Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Factory-to-Customer Shipping Route Efficiency Analysis for Nassau Candy Distributor: A Data-Driven Logistics Intelligence Study

Geetha S

Data Analytics Intern, Unified Mentor, Bengaluru, Karnataka, India
geethachandan87@gmail.com

Abstract
Nassau Candy Distributor ships confectionery products from 5 regional factories to customers across the United States and Canada. Although the organisation maintains detailed order and shipment records, logistics decisions have historically been made without route-level efficiency intelligence, leaving questions about which factory-to-customer lanes are reliable, which experience frequent delays, and where geographic bottlenecks exist largely unanswered. This study undertakes an exploratory data analysis of 10,194 orders spanning January 2024 to December 2025, covering 5 factories, 4 customer regions, and 15 products across 3 product divisions. Shipping Lead Time is computed as the interval between order date and ship date, and an Efficiency Score, a Delay Rate, and a Route Volume metric are derived for each factory-region lane. Findings indicate that the network records a mean lead time of 5.22 days and an overall delay rate of 23.2% against a 7-day threshold, with Standard Class shipping responsible for the substantial majority of delay incidents despite carrying 60% of total order volume. Route-level analysis identifies Sugar Shack's connection to the Gulf region as the network's least efficient lane, while geographic analysis reveals that proximity to a factory does not reliably predict shipping speed. A Streamlit-based interactive dashboard was developed to operationalise these findings for ongoing monitoring, supporting filtering by date range, region, ship mode, and lead-time threshold. The study concludes with 5 data-supported recommendations addressing ship-mode allocation, factory-specific performance gaps, routing inefficiency, congestion hotspots, and network resilience.

Keywords: Logistics Analytics, Shipping Lead Time, Route Efficiency, Supply Chain Optimisation, Exploratory Data Analysis, Geographic Bottleneck Analysis, Streamlit Dashboard, Nassau Candy Distributor, Confectionery Distribution, Data Analytics

1. Introduction

Nassau Candy Distributor operates as a national distributor, shipping confectionery products from factories to customers across multiple regions of the United States and Canada. In distribution operations of this kind, shipping efficiency has a direct bearing on customer satisfaction, unplanned delays increase operational cost, and inefficient routing reduces the scalability of the business. Despite the availability of rich order and shipment data, logistics decisions at Nassau Candy have, until this study, been made without route-level efficiency intelligence. The organisation currently lacks clarity on which factory-to-customer routes are consistently efficient, which routes experience frequent delays, how shipping performance varies by region, state, and shipping mode, and where operational bottlenecks exist geographically. Without this visibility, logistics optimisation remains reactive rather than data-driven. This study addresses that gap by transforming raw order and shipment data into route-level operational intelligence.

The objective of the study is threefold: first, to establish a clean and validated dataset of shipping lead times across the network; second, to benchmark performance across ship modes, factories, regions, and individual factory-region routes; and third, to translate these findings into an interactive dashboard and a set of actionable recommendations that support ongoing, data-driven logistics decision-making.

2. Literature Review

Shipping route efficiency and lead-time performance have received sustained attention in supply chain literature as organisations recognise that delivery speed and reliability are direct determinants of customer satisfaction and operational cost. Christopher (2016) established that logistics agility — the ability to match supply-side throughput to demand-side variability — depends critically on route-level intelligence rather than aggregate network metrics [1]. Ballou (2004) further argued that distribution network design decisions, including factory-to-customer routing, must be grounded in empirically measured performance data rather than geographic heuristics, a finding directly motivating the route-benchmarking framework developed in this study [2].

The role of shipping mode selection in lead-time outcomes is well established. Chopra et al. (2019) demonstrated that the choice of transport mode is among the most consequential decisions in outbound logistics, with faster modes reducing average lead time substantially but carrying higher per-unit cost [3]. Their work underscores the cost-time tradeoff that characterises the gap between Same Day and Standard Class shipping observed in this study. Bowersox et al. (2002) extended this analysis to show that high-volume distribution lanes are particularly sensitive to mode composition: when the dominant mode of a high-volume lane operates near its lead-time threshold, the aggregate delay rate of the network is disproportionately elevated, consistent with the Standard Class findings reported in Section 6 [4].

