Research article

Tomato Consumer Behavior in Hino City, Japan: Application of Multiple Correspondence and K-means Cluster Analyses

HETTIGE SAMITHA LAKSHAN GUNASEKARA*
Graduate School of International Food and Agricultural Studies,
Tokyo University of Agriculture, Tokyo, Japan
Email: 13823001@nodai.ac.jp

NINA N. SHIMOGUCHI
Faculty of International Agriculture and Food Studies, Tokyo University of Agriculture, Japan

RIKA TERANO
Faculty of International Agriculture and Food Studies, Tokyo University of Agriculture, Japan

RAMADHONA SAVILLE
Faculty of International Agriculture and Food Studies, Tokyo University of Agriculture, Japan

KATSUMORI HATANAKA
Faculty of International Agriculture and Food Studies, Tokyo University of Agriculture, Japan

Received 31 December 2023 Accepted 23 December 2024 (*Corresponding Author)

Abstract Japan is currently facing multiple challenges related to an aging society, including labor shortages in various sectors, especially in agriculture. Despite government support, new farmers struggle to discern and understand market trends, changing consumer behavior, and seasonal changes. This study focuses on tomatoes and aims to assist new farmers and producers in making strategic decisions by utilizing data science techniques to understand local market consumer behavior and trends. A questionnaire survey was used to identify consumer behavior based on demographic characteristics, consumer preferences, seasonal buying and consumption behaviors, spending price per purchase, and purchasing channels. We collected 316 valid data points using a questionnaire survey deployed in face-to-face interactions across multiple locations as well as online collection methods. Data was collected from May 01, 2022, to June 30, 2022. Multiple correspondence analysis and K-means cluster analysis were used to determine six consumer segments - convenient supermarket shoppers, premium salad shoppers, price-conscious tomato enthusiasts, diverse tomato tasters, quality-conscious shoppers, and gourmet tomato shoppers. The analyses revealed the unique characteristics of tomato purchasing and consumption behaviors, including variations in tomato type, consumption type, season, demographic factors, and preferred purchasing channels. This study further explored the factors influencing tomato purchasing decisions within each segment by integrating data-driven decision-making principles. This approach allows for actionable insights facilitated by understanding local consumers while empowering new tomato farmers to make informed decisions.

Keywords agribusiness, agriculture, consumer segmentation, data science, data-driven decision-making

INTRODUCTION

Tomatoes play an important role in the Japanese agriculture sector due to several factors. Tomatoes represent a high-value crop with a significant economic impact due to their substantial annual production in the year 2022 volume of 678,500 metric tons and extensive land area of 10,900 hectares dedicated to their cultivation (Kiyota et al., 2017; MAFF, 2023). The agricultural output value of tomatoes is 230.2 billion JPY, which accounts for approximately 10% of the total fresh vegetable agricultural output value of 2.2 trillion JPY by 2022 (Statistics of Japan, 2024). This makes tomatoes the largest economic contributor among all fresh vegetables, underscoring their critical importance in Japan's fresh vegetable sector. However, tomato farming in Japan faces challenges similar to other agricultural commodities because of the aging society and its decreasing labor force (MAFF, 2018). Tomato production often requires considerable manual labor for trellising, pruning, and harvesting. In response, the Japanese government introduced smart hydroponic greenhouses to reduce labor requirements, make efficient use of resources, and enhance productivity (MAFF, 2018). Smart hydroponic farming combines hydroponic cultivation techniques with advanced technologies, such as Internet of Things (IoT) sensors, automated nutrient delivery systems, and climate control, to enhance productivity and resource efficiency in a controlled environment (MAFF, 2018). Therefore, smart hydroponic farming has become increasingly popular in tomato farming to better manage the industry and aging workforce. Although the government assists new tomato farmers by providing subsidies, training, and access to smart agricultural technologies, such as hydroponic systems and IoT-enabled greenhouses, it is crucial to understand consumer behavior to create products that satisfy market demands (Nishimura, 2021). New farmers often need more experience and a deeper understanding of the local market dynamics. Their lack of experience makes it challenging to assess and react appropriately to customer trends and preferences. Another challenge is that each local tomato market has distinct features and customer habits (Olivera et al., 2011; Jąder, 2020). Therefore, understanding local tomato consumer behavior can be challenging for new farmers. To overcome these challenges, new farmers require access to comprehensive data on consumer behavior that encompasses insights into buying patterns, consumption habits, and personal preferences.

