



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 3, March 2026

Product Positioning Strategy Analysis: Evaluating Sales Performance using Tableau and Machine Learning Models

**G.Mahesh M.Tech¹, Bekkanti Ram Suresh², Bontha Ricky³, Syamala Charan Reddy⁴,
Chintamalla Karthik⁵**

Assistant Professor, Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India¹

U.G. Students, Department of CSE(Artificial intelligence & Machine Learning), Kallam Haranadhareddy Institute of
Technology, Guntur, Andhra Pradesh, India^{2,3,4,5}

ABSTRACT : This research introduces a web-based, data-centric platform designed to assess and enhance product placement tactics within retail settings. By merging Machine Learning methodologies with dynamic Tableau dashboards, the framework examines the effects of product location, pricing models, promotional efforts, rival pricing, shopper demographics, seasonal variations, and store traffic on sales outcomes. The server-side architecture, developed with Flask and Scikit-learn, implements and compares three distinct regression algorithms—Linear Regression, Random Forest, and Decision Tree—to forecast Sales Volume based on key retail parameters. Furthermore, clustering methods are utilized to categorize products by their sales performance. The system's predictive reliability is gauged using standard evaluation metrics, including the R^2 score, Mean Squared Error (MSE), and Mean Absolute Error (MAE), allowing for a comprehensive comparison of model performance. Leveraging Tableau Public, we construct interactive dashboards and narrative reports that convert raw analytical findings into practical business knowledge. This initiative illustrates the power of combining predictive analytics with business intelligence tools to facilitate informed, data-backed decision-making in contemporary retail strategy.

KEYWORDS: Tableau Visualization, Product Placement Analysis, Sales Forecasting, Machine Learning, Linear Regression, Random Forest, Decision Tree, Retail Analytics, Business Intelligence, Clustering, Feature Importance, Sales Optimization

I. INTRODUCTION

The emergence of Machine Learning (ML) and Business Intelligence (BI) technologies offers organizations a pathway beyond intuition-based strategies towards data-informed approaches. Machine Learning excels at uncovering subtle correlations, hidden patterns, and predictive connections within extensive retail datasets. Concurrently, Business Intelligence platforms like Tableau translate these complex analytical results into accessible, visual insights for stakeholders. The selection of appropriate ML algorithms is critical; hence, this study employs and compares three established techniques: Linear Regression for its interpretability, Decision Tree for its rule-based logic, and Random Forest for its robustness and accuracy through ensemble learning.

This project is centered on creating a predictive analytics framework to assess the influence of essential retail variables on sales figures. The system quantitatively examines the effects of: **Product Position** (e.g., front, middle, or back placement)

- Product Position (e.g., placement at the front, middle, or rear of the store)
- Promotion Strategy (including discounts and advertising campaigns)
- Foot Traffic (the volume of customers in the store)
- Product Price
- Competitor Pricing
- Consumer Demographics

- Product Category
- Seasonal Trends

By utilizing Linear Regression, Decision Tree, and Random Forest models, the platform forecasts Sales Volume based on these strategic elements and provides a comparative analysis of their performance. Additionally, clustering algorithms are employed to group products into distinct performance tiers, assisting retailers in pinpointing both their top-performing and underperforming segments.

To ensure these insights are accessible and actionable, the analytical outputs are seamlessly integrated with Tableau Public dashboards, offering interactive data exploration and visual storytelling. This fusion effectively closes the gap between complex machine learning outputs and practical retail strategy, empowering decision-makers to grasp market trends, anticipate outcomes, and fine-tune product placement for maximum effectiveness.

In essence, this project showcases how the synergy of multiple predictive models and data visualization can equip retailers to make intelligent, strategic choices that maximize profitability in a constantly evolving market..

II. LITERATURE SURVEY

Current academic and industry research underscores the critical function of Business Intelligence and visualization platforms in enhancing retail sales. Investigations, including those focused on Sales Strategy & KPI Analysis in Vending Machines and Retail Sales Dashboard Optimization, demonstrate that Tableau dashboards are effective tools for monitoring key performance indicators, examining sales patterns, and aiding strategic planning. These visual tools transform unprocessed retail data into comprehensible and valuable business narratives.

