Machine Learning Promoting Sustainable Customer Behavior and Product Pricing

Pourmahdi, Katayoon; Nyström, Anna-Greta; Majd, Amin

Published in:
Contemporary Issues in Industry 5.0

DOI:
10.1007/978-3-031-74779-3_11

Published: 18/01/2025

Document Version
Final published version

11

Machine Learning Promoting Sustainable Customer Behavior and Product Pricing

Katayoon Pourmahdi, Anna-Greta Nyström, and Amin Majd

11.1 Introduction

A more sustainable approach to marketing strategies helps firms protect the environment ecologically and leads to better organizational performance (Håkansson et al. 2005).

Sustainability concerns have become a highlighted topic influencing the marketing strategies of companies, particularly as regards product pricing strategies. Sustainable marketing is defined as a process involving the planning, implementation, and control of pricing, promotion, and distribution of products that reconciles ecological and economic factors (Fuller 1999; Sheth and Parvatiyar 1995). Although pricing strategies

K. Pourmahdi (✉) · A.-G. Nyström
Faculty of Social Sciences, Business and Economics, and Law, Åbo Akademi University, Turku, Finland
e-mail: katayoon.pourmahdi@abo.fi

A. Majd
Faculty of Engineering and Business, Turku University of Applied Sciences, Turku, Finland

P. Aaltonen and E. Kurvinen (eds.), *Contemporary Issues in Industry 5.0*, Technology, Work and Globalization, https://doi.org/10.1007/978-3-031-74779-3_11

are currently attracting attention from marketers and policymakers, few academic investigations specifically focus on pricing strategies. This topic has not been as extensively theoretically developed as other marketing subjects, such as promotion, product, and distribution (Hinterhuber 2004; Ingenbleek 2014; Liozu et al. 2012; Liozu and Hinterhuber 2013). Empirical studies show that only a few academic articles published in leading marketing journals have incorporated pricing strategies (Hinterhuber 2004; Liozu et al. 2012). Nevertheless, the pricing strategies of firms are becoming an important buying criterion for price-sensitive consumers (Belz and Peattie 2014; Hinterhuber 2004); in times of increasing energy prices and a focus on saving fuel energy, products are awarded labels indicating their energy efficiency levels, corresponding to different prices (Fuller 1999).

To address concerns regarding sustainable product pricing and price-sensitive customers' behavior, big data analytics powered by artificial intelligence (BDA-AI) is being used to assist companies in identifying influential factors affecting customer purchase decision-making (Dubey et al. 2020). More precisely, adapting Machine Learning (ML) methods offers a powerful tool for marketers to analyze and identify interactions within large quantities of data. Consequently, this chapter focuses on pricing decisions and sustainable consumer behavior, employing descriptive and predictive analytics using ML methods to visualize patterns and predict sustainable customer behavior. We aim to answer the following research questions:

What factors, in addition to price, influence customers' decision-making process when choosing sustainable food products that address sustainability concerns in their production processes? What is the likelihood of customers choosing sustainable food products over regular food products? The first research question aims to develop a theory-driven model for studying the topic, whereas the second research question showcases how descriptive and predictive analytics using ML methods can be deployed. From a theoretical standpoint, we review behavior theories to identify the main components affecting customer behaviors, such as attitudes, intentions, and habits. Then, using these theories, we propose a study model that showcases the inputs, exogenous factors, and hypothetical constructs of sustainable customer behavior that inform the empirical analysis. Finally, we develop a data-driven model utilizing descriptive and predictive analytics using ML methods, specifically visualization and logistic regression, to explore customers' purchasing behavior regarding sustainable products.

11.2 Consumer Purchase Behavior and Pricing Strategies: An Overview of Previous Research

This section presents an overview of theories related to consumer behavior in purchase decision-making, and pricing strategies. The review aims to explain the decision-making process, which will subsequently be examined using ML methods.

11.2.1 Consumer Behavior in Purchase Decision-Making

Success in creating a sustainable marketplace depends on developing a comprehensive understanding of consumer intention, their perception of purchasing sustainable products, and the barriers consumers encounter during the decision-making process to prevent sustainable consumption. By delving into the realm of consumer attitude, intention, habit, and behavior, researchers have endeavored to construct and evaluate models that forecast the drivers and barriers influencing sustainable consumer behavior (Ma et al. 2018). The following explanation of related theories gives a flavor of some important concepts related to consumer behavior.

According to the theory of planned behavior, the main determinant of consumer behavior is intention. Theoretically, a particular human behavior can be predicated by its intention, which is influenced by three core components: the subjective norm, the individual's behavioral attitudes (values, beliefs, and norms), and perceived behavioral control (Ajzen 1991). This theory was developed to predict an individual's intention to engage in a certain behavior at a specific place and time. Alphabet theory offers a detailed description of the relationship between habits, intentions, and the actual behavior of consumers. This theory was formulated through the synthesis of two environmental behavior theories: attitude–behavior–context theory and value–belief–norm theory. According to alphabet theory, knowledge, information-seeking, demographics, and context are the main components that affect human attitudes (Martínez-Carrasco Martínez et al. 2023; Sadeli et al. 2023; Taghikhah et al. 2021; Zepeda and Deal 2009). The rational choice theory is an early theory clarifying the understanding of the social, environmental, and economic behavior of customers. It has been used to describe the link between perception and human behavior in different contexts. According to this theory, an individual conducts a cost–benefit analysis before making an actual purchase decision (Zepeda and Deal 2009).

Inspired by the theories of planned behavior, alphabet theory, and rational choice theory, Fig. 11.1 illustrates a framework for capturing consumer behavior in purchase decision-making. Consequently, the components affecting individual attitudes are demographics, knowledge, information seeking, and context, while attitudes, habits, and context affect intention, intention, and habit subsequently affect behavior.

