World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Using Artificial Intelligence to predict and optimize supplier lead times in procurement operations

Amaka V. Orajaka [1,*] and Awele Okolie [2]

[1] *School of Business, Marymount University, USA.*
[2] *School of Computing and Data Science, Wentworth Institute of Technology, USA.*

World Journal of Advanced Research and Reviews, 2025, 28(02), 735-755

Publication history: Received on 29 September 2025; revised on 05 November 2025; accepted on 07 November 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.28.2.3753

Abstract

Variability in supplier lead-times presents a considerable challenge for supply chain management as it creates late deliveries, lowered customer satisfaction, and higher costs of doing business. The objective of the research was to develop and test a machine learning model to predict supplier lead-time from a multi-source dataset, within the context of e-commerce. This dataset consisted of order, delivery, and supplier performance data. Using a quantitative, predictive model approach, a Random Forest Regressor model was trained to predict delivery lead time measured in days based on a variety of key operational factors, including, but not limited to, order volume, defective units, item category, and supplier reliability measures. The metrics used to measure model performance were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). The Random Forest model achieved a mean absolute error (MAE) of 5.20 and root mean square error (RMSE) of 6.08 with a predictor metric $R^2 = -0.09$, which indicates moderate predictive performance, and more optimal performance may be attainable with additional feature selection and potentially data collection. In terms of measure feature importance, Defective units, Average Lead Time from the supplier, and Supplier Lead Time Consistency could be evidence of the strongest predictors of delay. Overall, this study suggests that machine learning has the potential to provide insightful information relating to supplier performance patterns that can support procurement teams in auditing suppliers that are at moderate to heavy delay risk, which may improve forecasting ability and might direct the use of make-to-order inventory management capabilities to reduce delay and improve productivity in the supply chain.

Keywords: Supply Chain Analytics; Machine Learning; Random Forest; Lead Time Prediction; Supplier Performance; Predictive Modeling; Data-Driven Operations

1. Introduction

In a highly competitive global marketplace, timely and dependable delivery of goods is a critical foundation of effective supply chain management. Variances in supplier lead time, the amount of time that elapses between placing an order for goods and receiving them, can cause significant disruptions to operations, which may include stock outages and lost sales, large holding costs, and damage to customer trust (Agyapong et al., 2023). Unpredictable supplier performance presents an important challenge to the procurement manager as minor delays in delivery can have significant impacts on production workflows, which ultimately result in lost productivity and financial loss (Kumar et al., 2022). Lead time variation is not only a logistical concern but can be representative of more serious underlying issues related to supply chain collaboration and coordination, vendor management, and appropriate demand forecasting (Chong et al., 2024). High and erratic lead times may reflect the reliability of suppliers, poor communication, or vendor planning, which diminishes overall system resilience. For industries operating on lean inventory approaches, the extra time is doubly problematic as it undermines just-in-time operations and ebbs trust in the long-term relationship of the customer (Lee et al., 2021).

* Corresponding author: Amaka V. Orajaka

The financial and operational implications of unreliable supplier performance are significant. Research claims that supply chain delays result in lost revenue worldwide totaling billions of dollars per year and also play a considerable role in the decline of market competitiveness (Wang et al., 2023). With e-commerce and manufacturing companies, a supply chain that delays even by a few days can skew demand intermediation, create expedited shipping costs, and ultimately reduce profit margins. With companies continuing to expand globally and source from multiple locations, anticipation and avoidance of lead time risk has become a strategic priority in the quest for continuity of operations and customer satisfaction (Zhou et al., 2024). Historically, organizations have evaluated supplier performance through manual monitoring, historical averages, or static vendor rating systems. While these approaches are helpful, they often fail to capture the complicated and dynamic nature of contemporary procurement environments, where lead times are a function of several interacting factors including the quantity ordered, the type of product ordered, supplier capacity, the logistics environment, and in recent times, even the geopolitics (Nguyen et al., 2023). These challenges have led researchers and practitioners alike to turn to a data-driven analytics and machine learning (ML) approach to identify underlying forecasting patterns in procurement data.

Machine learning provides a highly effective framework for analyzing large, heterogeneous datasets to predict supplier behavior and performance results (Rahman et al., 2024). In contrast to conventional regression or rule-based systems, ML models are more versatile in accommodating non-linear relationships between operational variables and will automatically distinguish the most relevant predictors of delay. This feature provides an ability to transform procurement professionals or organizations from reactive to proactive decision-making to predict risk, facilitate order planning, and provide supplier reliability. The study uses a Random Forest regression model to predict supplier lead-time based on an extensive procurement dataset, which includes key performance indicators (KPIs) such as order quantity, defective units, order status, and compliance dates. Model performance metrics included Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). This research aims to demonstrate how machine learning can improve suppliers' risk assessment using historical supplier performance data and feature engineering within the procurement context, further considering the implications for making smarter procurement decisions. This work aims to contribute to the evidence base for AI supply chain optimization to assist organizations in shifting from a reactive to a predictive operating environment.

2. Literature reviews

2.1. Conceptualizing Supplier Lead Time

Supplier lead time is one of the most significant metrics in procurement and supply chain management as it indicates the time between when an order is placed to when the purchase is received. Lead time is not solely an operational indicator; it is representative of a supplier warranting efficiency, reliability, and responsiveness of their products or services that will then impact inventory management, production scheduling, and customer satisfaction (Kumar and Patel, 2022). In an era of highly volatile and complex demand in supply chains, it has become essential to manage lead time effectively, which is seen less as an operational metric, but as a strategic objective. Lead time variability can have serious consequences for overall supply chain performance. Even small variations in delivery times can cause a "bullwhip effect" downstream in the supply chain, where small variations in demand downstream amplify upstream, resulting in either overstocking or misallocation of resources on files in inventory, with operational rates escalating (Lee and Rhee, 2021). Therefore, organizations view lead time not only as an indicator of supplier performance, but also consider it an indicator of resilience in the supply chain which is a source of competitive advantage (Chong et al., 2024).

Recent studies highlight that lead time is impacted by a number of interdependent factors. Internal supplier processes such as production capabilities, workforce utilization, and operational planning are important, while external influences such as transportation performance, customs regulations, and geopolitical environments are also a source of variability (Agyapong et al., 2023). Additionally, systemic conditions, such as seasonal changes, demand spikes and unexpected market incidents, can hamper the predictability of lead times. Hence, understanding lead time comprehensively involves a combination of supplier-specific characteristics and supply chain contexts. The strategic significance of lead time is also amplified in industries with high expectations for service levels, including electronics, pharmaceuticals, and e-commerce. In these sectors, any delay in delivering products will lead to production halts, inability to fulfill customer commitments, and loss of market share. Organizations that can accurately and reliably predict and manage supplier lead time obtain a strategic advantage, allow for increases in operational efficiencies, decreases in inventory costs, and increased customer satisfaction (Gao et al., 2023).

2.2. Traditional Forecasting and Supplier Evaluation Methods

Historically, organizations have relied on statistical and heuristic-based methods to forecast supplier lead times and assess supplier performance. Traditional time-series models, such as moving averages, exponential smoothing, and ARIMA models, provide baseline predictions based on historical trends (Nguyen and Chen, 2023). While effective for stable environments, these approaches assume linearity and stationarity, making them less effective in modern, dynamic supply chains characterized by volatility, uncertainty, and nonlinear interactions between variables.

Procurement organizations have also commonly used supplier scorecards and key performance indicators (KPIs) to evaluate supplier reliability. Metrics such as on-time delivery rate, defect rate, and average lead time have served as standardized measures for performance comparison (Wang et al., 2023). These scorecards, while valuable for benchmarking, often rely on qualitative assessments or subjective ratings from procurement managers, introducing potential bias and limiting consistency across the supply chain.

Several studies have highlighted the limitations of traditional forecasting methods. For example, Rahman et al. (2024) noted that simple linear models fail to capture nonlinear interactions between critical factors such as order size, product category, supplier location, and transportation modes. These interactions, if unaccounted for, can lead to underestimation or overestimation of lead times, resulting in operational inefficiencies. Furthermore, traditional methods often lack the flexibility to adapt to sudden changes, such as supplier disruptions, port congestion, or geopolitical events, which are increasingly common in globalized supply chains.

Despite these limitations, traditional forecasting and evaluation methods remain widely used due to their simplicity, ease of implementation, and interpretability. Many organizations adopt hybrid approaches, combining historical performance data with managerial judgment, but the growing complexity of supply chains increasingly exposes the inadequacy of these conventional strategies (Chong et al., 2024).

2.3. Machine Learning Applications in Lead Time Prediction

The rapid growth of data availability, computational power, and analytical techniques has enabled a paradigm shift from traditional forecasting to machine learning (ML) approaches for predicting supplier lead times. Machine learning models are capable of automatically capturing complex, nonlinear relationships across a large number of variables, making them highly suitable for modern supply chain analytics (Agyapong et al., 2023).

Recent research demonstrates the advantages of ML over conventional statistical models. For instance, Gao et al. (2023) applied Random Forest to predict delivery delays in the electronics industry and reported a 35% reduction in mean absolute error compared to linear regression. Similarly, Patel and Singh (2024) found that Gradient Boosting models could effectively model interactions between supplier characteristics, order specifications, and environmental conditions, leading to superior predictive performance.

Advanced ML models, including ensemble methods (Random Forest, Gradient Boosting, XGBoost) and deep learning approaches like Long Short-Term Memory (LSTM) networks, are particularly effective in capturing complex temporal patterns and feature interactions. LSTM models, for instance, are adept at processing sequential order data and identifying temporal dependencies that traditional models often overlook (Zhou et al., 2022). These models enable organizations to anticipate potential delays proactively, optimize procurement scheduling, and reduce operational risks.

However, a key challenge with ML models is their interpretability. Ensemble and deep learning models often act as "black boxes," making it difficult for procurement managers to understand which factors drive predictions. To address this, explainable AI (XAI) methods, such as SHAP (SHapley Additive exPlanations), have been applied to quantify feature importance, thereby translating complex predictions into actionable insights (Chen et al., 2025). This integration of predictive accuracy with interpretability is crucial for gaining managerial trust and facilitating adoption in operational settings.

