

American Journal of Economics and Business Management

Vol. 9 Issue 7 | **pp.** 480-486 | **ISSN:** 2576-5973

Available online @ https://www.globalresearchnetwork.us/index.php/ajebm

Article

An AI-Driven Methodology for Conducting Marketing Research in the Food Products Market: Evidence from Uzbekistan

Murtozayeva Dilnoza Boboniyoz kizi[1]

1. Department of Economy, Vatan University, Uzbekistan
* Correspondence: nnoza80@gmail.com

Abstract: The growing volume, variety, and velocity of consumer data have reduced the adequacy of marketing research approaches that rely exclusively on descriptive statistics and linear models. This article develops an artificial intelligence (AI)-driven methodology for conducting marketing research in the food products market, with a proposed empirical application to Uzbekistan. The framework integrates theory-guided survey measurement, structural equation modelling, unsupervised consumer segmentation, supervised purchase-intention prediction, and explainable AI. It evaluates the effects of perceived product quality, price perception, food safety, brand trust, product information, online reviews, acceptance of AI-based recommendations, privacy concerns, and environmental considerations on purchase intention. A cross-sectional survey of approximately 500 adult food purchasers is proposed. Conventional explanatory analysis is combined with K-means or hierarchical clustering, random forests and gradient boosting, while SHAP values make model predictions interpretable. The framework addresses a methodological gap between explanatory consumer research and prediction-oriented AI analysis. It offers a reproducible sequence for data collection, validation, segmentation, prediction, interpretation, and managerial translation. Because primary data collection is pending, the article presents a testable methodological protocol rather than fabricated empirical results. The approach can help food producers, retailers, and digital platforms generate more timely, granular, and accountable consumer insights while preserving theoretical validity and responsible data governance.

Keywords: artificial intelligence; marketing research; food products market; consumer behaviour; machine learning; consumer segmentation; purchase intention; explainable AI; Uzbekistan

Citation: Kizi M. D. B. An AI-Driven Methodology for Conducting Marketing Research in the Food Products Market: Evidence from Uzbekistan. American Journal of Economics and Business Management 2026, 9(7), 480-486 .

Received:4th July 2026
Revised: 10th July 2026
Accepted: 15th July 2026
Published: 23th July 2026

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

Food markets are characterised by frequent purchasing, heterogeneous preferences, strong price sensitivity, safety concerns, information asymmetry, and growing interaction between physical and digital retail channels. Consumers evaluate not only taste and price but also freshness, origin, ingredients, certifications, brand reputation, online reviews, convenience, and environmental attributes. These decisions generate increasingly complex datasets. Conventional marketing research remains essential for measuring attitudes and testing theory, yet it can be insufficient when relationships are nonlinear, interactions are numerous, or consumer groups differ substantially [1] .

Artificial intelligence expands the analytical capacity of marketing research. Machine-learning algorithms can identify latent patterns, classify consumers, forecast

behavioural outcomes, and process high-dimensional data. AI-supported marketing can improve targeting, personalisation, customer insight, and real-time decision-making. Nevertheless, prediction alone does not establish causal or theoretically meaningful relationships. A high-performing algorithm may be managerially unusable if its decisions cannot be explained, while an explanatory model may have limited out-of-sample predictive performance [2]. Marketing research therefore requires an integrated design that combines theoretical measurement, statistical explanation, machine-learning prediction, and interpretable outputs.

This need is particularly relevant in emerging markets. Uzbekistan's food retail system combines traditional markets, neighbourhood stores, supermarkets, and rapidly developing delivery and e-commerce channels. Consumer groups differ by region, income, digital experience, and shopping format. Yet a transferable methodological framework for using AI to examine these differences remains underdeveloped [3]. This study responds to that gap by proposing an AI-driven methodology for the general food products market and specifying its empirical implementation in Uzbekistan.

The study asks three questions. First, which product, market, digital, trust, and sustainability factors most strongly shape food-purchase intention? Second, can machine-learning methods reveal consumer segments and predict purchase intention more accurately than conventional models? Third, how can explainable AI translate complex predictions into transparent marketing decisions? The contribution is an integrated architecture that links structural equation modelling (SEM), clustering, supervised learning, and SHAP-based interpretation [4].