11.2.2 Capturing Consumer Behavior for Sustainable Product Purchases

Several theories and models, such as the Nicosia, Eagle, Kollat, and Blackwell models, have been proposed to explain consumer behavior and its influence on marketing strategies (Juan et al. 2017). A commonly used theory is the Howard–Sheth model, according to which consumer behavior theory, input, and external factors can provide various messages that can be crucial in customer purchasing decisions (Howard and Sheth 1969; Sheth 2011). The model claims that the effects of attitude on purchases are only possible through intention (Howard and Sheth 1969; Juan et al. 2017) and suggests three levels of consumer decision-making. These three categories are extensive problem-solving,

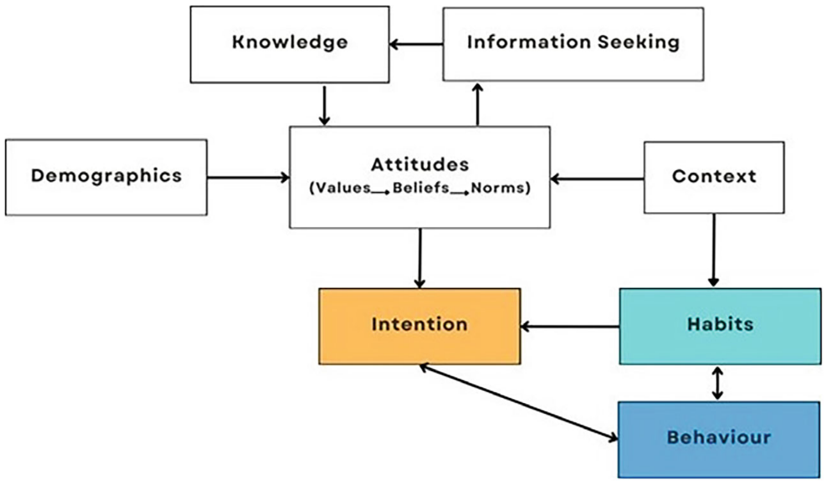

Fig. 11.1 Overview of consumer behavior in purchase decision-making

limited problem-solving, and habitual response behavior. These categories comprise four components of consumer behavior: input variables, hypothetical constructs, exogenous variables, and output variables. The input variables consist of three stages: significative, symbolic, and social. The output variables occur in a logical sequence, beginning with attention, brand comprehension, attitudes, and intentions, and ending with purchase. Hypothetical factors affect inputs and outputs' learning and perception constructs (Juan et al. 2017; Sheth 2011).

In line with consumer behavior theories and the logic of the Howard–Sheth model, we adopted four sets of dimensions in our study: input variables (product characteristics), hypothetical construct (customers' intentions, attitudes, and habits toward sustainable products), output variables (sales performance of sustainable products), and exogenous variables, which are not directly involved in decision making (demographic information). This study model simulates the real world and aims to attain a comprehensive understanding of the factors influencing sustainable consumer behavior. In Fig. 11.2, we identify the demographic information, product characteristics, and customer intentions, attitudes, and habits as determinants affecting sustainable consumer behavior.

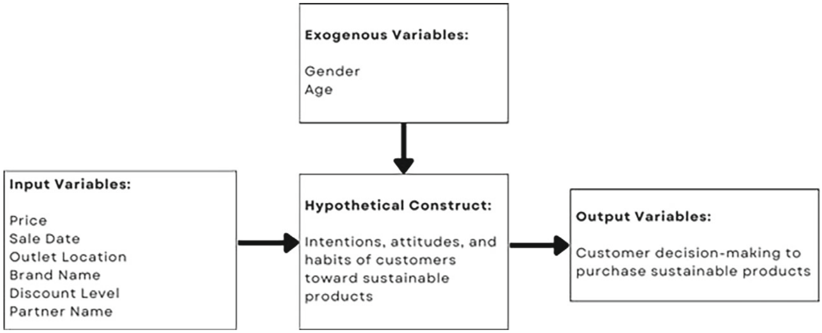

Fig. 11.2 Study model of customer behavior of sustainable product purchase

11.2.3 Pricing Strategies

Setting optimal pricing strategies requires the company to commit to objectives, actions, and operational strategies and to employ a set of control and review procedures. The marketing process of a company involves strategic choices that impact the type of pricing employed by the firm. Factors such as customer demographics, behavior, and product details all influence pricing. It is important to acknowledge the input required from departments such as accounting, research and development, sales, marketing, and manufacturing to implement optimal pricing strategies (Kotler and Keller 2016; Lancioni 2005). Moreover, the implementation of optimal pricing strategies is influenced by several key concepts, including market segmentation (Daraboina et al. 2024; Tynan and Drayton 1987), price elasticity of demand (Bijmolt et al. 2005; Whitaker 1988), and the concept of premium pricing (Anselmsson et al. 2014; Ashraf et al. 2017; Juan et al. 2017).

11.2.3.1 Pricing Strategy Development Based on Market Segmentation

Market segmentation enables marketers to address the diversity of consumers and their behaviors (Tynan and Drayton 1987). Sustainability market segmentation, in turn, divides heterogeneous markets into smaller ones. Thus, marketers vary their offerings to meet the evolving needs of customers regarding sustainable products. The four major variables used for market segmentation are geographic (customer's place of residence), behavioral (final decision to purchase), demographic (age, gender, and family size), and psychographic variables (attitude, intention, and habit) (Fuller 1999; Tynan and Drayton 1987). Considering consumers' perception of sustainability, which in turn influences their attitudes, intentions, and behaviors, the implementation of different pricing strategies based on the four major variables of sustainability market segmentation can create a win–win-win situation for consumers, firms, and society (Hempel 2024; Zhang and Zheng 2022). To clarify, traditionally, customers and businesses have been considered two parties in a competitive transactional game. In business transactions, price is a tentative quotation offered by the seller to a potential customer, which can be either accepted or refused. However, from a new competitive perspective, sustainable marketing considers the environment as a new party. All parties underpin transactions and aim for mutual success, making the integration of environmental costs into product prices a vital step (Fuller 1999).

11.2.3.2 Price Elasticity of Demand

By considering the price elasticity of products, decision-makers can calculate customers' willingness to pay for the product at different price points (Bijmolt et al. 2005). According to the price elasticity of demand devised by Marshall (2011), the formula for the coefficient of price elasticity of demand for products X_i (1, 2) is, $e_{(R)} = \frac{\frac{dQ}{Q}}{\frac{dP}{P}}$, where $\frac{dQ}{Q}$ represents the percentage change in demand for the good, and $\frac{dP}{P}$ is the percentage change in the price of the good. Although the demand for products generally moves in the opposite direction from their price, the impact of price changes can vary. The demand for some products is not significantly affected by changes in their prices, while the demand for others is highly responsive to price changes. The price elasticity of the demand for a product measures the percentage change in demand divided by the percentage change in its price. Products with high elasticity are considerably sensitive to price changes, whereas products with low elasticity are less responsive to price fluctuations (Auer and Papies 2020; Bijmolt et al. 2005; Ma et al. 2018; Marshall 2011; Whitaker 1988). The elasticities change according to the retailers, the manufacturer brand, location, time trend, stage of the product life cycle, household disposable income, inflation rate, and, importantly, the product category (Bijmolt et al. 2005).

