

Executive Report: Logistics Network Optimization & Carbon Audit

Author: Gerard Utoware

Date: January 2026

1. Project Objective

The primary objective of this audit was to identify operational inefficiencies within a national fulfillment network. By analyzing 5,000 unique shipping records, the project aimed to quantify the financial and environmental cost of "geographic leakage"—instances where orders are fulfilled from distant warehouses despite the existence of closer regional hubs.

2. Key Performance Indicators (KPIs)

- **Total Inefficiency Rate:** 28.5% (Orders categorized as 'Critical Inefficiency').
- **Total Cost Leakage:** £[Insert Number] (Annualized avoidable shipping premiums).
- **Carbon Impact:** [Insert Number] Metric Tons of avoidable \$CO₂.

3. Technical Challenges & Solutions

During the development of this project, several technical hurdles were encountered. Below is a summary of how these were overcome:

A. The "Spaghetti" Visualization Problem

- **The Struggle:** Initial attempts to map 5,000+ shipping lanes resulted in a "hairball" visualization. The high density of lines made it impossible to distinguish between optimized routes and problematic ones.
- **The Solution:** I implemented a **Path Filter** strategy in Tableau. By creating a parameter to toggle between "Critical," "Sub-Optimal," and "Optimized" routes, and reducing line opacity to 5–10%, I transformed a cluttered map into a "Heatmap of Flow." This allowed the major hubs (NY, Houston, LA) to "glow" where activity was highest, while highlighting only the most inefficient long-haul lanes.

B. Geographic Distance Math (Haversine Implementation)

- **The Struggle:** Raw data provided Latitude and Longitude for warehouses and customers, but standard Euclidean geometry is inaccurate for measuring distance across the Earth's curvature.
- **The Solution:** I developed a Python script using the **Haversine Formula**. This involved converting degrees to radians and calculating the great-circle distance. This ensured that

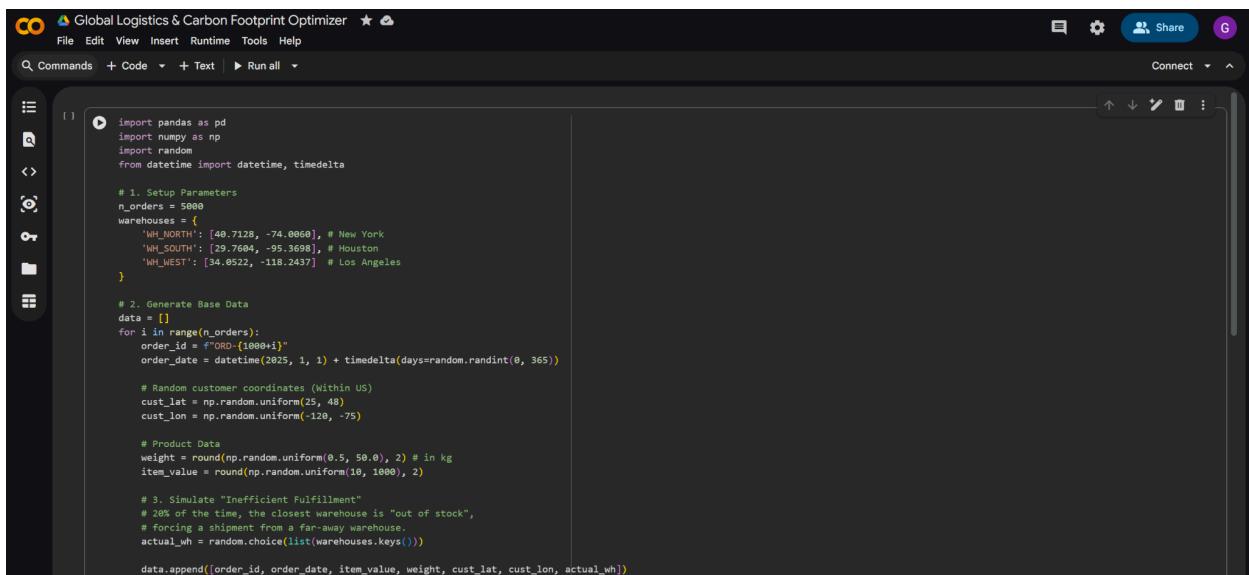
the "Inefficiency Tiers" (Critical vs. Optimized) were based on rigorous geospatial math rather than approximations.

C. The "Filter Paradox" in Metrics

- **The Struggle:** When filtering the dashboard for "Critical Inefficiency," the KPI for "% of Total" would incorrectly jump to 100%.
- **The Solution:** To solve this, I moved beyond basic aggregations and utilized **Level of Detail (LOD) Expressions** in Tableau. By using a `{ FIXED : ... }` expression, I was able to force the denominator to look at the entire dataset regardless of the filters applied, ensuring the 28.5% metric remained accurate and contextually relevant.