Puccinelli et al. (2009) highlight the pivotal importance of understanding consumer behavior in shaping market demand, encompassing insights into buying patterns, consumption habits, individual preferences, motivations, and decision-making processes. A comprehensive understanding of vegetable consumption is essential for agricultural producers to effectively meet consumer market demands. Various factors influence vegetable consumption, including personal, living and eating habits, and socio-demographic conditions (Olivera et al., 2011). Demographic factors, including consumer income, education, job status, family type, and place of residence, are crucial determinants of vegetable consumption (Jąder, 2020). The complexity of comprehending consumer behavior and its impact on vegetable consumption arises from these diverse influences. Numerous studies have explored the factors influencing vegetable consumption and provided insights into consumer behavior and food-choice motives. Consumer preferences play a significant role in shaping vegetable purchasing decisions, including monetary factors such as price and non-price factors such as product quality, shelf life, place of purchase, origin of vegetables, and awareness of safety concerns, all of which contribute to consumers purchasing decisions regarding fresh vegetables (Šebjan and Tominc, 2016; Singh and Raj, 2018). Furthermore, previous research has emphasized the importance of understanding local consumer behavior and its links to personal characteristics, such as eating habits, shopping preferences, and cooking practices (Mirosa and Lawson, 2012). In the Japanese market, consumers are sensitive to both price and quality. Consumers maintain a positive attitude towards vegetables, particularly tomatoes, as well as organic production and freshness which hold significant value (Narine et al., 2015; Yang et al., 2021). Segmentation is integral to consumer behavior analysis, enabling producers and stakeholders to effectively target specific consumer groups (Troy and Bogue, 2015). Categorical variables commonly encountered in consumer surveys represent qualitative characteristics or attributes useful for classifying or grouping individuals (Somya et al., 2022). Researchers can further understand the relationships and patterns in categorical data variables by applying data mining methods, including multiple correspondences and cluster analysis techniques (Causse et al., 2010; Mori et al., 2016).

OBJECTIVE

The primary goal of this research is to utilize data mining techniques to provide actionable insights to new farmers to support strategy planning and decision-making in the local tomato market by identifying consumer behavior. This study specifically aims to identify consumer behavior in the local tomato market through consumer segmentation using data mining techniques such as multiple correspondence analysis (MCA) and K-means cluster analysis.

METHODOLOGY

This case study focuses on N-farm, a new smart greenhouse tomato producer located in Hino City, Tokyo, which began production in 2019. N-farm was selected because it is a new smart greenhouse tomato producer, making it an ideal subject for studying the challenges and opportunities faced by new farmers. The primary target of N-farm is the local, Hino City tomato consumer market. This focused market provides data and an analysis of consumer behavior and preferences in a localized setting, making it more relevant and applicable to new farmers. N-farm produces multiple types of large, medium, and cherry tomatoes, and the diversity in production and sales allows for a comprehensive analysis of consumer preferences across these different tomato varieties. N-farm uses multiple distribution channels, including direct farmers shops (DFS), unmanned vegetable sale shops (UVSS), and online market channels, allowing the exploration of how different marketing and distribution channels affect consumer behavior and farm success. Farmers distribute their products under the brand name of N-farm. As a new farm in the tomato agribusiness targeting local consumers, N-farm struggles to understand local consumer behavior due to its limited experience. This lack of experience creates challenges in identifying critical consumer preferences such as preferred tomato varieties, purchasing channels, spending habits, and understanding seasonal trends in tomato buying and consumption. Therefore, to fulfill the primary goal of this study, we gathered information from consumers using multiple sources and locations, focusing on market channels where the products sold were centered on a newly established farmer, with the primary goal of identifying consumer behavior to better understand the local tomato market.