Further scholarly work has explored the application of predictive models for forecasting retail sales. This includes the use of regression techniques, ensemble methods like Random Forest, and rule-based models such as Decision Trees to anticipate product demand and evaluate the influence of pricing structures. Studies have shown that while linear models offer simplicity and interpretability, tree-based methods often capture non-linear relationships and feature interactions more effectively. Data mining methods, particularly clustering, have proven useful for identifying primary sales drivers and classifying products according to their market performance.

Much of the prior work has treated prediction, clustering, and visualization as separate tasks. This project distinguishes itself by integrating **multiple machine learning algorithms**—specifically Linear Regression, Decision Tree, and Random Forest—with Tableau's visualization capabilities into a unified, decision-support system specifically designed for strategic analysis in the retail sector.

FLASK:

Flask serves as the foundational backend framework for the Strategic Product Placement Analysis web application. It establishes the connection between the trained machine learning models and the user-facing interface, handling user requests and dynamically displaying prediction results from all three algorithms. Defined routes, such as /dashboard, /story, /model-performance**, **/sales-forecast, and /virtual-placement, manage core functionalities including sales prediction, product grouping, and performance ranking. Flask utilizes the Jinja2 templating engine to render HTML pages with dynamic content, allowing users to compare predictions from Linear Regression, Decision Tree, and Random Forest models side-by-side.

TABLEAU PUBLIC:

Tableau Public is employed to develop interactive, web-based dashboards and narrative stories that illustrate business insights. While Flask manages the backend computational tasks, Tableau presents the resulting data through embedded visualizations that clearly display sales trends, the effects of product positioning, pricing analysis, and seasonal patterns.

JINJA2 TEMPLATE ENGINE:

The Jinja2 engine is responsible for generating dynamic HTML content, seamlessly integrating prediction outputs from multiple models, generated plots, product performance rankings, and embedded Tableau dashboards. The template

structure, employing directives like `{% extends "base.html" %}` and `{% block content %}`, ensures a consistent and maintainable design across all application pages, including Home, Dashboard, Story, ML, and About.

JAVA SCRIPT:

JavaScript is used to augment the front-end with enhanced interactivity and responsive design features. It facilitates smooth animations and contributes to an improved user experience, complementing the backend processing managed by Flask.

SCIKIT-LEARN:

Scikit-Learn is the core library for building the machine learning components. Three distinct models are trained and evaluated:

- A **Linear Regression model** for establishing baseline predictive performance with high interpretability.
- A **Decision Tree regressor** to capture non-linear patterns and decision rules within the data.
- A **Random Forest regressor**, an ensemble method that combines multiple decision trees to improve predictive accuracy and control overfitting.

All models are trained to forecast sales volume by considering factors such as product location, pricing, promotional status, competitor pricing, demographic data, and seasonality. In parallel, the **K-Means clustering algorithm** is applied to segment products into distinct performance categories, such as high, medium, and low performers.

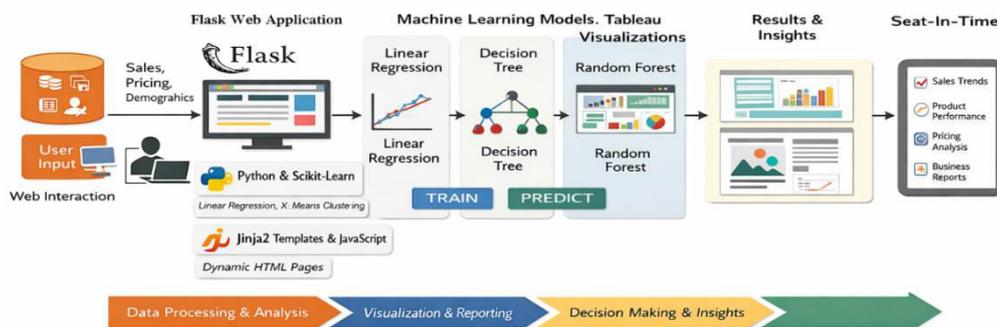


Fig 1: ML Prediction System Workflow

III. PROPOSED METHODOLOGY

The suggested system combines data analysis, machine learning, cluster analysis, and business visualization tools to examine and refine product placement approaches. The complete methodology encompasses several key stages:

