Trans-Border Freight Data Analysis

Matthew Adenyo

Email: adenyomatthew@gmail.com
GitHub: `https://github.com/matth-adenyo/TransBorder_Freight_Data_Analysis`

Overview

Trans-border freight (the movement of goods across international boundaries) is a cornerstone of the global economy. It facilitates trade, fosters economic growth, and promotes cultural exchange. In the United States, the Bureau of Transportation Statistics (BTS) provides comprehensive data that underscores the significance of trans-border freight, particularly with neighbouring countries Canada and Mexico.

While trans-border freight is economically beneficial, it also poses environmental challenges. The transportation sector is a significant contributor to greenhouse gas (GHG) emissions. According to the BTS, in 2013, transportation was responsible for about 27% of all GHG emissions in the United States, with medium and heavy-duty trucks accounting for a substantial portion.

The benefits of trans-border freight are undeniable, but, it is important to prioritize safety and sustainability to address environmental concerns and ensure the well-being of communities.

Problem Statement

This Trans-Border Freight Data Analysis project aims to enhance the efficiency, sustainability, and safety of freight across North American borders. Given the increasing volume of goods transported across various modes (vessel, rail, air, truck, etc), there is a need to identify inefficiencies, reduce environmental impact, and ensure optimal economic performance. The key objective here is to identify inefficiencies, recognize patterns, and propose actionable solutions to improve overall performance and sustainability while minimizing delays, costs, and environmental hazards.

Analytical Questions

- Which US states have the highest value of trade with Canada and Mexico?
- Which US port or district has the highest value of trade with Canada and Mexico?
- What is the contribution of trade value by country?
- Is there any difference in the distribution of trade value by Export and Import?
- What is the total trade value by mode of transportation?
- What are the top commodities transported across the U.S.-Canada and U.S.-Mexico borders, and how do they vary by mode of transportation?
- What are the most frequently used modes of transportation for trans-border freight?
- How has the volume of trans-border freight changed over the years?
- What are the primary trends in freight movement across different transportation modes over the past five years?

Analytical Questions Cont'd

- How do economic disruptions (e.g., trade policies, global events) impact freight movement?
- Which transportation mode has the highest inefficiencies in terms of cost, time, and environmental impact?
- Which ports or districts experience the highest congestion or delays?
- Which transportation modes have the highest carbon footprint based on freight weight and distance?
- Are there underutilised transportation modes or routes that could be optimised?
- What data-driven recommendations can be provided to improve cross-border freight transportation?

Overview of Dataset

The dataset covers from 2020 to 2024, capturing various aspects of goods movement such as type of trade (Export or Import), mode of transportation (Truck, Air, Vessel, etc), U.S. States, Mexican States, Canadian Province, commodity, U.S. port, among others between the U.S. - Canada, and U.S. - Mexico.

- 2020: contains 9 sub-folders from January to September
- 2021: contains 12 sub-folders from January to December
- 2022: contains 12 sub-folders from January to December
- 2023: contains 12 sub-folders from January to December
- 2024: contains 9 sub-folders from January to September

Each of these sub-folders contains 3 to 6 different CSV dataset files. The CSV files are named dot1_MMYY, dot2_MMYY, and dot3_MMYY for all sub-folders. Some years have additional CSV data files named dot1_ytd_MMYY, dot2_ytd_MMYY, and dot3_ytd_MMYY while others have dot1_YYYY, dot2_YYYY, and dot3_YYYY in addition, where **MM** is the month number and **YY** is the last two digits of the year.

Methods Used

Several methods and techniques were used in this analysis to extract insights from the dataset. Below is a list of the methods used, along with a brief overview of each:

- **Data Concatenation:** The process of combining two or more data sets into one.
- **Data Mapping:** The process of replacing coded values in a column with corresponding meaningful labels.
- **Data Aggregation and Grouping:** This method groups the data by specific columns (e.g., USASTATE, COUNTRY, DISAGMOT) and performs aggregate operations like summing or counting.
- **Data Sorting:** This helps sort the data based on a specific column in ascending or descending order.