11.2.3.3 Premium Pricing

From a sustainability perspective, the prices of sustainable products are higher than those of unsustainable products under normal competitive conditions (Ingenbleek 2015). The higher prices reflect the environmental costs, aiming to reduce the destruction and waste caused by production. This pricing strategy is known as *premium pricing*. Premium pricing refers to the practice of a retailer pricing a product or service above the market price in the same marketplace (Allsopp 2005; Fuller 1999; Juan et al. 2017).

According to Ottman (1993), customers are more likely to be receptive to green product prices when their primary needs for affordability, convenience, quality, and functionality are met. Additionally, as customers become more aware of environmental issues, the ecological attributes of products can influence their final purchase decisions and motivate them to pay premiums (Fuller 1999; Juan et al. 2017). In a typical market setting, customers seek products and services that meet their needs. It is important to recognize that a clean and habitable ecosystem is also a legitimate need; thus, customers must be aware of and prioritize the relationships between consumption decisions and environmental quality. To obtain environmental benefits, the five eco-cost drivers that may impact the unit cost structures are as follows: (1) product inputs of raw materials and energy; (2) process, facility, and management; (3) fugitive emission clean-up; (4) environmental legal action; and (5) routine regulatory compliance (Fuller 1999).

Primarily, the significance of the concepts above becomes even more important when considering sustainable products, which often have higher prices compared to regular products (Fuller 1999; Ingenbleek 2015). In this chapter, our results show that customers need to be segmented according to their needs and preferences. Companies must consider their profitability, costs, and external competitive dynamics. Additionally, product elasticity is measured to facilitate the implementation of optimal pricing decisions. Finally, customers must be convinced that the higher prices of sustainable products are a legitimate need, not only for our generation but also for future generations (Kotler and Armstrong 2010). To achieve this goal, in the following section, we analyze real-world retail data using ML methods to extract the significant features and position them effectively. We assume that ML methods assist in making effective pricing decisions by considering significant behavioral, demographic, and product features.

11.3 Descriptive, Predictive, and Prescriptive Analytics

The integration of ML in business analytics takes place through three distinctive analytical classifications (see Fig. 11.3): descriptive, predictive, and prescriptive analytics (Greasley 2019). While descriptive analytics (i.e., business intelligence) focuses on understanding past patterns and events, predictive analytics and prescriptive analytics are oriented toward the future, aiming to predict an outcome with a certain likelihood of accuracy. Concerning model development and analysis, both predictive and prescriptive analyses can be defined as methods utilizing historical data and using this data to make predictions. After importing the training dataset into ML algorithms, the prediction model can make predictions on new data. These analytics enable decision-makers to address questions such as "What will happen?" and "What next?" considering historical datasets and data-driven predictions. One should, however, note that descriptive analytics mainly deals with structured data, which is data that is processed by humans. Predictive and prescriptive analytics process and analyze both structured and unstructured data using computer algorithms (El Morr and Ali-Hassan 2019; Lone and Sofi 2022). The integration of ML into business operations provides a powerful tool for optimizing processes, solving complex problems, and making informed decisions within complex systems (Fishwick 1992). Interest in integrating ML algorithms into business operations has generated a new stream of literature highlighting the importance of developing specific capabilities in an organization before ML adoption. For instance, Keegan et al. (2022) highlight that firms face challenges preceding ML adoption in marketing. One specific challenge is the need to gain access to large high-quality datasets and acquire the necessary technological infrastructure for processing such data (procurement process). Furthermore, AI readiness and AI enablers have been proposed as core concepts in developing ML capabilities. For instance, Baabdullah et al. (2021) developed a conceptual model based on the technology-organization-environment framework (TOE) for understanding the impact of AI readiness and AI enablers on the acceptance of AI practices in the context of business-to-business small and medium enterprises.

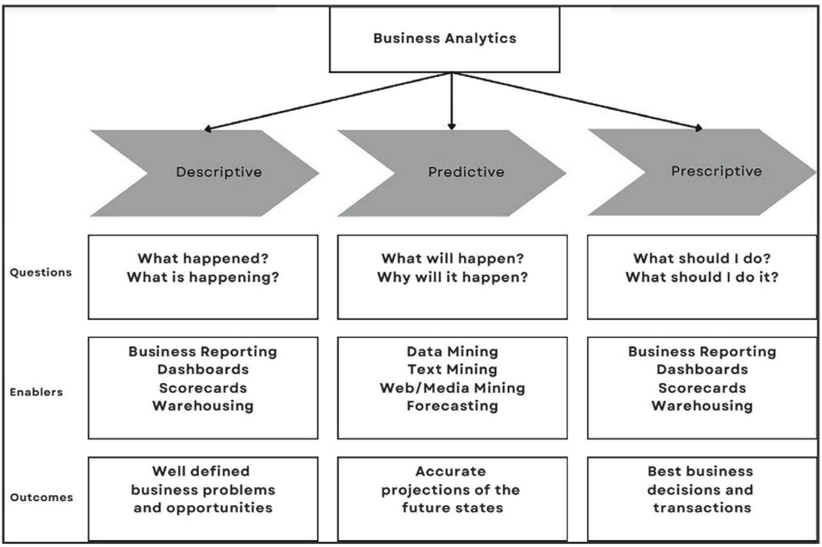

Fig. 11.3 Using ML in business analytics

The application of ML in the *food industry* effectively simulates its intangible aspects, including continuous feedback on the decision-making process despite limited information. This is particularly evident in the context of food policymaking, which requires long-term strategic decision-making regarding consumer preferences and future scenarios (Kler et al. 2022).

Integrating ML algorithms into the *retail industry* enables retailers to collect, curate, and analyze vast amounts of customer data and purchasing patterns. With this integration, retailers can predict demand patterns accurately. Furthermore, it enhances inventory management, demand forecasting, pricing optimization, and supply chain development. An overview of data analysis types, such as numeric, text, voice, and image/video data analysis, allows retailers to utilize and benefit from ML. The strategic adoption of ML in the retail industry is thus important for decision-makers but requires an understanding of how ML can specifically benefit their operations and customers rather than following trends or integrating ML into their businesses simply for public relations purposes (Shankar 2018).