The application of exploratory data analysis (EDA) to operational logistics problems was formalised by Tukey (1977), whose methods for visual pattern detection prior to formal modelling have become foundational to supply chain analytics [5]. Mentzer et al. (2001) subsequently argued that data-driven supply chain management requires systematic measurement of factory-to-customer performance lanes, distinguishing between network-level averages and the lane-specific variability that drives customer experience [6]. Geographic bottleneck analysis — identifying states and regions where shipment volume and delay rate jointly exceed acceptable thresholds — has been applied in distribution network research by Simchi-Levi et al. (2008), who found that congestion is rarely uniformly distributed and that a small number of high-traffic geographies typically account for a disproportionate share of aggregate network delay [7].

The translation of logistics analytics findings into interactive monitoring tools has gained traction as a discipline distinct from static reporting. Chen et al. (2016) showed that interactive dashboards built on real transactional data produce materially better operational outcomes than static reports, as they allow logistics managers to explore performance data contextually and apply real-time filters by route, mode, and geography [8]. The Streamlit framework has been widely adopted as a vehicle for this class of application, enabling Python-based EDA to be deployed as a shareable, stakeholder-facing monitoring tool without requiring front-end engineering (Streamlit Inc., 2024) [9]. This study follows that established approach, delivering an interactive 5-module dashboard that operationalises the route efficiency findings of the EDA into an ongoing management tool accessible to non-technical decision-makers at Nassau Candy Distributor.

3. Dataset Description

The dataset comprises 10,194 individual order records collected between January 2024 and December 2025. Each record represents a single product line item within a customer order and includes the fields described in Table 1. Five factories fulfil orders across four customer regions — Interior, Atlantic, Gulf, and Pacific — covering both United States states and Canadian provinces. 15 distinct products are organised into 3 product divisions: Chocolate, Sugar, and Other.

Table 1: Dataset Field Description

Field	Description
Row ID / Order ID	Unique row and order identifiers
Order Date / Ship Date	Date the order was placed and the date it was shipped
Ship Mode	Shipping method assigned to the order (Same Day, First Class, Second Class, Standard Class)
Customer ID, City, State/Province, Postal Code	Customer geographic identifiers
Region	Customer region: Interior, Atlantic, Gulf, or Pacific
Division, Product ID, Product Name	Product classification and identifiers
Sales, Units, Cost, Gross Profit	Order-level financial and quantity metrics

The five factories and their geographic coordinates are Lot's O' Nuts (Arizona), Wicked Choccy's (Georgia), Sugar Shack (North Dakota), Secret Factory (Iowa), and The Other Factory (Tennessee). Product-to-factory assignment follows a fixed mapping: Lot's O' Nuts and Wicked Choccy's exclusively produce Chocolate division items, Sugar Shack is the principal producer of Sugar division items, and Secret Factory and The Other Factory supplement the Sugar and Other divisions. This fixed mapping is a structural feature of the network that shapes several of the findings discussed in Section 6.

4. Analytical Methodology

4.1 Data Cleaning and Validation

Order Date and Ship Date fields were parsed and validated for consistency, geographic fields were standardised, and the dataset was checked for missing values; none were found across the 10,194 records. Shipping Lead Time was computed in days as Ship Date minus Order Date for every order. All data cleaning, aggregation, and statistical computation reported in this study, including the correlation and percentile-based analyses presented in Section 6, were performed using the pandas library (McKinney, 2022; Pandas Development Team, 2024) [10, 11].

4.2 Feature Engineering

Each shipment was flagged as delayed if its lead time exceeded 7 days, the default delivery-window benchmark used throughout the analysis. Shipments were also grouped by factory-to-region and factory-to-state/province pairings to support route-level analysis, and by ship mode for comparison across delivery methods.

4.3 Route Definition and Aggregation

Each route is defined as a Factory-to-Customer-Region pairing, yielding 20 distinct factory-region lanes across the 5 factories and 4 regions. For each route, total shipment volume, average lead time, and lead-time standard deviation (a measure of variability) were calculated.

