

Supply Chain Performance Analytics – SQL-Driven KPI Insights

1. Project Introduction

Title:

Supply Chain Performance Analytics – SQL-Driven KPI Insights

Objective:

The goal of this project is to analyse key supply chain performance indicators using SQL and Excel. The project uncovers insights into delivery performance, fulfilment efficiency, product profitability, and order priority effectiveness by extracting, transforming, and analysing data. These insights can help a retail company optimize its supply chain operations and make data-driven decisions.

Dataset:

A synthetic Global Superstore dataset (2,500 records) was created to mimic real-world supply chain data. The dataset includes the following fields:

- Order and shipping dates (order_id, order_date, ship_date)
- Shipping details (ship_mode, shipping_cost, order_priority)
- Customer and geographic information (customer_id, segment, country, city, state, region)
- Product information (product_id, category, sub_category, product_name)
- Financial metrics (sales, quantity, discount, profit)
- A precomputed shipping_delay (days between order and ship dates)

The dataset was carefully validated to ensure accuracy, with particular attention to verifying that all date fields were valid and that the precomputed shipping delay matched the expected values.

2. Database and Data Preparation

Database Creation

A MySQL database named **SupplyChainDB** was created. Within this database, the following table was defined:

```
CREATE TABLE Synthetic_Global_Superstore (  
  order_id VARCHAR(50) PRIMARY KEY,  
  order_date DATE,  
  ship_date DATE,
```

```
ship_mode VARCHAR(50),
shipping_cost DECIMAL(10,2),
order_priority VARCHAR(20),
customer_id VARCHAR(50),
segment VARCHAR(50),
country VARCHAR(50),
city VARCHAR(50),
state VARCHAR(50),
region VARCHAR(50),
product_id VARCHAR(50),
category VARCHAR(50),
sub_category VARCHAR(50),
product_name VARCHAR(100),
sales DECIMAL(10,2),
quantity INT,
discount DECIMAL(5,2),
profit DECIMAL(10,2),
shipping_delay INT
);
```

Data Import Process

The dataset was imported using MySQL Workbench's Table Data Import Wizard:

- Mapping Columns: Each column from the CSV file was manually mapped to the corresponding table column. This step was crucial because initial mismatches (e.g., region data showing as product IDs) were resolved by ensuring that the source and destination columns aligned correctly.
- Data Verification: After importing, a sample query was run:

```
SELECT * FROM Synthetic_Global_Superstore LIMIT 10;
```

This confirmed that all records were correctly imported and that key fields (such as order_date, ship_date, and shipping_delay) contained the expected values.

Data Cleaning

- **Validation:** A query verified that no orders had missing date fields:

```
SELECT order_id, order_date, ship_date  
  
FROM Synthetic_Global_Superstore  
  
WHERE order_date IS NULL OR ship_date IS NULL;
```

Since no nulls were found, the dataset was confirmed to be complete.

- **Shipping Delay:** As the CSV already contained a verified shipping_delay, no recalculation was needed. This field was retained to avoid redundancy and to preserve the integrity of the original data.
-

3. SQL Analysis

Four key SQL queries were developed to extract actionable insights:

3.1 Delivery Performance by Region

```
SELECT region,  
  
       AVG(shipping_delay) AS avg_shipping_delay,  
  
       COUNT(order_id) AS total_orders  
  
FROM Synthetic_Global_Superstore  
  
GROUP BY region  
  
ORDER BY avg_shipping_delay;
```

Explanation:

- **Group by Region:** Aggregates data per region.
- **Average Shipping Delay:** Computes mean shipping time (in days) for each region.
- **Total Orders:** Counts the number of orders per region.
- **Ordering:** Sorting by average delay highlights regions with faster and slower performance.

3.2 Fulfillment Efficiency by Shipping Mode

```
SELECT ship_mode,  
       AVG(shipping_delay) AS avg_delivery_time,  
       COUNT(order_id) AS order_count  
FROM Synthetic_Global_Superstore  
GROUP BY ship_mode  
ORDER BY avg_delivery_time;
```

Explanation:

- **Group by Shipping Mode:** Compares performance among different shipping methods.
- **Average Delivery Time:** Indicates the speed of order fulfilment.
- **Order Count:** Reflects the volume of orders per shipping mode.
- **Ordering:** Identifies the most efficient shipping method.

3.3 Product Profitability

```
SELECT product_id,  
       product_name,  
       SUM(sales) AS total_sales,  
       SUM(profit) AS total_profit,  
       (SUM(profit) / SUM(sales)) * 100 AS profit_margin  
FROM Synthetic_Global_Superstore  
GROUP BY product_id, product_name  
ORDER BY total_profit DESC;
```

Explanation:

- **Group by Product:** Aggregates financial data at the product level.
- **Total Profit and Sales:** Sums up key revenue and cost metrics.
- **Profit Margin:** Computes profitability as a percentage.
- **Ordering:** Highlights the most profitable products.

