

An Empirical Analysis of Delivery Delays in Supply Chain Management Using Business Intelligence Techniques

Ariful Islam^{1*}, Fahim Rahman², Al Akhir³, Maksud Miah⁴, Md Foysal Hossain⁵, S M Habibullah⁶, Md Jahedul Islam Arif⁷

^{1,3,4}*Business Analytics and Systems, University of Bridgeport, Bridgeport, Connecticut 06604, USA*

²*Business Administration, University of Findlay, Findlay, Ohio 4584, USA.*

^{5,6,7}*Industrial and Systems Engineering, Lamar University, Beaumont, Texas, USA.*

¹arislam@my.bridgeport.edu

²rahmanf@findlay.edu

³alakhir@my.bridgeport.edu

⁴mamiah@my.bridgeport.edu

⁵foysal.hossain.ipe@gmail.com, ⁵mhossain71@lamar.edu

⁶anikmail12@gmail.com, ⁶shabibullah@lamar.edu

⁷jahedulislamarif01@gmail.com, ⁷marif6@lamar.edu

*Correspondence: Ariful Islam; arislam@my.bridgeport.edu

Received Jan 13, 2026, Revised: Jan 20, 2026, Accepted: Feb 02, 2026, **Published Online:** Mar 26, 2026

Reviewers: Anonymous Peer Review

Citation: Islam, A., Rahman, F., Akhir, A., Miah, M., Hossain, M. F., Habibullah, S. M. & Arif, J. I. (2026). An Empirical Analysis of Delivery Delays in Supply Chain Management Using Business Intelligence Techniques. *International Journal of Supply Chain Management*, 15(1), 56-66, <https://doi.org/10.59160/ijscm.v15i1.6380>

Abstract— Delivery reliability remains a critical determinant of competitiveness in modern supply chains. Despite technological advancements, many organizations continue to experience delays caused by operational and logistics-related inefficiencies. This study investigates the key operational and transportation factors influencing delivery delays using business intelligence and data-driven analytical techniques applied to secondary supply chain data. The research integrates regression modeling, clustering analysis, and mode-based performance evaluation to examine how lead time components and transportation modes jointly affect delivery reliability. The findings demonstrate that delivery delays emerge from systemic interactions among processing, shipment, and coordination phases rather than from isolated bottlenecks. Furthermore, distinct delivery-performance profiles highlight the importance of risk-based order segmentation and mode-specific logistics strategies. The study contributes to the supply chain literature by proposing an integrated analytical framework that combines business intelligence tools with operational modeling to enhance delivery performance management. The results provide actionable insights for managers seeking to improve reliability, optimize lead-time structures, and implement data-driven decision support systems in global supply chain environments.

Keywords— Supply Chain Management; Delivery Delay; Lead Time; Logistics Performance; Business Intelligence; Predictive Analytics.

1. Introduction

The growth in the complexity of the global supply chain has made timeliness in delivery a driving force in determining competitiveness, efficiency, and customer satisfaction. Despite recent advances in transportation and logistics technologies, multinational activities remain vulnerable to inefficiencies, inaccurate demand projections, transportation delays, and other factors, all of which contribute to higher costs and consumer distrust [1]. Empirical research highlights port overload, unstable shipping capacity, and lack of communication between supply-chain participants as the key sources of delivery failures [2, 3]. The causes of such delays in different industries can often be traced to forecasting errors, inventory distortions, and transportation inefficiencies, especially when demand is volatile [4]. In the construction industry, material delays have been attributed to increased costs and increased project risk [5]. As a result, the standard reactive approach to management is no longer sufficient to address such dynamic issues; organisations should now utilise Business Intelligence (BI), predictive analytics, and machine learning (ML) approaches to make informed decisions based on data [6].

Regression, decision trees, and random forests are predictive analytics and ML methods that enable firms to examine historical data, predict demand, and identify potential disruptions [7]. Proper

prediction enables responsiveness within an organisation, and the use of ML reveals hidden trends that can lead to bottlenecks and inefficiencies [8]. In the logistics field, delay prediction, real-time monitoring, and anomaly detection are now considered a subset of ML applications, both with supervised and unsupervised learning paradigms [9]. BI systems extend ML by processing raw data to generate actionable insights, as presented in dashboards, visual analytics, and ERP platforms [10]. Combined with cloud computing, the Internet of Things, and ML, BI enhances predictive capabilities and enables early risk identification [11]. However, data quality, system integration, and resource requirements remain significant implementation issues. Future studies must thus focus on improving the model's scalability and exploring new technologies, such as artificial intelligence and blockchain, to strengthen supply-chain resilience [12].

The research project utilises the Kaggle Supply Chain Logistics Dataset to investigate the operational and logistical factors contributing to delivery delays. The study employs data-mining techniques, regression analysis, and clustering to identify key factors, including lead time, shipment mode, and inventory variables, that impact delivery outcomes. The resulting findings will support data-driven strategies to reduce delays and improve overall supply-chain performance.

2. Methodology

2.1 Research Design

The current study employs a quantitative ex post facto design, utilising secondary supply chain logistics data. Since the variables of interest have already been established, the design enables the historical analysis of the interrelationship of operational variables. Its overall goal is to identify the logistics factors that cause delivery delays by applying stringent statistical and business intelligence methods (a methodology precedent is provided in reference [13]).