Decision-making through ML algorithms and computational methods has been demonstrated by, for instance, Huiru et al. (2018), who employed experiments and mathematical analysis to show how customers shift their decisions toward other original alternatives when there is another option. Furthermore, realizing the benefits of AI requires interactive collaboration among suppliers, customers, and AI in the development of value co-creation practices. This involves adopting service-dominant logic and expanding critical capabilities in business-to-business (B2B) marketing (Paschen et al. 2021).

To summarize, the primary motivation for employing ML methods in the context of sustainable consumer behavior is the ability of these methods to capture the complexities arising from interactions among multiple agents (Huiru et al. 2018). However, in our study, the objective of the data-driven model is to examine customer decision-making using ML algorithms and explore how organizations can influence their customers' behavior to purchase sustainable products rather than unsustainable protein-based substitutes. The focus of this model is on predicting the demand of consumers for sustainable products while changes in prices and the consumers' perception of greenness occur in the system.

11.4 Descriptive Analysis of Retail Data: Vegetarian/Vegan Versus Meat-Based Purchases

Based on the overview of pricing strategies, consumer behavior, and especially the identified factors illustrated in Figs. 11.1 and 11.2, this section develops a framework for further elucidating the factors influencing sustainable food purchasing decisions and subsequently sustainable consumption. To exemplify how ML methods can aid in decision-making, we assume that customers plan to purchase protein-based products, which can be either vegetarian/vegan, or meat-based products. A protein-based food dataset from a food retailer was then used, including information on customer clusters and their sales performance over 5 years for two types of products, and the occurrence of nearly one million sales. There are fifteen customer clusters and two protein-based product types: vegetarian/vegan (P1) and meat-based (P2). In this study, the sustainability concerns are limited to the production process of vegetarian/vegan products, i.e., the product, its production process, or the packaging highlighting green consumerism and sustainable action. When consumers, acting as autonomous entities, arrive at the marketplace, they find both P1 and P2 items available in the retail store. Each consumer has a product preference that reflects their perception of sustainable consumption. During the decision-making phase, consumers decide whether to buy P1 or P2 items and consider various factors, such as the price. First, we evaluated the likelihood of each customer cluster purchasing each product category. Then, logistic regression modeling was applied to historical sales performance data to predict the purchase performance of new customers. This analysis aims to provide insights into purchasing patterns among different customer groups and forecast future trends in sustainable food consumption based on the developed framework of descriptive and predictive-analytic methods.

As depicted in Fig. 11.4, the prices of protein-based products have increased over time, and the upward slope of meat-based products is sharper than that of vegetarian/vegan products. However, the average price for vegetarian/vegan products is higher than their meat-based counterparts. Additionally, it is observed that the demand for meat-based products is higher than for vegetarian/vegan products. To understand demand and price changes, we refer to the price elasticity of demand for a product, which measures the degree of demand response to changes in an economic factor. Contrary to common belief, it does not mean that a lower price is more appealing to customers (Bijmolt et al. 2005). Therefore, it is assumed that consumers can be persuaded to purchase protein-based products at higher prices, considering factors such as income level, family size, living location, educational level, age, and gender.

The distribution of fifteen customer clusters across five popular outlets is depicted in Fig. 11.5, illustrating the dominant customer clusters in the market. Figure 11.6 and Fig. 11.7 illustrate the comprehensive purchasing trends of each customer group for these two products during

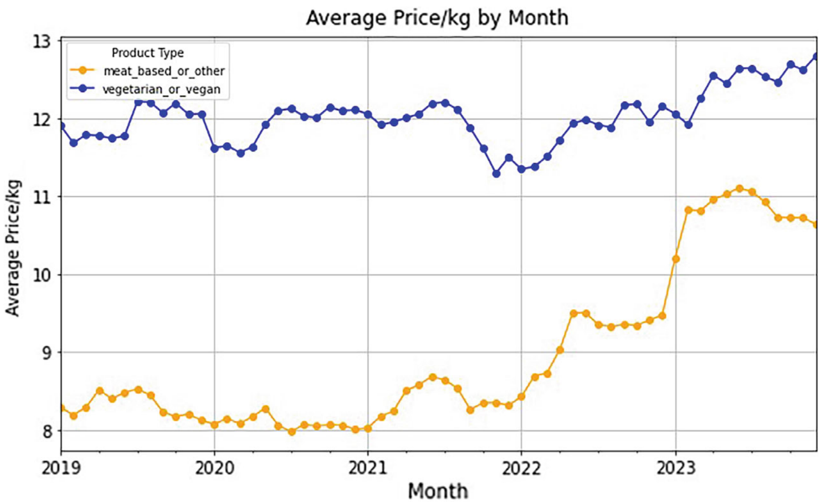

Fig. 11.4 Price changes for meat-based and vegetarian/vegan products over time

five time periods. There is a higher likelihood that females aged 25–44 are interested in vegetarian/vegan products. Thus, when measuring the correlation between customer groups, including age and gender, and sales performance, their correlation coefficient is 0.297, indicating a positive correlation. Additionally, the p-value is significantly less than 0.05, suggesting strong evidence that there is a correlation between the customer group and the sales performance of vegetarian/vegan products (Table 11.1).

The sales performance of vegetarian/vegan products has declined over time, which suggests that the higher prices of vegetarian/vegan products during this period must have affected price-sensitive customers. We can describe this trend in customer purchase performance by employing the salience theory established by Bordalo et al. (2012). The assumption is that there are two types of consumers in the market: one type of consumer is more sensitive to price while the other is more sensitive to greenness and these decision-makers assign higher importance to the product's salient attribute (Balcombe et al. 2021; Bordalo et al.