AI in marketing encompasses computational systems that support sensing, learning, reasoning, recommendation, and decision-making. Applications include demand forecasting, targeting, recommendation systems, customer-service automation, sentiment analysis, and customer analytics. Davenport et al. (2020) explain how AI changes marketing strategy and customer interaction, while Huang and Rust (2021) distinguish mechanical, thinking, and feeling intelligence across marketing tasks. Kumar et al emphasise the use of AI throughout marketing functions and the need to align algorithms with managerial objectives [5].

For consumer research, AI has two complementary roles. It can be the object of inquiry for example, consumers' acceptance of AI recommendations and an analytical instrument for identifying behavioural patterns. The present framework includes both roles: AI-based recommendations constitute a consumer-perception construct, while machine learning is used to segment respondents and predict purchase intention [6].

Food-purchasing decisions combine functional, economic, informational, social, and ethical considerations. Perceived quality reflects an overall judgement of product excellence, while price operates as both economic sacrifice and quality signal. Safety information and clear labelling reduce risk and information asymmetry. Brand trust simplifies choice where quality cannot be fully evaluated before consumption. Online reviews provide electronic word-of-mouth signals, while environmental concerns increasingly shape preferences for packaging, local origin, and responsible production. Food-values research identifies price, safety, taste, nutrition, naturalness, origin, fairness, and environmental impact as distinct drivers of choice [7], [8].

Purchase intention is the primary behavioural outcome. Consistent with the Theory of Planned Behavior, intention is a proximal antecedent of behaviour, although situational constraints may prevent intention from becoming an actual purchase. The framework uses purchase intention as a prediction target while recognising that later research should validate it against transaction data.

AI recommendation systems can reduce search costs and expose consumers to relevant products. Their value depends on perceived accuracy, usefulness, and convenience. Yet personalisation requires data, creating a trade-off between relevance and

privacy [9]. Consumers may resist recommendations when data practices appear intrusive or opaque. Trust can reduce uncertainty, but responsible AI also requires meaningful consent, data minimisation, security, and understandable explanations. The model therefore examines both acceptance of AI recommendations and privacy concerns instead of treating adoption as uniformly beneficial.

Explanatory and predictive models answer different questions. SEM estimates theory-specified relationships among latent constructs, whereas machine learning prioritises generalisation to unseen observations. Shmueli cautions that explanatory power and predictive accuracy are not equivalent. This distinction is central to the proposed methodology: SEM tests hypotheses; cross-validated algorithms assess prediction; and explainable AI connects performance to actionable insight. SHAP values allocate a prediction among input features and permit both global and individual-level interpretation[10].

The framework positions purchase intention (BI) as the principal endogenous construct. Its antecedents are perceived product quality (PQ), price perception (PP), food safety (FS), brand trust (BT), product information (PI), online reviews (OR), acceptance of AI-based recommendations (AIR), privacy concerns (PC), and environmental considerations (EC). Privacy concerns are also expected to reduce acceptance of AI recommendations.

H1. Perceived product quality has a positive effect on food-purchase intention.

H2. Favourable price perception has a positive effect on food-purchase intention.

H3. Confidence in food safety has a positive effect on food-purchase intention.

H4. Brand trust has a positive effect on food-purchase intention.

H5. The availability and clarity of product information have a positive effect on food-purchase.

H6. Online consumer reviews have a positive effect on food-purchase intention.

H7. Acceptance of AI-based personalised recommendations has a positive effect on food-purchase.

H8. Privacy concerns have a negative effect on acceptance of AI-based personalised recommendations.

H9. Environmental considerations have a positive effect on food-purchase intention.

2. Materials and Methods

The empirical study will use a quantitative, cross-sectional survey. The target population consists of residents of Uzbekistan aged 18 years or older who participate in household food-purchasing decisions. Screening questions should confirm eligibility. A mixed online and face-to-face strategy is recommended to avoid excluding consumers with limited digital access. Participation must be voluntary, anonymous, and based on informed consent; no directly identifying data are required.