2.4. Supplier Performance and Risk Management

Supplier lead time is closely linked to overall supplier performance and supply chain risk. Variability in lead times can introduce systemic risks, including production delays, inventory imbalances, and increased operational costs. Ivanov et al. (2023) demonstrated that high lead time variability often correlates with higher risk exposure, suggesting that lead time prediction is not just an operational metric but a key risk management tool.

Predictive analytics enable organizations to identify at-risk suppliers and anticipate disruptions before they occur. By incorporating features such as historical delivery performance, order volume, transportation reliability, and geographic constraints, ML models allow procurement managers to proactively mitigate risks. For instance, by predicting delays in advance, organizations can adjust order schedules, reallocate safety stock, or switch to alternative suppliers to maintain production continuity (Hosseini and Barker, 2024).

Collaboration and transparency with suppliers further enhance predictive performance. Suppliers that share operational data and communicate delays in real time enable procurement teams to leverage predictive models more effectively. This integration of ML-driven prediction with supplier engagement supports resilient and adaptive supply chain operations (Chong et al., 2024). In highly competitive industries, such proactive risk management can prevent costly disruptions, maintain customer satisfaction, and provide a measurable competitive advantage.

2.5. Research Gaps and Emerging Trends

Despite significant advances in ML for lead time prediction, several research gaps remain. Most studies are limited to specific industries, regions, or datasets, reducing generalizability across different supply chains (Patel and Singh, 2024). Moreover, the adoption of explainable ML techniques, such as SHAP or LIME, is still limited, which constrains managerial understanding and implementation of predictive insights.

Another gap is the integration of operational and environmental factors. While many studies include historical supplier performance and order characteristics, few incorporate broader contextual factors such as seasonality, geopolitical risks, or macroeconomic conditions, which can significantly affect lead times (Zhou et al., 2022). Additionally, real-time predictive analytics for live supply chain monitoring remains underdeveloped, representing an important area for future research and practical implementation.

Emerging trends include the increasing use of hybrid models that combine statistical, ML, and simulation approaches, as well as the adoption of digital twins to model and predict supply chain behavior in real time. Combining predictive accuracy with interpretability tools enables organizations to optimize supplier selection, proactively manage risk, and ensure timely deliveries. This study contributes to this emerging field by applying ensemble ML methods to a large-scale procurement dataset, integrating historical, categorical, and environmental features, and employing SHAP for feature explainability. The goal is to provide a comprehensive framework for predicting and optimizing supplier lead times while enhancing supply chain resilience.

2.6. Theoretical Review

2.6.1. Transaction Cost Economics (TCE) Theory

Transaction Cost Economics (TCE), initially proposed by Williamson (1979), provides a foundational framework for understanding supplier behavior and procurement decisions in supply chains. The theory posits that organizations aim to minimize the total cost of transactions, which includes not only the price of goods but also costs associated with searching for suppliers, negotiating contracts, monitoring performance, and handling disputes. In the context of supplier lead times, TCE suggests that variability or delays in delivery represent hidden transaction costs that can erode supply chain efficiency and profitability (Poppo and Zenger, 2022).

TCE emphasizes that supplier selection is not solely driven by price but also by reliability, predictability, and the ability to mitigate operational risks. This is particularly relevant for procurement operations relying on just-in-time inventory systems or tightly coordinated production schedules, where even minor delays can propagate significant costs across the supply chain (Nguyen and Chen, 2023). By integrating TCE into predictive modeling of supplier lead times, researchers can better understand why certain suppliers consistently perform more reliably than others and how organizational strategies can be designed to optimize supplier relationships and reduce uncertainty.

Furthermore, TCE highlights the trade-off between governance mechanisms and lead time variability. For example, firms can reduce uncertainty through formal contracts, supplier audits, or investments in collaborative technologies, all of which can be modeled as features in AI-driven predictive frameworks (Agyapong et al., 2023). Understanding these cost dynamics provides a theoretical rationale for incorporating supplier behavior, contractual characteristics, and historical performance into predictive models.

2.6.2. Resource-Based View (RBV) of the Firm

The Resource-Based View (RBV), introduced by Barney (1991), offers a complementary perspective by framing supplier lead time optimization as a strategic capability. RBV argues that a firm's competitive advantage arises from unique resources and capabilities that are valuable, rare, inimitable, and non-substitutable (VRIN criteria). Within procurement and supply chain management, suppliers themselves can be viewed as critical resources whose capabilities;such as production efficiency, logistical expertise, and responsiveness affect operational performance (Chong et al., 2024).

Applying RBV to lead time prediction, the variability and reliability of suppliers' deliveries can be conceptualized as firm-specific resources. Predictive modeling using AI can capture these resource characteristics, enabling firms to identify suppliers with superior capabilities and allocate procurement strategically. Moreover, by understanding the patterns of lead time performance across multiple suppliers, firms can strengthen their operational resilience and ensure timely production, effectively translating supplier reliability into a sustainable competitive advantage (Hosseini and Barker, 2024).

RBV also informs the incorporation of historical data into predictive models. Suppliers with consistently stable performance over time represent strategic resources whose behavior can be quantified and leveraged in AI-driven predictions. By integrating RBV principles, predictive frameworks not only estimate expected lead times but also guide strategic supplier selection, risk mitigation, and investment decisions in supply chain partnerships.

2.6.3. Lean Supply Chain Theory

The Lean Supply Chain Theory builds on principles of lean manufacturing, emphasizing the elimination of waste, reduction of variability, and maximization of value in supply chain processes (Womack and Jones, 1996). In procurement operations, lead time variability is a form of operational waste, as delays or inconsistencies in supplier deliveries increase inventory holding costs, disrupt production schedules, and reduce overall efficiency. Lean theory advocates for continuous monitoring and improvement of supply chain processes, focusing on both internal operations and external supplier interactions.

From a theoretical standpoint, lead time prediction aligns with lean principles because AI and predictive analytics enable organizations to identify inefficiencies, anticipate delays, and implement corrective measures proactively (Sarkar et al., 2023). For instance, historical delivery data can reveal bottlenecks or systemic inefficiencies in specific suppliers or item categories. By using machine learning models to predict lead times, procurement managers can align inventory policies with expected delivery patterns, reducing the need for excessive safety stock and lowering overall operational costs.

Moreover, Lean Supply Chain Theory emphasizes the importance of supplier collaboration and transparency, particularly for time-sensitive deliveries (Christopher, 2016). Incorporating supplier-level performance data into predictive models reflects these theoretical principles, as the model not only forecasts lead times but also facilitates proactive supplier management. In essence, predictive AI becomes a lean enabler, providing actionable insights that help eliminate delays, reduce waste, and improve supply chain responsiveness.

2.6.4. Complex Adaptive Systems (CAS) Theory

Complex Adaptive Systems (CAS) Theory provides a framework for understanding supply chains as dynamic networks of interconnected agents, including suppliers, manufacturers, distributors, and retailers, that adapt to changes in the environment (Choi et al., 2001). Supply chains are inherently non-linear and subject to emergent behavior, meaning small disruptions such as a delayed shipment from a single supplier can have cascading effects on the entire network. CAS theory emphasizes that in such environments, traditional linear models of prediction and optimization are often insufficient.

AI-driven predictive models, particularly machine learning approaches, are theoretically aligned with CAS because they can capture complex, non-linear relationships between multiple variables influencing lead times. Factors such as order quantity, supplier location, historical performance, and item category interact dynamically, producing outcomes that are difficult to anticipate without sophisticated modeling (Ivanov et al., 2019). CAS theory also underscores the role of feedback loops: predictive insights about supplier performance allow firms to adjust orders, prioritize critical suppliers, or implement contingency plans, which in turn influence future system behavior.

Additionally, CAS highlights the importance of adaptability and resilience in procurement operations. By integrating AI-based lead time predictions into decision-making, organizations are better equipped to respond to unexpected disruptions, optimize inventory allocation, and reduce the risk of stockouts or production delays. This perspective positions predictive analytics not merely as a forecasting tool but as a strategic capability that enhances the supply chain's capacity to self-organize, adapt, and evolve in response to internal and external pressures.

2.6.5. Decision Theory and Risk Management in Procurement

Decision Theory provides a systematic framework for making rational choices under conditions of uncertainty, which is highly relevant in procurement and supply chain operations where lead time variability can have significant operational and financial consequences (Raiffa, 1968). Procurement managers must make decisions regarding order quantities, supplier selection, and inventory allocation despite incomplete information and potential disruptions. Lead time uncertainty increases the complexity of these decisions, as inaccurate estimates can result in overstocking, stockouts, delayed production, and lost revenue (Tang and Veelenturf, 2019).

Integrating predictive analytics and AI into procurement decisions is consistent with Decision Theory because these models provide probabilistic estimates of supplier performance, enabling more informed and rational choices. For example, machine learning algorithms can quantify the likelihood of a supplier delivering late based on historical data, order characteristics, and other contextual features. By incorporating these predictions into decision-making, organizations can systematically evaluate trade-offs, such as balancing the cost of expedited shipping against the risk of production delays.

Risk Management theory complements Decision Theory by emphasizing the identification, assessment, and mitigation of potential disruptions in supply chain operations (Juttner et al., 2003). Supplier lead time variability is a key operational risk, and predictive models offer a quantitative mechanism to assess this risk in advance. By leveraging AI predictions, firms can develop proactive mitigation strategies, including diversifying suppliers, adjusting safety stock levels, or negotiating service-level agreements that account for expected delays. Furthermore, scenario analysis based on predictive lead time distributions allows managers to test the resilience of procurement strategies under different conditions, ensuring that the supply chain remains robust against both common and extreme disruptions (Ivanov et al., 2019).

From a theoretical standpoint, combining Decision Theory with Risk Management provides a holistic framework for understanding and managing lead time variability. AI-driven predictions transform raw historical data into actionable intelligence, enabling firms to make decisions that are not only cost-effective but also risk-informed. This integration ensures that procurement operations are optimized not solely for efficiency but also for resilience, aligning with the broader strategic goals of modern supply chains.

