Tab 1

**Freight Analytics POC**

The dataset you provided, freight.csv, contains information about various freight shipments, likely from a logistics or transportation company.

Description of the dataset

Imagine it's a detailed logbook for packages being shipped. Each row in this "logbook" represents a single shipment and records various details about it. For example, it tells you:

* **When** the shipment happened (Year, Month, DayofMonth, DayOfWeek).
* **When it was supposed to leave** (Planned\_Shipment\_Time) and **when it actually left** (Actual\_Shipment\_Time).
* **When it was supposed to arrive** (Planned\_Delivery\_Time).
* **Which company** (carrier) transported it (Carrier\_Name, Carrier\_Num).
* **How long it was supposed to take** (Planned\_TimeofTravel).
* **Whether it was delayed** and by **how much** (Shipment\_Delay).
* **If it arrived on time or late** (Delivery\_Status, where 0 means on-time and 1 means delayed).
* **Where it started** (Source) and **where it was going** (Destination).
* **How far** it traveled (Distance).

**Data Analysis Done using Python**

The data analysis performed using Python on this dataset focused on understanding two key aspects of the freight operations: Shipment Volume and Delay Patterns.

Here's how the analysis was done:

1. **Data Preparation:** Before any analysis, the dataset was cleaned. This involved handling any missing information in crucial columns like Shipment\_Delay or Delivery\_Status to ensure that calculations were accurate.
2. **Shipment Volume Analysis:**
   * **What was analyzed:** We looked at how many shipments there were in total. We then broke this down to see:
     + Which **carriers** transported the most packages.
     + What were the **busiest routes** (from specific origins to specific destinations).
     + How shipment volumes varied across different **days of the week** (e.g., are Tuesdays busier than Saturdays?).
     + How shipment volumes changed from **month to month**, which can reveal seasonal trends.
   * **Why it's useful:** This analysis helps understand operational capacity, identify popular routes, and forecast demand.
3. **Delay Pattern Analysis:**
   * **What was analyzed:** This part focused on understanding why and when shipments were delayed. We calculated:
     + The **average delay** across all shipments.
     + The overall **on-time delivery rate** (what percentage of shipments arrived on schedule).
     + The **average delay for each carrier**, to see which carriers are most punctual or prone to delays.
     + The **average delay for specific routes**, to identify problematic lanes.
     + How delays varied by **day of the week** and **month**, to pinpoint times when delays are more common.

Note :- The original Data set had 3.6 Million records, however for this POC, the analysis was conducted on 905,300 records.

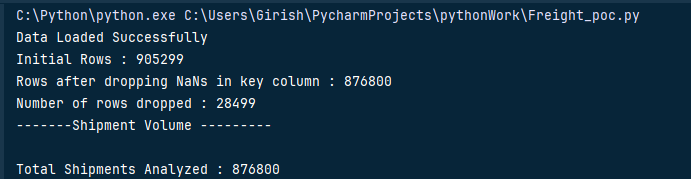
How Python helped :-

Python, with its powerful libraries like **Pandas** and **Matplotlib/Seaborn**, allowed us to:

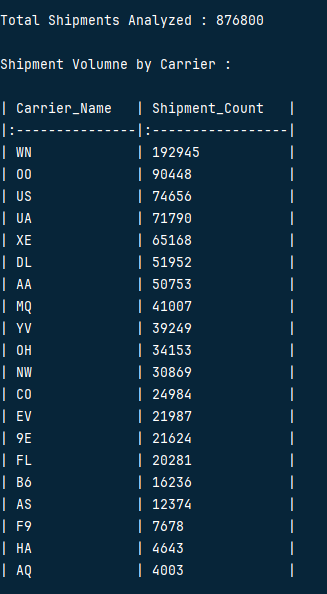
* Efficiently load and clean the large dataset.
* Perform complex calculations and aggregations (like averaging delays or counting shipments by category).
* Create clear visual summaries (like bar charts and pie charts) of the findings, making the insights easy to understand at a glance.

**Snapshots of Analysis**

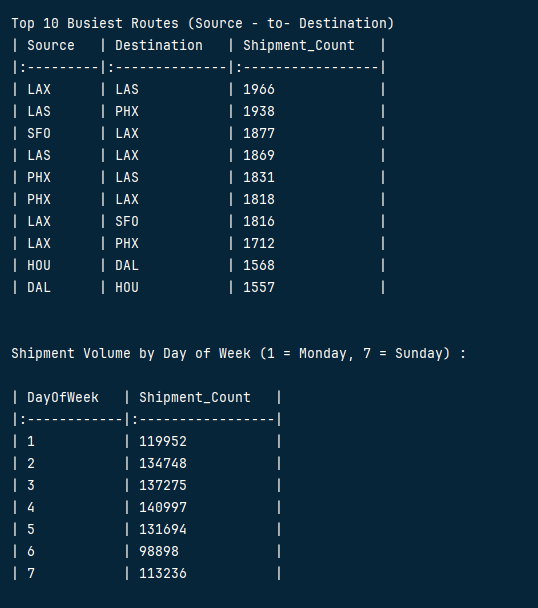
Initial snapshot after Data has been Loaded.



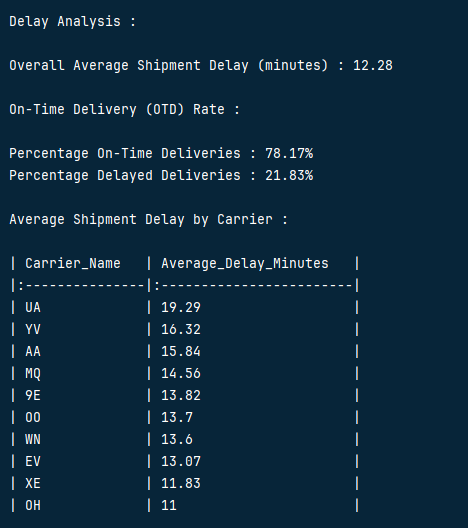
***Volume by Carrier***

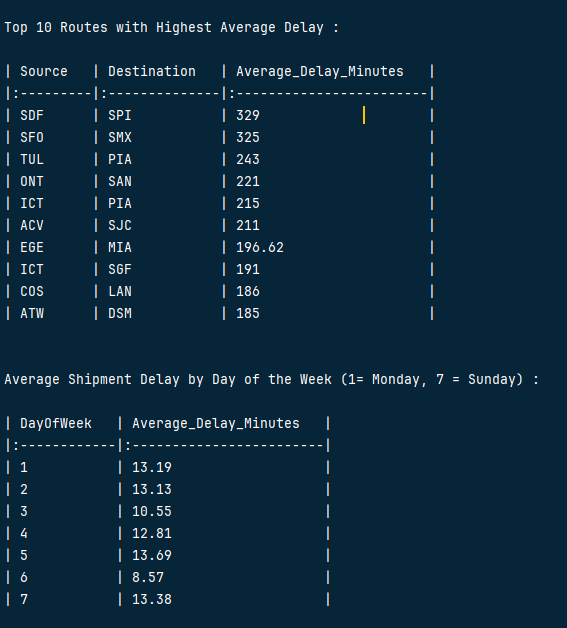


***Busiest Routes and Shipment Volume by Week***



***Delay Analysis***





***Graphical Representation of Analysis***

