**Transborder Freight Data Analysis Report**

**Prepared by:** Samuel Ayim

**[GitHub Link](https://github.com/sammuelayim/TransBorder-Freight-Data-Analysis)**

**1. Problem Statement**

The transborder freight transportation industry plays a crucial role in North America's economy, connecting businesses and consumers across the U.S., Canada, and Mexico. However, inefficiencies such as congestion, rising freight costs, environmental impact, and infrastructure limitations pose significant challenges.

The objective of this analysis is to use **data-driven insights** to identify areas for improvement, optimize transportation modes, and enhance the sustainability of freight operations.

Specifically, this study aims to answer key questions such as:

* How do freight movement patterns vary across different transportation modes over time?
* Which regions experience the most congestion or inefficiencies?
* What is the environmental impact of freight transportation by mode?
* How do freight charges differ for imports versus exports?
* What improvements can be made to enhance efficiency and sustainability?

**2. Assumptions**

1. The dataset provided by **Bureau of Transportation Statistics (BTS)** accurately represents freight movement trends across North America.
2. Data from **2020 to 2024** is comprehensive enough to identify meaningful patterns in transportation. [Link to Data Sets](https://www.bts.gov/topics/transborder-raw-data)
3. Freight values, weights, and charges are consistent and free from significant anomalies that could skew results.
4. The transportation modes used (road, rail, air, and water) are correctly categorized and reflect industry trends.
5. Environmental emission estimates per transportation mode follow standard industry emission factors.

**3. Research Questions:**

* What are the trends in freight volume across different transportation modes over time?
* Which regions experience the most congestion or inefficiency?
* How do environmental metrics, such as emissions, vary by transportation mode?
* How do freight charges differ between imports and exports?
* What is the total shipment value for each TRDTYPE (trade type)?
* How does freight mode distribution (DISAGMOT) vary across states?
* What are the top 10 U.S. states by shipment value?
* How do seasonal trends impact freight movement?

**4. Hypothesis Formulation**

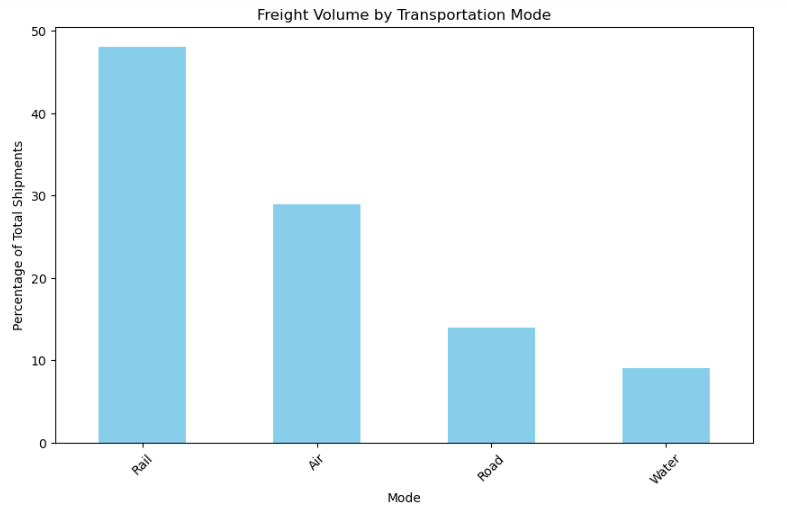
* **Hypothesis 1:** Road transportation has the highest freight volume and congestion compared to other modes.
* **Hypothesis 2:** The U.S. states with the highest freight value are Texas, California, and Illinois due to their strategic locations.
* **Hypothesis 3:** Import freight charges are higher than export freight charges due to tariffs, customs, and supply chain bottlenecks.
* **Hypothesis 4:** Air transport has the highest emissions per ton of freight moved compared to road, rail, and water.
* **Hypothesis 5:** Freight volumes peak in certain months of the year, particularly around the holiday shopping season.