3.4 Delivery Time by Order Priority (Optional KPI)

```
SELECT order_priority,  
       AVG(shipping_delay) AS avg_shipping_delay,  
       COUNT(order_id) AS total_orders  
FROM Synthetic_Global_Superstore  
GROUP BY order_priority  
ORDER BY avg_shipping_delay;
```

Explanation:

- **Group by Order Priority:** Analyses performance based on order urgency.
 - **Average Shipping Delay:** Measures fulfilment speed per priority level.
 - **Total Orders:** Indicates the distribution of orders by priority.
 - **Ordering:** Helps assess if high-priority orders are indeed processed faster.
-

4. Exporting and Visualizing the Data

The query results were exported to CSV files using MySQL Workbench's export functionality. Each query was run, and the resulting dataset was saved as follows:

- **delivery_performance_by_region.csv**
- **fulfillment_efficiency_by_ship_mode.csv**
- **product_profitability.csv**
- **delivery_by_order_priority.csv**

Creating Visualizations in Excel

A separate Dashboard sheet was created in Excel to consolidate the visualizations. Dynamic links were set up by copying charts from their source sheets and pasting them into the Dashboard sheet so that any updates to the underlying data automatically refresh the visualizations.

4.1 Visualization 1: Delivery Performance by Region

- **Chart Type:** Combo Chart
 - **X-Axis:** Regions

- **Primary Y-Axis:** Total Orders (displayed as columns)
- **Secondary Y-Axis:** Average Shipping Delay (displayed as a line)
- **Chart Title:** "Delivery Performance by Region: Total Orders & Average Shipping Delay"

4.2 Visualization 2: Fulfilment Efficiency by Shipping Mode

- **Chart Type:** Either a Combo Chart or Two Separate Charts:
 - **Option A:** A combo chart with shipping mode on the x-axis, where total orders are shown as columns (primary axis) and average delivery time as a line (secondary axis).
 - **Option B:** Two side-by-side charts (one showing order count by shipping mode and the other showing average delivery time by shipping mode).
- **Chart Title:** "Fulfillment Efficiency by Shipping Mode"

4.3 Visualization 3: Product Profitability

- **Chart Type:** Column Chart (filtered to the top 10 products)
 - **X-Axis:** Product Names (filtered to top 10 by total profit)
 - **Primary Y-Axis:** Total Profit
 - **Secondary Y-Axis:** Average Profit Margin (displayed as a line)
- **Filtering:**

In the PivotTable, the top 10 products were filtered by right-clicking on a product name → Filter → Top 10 Items, using the Sum of total_profit.
- **Chart Title:** "Top Products by Profit and Profit Margin"

Note: Profit margin values were kept as whole numbers (e.g., 26.26) with clear labelling, as they already represent percentages.

4.4 Visualization 4: Delivery Time by Order Priority

- **Chart Type:** Separate Visualizations
 - **Chart A (Pie Chart):**
 - **Data:** Order Distribution by Priority
 - **Chart Title:** "Order Distribution by Priority"
 - **Chart B (Bar or Column Chart):**
 - **Data:** Average Shipping Delay by Order Priority
 - **Chart Title:** "Average Shipping Delay by Order Priority"

- **Approach:**
Creating separate charts avoids clutter and ensures that each KPI is communicated clearly.
-

5. Conclusion

Key Insights

- **Delivery Performance:**
Certain regions show faster average shipping delays and higher order volumes, indicating more efficient logistics.
- **Fulfillment Efficiency:**
Specific shipping methods (e.g., "First Class" or "Same Day") may offer a balance between speed and volume, informing potential logistics improvements.
- **Product Profitability:**
The analysis identifies the top products by profit and profit margin, guiding targeted marketing and inventory focus.
- **Order Priority:**
The order distribution and shipping delay analysis by order priority offer insights into whether high-priority orders are fulfilled more quickly than lower-priority ones.

Recommendations

- **Optimize Logistics in Slower Regions:**
Investigate and enhance processes in regions with longer shipping delays.
- **Invest in Efficient Shipping Methods:**
Consider expanding the use of faster shipping modes for high-margin or high-priority orders.
- **Focus on Top-Performing Products:**
Use profitability insights to prioritize inventory and marketing for the most profitable product segments.
- **Improve Priority Handling:**
Ensure that the processing of high-priority orders is optimized for faster fulfillment.

Future Work

- **Integrate Inventory Data:**
Incorporate metrics such as stock levels and stockout rates for a more comprehensive supply chain analysis.
- **Include Supplier Performance:**
Add supplier data to analyze on-time delivery and overall supplier efficiency.

- **Expand Delivery Metrics:**

Track end-to-end delivery times (from order placement to final delivery) for an even more detailed analysis.