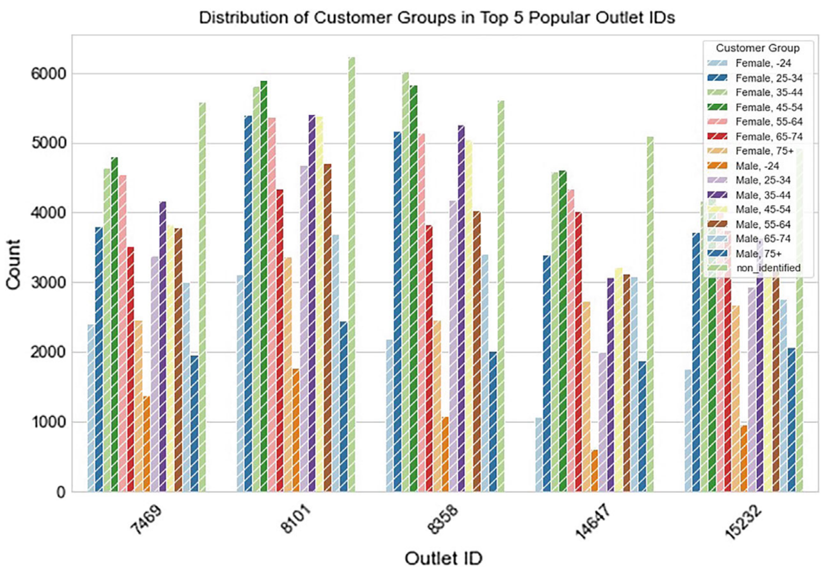

Fig. 11.5 Distribution of customer groups in the top five popular outlet IDs

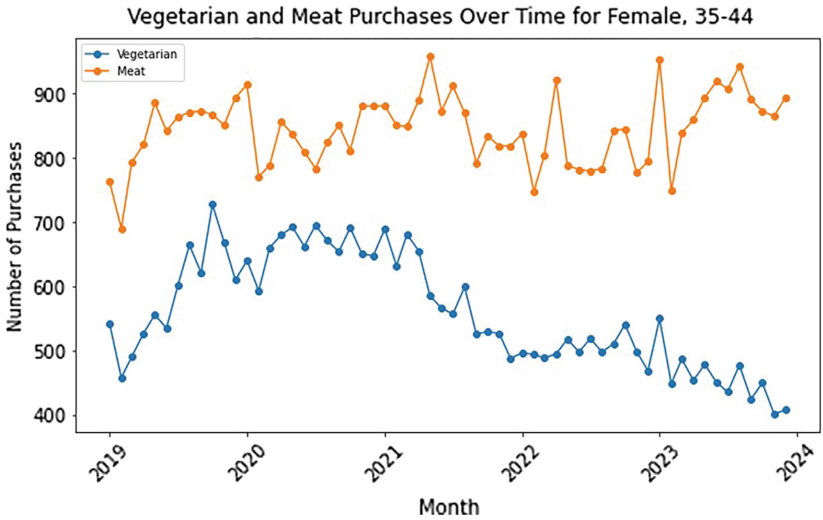

Fig. 11.6 Monthly sales performance of product types for females aged 35–44

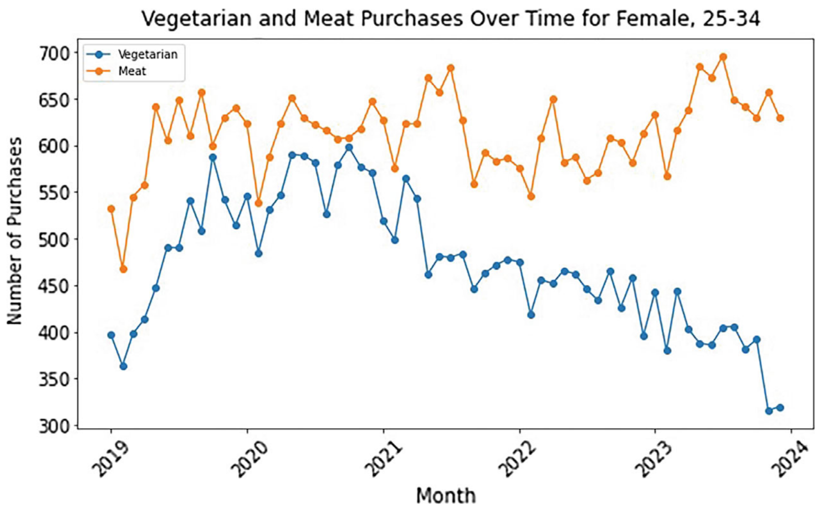

Fig. 11.7 Monthly sales performance of product types for females aged 25–34

Table 11.1 Sample generation for the literature review

Customer group	Meat-based purchase (%)	Vegetarian/vegan purchase (%)
Female, – 24	64.41382	35.58618
Female, 25–34	62.35827	37.64173
Female, 35–44	65.9257	34.0743
Female, 45–54	70.14261	29.85739
Female, 55–64	72.65657	27.34343
Female, 65–74	74.59792	25.40208
Female, 75 +	79.20754	20.79246
Male, – 24	80.66654	19.33346
Male, 25–34	70.07245	29.92755
Male, 35–44	70.81305	29.18695
Male, 45–54	75.3744	24.6256
Male, 55–64	80.35555	19.64445
Male, 65–74	80.77859	19.22141
Male, 75 +	86.04237	13.95763

2012). Type one (p_k) gives a higher weight to price in their decision-making process, while greenness receives the top ranking for type two (q_k). According to the salient thinker, the evolution of weight given by a customer's utility (U^{LT}) for greenness and price over time can be described using the following utility formulas: (1) price-sensitive and (2) green-sensitive customers. (1) $U^{LT}\ (q_k) = \theta_1 \left(\frac{\delta}{\delta\theta_1+\theta_2}\right) q_k - \theta_2 \left(\frac{1}{\delta\theta_1+\theta_2}\right) p_k$ and (2) $U^{LT}\ (q_k) = \theta_1 \left(\frac{1}{\theta_1+\delta\theta_2}\right) q_k - \theta_2 \_ \left(\frac{\delta}{\theta_1+\delta\theta_2}\right) pk$. In this context, θ_1 and θ_2 denote the utility weights, and their sum is equal to 1. δ captures the degree of consumers' salient thinking: δ ($0 < \delta < 1$). The smaller δ, the higher the level of consumers' salient thinking. If customers prefer greenness, the provided equations indicate that the weight of greenness increases over time as $\hat{\theta}_1^k = \theta_1\left(\frac{1}{\theta_1+\delta\theta_2}\right) > \theta_1$, and simultaneously, the weight of the price decreases over time as $\hat{\theta}_2^k = \theta_2\left(\frac{\delta}{\theta_1+\delta\theta_2}\right) < \theta_2$, (Meng et al. 2022). Subsequently, consumers' decisions can change over time based on changes in various factors, including their perception of the salient attributes of the products (Herweg and Müller 2021). To validate the factors that have a statistically significant impact on sales performance (Y) (see Fig. 11.2), factor analysis was

Table 11.2 Influential factors that affect sales performance

	Exogenous variables	Analysis method	*P*-value
X1	Customer group (age, gender)	ANOVA	0.0000
	Input variables	Analysis method	*P*-value
X2	Sale date	ANOVA	4.79e-53
X3	Outlet location	ANOVA	6.40e−37
X4	Brand name	ANOVA	0.0000
X5	Discount level	ANOVA	0.0000
X6	Name of partnering entity	ANOVA	0.0000
X7	Consumer package size	Linear Regression	0.0000
X8	Price	Linear Regression	0.0000

employed to analyze the reliability of each factor in the given dataset. Analysis of Variance (ANOVA) is used for categorical variables, and linear regression is used for numerical variables in the provided code. Based on the results of the P-values, it concludes that variables X1 to X8 have a significant impact on the sales performance of sustainable products (see Table 11.2).