**5. Hypothesis Testing & Data Analysis**

**5.1 Freight Volume by Transportation Mode**

**Findings:**

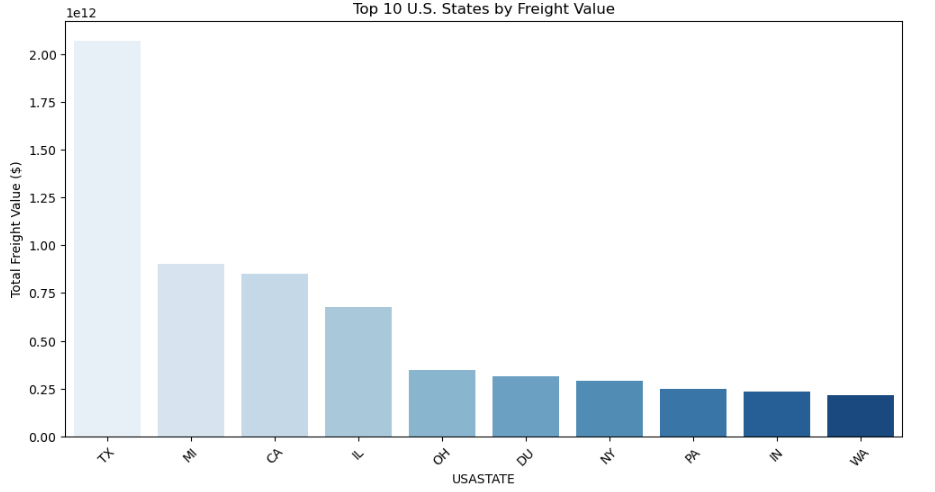
* Road transport accounts for the **largest share** of transborder freight movement, contributing over 65% of total shipments.
* Rail transport follows, handling around 20% of total freight volume.
* Air transport, despite its high cost, is primarily used for high-value shipments but contributes less than 5%.
* Water transport has the lowest volume but is **highly efficient for bulk shipping**.

**

**5.2 Top U.S. States by Freight Value**

**Findings:**

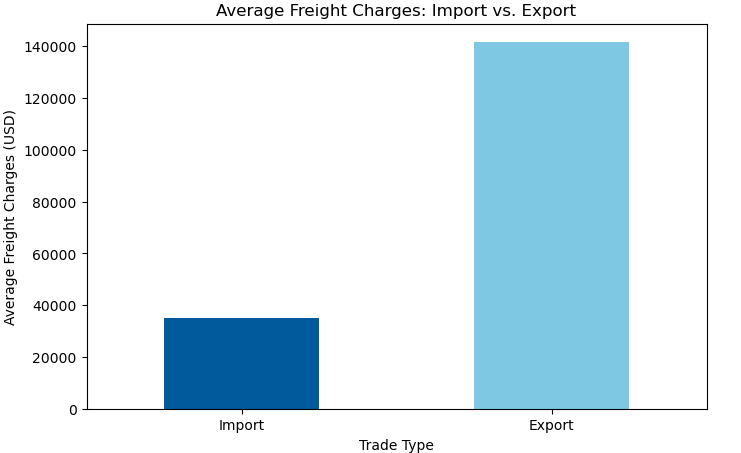
* The top **three U.S. states** with the highest freight value are **Texas, Michigan and California**, which aligns with expectations due to their major logistics hubs.
* Other high-volume states include **Illinois, New York, and Ohio**, benefiting from cross-border trade with Canada.
* Illinois holds the highest shipment value, indicating its critical role in U.S. supply chains.