4.4 Efficiency Benchmarking

Routes were rank-ordered by average lead time, and an Efficiency Score was computed on a 0-to-100 scale, normalised such that a score of 100 corresponds to the network's shortest average lead time. The top 10 and bottom 10 routes were identified, and performance was compared across ship modes within each route.

4.5 Geographic Bottleneck Analysis

State- and province-level average lead time and delay rate were computed to identify regions of high average lead time and regions combining high shipment volume with elevated delay rates, surfacing congestion-prone states. A distance-based proximity analysis additionally compared each state's geographically nearest factory against the factory that actually fulfilled the majority of its orders.

4.6 Key Performance Indicators

Table 2: Key Performance Indicators

KPI	Description
Shipping Lead Time	Ship Date minus Order Date, in days
Average Lead Time	Mean shipping duration per route, factory, or ship mode
Route Volume	Number of orders per factory-region route
Delay Frequency	Percentage of shipments exceeding the 7-day threshold
Route Efficiency Score	Normalised lead-time performance, 0 to 100

5. Results and Discussion

5.1 Descriptive Overview

Across the full dataset of 10,194 orders, shipping lead time ranges from 1 to 19 days, with a mean of 5.22 days, a median of 5 days, and a standard deviation of 3.07 days. The overall delay rate against the 7-day threshold is 23.2%. Total network revenue across the observation window is $141,783.63, with total gross profit of $93,442.80. The lead-time distribution is right-skewed: a cluster of shipments completes in a single day, reflecting Same Day and First Class deliveries, while a secondary cluster of slower shipments — predominantly Standard Class — extends the distribution's upper tail toward 19 days.

5.2 Ship Mode Performance Analysis

Table 3 summarises performance across the 4 shipping methods. Same Day and First Class deliver with zero recorded delays and minimal variability, but together account for only 20.6% of total shipment volume. Standard Class, by contrast, carries 60% of total order volume (6,120 shipments) and records a 36.9% delay rate — its 6.53-day mean lead time sits close enough to the 7-day threshold that ordinary shipment-to-shipment variability is sufficient to push a large share of orders into delay. Second Class occupies an intermediate position, with a 4.46-day average lead time and a 5.5% delay rate.

Table 3: Ship Mode Performance Summary

Ship Mode	Shipments	Avg Lead Time (days)	Delay Rate
Same Day	547	1.0	0.0%
First Class	1,548	2.5	0.0%
Second Class	1,979	4.46	5.5%
Standard Class	6,120	6.53	36.9%

The core operational implication is that the network's highest-volume mode also exhibits the highest delay rate. Because faster modes demonstrate delay rates at or near zero, the pattern does not indicate a network-wide capability constraint; rather, it points to a routing allocation imbalance in which a disproportionate share of orders defaults to Standard Class. A structured review of order-to-mode assignment logic is therefore likely to yield the most measurable improvement available within the existing network.

5.3 Regional Shipping Performance

At the regional level, performance is remarkably uniform: average lead times across Pacific, Gulf, Atlantic, and Interior range narrowly between 5.15 and 5.34 days, with delay rates clustered between 23% and 24%. This indicates that shipping performance at the regional level of aggregation is driven primarily by ship-mode composition, which is broadly consistent across regions, rather than by geography itself. Commercially, the regions differ more meaningfully: Pacific leads in both shipment volume (3,253 orders) and total sales ($46,301.53), followed by Atlantic (2,986 shipments, $41,197.24), while Gulf is the smallest region by volume (1,620 shipments) despite a competitive average lead time. Regional uniformity in lead time should not be interpreted as evidence of equal performance across the network as a whole, since state-level analysis (Section 6.6) reveals meaningful variation beneath the regional surface.