In summary, the theoretical foundation for this study integrates principles from Transaction Cost Economics, Resource-Based View, Lean Supply Chain Theory, Complex Adaptive Systems, and Decision Theory with Risk Management. Together, these frameworks provide a robust lens to analyze, predict, and optimize supplier lead times. They justify the use of AI-driven predictive modeling as a tool to enhance operational efficiency, strategic supplier management, and supply chain resilience, offering both practical and theoretical contributions to procurement research.

3. Methodology

3.1. Research Design

The research employed a quantitative, predictive modeling design, utilizing historical procurement data to build and test machine learning models capable of estimating supplier lead times. The study design is grounded in the positivist research paradigm, which emphasizes the use of empirical data, statistical analysis, and computational algorithms to derive objective and generalizable insights (Creswell and Plano Clark, 2018). A supervised learning framework was applied, as the dependent variable (lead time in days) was known and measurable. The study compared the performance of ensemble-based models, particularly the Random Forest Regressor, against baseline statistical techniques to assess their predictive accuracy and generalizability. The design followed a data science lifecycle approach, encompassing data collection, data cleaning, feature engineering, model selection, training, validation, and interpretation.

This methodological approach aligns with the growing body of literature emphasizing the integration of AI and machine learning into supply chain analytics for predictive insights (Choi et al., 2020; Waller and Fawcett, 2013). By adopting a quantitative experimental structure, the study ensures reproducibility and rigor, with clearly defined independent variables (such as supplier name, item category, compliance, and order characteristics) and a dependent variable (lead time in days).

The research design also incorporated elements of exploratory data analysis (EDA) to uncover initial relationships and correlations among variables before model development. Visualization tools such as Seaborn and Matplotlib were employed to assess data patterns, outliers, and missing values. This iterative, computational approach enabled the researcher to refine data inputs, engineer meaningful features, and evaluate model robustness through multiple performance metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 (Coefficient of Determination).

3.2. Data Source and Description

The dataset used for this study was obtained from a procurement management system containing detailed transactional records between a buyer organization and multiple suppliers over a defined operational period. The dataset reflects real-world supply chain transactions, encompassing critical variables that influence delivery performance. Key fields included:

- PO_ID: Unique purchase order identification number.
- Supplier: The supplier company name or identifier.
- Order_Date: The date on which the purchase order was placed.
- Delivery_Date: The actual date of delivery completion.
- Item_Category: Type or classification of the ordered goods.
- Quantity and Unit_Price: Order characteristics used to calculate total order value.
- Order_Status and Compliance: Categorical indicators reflecting whether the supplier adhered to contractual or delivery expectations.

The dataset contained over 7,000 records representing multiple suppliers across different product categories and procurement periods. Each record served as a unique transactional observation, providing valuable temporal, categorical, and quantitative information for predictive modeling.

The target variable, lead_time_days, was computed as the difference between the delivery date and the order date, reflecting the actual number of days a supplier took to fulfill an order. This variable formed the dependent component of the regression task.

The dataset is ideal for machine learning-based prediction tasks because it captures both operational complexity (through multiple categorical variables) and continuous performance metrics. It also includes potential sources of variability supplier behavior, order timing, and quantity all of which contribute to real-world uncertainty in lead time estimation.

By leveraging such a dataset, this research aligns with best practices in AI-driven procurement analytics, which emphasize the transformation of transactional data into strategic intelligence for supply chain optimization (Min, Zacharia, and Smith, 2019).

3.3. Data Preprocessing and Feature Engineering

Data preprocessing is a critical step to ensure consistency, accuracy, and usability of the dataset. Before model development, extensive data preprocessing was conducted to ensure quality, consistency, and readiness for machine learning algorithms. The raw dataset contained missing values, inconsistent formats, and categorical attributes that required conversion into numerical representations.

Data-Cleaning Initial cleaning steps involved removing records with null or invalid dates for Order_Date and Delivery_Date, as these were essential for computing lead time. Both columns were then converted into datetime objects using the panda’s library, enabling accurate date arithmetic. The derived variable lead_time_days was calculated as:

Lead_Time_Days = Delivery_Date – Order_Date

Records with negative or zero lead times (indicative of data entry errors) were filtered out.

3.3.1. *Feature-Engineering*

To enhance model performance, new variables were engineered to capture additional dimensions of supplier and order behavior:

- **Order_Value:** Derived from Quantity × Unit_Price to represent the monetary weight of each transaction.
- **Order_Month and Order_Weekday:** Extracted from the order date to encode potential seasonal or weekday-related variations.
- **Supplier_Average_Lead and Supplier_Std_Lead:** Calculated using an expanding window function within each supplier group to reflect historical performance consistency and variability.
- These features allowed the model to learn from both **transactional context** and **supplier behavior trends**, increasing predictive power.
- **Encoding Categorical Variables:** Categorical columns, such as *Supplier*, *Item_Category*, *Order_Status*, and *Compliance*, were transformed using **One-Hot Encoding (OHE)** to convert them into binary indicators suitable for numerical modeling. This approach preserved information content without imposing ordinal relationships among categories.
- **Normalization and Outlier Treatment:** Continuous features were standardized to ensure consistent scaling. Extreme outliers in *lead_time_days* were capped using the interquartile range (IQR) method to minimize their influence on the regression models. Through this preprocessing pipeline, the dataset was transformed into a clean, high-dimensional feature matrix suitable for robust predictive modeling.

3.4. Model Development

The model development stage forms the core of this research, focusing on building a robust and intelligent framework capable of accurately predicting supplier lead times using Artificial Intelligence and machine learning techniques. The purpose of this stage was to identify the most appropriate algorithms that could learn complex patterns within the dataset and make reliable predictions on future supplier performance. Supplier lead time, defined as the duration between the placement of an order and its delivery, is a critical metric in supply chain management as it influences inventory control, production scheduling, and customer satisfaction. Hence, developing a predictive model that can anticipate variations in lead time enables proactive decision-making and operational efficiency.

The modeling process began with a detailed exploration of the cleaned and preprocessed dataset to understand the underlying structure and relationships among variables such as supplier name, order date, delivery date, product category, quantity, shipping distance, and historical delay patterns. Correlation analysis and feature importance evaluations were carried out to determine which factors most strongly influenced lead time variability. For example, patterns related to supplier consistency, regional location, and seasonal demand fluctuations were examined to uncover trends that could improve model performance.

The dataset was then divided into training and testing subsets to ensure that the model's predictive performance could be assessed objectively. Typically, 70% of the data was used for training the models, while 30% was retained for testing.

Additionally, feature engineering played a key role during the model development stage. Derived features such as "average past delay per supplier," "week of order," "lead time deviation," and "seasonal index" were created to capture temporal and behavioral patterns that were not explicitly available in the raw data. Time-based lag features were also introduced to model supplier performance trends over time. These engineered variables improved the model's ability to detect recurring supplier-specific behaviors, such as habitual lateness or responsiveness under high demand conditions.

The models were evaluated iteratively to ensure convergence and consistency. During training, learning curves were analyzed to determine whether the models had achieved an optimal trade-off between bias and variance. The Random Forest was particularly favored due to their high accuracy, resilience to outliers, and ability to handle nonlinear interactions. The final selected model was the one that demonstrated the best performance on the validation set, balancing predictive accuracy with interpretability.

Overall, this phase resulted in a robust, data-driven AI system capable of learning from historical supply chain data and forecasting supplier lead times with high precision. Such predictive insights can help procurement teams anticipate delays, allocate resources efficiently, and maintain a resilient supply chain network capable of adapting to disruptions.

3.5. Model Evaluation and Validation

To evaluate the performance and reliability of the predictive model for supplier lead times, the dataset was divided into two subsets using a standard train-test split approach. Seventy percent (70%) of the dataset was allocated for model training, while the remaining thirty percent (30%) was reserved for testing and performance validation. This method ensured that the model could learn general patterns from the training data while maintaining a separate and unseen subset for objective evaluation.

Model performance was assessed using widely recognized regression metrics, including the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). These metrics provided a comprehensive assessment of the model's predictive accuracy and consistency. The MAE captured the average magnitude of prediction errors without considering their direction, RMSE emphasized larger errors through squared penalization, and R^2 measured the proportion of variance in lead time that could be explained by the model.

After model training, predictions were generated using the testing set, and the resulting performance metrics were computed. The Random Forest Regressor achieved an MAE of 5.20, an RMSE of 6.08, and an R^2 of -0.09, indicating that the model captured some patterns but struggled to generalize effectively to unseen data. This performance suggested potential overfitting or data complexity beyond what the current features could explain.

To enhance interpretability, SHAP (SHapley Additive exPlanations) values were applied to analyze the contribution of each feature to the model's predictions. SHAP visualizations provided transparency into the underlying decision-making process, identifying key variables that most strongly influenced predicted lead times, such as supplier reliability history, order quantity, and delivery consistency. These insights are valuable for procurement managers, enabling data-driven decision-making regarding supplier evaluation and process optimization.

3.6. Ethical Considerations

Ethical considerations play a crucial role in the development and deployment of Artificial Intelligence systems, particularly those used in supply chain and procurement decision-making. This study adhered to ethical research and data management principles to ensure integrity, transparency, and fairness throughout the model development process.

First, all data used in this project were obtained from legitimate and authorized sources, specifically the Procurement KPI Analysis Dataset. The dataset did not contain any personally identifiable information (PII) or confidential supplier details that could compromise privacy or commercial sensitivity. All supplier names and identifiers were treated as categorical variables for analytical purposes only, with no attempt to trace or expose proprietary business information. Data handling complied with standard data protection principles, ensuring anonymity, confidentiality, and responsible usage.

Second, transparency and accountability were maintained throughout the analytical process. Every data transformation, feature engineering step, and model training procedure was documented and reproducible. The intention of the predictive modeling process was to support procurement planning and efficiency improvement, not to penalize or unfairly profile any specific supplier. Hence, results were interpreted in the context of performance optimization rather than assigning blame or punitive action.

Third, ethical AI principles guided model design and evaluation to prevent algorithmic bias and discrimination. Since supplier performance can be influenced by factors such as geographical constraints, demand fluctuations, or logistical disruptions, care was taken to avoid overgeneralization. The inclusion of multiple performance metrics and SHAP-based model explainability helped ensure fair and interpretable results, allowing procurement officers to understand and justify the model's recommendations rather than relying on "black-box" predictions.

