**From Chaos to Control: The Power of Data in Freight Logistics Optimization**

**1.0 Introduction**

A diagram of a pyramid

AI-generated content may be incorrect.

**Figure 1. Framework for Freight Logistics Optimization: Highlighting key focus areas, including freight movement analysis, operational efficiency, environmental impact assessment, safety and risk evaluation, and economic resilience analysis, to drive sustainable and resilient supply chain systems.**

**1.1 Background Information**

Freight transportation is a cornerstone of global trade that facilitates the movement of goods across supply chains and ensuring economic stability. However, the sector faces numerous challenges. These challenges include operational inefficiencies, rising costs, safety concerns, and environmental impacts. Addressing these challenges requires a data-driven approach to optimize logistics, reduce risks, and promote sustainability.

This project utilizes data from the Bureau of Transportation Statistics (BTS) and follows the CRISP-DM framework to systematically explore freight logistics data. It aims to generate actionable insights that enhance efficiency and resilience in freight operations based on analysed patterns in freight movement, assessed environmental impacts, and evaluated economic disruptions.

Recent global events, such as the COVID-19 pandemic and geopolitical tensions, have demonstrated the vulnerability of freight systems. Studies indicate that the pandemic significantly disrupted freight transportation, reinforcing the need for adaptive and robust logistics networks (Schofer et al., 2022). Additionally, geopolitical factors have caused disruptions in global supply chains, underscoring the necessity of proactive freight management strategies (Lawrence et al., 2024; Zheng, 2024).

The primary objective of this project is to leverage advanced analytics to enhance freight transportation systems by identifying inefficiencies, assessing risk factors, and recommending sustainability strategies. These insights will be particularly valuable for logistics managers, policymakers, and business analysts seeking to navigate an increasingly complex economic and regulatory landscape.

**1.2 Key Factors in Freight Transportation Analysis**

The analysis focuses on key categories influencing freight movement efficiency, cost, and sustainability:

* **Freight Movement Patterns:** Analyzing shipment trends across transportation modes and trade types.
* **Operational Efficiency:** Identifying logistical bottlenecks and cost inefficiencies.
* **Environmental Impact:** Evaluating carbon emissions and sustainability measures.
* **Safety and Risk Factors:** Assessing hazards and developing risk mitigation strategies.
* **Economic Disruptions:** Measuring the impact of global events on freight transport.

To effectively analyse these key categories, the following features are identified as crucial to the study:

**1. Shipment Identification & Trade Information**

|  |  |
| --- | --- |
| **Feature** | **Description** |
| **YEAR** | **Year** of the shipment (four-digit AD format). |
| **MONTH** | **Month** of the shipment (1 - 12). |
| **TRDTYPE** | Trade type, indicating whether the shipment is an **export** or **import**. |
| **COMMODITY2** | **2-digit commodity code** categorizing the type of goods being transported. |

**2. Shipment Origin & Destination Details**

|  |  |
| --- | --- |
| **Feature** | **Description** |
| **USASTATE** | U.S. state code where the freight originates or arrives. |
| **MEXSTATE** | Mexican state code, applicable when the shipment involves Mexico. |
| **CANPROV** | Canadian province code, applicable when the shipment involves Canada. |
| **COUNTRY** | Country code indicating the international origin or destination of the shipment (Canada: **1220**, Mexico: **2010**). |
| **DEPE** | Port or district code representing the shipment's processing location. |

**3. Transportation & Logistics Information**

|  |  |
| --- | --- |
| **Feature** | **Description** |
| **DISAGMOT** | Mode of transportation code specifying how the freight is transported (**Vessel, Air, Mail, Truck, Pipeline, Other, Foreign Trade Zones (FTZs)**). |
| **CONTCODE** | Indicates whether the shipment is **containerized (X)** or **non-containerized (0)**. |

**4. Economic & Cost Factors**

|  |  |
| --- | --- |
| **Feature** | **Description** |
| **VALUE** | Total **value of goods** being shipped, measured in U.S. dollars (USD). |
| **SHIPWT** | Total **shipping weight** of the goods in kilograms (Kg). |
| **FREIGHT\_CHARGES** | **Freight cost** associated with transporting the shipment (in USD). |
| **DF** | Indicates whether the **merchandise was produced in the U.S. (1 = Domestic) or outside the U.S. (2 = Foreign)**. |

By categorizing these expected features, the analysis remains focused on the key aspects of freight transportation:

* **Time-Based Features** (e.g., YEAR, MONTH) allow analysts to study shipment trends over time and assess seasonal fluctuations.
* **Origin & Destination Features** (e.g., USASTATE, COUNTRY) enable tracking of freight movement and identifying regional trade patterns.
* **Logistics Features** (e.g., DISAGMOT, CONTCODE) provide insights into transportation methods, efficiency, and infrastructure utilization.
* **Economic & Cost Features** (e.g., VALUE, SHIPWT, FREIGHT\_CHARGES) help in assessing financial impacts, cost optimization, and economic disruptions.