5.4 Factory Performance Analysis

Factory-level analysis (Table 4) indicates that Secret Factory, The Other Factory, and Lot's O' Nuts are largely comparable with respect to lead time, each recording an average of between 5.16 and 5.18 days, whereas Wicked Choccy's trails marginally at 5.27 days. Sugar Shack constitutes the clear outlier, exhibiting an average lead time of 6.09 days alongside a delay rate of 33%, the highest observed across the network by a substantial margin. From a commercial standpoint, Lot's O' Nuts emerges as the dominant contributor, generating $76,340 in total sales and $52,771 in gross profit across 5,692 shipments, a consequence of its exclusive mandate over 3 of the 5 Chocolate division products.

Table 4: Factory Performance Summary

Factory	Shipments	Avg Lead Time (days)	Delay Rate
Lot's O' Nuts	5,692	5.18	23%
Wicked Choccy's	4,152	5.27	24%
Secret Factory	217	5.16	20%
The Other Factory	100	5.17	23%
Sugar Shack	33	6.09	33%

Sugar Shack's elevated lead time and delay rate together establish it as the factory of highest priority for operational review, irrespective of a comparatively modest shipment volume of 33 orders (0.3% of total network volume). A factory reach analysis further indicates that Sugar Shack serves fewer states and provinces than any other factory in the network — 15 in total — while simultaneously recording the highest delay rate, indicating that its operational impact, though limited in scope, is concentrated rather than dispersed.

5.5 Route Efficiency Benchmarking

Across the 20 factory-region routes, efficiency scores span the full 0-to-100 range against a portfolio average lead time of approximately 5.2 days. The most efficient route, The Other Factory to Gulf, averages 3.63 days, while the least efficient, Sugar Shack to Gulf, averages 8.25 days — a 4.62-day gap within the same network. Critically, this variation is lane-specific rather than factory-wide: The Other Factory leads the network on its Gulf route yet ranks near the bottom on its Interior route, confirming that lane-level intervention, rather than factory-wide initiatives, is where targeted improvement is most likely to deliver measurable gains.

The network's operational demands are concentrated disproportionately on its highest-volume routes rather than its fastest ones. Lot's O' Nuts to Pacific (1,813 shipments) and Lot's O' Nuts to Atlantic (1,661 shipments) constitute the 2 highest-volume lanes in the network, each recording efficiency scores in the 67-to-69 range and delay rates of approximately 22%. No high-volume route in the network achieves an efficiency score above 70.

Table 5: Top and Bottom Routes by Average Lead Time

Route	Avg Lead Time (days)	Delay Rate
The Other Factory → Gulf	3.63	15.8%
Sugar Shack → Pacific	4.67	0.0%
Secret Factory → Atlantic	4.94	19.4%
The Other Factory → Interior	7.00	36.4%

Route	Avg Lead Time (days)	Delay Rate
Sugar Shack → Atlantic	6.33	38.9%
Sugar Shack → Gulf	8.25	50.0%

Standard deviation analysis provides a further basis for evaluating reliability beyond average lead time. Sugar Shack to Gulf is not only the slowest route in the network but also the most variable, with a standard deviation of approximately 5.91 days, meaning individual shipments on this lane range from as low as 1 day to as high as 14 days. This combination of elevated average lead time and elevated variability reflects a doubly elevated reliability risk that distinguishes this route from the rest of the network, where the remaining 19 routes range from 2.08 to 4.00 days of standard deviation.

5.6 Geographic Bottleneck Analysis

State-level analysis covering 49 United States states (9,994 orders; Canadian provinces are analysed separately given that all factories are US-based) finds that the correlation between distance to the nearest factory and average lead time is near zero ($r = 0.16$), and the correlation for non-nearest-factory routes is similarly negligible ($r = -0.03$). This indicates that factory-to-customer distance has no meaningful relationship with average lead time across the network — a finding the study terms an unrealised proximity advantage, since geographic closeness would be expected to translate into faster delivery but does not do so in practice. In 19 of 35 comparable states (54%), the non-nearest factory fulfils orders at the same speed or faster than the geographically nearest factory, indicating that this expected distance advantage is not materializing in practice, and likely reflects other operational factors that further analysis would need to identify.

8 states sit simultaneously above the 60th percentile on both shipment volume and delay rate, representing the network's highest-priority congestion hotspots (Table 6). Texas is the most commercially significant of these, with 985 shipments and a 26% delay rate; because of its scale, improving fulfilment performance in Texas alone would meaningfully reduce the network's overall delay count.