Data Collection

Primary data were collected using a questionnaire survey. The questionnaire survey was conducted from May 01, 2022, to June 30, 2022, using a combination of face-to-face and online methods. Consumer responses were collected from two direct farmers' shops (DFS), and an unmanned vegetable sales shop (UVSS) located in Hino City. The study placed information on the tomato shelf in the DFSS and UVSS outlets and used QR codes to facilitate online participation and data collection. Furthermore, data for the questionnaire survey were also gathered using social network services (SNS) and the N-farm's business website. This multi-method approach was used to ensure a diverse and representative data set. The questionnaire used identical survey questions and standardized instructions to ensure consistency across both methods. During face-to-face data collection, surveyors were available to clarify any questions as needed, whereas online participants completed the survey independently. These differences in data collection may have influenced response patterns. However, both methods were deemed equally valid for the study due to the standardized questionnaire design and consistent implementation steps across both approaches. This survey was not piloted before deployment due to monetary, human resource, and time constraints. Before collecting responses, all participants were informed of the purpose of the study, their consent was obtained, and their privacy and confidentiality were protected.

Questionnaire Design

The questionnaire survey was designed to cover all relevant aspects of consumer demographics and behavior to ensure that the data accurately captured consumer characteristics and behavior. The questionnaire included 14 questions regarding consumer demographics, tomato purchasing patterns, and consumption behavior. Participants were asked about their age, gender, family type, and annual household income. The survey included questions on place of purchase, reasons for selecting a particular place, and spending per purchase of tomatoes. Seasonal tomato purchasing behavior was explored by categorizing tomato types, large, medium, and cherry, and consumption preferences, salad, cooked dishes, and lunch boxes, across the spring, summer, autumn, and winter seasons.

Data Preprocessing and Analysis

Before the analysis, 470 responses were collected, with 65% obtained through face-to-face interviews and 35% submitted online via QR codes, SNS, and the N-farm website. Data cleaning excluded 33% of the dataset due to missing responses (70% of the missing values related to annual household income) and invalid answers from respondents who did not follow instructions. The exclusion of 33% of responses highlights areas for improvements in future studies by conducting pilot studies and implementing measures to ensure complete responses, particularly for sensitive questions such as annual household income. The resultant 316 responses were used for the analysis. Descriptive statistics were used to summarize and present the main features of the dataset. The MCA method investigates the relationships between categorical variables, revealing consumer behavior patterns. K-means cluster analysis was used to identify distinct groups of consumers based on their shared characteristics. The statistical method of the chi-square test was employed to evaluate differences among multiple variables and to assess the significance of the identified clusters. Data analysis was conducted using the RStudio (version 4.2.2) statistical program.

Multiple Correspondence Analysis

MCA is a statistical technique used to analyze patterns and relationships in categorical data (Greenacre and Blasius, 2006). It transforms the original variables into a low-dimensional space, making complex relationships easier to visualize and interpret. The procedure begins by creating a contingency table with a correspondence matrix that represents the relationships between the categories and observations. The matrix is subjected to singular value decomposition, yielding principal coordinates that identify patterns and associations (Greenacre and Blasius, 2006; Hoffman and De Leeuw, 1992). In our study, demographic and tomato buying and consumption behavior data were used to perform the MCA analysis to determine consumer behavior. We used MCA to analyze the questionnaire responses and assign numerical values to the categorical data variables noted above.

K-means Cluster Analysis

K-means clustering is an iterative algorithm that partitions a dataset into K clusters based on the similarities between data points. This algorithm aims to minimize the within-cluster sum of squares (WCSS), which is the sum of the squared distances between each data point and the centroids of its assigned cluster (Lloyd, 1982). Eq. (1) presents the key equations for the K-means clustering algorithm.

$$WCSS = \sum_{k=1}^{K} \sum_{i=1}^{n_k} ||x_i - m_k||^2 \tag{1}$$

Where, K is the total number of clusters, and n_k is the number of data points in a cluster. The x_i is the i-th data point, and m_k is the centroid of the cluster. In the K-means cluster analysis, the number of clusters must be decided before the analysis is performed. The “elbow method” is used to determine the optimal number of K-means clusters. The elbow method involves running the K-means algorithm for different values of K and plotting the corresponding WCSS. A scree plot was used to visualize and identify the optimal number of clusters. The goal is to identify the “elbow point” or sharp bend in the plot, where the rate of decrease in WCSS encounters a sharp bend. This point indicates a significant drop in the variance, which explains the optimal number of clusters. Applying this method ensures that the clustering results are meaningful and interpretable, representing the underlying data structure (Syakur et al., 2018; van de Velden et al., 2017).