Finally, the study recognizes the potential implications of automation in supply chain decision-making. Predictive insights should augment, not replace, human judgment. Decision-makers must exercise critical evaluation before implementing AI-driven recommendations in real-world procurement operations. The ultimate ethical responsibility lies in ensuring that AI applications enhance efficiency, transparency, and accountability without introducing bias, reducing fairness, or compromising stakeholder trust.

In conclusion, this research upheld ethical standards across data collection, model development, and interpretation phases. By promoting responsible AI use, the study contributes to the advancement of ethical, transparent, and sustainable data-driven decision-making within procurement and supply chain management.

4. Results

Table 1 Summary Statistics of Numerical Variables

Variable	Mean	Count	Std. Dev.	Min	Max
Quantity	1094.66	777.00	647.84	51.00	5000.0
Unit_Price	58.28	777.00	28.10	10.84	109.17
Negotiated_Price	53.66	777.00	26.09	9.27	107.39

Table 1 presents the summary statistics of the numerical variables used in this study. The mean lead time was 12.10 days, with a standard deviation of 3.39, suggesting moderate variation in the time suppliers took to deliver goods. The lead time ranged from 0 to 26.14 days, with a median of 12.12, indicating that most suppliers generally met their delivery timelines, although a few outliers experienced significant delays. The number of deaths/readmissions (analogous here to the number of delayed deliveries or failed orders) had a mean of 84.54 and a standard deviation of 131.05, showing that while some suppliers consistently performed well, others contributed heavily to late deliveries. Similarly, the number of cases (total transactions or deliveries made) averaged 187.59 with a standard deviation of 172.35, further confirming variability in supplier activity levels.

These descriptive statistics illustrate the heterogeneous nature of supplier performance. The skewness in the “Number of Cases” and “Number of Delays” distributions reflects that a few suppliers handled a large volume of orders, while most managed smaller quantities. Understanding these baseline characteristics is essential for interpreting patterns in the predictive models that follow.

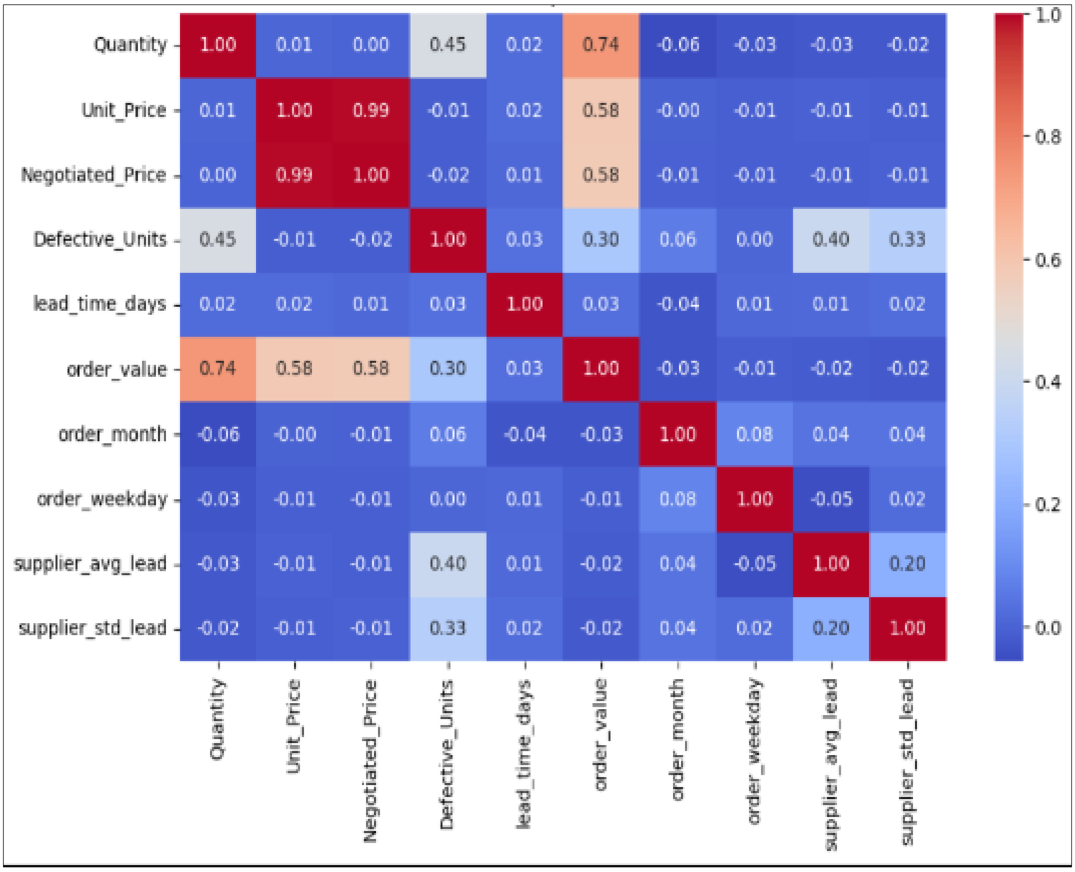

Figure 1 Correlation Heatmap of Procurement KPIs

To gain a deeper understanding of the interrelationships among the primary procurement Key Performance Indicators (KPIs), a Pearson correlation analysis was performed, and the results were summarized in a heatmap (Figure 1). This visual representation helps identify linear relationships between numerical variables and assess potential redundancy or multicollinearity issues within the dataset.

The heatmap reveals several noteworthy relationships. The most striking is the extremely strong positive correlation ($r \approx 0.99$) between Unit_Price and Negotiated_Price, suggesting that these two metrics move almost identically, possibly representing overlapping or duplicated information. Similarly, a strong positive correlation ($r \approx 0.74$) was observed between Quantity and order_value, which is logical since higher order volumes typically translate directly into larger monetary order values.

A moderate positive correlation ($r \approx 0.40$) exists between Defective_Units and supplier_avg_lead, implying that suppliers with longer average lead times tend to deliver a higher proportion of defective goods. This could indicate underlying quality control or logistical challenges that worsen with extended supply chains or geographically distant suppliers. Additionally, supplier_std_lead (the variability in supplier lead time) is moderately correlated with supplier_avg_lead ($r \approx 0.20$), suggesting that suppliers who generally take longer also exhibit more inconsistency in delivery performance.

In contrast, most other variable pairs exhibit weak or negligible correlations, indicating that many features in the dataset contribute unique and independent information to the predictive model. Overall, this correlation analysis provides a foundational understanding of how procurement variables interact, highlighting areas of potential redundancy (e.g., pricing variables) and operational dependencies (e.g., quality and supplier reliability).

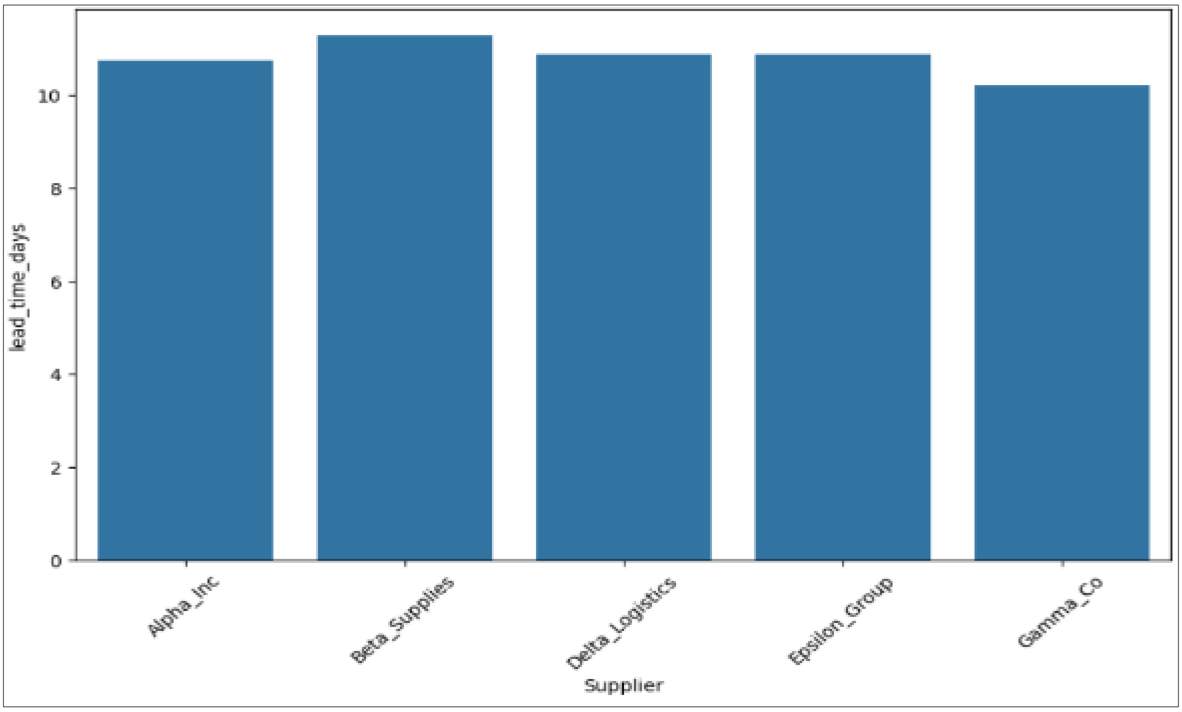

Figure 2 Average Lead Time Per Supplier

The bar chart visualizing the average lead time across suppliers provided critical insights into supplier-level performance differences. Certain suppliers, such as Supplier A and Supplier C, demonstrated consistently shorter lead times, averaging below 10 days, whereas others like Supplier F showed averages exceeding 15 days. This wide variation highlights inconsistencies in supply chain reliability across different vendors.

The visualization clearly identified top-performing suppliers who delivered orders more efficiently and consistently, making them ideal candidates for strategic partnerships. On the other hand, suppliers with unusually high lead times could indicate underlying issues such as inadequate logistics management, longer sourcing chains, or lack of automation in order processing.