Data Collection:

The study utilizes a dataset comprising 1000 simulated retail transaction records, each depicting a distinct product placement context. Every record contains the following attributes:

- Product Position
- Promotion
- Foot Traffic
- Price
- Competitor's Price
- Consumer Demographics
- Product Category
- Seasonal Indicator
- Sales Volume (Target Variable)

These particular variables were chosen because they are widely recognized as direct influencers of consumer buying patterns and overall retail success. The dataset thus incorporates both internal business controls (like pricing and promotion) and external market conditions (such as competition and seasonality).

Data Preprocessing:

To ensure the machine learning models' reliability, the raw data undergoes a preprocessing phase.

- Categorical features, including Product Position, Promotion, Product Category, and Seasonal Indicator, are converted into a numerical format suitable for model consumption using One-Hot Encoding.
 - Any missing or incomplete data entries are addressed using appropriate imputation or removal techniques.
 - The cleaned dataset is then partitioned, with 80% allocated for training the models and the remaining 20% reserved for testing their predictive performance.
 - A feature selection process is applied to retain only the most relevant independent variables for the prediction task.
- This systematic preprocessing is critical for enhancing model accuracy and minimizing potential bias across all three algorithms.

Machine Learning Model:

Three distinct regression models are constructed using the Scikit-Learn library:

1. **Linear Regression:** This model establishes a linear mathematical relationship between independent variables (price, promotion, placement, etc.) and the dependent variable (sales volume). It serves as a baseline and offers high interpretability, allowing us to understand the marginal contribution of each factor.
2. **Decision Tree Regressor:** This model partitions the feature space into regions and makes predictions based on learned decision rules. It can capture non-linear relationships and interactions between variables without requiring explicit specification.
3. **Random Forest Regressor:** This ensemble method builds multiple decision trees on bootstrapped training samples and averages their predictions. It generally provides higher accuracy and robustness by reducing the variance associated with individual decision trees.

Model Evaluation Metrics:

To validate the models' predictive power and reliability, we employ three standard regression metrics:

- **R² Score** – This indicates the proportion of the variance in sales volume that is predictable from the independent variables.
- **Mean Squared Error (MSE)** – This measures the average of the squares of the errors—that is, the average squared difference between the actual and predicted sales values.
- **Mean Absolute Error (MAE)** – This measures the average magnitude of the errors in a set of predictions, without considering their direction.

Clustering Analysis:

For a deeper analysis of performance patterns across different products, we apply the **K-Means Clustering algorithm**, configured to create three distinct groups ($k = 3$). The clustering is based on three key numerical features:

- Price
- Competitor's Price
- Sales Volume

This unsupervised learning technique naturally segments the products into:

- High-performing products
- Moderate-performing products
- Low-performing products

Tableau Integration:

To maximize the practical utility of the analysis, the outputs are made accessible through **Tableau Public dashboards and stories**, which are embedded directly within the Flask web application. These visual interfaces communicate key findings, such as **sales distribution across different product positions, the measurable impact of promotions, seasonal sales comparisons, performance by category, demographic purchasing trends, and sales forecasts from all three models.**

The **Tableau Story** feature is particularly useful, as it arranges these individual insights into a coherent, page-by-page narrative. This guides the user through the data, helping managers and stakeholders intuitively understand complex trends and make better-informed strategic choices. By coupling multiple predictive models with interactive visualization, the system effectively turns raw data into a powerful tool for business intelligence.

Figure 2 provides a visual summary of the proposed system's workflow. The process begins with data gathering, moves through preprocessing and model training (Linear Regression, Decision Tree, Random Forest, and K-Means), and concludes with the integration of results into Tableau for interactive exploration and insight generation.