RESULTS AND DISCUSSION

The demographic characteristics of the participants are presented in Table 1. The average age of participants was 51 years. The largest age group was between 40 and 59 years, accounting for 49.4% of the respondents, followed by those aged 60 to 79 years. The age group between 20 to 39 years accounts for 20.9%, while the age groups below 19 years and above 80 years represent 1.2% and 1.6% of the respondents, respectively. Regarding gender distribution, 73% of respondents were female, and 27% of respondents were male. When considering family type, 59% were married with children, 24% were married with no children, and 17% were single. When analyzing annual household income, an income of more than 7.1 million JPY was reported by 38% of respondents, followed by 24% for 5.1 to 7 million JPY, 27% for 3.1 to 5 million JPY, and 11% for 2.1 to 3 million JPY. No respondents reported income below 2 million JPY. Furthermore, by combining the annual household income of respondents earning more than 7.1 million JPY or 5.1 to 7 million JPY, 62% of respondents had an annual household income exceeding the national average of 5.25 million JPY (MIAC, 2023).

Table 1 Descriptive statistic results for demographic variables (n=316)

Data variable	Percentage (%)
Age	
0 to 19	1.2
20 to 39	20.9
40 to 59	49.4
60 to 79	25.9
80 to 99	1.6
Gender	
Male	27.0
Female	73.0
Family type	
Single	17.0
Married with no children	24.0
Married with children	59.0
Annual household income	
Below 2 million JPY	0.0
2.1 to 3 million JPY	11.0
3.1 to 5 million JPY	27.0
5.1 to 7 million JPY	24.0
More than 7.1 million JPY	38.0

Source:2022 Survey conducted by author

Results for Multiple Correspondence Analysis

The MCA results are shown in the biplot in Fig. 1. The eigenvalues represent the variance explained by each dimension extracted from the analysis. The first dimension, Dim 1, had an eigenvalue of 0.242, representing 5.18% of the total variance. The second dimension, Dim 2, had an eigenvalue of 0.204, which explains 4.38% of the total variance. The cumulative percentage of the variance indicates the cumulative contribution from each dimension. The Dim 1 and Dim 2 dimensions explain 9.56% of the total variance. A biplot is a graphical representation and a powerful tool for interpretation that combines variables and individual coordinates in a single plot, providing a comprehensive visualization of the relationships between variables and individuals in an MCA. As shown in Fig. 1, the biplot displays the relationships between the variables and individuals along the extracted dimensions. Data points represent variables based on their correlations with each dimension, indicating the direction and strength of their associations. For Dim 1, the variables of seasonal eating types of tomatoes in the lunch box, salad, cooked version, and tomato buying place of DFS and UVSS were positively associated, suggesting a link between the variables and consumer buying behavior factors. Meanwhile, variables of spring, autumn, and summer seasonal tomato buying types of cherry, medium, and large size tomatoes and a separate group of undecided seasonal tomato buying characteristics point in the direction of Dim 2, implying an association with preferences related to seasonal tomato buying behavior.

Fig. 1 Multiple correspondence analysis biplots for tomato consumer behavior analysis

