IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 5, Ser. I (Sep. – Oct. 2026), PP 07-10
www.iosrjournals.org

Data Science and Marketing Analytics

Somnath Kar[1], Bikash Pradhan[2], Manas Napit[3]
[1]*Department of Computer Applications (MCA), Kolhan University, Chaibasa, Jharkhand, India*
[2]*Department of Computer Applications (MCA), Kolhan University, Chaibasa, Jharkhand, India Email:*
[3]*Department of Computer Applications (MCA), Kolhan University, Chaibasa, Jharkhand, India*

Abstract: *Data Science and Marketing Analytics has become an important area for organizations that want to make evidence-based business decisions. Modern organizations collect large volumes of customer, transaction, web, social-media and campaign data. Data science methods for cleaning, integrating, analyzing and modelling these data, while marketing analytics converts analytical findings into practical decision related to customer acquisition, retention, segmentation, pricing, promotion and campaign performance. This thesis presents a conceptual and practical framework for applying data science techniques to marketing analytics. Data collection preprocessing exploratory data analysis customer segmentation predictive modelling campaign measurement and visualization. The study also explains how descriptive, diagnostic, predictive and prescriptive analytics can work together. The proposed framework can help organization improve customer understanding, optimize marketing resources and build measurable, data driven strategies.*

Keywords: *data science, marketing analytics, machine learning, customer segmentation, predictive analytics, customer lifetime value, business intelligence, marketing optimization, artificial intelligence.*

Date of Submission: 26-08-2026 Date of Acceptance: 06-09-2026

I. INTRODUCTION

Marketing has changed significantly with the growth of digital platforms. Customers interact with organizations through websites mobile applications social networks email, search engines and physical channels. Each interaction can produce useful information. However, collecting data alone does not create business value. Organizations need analytical techniques that transform raw data insights and decisions.

Data science combines statistics, mathematics, programming, machine learning and domain knowledge. Marketing Analytics applies these analytical capabilities to marketing problems. Typical questions include Who are the most valuable customers. Which customers may leave? Which campaign produces the best response. Which products frequently purchased together. Which customer segments should receive a particular offer.

The purpose of this research is to examine how data science methods can support marketing decisions and to propose an end-to-end framework that connects business objectives with data, analytical models and measurable marketing outcomes.

II. OBJECTIVES AND SCOPE

The major objectives of study are -

(1) to understand role of data science in marketing analytics.
(2) to identify important marketing data sources;
(3) to explain data preprocessing and exploratory analysis.

(4) to study customer segmentation techniques.
(5) to examine predictive models for marketing decisions.
(6) to evaluate campaign performance using analytical measures.
(7) to propose a practical data-driven marketing framework.

The scope includes customer analytics, segmentation, campaign analysis, customer lifetime value, churn prediction, recommendation concepts and marketing dashboards. The study is designed as an academic framework and can be implemented with tools such as python, SQL, spreadsheets and visualization platforms.

III. LITERATURE REVIEW

Research in marketing analytics has increasingly focused on using customer-level data to understand behavior and improve marketing effectiveness. Traditional marketing research relied heavily on surveys, expert judgement and aggregate measures. Digital transformation has increased the availability of behavioral data and created opportunities for near-real-time analysis.

Customer segmentation is a foundational marketing analytics activity. Clustering techniques such as K-Means can group customers according to similarities in behavior.

Predictive analytics can estimate
outcomes such as response probability, churn risk or expected customer value. Machine learning methods can identify complex relationships may not be obvious through simple descriptive statistics.

A recurring theme in the literature is that analytical accuracy must be connected to business usefulness. A model that performs well statistically may still have limited value if its predictions cannot be converted into an actionable marketing decision. Therefore, model evaluation should include both technical metrics and business metrics.

IV. RESEARCH METHODOLOGY

The proposed methodology follows a structured data science lifecycle. First, the business problem is defined in measurable terms. Second, relevant data are collected from sources such as customer databases, sales records, campaign systems, websites and digital channels. Third, data quality assessed and preprocessing is performed.

Preprocessing handling missing values, removing duplicates, correcting inconsistent formats, treating outliers and transforming variables. Exploratory Data Analysis is then used to understand distributions, correlations, customer behavior and campaign patterns. Visualization is useful for communicating these findings to decision makers.

For segmentation, customer attributes such as recency, frequency and monetary value can be analyses. K-Means clustering can then be applied to create behavioral groups. For prediction, algorithms such as logistic regression decision trees random forests or gradient-boosting methods may be considered depending on the problem. Models should be trained using appropriate validation procedures and evaluated with suitable metrics.

V. DATA SCIENCE TECHNIQUES IN MARKETING

Descriptive analytics answers what happened. Examples include monthly sales, conversion rate, customer count and campaign response rate. Diagnostic analytics examines why an outcome occurred by comparing customer groups, channels, products or time periods.

Predictive analytics estimates what is likely to happen. Churn prediction can identify customers with a higher probability of leaving. Response modelling can estimate which customers are more like to responds to a campaign. Customer Life time Value models can estimate the future economic value of a customer.

Prescriptive analytics goes a step further by recommending actions. For example, a company may use customer propensity scores to prioritize marketing contacts, select suitable offers or allocate budgets across channels. The recommended action should always be evaluated against business constraints, cost and expected benefit.

VI. CUSTOMER SEGMENTATION

Customer segmentation divides a broad customer base into meaningful groups. A common analytical approach uses Recency, Frequency and Monetary (RFM) variables. Recency represents how recently a customer purchased, Frequency represents purchase frequency, and Monetary represents spending value.