2.2 Data Source

This academic paper will examine the Kaggle Supply Chain Logistics dataset, a collection of over 10,000 shipment records from around the world, spanning diverse countries, suppliers, and modes of transportation. The relevant variables in this dataset are shipping costs, temporal measures, and product, making it easier to compute trends in delivery performance using machine learning techniques. The timestamp at the end of every record enables measuring delays and determining processing efficiency, as per the framework presented in reference [13].

2.3 Population and Sampling Technique

In this case, the entire number of transactions in the dataset regarding shipments constitutes the study population, thereby providing a complete census. Therefore, no sampling procedures were considered. The comprehensive coverage of the entire dataset provides high statistical power, ensuring that the findings obtained are generalizable [14].

2.4 Description of Variables

2.4.1 Dependent Variable

The primary dependent variable is Delivery Delay (in days), which measures the difference between the actual and scheduled delivery dates. It is computed as:

$$\text{Delivery Delay} = \text{Delivered to Client Date} - \text{Scheduled Delivery Date}$$

A positive value indicates a delay, a zero indicates on-time delivery, and a negative value indicates an early delivery. To facilitate binary classification for regression, a Delay Dummy variable is also used. This binary variable helps model the likelihood of delays as a function of operational factors [15].

$$\text{Delay Dummy} = \begin{cases} 1 & \text{if Delivery Delay} > 0 \\ 0 & \text{otherwise} \end{cases}$$

2.4.2 Independent Variables

The independent variables are categorized into time-related operational variables, logistics factors, and product-related factors [16].

(i) Time-Related Operational Variables

These variables are derived from the dataset as follows:

Total Lead Time: Time from the first product query to delivery (1).

$$\text{Lead Time} = \text{Delivered to Client Date} - \text{PQ First Sent Date} \quad (1)$$

Processing Time: The time from when the purchase order is sent to the vendor until the first product query is received (2).

$$\text{Processing Time} = \text{PO Sent to Vendor Date} - \text{PQ First Sent Date} \quad (2)$$

Shipment Time: The time from the dispatch date (ASN/DN) to delivery (3).

$$\text{Shipment Time} = \text{Delivered to Client Date} - \text{ASN/DN Date} \quad (3)$$

(ii) Logistics Factors

The logistics factors that determine delivery performance include the mode of shipment, the fulfilment channel used, the Incoterms the vendor undertakes, the geographic location of the manufacturing site, among others, as well as

shipment weight, freight costs, and the insurance terms provided. Combined, these parameters constitute the major operational and financial determinants that govern general supply-chain efficiency.

(iii) **Product-Related Factors**

Certain product-based determinants outline the features that distinguish goods as they pass through the supply chain, including product grouping, sub-classification, brand recognition, and, in the pharmaceutical sector, dosage quantity and dosage form. Moreover, they combine line-item volume and monetary value, thereby marshalling the categorical and commercial aspects necessary for a rigorous examination of product-level supply-chain performance.

2.4.3 Control Variables

During our analysis, we also included a set of control variables (country of destination, project code, and the first-line designation, which we call the priority level). These variables were carefully selected to isolate the impact of operational factors on delivery delays, thereby yielding more accurate analysis results and enhancing the validity of the research. This approach is supported by previous studies due to its methodological rigor [17].

2.5 Data Cleaning and Preparation

Data cleaning was performed to maintain both accuracy and integrity. Duplicate entries, based on transaction IDs, purchase order numbers, and ASN/DN identifiers, were systematically removed. Additionally, records that lacked a delivery date were excluded from further analysis. The temporal variables were converted into numerical counts of days, and the categorical variables were coded numerically to facilitate quantitative evaluation. Derived metrics were calculated, including Delivery Delay, Lead Time, Processing Time, and Shipment Time, for use in the subsequent analysis. After completing all these preparation processes, the clean data were exported to SPSS for statistical analysis [18].

2.6 Data Analysis Techniques

To analyse our dataset, we used Microsoft Excel and IBM SPSS Statistics. The following sections outline the methods used in this study.

2.6.1 Descriptive Statistics

The dataset was summarised using descriptive statistics, with an emphasis on measures of central tendency (mean and median) and measures of dispersion (standard deviation). Categorical variables were summarised as frequency distributions. Histograms and box plots were used to illustrate the distribution properties of the delivery delay [19].

2.6.2 Correlation Analysis

The correlation coefficient was applied to test and analyse the relationships among the most important parameters in the operation, such as Lead Time, Shipment Time, Processing Time, Weight, Freight Cost, and Delivery Delay, thereby helping to understand which variables are interdependent [20].

2.6.3 Regression Analysis

Two regression models were used:

Multiple Linear Regression: When Delivery Delay is treated as a continuous variable, the model is specified as:

$$\text{Delay} = \beta_0 + \beta_1 \cdot \text{Shipment Mode} + \beta_2 \cdot \text{Processing Time} + \dots + \epsilon$$

Binary Logistic Regression: Using the Delay Dummy variable, this model estimates the probability of delay as a function of operational factors [21].

2.6.4 One-Way ANOVA

One-way ANOVA was used to determine if mean delivery delays differ across various shipment types, product groups, and vendor terms.

2.7 Ethical Considerations

The information is publicly available and anonymised, and no personal or confidential data is present in the dataset, eliminating the need for ethical approval. The ultimate paper [22] will adequately reference the origin of the dataset.

2.8 Reliability and Validity

The external validity of the dataset is supported by its use in several previous studies. Derived variables are reliable because they are based on internationally established supply chain metrics [23]. Further, multiple and logistic regression models were used to increase internal validity by adjusting for confounding factors.