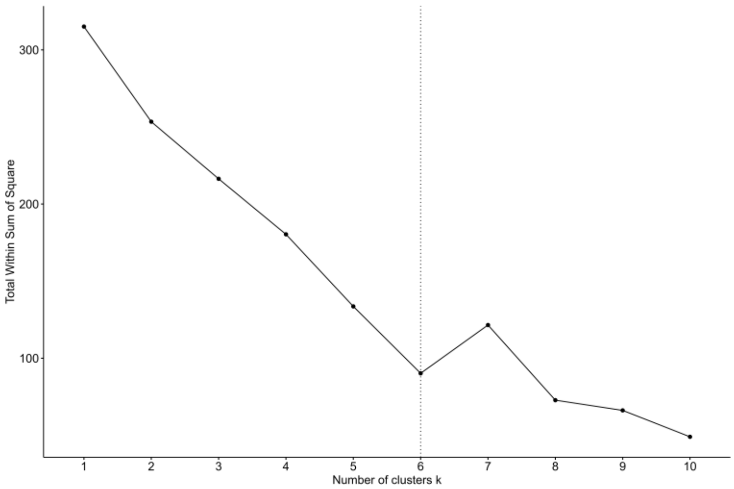

Fig. 2 Optimal number of clusters using the elbow method

Consumer Segment Profiling

After completing the MCA, the transformed data were used for K-means clustering to identify segments. K-means clustering is a quantitative method for categorizing and grouping similar consumers based on their coordinates in an MCA-derived space. Using the elbow method, as shown in Fig. 2, this study identified six as the optimal number of clusters for K-means cluster analysis. The p-values obtained from the chi-square analysis test results emphasized the statistical significance of factors influencing tomato-related behaviors among the six consumer segment characteristics. A significant difference was identified in the purchasing location (p-value < 0.001) and ranged across direct farmers' shops (DFS), unmanned vegetable sale shops (UVSS), supermarkets, and online marketing channels. The reason for purchase location selection was significant (p-value < 0.001), which can be explained by the fact that consumers were strongly motivated to choose specific purchasing locations, influenced by price, quality, and convenience. The spending price per purchase was significant (p-value < 0.001) and reveals variations in the amount consumers are willing to spend on tomatoes, suggesting differences in price sensitivity and budget allocation among segments.

Regarding seasonal tomato purchasing behavior (spring, summer, autumn, and winter), purchasing by tomato type (large, medium, and cherry) and overall consumption behavior of tomatoes (salad, cooked dishes, and lunch boxes) were significant (p-value < 0.001). Preferences for different types of tomatoes varied significantly by season, and there were notable differences in tomato consumption across seasons. These differences were statistically significant (p-value < 0.001). They were identified through a chi-square test analysis, confirming the validity of the segmentation approach and highlighting the complex relationship between behaviors and influencing factors.

Consumer Segments

Six unique consumer segments were identified using K-means cluster analysis. These segments provide valuable insights into the dataset's diverse consumer behaviors and preferences. The characteristics of each identified consumer segment are explained below.

Convenient supermarket shoppers (n=159, 50.3%): Convenient supermarket shoppers have an annual household income of more than 7.1 million JPY, and this segment prefers the convenience of shopping at supermarkets, 55.5%, with 47.8% of the segment reporting that they paid between 101 and 300 JPY per tomato purchase, 89.3% of the segment preferred cherry tomatoes throughout the year. Before purchasing tomatoes, 70.4% of the segment checked their prices.

Premium salad shoppers (n=100, 31.6%): Premium salad shoppers have an annual household income between 5.01 to 7 million JPY, and they prefer to buy tomatoes from supermarkets, 29% of respondents. Quality was an essential factor in this purchasing location for 41% of the segment, and 52% of the segment reported that they paid between JPY 501 and 700 per tomato purchase. The segment has a year-round preference for medium-sized tomatoes, with 88% of consumers enjoying tomatoes in salads during summer and 50% of segment respondents checking tomato prices before purchasing.

Price-conscious tomato enthusiasts (n=26, 8.2%): Price-conscious tomato enthusiasts have an annual household income between 5.01 to 7 million JPY, and they prefer to buy tomatoes from DFS. In this segment, 50% paid between 101 and 300 JPY per purchase. Seasonal preferences are more neutral, but with a strong preference for tomato salads in spring, accounting for 50% of consumption. A notable feature is their purchasing behavior, with 61.5% of this segment checking the price before purchasing tomatoes.

Diverse tomato tasters (n=13, 4.1%): Diverse tomato tasters have an annual household income between 5.01 to 7 million JPY, and they show a more diverse approach to tomato shopping. In this segment, 46.1% buy tomatoes from both DFS and supermarkets. Additionally, 61.5% paid between 301 to 500 JPY per tomato purchase. Seasonal preferences reflect a desire for variety, with a preference for cherry and large-sized tomatoes in spring and cherry and medium-sized tomatoes in summer and autumn. During the winter months, only cherry tomatoes were purchased. Most of this segment did not consider any tomato variables before their purchase.

