CENTRAL ASIAN JOURNAL OF INNOVATIONS ON TOURISM MANAGEMENT AND FINANCE

https://cajitmf.centralasianstudies.org/index.php/CAJITMF
Volume: 07 Issue: 04 | October 2026 ISSN: 2660-454X

Article

A Big Data–Driven Methodology for Conducting Marketing Research in the Food Products Market: A Conceptual Framework for Uzbekistan

Murtozayeva Dilnoza Boboniyoz kizi*[1]

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

Abstract: *The digitalisation of food production, distribution, retailing, and consumption has created large volumes of heterogeneous and rapidly changing data. Point-of-sale transactions, loyalty programmes, e-commerce clickstreams, delivery applications, social media, online reviews, mobile location records, Internet of Things sensors, weather information, and supply-chain systems can provide a more continuous view of food markets than conventional surveys alone. However, the availability of data does not automatically produce valid marketing knowledge. This article develops a Big Data–driven methodology for conducting marketing research in the food products market, with a proposed application to Uzbekistan. The framework organises the research process around five properties of Big Data volume, velocity, variety, veracity, and value and integrates transactional, behavioural, attitudinal, contextual, and operational data. It specifies six methodological layers: research-problem formulation, data-source mapping, data governance and integration, analytical modelling, insight validation, and managerial activation. The proposed design combines descriptive, diagnostic, predictive, and prescriptive analytics to support market monitoring, demand forecasting, consumer segmentation, price and promotion evaluation, assortment decisions, and early detection of changes in consumer sentiment. Particular attention is given to data quality, identity resolution, selection bias, privacy, security, algorithmic fairness, and the difference between correlation and causality. Since an integrated empirical database is not yet available, the paper presents a conceptual and operational protocol rather than fabricated results. The framework contributes by adapting Big Data marketing research to the distinctive characteristics of food markets and emerging-market data environments.*

Keywords: *Big Data; marketing research; food products market; consumer analytics; data integration; demand forecasting; real-time analytics; food retail; Uzbekistan*

Citation: Boboniyoz kizi, M. D. A Big Data–Driven Methodology for Conducting Marketing Research in the Food Products Market: A Conceptual Framework for Uzbekistan. Central Asian Journal of Innovations on Tourism Management and Finance 2026, 7(4), 146-153.

Received: 14th Dec 2025
Revised: 28th Jan 2026
Accepted: 11th Feb 2026
Published: 20th July 2026

1. Introduction

Marketing research in food markets has traditionally relied on questionnaires, interviews, focus groups, retail audits, and official statistics. These instruments remain valuable because they measure attitudes, motivations, satisfaction, and intentions. Nevertheless, they often provide periodic snapshots, depend on self-reports, and may not capture rapidly changing behaviour. Food markets are especially dynamic because products are purchased frequently, many items are perishable, prices and availability vary, and purchasing is influenced by seasonality, income, promotions, weather, health concerns, and social trends.

Digital transactions have changed the informational environment. Supermarkets generate item-level sales records; loyalty programmes connect purchases over time; e-commerce platforms observe searches, clicks, baskets, substitutions, and abandoned orders; delivery applications record time and location; online reviews capture consumer language; and supply-chain systems provide data about inventory, traceability, temperature, and delivery [1]. When integrated responsibly, these sources allow marketing researchers to analyse what consumers say, what they search for, what they purchase, when and where they purchase it, and the conditions surrounding the decision.

Big Data is not defined by size alone. It is commonly characterised by volume, velocity, variety, veracity, and value. Volume concerns scale; velocity concerns the speed of generation and required response; variety concerns structured and unstructured formats; veracity concerns accuracy and reliability; and value concerns the conversion of data into beneficial decisions. In food markets, these properties create opportunities but also methodological difficulties. Transaction data can be behaviourally precise yet lack motivations [2], [3]. Social-media data are rich in language but may not represent the population. Sensor records can improve traceability but may be incomplete or incompatible across organisations.

