

# Technical Documentation: Supply Chain Risk Architecture

This document details the end-to-end pipeline, from local data ingestion to statistical modeling, SQL engineering, and interactive visualization.

---

## 1. Data Prep & Statistical Modeling (Python)

**Environment:** Local Laptop (Jupyter/VS Code) using `os` for local file management.

**Status:** The dataset was pre-cleaned in Excel; Python was used strictly for advanced analytics.

- **Inquiry & Ingestion:**
    - Used the `os` library to navigate local directories and ingest the cleaned Kaggle dataset from the laptop's file system.
    - No further data cleaning (trimming, duplicate checks, or null handling) was performed in Python, as the data was validated as clean during the Excel phase.
  - **Statistical Risk Model:**
    - Developed a **Volatility Index** based on the Standard Deviation of `Actual_Lead_Time`.
    - **Risk Grading:** Automatically categorized suppliers into **High, Medium, and Low** risk tiers based on their volatility scores. This transformed raw logistics performance into a qualitative risk metric.
- 

## 2. Data Engineering & Table Refinement (SQL/BigQuery)

Data was uploaded to Google BigQuery to structure it for a professional analytics environment.

### The Three-Step Transformation:

1. **Dim\_Suppliers:** A dimension table containing unique supplier attributes, including the Python-generated `Volatility_Index` and `Risk_Grade`.
2. **Fact\_Orders:** A transactional table containing order IDs, values, shipping modes, and lead times.
3. **Executive Summary View:** A specific SQL view created to aggregate total spend and Value at Risk (VaR) by category.

### The Purpose of the SQL Executive Summary View:

Even though this view was not imported into Power BI, it served a critical role in the **Data Validation** phase. It allowed for high-level "sanity checks" in SQL to ensure that the Python calculations were correctly joined and aggregated before moving into the visualization stage. It acts as a permanent, server-side record of risk metrics for non-Power BI users (e.g., those using SQL or Looker Studio).

### 3. Analytics & Visualization (Power BI)



**Data Connection:** Due to BigQuery Sandbox limitations (lack of automated refresh/connector access), Dim\_Suppliers and Fact\_Orders were exported from SQL and manually uploaded to Power BI as CSV/Excel files.

#### Data Modeling & Time Intelligence

- **Date Table:** Built a custom **Date Table** using DAX and used the "**Mark as Date Table**" feature. This allows for accurate time-intelligence calculations (Quarterly and Monthly trends) through a One-to-Many relationship with Fact\_Orders.
- **The "No-Summary-Table" Strategy:** The Executive Summary View was intentionally left out of Power BI. By using raw Fact and Dim tables, we enabled **Full Interactivity**.

#### Core DAX Measures (Dynamic Logic)

To replace the static SQL view, the following DAX measures were created to ensure the dashboard responds to every click:

- **Total Portfolio Value:** `SUM(Fact_Orders[Order_Value_USD])`
  - **Value at Risk (VaR):** `CALCULATE([Total Portfolio Value], Dim_Suppliers[Risk_Grade] = "High")`
  - **Risk Exposure %:** `DIVIDE([Value at Risk], [Total Portfolio Value], 0)`
  - **Risk Status:** A dynamic text indicator that categorizes the "Health" of a selection (e.g., "  CRITICAL EXPOSURE" or "  STABLE") based on the current Risk Exposure %.
- 

#### Summary of Process Flow

1. **Excel:** Initial data cleaning and validation.
  2. **Python:** Statistical volatility modeling using local files (os library).
  3. **BigQuery SQL:** Normalization into Fact/Dim tables and creation of a validation Executive Summary View.
  4. **Power BI:** Manual data upload, Date Table creation, and DAX engineering for a fully interactive, filter-responsive risk dashboard.
- 

#### Final Project Conclusion

This architecture proves that even with tool limitations (Sandbox constraints), a robust data model can be built. By utilizing DAX instead of static summary tables, the dashboard provides a "drill-down" experience where users can analyze risk at the category, shipping mode, or monthly level.