Quality conscious shoppers (n=11, 3.48%): Quality-conscious shoppers have an annual household income between 3.01 to 5 million JPY, and they prefer to buy tomatoes from DFS and supermarkets. They preferred the chosen locations because the tomatoes were of high quality. 54.5% of the segment spends between 301 to 500 JPY per tomato purchase. Cherry tomatoes are preferred in all seasons. Additionally, 54.5% of the segment enjoy tomatoes in lunchboxes, cooked dishes, and salads. Within the segment, 63.5% check the color, and 54.5% check the price before purchasing tomatoes.

Gourmet tomato shoppers (n=7, 2.2%): Gourmet tomato shoppers have an annual household income of more than 7.1 million JPY. The segment primarily shops at supermarkets, 42.8%, and 71.4% are willing to spend between JPY 301 to 500 JPY for each tomato purchase. Seasonal preferences include cherry and large tomatoes, and 85.7% of the segment preferred to consume tomatoes in lunch boxes. 71.4% of the segment stated that they checked the price before purchasing tomatoes.

The analysis of tomato consumers' demographic characteristics provides N-farm with an explanation of tomato purchases in Hino City. The findings revealed that the main tomato consumers are females aged 40 to 59 years, married with children, and with a household income above the national average. This demographic profile highlights a key target audience for N-farm's marketing efforts. However, treating consumers as a homogeneous group based on demographics alone is insufficient to gain deeper market insights. Therefore, to address these limitations, consumer preferences, consumption habits, and purchasing behaviors were further analyzed by integrating them with the demographic characteristics using MCA and K-means cluster analysis. This method identified six consumer segments, convenient supermarket shoppers, premium salad shoppers, price-conscious tomato enthusiasts, diverse tomato tasters, quality-conscious shoppers, and gourmet tomato shoppers, providing a more comprehensive understanding of the interdependence among the categorical variables of consumer demographic characteristics and seasonal tomato buying and consumption behaviors. Additionally, they revealed heterogeneity at the cluster level, offering deeper insights into the diversity and relationships within the data set. Among the identified consumer segments, cherry tomatoes emerged as the preferred variety across all segments, reflecting a universal appeal, while medium-sized tomatoes showed a demand influenced by seasonal trends. Based on the findings, the results reflect diverse consumers by identifying unique characteristics and preferences, allowing N-farm to offer valuable insights for strategic decision-making that can substantially enhance its operations and market position by following primary strategic takeaways and practical applications. This can be used to move beyond one-size-fits-all approaches and tailor their offerings to specific consumer clusters, thereby increasing the relevance and appeal of products. For example, N-farm needs to allocate a significant portion of the greenhouse space to cultivate cherry tomatoes based on the results of this study. When considering tomato distribution channels, N-farm can supply DFS with good quality and affordable prices while ensuring a strong presence in supermarkets to cater to the preferences of convenient supermarket shops in the local area for marketing and sales strategies. Furthermore, pricing is an important strategy to compete with other tomato producers for price-conscious tomato enthusiasts. N-farm can embrace dynamic pricing strategies, allowing adjustments based on fluctuating seasonal demand, promotions, and consumer discounts. Establishing strategic partnerships between supermarkets and DFS locations will ensure a strong market. Furthermore, it is suggested that N-farm needs to implement consistent quality control measures by using technological integration throughout the cultivation process to meet the expectations of quality-conscious consumers, and so that farmers remain up to date on changing consumer preferences and adjusting cultivation strategies accordingly. Therefore, the findings of this study provide insight into the local tomato market by identifying six distinct consumer clusters and explaining valuable information that can significantly inform strategic decision-making for farmers of N-farm, particularly consumer segmentation, price differentiation, market channel analysis, and seasonal planning.

CONCLUSION

Tomatoes are a critical food product for Japanese agriculture because of their economic importance, large output volumes, and large cultivated areas. However, the sector faces challenges from an aging population and shrinking labor force, requiring significant physical labor. Smart hydroponic farming has also gained popularity as a long-term option owing to its controlled environment and reduced physical labor requirements. New tomato producers frequently fail to understand and respond to local consumer behavior owing to a lack of experience and market expertise. The main goal of this study was to provide actionable insights derived from data mining techniques to new farmers with strategy planning and decision-making for the local tomato market by identifying consumer behavior. The findings highlight the importance of consumer behavior analyses in aligning production with market demand. This study helped new farmers, N-farm, to understand the demographic characteristics of the local tomato market of Hino City, Tokyo, that primary tomato consumers are predominantly middle-aged to older females who are married with children and generally have incomes higher than the national average. Six consumer segments were identified, each with their characteristics and preferences. Findings from this research will help new farmers, like those at N-farm, organize their agricultural and marketing activities more strategically. Furthermore, based on the findings of this study, focusing on cherry tomatoes allows N-farm to cater to clients' desires for quality and convenience while having a diversified product line that appeals to a larger market. Price sensitivity should be addressed by providing inexpensive yet high-quality tomatoes. Reaching different consumer demographics requires effective marketing and distribution tactics such as targeted promotions and sales channels. Furthermore, using innovative technologies in farming and distribution processes can improve efficiency and product quality, ensuring that N-farm remains competitive and meets the changing needs of consumers.

