

Comprehensive Project Report: Supply Chain Performance & Bottleneck Analysis

Project Overview & Goal

This project focused on conducting a comprehensive **Supply Chain Performance and Bottleneck Analysis** to identify inefficiencies, assess quality, and optimize operational processes within a simulated product supply chain. The primary goal was to transform raw supply chain data into actionable insights, enabling data-driven decision-making for improved efficiency and reduced costs. This project demonstrates an end-to-end analytical workflow, from data preparation and engineering to advanced visualization and strategic recommendations.

Data Acquisition & Augmentation

The project began with a small, initial dataset downloaded from Kaggle in a CSV format (supply_chain_data.csv), containing 100 rows of various supply chain metrics. To ensure the analysis was robust and representative of real-world scenarios in a large company, I decided that it was crucial to work with a larger volume of data. I used Gemini to help me create a python script to run in VS code to generate 10,000 more rows of similar data consistent with that of the original dataset.

Data Augmentation using Python

This script analyzed the patterns, ranges, and categorical distributions within the original 100 rows and then generated 10,000 additional, random rows that adhered to these learned characteristics. This process resulted in a significantly expanded dataset of 10,100 rows (LargerSupplyChainData.csv) therefore providing a more realistic foundation for analysis.

Data Preparation & Feature Engineering in MySQL Workbench

The augmented dataset was then loaded into MySQL Workbench. I then started the data preparation and feature engineering process which was performed directly within the database environment using SQL queries.

Key SQL Transformations & Engineered Features:

1. Key SQL Transformations & Engineered Features:

1. **Total_Cost and Gross_Margin Calculation:** New columns were added to the table. Total_Cost was calculated as the sum of Manufacturing costs and Shipping costs. Gross_Margin was calculated using the formula $(\text{Revenue generated} - \text{Total_Cost}) / \text{Revenue generated} * 100$, providing a key profitability metric.
2. **Total_Lead_Time:** A new column was created by summing Lead time, Manufacturing lead time, and Shipping times to represent the overall duration of the supply process.
3. **Cost_Category Binning:** A Cost_Category column was created to segment Total_Cost into categorical ranges like 'Low-Cost', 'Mid-Range', and 'High-Cost', enabling easier analysis and visualization.
4. **Rounding Numerical Values:** Many of the Columns such as Gross_Margin and Shipping costs were rounded to two decimal places using the ROUND() function for improved readability and presentation in the dashboard.

Data Loading & Transformation in Qlik Sense

The cleaned and engineered dataset from MySQL was then loaded into Qlik Sense. A crucial transformation was performed directly within the Qlik Sense Data Load Editor to prepare the data for bottleneck analysis.

Unpivoting Time Metrics using CROSSTABLE

To effectively visualize the average time for each stage of the supply chain (Lead times, Manufacturing lead time, Shipping times) in a single bar chart, the data needed to be "unpivoted." The CROSSTABLE function was used for this in the Qlik Sense script. This transformed the columns Lead times, Manufacturing lead time, and Shipping times into two new columns:

- **Metric_Name:** Containing the names of the original time columns (e.g., 'Lead times', 'Manufacturing lead time').
- **Metric_Value:** Containing the corresponding time values.

A new table named BottleneckData was created, which is perfectly structured for a bar chart comparing average times across different stages.

Overcoming Scripting Challenges:

This process involved iterative debugging and diagnosing with gemini and chat GPT to refine the Qlik Sense load script:

- **Initial FROM (csv) errors:** Resolved by using the explicit (txt, utf8, embedded labels, delimiter is ',') format.
- **Unexpected token: 'Set' errors:** Addressed by ensuring the script was placed in a new, clean section of the Data Load Editor, preventing interference from Qlik's auto-generated code.
- **Table not found errors:** Resolved by ensuring the CROSSTABLE function correctly referenced the loaded data, demonstrating robust data loading practices.

Qlik Sense Dashboard Development & Visualizations

The core of the project's output is an interactive Qlik Sense dashboard titled "**Supply Chain Performance & Bottleneck Analysis.**" This dashboard is designed to provide quick, actionable insights into key operational areas, enabling stakeholders to understand performance, identify issues, and drive improvements.