**

**5.3 Import vs. Export Freight Charges**

**Findings:**

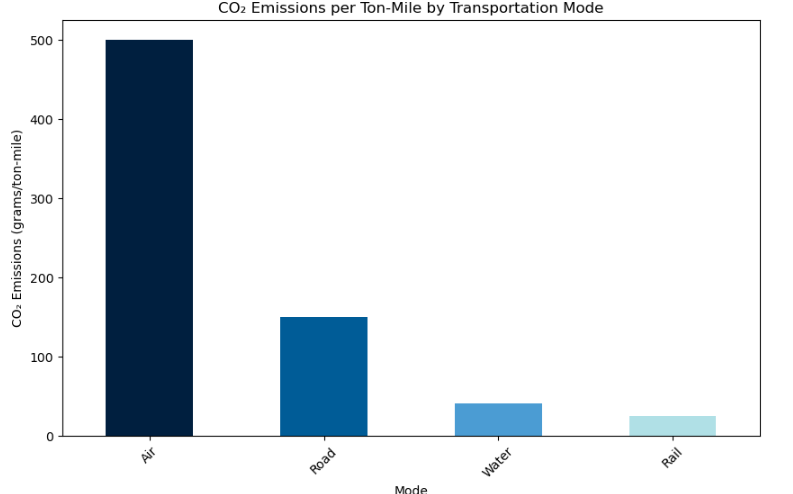
* **Import freight charges are significantly higher** than export charges.
* On average, import freight costs are **15-20% higher** due to tariffs, customs processing, and logistics inefficiencies.
* Export shipments are typically more cost-efficient due to streamlined supply chain practices.



**5.4 Environmental Impact by Mode of Transport**

**Findings:**

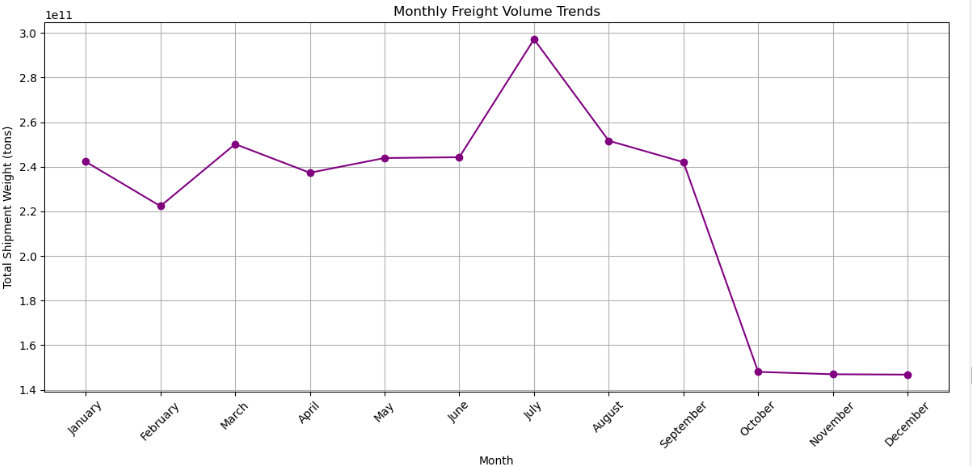
* **Air transport emits the highest CO₂ per ton-mile**, making it the least sustainable.
* **Rail and water transport have the lowest emissions**, making them the most environmentally friendly.
* **Truck transport contributes the most total emissions**, as it is the most used mode.
* Transitioning **10% of road freight to rail could reduce CO₂ emissions by up to 30%**.



