v. (2013). Associations of organic produce consumption with socioeconomic status and the local food environment: Multi-Ethnic study of atherosclerosis (MESA). *PLoS ONE*, 8(7):1-8. <https://doi.org/10.1371/journal.pone.0069778>
- Dardak, R. A., Abidin, A. Z. Z. and Kasim Ali, A. (2009). Consumers' perceptions, consumption and preference on organic product: Malaysian perspective (Persepsi, penggunaan dan kecenderungan pengguna terhadap produk organik: Perspektif rakyat Malaysia). *Economic and Technology Management Review*, 4:95-107.
- De-Groote, H. and Kimenju, S. C. (2008). Comparing consumer preferences for color and nutritional quality in maize: Application of a semi-double-bound logistic model on urban consumers in Kenya. *Food Policy*, 33(4): 362–370. <https://doi.org/10.1016/j.foodpol.2008.02.005>
- Di Franco, G. (2016). Multiple correspondence analysis: one only or several techniques? *Quality and Quantity*, 50(3): 1299–1315. <https://doi.org/10.1007/s11135-015-0206-0>
- Díaz-Pérez, M., Carreño-Ortega, Á., Salinas-Andújar, J. A. and Callejón-Ferre, Á. J. (2019). Application of logistic regression models for the marketability of cucumber cultivars. *Agronomy*, 9(1): 17. <https://doi.org/10.3390/agronomy9010017>
- Doğan, G. H., and Gürel, E. (2016). The determination of attitudes and behaviors in organic product consumption of consumers living of central district of Kırşehir Province. *Journal of Agricultural Faculty of Gaziosmanpaşa University*, 33(2016-2): 147–156. <https://doi.org/10.13002/jafag1033>
- Echeverría, R., Montenegro, A. B., Albarrán, E. S. and Charry, L. (2022). Consumer willingness to pay for cheese with a social sustainability attribute. *Ciência Rural*, 52(5):1-8. <https://doi.org/10.1590/0103-8478cr20210281>
- Eldesouky, A., Mesias, F. J. and Escribano, M. (2020). Perception of Spanish consumers towards environmentally friendly labelling in food. *International Journal of Consumer Studies*, 44(1): 64–76. <https://doi.org/10.1111/ijcs.12546>
- Fahmida, A., Palash, M. S., Alam Monirul, G. M. and Amin, M. R. Md. (2020). Young consumers' eco-friendly food purchasing consciousness-behaviorgap in Bangladesh. *The Bangladesh Journal of Agricultural Economics*, 41(2): 29–44. <https://www.researchgate.net/publication/348658012>
- Fox, J. and Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417):178–183.
- Fox, J. and Weisberg, S. (2019). *An R Companion to Applied Regression Third Edition*. SAGE Publications Inc. https://toc.library.ethz.ch/objects/pdf03/z01_978-1-5443-3647-3_01.pdf.
- Greenacre, M. (2015). Correspondence Analysis. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 1-5). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.42005-2>
- Husson, F., Le, S. and Pagès, J. (2017). *Exploratory multivariate analysis by example using R*. CRC Press.
- IBM (2022). *Machine Learning*. <https://www.ibm.com/cloud/learn/machine-learning>