Given the technical nature of the methodologies employed in this study, it is critical to develop services that bridge the gap between complicated data analyses and practical applications. As a suggestion, agricultural cooperatives can participate, combine resources to access modern data analysis services, and share insights with their members. Furthermore, creating user-friendly software solutions would enable farmers to enter data and obtain insights without fully understanding the underlying algorithms. We suggest that future research should include longitudinal surveys to examine changes in consumer segments and purchase patterns, as well as an analysis of how seasonal variations, economic changes, and technological improvements affect consumer behavior. Continuous monitoring and analyses will offer farmers a more dynamic grasp of consumer trends, allowing them to respond quickly to dynamic market trends.

ACKNOWLEDGEMENTS

The authors express gratitude to N-farm owners and Tokyo University of Agriculture students, who supported their active involvement in planning and execution, and respondents who supported this research with valuable input in shaping the study's outcomes.

REFERENCES

Causse, M., Friguet, C., Coiret, C., Lépicier, M., Navez, B., Lee, M., Holthuysen, N., Sinesio, F., Moneta, E. and Grandillo, S. 2010. Consumer preferences for fresh tomato at the European scale, A common segmentation on taste and firmness. Journal of Food Science, 75 (9), S531-S541, Retrieved from DOI https://doi.org/10.1111/j.1750-3841.2010.01841.x

Greenacre, M. and Blasius, J. 2006. Multiple correspondence analysis and related methods. Behavioral Sciences, Bioscience, Mathematics & Statistics, Chapman and Hall/CRC, ISBN 9780429141966, USA.

Hoffman, D.L. and Leeuw, J.D. 1992. Interpreting multiple correspondence analysis as a multidimensional scaling method. Marketing Letters, 3, 259-272, Retrieved from DOI https://doi.org/10.1007/BF00994134

Jąder, K. 2020. Regional diversification in the consumption of fresh fruit, vegetables and their processed products in Poland. Annals of the Polish Association of Agricultural and Agribusiness Economists, XXII (1), 126-136, Retrieved from DOI https://doi.org/10.5604/01.3001.0013.9605

Kiyota, K., Yoshimitsu, M., Satsuki-Murakami, T., Akutsu, K., Kajimura, K. and Yamano, T. 2017. Detection of the tomato allergen Sola l 1 and evaluation of its reactivity after heat and papain treatment. Food and Agricultural Immunology, 28 (6), 1450-1459, Retrieved from DOI https://doi.org/10.1080/09540105.2017.1347914

Lloyd, S. 1982. Least squares quantization in PCM. IEEE Transactions on Information Theory, 28 (2), 129-137, Retrieved from DOI https://doi.org/10.1109/tit.1982.1056489

Ministry of Agriculture, Forestry and Fisheries (MAFF). 2018. Situation of greenhouse horticulture. Ministry of Agriculture Forestry and Fisheries, Retrieved from URL https://www.maff.go.jp/e/policies/agri/attach/pdf/index-16.pdf

Ministry of Agriculture, Forestry and Fisheries (MAFF). 2023. Acreage, yield, and shipment volume of designated vegetables (spring, summer, and autumn vegetables) for 2022. Ministry of Agriculture, Forestry and Fisheries, Retrieved from URL https://www.maff.go.jp/j/tokei/kekka_gaiyou/sakumotu/sakkyou_yasai/r5/shitei_yasai/index.html

Ministry of Internal Affairs and Communications (MIAC). 2023. Statistics bureau of Japan. Ministry of Internal Affairs and Communications, Retrieved from URL https://www.stat.go.jp/english/data/kakei/156n.html