A target of 500 completed questionnaires is recommended, with at least 400 valid cases after screening. Quotas for gender, age, and broad geographical area should improve coverage. If a probability sampling frame is unavailable, the design must be described accurately as quota-assisted convenience sampling and should not be presented as nationally representative. A pilot test with 30–50 respondents should assess clarity, completion time, and item reliability.

All constructs will be measured with four reflective items on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Items are conceptually adapted from consumer-behaviour, technology-acceptance, trust, privacy, and food-choice research. The English instrument should be translated into Uzbek and Russian and independently back-

translated. Review by marketing, food-market, and quantitative-method specialists should establish content validity.

Table 1. Research Constructs and Operational Definitions

Construct	Code	Operational meaning
Perceived product quality	PQ	Quality, freshness, taste, and appearance
Price perception	PP	Value for money and promotional sensitivity
Food safety	FS	Expiry, certification, producer, and ingredient information
Brand trust	BT	Reliability, confidence, and repurchase willingness
Product information	PI	Packaging, nutrition, and origin information
Online reviews	OR	Ratings and electronic word of mouth
AI recommendation acceptance	AIR	Usefulness and willingness to use personalisation
Privacy concerns	PC	Collection, storage, consent, and security concerns
Environmental considerations	EC	Packaging, locality, impact, and producer responsibility
Purchase intention	BI	Future purchase, repurchase, and recommendation intention

Screening should remove duplicates, ineligible cases, failed attention checks, excessive missingness, and implausibly short completion times. Exclusion rules should be determined before hypothesis testing and reported transparently. Internal consistency should be assessed with Cronbach's alpha and composite reliability, generally seeking values of 0.70 or above. Confirmatory factor analysis should evaluate standardised loadings, model fit, convergent validity, and discriminant validity. Average variance extracted should normally exceed 0.50, while discriminant validity can be examined through the heterotrait–monotrait ratio. Common-method risk should be reduced procedurally through anonymity, clear wording, separation of predictor and outcome blocks, and varied item ordering.

SEM is proposed for H1–H9. Covariance-based SEM may be used where distributional conditions, specification, and sample size are adequate; PLS-SEM is an alternative where prediction is emphasised or assumptions are restrictive. The method must be justified before analysis. Results should include path coefficients, standard errors, confidence intervals, significance levels, explained variance, effect sizes, and appropriate fit or predictive-quality measures. Demographic and shopping-channel controls may be included, but exploratory controls must be distinguished from preregistered hypotheses.

Segmentation should use standardised construct scores rather than demographics alone. K-means, hierarchical clustering, and, where appropriate, Gaussian mixture models can be compared. Cluster selection should consider silhouette coefficient, within-cluster dispersion, resampling stability, parsimony, and managerial interpretability. Profiles should then be described using construct means, demographics, shopping frequency, and preferred channels. Labels such as "quality-and-safety seekers" or "digitally receptive shoppers" must emerge from observed profiles rather than be imposed in advance.

Purchase intention should preferably remain a continuous composite score to avoid arbitrary information loss. Multiple regression provides a transparent benchmark. Random forest and gradient boosting, including XGBoost, can capture nonlinearities and interactions. Data should be divided into training and untouched test sets, with hyperparameters tuned only within training data through repeated k-fold cross-validation. Preprocessing must be fitted inside each training fold to prevent leakage.

Continuous prediction should be evaluated using out-of-sample R-squared, root mean squared error, and mean absolute error. If a defensible categorical outcome is used, precision, recall, F1-score, ROC-AUC, and class balance should supplement accuracy. All models must use identical resampling splits, and uncertainty should be reported through confidence intervals or repeated-resampling distributions.

SHAP analysis should be applied to the best-performing tree-based model to estimate global importance, effect direction, nonlinear thresholds, and individual prediction explanations. Partial-dependence or accumulated-local-effect plots may provide robustness checks. Predictive importance must not be described as causal. Stability checks should compare feature rankings across resamples and algorithms. Potential algorithmic bias should be evaluated by testing whether prediction errors differ systematically across gender, age, region, or income groups.