11.4.1 Predictive Analytics on Purchase Data: Logistic Regression Development

In the past few years, the logistic regression model has been widely employed to examine sales performance and customer decision-making (Fadlalla 2005). Logistic regression utilizes a binary dependent variable (sales performance, 0 or 1) to determine whether each customer group, in a specific scenario, comes to the marketplace and makes a final decision to purchase vegetarian/vegan (0) or meat-based products (1). It attempts to predict the probability of this binary outcome. Here, logistic regression offers a powerful tool for predicting our binary target, the named sales performances of vegetarian/vegan or meat-based products. In our analysis, the data set has been separated into inputs and targets. We employed logistic regression to predict the binary target, which is the sales performance, based on the input variables (X1:X8) that have higher P-values extracted from the dataset within the previous five time periods. After feature selection, the extracted independent observations include price, consumer package size, customer group (including age and gender), discount class, brand name, outlet location, partner name, and sale date. To divide the data, 80% of the one million sales performances were used for training, while the remaining 20% was reserved for testing which would evaluate the model's performance on unseen data. In our analysis, 20% of the one million sales performances were set aside as unseen data, meaning they were not used during the training phase.

The classification report on logistic regression (see Table 11.3) demonstrates that the model achieved high precision, recall, and an F1-score for both classes (0 and 1), with an overall accuracy of 91%. The model attempts to predict the decision of new customers entering the market. When we input the information on the relevant features of the new customer, such as the customer group, the consumer package size, and the discount class into the trained model, it will predict the probability of the customer belonging to each class of the binary dependent variable (e.g., sales performance being 0 or 1). To analyze a practical scenario, we considered a male customer aged thirty-five who enters a marketplace at a certain location. He notices a product priced at 9.50 euros, which provides an example for our analysis. The product is not on discount and has a specific brand name. The male customer decides to behave as code 1, meaning he buys the meat-based product (among other goods purchased in the retail store). The logistic regression model predicts that 30.31% of male customers aged thirty-five are to be classified as class 0, meaning that roughly a third of new customers in the chosen demographic target group decide to purchase vegetarian/vegan products, while the future sales for the meat-based products are indeed more than twice the percentage (60.31%) predicted for class 0.

In Fig. 11.8, the Receiver Operating Characteristic Curve (ROC) visualizes the logistic regression of this study, proving the model's accuracy in

Table 11.3 Classification report of logistic regression

Class	Precision	Recall	F1-score	Support
0	0.88	0.86	0.87	63,779
1	0.93	0.94	0.93	120,506
Accuracy			0.91	

classifying the data. The ROC score of 0.98% represents the performance of the model in classifying the positive and negative samples.

In addition, the coefficient of the price variable for the given dataset is approximately − 0.18. This shows that as the average kilogram price of a product increases, the likelihood of a positive sales performance decreases. In our case study, this implies that consumers are indeed price-sensitive, meaning they are influenced by changes in the price of both sustainable and unsustainable products, but only if the price changes by 0.18%. To summarize the predictive analysis, Table 11.4 presents an example of the predictive results of customer sales performance concerning sustainable food products, with a focus on the most popular outlet and a female customer group aged 55–64. The results indicate that customers visiting stores in different locations make different decisions. The accuracy of this prediction is nearly 92%. It is concluded that customers are evolving in their perception of the salient attributes of

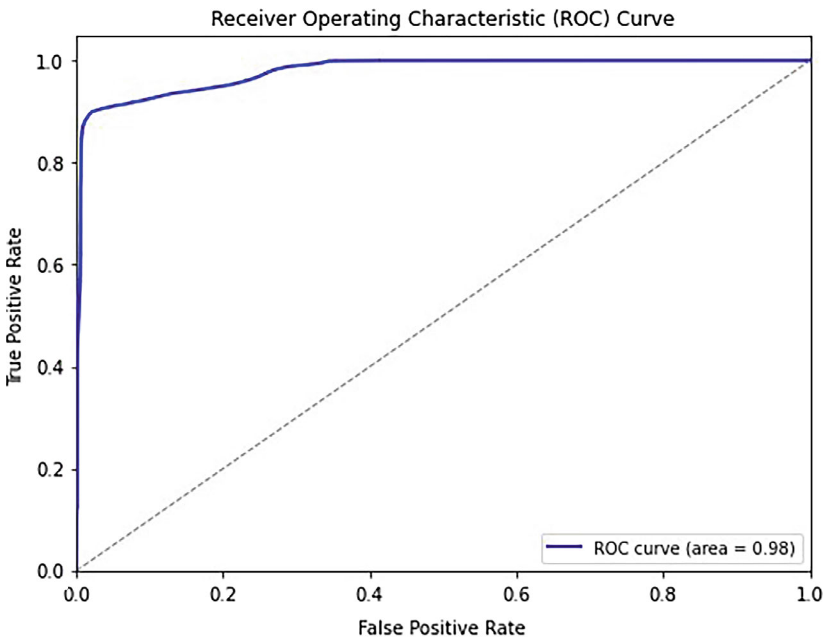

Fig. 11.8 ROC curve of study model

Table 11.4 Predicted sales performance results for female customers aged 55–64 at a popular outlet

	Predicted 0 (%)	Predicted 1 (%)
Actual 0	41.10	3.20
Actual 1	4.65	51.15

products. As can be seen in Table 11.3, the difference between selecting meat-based products and vegetarian/vegan is narrow in this outlet, whereas Table 11.2 indicates that, overall, customers have a greater preference for meat-based products than this difference suggests. Based on the given dataset, this prediction implies that customer behavior varies across outlets based on their perceptions of sustainability, as well as other factors such as prices, age, and income.