Methods Used Cont'd

- **Data Filtering:** Help to filter the dataset based on specific conditions (e.g., filtering data for specific years or transportation modes).
- **Data Visualization:** To create visualizations like line plots, bar charts, and histograms to represent trends and patterns in the data. Visualization helps in understanding the data and communicating insights effectively.
- **Trend Analysis:** This method involves analyzing how a metric (e.g., trade value) changes over time. It is used to identify trends and patterns over the years.
- **Inefficiency Analysis:** This involves deriving metrics that indicate inefficiencies (e.g., high freight charges relative to shipment weight). This helps in identifying areas for improvement.

Key Insights - Trade Distribution

US states with the highest trade value:

- Texas, Michigan, and California dominate cross-border trade.

Major trade partners:

- Canada accounts for 49.45% of trade values while Mexico accounts for 50.55%.
- Canada contributes over $5,131 billion trade value in Import and over $4,362 billion in Export while Mexico contributes over $5,739 billion in Import and over $3,965 billion in Export.

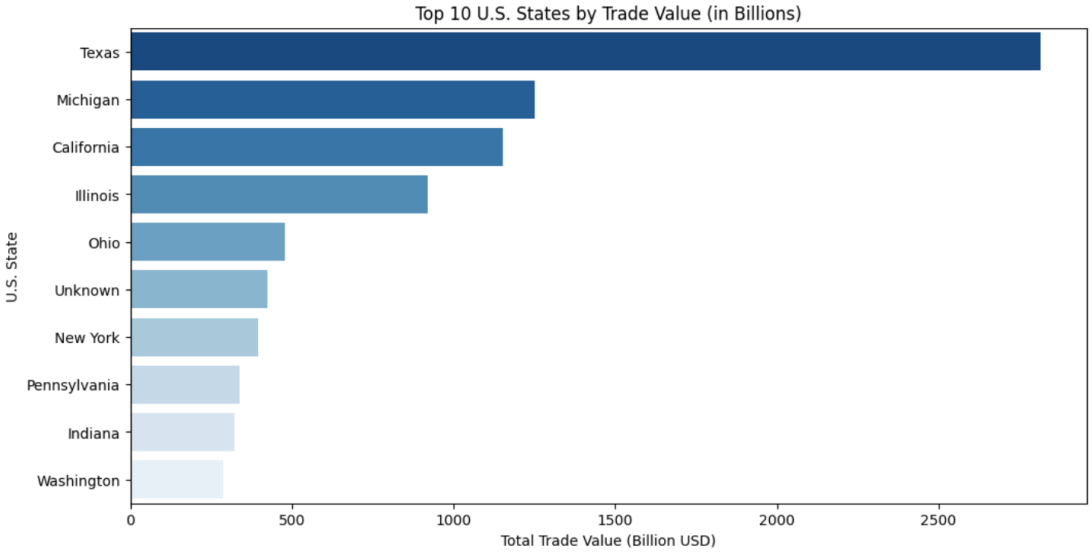
Top ports handling the highest trade value:

- Laredo, Detroit, and port huron lead in trade value.