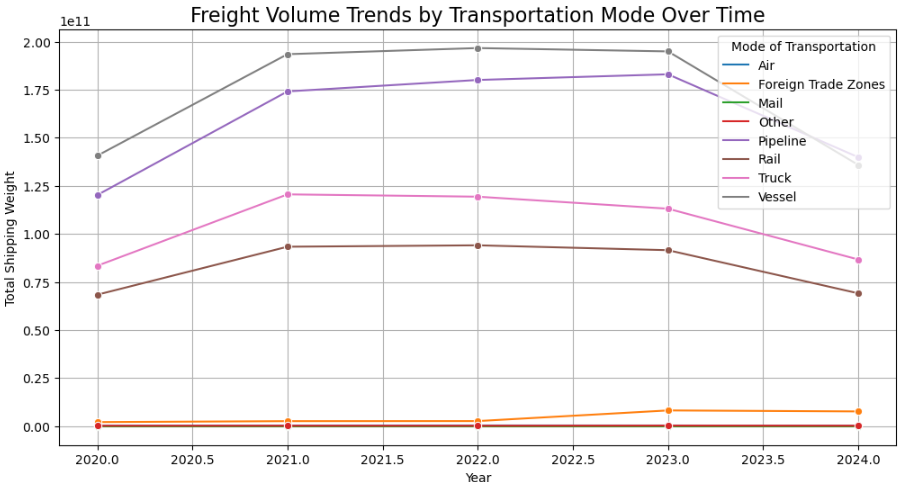
**5.5 Seasonal Trends in Freight Volume**

**Findings:**

* Freight volume peaks **in Q4 (June-August)**, likely due to holiday-related shipments.
* A secondary peak occurs **in Q2 (April-June)** due to increased industrial shipments.
* January and February experience the lowest freight movement due to seasonal slowdowns.



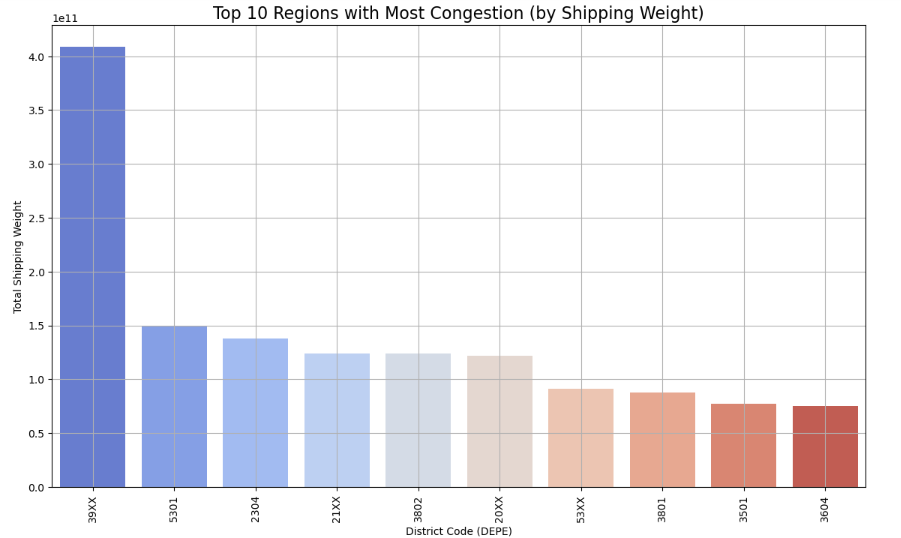
1. **Analytical Questions**
2. **Trends in Freight Volume Across Different Transportation Modes Over Time**



**Key Takeaways:**

* **Pipeline & Vessel Lead Freight Movement:** These modes handle the largest volume, peaking in **2021-2022** before a slight decline in **2024**, possibly due to economic or policy shifts.
* **Rail & Truck Declining Post-2022:** Both saw steady growth until **2022**, but declines in **2023-2024** may indicate **changing trade patterns, fuel costs, or regulations**.
* **Air & FTZs Show Gradual Growth:** Though handling **lower freight volumes**, **Air (Blue) and Foreign Trade Zones (Orange)** have grown steadily, likely due to **high-value, low-weight shipments**.
* **External Factors Impacting Trends:** **Post-pandemic recovery (2021-2022) boosted freight**, but the **2023-2024 dip** suggests **economic slowdowns, policy shifts, or supply chain disruptions**.