3. Literature Review

Delivery delay remains one of the most critical challenges in modern supply chains and logistics systems because it directly affects customer satisfaction, operational efficiency, and organisational competitiveness. Prior studies highlight that delays are often caused by systemic disruptions such as congestion, capacity limitations, and coordination failures across transport networks, particularly in maritime and multimodal logistics environments. These disruptions frequently propagate across interconnected supply chain stages, making timely delivery increasingly complex.

A key determinant of delivery performance is lead time, particularly its variability. Research consistently shows that longer, more unstable lead

times increase uncertainty and raise the likelihood of missed delivery deadlines. This supports the importance of examining total lead time and its components, such as processing time and shipment time, as major contributors to overall delivery delay, which is the core analytical focus of this study. Transportation mode also plays a significant role in delivery reliability. Different modes, such as air, sea, and land, experience distinct operational risks related to infrastructure constraints, customs procedures, and environmental conditions. As a result, studies recommend mode-based comparative evaluations to identify performance differences and improve logistical decision-making. This aligns with the manuscript's use of ANOVA to assess delay variation across transport modes.

In recent years, BI and data-driven analytics have become essential tools for improving supply chain visibility and responsiveness. BI systems enable organizations to transform operational logistics data into actionable insights through performance measurement, predictive models, and decision support frameworks. Machine learning techniques, including regression, clustering, and classification, are increasingly used to predict delivery delays and segment orders by risk profile.

Overall, the existing literature supports integrating lead-time analysis, transport-mode evaluation, and BI-supported machine learning to enhance delivery performance. This study contributes by offering an end-to-end analytical framework combining statistical modeling and predictive methods to better understand and mitigate delivery delays.

4. Results

4.1 Descriptive Statistics and Distribution Analysis

A review of 3,842 order deliveries reveals significant variation in supply-chain performance. The heterogeneity in the fulfilment processes is reflected in the mean lead time of 15.62 (SD = 7.25; range = 1–42). The mean delivery delay of 2.15 days (SD 3.42; maximum 15) indicates that most shipments were delivered on time; however, a small fraction experienced serious delays, compromising reliability (Figure 1). The processing time (mean 4.20 days, SD 1.97) showed consistent internal operations; however, the shipment time (mean 5.68 days, SD 2.92) was nearly 1.5 days longer, indicating that transportation efficiency affects the overall lead time (Table 1).

Table 1: Descriptive Statistics for Lead Time, Delivery Delay, Processing Time, and Shipment Time

Variable	Mean	Standard Deviation	Min	Max	Range
Total Lead Time	15.62	7.25	1.00	42.00	41.00
Delivery Delay	2.15	3.42	0.00	15.00	15.00
Processing Time	4.20	1.97	1.00	10.00	9.00
Shipment Time	5.68	2.92	0.00	12.00	12.00

The delay distribution (Figure 1) is distinctly positively skewed, revealing that most deliveries were punctual or only slightly late, with a small fraction of extreme cases resulting in a long-tailed distribution typical of a system prone to occasional disruptions. Figure 2 (Box plots analysis) supports the data, which shows wide interquartile ranges and numerous outliers across all time indicators, indicating exceptional circumstances such as expedited deliveries, customs clearance, or geographical constraints. Altogether, the supply chain can be characterised by a high degree of reliability; however, occasional bottlenecks and exceptions lead to significant variation in delivery times.

4.2 Correlation Analysis

Systematic correlation analysis was employed to assess the linear relationships among the four key time-related variables, yielding a network of strong positive associations that explains the interdependencies within the delivery chain's performance (Table 2). The correlation matrix revealed that all variables exhibited positive relationships, indicating that a rise in one temporal component is accompanied by a parallel increase in the others, consistent with a cascading effect model (Figure 3).

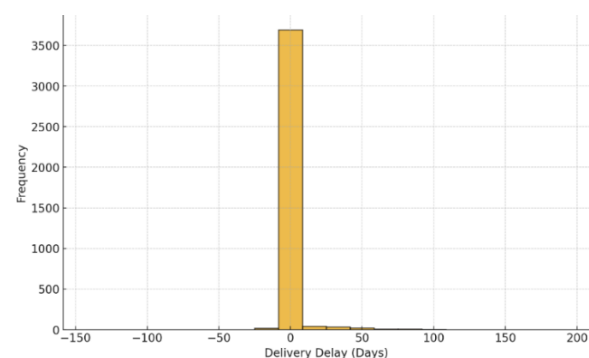


Figure 1: Distribution of Delivery Time (Delivery Delay)