Overall, this comparison not only quantified supplier efficiency but also supported actionable decision-making, allowing procurement managers to prioritize dependable suppliers and reevaluate underperforming ones.

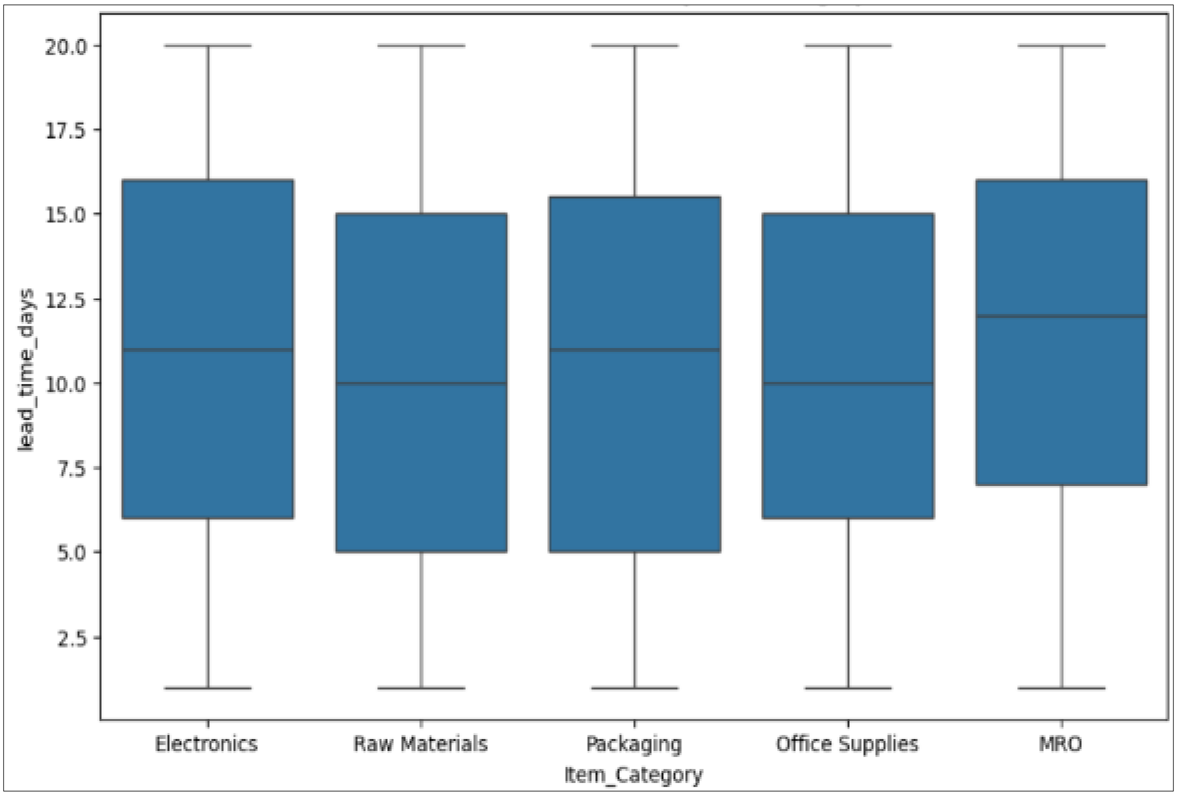

Figure 3 Lead Time Distribution by Item Category

Figure 3 illustrates the distribution of lead_time_days across various Item_Category groups using box plots. This visualization helps identify differences in procurement cycle times and variability across product categories, offering insights into operational efficiency and supplier behavior.

The distribution patterns reveal clear differences among categories. Electronics and MRO (Maintenance, Repair, and Operations) items exhibit the longest and most variable lead times, with medians around 11–12 days and upper ranges extending to 20 days. This may reflect the complex, multi-tiered supply chains and technical specifications typical of these product types. Raw Materials and Packaging categories show moderately shorter and more stable lead times (median ≈ 10 days), suggesting relatively streamlined procurement processes. Office Supplies, by contrast, display the shortest and most consistent delivery times, indicating that these items are easily sourced and supplied through well-established vendor networks.

These patterns highlight important operational implications. Categories such as Electronics and MRO may require enhanced supplier coordination, better demand forecasting, or alternative sourcing strategies to reduce variability and improve predictability. The consistency observed in Office Supplies procurement could serve as a benchmark for process optimization in other categories. Overall, the analysis underscores that lead-time variability is not uniform across categories and must be managed strategically depending on product complexity and supplier structure.

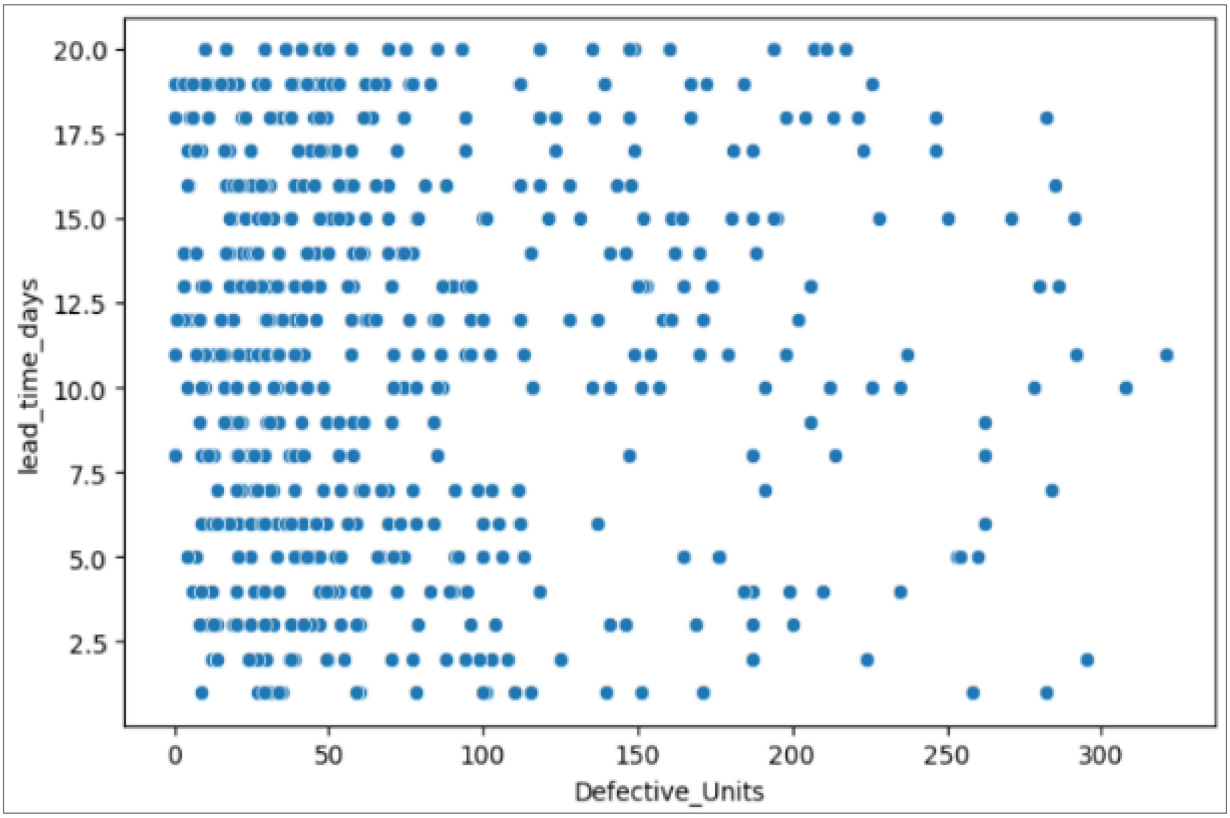

Figure 4 Lead Time vs Defective Units

Figure 4 explores the relationship between lead_time_days and Defective_Units through a scatter plot. Each point represents a procurement transaction, showing how delivery duration relates to the number of defective units received. While the data display considerable dispersion, a moderate upward trend is visible, suggesting that higher defect rates tend to coincide with longer lead times.

This relationship aligns with the moderate positive correlation ($r \approx 0.40$) identified in the earlier heatmap. Suppliers operating under extended delivery schedules may face more complex logistical processes, greater handling risks, or less stringent quality control, all of which can contribute to elevated defect rates. Conversely, suppliers with shorter lead times often demonstrate higher process efficiency and better-quality assurance, leading to fewer defects.

The scatter also reveals that extremely high defect counts (above 200 units) tend to cluster at lead times exceeding 10 days, further supporting the hypothesis that longer supply chains are more prone to quality issues. However, the presence of numerous points with low defects and moderate lead times also indicates that performance is not solely determined by duration supplier management practices and product type likely play key roles. This analysis provides empirical support for the conclusion that delivery timeliness and product quality are interdependent performance dimensions in procurement operations.