11.5 Concluding Discussion

This chapter links three distinct research areas that have not been previously synthesized in a modeling study, namely: (1) optimal pricing decisions for sustainable products, (2) sustainable consumer behavior, and 3) ML methods in marketing. This chapter explores decision-making based on a data-driven model and examines how a retailer, taking a pricing strategy and other features into account, can analyze and predict customers' purchase behavior when choosing sustainable products over unsustainable protein-based substitutes. We applied this model in the context of food-purchasing behavior, providing insights into consumers' sustainable choices and preferences. We rooted our descriptive and predictive analysis model on a wide understanding of consumer behavior, purchase decision-making toward sustainable customer behavior, and pricing decisions to identify factors that impact the choice of a product (green product versus non-green product, or vegetarian/vegan versus meat-based products). We make an important contribution to the modeling of pricing strategies based on ML, aiming to facilitate the understanding of how to utilize big data to predict purchase decisions; thus demonstrating how to facilitate managerial decision-making and impact pricing strategies to nudge consumer purchases toward sustainable consumption and green consumerism (Sharma and Joshi 2017). To summarize, this study underlines the complexity of promoting and predicting sustainable consumer behavior based on historical data. We have demonstrated how descriptive and predictive analytics using ML methods aid in identifying both current and future purchase trends and provided an extensive overview of the antecedents for those purchases to take place. We contribute to marketing research by bringing pricing strategies to the fore and using ML as the basis for empirical analysis of sustainable consumption within the food retail industry.

We suggest that future studies should focus on the behavioral factors that impact consumers' decision-making process when selecting sustainable products over unsustainable ones, even when the former is more costly. In addition to product specifications and consumer demographics, identifying and measuring intentions, attitudes, habits, and, subsequently, customer behavior will give marketers a comprehensive understanding of the company and customers' positions. Thus, data on individual characteristics and attitudes toward sustainability will help marketers achieve more detailed descriptions and predictions using ML methods. Eventually, data analyzed using ML methods may assist decision-makers, and as our case suggests retailers, to persuade customers that it is worth paying a higher price premium for sustainable products (Anselmsson et al. 2014; Ashraf et al. 2017), especially when ecological concerns are highlighted in their marketing.

Acknowledgements Katayoon Pourmahdi acknowledges Åbo Akademi University and the Turun Kauppaopetussäätiö research Grant No. 2023031, Turku, Finland.

References

Icek Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211, 1991. https://doi.org/10.1016/0749-5978(91)90020-T.

James Allsopp. Premium pricing: Understanding the value of premium. *Journal of Revenue and Pricing Management*, 4(2):185–194, 2005. https://doi.org/10.1057/palgrave. rpm.5170138.

Johan Anselmsson, Niklas Vestman Bondesson, and Ulf Johansson. Brand image and customers' willingness to pay a price premium for food brands. *Journal of Product & Brand Management*, 23(2):90–102, 2014. https://doi.org/10.1108/JPBM-10-2013-0414.

Sajid Farooq Ashraf, Cheng Li, and Bilal Mehmood. A study of premium price brands with special reference to willingness of customer to pay. *International Journal of Academic Research in Business and Social Sciences*, 7(7):619–639, 2017. https://doi.org/10.6007/IJARBSS/v7-i7/3126.

Jens Auer and Dominik Papies. Cross-price elasticities and their determinants: A meta-analysis and new empirical generalizations. *Journal of the Academy of Marketing Science*, 48(3):584–605, 2020.

Abdullah M. Baabdullah, Ali A. Alalwan, Emma L. Slade, Rohit Raman, and Khaled F. Khatatneh. Smes and artificial intelligence (ai): Antecedents and consequences of ai-based b2b practices. *Industrial Marketing Management*, 98: 255–270, 2021. https://doi.org/10.1016/j.indmarman.2021.09.003.

Kelvin Balcombe, Iain Fraser, and Aditi Sharma. An econometric analysis of salience theory. *Bulletin of Economic Research*, 73(4):545–554, 2021. https://doi.org/10.1111/boer. 12264.

Frank-Martin Belz and Ken Peattie. *Sustainability Marketing, A Global Perspective*. John Wiley & Sons, second edition, 2014.

Tammo H. A. Bijmolt, Harald J. Van Heerde, and Rik G. M. Pieters. New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, 42(2):141–156, 2005. https://doi.org/10.1509/jmkr.42.2.141.62296.

Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Salience in experimental tests of the endowment effect. *American Economic Review*, 102(3):47–52, 2012. https://doi.org/10.1257/aer.102.3.47.

Ravi Daraboina, Orhan Cooper, and Mehdi Amini. Segmentation of organic food consumers: A revelation of purchase factors in organic food markets. *Journal of Retailing and Consumer Services*, 78:103710, 2024.

Rameshwar Dubey, Angappa Gunasekaran, Stephen J. Childe, David J. Bryde, Mihalis Giannakis, Cyril Foropon, David Roubaud, and Benjamin T. Hazen. Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations. *International Journal of Production Economics*, 226:107599, 2020. https://doi.org/10.1016/j.ijpe.2019.107599.

Christo El Morr and Hossam Ali-Hassan. Descriptive, predictive, and prescriptive analytics. In *Analytics in Healthcare*, pages 31–55. Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-04506-7_3.

Adam Fadlalla. An experimental investigation of the impact of aggregation on the performance of data mining with logistic regression. *Information & Management*, 42(5):695–707, 2005. https://doi.org/10.1016/j.im.2004.04.005.

Paul A. Fishwick. An integrated approach to system modeling using a synthesis of artificial intelligence, software engineering and simulation methodologies. *ACM Transactions on Modeling and Computer Simulation*, 2(4):307–330, 1992. https://doi.org/10.1145/149516.149530.

Donald A. Fuller. *Sustainable Marketing: Managerial-Ecological Issues*. SAGE Publications, 1999. https://doi.org/10.4135/9781452220611.

Andrew Greasley. *Simulating Business Processes for Descriptive, Predictive, and Prescriptive Analytics*. Pearson Education, 2019.

Carl Hempel. A segmentation and characterization of consumers affected by rising food prices. *Appetite*, 194:107192, 2024.

Fabian Herweg and Daniel Müller. A comparison of regret theory and salience theory for decisions under risk. *Journal of Economic Theory*, 193:105226, 2021. https://doi.org/10.1016/j.jet.2021.105226.

Andreas Hinterhuber. Towards value-based pricing—an integrative framework for decision making. *Industrial Marketing Management*, 33(8):765–778, 2004. https://doi.org/10.1016/j.indmarman.2003.10.006.

John A. Howard and Jagdish N. Sheth. *The Theory of Buyer Behavior*. John Wiley & Sons, New York, 1969.