These structured categories ensure the analysis remains comprehensive, data-driven, and aligned with the core objectives of optimizing freight transportation efficiency, reducing costs, and mitigating risks.

**1.3 Hypothesis**

To establish a structured analytical approach, the following hypotheses are formulated:

* **Null Hypothesis (H₀)**: Transportation mode selection, risk management strategies, and sustainability measures have no statistically significant effect on freight transportation efficiency, environmental impact, or resilience.
* **Alternative Hypothesis (H₁)**: Optimized transportation mode selection, risk management strategies, and sustainability measures significantly improve freight transportation efficiency, reduce environmental impact, and enhance resilience.

By testing these hypotheses, the study aims to quantify the impact of strategic interventions in freight logistics.

**1.4 Key Business Questions**

The analysis is structured around critical business questions that guide insights and decision-making:

1. **What are the key patterns and trends in freight movement across different transportation modes and trade types?**

**Purpose**: To identify inefficiencies and areas for optimization in logistics and supply chain management.

1. **How do different transportation modes impact freight costs and operational efficiency?**

**Purpose**: To determine the most cost-effective and efficient transportation methods for different types of goods.

1. **What are the primary contributors to carbon emissions in freight transportation, and how can they be reduced?**

**Purpose**: To assess environmental impact and explore sustainable alternatives.

1. **What are the most common risks and safety concerns in freight transport, and how can they be mitigated?**

**Purpose**: To enhance freight safety by implementing proactive risk management strategies.

1. **How have economic disruptions (e.g., COVID-19, geopolitical tensions) affected freight movement, and what strategies can improve resilience?**

**Purpose**: To develop adaptive strategies that mitigate the impact of economic fluctuations on freight operations.

1. **How do domestic and foreign freight patterns compare, and what implications do they have for trade policies?**

**Purpose**: To analyse the effects of domestic vs. international freight trends on economic resilience and strategic decision-making.

By structuring the analysis around these questions, the project builds a logical progression toward validating or rejecting the null hypothesis.

**2.0 Methodology**

A structured approach is adopted to ensure that each step aligns with the project’s goal of improving freight transportation through data-driven insights. The methodology involves data exploration, feature analysis, and the development of interactive dashboards using Power BI for effective data visualization and stakeholder engagement.

**2.1 Exploratory Data Analysis (EDA) and Data Cleaning**

**2.1.1 Data Quality Assessment & Exploration**

* **Assess Data Structure**: Used methods like .info(), .head(), and .describe() to explore the dataset characteristics.
* **Check Duplicates**: Removed duplicate records to ensure the dataset integrity.
* **Validate Data Consistency**: Ensured values align with expected ranges (e.g., shipment months, trade types).
* **Identify Missing Values**: Verified and addressed missing data points.

**2.1.2 Features Analysis**

* **Univariate Analysis:** Visualized individual feature distributions using histograms, bar charts, and box plots.
* **Bivariate & Multivariate Analysis:** Examined relationships between features, such as trade type vs. freight charges.
* **Feature Engineering:** Converted encoded variables to readable text, normalized numerical values, and created new features like shipment seasonality indicators and economic impact scores.

**2.2 Interactive Dashboard Development using Power BI**

* **Data Integration:** Imported and structured data within Power BI for efficient visualization.
* **Dashboard Design:** Created interactive charts, maps, and reports to highlight trends and key insights.
* **Stakeholder Engagement:** Ensured visualizations are tailored to logistics managers, policymakers, and analysts for informed decision-making.

**2.3 Protocol for Replication**



**Figure 2. Illustration of the data refinement process, showcasing key steps from data quality assessment and duplicate removal to feature analysis and interactive dashboard development for delivering actionable insights.**

**References**

Lawrence, M., Homer-Dixon, T., Janzwood, S., Rockstöm, J., Renn, O., & Donges, J. F. (2024). Global polycrisis: the causal mechanisms of crisis entanglement. *Global Sustainability*, *7*, e6. https://doi.org/10.1017/SUS.2024.1

Schofer, J. L., Mahmassani, H. S., & Ng, M. T. M. (2022). Resilience of U.S. Rail Intermodal Freight during the Covid-19 Pandemic. *Research in Transportation Business & Management*, *43*, 100791. https://doi.org/10.1016/J.RTBM.2022.100791

Zheng, Z. (2024, December 19). *Enhancing supply chain resilience amid rising global risks*. UN Trade and Development (UNCTAD) Transport and Trade Facilitation Newsletter. https://unctad.org/news/enhancing-supply-chain-resilience-amid-rising-global-risks