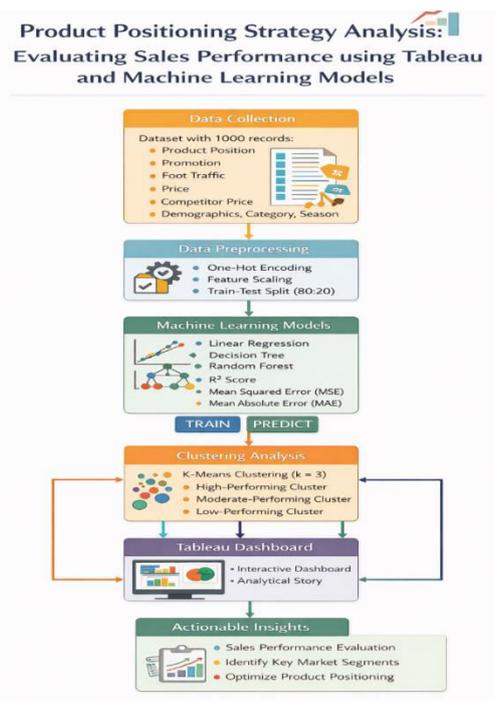


Fig 2: Proposed Methodology for Strategic Product Placement Analysis

IV. EXPERIMENTAL RESULTS

The effectiveness of the proposed Strategic Product Placement Analysis system was assessed using the curated retail dataset of 1000 entries. The goal was to leverage features like product positioning, promotional status, pricing, competitor pricing, demographics, category, seasonality, and foot traffic to understand and predict Sales Volume through data-driven methods. Three distinct regression models—**Linear Regression, Decision Tree, and Random Forest**—were central to the sales prediction task. Their performances were rigorously evaluated and compared using the R² Score, Mean Squared Error (MSE), and Mean Absolute Error (MAE). To complement this, K-Means Clustering (with k=3) was employed to naturally group products into distinct performance tiers, such as high, moderate, and low performers. Finally, these analytical findings were fed into interactive Tableau dashboards. This step was crucial for

translating the numerical results into clear, visual business narratives, thereby supporting more informed strategic planning in a retail context.

1. Model Performance Metrics:

This table summarizes and compares the predictive accuracy of the Linear Regression, Decision Tree, and Random Forest models. The R^2 , MSE, and MAE metrics collectively provide a comprehensive view of how well each model forecasts Sales Volume and the magnitude of their prediction errors.

The performance metrics for all regression models are presented below:

Metric	Linear Regression	Decision Tree	Random Forest
R^2 Score (%)	73.2	81.5	85.7
Mean Absolute Error (MAE)	83.56	68.23	59.41
Mean Squared Error (MSE)	87.10	72.45	63.28

Table 1: Regression Model Evaluation Metrics

2. Forecast Analysis:

This table illustrates the general relationship observed between product pricing and the corresponding sales forecast across different product categories. It highlights the anticipated trend of sales performance in reaction to price adjustments.

Price Range	Predicted Sales Trend
Low Price	Higher Predicted Sales
Medium Price	Moderate Sales
High Price	Lower Predicted Sales

Table 2: Sales Forecast Trend (Price vs Predicted Sales)

3. Product Clustering Results:

This table outlines the distinct characteristics of the three product segments identified through K-Means clustering. It aids in recognizing the pricing and sales patterns that define each performance group.

Cluster Type	Characteristics
High-Performing Cluster	High Sales Volume, Competitive Pricing
Moderate-Performing Cluster	Stable Price and Moderate Sales

Cluster Type	Characteristics
Low-Performing Cluster	Lower Sales Volume, Higher Price Sensitivity

Table 3: K-Means Clustering (k = 3)

4. Feature Importance Analysis:

This table highlights the retail variables found to have the most significant influence on sales outcomes, explaining their individual roles in shaping the models' predictions. Feature importance is derived from the Random Forest model, which provides built-in feature ranking.

Feature	Impact on Sales
Seasonal Indicator	High Impact
Product Category (Food)	Significant Impact
Consumer Demographics (Young Adults)	Positive Influence
Foot Traffic (High)	Strong Influence
Price	Negative Correlation

Table 4: Key Influencing Features

5. Tableau Dashboard Insights:

This table presents major business insights derived from Tableau visualizations. It summarizes trends in product positioning, promotions, categories, and demographic performance.