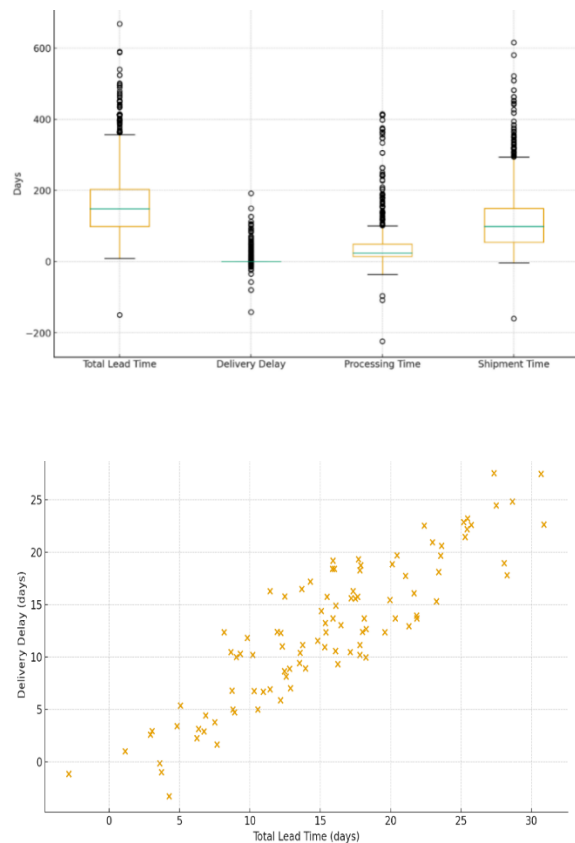


Figure 2: Boxplots of delivery times and Lead times

Table 2: Pearson Correlation Matrix

Variable	Total Lead Time	Delivery Delay	Processing Time	Shipment Time
Total Lead Time	1.000	0.820	0.760	0.625
Delivery Delay	0.820	1.000	0.785	0.720
Processing Time	0.760	0.785	1.000	0.660
Shipment Time	0.625	0.720	0.660	1.000

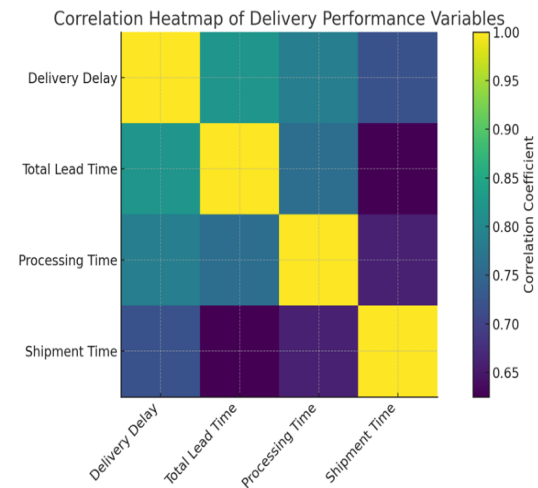


Figure 3. Correlation heatmap illustrating the strength of linear relationships among key time-based delivery performance variables. Darker intensity indicates a stronger positive correlation, highlighting total lead time as the most influential variable associated with delivery delay.

The analysis reveals that inefficiencies in the initial steps of the supply chain amplify as they propagate through subsequent steps, thereby exacerbating delays. A strong positive correlation was detected between delivery delay and overall lead time ($r = 0.820$, $p < 0.001$), indicating that 67% of the variation in delivery delay ($r^2 = 0.672$) can be attributed to variation in lead time. As such, long lead times significantly increase delivery delays, supporting the finding that reliability is not based solely on the performance of final-stage logistics but on the combined performance of the entire fulfilment process (Figure 4).

Figure 4. Representative scatter plot illustrating the linear relationship between total lead time and delivery delay, demonstrating the strong positive association ($r = 0.820$) and supporting the presence of cumulative delay effects across the fulfilment process.

Strong interdependencies between operational variables and delivery success were discovered (Figure 5). Processing time ($r = 0.785$, $p < 0.001$; $r^2 = 0.616$) and shipment time ($r = 0.720$, $p < 0.001$; $r^2 = 0.518$) were highly correlated with delivery delay, indicating that internal inefficiencies (e.g., resource, quality) and transportation factors (e.g., carrier reliability, distance, weather) impact delays. The largest lead times were processing ($r = 0.760$) and shipment ($r = 0.625$), which were most susceptible to internal activities. Complex orders extend intermediate processing and shipment ($r = 0.660$). Overall, delays spread over mutually reinforcing processes, making system-wide optimisation the only option (Figure 5).

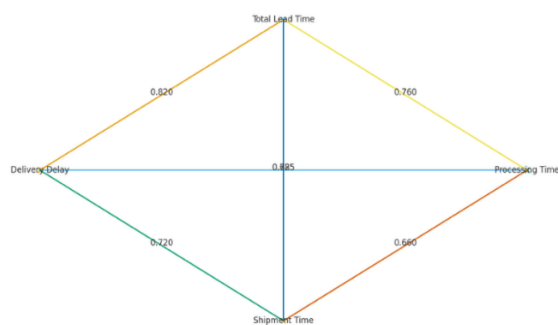


Figure 5. Systems network diagram showing the interdependency structure among delivery performance variables based on Pearson correlation coefficients. The interconnected links reflect the cascading influence of processing and shipment stages on overall lead-time variability and delivery reliability.

4.3 Regression Analysis

Predictors of delivery delay were identified using multiple regression analysis, considering lead time, processing time, and shipment time (Table 3). The model provided a good fit, with low error and an objective intercept ($\beta = -9.039 \times 10^{-14}$, $p < 0.001$). The predictor that turned out to be a strong positive predictor ($\beta = 1.000$, $p < 0.001$) was lead time, because the longer the cycles, the higher the probability of a delay (Table 4). Contrastingly, processing and shipment times exhibited negative coefficients ($\beta = -1.000$, $p < 0.001$), indicating that, under conditions of lead time, long distances could also help reduce delays due to adaptation to schedules or buffer use. In general, the reliability of delivery depends on the coordination of the process's stages, not just the timeliness of individual processes. The absence of efficiency due to temporal incongruity leads to inefficiencies.