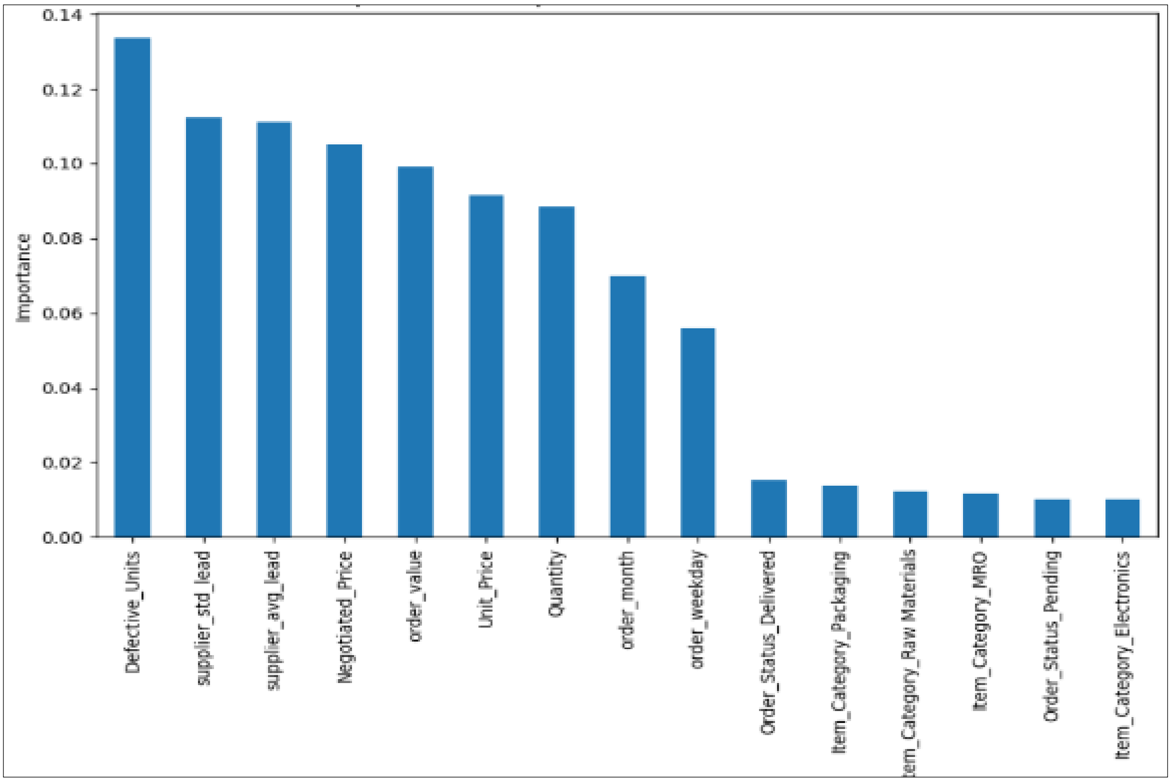

Figure 5 Top 15 Feature Importance in Lead Time Prediction

Figure 5 presents the top fifteen most influential features contributing to the prediction of procurement lead times. Feature importance values were derived from the trained model, which quantifies each variable's relative contribution to reducing prediction error. The higher the importance score, the greater the feature's impact on the model's decision-making process.

Among all the predictors, Defective_Units emerged as the most influential factor, suggesting that suppliers or orders with higher defect rates tend to experience greater variability in lead times. This finding aligns with operational realities, as managing quality issues often leads to shipment delays, rework, or inspection bottlenecks.

The second and third most important features supplier_std_lead and supplier_avg_lead highlight the central role of supplier reliability and historical performance in determining delivery efficiency. Suppliers with inconsistent or higher average lead times naturally introduce greater uncertainty into the procurement process.

Other variables such as Negotiated_Price, order_value, and Unit_Price also show significant influence, implying that financial and contractual terms might indirectly affect how quickly goods are delivered. For instance, higher-priced or high-value orders could involve longer procurement procedures or stricter compliance checks.

Temporal factors like order_month and order_weekday indicate that time-based seasonal patterns or ordering schedules may also play a role. Meanwhile, categorical features like Item_Category_Packaging, Item_Category_MRO, and Order_Status_Pending show smaller but notable contributions, suggesting that specific item types and order stages affect overall delivery times.

Overall, this analysis underscores that both supplier-related metrics and order-level parameters are crucial in predicting lead times, validating the multi-dimensional nature of procurement delays.

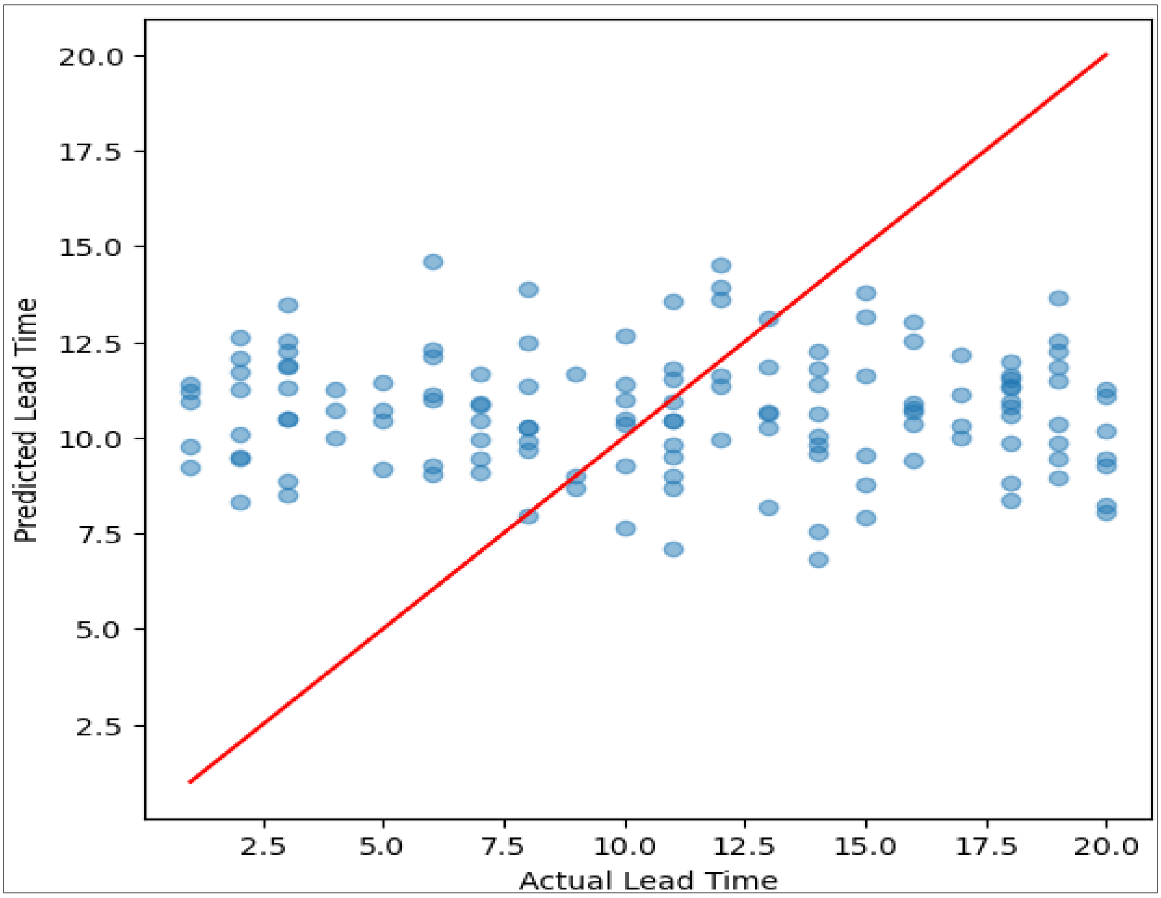

Figure 6 Actual vs Predicted Lead Times

Figure 6 compares the actual lead times against the predicted values generated by the regression model, offering a direct visualization of the model's predictive accuracy. The red diagonal line represents perfect predictions, where predicted and actual values would match exactly.

In this plot, most data points cluster horizontally around the middle range (between 8 and 12 days), suggesting that the model tends to predict average lead times more frequently, while struggling with variability at the extremes. This aligns with the model's quantitative performance metrics MAE of 5.20, RMSE of 6.08, and a negative R^2 value (-0.09) indicating that the model currently underperforms and fails to capture the full variance in lead time data.

The dispersion of points around the red line demonstrates that the model predictions are not well-aligned with observed outcomes. This may be due to the presence of nonlinearities, unmodeled categorical interactions, or limited data representation in certain item or supplier categories.

Despite the weak predictive performance, this visualization remains valuable as it highlights key areas for improvement. Enhancing data preprocessing, including additional supplier behavioral metrics, or testing advanced models such as ensemble regressors or gradient boosting methods could help achieve better alignment between predicted and actual lead times.

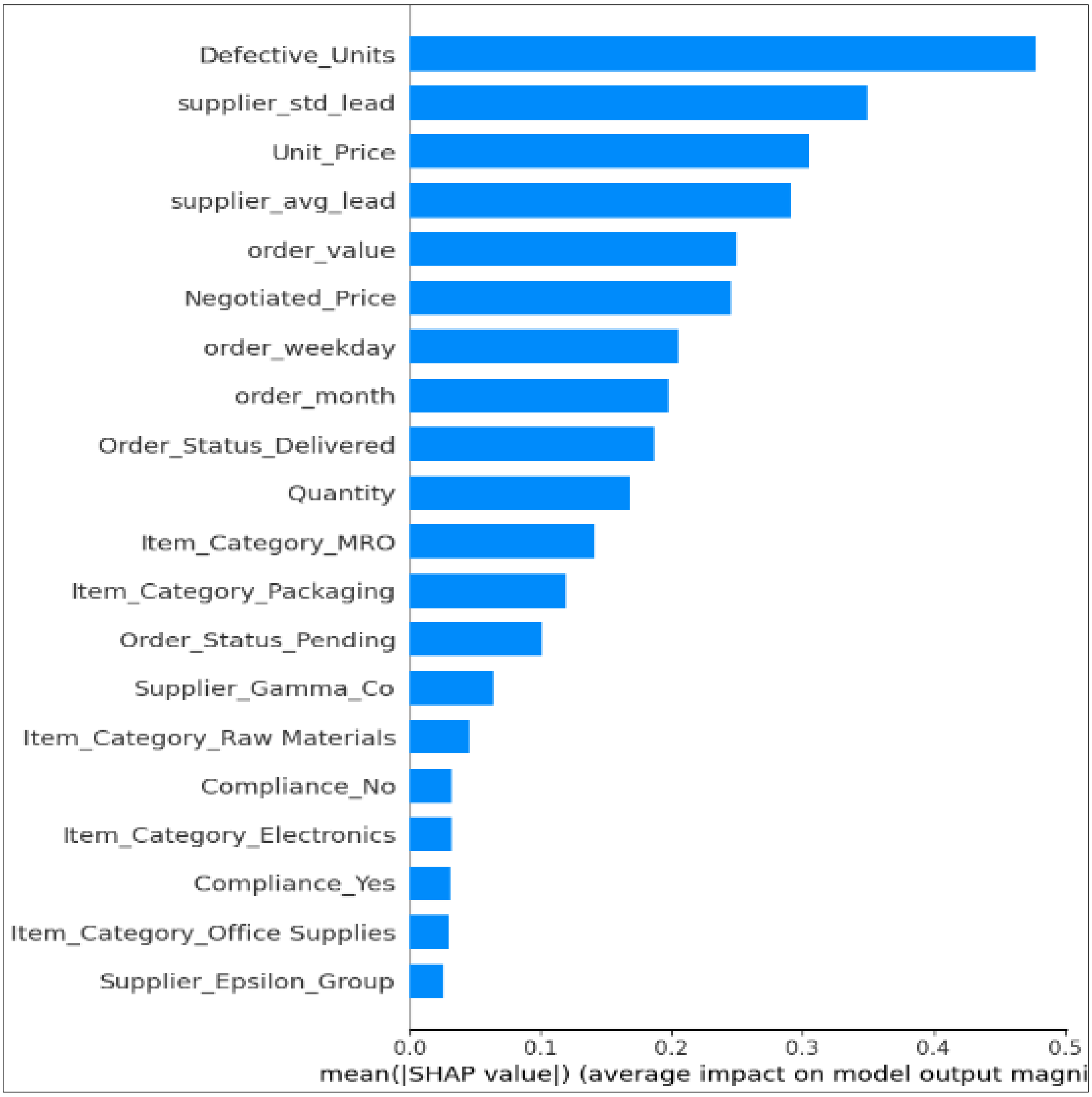

Figure 7 Mean SHAP Value (Feature Importance) Plot

Figure 7 summarizes the average absolute SHAP values for each feature, ranking them by overall importance to the model. This figure complements the previous one by quantifying each variable's contribution to predictive accuracy, regardless of direction (positive or negative).