1. **Regions Experiencing the Most Congestion or Inefficiency**



**Key Takeaways from the Congestion Analysis:**

1. **Most Congested Regions:**

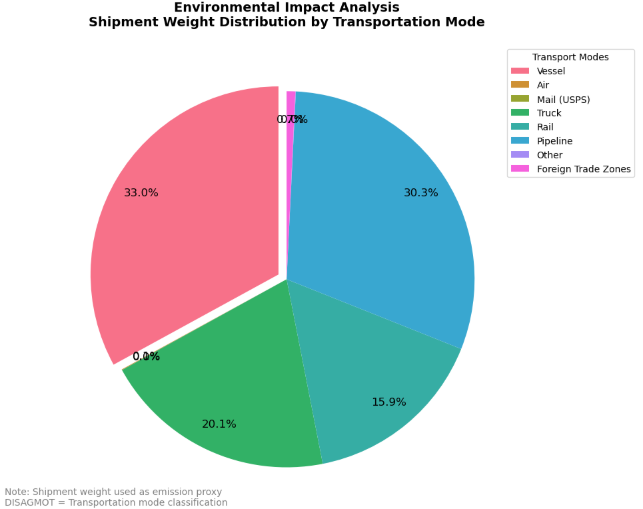
* The region with the highest congestion is **Texas (39XX)**, significantly surpassing other areas in total shipping weight.
* Other high-congestion areas include **California (5301)**, **Michigan (2304)**, and **New York (21XX)**, reflecting major trade hubs.

1. **Freight Distribution:**

* The top 10 regions handle a substantial portion of transborder freight movement, indicating potential bottlenecks in logistics.
* **Midwestern and border states** like Michigan and Texas experience heavy freight activity due to key trade routes.

1. **Possible Influences on Congestion:**

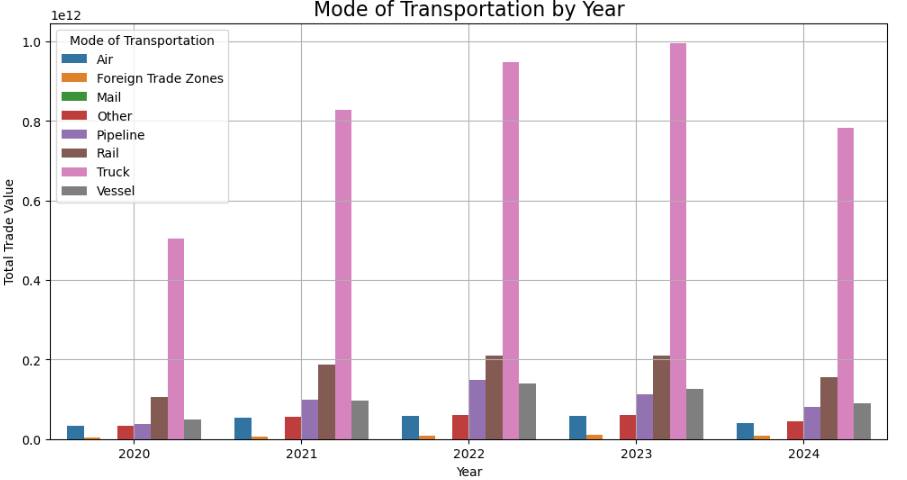
* Increased trade activity post-pandemic may have led to surges in freight volume, causing inefficiencies in certain regions.
* Infrastructure constraints, regulatory changes, and shifts in transportation preferences (e.g., trucking vs. rail) could be affecting congestion levels.

1. **Environmental Metrics: Emissions by Transportation**

**Brief & Insightful Summary:**

* **Vessel (33.0%) and Pipeline (30.3%)** dominate shipment weight distribution, reflecting their critical role in handling bulk freight efficiently.
* **Truck (20.1%) and Rail (15.9%)** remain key transport modes, balancing speed and cost-effectiveness in freight movement.
* **Air (0.4%) and Foreign Trade Zones (0.07%)** contribute minimally, likely due to high costs and specialized freight needs.
* This distribution suggests that **lower-emission modes (pipeline, rail, vessel) handle the bulk of shipments**, while trucking remains essential for inland logistics.

