**📦 Supply Chain Analytics Project – Data Overview & Exploration**

**Overview of the Project:**

The "**Dynamic Supply Chain Logistics Analytics**" project aims to **analyze** and **optimize** **supply chain operations using a large-scale logistics dataset**. This dataset reflects real-world supply chain activities including production, shipping, supplier performance, inventory, and customer behavior. The goal is to use **Python** and **MySQL** to extract, clean, and analyze data, then communicate findings via **dashboards** and **reporting**.

This project is conducted as part of the Data Analysis Diploma's final requirement and is meant to simulate a real freelance or industry project scenario where business decisions depend on data-driven insights.

**Problem Statement:**

Supply chain systems are **increasingly complex**, with issues such as **delayed shipments**, **inefficient** **supplier management**, **poor inventory forecasting**, and **rising logistics costs**. Without a data-driven approach, businesses struggle to understand **where inefficiencies lie**.

**Project Goals:**

* Identify key performance trends in logistics and inventory data.
* Analyze supplier performance and risk.
* Forecast demand and shipping delays using historical data.
* Build actionable dashboards for decision-makers.
* Gain hands-on experience in end-to-end data projects using MySQL and Python.

**🧾 Dataset Columns:**

|  |  |
| --- | --- |
| Column Name | Description |
| timestamp | The date and time at which a logistics event or data point was recorded. This acts like the heartbeat of the entire system, allowing us to track events over time. |
| vehicle\_gps\_latitude | The real-time latitude location of the vehicle involved in delivery or shipping. Used for tracking movements. |
| vehicle\_gps\_longitude | The corresponding longitude for the GPS coordinate — helps create a full map of the journey. |
| fuel\_consumption\_rate | Measures how much fuel the vehicle is using. Useful for cost and sustainability analysis. |
| eta\_variation\_hours | The difference between actual and expected arrival time, measured in hours. Key to evaluating delivery performance. |
| traffic\_congestion\_level | Indicates how bad traffic was during the delivery — essential for route optimization. |
| warehouse\_inventory\_level | It shows how fully the warehouse was at that time — critical for demand planning and restocking decisions. |
| loading\_unloading\_time | Time spent loading or unloading goods. High times might point to inefficiencies at the dock. |
| handling\_equipment\_availability | Whether tools like forklifts and cranes were available — important for warehouse efficiency. |
| order\_fulfillment\_status | Whether a customer order was fulfilled on time or not. Affects satisfaction and KPIs. |
| weather\_condition\_severity | Describes how extreme the weather was — storms, snow, rain, etc. This influences safety and delays. |
| port\_congestion\_level | How busy the port was — a major factor in global shipping delays. |
| shipping\_costs | The total cost of moving goods from origin to destination — critical for financial optimization. |
| supplier\_reliability\_score | Rates how dependable a supplier is based on past performance — higher = more trustworthy. |
| lead\_time\_days | Time in days between placing and receiving an order. Vital for supply chain planning. |
| historical\_demand | How many units were previously requested — helps with forecasting future needs. |
| iot\_temperature | IoT sensor data reporting the temperature during shipping — essential for perishable goods. |
| cargo\_condition\_status | A status update on whether the cargo arrived in good condition or not. |
| route\_risk\_level | How risky the transportation route was, considering weather, traffic, theft, etc. |
| customs\_clearance\_time | Time spent waiting for customs approval. Bottlenecks here can slow down international shipments. |
| driver\_behavior\_score | Score based on the driver's performance — speed, braking, safety compliance, etc. |
| fatigue\_monitoring\_score | IoT-driven estimate of driver tiredness — higher fatigue increases risk. |
| disruption\_likelihood\_score | Predicts the chances of a supply chain disruption. Preventative planning relies on this. |
| delay\_probability | Probability that a shipment will be delayed based on all real-time and historical factors. |
| risk\_classification | Categorize the shipment or route into risk levels (e.g., Low, Medium, High). |
| delivery\_time\_deviation | The actual difference between scheduled and actual delivery time. Like ETA variation, but from a scheduling standpoint. |

**Task Assignment & Roles**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Team Member | Main Task | % | Secondary task | % |
| Ziad Khaled Masoud | Build Data Model, Data Cleaning and Preprocessing | **100%** | Visualization Dashboard and Final Presentation | **Not started** |
| Abdelrhman Ahmed | Analysis Questions Phase | **100%** | Visualization Dashboard and Final Presentation | **Not started** |
| Sama Mahmoud | Analysis Questions Phase | **100%** | Forecasting Questions Phase | **Pending** |
| Seif El-Deen Ahmed | Forecasting Questions Phase | **In progress** | Analysis Questions Phase | **Pending** |
| Mohamed Hymeda | Forecasting Questions Phase | **85%** | Analysis Questions Phase | **Not started** |
| Reneh Romany | Visualization Dashboard and Final Presentation | **Not**  **started** | Build Data Model, Data Cleaning and Preprocessing | **100%** |
| Mohamed Rafaat | Visualization Dashboard and Final Presentation | **Not started** | Forecasting Questions Phase | **Pending** |