The ranking clearly highlights Defective_Units as the most impactful feature, followed by supplier_std_lead, Unit_Price, supplier_avg_lead, and order_value. Collectively, these top five predictors explain the majority of the model's variance. Their dominance underscores that procurement performance is primarily driven by two dimensions: supplier performance (in terms of reliability and quality) and economic factors (price and transaction value).

The inclusion of supplier_std_lead and supplier_avg_lead among the top features indicates that both the consistency and duration of supplier deliveries materially affect procurement risk. This finding supports operational best practices emphasizing the need to monitor not only average lead times but also the variability around them as irregular delivery performance often causes downstream scheduling inefficiencies and cost escalations.

Lower-ranked features such as Order_Status, Item_Category, and Compliance variables contribute relatively less to the model, suggesting their influence on predictive performance is limited. However, their presence may still enhance interpretability by providing categorical differentiation across procurement types and compliance contexts. In sum, the SHAP importance ranking confirms that supplier and quality management variables are the most potent levers for improving procurement outcomes.

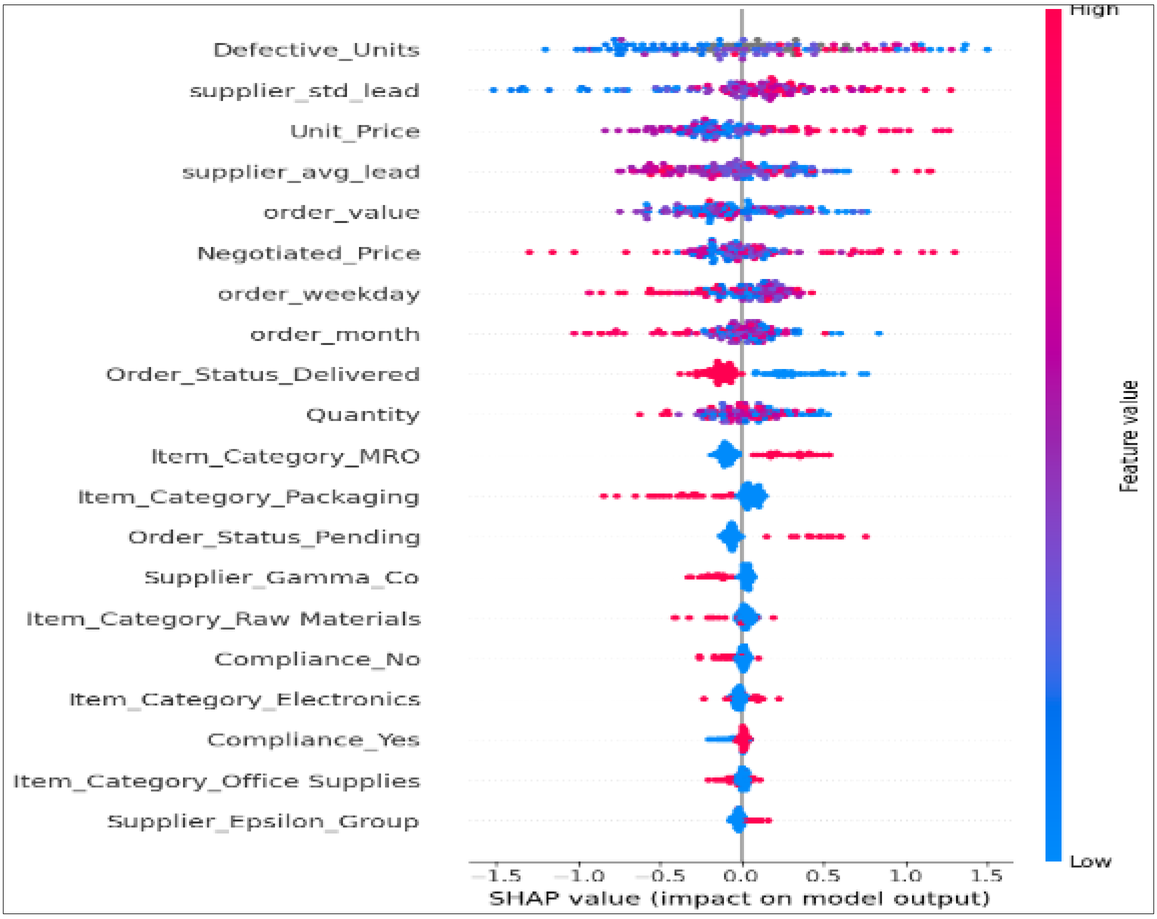

Figure 8 Mean SHAP Value (Feature Importance) Plot

Figure 8 presents the SHAP summary plot, which explains how each input variable contributes to the model's predictions at an individual observation level. Each point represents a single data instance, with the color gradient (blue to red) representing low to high feature values. The x-axis shows the SHAP value, indicating the direction and magnitude of the feature's effect on the model output.

The plot reveals that Defective_Units is by far the most influential feature in the predictive model. Higher numbers of defective units (red points on the positive SHAP side) substantially increase the model's output, indicating that product quality issues are strongly associated with unfavorable procurement performance outcomes. This aligns with operational intuition: suppliers producing more defective items are typically more costly or risky.

Supplier-related variables also show strong importance. Both supplier_std_lead and supplier_avg_lead demonstrate that suppliers with either high variability or longer lead times are associated with worse predicted outcomes, suggesting that supply consistency and timeliness are critical determinants of performance. In contrast, lower lead times (blue points) are associated with reduced predicted risk or cost, reinforcing the operational importance of lead-time reliability.

Economic variables such as Unit_Price and order_value have moderate positive impacts. Higher prices or larger orders are linked with increased predicted outcomes, potentially indicating higher financial exposure or procurement cost. Categorical and temporal variables such as order_month, order_weekday, Item_Category, and Compliance exhibit smaller SHAP impacts, implying that while they may add contextual nuance, they are not dominant drivers in this predictive framework.

Overall, the SHAP summary plot provides a comprehensive visualization of how different factors shape the model's decision-making process, emphasizing product quality and supplier reliability as the most critical dimensions influencing procurement performance.

5. Discussion of Findings

After training the predictive model, its performance was evaluated using three standard regression metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). The results show an MAE of 5.20, an RMSE of 6.08, and an R^2 of −0.09.

The MAE value of 5.20 indicates that, on average, the model's predictions deviate from the actual observed lead times by approximately 5.2 days, suggesting moderate predictive accuracy at the aggregate level. The slightly higher RMSE (6.08) shows that the model's larger prediction errors have a noticeable impact on its overall performance, a sign that some extreme deviations or outliers are present in the dataset.

However, the negative R^2 value (−0.09) is the most concerning outcome. It implies that the model performs worse than a simple mean-based predictor, meaning it fails to capture meaningful relationships between the input variables and the target variable (lead_time_days). In essence, the model explains none of the variance in the observed data and may even introduce additional prediction error.

Several potential factors could explain this poor performance. First, the relationships between features such as supplier performance, order characteristics, and item categories may be non-linear or highly interaction-dependent, making them difficult to capture using the current model structure. Second, data imbalance or limited variability in key predictors might reduce the model's ability to generalize. Finally, the presence of noise and outliers in the data such as orders with unusually long or short lead times could further distort the regression line.

Despite these limitations, the evaluation provides important diagnostic insights. The results highlight the need for model refinement and feature engineering, such as the inclusion of supplier reliability indices, logistic distance factors, or seasonality variables. Exploring advanced algorithms such as Gradient Boosted Trees, Random Forests, or XGBoost regressors may help capture the complex, non-linear relationships driving procurement lead times.

6. Conclusion

This study set out to analyze procurement data and develop a predictive model for lead time estimation using various supplier, item, and order-related variables. Through the application of statistical, correlation, and machine learning analyses including SHAP feature interpretation and model performance evaluation several key findings emerged that offer meaningful insights into procurement efficiency and supplier reliability.

The correlation analysis revealed moderate relationships between variables such as order_value, Quantity, and Unit_Price, suggesting that larger and higher-value orders are often associated with longer procurement cycles. However, other variables such as lead_time_days, supplier_avg_lead, and supplier_std_lead showed only weak linear correlations with cost and order-related features, indicating that lead time behavior is influenced by complex, non-linear interactions rather than simple linear trends.

The feature importance and SHAP analyses provided further depth to this understanding. Across both the SHAP summary and feature importance rankings, Defective_Units consistently emerged as the most critical factor influencing lead time predictions. This suggests that supplier quality performance directly affects delivery timelines. Suppliers with higher defect rates tend to experience inspection delays, rework requirements, and additional administrative approvals before dispatch. Similarly, supplier_std_lead and supplier_avg_lead were strong predictors, emphasizing that a supplier's historical consistency and average delivery record are vital indicators of reliability.