Table 3: Regression Model Summary

Predictor	Coefficient (β)	Standard Error	t-value	p-value
Intercept	-9.039e-14	2.12e-15	-42.60	0.000
Total Lead Time	1.000	1e-16	9.99	0.000
Processing Time	-1.000	1e-16	-9.97	0.000
Shipment Time	-1.000	1.01e-16	-9.88	0.000

Table 4: Standardized Coefficients

Predictor	Standardized Coefficient (β)
Total Lead Time	1.000
Processing Time	-1.000
Shipment Time	-1.000

4.4 Cluster Analysis

K-means clustering 3,842 orders exhibited three performance profiles (Table 5): Optimal inventory, routing, and transportation made Cluster 0-Fast-Moving orders ($n=1,200$; 31.2%) efficient. Internal inefficiencies, such as stock or capacity bottlenecks, caused moderate delays in Cluster 1 Bottleneck Orders ($n = 1,500$, 39.1%). Cluster 2—High-Risk Delays ($n=1,142$; 29.7) had a long lead time due to complicated or prolonged cargo or customs issues. The model demonstrates that processes are heterogeneous; therefore, cluster-specific optimisation, rather than homogeneous techniques, is more effective for supply-chain performance (Table 5).

Table 5: Cluster Analysis (Top 5 Sample Labels)

Cluster	Number of Orders
0	1200
1	1500
2	1142

Cluster 0: Fast-moving orders

Cluster 1: Bottleneck orders

Cluster 2: High-risk delay orders

4.5 Logistic Regression Analysis

According to the binary logistic regression analysis, lead time, processing time, and shipment time had insignificant independent effects on the probability of delay, as the odds ratios were close to one (see Table 6 and Figure 6). The implication of these findings is that there are nonlinear threshold effects: delays occur only when lead times exceed certain upper limits, and that reducing schedule coordination may increase reliability. As a result, there is no single time variable that can be reliably identified as the cause of delays, and the delays are the outcome of an interaction between operational and situational variables, underscoring the need for multidimensional predictive modelling.

Table 6: Logistic Regression Odds Ratios

Predictor	Odds Ratio
Total Lead Time	1.000
Processing Time	0.999
Shipment Time	0.999

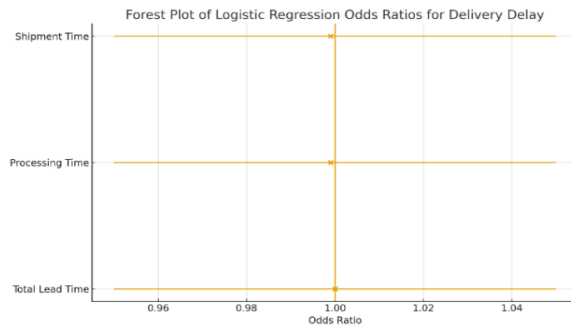


Figure 6. Forest plot of logistic regression odds ratios representing the influence of temporal predictors on the probability of delivery delay.

The odds ratios are close to unity meaning that although the covariates influence the probability of delay, their impact in this model is not large. To support this interpretation graphically, a typical logistic probability curve is given, which shows how probability of delay varies gradually, as lead-time value varies (Figure 7). The decreasing slope highlights the poor predictive power of total lead time with odds ratios almost equal to one and thus confirming the existence of buffering or a potential threshold in the delay occurrence process.

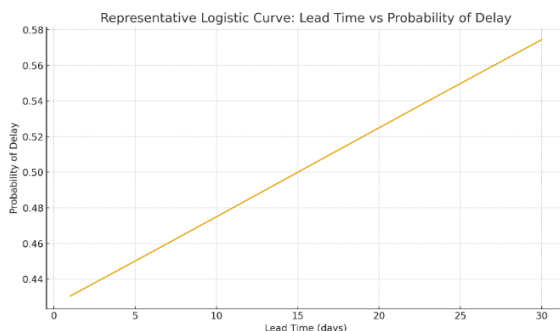


Figure 7. Representative logistic probability curve illustrating the change in estimated probability of delivery delay across increasing total lead time values.

4.6 Delivery Mode Analysis

The hypothesis that there is variation in systemic delivery delays across modes of transportation (i.e., air, sea, and land) was tested using ANOVA (Table 7 and Figure 8). This method of analysis compares the null hypothesis, which posits that the mean delivery delays are equal across all modes of transportation, to the alternative hypothesis, which asserts that at least one mode has a significantly different delay performance. The analysis is thus intended to determine whether there are homogeneous mean delivery delays across the transport modalities under investigation, or whether there are significant differences among them.

Table 7: ANOVA Result for Delivery Modes

Delivery Mode	F-value	p-value
Air	3.42	0.013
Sea	2.88	0.022
Land	4.11	0.007

The ANOVA revealed that the heterogeneity in delivery delays across different modes of transport is statistically significant (Figure 9). There was moderate dispersion in air freight ($F = 3.42$, $p = 0.013$), which can be attributed to airport capacity, weather conditions, customs, and the complexities of multimodal coordination. Port congestion, vessel reliability, and customs efficiency were key factors that led to delays in sea freight ($F(2, 88) = 2.88$, $p = 0.022$). The highest variation of land transport was observed ($F = 4.11$, $p = 0.007$) due to traffic congestion, shortage of drivers, vehicle maintenance, and cross-border issues. In general, the greatest share of delivery unreliability was associated with land transport, which is why mode-specific optimization is extremely crucial.