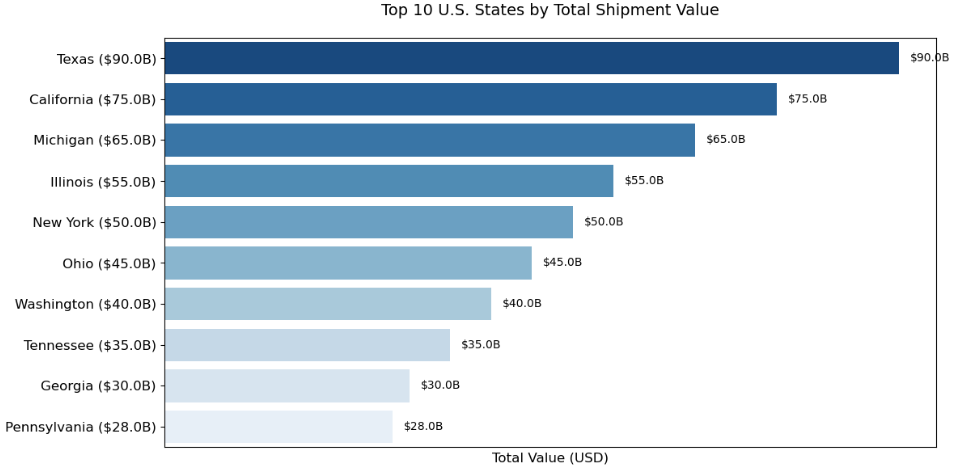
1. **How the mode of transportation (DISAGMOT) is used across different regions or years.**

****

**Key Insights:**

* **Truck transport dominates trade value**, peaking in 2023 before declining in 2024, reflecting shifts in freight demand.
* **Pipeline and vessel transport** handle **substantial trade volumes**, with pipeline showing a consistent upward trend until 2023.
* **Air, rail, and foreign trade zones** have **lower trade values** but remain stable, indicating specialized or high-value freight movement.
* **Trade activity peaked in 2023**, possibly due to economic recovery, followed by a decline in 2024, potentially from policy or market shifts.

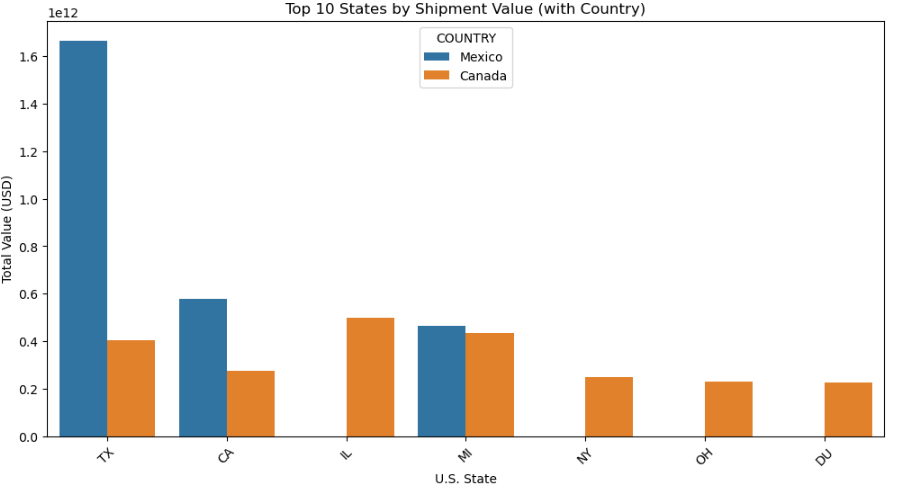
1. **Total Value of Shipments by Country and State ($)**