After preparing the variables, clustering can be performed. The resulting clusters should be profiled using averages, distributions and business characteristics. A possible interpretation may include high-value loyal customers, recent customers, occasional customers and inactive customers. These labels should be based on the actual characteristics of the dataset rather than assumptions.

Segmentation can improve personalization because different groups can receive different messages, offers and retention strategies. However, organizations should periodically review segments because customer behavior changes over time.

VII. CAMPAIGN AND CUSTOMER ANALYTICS

Marketing campaign performance can be measured through metrics such as impressions, clicks, leads, conversions, cost per acquisition, return on marketing investment and customer retention. The choice of metric depends on the campaign objective.

A useful analytical process compares campaign outcomes across customer segments and channels. A/B testing can be used to compare alternative messages, offers or landing pages. Statistical testing and controlled experimentation can reduce the risk of interpreting random variation as a real marketing effect.

Customer analytics can also connect campaign interactions with downstream outcomes such as repeat purchase and customer lifetime value. This helps organizations move beyond short-term engagement metrics and evaluate longer-term business impact.

VIII. PROPOSED SYSTEM

The proposed framework consists of six layers: Business Objective, Data Layer, Data Preparation, Analytics and Machine Learning, Visualization, and Decision/Action. The business objective defines the question. The data layer integrates relevant sources. Data preparation improves reliability. Analytical models generate insights. Dashboards communicate results. Finally, marketing teams act on the findings and measure the outcome.

A practical implementation may use SQL for data extraction, Python for preprocessing and modelling, and a dashboarding tool for visualization. The workflow should be repeatable so that updated data can be processed without rebuilding the complete analysis manually.

IX. RESULTS AND DISCUSSION

For an academic implementation, the results section should report the actual dataset characteristics, preprocessing steps, model parameters and evaluation scores obtained during experimentation. Illustrative results must not be presented as real findings unless they are calculated from a defined dataset. A suitable evaluation table can include segmentation quality, classification accuracy precision recall F1-score ROC-AUC campaign conversion rate and return on investment, depending on the selected research problem. Business interpretation is essential: for example, a small improvement in response prediction may be valuable if it reduces campaign cost or improves conversion among a high-value segment.

X. ETHICAL, PRIVACY AND SECURITY CONSIDERATIONS

Marketing analytics often uses personal and behavioral information. Data should therefore be collected and processed lawfully, fairly and transparently. Organizations should apply appropriate accesscontrols, data minimization, retention policies and security measures. Analytical models can also create bias if training data are incomplete or systematically unrepresentative. Model outputs should be monitored, validated and reviewed by responsible teams. Personalization should not become intrusive, discriminatory or manipulative. Privacy-preserving practices and responsible governance are important components of a sustainable analytics program.

XI. FUTURE SCOPE

- **AI Integration:** Artificial intelligence automates basic data tasks, shifting the analyst's role from manual coding to validating AI results and making smart choices.
- **Personalized Marketing:** Brands use predictive analytics to anticipate consumer behaviour and deliver hyper-targeted campaigns.
- **Specialized Roles:** The general "data scientist" title is splitting into focused tracks like machine learning engineering, data engineering, and customer analytics.
- **High Market Demand:** Millions of new jobs are opening globally and locally, making cross-disciplinary skills in both marketing and statistics very valuable.

XII. CONCLUSION

Data Science and Marketing Analytics have become essential components of modern business strategy. The rapid growth of digital technologies, social media platforms, e-commerce, and online customer interactions has generated enormous amounts of data. Organizations can use Data Science techniques to collect, process, analytics, and interpret this data, enabling them to make more accurate and effective marketing decisions. This study concludes that the integration of Data Science with Marketing Analytics helps organizations better understand customer behaviour, preferences, needs, and purchasing patterns. Techniques such as machine learning, predictive analytics, data mining, customer segmentation, sentiment analysis, and recommendation systems enable businesses to develop personalized marketing strategies and improve customer experiences.

References

[1]. Berman, R. (2018) Beyond the last touch Attribution in online advertising. Marketing Science.

[2]. Bruce, P, Bruce, A., & P. (2020). Practical statistics for data scientists: 50+ essential concepts using R and Python (2nd ed.). O'Reilly Media.

[3]. Davenport, T. H & Harris, J. G. (2017). Competing on analytics: The new science of winning (Updated ed.). Harvard Business Review Press.

[4]. Han, J. Kamber, M. & Pei, J. (2012). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann.
[5]. James, G, Witten, D., Hastie, T., & R. (2021). An introduction to statistical learning: With applications in R (2nd ed.). Springer.
[6]. Kotler, P. Keller, K. L. & Chernev, A. (2022). Marketing management (16th Global ed.). Pearson.
[7]. Provost, F. & Fawcett, T. (2013). Data science for business. What you need to know about data mining and data-analytic thinking. O'Reilly Media.
[8]. Shmueli, G., Bruce, P. C, k, P., & Patel, N. R. (2020). Data mining for business analytics Concepts, techniques and applications in Python (1st ed.). Wiley.

Author's Short Profile:

Dr. Somnath Kar is an Assistant Professor at University Department of Physics, Kolhan University, Chaibasa and also the Coordinator of University Department of MCA, Kolhan University, Chaibasa.

Bikash Pradhan is pursuing the Master of Computer Applications (MCA) at Kolhan University, Chaibasa. This paper is drawn from her MCA thesis on Data science & Marketing, prepared under the guidance of Dr. Somnath Kar.

Manas Napit is pursuing the Master of Computer Applications (MCA) at Kolhan University, Chaibasa. This paper is drawn from her MCA thesis on Data science & Marketing, prepared under the guidance of Dr. Somnath Kar