3. Results and Discussion

No empirical coefficients are reported because primary data have not yet been collected. This is a deliberate safeguard against fabricated findings. After collection, the results section should begin with the final sample profile and transparent response-screening flow. It should report measurement reliability and validity, hypothesis tests, out-of-sample predictive comparisons, cluster stability and profiles, and SHAP explanations.

The design enables several forms of evidence. SEM may show whether quality, safety, trust, or AI acceptance has a statistically supported association with intention. Machine learning may reveal nonlinearities for example, that privacy becomes particularly important beyond a threshold or interactions between digital experience and recommendation acceptance. Clustering may show that a variable with a modest average effect is highly important for one segment. These outputs should be interpreted jointly, not treated as interchangeable [11].

The methodology connects theoretical validity with predictive usefulness. Surveys are sometimes criticised for aggregate relationships that obscure heterogeneity, while AI applications are criticised for weak theoretical grounding and opacity. The framework addresses both concerns. Theory determines what is measured and why; SEM evaluates the nomological structure; clustering detects heterogeneity; supervised learning tests generalisation; and explainability converts complex models into intelligible evidence.

This integration is valuable in food markets because purchase decisions reflect recurring trade-offs [12]. Price might dominate average predictions, yet a high-value segment may respond more strongly to safety information and trusted brands. Another group may accept personalised recommendations only when privacy protections are clear. A mass-market strategy can therefore be replaced by differentiated, evidence-based actions.

The Uzbekistan context extends research beyond heavily studied Western and East Asian digital markets. Its coexistence of traditional and modern retail channels permits examination of whether AI readiness and information use vary by shopping environment. Contextual contribution should nevertheless not be overstated: empirical generalisation depends on sampling quality and replication [13].

The study conceptualises AI not only as a marketing technology but also as part of research methodology. It connects intention-based consumer theory with food-choice attributes, digital influence, and technology acceptance. It operationalises the distinction

between explanation and prediction by assigning SEM and machine learning complementary functions. Explainable AI introduces an interpretation layer for nonlinear and segment-specific patterns without misrepresenting feature importance as causality. Privacy and fairness checks position responsible governance as a component of research quality.

Food producers can identify which combinations of quality, safety, origin, packaging, and brand information matter to specific groups. Retailers can improve assortments, promotions, and channel communication by comparing segments instead of demographic averages [14], [15]. Digital platforms can refine recommendations while monitoring privacy concerns and prediction errors. Model explanations should support, not automate, managerial judgement; predictive performance does not justify opaque data collection or discriminatory targeting.

4. Conclusion

The cross-sectional design cannot establish temporal causality and uses self-reported intentions rather than observed purchases. Quota-assisted convenience sampling may limit representativeness. A general food-products frame increases coverage but may conceal differences among fresh produce, dairy, meat, and packaged food. Future studies should use longitudinal panels, retailer transactions, online reviews, experiments, and category-specific samples. Multilingual sentiment analysis may add unstructured evidence but must be validated for Uzbek and Russian. Replication across Central Asia would clarify contextual boundaries.

Model accuracy may deteriorate when prices, expectations, or channels change. Researchers should document model versions, training periods, missing-data procedures, and subgroup performance. Where automated recommendations are deployed, consumers should receive meaningful information about data use and accessible choices.

This article proposes an integrated AI-driven methodology for marketing research in the food products market. It combines a theory-based consumer survey, measurement validation, SEM, machine-learning segmentation, predictive modelling, SHAP-based explanation, and responsible-data checks. Explanation identifies theoretically meaningful relationships, prediction tests generalisation, clustering reveals heterogeneity, and explainability supports transparent interpretation. Applied to Uzbekistan, the framework can generate evidence on quality, price, food safety, trust, information, online reviews, AI recommendations, privacy, sustainability, and purchase intention. The next stage is to pilot the questionnaire, collect genuine responses, preregister analytical decisions, and populate the Results and Discussion sections with reproducible empirical evidence.

