Business Challenge:

Delivery failures and delays are one of the biggest drivers of customer dissatisfaction and revenue leakage in logistics. While current systems can report on how many deliveries failed, they provide little clarity on **why** they failed. Operations managers must manually investigate across siloed systems — order logs, fleet reports, warehouse records, and customer complaints — making the process reactive, time-consuming, and error-prone.

Problem in Detail:

- **Order & Shipment Data** shows timestamps of pickup, transit, and delivery but rarely connects them to external conditions.
- **Fleet & Driver Logs** capture GPS traces and driver notes, but this unstructured information is not analyzed systematically.
- **Warehouse Data** records stockouts and delays in order preparation, yet is not linked to downstream delivery performance.
- **Customer Feedback** highlights dissatisfaction (e.g., "my order was late" or "the driver couldn't find my address"), but is unstructured and difficult to aggregate.
- **Contextual Data** such as traffic and weather, which often explain delays, is not integrated into root cause reporting.

This fragmented view prevents the business from identifying systemic issues, such as recurring bottlenecks at specific warehouses, or delays correlated with particular delivery windows.

Strategic Need:

The company needs a system that can:

1. **Aggregate Multi-Domain Data** — orders, fleet logs, warehouse dispatch times, external conditions, and customer complaints.
2. **Correlate Events Automatically** — e.g., link traffic spikes with late deliveries or stockouts with order cancellations.
3. **Generate Human-Readable Insights** — provide narrative explanations instead of raw dashboards.
4. **Surface Actionable Recommendations** — suggest operational changes (e.g., rescheduling, staffing adjustments, address verification).

Sample Data Set:

third-assignment-sample-data-set

Expected Output:

- A word document with write up about how this problem can be solved. A simple diagram can be added to explain better.
- A simple sample program which can aggregate the data and demo the outcome. Need not to be a UI or full fledge backend service.
- Demo this working program from your local system with example use cases. Record a video with your voice explaining the demo.
- Record output for sample use cases and save into a doc and share.
- Email the output details: A GitHub repo link, a folder link containing the doc and recording.

Sample use cases:

1. Why were deliveries delayed in city X yesterday?
2. Why did Client X's orders fail in the past week?
3. Explain the top reasons for delivery failures linked to Warehouse B in August?
4. Compare delivery failure causes between City A and City B last month?
5. What are the likely causes of delivery failures during the festival period, and how should we prepare?
6. If we onboard Client Y with ~20,000 extra monthly orders, what new failure risks should we expect and how do we mitigate them?

Points to consider:

- In case you need any support for any API key or resource, message/email me.
- The maximum spend one person can do for this assignment is $5.