Wang Huiru, Song Jinhui, Feng Jianying, Fan Huiru, Zhao Zhijian, and Mao Weisong. An agent-based modeling and simulation of consumers' purchase behavior for wine consumption. *IFAC-PapersOnLine*, 51(17):843–848, 2018. https://doi.org/10.1016/j.ifacol.2018.08.089.

Håkan Håkansson, Debbie Harrison, and Alexandra Waluszewski. *Rethinking Marketing: Developing a New Understanding of Markets*. Wiley, 2005. Available at: https://books.google.fi/books?id=42zYx0E8OWsC.

Paul Ingenbleek. The theoretical foundations of value-informed pricing in the service-dominant logic of marketing. *Management Decision*, 52(1):33–53, 2014. https://doi.org/10.1108/MD-03-2013-0128.

Paul Ingenbleek. Price strategies for sustainable food products. *British Food Journal*, 117(2):915–928, 2015. https://doi.org/10.1108/BFJ-02-2014-0066.

Yi-Kai Juan, Yuan-Hsiang Hsu, and Xiaoling Xie. Identifying customer behavioral factors and price premiums of green building purchasing. *Industrial Marketing Management*, 64:36–43, 2017. https://doi.org/10.1016/j.indmarman.2017.03.004.

Brendan J. Keegan, Ana I. Canhoto, and Dorothy A. Yen. Power negotiation on the tango dancefloor: The adoption of ai in b2b marketing. *Industrial Marketing Management*, 100:36–48, 2022. https://doi.org/10.1016/j.indmarman.2021.11.001.

Rahul Kler, Galal Elkady, Kailash Rane, Aman Singh, Md. Shahin Hossain, Deepak Malhotra, Suman Ray, and Kamal K. Bhatia. Machine learning and artificial intelligence in the food industry: A sustainable approach. *Journal of Food Quality*, 2022:1–9, 2022. https://doi.org/10.1155/2022/8521236.

Philip Kotler and Gary Armstrong. *Principles of Marketing*. Pearson Education, 13th edition, 2010.

Philip Kotler and Kevin Lane Keller. *Marketing Management*. Pearson, 15th edition, 2016.

Richard A. Lancioni. A strategic approach to industrial product pricing: The pricing plan. *Industrial Marketing Management*, 34(2):177–183, 2005. 10.1016/j. indmarman.2004.07.015.

Stephan M. Liozu and Andreas Hinterhuber. Pricing orientation, pricing capabilities, and firm performance. *Management Decision*, 51(3):594–614, 2013.

Stephan M. Liozu, Andreas Hinterhuber, Richard Boland, and Sheri Perelli. The conceptualization of value-based pricing in industrial firms. *Journal of Revenue and Pricing Management*, 11(1):12–34, 2012. https://doi.org/10.1057/rpm.2011.34.

Khursheed Lone and Shabir Ahmad Sofi. Descriptive, predictive, and prescriptive analytics in healthcare. In *Data Science for Effective Healthcare Systems*, pages 89–103. Chapman and Hall/CRC, 2022.

Pan Ma, Chao Zhang, Xiao Hong, and Huan Xu. Pricing decisions for substitutable products with green manufacturing in a competitive supply chain. *Journal of Cleaner Production*, 183:618–640, 2018. https://doi.org/10.1016/j.jclepro.2018.02.152.

Alfred Marshall. *Principles of Economics*. The Online Library of Liberty, 2011.

Laura Martínez-Carrasco Martínez, Margarita Brugarolas Mollá-Bauzá, and José Antonio Sánchez-Zapata. In search of a sustainable alternative for meat production: Understanding the purchase intention of meat from transhumance origin. *Renewable Agriculture and Food Systems*, 38(e20), 2023. https://doi.org/10.1017/S1742170523000121.

Zhiyi Meng, Na Zhao, Bin Shen, and Chang Zhai. Optimal pricing strategy for green products under salience theory. *Economic Research-Ekonomska Istraživanja*, 35(1):2960–2983, 2022. https://doi.org/10.1080/1331677X.2021.1985572.

Jacquelyn A. Ottman. *Green Marketing*. NTC Business Books, 1993. https://cir.nii.ac.jp/crid/1130282272598906368.

Jeannette Paschen, Ulrich Paschen, Erol Pala, and Jan Kietzmann. Artificial intelligence (ai) and value co-creation in b2b sales: Activities, actors and resources. *Australasian Marketing Journal*, 29(3):243–251, 2021. https://doi.org/10.1016/j.ausmj.2020.06.004.

Agriani H. Sadeli, Tomy Perdana, Yosini Deliana, and Bhakti S. Onggo. Consumers' purchase behavior in short food supply chains using social commerce in indonesia. *Journal of Cleaner Production*, 386:135812, 2023. https://doi.org/10.1016/j.jclepro.2022.135812.

Venkatesh Shankar. How artificial intelligence (ai) is reshaping retailing. *Journal of Retailing*, 94(4):vi–xi, 2018. https://doi.org/10.1016/S0022-4359(18)30076-9.

Aasha Sharma and Seema Joshi. Green consumerism: Overview and further research directions. *International Journal of Process Management and Benchmarking*, 7(2): 206, 2017. https://doi.org/10.1504/IJPMB.2017.083106.

Jagdish N. Sheth. *Models of Buyer Behavior: Conceptual, Quantitative, and Empirical*. Marketing Classics Press, 2011.

Jagdish N. Sheth and Atul Parvatiyar. The evolution of relationship marketing, 1995. Unpublished work or replace with publication details if available.

Firouzeh Taghikhah, Tatiana Filatova, and Alexey Voinov. Where does theory have it right? a comparison of theory-driven and empirical agent based models. *Journal of Artificial Societies and Social Simulation*, 24(2), 2021. https://doi.org/10.18564/jasss.4573.

A. Caroline Tynan and Jennifer Drayton. Market segmentation. *Journal of Marketing Management*, 2(3):301–335, 1987.

John K. Whitaker. The distribution theory of marshall's principles. In A. Asimakopulos, editor, *Theories of Income Distribution*, pages 105–132. Springer Netherlands, 1988. https://doi.org/10.1007/978-94-009-2661-5_5.

Lydia Zepeda and David Deal. Organic and local food consumer behaviour: Alphabet theory. *International Journal of Consumer Studies*, 33(6):697–705, 2009. https://doi.org/10.1111/j.1470-6431.2009.00814.x.

Qing Zhang and Yini Zheng. Pricing strategies for bundled products considering consumers' green preference. *Journal of Cleaner Production*, 344:130962, 2022. https://doi.org/10.1016/j.jclepro.2022.130962.