Research on Big Data analytics shows potential for customer-behaviour analysis, demand prediction, operational visibility, and organisational performance. Food-sector reviews also identify applications in safety, waste reduction, forecasting, traceability, and consumer service. However, research often concentrates on individual technologies or supply-chain applications without presenting an integrated marketing-research process that connects business questions, multiple data sources, validity, governance, and managerial decisions [4].

This article therefore develops a Big Data–driven methodology for marketing research in the food products market. It asks: (1) which data sources should be integrated to provide a multidimensional view of consumers and markets; (2) how should quality, privacy, representativeness, and analytical validity be managed; and (3) how can Big Data outputs be translated into marketing decisions? The proposed framework is designed for empirical implementation in Uzbekistan while remaining adaptable to other emerging markets [5], [6].

Traditional marketing research usually begins with a defined research question and purposefully collected data. Big Data research often begins with data generated as a by-product of transactions or digital interaction. This difference creates both opportunity and risk. Behavioural records reduce recall error and permit longitudinal analysis, but variables may not correspond directly to theoretical constructs [7]. Consequently, Big Data should complement rather than automatically replace theory-guided primary research.

Marketing analytics converts marketplace data into insight for decisions. Erevelles et al. explain that consumer Big Data can support value creation through the generation, processing, and use of information. Wedel and Kannan emphasise analytics for data-rich environments, including personalisation, marketing-mix decisions, and scalable modelling. Dynamic-capabilities theory further suggests that organisations gain advantage when they can sense market change, seize opportunities, and reconfigure resources [8], [9]. From this perspective, data infrastructure alone is not a capability; value arises from the coordinated ability to acquire, integrate, analyse, interpret, and act on data .

Volume is generated through high-frequency purchases, stock movements, online interactions, and sensor observations.

Velocity reflects continuous sales, price changes, delivery events, and social conversations.

Variety includes numerical transactions, text, images, locations, timestamps, weather, and sensor signals.

Veracity is challenged by missing product identifiers, duplicate customers, false reviews, inconsistent units, and sensor errors.

Value depends on whether analysis improves decisions such as forecasting, targeting, pricing, assortment, communication, and waste reduction.

Food markets add distinctive characteristics. Perishability makes time and inventory essential. Product substitution occurs when preferred items are unavailable. Quality can change during storage and transport. Purchases are often made for households, meaning that the buyer and consumer may differ. Cultural and regional preferences are strong, and informal retail may be poorly represented in digital datasets. A suitable methodology must therefore combine market, consumer, product, channel, and contextual information.

The methodology incorporates four levels of analytics. Descriptive analytics asks what occurred for example, changes in category sales. Diagnostic analytics asks why it occurred by examining prices, promotions, stock-outs, channels, or sentiment. Predictive analytics estimates future demand, churn, response, or category growth. Prescriptive analytics recommends actions under constraints, such as promotion timing or assortment allocation [10]. These levels are cumulative: prescriptive recommendations are unreliable when descriptive data or predictive models are weak.

2. Materials and Methods

Research

Three gaps motivate the framework. First, food-market datasets are frequently analysed separately, preventing a unified consumer and market view. Second, technical performance is often prioritised over measurement validity, representativeness, and causal interpretation. Third, organisational and ethical requirements are frequently treated as implementation issues rather than components of research methodology.

The following propositions guide later empirical testing:

P1. Integration of transactional, behavioural, attitudinal, contextual, and operational data improves the completeness of marketing insight compared with reliance on a single data source.

P2. Big Data veracity positively influences the reliability and managerial usefulness of marketing research results.

P3. Real-time analytical capability improves the speed of detecting changes in food demand and consumer sentiment.

P4. Advanced analytical capability improves demand forecasting, consumer segmentation, and evaluation of marketing actions.

P5. Data-governance capability strengthens the positive relationship between Big Data analytics and trust in marketing decisions.

P6. The effect of Big Data analytics on marketing performance is mediated by organisational insight-activation capability.