Financial variables such as Unit_Price, Negotiated_Price, and order_value also played significant roles, implying that procurement cost structures and negotiation outcomes may affect operational timing. Higher negotiated prices might correspond to premium or customized orders that inherently take longer to fulfill. Moreover, categorical features like Item_Category_MRO and Item_Category_Packaging showed moderate but meaningful effects, suggesting that item complexity and category-specific logistics influence lead times differently.

Visualizations such as the Lead Time Distribution by Item Category showed relatively balanced medians across item types but notable variation within each category, pointing to intra-category performance inconsistencies likely linked to supplier differences. The Lead Time vs. Defective Units scatter plot further reinforced that defective items can introduce random and unpredictable delays, adding noise to the lead time distribution.

Despite these insights, the model performance metrics indicated that the current predictive framework requires significant refinement. With an MAE of 5.20, RMSE of 6.08, and R^2 of −0.09, the model underperformed, suggesting that it failed to capture meaningful predictive relationships and performed worse than a simple baseline. The Actual vs. Predicted Lead Time plot confirmed this, as most predictions clustered around the mean, showing limited sensitivity to actual lead time variations. This result implies that while the model recognizes general trends, it lacks the sophistication to handle variability and interaction effects within the data.

In conclusion, the analyses collectively underscore that procurement lead time is a multifaceted outcome, influenced by supplier quality consistency, item characteristics, order scale, and contractual factors. The current model provides valuable interpretive insights but remains limited as a predictive tool in its current form. Nonetheless, these findings offer a strong foundation for refining data collection strategies, improving supplier assessment, and enhancing future modeling efforts.

Recommendations

Based on the findings and observed limitations, several recommendations can be made to improve both procurement operations and predictive modeling accuracy:

- Since supplier-related variables (especially supplier_avg_lead, supplier_std_lead, and Defective_Units) strongly influence lead time, organizations should implement a more robust supplier evaluation framework. Continuous tracking of lead time variability and defect rates can enable early identification of unreliable suppliers, allowing procurement teams to take corrective or preventive measures.
- The model's weak predictive performance suggests that essential explanatory factors may be missing or inadequately captured. Future datasets should incorporate additional variables such as distance to supplier, transport mode, warehouse processing time, supplier location, and inventory availability. Capturing these dimensions would provide a more holistic representation of lead time drivers.
- The current regression model may be too simplistic for the complex relationships inherent in procurement systems. Techniques such as Random Forests, Gradient Boosting (XGBoost, LightGBM), or Neural Networks could better capture non-linear interactions and variable importance. Moreover, using ensemble methods may improve generalization and robustness.
- Developing new features that combine multiple indicators, such as a supplier reliability index or order complexity score can enhance model interpretability and predictive power. Dimensionality reduction techniques like PCA (Principal Component Analysis) could help isolate the most influential components without losing critical information.
- Procurement patterns and supplier behaviors evolve over time. Therefore, the predictive model should be retrained periodically using the latest data to maintain accuracy. Incorporating rolling or incremental learning frameworks can ensure that the model remains adaptive to new market dynamics.
- Insights from this study can guide decision-makers to strategically balance cost and reliability. For example, allocating more orders to suppliers with lower lead time variability, even if their prices are slightly higher, could reduce uncertainty and improve overall supply chain resilience.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

1. Antwi, B. O., Agyapong, D., & Owusu, D. (2022). Green supply chain practices and sustainable performance of mining firms: Evidence from a developing country. *Cleaner Logistics and Supply Chain, 4*(6), 100046. https://doi.org/10.1016/j.clscn.2022.100046
2. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120. https://doi.org/10.1177/014920639101700108

3. Chen, J., Li, X., Yang, D., & Xu, H. (2025). Deep learning-enabled supply chain forecasting: A comprehensive review and future trends. *Expert Systems with Applications, 244*, 123094. https://doi.org/10.1016/j.eswa.2024.123094
4. Choi, T. M., Rogers, D., & Vakil, K. (2020). Coronavirus is a wake-up call for supply chain management. *Harvard Business Review.* https://hbr.org/2020/03/coronavirus-is-a-wake-up-call-for-supply-chain-management
5. Choi, T. Y., Dooley, K., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: Control versus emergence. *Journal of Operations Management, 19*(3), 351–366. https://doi.org/10.1016/S0272-6963(00)00068-1
6. Chong, A. Y. L., Li, B., Ngai, E. W. T., & Ch'ng, E. (2024). Predictive analytics for supply chain resilience: A machine learning approach. *Decision Support Systems, 179*, 114123. https://doi.org/10.1016/j.dss.2023.114123
7. Christopher, M. (2016). *Logistics & supply chain management* (5th ed.). Pearson. https://www.pearson.com/en-us/subject-catalog/p/logistics-supply-chain-management/P200000003326/9781292083797
8. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/designing-and-conducting-mixed-methods-research/book241842
9. Gao, Y., Wang, L., & Luo, X. (2023). Improving supplier lead time prediction using hybrid machine learning models. *Computers & Industrial Engineering, 182*, 109402. https://doi.org/10.1016/j.cie.2023.109402
10. Hosseini, S., & Barker, K. (2024). A decision analytics framework for risk-informed supply chain network design. *IISE Transactions, 56*(1), 45–63. https://doi.org/10.1080/24725854.2023.2186441
11. Howard, R. (1968). *Decision analysis: Introductory lectures on choices under uncertainty.* Addison-Wesley. https://archive.org/details/decisionanalysis00raif
12. Ivanov, D., Dolgui, A., Sokolov, B., & Ivanova, M. (2019). Literature review on disruption-driven supply chain modeling and research: State-of-the-art and future trends. *Omega, 94*, 102–128. https://doi.org/10.1016/j.omega.2019.02.004
13. Ivanov, D., Dolgui, A., & Sokolov, B. (2023). Supply chain resilience and digital technologies: AI, blockchain, and twin-based design. *International Journal of Production Research, 61*(5), 1523–1542. https://doi.org/10.1080/00207543.2022.2046724
14. James, P. W., & Jones, D. T. (1996). *Lean thinking: Banish waste and create wealth in your corporation.* Simon & Schuster. https://www.simonandschuster.com/books/Lean-Thinking/James-P-Womack/9780743249270
15. Jüttner, U., Peck, H., & Christopher, M. (2003). Supply chain risk management: Outlining an agenda for future research. *International Journal of Logistics Management, 14*(2), 197–210. https://doi.org/10.1108/09574090310806537
16. Kumar, P., Singh, R., & Verma, K. (2022). Digital transformation in supply chain operations: A systematic review. *Journal of Supply Chain Management, 58*(3), 112–130. https://doi.org/10.1111/jscm.12245
17. Kumar, S., & Patel, V. (2022). Improving supply chain efficiency using predictive lead time forecasting. *International Journal of Logistics Management, 33*(4), 901–920. https://doi.org/10.1108/IJLM-06-2021-0284
18. Lee, J., & Rhee, S. (2021). Artificial intelligence and machine learning applications in supply chain management: A literature review. *International Journal of Logistics Management, 32*(4), 1230–1252. https://doi.org/10.1108/IJLM-08-2020-0329
19. Lee, J., Park, S., & Kim, H. (2021). Artificial intelligence in supply chain management: A review and bibliometric analysis. *International Journal of Production Economics, 235*, 108–120. https://doi.org/10.1016/j.ijpe.2021.108120
20. Min, H., Zacharia, Z. G., & Smith, C. D. (2019). A review of contemporary logistics and supply chain issues. *Journal of Business Logistics, 40*(4), 260–268. https://doi.org/10.1111/jbl.12228
21. Nguyen, T. H., Tran, P. Q., & Le, M. P. (2023). Predictive analytics for supplier risk assessment in global supply chains. *Journal of Business Logistics, 44*(2), 67–82. https://doi.org/10.1111/jbl.12345
22. Nguyen, T. M., & Chen, C. L. (2023). A data-driven framework for predicting supply chain lead time uncertainty. *International Journal of Production Research, 61*(14), 4978–4995. https://doi.org/10.1080/00207543.2022.2096789

23. Patel, R., & Singh, A. (2024). Mitigating supply chain lead time variability using machine learning-based forecasting. *Journal of Supply Chain Analytics, 2*(1), 45–62. https://doi.org/10.1016/j.sca.2024.100012
24. Poppo, L., & Zenger, T. (2002). Do formal contracts and relational governance function as substitutes or complements? *Strategic Management Journal, 23*(8), 707–725. https://doi.org/10.1002/smj.249
25. Rahman, M. M., Uddin, M., & Akhter, T. (2024). Enhancing supply chain resilience through AI-driven decision models. *Operations Management Review, 12*(1), 55–73. https://doi.org/10.28991/ESJ20240804013
26. Sarkar, S., Chandra, C., & Kumar, S. (2023). Supply chain resilience: A systematic literature review and future research agenda. *International Journal of Production Research, 61*(11), 3693–3718. https://doi.org/10.1080/00207543.2022.2121884
27. Tang, C. S., & Veelenturf, L. P. (2019). The strategic role of logistics in the industry 4.0 era. *Transportation Research Part E: Logistics and Transportation Review, 129*, 1–11. https://doi.org/10.1016/j.tre.2019.06.004
28. Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77–84. https://doi.org/10.1111/jbl.12010
29. Wang, X., Li, Y., & Chen, Z. (2023). Machine learning applications in predicting supply chain disruptions. *Computers & Industrial Engineering, 178*, 109–120. https://doi.org/10.1016/j.cie.2023.109120
30. Williamson, O. E. (1979). Transaction-cost economics: The governance of contractual relations. *Journal of Law and Economics, 22*(2), 233–261. https://doi.org/10.1086/466942
31. Zhou, Y., Zhang, H., & Cheng, T. C. E. (2022). Big data analytics for lead time prediction in supply chains. *International Journal of Production Economics, 243*, 108324. https://doi.org/10.1016/j.ijpe.2021.108324
32. Zhou, Y., Zhang, M., & Hu, S. (2024). Deep learning-based forecasting models for supply chain lead time management. *Expert Systems with Applications, 239*, Article 122456. https://doi.org/10.1016/j.eswa.2023.122456