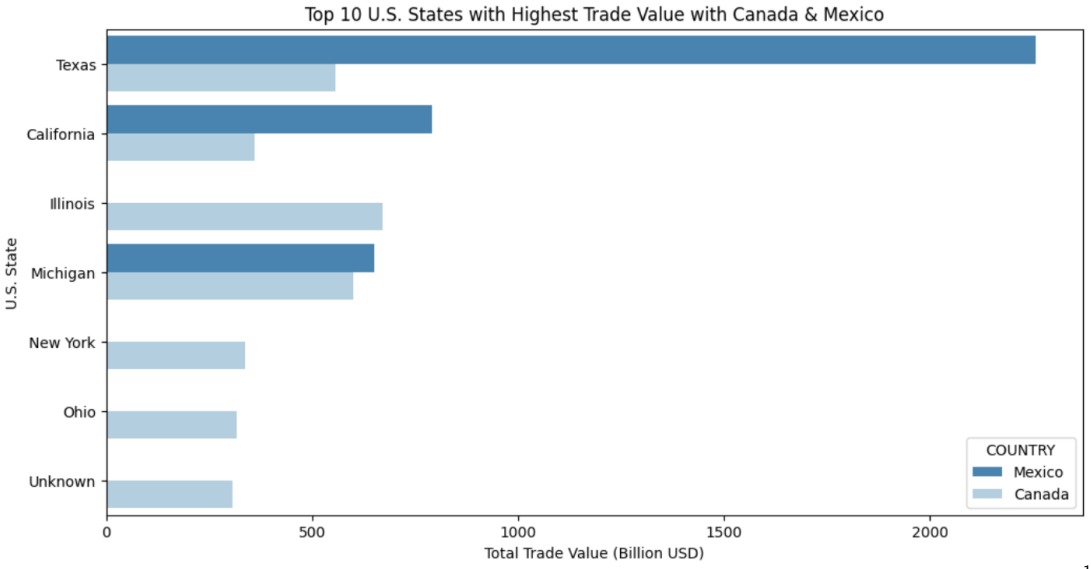
Top Commodity by Trade Value:

- Vehicles, other than railway or tramway rolling stock, and parts and accessories thereof lead in trade value.
- By mode of transport, Nuclear reactors, boilers, machinery and mechanical appliances dominate through Truck followed by Electrical machinery and equipment also through Trucks.