Big Data research can reproduce exclusion when digital traces overrepresent urban, higher-income, or platform-active consumers. Loyalty data exclude non-members; online reviews overrepresent people motivated to post; and informal markets may remain invisible. Researchers must describe the population represented by each source and avoid presenting digital coverage as universal coverage.

Privacy risks arise when records are linked across sources. Purpose limitation, informed consent where required, minimisation, pseudonymisation, access control, retention limits, and security are essential. Algorithmic models should be checked for uneven error or impact across relevant groups. Researchers must also avoid causal claims from correlations and should document uncertainty, data revisions, and model drift.

Proposed Big Data–Driven Marketing Research Methodology

The process begins with a decision, not a dataset. Researchers should specify the target market, unit of analysis, time horizon, outcome, and intended decision. Examples include weekly demand forecasting, identification of price-sensitive households, evaluation of promotion incrementality, detection of emerging product attributes, or prediction of customer inactivity. A data source should be included only when it contributes to the research question and has a lawful and defensible purpose.

3. Results and Discussion

Dickey-The framework uses five complementary data domains.

Table 1. Major Data Domains and Their Contribution to AI-Driven Marketing Research

Data domain	Examples	Marketing contribution
Transactional	POS sales, prices, baskets, loyalty purchases, returns	Reveals observed demand, frequency, value, and product combinations
Digital behavioural	Searches, clicks, dwell time, carts, abandoned orders	Captures the path to purchase and unmet interest
Attitudinal and textual	Surveys, ratings, reviews, service messages, social posts	Explains perceptions, satisfaction, motives, and sentiment
Contextual	Weather, holidays, geography, inflation, demographics	Explains external variation and improves forecasting
Operational	Inventory, stock-outs, delivery, traceability, sensor and waste records	Separates demand from availability and links marketing with fulfilment

The architecture should distinguish first-party data collected directly by a firm, second-party data shared by an authorised partner, and third-party or open data. Provenance, ownership, update frequency, permitted use, retention, and quality should be documented for each source.

Governance begins before ingestion. The organisation should establish a data dictionary, product taxonomy, access roles, retention schedule, consent rules, security controls, and audit trail. Personal data should be minimised and pseudonymised where individual-level linkage is necessary. Sensitive attributes should not be inferred or used merely because prediction is technically possible.

Integration requires common keys for products, stores, channels, geography, and time. Customer identity resolution is particularly sensitive and should use transparent, authorised rules. Product records require harmonisation of spelling, package size, units, brands, categories, and changing stock-keeping-unit identifiers. Event timestamps should be converted to a consistent standard.

Quality assessment should cover completeness, accuracy, consistency, timeliness, uniqueness, validity, and lineage. Dashboards should track missingness, duplicates, unusual price or volume changes, broken sensor sequences, and schema changes. Veracity must be evaluated at source, variable, and analytical-output levels. A larger dataset does not compensate for systematic bias.

The analytical pipeline contains six stages. First, exploratory analysis profiles distributions, trends, seasonality, missingness, and relationships. Second, descriptive dashboards monitor category sales, prices, promotion, channel, geography, and stock availability. Third, segmentation uses behavioural measures such as recency, frequency, monetary value, category diversity, price sensitivity, channel use, and promotion responsiveness. Fourth, forecasting models estimate demand at appropriate product, location, and time levels. Fifth, text analytics extracts topics and sentiment from reviews and customer messages. Sixth, causal or quasi-experimental methods evaluate marketing interventions where randomisation is not available.

Algorithms should be selected according to the question rather than fashion. Time-series models may be suitable for stable seasonal series, while tree-based or deep-learning models may help when nonlinear contextual effects are important. Clustering can support segmentation, association rules can identify basket relationships, and natural-language processing can analyse text. Baseline models must always be reported so that analytical complexity demonstrates incremental value.

Validation should address technical performance, construct meaning, external validity, and decision relevance. Forecasts require rolling-origin or time-based validation rather than random splits that leak future information. Classification models should report discrimination, calibration, and subgroup error. Segments should be stable across samples and meaningful to managers. Text models should be manually validated on a labelled subset, particularly for Uzbek and Russian language content.