REFERENCES

[1] O. Adekanle, D. A. Ndububa, S. A. Olowookere, O. Ijarotimi, and K. T. Ijadunola, "Knowledge of hepatitis B virus infection, immunization with hepatitis B vaccine, risk perception, and challenges to control hepatitis among hospital workers in a Nigerian tertiary hospital," Hepatitis Research and Treatment, vol. 2015, Art. no. 439867, 2015.

[2] T. O. Ajayi, S. O. Awojide, and M. C. Nkamigbo, "Knowledge, perception and preventive practices of hepatitis B infection among young adults in Lagos Island Local Government Area, Lagos, Nigeria: A cross-sectional study," Discover Public Health, vol. 23, Art. no. 381, 2026.

[3] I. A. Audu, E. B. Banwat, D. Z. Egah, and S. S. Gomerep, "Knowledge and awareness of hepatitis B virus infection among adults in North-Central Nigeria," African Health Sciences, vol. 22, no. 2, pp. 323–332, 2022.

[4] O. D. Bamidele, W. O. Adebimpe, and E. K. Oladipo, "Knowledge and perception of viral hepatitis among patients attending tertiary healthcare facilities in Nigeria," Pan African Medical Journal, vol. 42, Art. no. 115, 2022.

[5] V. L. Champion and C. S. Skinner, "The Health Belief Model," in Health Behavior and Health Education: Theory, Research, and Practice, 4th ed., K. Glanz, B. K. Rimer, and K. Viswanath, Eds. San Francisco, CA, USA: Jossey-Bass, 2008, pp. 45–65.

[6] Federal Ministry of Health, National Viral Hepatitis Strategic Plan 2021–2025. Abuja, Nigeria: Federal Ministry of Health, 2021.

[7] M. A. Hassan-Kadle, A. A. Osman, and A. M. Ibrahim, "Knowledge and awareness of viral hepatitis among hospital visitors: A cross-sectional study," BMC Public Health, vol. 23, Art. no. 1684, 2023.

[8] A. Kumah, M. Asare, J. Osei-Yeboah, and E. O. Anto, "Public awareness and knowledge of hepatitis B infection in sub-Saharan Africa: A systematic review," PLOS ONE, vol. 18, no. 7, Art. no. e0288463, 2023.

[9] B. M. Musa, S. Bussell, M. M. Borodo, A. A. Samaila, and O. L. Femi, "Prevalence of hepatitis B virus infection in Nigeria: A systematic review and meta-analysis," Nigerian Journal of Clinical Practice, vol. 24, no. 8, pp. 1095–1106, 2021.

[10] T. T. Nguyen, M.-L. McLaws, and G. J. Dore, "Knowledge, attitudes and practices regarding viral hepatitis: A systematic review," Journal of Viral Hepatitis, vol. 27, no. 11, pp. 1183–1195, 2020.

[11] A. T. Olayinka, B. O. Musa, and B. B. Afolabi, "Hepatitis B virus infection in Nigeria: Epidemiology, prevention and control challenges," Journal of Public Health in Africa, vol. 13, no. 2, Art. no. 2120, 2022.

[12] D. M. Razavi-Shearer et al., "Global prevalence, treatment, and prevention of hepatitis B virus infection in 2022: A modelling study," The Lancet Gastroenterology & Hepatology, vol. 7, no. 9, pp. 796–808, 2022.

[13] D. M. Razavi-Shearer et al., "Global prevalence, cascade of care, and prophylaxis coverage of hepatitis B in 2022: A modelling study," The Lancet Gastroenterology & Hepatology, vol. 8, no. 10, pp. 879–907, 2023.

[14] M. B. Shiferaw, A. Endalamaw, and W. A. Bayih, "Knowledge, attitude and practice toward hepatitis B virus infection among adults in sub-Saharan Africa: A systematic review and meta-analysis," BMC Infectious Diseases, vol. 21, Art. no. 742, 2021.

[15] World Health Organization, Global Hepatitis Report 2024: Action for Access in Low- and Middle-Income Countries. Geneva, Switzerland: World Health Organization, 2024.