Table 6: Congestion Hotspot States (Above 60th Percentile on Volume and Delay Rate)

State	Shipments	Delay Rate
Texas	985	26%
Pennsylvania	587	25%
Washington	506	25%
Arizona	224	27%
Tennessee	183	27%
Indiana	149	26%
New Jersey	130	26%
Wisconsin	110	27%

At the country level, Canadian provinces Quebec and Ontario record the highest combination of volume and delay rate among the network's slowest-performing geographies, each at a 34% delay rate, making them the most data-supported targets for operational intervention.

5.7 Customer Segmentation by Order Value

Orders were segmented into High, Medium, and Low value tiers using the 75th and 25th sales percentiles as thresholds. Average lead time is effectively flat across all 3 tiers, at approximately 5.2 to 5.3 days, and delay rates show no substantial separation by tier (22.9% for Low and Medium value, 24.4% for High value). This suggests that order value is not a strong driver of fulfilment performance in this network: high-value orders do not appear to be prioritised for faster shipping, nor low-value orders deprioritised. This finding narrows the scope for operational intervention by ruling out order-value segmentation as a likely factor.

5.8 Monthly Trends

Average lead time is stable across all 24 months of the observation window (January 2024 to December 2025), oscillating narrowly between 4.8 and 5.8 days with no discernible seasonal pattern. Monthly delay rates range between 20% and 29%, meaning that roughly 1 in 4 orders misses its delivery window in any given month, without exception. Delay rates show no sign of improvement over the 2-year period: figures from late 2025 are nearly identical to those recorded in early 2024, indicating that the delay pattern is structural in nature and unlikely to resolve without targeted operational intervention.

5.9 Cross-Variable Correlation Validation

A correlation matrix across lead time, sales, units, gross profit, and cost serves as a methodological cross-check on findings established in earlier sections. Lead time records a near-zero correlation (0.02) with every financial variable, confirming that shipping speed has no measurable bearing on order size or profitability. By contrast, sales, gross profit, and cost form a tightly correlated cluster (correlations of 0.88 to 0.98), consistent with a structurally sound, volume-driven distribution business. This cross-validation strengthens confidence that fulfilment performance in this network is more plausibly explained by operational factors than by the financial characteristics of individual orders.

6. Streamlit Dashboard Implementation

To operationalise the findings of this study for ongoing use, an interactive web application was developed using the Streamlit framework [9]. The dashboard is organised into 5 tabs — Route Efficiency, Geographic Coverage, Ship Mode Analysis, Factory Intelligence, and Trends & Deep-Dive — consistent with the analytical structure of the underlying study. Users are able to filter results by date range, region, ship mode, factory, and a configurable lead-time delay threshold, allowing the dashboard to be used as a live monitoring tool rather than a static report.

The Geographic Coverage tab presents a United States choropleth map of average shipping lead time by state, alongside underserved-region and bottleneck-state identification. The Ship Mode Analysis tab visualises lead-time and delay-rate differences across the 4 shipping methods, the Factory Intelligence tab surfaces factory-level reach, clustering risk, and revenue/profit metrics, and the Trends & Deep-Dive tab supports monthly trend analysis, cross-variable correlation validation, and order-level shipment timeline drill-down. All summary tables produced during the exploratory analysis, including region, factory, route, and ship-mode aggregates, are exportable directly from the dashboard for downstream reporting use.

6.1 Architecture Overview

The analytical findings from this study were operationalised into a single-file, five-tab Streamlit web application (nassau_candy_dashboard.py). The application loads and caches the underlying order-level dataset to ensure responsive interactive filtering. Five cross-cutting sidebar filters — Date Range, Region, Ship Mode, Factory, and a configurable Delay Threshold slider — propagate across all five tabs simultaneously.