Triangulation is central. Transaction patterns should be compared with survey attitudes; online sentiment should be compared with sales and service records; observed declines should be separated from stock-outs; and algorithmic findings should be reviewed by domain experts. Where the objective is causal evaluation, controlled experiments are preferred. Observational attribution should use credible comparison groups, difference-in-differences, matching, or other appropriate designs and should state assumptions explicitly.

Analytical outputs create value only when linked to decisions. Each result should identify the responsible decision owner, action threshold, refresh frequency, expected benefit, potential harm, and monitoring indicator. Demand forecasts may guide inventory and communication; segments may guide differentiated offers; sentiment changes may trigger product or service investigation; and promotion analysis may reallocate spending.

Activation should include feedback. The system must record which action was taken and whether outcomes improved. This converts marketing research from a one-time report into a learning cycle. Models should be recalibrated when consumer behaviour, prices, assortment, channels, or data-generation processes change.

Empirical Implementation in Uzbekistan

An initial implementation can be conducted with one cooperating supermarket chain, online grocery platform, or food producer with direct sales. A twelve- to twenty-four-month period is preferable to capture seasonality. The minimum viable dataset may combine item-level weekly sales, price, promotion, store or channel, inventory availability, product characteristics, and public contextual data. If authorised loyalty or e-commerce data are available, behavioural segmentation can be added. A complementary consumer survey can measure attitudes that transactions cannot reveal [11].

The empirical study can be organised into four modules:

1. **Market-monitoring module:** category, brand, price, promotion, region, channel, and seasonal trends.
2. **Consumer-insight module:** RFM behaviour, basket composition, channel preference, survey attitudes, review topics, and sentiment.
3. **Predictive module:** product–location–time demand forecasts and response or churn prediction.
4. **Decision module:** promotion evaluation, assortment recommendations,

communication priorities, and monitoring of implementation outcomes.

Table 3. Evaluation Dimensions and Representative Performance Indicators

Dimension	Illustrative indicators
Data quality	Completeness, duplicate rate, valid identifiers, latency, lineage coverage
Analytical quality	MAE/RMSE, MAPE where appropriate, AUC, calibration, cluster stability
Insight quality	Relevance, interpretability, timeliness, manager confidence
Marketing outcomes	Sales lift, retention, basket value, response rate, forecast improvement
Operational outcomes	Stock-out rate, waste, inventory turnover, service level
Governance outcomes	Consent coverage, access violations, deletion compliance, subgroup error gaps

4. Discussion

The proposed framework changes marketing research in three ways. First, it moves from periodic data collection toward continuous market sensing while retaining purposefully designed surveys for latent attitudes. Second, it treats data engineering and governance as part of methodology rather than invisible technical preparation [12]. Third, it links analysis to controlled activation and monitoring, enabling research findings to be evaluated through subsequent outcomes.

For food markets, integrated data can distinguish apparent demand from availability. A sales decline may reflect reduced preference, higher prices, a stock-out, poor delivery coverage, or a competing promotion. Transaction data alone cannot resolve all explanations. By combining operational, contextual, behavioural, and attitudinal sources, the researcher can develop a more credible diagnosis.

Big Data also improves temporal and spatial granularity. Managers can identify regional differences, seasonal patterns, and rapidly emerging concerns [13]. Yet granularity increases privacy risks and the possibility of unstable small-group conclusions. Responsible aggregation, access control, minimum-cell rules, and subgroup validation should therefore accompany analytical detail.

The study contributes by integrating the five-V conception of Big Data with a complete marketing-research cycle. It extends a technology-centred view toward a capability-based model in which data, analytical skills, governance, domain knowledge, and activation jointly produce value [14]. It also distinguishes behavioural trace data from attitudinal measures and explains why neither is sufficient alone. Finally, it embeds veracity, representativeness, causal validity, privacy, and fairness within the research design.

Food producers can use integrated data to understand product combinations, regional preferences, and changes in quality perceptions. Retailers can improve forecasting, assortment, pricing, promotion, and communication. Digital platforms can examine search-to-purchase gaps, substitution, and delivery experience. Public institutions may use appropriately aggregated and governed information to monitor food-market conditions without exposing individual consumers [15].