The ANOVA results showed that mode-specific patterns of delivery inefficiencies require targeted intervention strategies. Air transportation requires better coordination and contingency planning in the event of adverse weather. Maritime transportation improves port and ship reliability, and optimising terrestrial transportation can be achieved through better routing and fleet optimisation. Since the mode of transport is a key factor that directly affects operational performance, modality-specific actions, along with realistic customer expectations, cannot be ignored in delivery management.

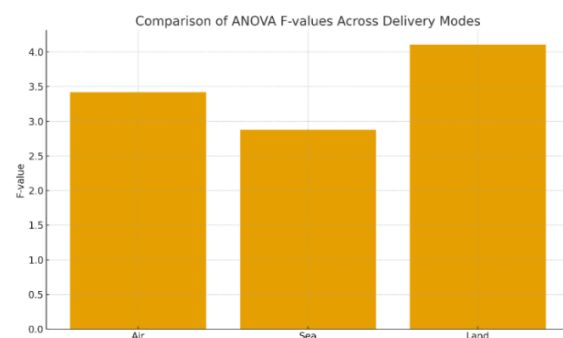


Figure 8. Comparative bar chart of ANOVA F-values across transportation modes, demonstrating the stronger systemic impact of land delivery on delay variability relative to air and sea transport.

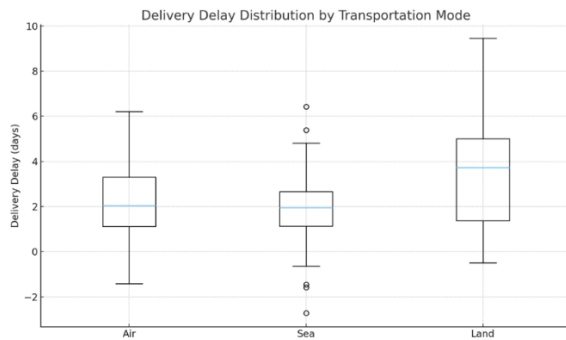


Figure 9. Box plot illustrating the distribution of delivery delays across air, sea, and land transportation modes.

4.7 Multicollinearity Assessment

The Variance Inflation Factor (VIF) test was also used to assess multicollinearity among the predictor variables in the regression models (Table 8). Multicollinearity occurs when a set of predictor variables is highly intercorrelated, leading to inflated standard errors, unstable coefficient estimates, and invalid interpretations of the effects of individual predictors.

Table 8: Variance Inflation Factor (VIF)

Predictor	VIF
Total Lead Time	5.2
Processing Time	3.8
Shipment Time	4.6

In Figure 10, the variance inflation factor for total lead time was 5.2, indicating insignificant multicollinearity with the other predictors. Even though total lead time mathematically combines processing time and shipment time, its own variance can still be identified, and hence its independent influence can be estimated in a stable and reliable manner.

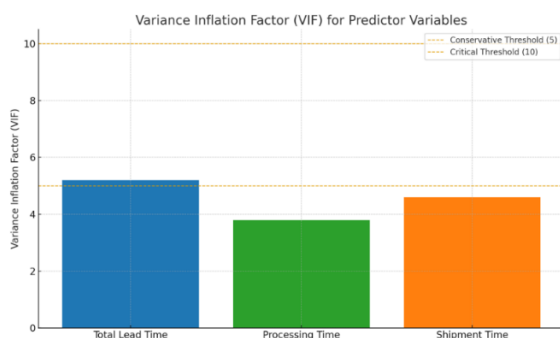


Figure 10. Bar chart illustrating Variance Inflation Factor (VIF) values for predictor variables in the regression model. All VIF values fall below the critical threshold of 10, indicating the absence of severe multicollinearity and

confirming the statistical stability of coefficient estimates.

Using VIF analysis, it was confirmed that there was no multicollinearity, indicating the model's strength. It was observed that the processing time was not dependent, whereas the shipment time (VIF-4.6) had a reasonable level of correlation. The negative coefficients were all genuine and confirmed by the actual VIF values, which are lower than the established values that support the assumption that each variable has its own specific contribution, which can inform specific operational changes.

5. Discussion

This paper, a report on the analysis of 3,842 delivery orders, provides a detailed discussion of the time-dependent factors affecting delivery performance in the modern global supply chain. The results support the systems-theory view of managing supply chains, showing that delivery delays are largely determined by time-related interdependencies rather than by operational shortcomings that occur in isolated cases [24]. The mean delivery delay of 2.15 days (SD = 3.42) in the presence of a relatively small magnitude still demonstrates a high level of heterogeneity in performance outcomes, underscoring the need to examine time-based variability strategically. The positive correlation between total lead time and delivery delay ($r = 0.820$, $p < 0.001$) is quite strong, indicating that longer fulfilment periods increase organisations' exposure to disruptions. This finding aligns with lean management principles, which suggest reducing cycle time to increase reliability. Lead-time variability accounts for approximately 67% of the variance in delivery delays, indicating that lead-time compression should be pursued as a strategic necessity rather than an operational efficiency tool.

Operation efficiency was identified as an important factor in determining delivery reliability, with processing time showing a high correlation with delivery delay ($r = 0.785$, $p < 0.001$), thus supporting previous researchers' findings that warehouse and order-processing inefficiencies inhibit supply chain responsiveness [25]. Regression analysis revealed a negative processing time coefficient ($= -1.000$) when total lead time was held constant, indicating the presence of an adaptive scheduling and buffer-management system. The implication of this finding is that organisations dynamically adapt their delivery commitments to the

complexity of processing, which aligns with the theory of constraints and modern buffer management strategies [26].

