

Executive Summary

This project presents a comprehensive analysis of production risk and logistics performance based on synthetic operational data from a pharmaceutical company.

The objective was to design and implement a complete end-to-end workflow — from ETL and data modeling to visualization — to support executive and operational decision-making.

The final deliverable is a set of interactive Tableau dashboards that enable stakeholders to identify:

- Critical production batches
- Economic impact associated with operational risk
- Delivery performance by region
- Logistics delay patterns

Data & Model Design

The data was structured following a **star schema**, optimized for business intelligence and visualization tools:

- **fact_batches**: production and quality metrics at batch level
- **fact_shipments**: delivery performance and logistics delay metrics
- **dim_plants**: descriptive information about manufacturing plants and regions

Logical relationships were defined to allow **flexible aggregation**, avoiding rigid joins and improving performance and usability within Tableau.

Key KPIs

Production & Risk

- % Critical Batches
- Aggregated Batch Risk Score
- Average Reject Rate
- Cost at Risk

Logistics

- % On-Time Shipments
- Average Delivery Delay by Region
- Shipment Delay Distribution

Key thresholds and metrics were designed to be **dynamically adjustable** through filters, without requiring changes to the ETL process.

Delivered Dashboards

Executive Overview

High-level view for stakeholders:

- Core production and logistics KPIs
- Time-based trends
- Plant ranking by risk level

Production & Quality Analysis

Operational-focused dashboard:

- Batch-level detail
- Reject rate by plant
- Production risk distribution
- Interactive filters for rapid exploration

Logistics Performance

Delivery and distribution analysis:

- % On-Time Shipments
- Average delivery delays by region
- Shipment delay distribution
- Identification of critical regions

Key Insights

- More than 69% of production batches present some level of criticality, resulting in a significant economic impact.
- Only 52.8% of shipments are delivered on time, with certain regions concentrating the majority of delays.
- Operational risk is not evenly distributed; a small number of plants and regions account for most of the impact, highlighting clear opportunities for optimization.

Key Technical Decisions

- Use of clean, normalized CSV datasets to ensure portability and reproducibility.
- Clear separation between:
 - metrics calculated during ETL
 - dynamic metrics calculated in Tableau
- Implementation of filters and action-based interactivity to support exploratory analysis, avoiding static dashboards.

Next Steps (Future Extensions)

- Automate data ingestion pipelines.
- Incorporate forecasting models for risk and delivery delays.
- Define alerting mechanisms based on critical thresholds.
- Integrate logistics costs and scenario simulations.

Technologies Used

- Python (ETL and data preparation)
- Tableau Desktop / Tableau Public
- GitHub (version control and documentation)

Final Note

This project was developed using an end-to-end analytical approach, replicating a real-world business intelligence scenario — from raw data to dashboards ready for decision-making.