Implementation should begin with a bounded decision problem rather than an expensive enterprise-wide platform. Organisations should establish product identifiers, quality rules, and governance responsibilities before adopting complex algorithms. Collaboration between marketing specialists, data engineers, statisticians, food experts, legal professionals, and decision owners is necessary. Skills and organisational routines are as important as infrastructure.

5. Conclusion

This article proposes a methodology and does not report an implemented integrated database. Its practical performance must therefore be tested with real organisational data. Data availability, standardisation, and institutional readiness may vary across Uzbekistan's food sector. The framework may require simplification for small enterprises and stricter controls for highly granular customer data.

Future research should compare Big Data results with conventional survey estimates, test the incremental forecasting value of each data source, evaluate cross-channel identity methods, and develop validated Uzbek-language text models. Longitudinal studies should examine whether analytical capability improves marketing and sustainability performance. Research should also evaluate whether data-driven personalisation benefits consumers equitably.

This article develops a Big Data–driven methodology for marketing research in the food products market. The framework integrates problem formulation, source mapping, governance, data integration, descriptive and advanced analytics, triangulation, activation, and continuous learning. It recognises that volume and computational power are insufficient without veracity, theoretical meaning, representativeness, privacy, and managerial relevance. Applied responsibly in Uzbekistan, the methodology can support timely market monitoring, deeper consumer insight, stronger demand forecasting, better segmentation, and evidence-based marketing decisions. The next stage is to establish an authorised data partnership, build a minimum viable integrated dataset, and empirically evaluate analytical and managerial improvements.

REFERENCES

[1] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," International Journal of Information Management, vol. 35, no. 2, pp. 137–144, 2015.

[2] M. Wedel and P. K. Kannan, "Marketing analytics for data-rich environments," Journal of Marketing, vol. 80, no. 6, pp. 97–121, 2016.

[3] N. Rustamova, "Improving the methodology of conducting marketing research in municipal markets for goods and services," Russian Journal of Economics and Law, vol. 1, pp. 89–96, 2025.

[4] D. B. Murtozaeva, "Enhancing the effectiveness of marketing research in the food market under digital transformation," 2025.

[5] V. Kumar, A. Dixit, R. Javalgi, and D. Dass, "Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing," Journal of the Academy of Marketing Science, vol. 44, no. 1, pp. 24–45, 2016.

[6] E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York, NY, USA: W. W. Norton & Company, 2014.

[7] T. H. Davenport and J. G. Harris, Competing on Analytics: The New Science of Winning. Boston, MA, USA: Harvard Business Review Press, 2017.

[8] V. Mayer-Schönberger and K. Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think. Boston, MA, USA: Houghton Mifflin Harcourt, 2013.

[9] P. C. Verhoef, E. Kooge, and N. Walk, Creating Value with Big Data Analytics: Making Smarter Marketing Decisions. London, U.K.: Routledge, 2016.

[10] D. Chaffey and F. Ellis-Chadwick, Digital Marketing: Strategy, Implementation and Practice, 8th ed. Harlow, U.K.: Pearson, 2022.

[11] P. Kotler, G. Armstrong, S. Adam, and L. Denize, Principles of Marketing, 9th Asia-Pacific ed. Melbourne, Australia: Pearson, 2021.

[12] K. C. Laudon and C. G. Traver, E-Commerce: Business, Technology, Society, 18th ed. Boston, MA, USA: Pearson, 2023.

[13] T. H. Davenport, Big Data at Work: Dispelling the Myths, Uncovering the Opportunities. Boston, MA, USA: Harvard Business Review Press, 2014.

[14] S. Akter and S. F. Wamba, "Big data analytics in e-commerce: A systematic review and agenda for future research," Electronic Markets, vol. 26, no. 2, pp. 173–194, 2016.

[15] S. F. Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, “How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study,” International Journal of Production Economics, vol. 165, pp. 234–246, 2015.