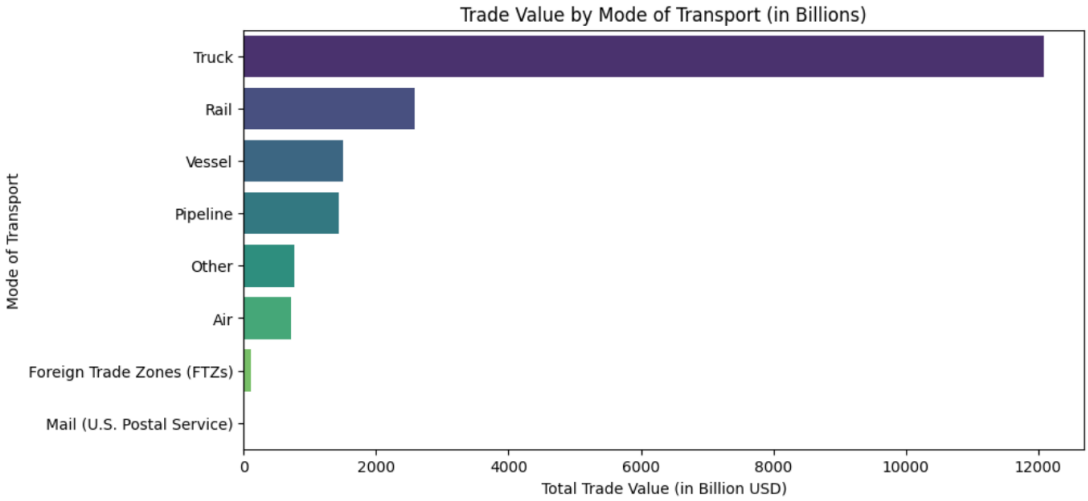
Key Insights - Trade Distribution



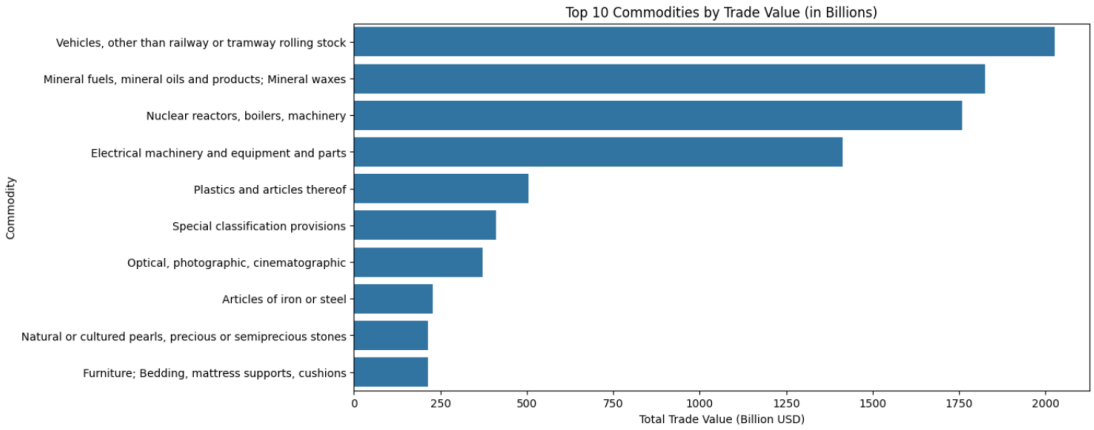
Key Insights - Trade Distribution



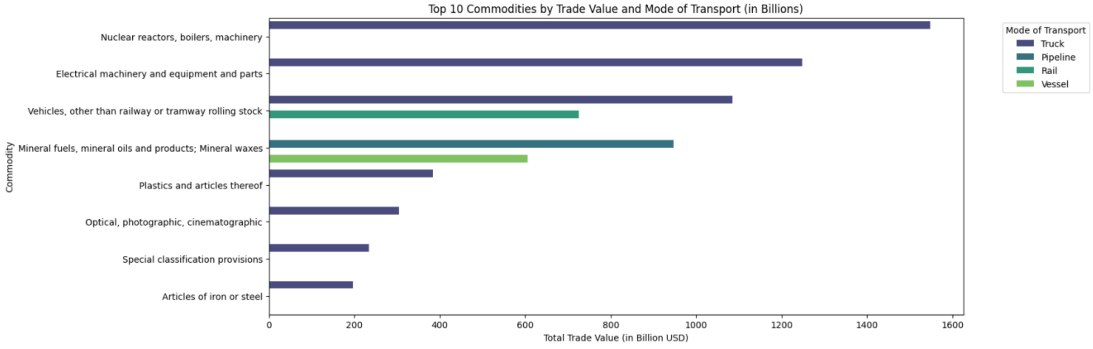
Key Insights - Trade Distribution



Key Insights - Trade Distribution



Key Insights - Trade Distribution



Key Insights - Trade Mode Analysis

Most Frequently Used Modes:

- Trucks account for more than 65% of transborder freight while air transport accounts for more than 17%.
- Vessel and Rail contribute significantly to trade efficiency.

Mode Contribution by Trade Value:

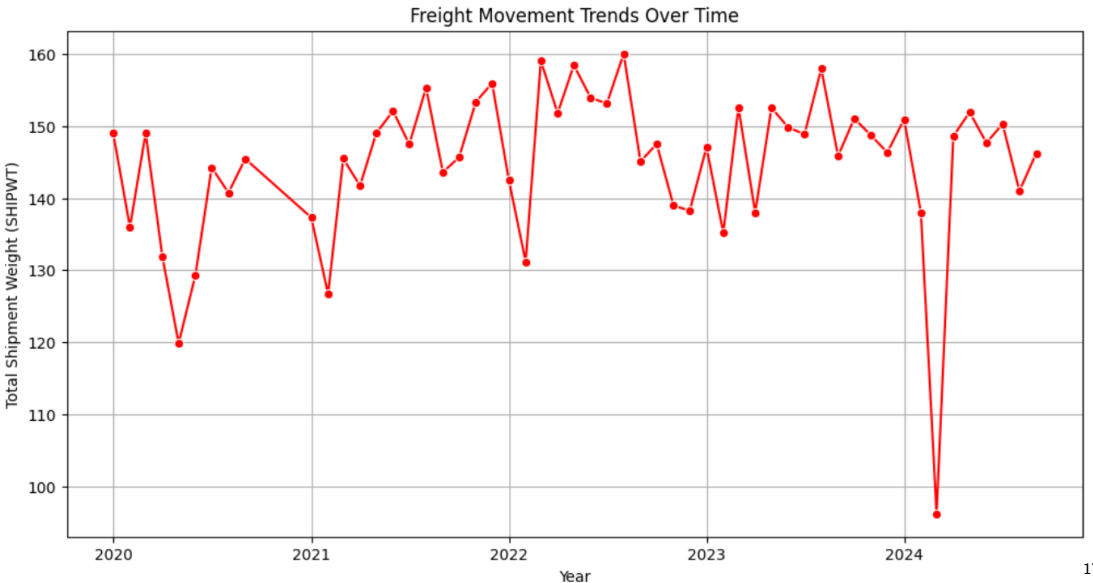
- Truck transport has the highest trade value with over $12,000 billion and the highest freight charge for zero-weight shipments with over $29 million.
- Pipeline shipment between U.S. - Canada is the most expensive transport.
- Cost per unit weight for freight between U.S.-Canada freight are the most expensive one.