Table 7: Streamlit Dashboard Tab Architecture

Tab	Key Visualisations / Features
Route Efficiency	Route Performance Leaderboard heatmap; Efficiency Score vs Shipment Volume scatter; Top 10/Bottom 10 Route Leaderboard; Lead Time Variability by Route; Lead Time vs Profit Margin scatter; Distance vs Lead Time Correlation; Strategic Assessment cards
Geographic Coverage	Shipping efficiency choropleth; Region Metrics Summary; Factory Shipment Volume Breakdown; Factory Network map; Factory Reach & Expansion Indicators; Network Coverage Deep Dive; Factory Expansion Opportunity; Underserved Regions; Top 10 Bottleneck States; High Volume + High Delay States; Customer Segmentation by Order Value; Geographic Network Assessment cards
Ship Mode Analysis	Ship Mode Performance Comparison; Lead Time Variability & Distribution by Ship Mode; Ship Mode × Region and × Factory heatmaps; Ship Mode Risk & Efficiency Breakdown
Factory Intelligence	Product-Level Lead Time; Factory Revenue & Profit; Factory Reach Index; Factory Clustering Risk heatmap; Coast Exposure to Central Factory Cluster; Factory Network Assessment cards
Trends & Deep-Dive	Monthly Shipping Trends; Cross-Variable Correlation Matrix; 3-Month Shipment Volume Forecast; Order-Level Shipment Timeline; Division Performance Over Time

6.2 Technology Stack

• Python — core analytical and application language
• Streamlit — web application framework and interactive dashboard engine
• Pandas — data loading, cleaning, and aggregation
• NumPy — numerical operations, distance calculations, and threshold logic
• Plotly (Express, Graph Objects, Subplots) — interactive charts, choropleths, and heatmaps

The dashboard implements session-state-based filter management, with a reset control available for each of the five filter dimensions to allow users to return to the full-portfolio view without manual deselection. All summary tables produced during the exploratory analysis, including region, factory, route, and ship-mode aggregates, are exportable directly from the dashboard for downstream reporting use. All interactive visualisations, including choropleth maps, heatmaps, and scatter plots, were rendered using the Plotly graphing library (Plotly Technologies Inc., 2024) [12].

7. Discussion

The findings from this analysis converge on three overarching strategic conclusions for Nassau Candy Distributor.

First, the network's primary reliability risk is a routing allocation imbalance rather than a capacity constraint. Standard Class shipping carries 60% of total shipment volume and records a 36.9% delay rate — the highest of any shipping mode — while Same Day and First Class deliveries, which together account for only 20.6% of volume, record delay rates at or near zero. Because faster modes are capable of near-perfect on-time performance, the pattern does not indicate a network-wide capability constraint; rather, it indicates that a disproportionate share of orders defaults to the slowest available mode. A structured review of order-to-mode assignment logic therefore represents the single most measurable lever available within the existing network.

Second, geographic proximity is not translating into the delivery-speed advantage that would ordinarily be expected. Distance to the nearest factory shows only a weak correlation with average lead time ($r \approx 0.16$), and in 54% of comparable states, the non-nearest factory fulfils orders at the same speed or faster than the geographically nearest one. This unrealised proximity advantage suggests that operational factors not yet isolated in this analysis — rather than distance itself — better explain lead-time variation, and that further diagnostic work would be needed before geography-based interventions could be justified.

Third, Sugar Shack represents a concentrated rather than diffuse reliability risk. Despite accounting for only 0.3% of total network shipment volume, Sugar Shack records both the longest average lead time (6.09 days) and the highest delay rate (33%) of any factory, and serves the fewest states of any factory in the network. Because its footprint is narrow, operational review targeted specifically at Sugar Shack is likely to be a low-cost, high-clarity starting point relative to network-wide initiatives.

8. Recommendations

The following 5 recommendations are derived directly from the analytical findings of Section 6 and are prioritised by the scale of their potential operational impact.

8.1 Address Standard Class delay concentration as the primary network-wide risk

Standard Class accounts for 60% of all shipments and a 36.9% delay rate, making it the sole mode responsible for substantially all delay activity across every factory. A structured review of order-to-mode assignment logic, with the objective of shifting eligible orders to Second Class or First Class where operationally feasible, represents the single highest-leverage intervention available within the current network.

