**TRANSBORDER FREIGHT DATA ANALYSIS**

– omarieben7@gmail.com

### Introduction

This document outlines the analysis conducted on the TransBorder Freight data provided by the Bureau of Transportation Statistics (BTS).

The primary objectives were to:

* uncover freight movement patterns,
* identify inefficiencies,
* analyze environmental impacts
* safety and risk assessment
* economic disruptions
* propose actionable recommendations to improve cross-border freight transportation.

### Key Analytical Questions

1. **What are the top modes of transportation by freight volume and value?**
2. **How does freight movement vary across states and countries?**
3. **What is the trend of freight charges over time?**
4. **Which states or countries contribute the most to freight weight/value?**
5. **What patterns exist between shipment weight and charges?**
6. **What are the environmental implications of freight transportation?**
7. **How can inefficiencies in the freight system be reduced?**

### CRISP-DM Framework

The analysis followed the CRISP-DM methodology, which includes the following stages:

1. Business Understanding

The objectives were defined, and the above analytical questions were formulated to guide the analysis process.

2. Data Understanding

The dataset consisted of 5,242,875 records and 14 variables. Key variables included:

* TRDTYPE: Transaction type.
* USASTATE, MEXSTATE, CANPROV: Regions involved in freight movement.
* VALUE: Value of goods in USD.
* SHIPWT: Shipment weight.
* FREIGHT\_CHARGES: Associated transportation costs.
* MONTH, YEAR: Temporal attributes.

The dataset covered multiple years and regions, with data on road, rail, air, and water transport modes.

Summary statistics revealed:

* **Total Value:** $14.57 trillion USD
* **Total Shipment Weight:** 6.2 trillion kilograms
* **Total Freight Charges:** $191.93 billion USD

#### 3. Data Preparation

* Empty values in columns such as MEXSTATE, CANPROV, DF were identified and addressed.
* The data was cleaned and aggregated to provide meaningful insights at state, country, and mode-of-transport levels.
* Converting categorical and numerical variables to appropriate data types.