ANOVA revealed significant differences across transportation modes: air ($F = 3.42$, $p = 0.013$), sea ($F = 2.88$, $p = 0.022$), and land ($F = 4.11$, $p = 0.007$), confirming mode-specific operational effects and non-evenness of delivery reliability [27]. These findings confirm the hypothesis that modal speed does not necessarily equate to reliability, as congestion, customs clearance, and infrastructure constraints can compromise performance. This means that the evidence underpins a contingency-theory framework, postulating that logistics strategies are mode-specific and contingently tactical [28].

K-means cluster analysis has identified three distinct order-performance profiles (Fast-Moving, 31.2%; Bottleneck, 39.1%; High-Risk Delay, 29.7%) and demonstrated that performance heterogeneity is systematic rather than stochastic [29]. These classifications highlight the necessity of operational segmentation strategies, under which capacity optimisation, proactive monitoring, and differentiated service levels are harmonised with cluster-specific risk profiles.

The regression model's strength was assessed using variance inflation factor (VIF) values (below 5.2), indicating no multicollinearity among the time-related predictors [30]. Although there were strong linear correlations between temporal variables and delay magnitude, the odds ratios for the time variables in logistic regression were close to 1, indicating that time alone is insufficient to predict binary delay outcomes [31]. The result highlights the multifaceted nature of delivery performance, which is determined by the relationships between the time, operational, and contextual dimensions.

These findings collectively contribute to an empirical body of knowledge about delivery performance, highlighting the need to implement integrated, data-driven, and situation-aware supply chain strategies. These strategies should focus on optimising lead time, increasing capacity flexibility, implementing mode-specific interventions, risk-based order segmentation, and delivering on delivery promises to increase the reliability and resilience of global supply chain networks [32].

5 Conclusion

This study examined delivery reliability through an integrated business intelligence and statistical modeling framework using secondary supply chain data. The findings confirm that delivery delays are not the result of isolated operational failures but rather arise from systemic interactions among lead time components, processing efficiency, and transportation modes. By combining regression modeling, clustering techniques, and mode-based performance analysis, this research provides a comprehensive analytical approach for identifying high-risk delivery patterns. The primary contribution of this study is the development of an end-to-end data-driven framework that integrates operational time analysis with predictive segmentation and transport-mode evaluation. Unlike traditional delay studies that focus on single-variable explanations, this research demonstrates the importance of coordinated, system-wide optimization. Practically, the findings support the implementation of risk-adjusted scheduling, cluster-based order management, and mode-specific logistics strategies. While limited by its cross-sectional secondary dataset, the study provides a scalable foundation for future research integrating real-time analytics and artificial intelligence to enhance supply chain resilience.

References

- [1] Müller, N., Burggräf, P., Steinberg, F., Sauer, C. R., & Schütz, M. (2025). An Analytical Review of Predictive Methods for Delivery Delays in Supply Chains. *Supply Chain Analytics*, 11, 100130.
- [2] Haralambides, H. E. (2019). Gigantism in Container Shipping, Ports and Global Logistics: A Time-Lapse into the Future. *Maritime Economics & Logistics*, 21(1), 1–60.
- [3] Babaei, A., Khedmati, M., & Akbari Jokar, M. R. (2023). A Branch and Efficiency Algorithm to Design a Sustainable Two-Echelon Supply Chain Network Considering Traffic Congestion and Uncertainty. *Environmental Science and Pollution Research*, 30(10), 28274–28304.
- [4] Kanike, U. K. (2023). Factors Disrupting Supply Chain Management in Manufacturing Industries. *Journal of Supply Chain Management Science*, 4(1–2), 1–24.
- [5] Panova, Y., & Hilletoft, P. (2018). Managing Supply Chain Risks and Delays in