- İnan, R., Bekar, A. and Uurlu, H. (2021). An assessment of consumers of organic food purchase behavior and attitudes. *Journal of Tourism and Gastronomy Studies*, 9(1): 220–235. <https://doi.org/10.21325/jotags.2021.786>
- Ingwersen, E. W., Stam, W. T., Meijjs, B. J. V., Roor, J., Besselink, M. G., Groot Koerkamp, B., de Hingh, I. H. J. T., van Santvoort, H. C., Stommel, M. W. J. and Daams, F. (2023). Machine learning versus logistic regression for the prediction of complications after pancreatoduodenectomy. *Surgery*, 174(3): 435-440. <https://doi.org/10.1016/J.SURG.2023.03.012>
- Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: A review and recent developments. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 374, Issue 2065). Royal Society of London. <https://doi.org/10.1098/rsta.2015.0202>
- Kashif, U., Hong, C., Naseem, S., Khan, W. A. and Akram, M. W. (2020). Consumer preferences toward organic food and the moderating role of knowledge: A case of Pakistan and Malaysia. *Ciencia Rural*, 50(5): 1–13. <https://doi.org/10.1590/0103-8478cr20190842>
- Kassambara, A. and Mundt, F. (2020). Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>
- Klopčič, M., Kuipers, A. and Hocquette, J. F. (2012). *Consumer attitudes to food quality products: Emphasis on Southern Europe* (EAAP publication no. 133). Wageningen Academic Publishers. <https://edepot.wur.nl/242193>
- Nam, K., Lim, H. and Ahn, B. il. (2020). Analysis of consumer preference for milk produced through sustainable farming: The case of mountainous dairy farming. *Sustainability*, 12(7): 1-15. <https://doi.org/10.3390/su12073039>
- Nascimento, A. G. M., Toledo, B. S., Guimarães, J. T., Ramos, G. L. P. A., da Cunha, D. T., Pimentel, T. C., Cruz, A. G., Freitas, M. Q., Esmerino, E. A. and Mársico, E. T. (2022). The impact of packaging design on the perceived quality of honey by Brazilian consumers. *Food Research International*, 151: 110887. <https://doi.org/10.1016/J.FOODRES.2021.110887>
- Newbold, P. (1995). *Statistics for Business and Economics*. Prentice-Hall International.
- Onianwa, O., Wheelock, G. and Mojica, M. (2005). An analysis of the determinants of farmer-to-consumer direct-market shoppers. *Journal of Food Distribution Research*, 36(1): 1-5.
- Oraman, Y., Unakitan, G., Yılmaz, E. and Başaran., B. (2011). Analysis of the factors affecting consumer's some traditional food products preferences by multidimensional scaling method. *Journal of Tekirdag Agricultural Faculty*, 8 (1): 33-40. <https://dergipark.org.tr/pub/jotaf/issue/19043/201415>
- Petrescu, D. C., Vermeir, I. and Petrescu-Mag, R. M. (2020). Consumer understanding of food quality, healthiness, and environmental impact: A cross-national perspective. *International Journal of Environmental Research and Public Health*, 17(1): 1-20. <https://doi.org/10.3390/ijerph17010169>
- Polimeni, J. M., Iorgulescu, R. I. and Mihnea, A. (2018). Understanding consumer motivations for buying sustainable agricultural products at Romanian farmers markets. *Journal of Cleaner Production*, 184: 586–597. <https://doi.org/10.1016/J.JCLEPRO.2018.02.241>
- R Core Team. (2020). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.r-project.org/>
- Ramya, N. and Ali, S. M. (2016). Factors affecting consumer buying behavior. *International Journal of Applied Research*, 2(10): 76–80. www.allresearchjournal.com
- Revelle, W. (2019). Package “psych”: Procedures for Psychological, Psychometric, and Personality Research. <https://CRAN.R-project.org/package=psych> Version = 1.9.12.
- Ripoll, G., Joy, M. and Panea, B. (2018). Consumer perception of the quality of lamb and lamb confit. *Foods*, 7(5): 1-15. <https://doi.org/10.3390/foods7050080>
- Rodríguez-Bermúdez, R., Miranda, M., Orjales, I., Ginzo-Villamayor, M. J., Al-Soufi, W. and López-Alonso, M. (2020). Consumers' perception of and attitudes towards organic food in Galicia (Northern Spain). *International Journal of Consumer Studies*, 44(3): 206–219. <https://doi.org/10.1111/ijcs.12557>
- Roitner-Schobesberger, B., Darnhofer, I., Somsook, S. and Vogl, C. R. (2008). Consumer perceptions of organic foods in Bangkok, Thailand. *Food Policy*, 33(2): 112–121. <https://doi.org/10.1016/J.FOODPOL.2007.09.004>
- Shiokawa, Y., Date, Y. and Kikuchi, J. (2018). Application of kernel principal component analysis and computational machine learning to exploration of metabolites strongly associated with diet. *Scientific Reports*, 8(1): 1-8. <https://doi.org/10.1038/s41598-018-20121-w>
- Shrestha, A., and Baral, S. (2019). Consumers' willingness to pay for organic agriculture products: a case study of Nepalgunj city, Banke. *International Journal of Agriculture, Environment and Food Sciences*, 3(2): 58–61. <https://doi.org/10.31015/jaefs.2019.2.2>
- Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunanathan, S., Bandeen-Roche, K., Béland, F. and Bergman, H. (2010). Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *Journal of Clinical Epidemiology*, 63(6): 638–646. <https://doi.org/10.1016/j.jclinepi.2009.08.008>
- Vindigni, G., Peri, I., Consentino, F., Selvaggi, R. and Spina, D. (2022). Exploring consumers' attitudes towards food products derived by new plant breeding techniques. *Sustainability*, 14(10): 1-14. <https://doi.org/10.3390/su14105995>

Witek, L. and Kuźniar, W. (2021). Green purchase behavior: The effectiveness of sociodemographic variables for explaining green purchases in emerging market. *Sustainability*, 13(1): 1–18. <https://doi.org/10.3390/su13010209>

Yılmaz, E. Oraman, Y. and İnan, İ, H. (2009). Determination of consumer behaviour dynamics relating to food products: Trakya Region Example. *Journal of Tekirdag Agricultural Faculty*, 6(1): 1-10. <https://dergipark.org.tr/tr/pub/jotaf/issue/19049/201465>