8.2 Prioritise Sugar Shack's Standard Class performance for operational review

Sugar Shack records the network's highest overall delay rate (33%) and highest average lead time (6.09 days). Its connection to the Gulf region (8.25 days, 50% delay rate) is the single highest-risk route in the network and should be the primary focus of operational review, with its Atlantic connection (6.33 days, 38.9% delay) as a secondary priority.

8.3 Investigate the unrealised proximity advantage at nearest factories

In 54% of comparable states, the non-nearest factory fulfils orders at the same speed or faster than the geographically closest one, indicating that nearest factories are not converting their proximity advantage into measurable delivery-speed improvements. A factory-level processing audit, particularly for factories serving states where non-nearest routing outperforms, is warranted.

8.4 Target congestion hotspot states for operational review

8 states combine high shipment volume with elevated delay rates, led by Texas, Pennsylvania, and Washington. These are commercially active, high-traffic states where delay is a systemic rather than incidental pattern; improving fulfilment performance in these 3 states alone would meaningfully reduce the network's overall delay count.

8.5 Evaluate factory capability expansion to address concentration risk

Lot's O' Nuts and Wicked Choccy's, as the sole producers of Chocolate division products, together handle the dominant share of network shipments, while Sugar Shack, Secret Factory, and The Other Factory are geographically clustered in the central United States. Expanding factory product capabilities and evaluating coastal distribution capacity should be assessed as medium-term strategic priorities to improve both network resilience and geographic coverage.

9. Conclusion

This study establishes a data-driven understanding of factory-to-customer shipping route efficiency for Nassau Candy Distributor. Transforming 10,194 order and shipment records into route-level operational intelligence, the analysis identifies Standard Class shipping as the dominant driver of network-wide delay, Sugar Shack as the factory most in need of operational review, and a measurable gap between geographic proximity and delivery performance. State-level congestion hotspots and factory concentration risks are similarly quantified, giving decision-makers specific, prioritised targets rather than broad guidance. The accompanying dashboard extends these findings into an ongoing monitoring capability, tracking route efficiency, ship-mode performance, and geographic bottlenecks as conditions evolve — supporting a shift from reactive to proactive logistics decision-making, with the potential to improve delivery reliability across the network.

Acknowledgement

I thank Unified Mentor for providing the project specification, dataset, and analytical framework that guided this research.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

References

1. Christopher M., Logistics and Supply Chain Management, FT Publishing International, 2016, 5th Edition, 1-328.
2. Ballou R.H., Business Logistics / Supply Chain Management, Pearson Prentice Hall, 2004, 5th Edition, 1-816.
3. Chopra S., Meindl P., Supply Chain Management: Strategy, Planning, and Operation, Pearson Education, 2019, 7th Edition, 1-528.
4. Bowersox D.J., Closs D.J., Cooper M.B., Supply Chain Logistics Management, McGraw-Hill, 2002, 1st Edition, 1-656.
5. Tukey J.W., Exploratory Data Analysis, Addison-Wesley, 1977, 1st Edition, 1-688.
6. Mentzer J.T., DeWitt W., Keebler J.S., Min S., Nix N.W., Smith C.D., Zacharia Z.G., Defining Supply Chain Management, Journal of Business Logistics, 2001, Vol. 22(2), 1-25.
7. Simchi-Levi D., Kaminsky P., Simchi-Levi E., Designing and Managing the Supply Chain, McGraw-Hill, 2008, 3rd Edition, 1-498.
8. Chen H., Chiang R.H.L., Storey V.C., Business Intelligence and Analytics: From Big Data to Big Impact, MIS Quarterly, 2016, Vol. 36(4), 1165-1188.
9. Streamlit Inc., Streamlit Documentation, 2024, https://docs.streamlit.io
10. McKinney W., Python for Data Analysis, O'Reilly Media, 2022, 3rd Edition, 1-579.
11. Pandas Development Team, Pandas Documentation, 2024, https://pandas.pydata.org/docs
12. Plotly Technologies Inc., Plotly Python Open Source Graphing Library, 2024, https://plotly.com/python