Mirosa, M. and Lawson, R. 2012. Revealing the lifestyles of local food consumers. British Food Journal, 114 (6), 816-825, Retrieved from DOI https://doi.org/10.1108/00070701211234345

Mori, Y., Kuroda, M. and Makino, N. 2016. Joint Dimension Reduction and Clustering. In: Mori, Y., Kuroda, M. and Makino, N. (Eds.) Nonlinear Principal Component Analysis and Its Applications, SpringerBriefs in Statistics (JSSRES). Springer, Singapore, 57-64, Retrieved from DOI https://doi.org/10.1007/978-981-10-0159-8_6

Narine, L., Ganpat, W. and Seepersad, G. 2015. Demand for organic produce, Trinidadian consumers' willingness to pay for organic tomatoes. Journal of Agribusiness in Developing and Emerging Economies, 5 (1), 76-91, Retrieved from DOI https://doi.org/10.1108/jadee-04-2013-0015

Nishimura, T. 2021. The effect of greenhouse pollination methods on consumers' willingness to pay for tomatoes in Japan. Journal of Agricultural and Applied Economics, 53 (2), 186-208, Retrieved from DOI https://doi.org/10.1017/aae.2020.33

Olivera, D.F., Vina, S.Z., Ferreyra, R.M., Mugridge, A., Mascheroni, R.H. and Chaves, A.R. 2011. Integration of an objective approach focused on the consumer to evaluate the quality of processed brussels sprouts. Food and Nutrition Sciences, 2 (9), 1011-1017, Retrieved from DOI https://doi.org/10.4236/fns.2011. 29135

Puccinelli, N., Goodstein, R.C., Grewal, D., Price, R., Raghubir, P. and Stewart, D. 2009. Customer experience management in retailing, Understanding the buying process. Journal of Retailing, 85 (1), 15-30, Retrieved from DOI https://doi.org/10.1016/j.jretai.2008.11.003

Šebjan, U. and Tominc, P. 2016. Young customers' organoleptic assessment of tomatoes with different geographic origins, A preliminary study. British Food Journal, 118 (4), 871-884, Retrieved from DOI https://doi.org/10.1108/bfj-07-2015-0259

Singh, K. and Neeraj. 2018. A study on factors affecting consumers decision to purchase vegetables. International Journal of Current Microbiology and Applied Sciences, 7 (2), 1211-1222, Retrieved from DOI https://doi.org/10.20546/ijcmas.2018.702.149

Somya, R., Winarko, E. and Priyanta, S. 2022. A hybrid recommender system based on customer behavior and transaction data using generalized sequential pattern algorithm. Bulletin of Electrical Engineering and Informatics, 11 (6), 3422-3432, Retrieved from DOI https://doi.org/10.11591/eei.v11i6.4021

Statistics of Japan. 2024. Statistics of Japan. Retrieved from URL https://www.e-stat.go.jp/en/stat-search?page=5&query=agricultural%20output%20value%20&layout=dataset&metadata=1&data=1

Syakur, M.A., Khotimah, B.K., Rochman, E.M.S. and Satoto, B.D. 2018. Integration K-means clustering method and elbow method for identification of the best customer profile cluster. IOP Conference Series, Materials Science and Engineering, 336, 012017, Retrieved from DOI https://doi.org/10.1088/1757-899X/336/1/012017

Troy, A-J. and Bogue, J. 2015. Simulation modelling, Identification of market segments for novel beta-glucan enriched breads. British Food Journal, 117 (2), 943-962, Retrieved from DOI https://doi.org/10.1108/bfj-12-2013-0363

van de Velden, M., D'Enza, A.I. and Palumbo, F. 2017. Cluster correspondence analysis. Psychometrika, 82 (1), 158-185, Retrieved from DOI https://doi.org/10.1007/s11336-016-9514-0

Yang, S-H., Panjaitan, B.P., Ujiie, K., Wann, J-W. and Chen, D. 2021. Comparison of food values for consumers' preferences on imported fruits and vegetables within Japan, Taiwan, and Indonesia. Food Quality and Preference, 87, 104042, Retrieved from DOI https://doi.org/10.1016/j.foodqual.2020.104042