Key Insights - Trend Analysis

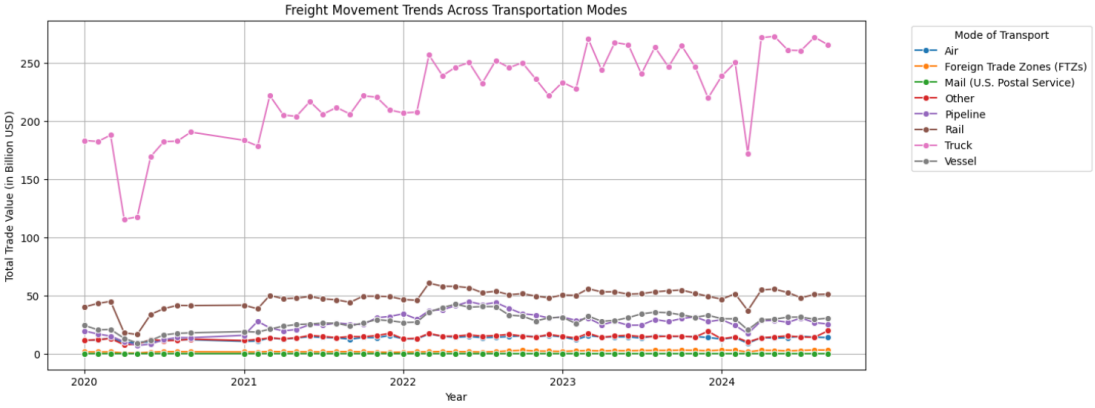
Seasonal Variation:

- There is a noticeable seasonal or operational variation in freight movement.
- 2020 and 2024 receive a sharp drop in freight movement possibly due to COVID-19 impact and trade policy.
- Despite fluctuations, truck freight trade has remained above $150 billion and even crossed $250 billion in some periods.
- Vessel transport has shown a gradual upward trend, indicating an increasing reliance on shipping trade.
- While rail transport is relatively stable but has periodic spikes.

Key Insights - Trend Analysis



Key Insights - Trend Analysis



Key Insights - Environmental Impact

High Carbon Footprint:

- Trucks have the largest carbon footprint contribution due to fuel consumption.

Congestion and Delays:

- Ports in Chicago experience bottlenecks receiving majorities of its goods through pipelines affecting efficiency.

Underutilized Transport Modes:

- Hawaii to Mexico is the Least Utilized Route having the lowest shipment weight.
- Alaska to Mexico is the 10th most underutilized route.

Recommendations

Optimize Trade Routes:

- Reduce congestion by investing in rail infrastructure and underutilized routes.

Sustainable Transport Solutions:

- Encourage greener fuel alternatives and carbon reduction strategies.
- Optimized rail transport for cost-effective trade.

Technology Integration:

- Implement AI-driven logistics for better freight management.

Policy Adjustments:

- Harmonize trade policies to minimize delays and inefficiencies.

Conclusion

- Addressing inefficiencies in trans-border freight is key to improving sustainability and economic performance.
- Data-driven strategies can enhance trade flow, reduce environmental harm, and improve cost-effectiveness.
- Further research and technological advancements are essential for long-term improvements.

The End

Email: adenyomatthew@gmail.com

GitHub: `https://github.com/matth-adenyo/TransBorder_Freight_Data_Analysis`