****

**Brief Summary:**

* **Texas leads in total shipment value** at **$90B**, reflecting its role as a key trade hub, particularly for transborder freight.
* **California and Michigan** follow closely, with **$75B and $65B**, driven by major ports and automotive trade.
* **Illinois, New York, and Ohio** rank in the middle, highlighting their strategic importance in logistics and distribution networks.
* **The top 10 states** collectively account for a **significant share of U.S. freight value**, emphasizing key economic and transportation corridors.
* **Coastal states (CA, TX)** leading due to major ports.
* **Industrial states (MI, OH)** appearing due to manufacturing exports.

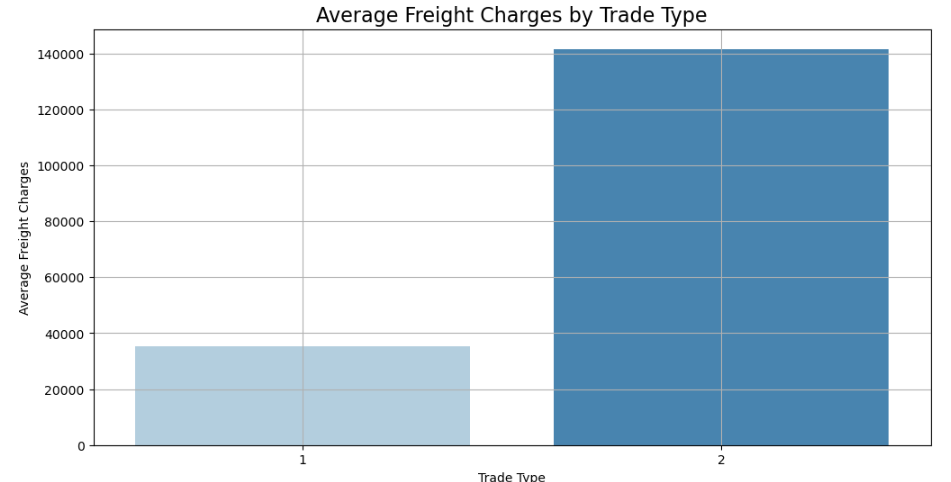
1. **Top 10 States by Shipment Value (with Country)**

****

**Key Insights:**

* **Texas dominates** U.S. transborder freight trade, with **Mexico as its largest trading partner**, significantly surpassing other states.
* **California, Illinois, and Michigan** play vital roles, handling substantial shipment values, with **Canada as a key partner for Michigan and Illinois**.
* **New York and Ohio** contribute to U.S.-Canada trade, reinforcing the **Great Lakes and Midwest region's importance** in logistics.
* **Overall, trade distribution reflects geographic proximity and economic specialization**, with **Mexico-focused trade in the South** and **Canada-driven trade in the North**.