Key Dashboard Components & Insights:

- **Executive Summary KPIs:**
 - **"Estimated Cost of Defect":** Provides a high-level financial impact of quality issues, making the business case for improvements clear.
 - **"Avg Shipping Time (Days)":** Shows the overall efficiency of product delivery.
 - **"Avg Cycle Time Per Order":** Quantifies the end-to-end duration of a typical order, providing a high-level view of process efficiency.
 - **"Percent of Inspections Failed":** A critical quality metric that highlights the proportion of products failing inspection.

These KPI's offer an at-a-glance overview of critical supply chain health metrics for quick executive review.

- **Summary of Key Findings (Text Box):**

This text box acts as a clear executive summary, directly stating the most important insights derived from the data.

- **"Average Lead Time by Supply Chain Stage" (Bar Chart):**

Visually compares the average duration of each key stage in the supply chain process (Manufacturing lead time and Shipping times). Directly identifies the primary bottleneck. In this analysis, **"Manufacturing lead time"** was clearly identified as the longest stage, indicating a critical area for process optimization.

- **"Supplier Performance" (100% Stacked Bar Chart):**

Displays the percentage breakdown of inspection results (Fail, Pass, Pending) for each supplier. Crucially highlights suppliers with a higher *proportion* of failed or pending inspections, enabling targeted supplier management and quality improvement initiatives. This chart is crucial for assessing supplier reliability.

- **"Product Quality Performance" (Bar Chart):**

Compares the average defect rates across different Product type categories (e.g., Cosmetics, Haircare, Skincare). Identifies specific product categories with higher defect rates, prompting investigation into their manufacturing processes, materials, or design. For instance, "Haircare" products were identified as having a higher average defect rate.

- **"Carrier Performance: Shipping Time vs Cost" (Scatter Plot):**

Visualizes the trade-off between average shipping time and average shipping costs for different Shipping carriers. Aids in strategic decision-making for carrier selection, allowing businesses to choose carriers that align with their priorities (e.g., lowest cost, fastest delivery, or a balance).

The entire dashboard is interactive, allowing users to click on any element (e.g., a specific supplier, a product type) to filter all other charts and explore related insights.

Key Insights & Business Recommendations

The comprehensive analysis, as summarized on the dashboard, revealed several critical areas for supply chain optimization:

- **Manufacturing Bottleneck:** The consistently long "Manufacturing lead time" indicates a need for process re-engineering, automation, or resource optimization within the manufacturing stage.
 - **Recommendation:** Investigate manufacturing processes for inefficiencies. This could involve process mapping, Lean Six Sigma methodologies, or technology upgrades to reduce production time.
- **Target Haircare Quality:** The higher defect rates in "Haircare" products, as shown in the "Product Quality Performance" chart, require a focused investigation into their specific production lines, material sourcing, or design.
 - **Recommendation:** Conduct a root cause analysis for haircare product defects. This might involve reviewing raw material quality, manufacturing line issues, or packaging processes specific to haircare.
- **Proactive Supplier Management:** The "Supplier Performance" chart highlights which suppliers have the highest percentages of failed or pending inspections.
 - **Recommendation:** Implement a supplier performance review program. Engage with high-failure-rate suppliers to address quality issues, consider alternative suppliers, or adjust order volumes based on performance. Address "Pending" inspections promptly to avoid potential hidden quality problems.
- **Strategic Carrier Selection:** The cost-time trade-offs among carriers, as visualized in the "Carrier Performance" scatter plot, enable informed decisions for optimizing delivery based on urgency and budget.
 - **Recommendation:** Optimize carrier selection based on specific shipment needs. For urgent deliveries, a faster (potentially more expensive) carrier might be chosen, while for less time-sensitive shipments, a lower-cost carrier could be preferred.

Conclusion

This project leveraged data analytics to provide a holistic view of supply chain performance. By cleaning and engineering features in **MySQL**, augmenting the dataset with **Python**, and building an interactive dashboard in **Qlik Sense**, key bottlenecks and quality issues were identified. This was then translated from complex data into clear, actionable insights and strategic recommendations.