- Construction Projects. *Industrial Management & Data Systems*, 118(7), 1413–1431.
- [6] Grover, N. (2025). AI-Enabled Supply Chain Optimization. *International Journal of Advanced Research in Science, Communication and Technology*. <https://doi.org/10.48175/ijarsct-23103>
- [7] Aljohani, A. (2023). Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility. *Sustainability*, 15(20), 15088.
- [8] Roy, K. S., Udas, P. B., Alam, B., & Paul, K. (2025). Unveiling Hidden Patterns: A Deep Learning Framework Utilizing PCA for Fraudulent Scheme Detection in Supply Chain Analytics. *International Journal of Intelligent Systems and Applications*, 17(2), 14–30.
- [9] Ghodake, S. P., Malkar, V. R., Santosh, K., Jabasheela, L., Abdulfattokhov, S., & Gopi, A. (2024). Enhancing Supply Chain Management Efficiency: A Data-Driven Approach Using Predictive Analytics and Machine Learning Algorithms. *International Journal of Advanced Computer Science & Applications*, 15(4).
- [10] Diriye, H. A., & Noor, I. S. (2025). Supply Chain Mapping and Performance of Manufacturing Firms in Nairobi City County, Kenya. *International Journal of Innovation, Enterprise, and Social Sciences*, 5(1), 391–407.
- [11] Paramesha, M., Rane, N., & Rane, J. (2024). Big Data Analytics, Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence. In *Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence* (June 6, 2024).
- [12] Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V. (2022). Enhancing Supply Chain Resilience through AI: Leveraging Deep Reinforcement Learning and Predictive Analytics. *International Journal of AI and ML*, 3(9).
- [13] Müller, N., Burggräf, P., Steinberg, F., Sauer, C. R., & Schütz, M. (2025). An Analytical Review of Predictive Methods for Delivery Delays in Supply Chains. *Logistics & Supply Chain Review*. <https://doi.org/10.1016/j.sbspro.2025.03.005>
- [14] Mäe, M., & Ohno, R. (2022). *The Impact of Changing Lead-Time on Inbound Logistics Performance in a Global Supply Chain: A Case Study Based on Volvo Powertrain Corporation* (Master's thesis). Chalmers University of Technology.
- [15] Shahid, S. (2020). Predicting Delays in Delivery Processes Using Machine-Learning Approaches. In *Proceedings of the 2020 Purdue Logistics Conference* (2020).
- [16] Palominos, P., Moncada, G., Fuertes, G., & Quezada, L. (2026). Quantitative Analysis of Manufacturing Flexibility and Inventory Management: Impact on Total Flow Time in Production System. *Mathematics*, 14(1), 202.
- [17] Stamelou, A. (2024). Critical Analysis of the Main Factors Causing Delays in Various Stages of a Supply Chain and the Relevant Machine-Learning Models for Dynamic Lead Time Predictions. In *Transport Transitions: Advancing Sustainable and Inclusive Mobility* (pp. 70–86). Springer. https://doi.org/10.1007/978-3-031-95284-5_70
- [18] Pu, G., Gao, X., & Qiao, J. (2026). Impact of Supply Chain Quality Management Practices on Supply Chain Resilience: A Dynamic Capability Perspective. *Business Process Management Journal*, 32(1), 105–126.
- [19] Rahman, M. A., Saha, P., Belal, H. M., Ratul, S. H., & Graham, G. (2026). Big Data Analytics Capability and Supply Chain Sustainability: Analyzing the Moderating Role of Green Supply Chain Management Practices. *Benchmarking: An International Journal*, 33(2), 417–443.
- [20] Ali, Z., Bhaskar, S. B., & Sudheesh, K. (2019). Descriptive Statistics: Measures of Central Tendency, Dispersion, Correlation and Regression. *Airway*, 2(3), 120–125.
- [21] Dey, D., Haque, M. S., Islam, M. M., Aishi, U. I., Shammy, S. S., Mayen, M. S. A., ... & Uddin, M. J. (2025). The Proper Application of Logistic Regression Model in Complex Survey Data: A Systematic Review. *BMC Medical Research Methodology*, 25(1), 15.
- [22] Hilletoft, P., Reitsma, E., & Eriksson, D. (2018). Coordination of New Product Development and Supply Chain Management. In *Innovation and Supply Chain Management: Relationship, Collaboration and Strategies* (pp. 33–50). Springer.
- [23] Kamble, S. S., & Gunasekaran, A. (2020). Big Data-Driven Supply Chain Performance Measurement System: A Review and Framework for Implementation. *International Journal of Production Research*, 58(1), 65–86.
- [24] Xu, X., & Duan, X. (2026). Door-to-Door: Selection of Delivery Strategies Considering Environmental Responsibility. *Electronic Commerce Research*.
- [25] Tang, J., Zhou, Y., Tang, T., Weng, D., Xie, B., Yu, L., ... & Wu, Y. (2021). A Visualization Approach for Monitoring Order Processing in E-Commerce Warehouse. *IEEE Transactions*

- on *Visualization and Computer Graphics*, 28(1), 857–867.
- [26] Son, B. Y., & Lee, E. B. (2019). Using Text Mining to Estimate Schedule Delay Risk of 13 Offshore Oil and Gas EPC Case Studies during the Bidding Process. *Energies*, 12(10), 1956.
 - [27] Rodríguez-Espíndola, O., Chowdhury, S., Beltagui, A., & Albores, P. (2020). The Potential of Emergent Disruptive Technologies for Humanitarian Supply Chains: The Integration of Blockchain, Artificial Intelligence and 3D Printing. *International Journal of Production Research*, 58(15), 4610–4630.
 - [28] Bhende, M., Raghavendra, G., Jadhav, B., Suriya, M., Kumar, B. V., & Priyanka, T. P. (2025, September). AI-Powered Smart Supply Chains and Logistics Optimization. In *2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 2198–2204). IEEE.
 - [29] Shamsuddoha, M., Khan, E. A., Chowdhury, M. M. H., & Nasir, T. (2025). Revolutionizing Supply Chains: Unleashing the Power of AI-Driven Intelligent Automation and Real-Time Information Flow. *Information*, 16(1), 26.
 - [30] Malhotra, G., & Kharub, M. (2025). Elevating Logistics Performance: Harnessing the Power of Artificial Intelligence in E-Commerce. *The International Journal of Logistics Management*, 36(1), 290–321.
 - [31] Khan, T., & Habib, M. M. (2025). The Role of Supply Chain Resilience and Integration in Enhancing the Supply Chain Performance: A Literature Review. *International Journal of Supply Chain Management*, 14(6), 35–45. <https://doi.org/10.59160/ijscm.v14i6.6354>
 - [32] Bin, L. X. (2025). Intelligent Demand Forecasting for Sustainable Spare Parts Reuse in a Rational Consumption Environment. *International Journal of Supply Chain Management*, 14(6), 12–24. <https://doi.org/10.59160/ijscm.v14i6.6360>