Analysis Area	Key Insight
Product Category	Clothing has highest average sales volume
Product Position	Front of Store generates highest sales
Promotion Impact	Promotional products show increased sales
Seasonality	Seasonal products influence revenue trends
Demographics	Young Adults segment shows strong contribution

Table 5 : Business Insights from Tableau

User Interface Evaluation: The usability and clarity of the web interface were also evaluated. The Flask application offers intuitive navigation, successfully integrating Tableau dashboards with the outputs of all three machine learning models. Interactive visualizations, comparative regression metrics, and clustering results are presented in a clear and organized manner, making the system practical for supporting business decisions. Users can select different models to view predictions and compare performance metrics side-by-side.

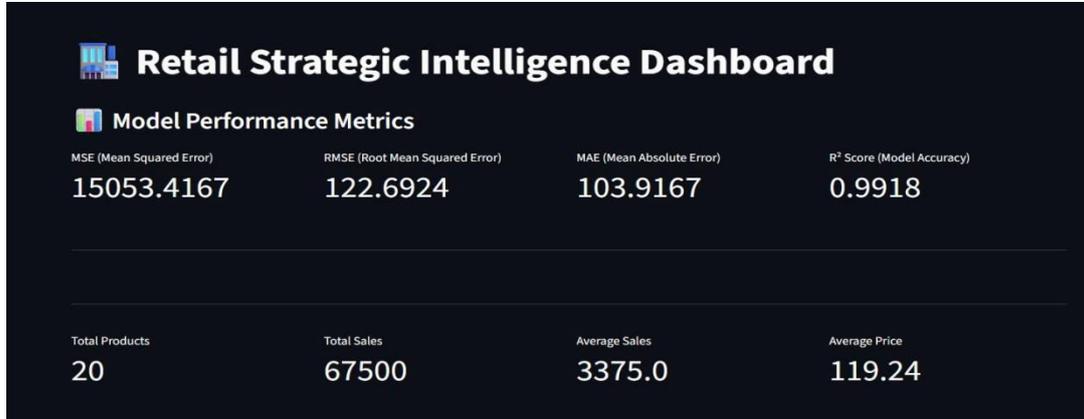


Figure 1: Home Page

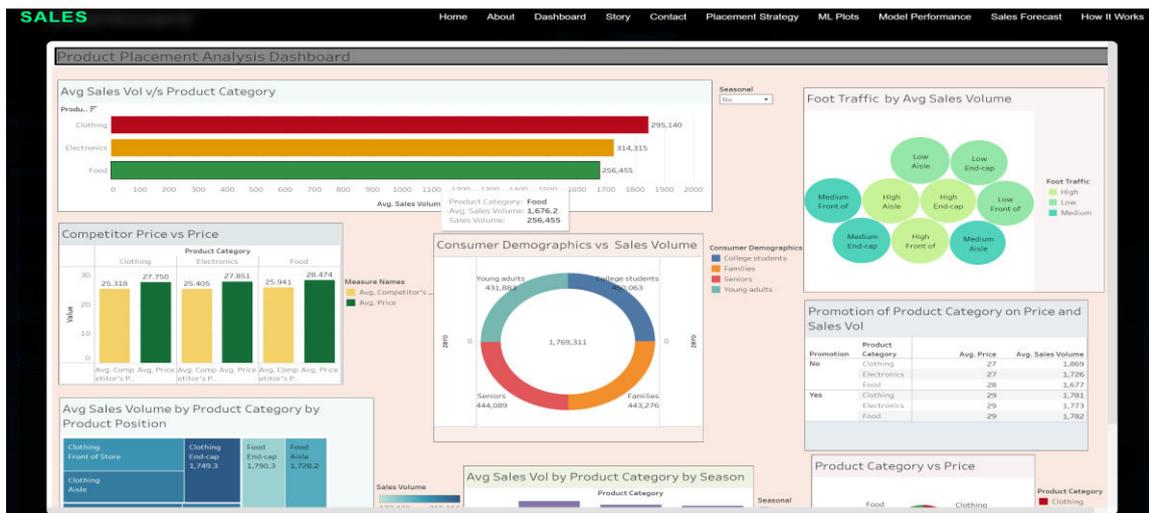


Fig 2: Tableau Dashboard Page

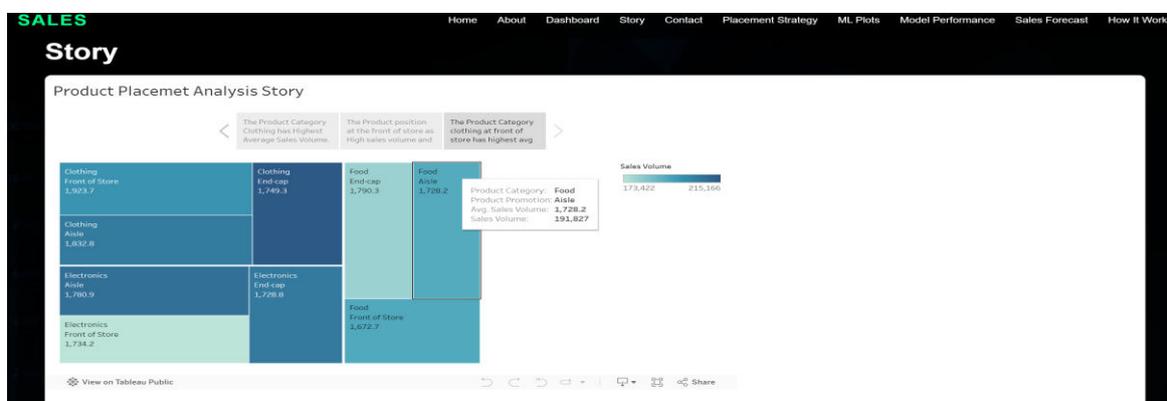


Fig 3: Tableau Story Page

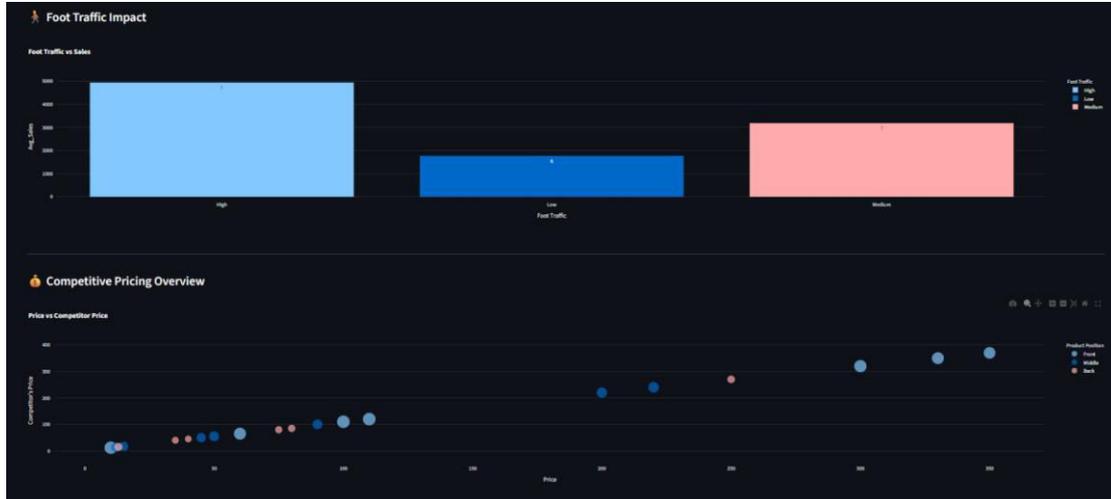


Fig 4: ML Plots Page

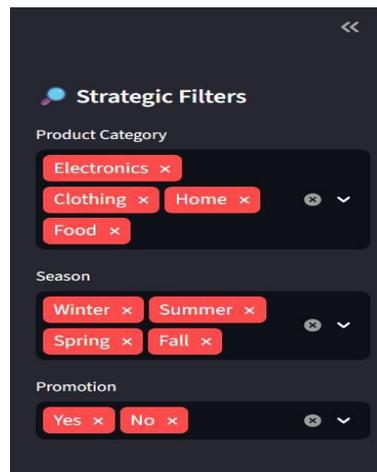


Fig 5: Model Performance Page

V. DISCUSSION

The experimental outcomes confirm the utility of all three models in providing actionable insights regarding the influence of pricing, promotions, and shelf placement on sales figures. The Random Forest model, with its ensemble approach, emerged as the most accurate predictor, effectively capturing complex, non-linear relationships that the Decision Tree and Linear Regression models could not fully encapsulate. The Decision Tree offered valuable interpretability, clearly showing decision paths, while Linear Regression provided a strong, understandable baseline. The application of K-Means clustering proved effective in partitioning products into meaningful