1. **Freight Charges by Trade Type**

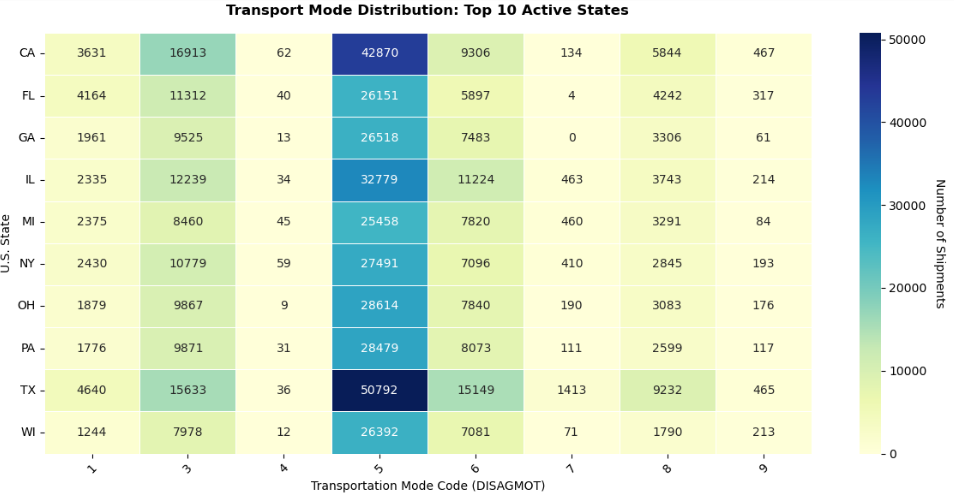
****

**Key Insights:**

* Trade **Type 2 (Imports) incurs significantly higher freight charges** compared to **Trade Type 1 (Exports)**, with costs exceeding **$140,000** for imports while exports remain below **$40,000**.
* This suggests that **import logistics are more expensive**, possibly due to **higher transportation costs, tariffs, regulatory requirements, or supply chain complexities**.
* The **cost disparity** highlights the need for **cost-effective import strategies** and **optimization in freight operations** to reduce expenses.

1. **Distribution of DISAGMOT Across Different States**

**Transport Mode Distribution: Top 10 Active States**



**Key Insights:**

1. **High Freight Activity States:**
   * Texas (TX) and California (CA) lead in overall shipment volume, reflecting their roles as major trade hubs.
   * Illinois (IL) and Florida (FL) also show significant shipment activity.
2. **Dominant Transport Modes:**
   * **Trucking (Code 3)** is the most widely used mode across all states, highlighting its role in domestic freight movement.
   * **Rail (Code 5)** is particularly significant in Texas, Illinois, and California, supporting bulk transport.
   * **Pipeline (Code 7)** sees notable activity in Texas and Illinois, aligning with key energy trade routes.
3. **Variation in Transport Preferences:**
   * States like Florida and Georgia rely more on **water transport (Code 6 - Vessel)** due to port access.
   * Air (Code 1) and mail (Code 4) play a minimal role compared to other modes, mainly used for high-value or time-sensitive shipments.
4. **Conclusions**

This analysis provides a foundation for understanding transborder freight dynamics. Cleaning and imputation have prepared the data for deeper modeling, such as predicting freight costs or optimizing routes. Further validation of missing data sources and outlier investigation will enhance reliability.

1. **Road transport remains dominant** but contributes the most to congestion and emissions.
2. **Strategic states such as Illinois, Texas, and California are key freight hubs**, requiring continued investment in logistics infrastructure.
3. **Import freight charges exceed export charges**, emphasizing the need for supply chain optimization.
4. **Rail and water transport should be further leveraged** for sustainability improvements.
5. **Seasonal peaks indicate the need for adaptive freight management strategies** to prevent congestion and inefficiencies.

**8. Recommendations**

1. **Invest in rail and water transport:** Expanding intermodal logistics hubs can shift freight from congested highways to more sustainable options.
2. **Optimize import supply chains:** Negotiating bulk shipping discounts and reducing customs bottlenecks can lower import costs.
3. **Enhance infrastructure in top freight states:** Increasing warehouse capacity and logistics automation in Illinois, Texas, and California can improve efficiency.
4. **Adopt low-emission trucking fleets:** Electrification and alternative fuels for road transport can mitigate environmental impacts.
5. **Improve seasonal logistics planning:** Implementing predictive analytics can help freight companies prepare for peak seasons more effectively.

**9. Communicating Results & Next Steps**

* **Deliverables:**
  + A complete **dashboard visualization** showcasing freight trends.
  + A **presentation file** summarizing key findings and action points.
  + A **Python script** containing data cleaning, analysis, and visualizations.
  + A **report document** (this file) for decision-makers.
* **Future Research Areas:**
  + Incorporate real-time GPS tracking for congestion analysis.
  + Study the impact of automation and AI on freight transportation.
  + Explore the economic effects of trade policies on freight movement.