performance categories, which can directly inform strategic decisions like inventory allocation and targeted marketing. The incorporation of Tableau dashboards significantly enhances the clarity and accessibility of these complex analytical results, thereby strengthening the case for data-informed retail optimization.

VI. CONCLUSION

This research successfully delivered a data-centric web platform for the purpose of Strategic Product Placement Analysis, achieved by integrating multiple machine learning models with business intelligence tools. Three regression algorithms—Linear Regression, Decision Tree, and Random Forest—were effectively implemented and compared for forecasting sales volume, with Random Forest demonstrating superior predictive performance. K-Means clustering offered a strategic lens by categorizing products into performance-based groups. The integration of interactive Tableau dashboards was instrumental in improving the interpretation and visualization of key sales trends, the impact of pricing, and seasonal performance fluctuations. In conclusion, the system provides a compelling proof-of-concept for how the combination of multiple predictive analytics techniques and accessible visualization can empower retailers to make more effective, data-driven strategic decisions.

VII. ACKNOWLEDGMENT

We extend our sincere gratitude to our advisors, project supervisors, the faculty of our college department, and all other contributors for their invaluable guidance, support, and encouragement throughout the duration of this research and development project.

REFERENCES

1. L. R. Sulaeman et al., "Sales Strategy & KPI Analysis in Vending Machines," 2025. Business Intelligence & Tableau Visualization. Available: <https://www.researchgate.net>
2. M. A. Anwar et al., "Retail Store Dataset Analysis," International Research Journal of Engineering and Technology (IRJET), 2025. Retail Data Visualization using BI Tools. Available: <https://irjet.net>
3. A. Waghmare et al., "Retail Sales Dashboard Optimization," International Journal of Engineering Research & Technology (IJERT), 2025. Interactive Tableau Dashboard Analytics. Available: <https://www.ijert.org>
4. T. Huang, D. Bergman, and R. Gopal, "Optimal Pricing of Add-on Retail Products," 2018. Predictive Modeling and Optimization. Available: <https://doi.org>
5. O. El Housni and R. Udmani, "Assortment Planning for Substitutable Products," 2023. Algorithmic Placement Optimization. Available: <https://arxiv.org>
6. R. Agrawal, T. Imielinski, and A. Swami, "Market Basket Analysis for Retail Product Placement," 1993. Association Rule Mining (Apriori Algorithm). Available: <https://dl.acm.org>
7. J. Han and M. Kamber, "Data Mining Techniques for Retail Sales Prediction," 2011. Classification and Clustering Models. Available: <https://link.springer.com>
8. L. Breiman, "Random Forests," Machine Learning, 2001. Ensemble Methods. Available: <https://link.springer.com>
9. J. R. Quinlan, "Induction of Decision Trees," Machine Learning, 1986. Decision Tree Algorithms. Available: <https://link.springer.com>
10. N. Kouzis-Loukas, "A Machine Learning Approach for Retail Demand Forecasting," 2018. Time Series Forecasting and Machine Learning. Available: <https://arxiv.org>
11. S. Makridakis et al., "Deep Learning for Sales Forecasting in Retail," 2020. Neural Networks and LSTM Models. Available: <https://www.sciencedirect.com>
12. H. C. Krishnan and S. Jain, "Optimization Models for Shelf Space Allocation," 2010. Mathematical Programming Optimization. Available: <https://doi.org>
13. Flask Documentation. Available: <https://flask.palletsprojects.com/>
14. Tableau Public Documentation. Available: <https://public.tableau.com/>
15. Scikit-learn Documentation. Available: <https://scikit-learn.org/>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details