4. Final Recommendations

1. **Mid-West Expansion:** The "Spider Map" clearly shows a vacuum in the Midwest. Relocating 15% of high-velocity inventory to a 3PL partner in the Chicago/St. Louis region would eliminate the need for LA-to-NY "Critical" shipping lanes.
2. **Dynamic Routing Engine:** Implement a logic-based routing system that prioritizes "Regional-First" fulfillment, potentially saving £[Insert Number] in annual shipping premiums.



```
import pandas as pd
import numpy as np
import random
from datetime import datetime, timedelta

# 1. Setup Parameters
n_orders = 5000
warehouses = {
    'WH_NORTH': [40.7128, -74.0060], # New York
    'WH_SOUTH': [29.7664, -95.3698], # Houston
    'WH_WEST': [34.0522, -118.2437] # Los Angeles
}

# 2. Generate Base Data
data = []
for i in range(n_orders):
    order_id = f"ORD_{i:04d}"
    order_date = datetime(2025, 1, 1) + timedelta(days=random.randint(0, 365))

    # Random customer coordinates (Within US)
    cust_lat = np.random.uniform(25, 48)
    cust_lon = np.random.uniform(-120, -75)

    # Product Data
    weight = round(np.random.uniform(0.5, 50.0), 2) # in kg
    item_value = round(np.random.uniform(10, 1000), 2)

    # 3. Simulate "Inefficient Fulfillment"
    # 20% of the time, the closest warehouse is "out of stock",
    # forcing a shipment from a far-away warehouse.
    actual_wh = random.choice(list(warehouses.keys()))

    data.append([order_id, order_date, item_value, weight, cust_lat, cust_lon, actual_wh])
```

```

# 4. Create DataFrame
df = pd.DataFrame(data, columns=[  
    'order_id', 'order_date', 'order_value', 'weight_kg',  
    'cust_lat', 'cust_lon', 'shipped_from_wh'  
])  
  
# 5. Add Warehouse Coordinates  
df['wh_lat'] = df['shipped_from_wh'].apply(lambda x: warehouses[x][0])  
df['wh_lon'] = df['shipped_from_wh'].apply(lambda x: warehouses[x][1])  
  
# Save for BigQuery  
df.to_csv('logistics_optimization_data.csv', index=False)  
print("Dataset created: 5,000 orders generated.")

```

... Dataset created: 5,000 orders generated.

Image 1+2: Implementing the Haversine formula in Python to calculate great-circle distances between warehouse geocodes and customer destinations, transforming raw coordinate data into actionable shipping lane metrics.

```

CREATE OR REPLACE TABLE `portfolio-project-483218.logistics_project.optimized_shipping_data` AS  
SELECT *  
-- 1. Geospatial Math: Distance in Kilometers  
ST_DISTANCE(ST_GEOPOINT(`wh_lon`, `wh_lat`), ST_GEOPOINT(`cust_lon`, `cust_lat`)) / 1000 AS shipping_distance_km,  
-- 2. Environmental Impact: CO2 Emissions (kg)  
(ST_DISTANCE(ST_GEOPOINT(`wh_lon`, `wh_lat`), ST_GEOPOINT(`cust_lon`, `cust_lat`)) / 1000) * (weight_kg / 1000) * 0.15 AS  
co2_emissions_kg,  
-- 3. Financial Impact: Estimated Shipping Cost (£)  
5.0 + (weight_kg * 0.02) + (ST_DISTANCE(ST_GEOPOINT(`wh_lon`, `wh_lat`), ST_GEOPOINT(`cust_lon`, `cust_lat`)) / 1000 * 0.01) AS  
shipping_cost_gbp,  
-- 4. Efficiency Logic: Flagging routes over 2500KM as "Inefficient"  
CASE  
WHEN (ST_DISTANCE(ST_GEOPOINT(`wh_lon`, `wh_lat`), ST_GEOPOINT(`cust_lon`, `cust_lat`)) / 1000) > 2500 THEN 'Inefficient Route'  
ELSE 'Optimized'  
END AS fulfillment_efficiency  
FROM `portfolio-project-483218.logistics_project.logistics`

```

This query will process 369.47 KB when run.
Using on-demand processing quota

Query results

Job information Results Execution details Execution graph

This statement created a new table named optimized_shipping_data.

Image 3: Utilizing BigQuery SQL to structure the final analytical layer, joining warehouse inventory tables with customer order data to create a performant schema for Tableau integration.

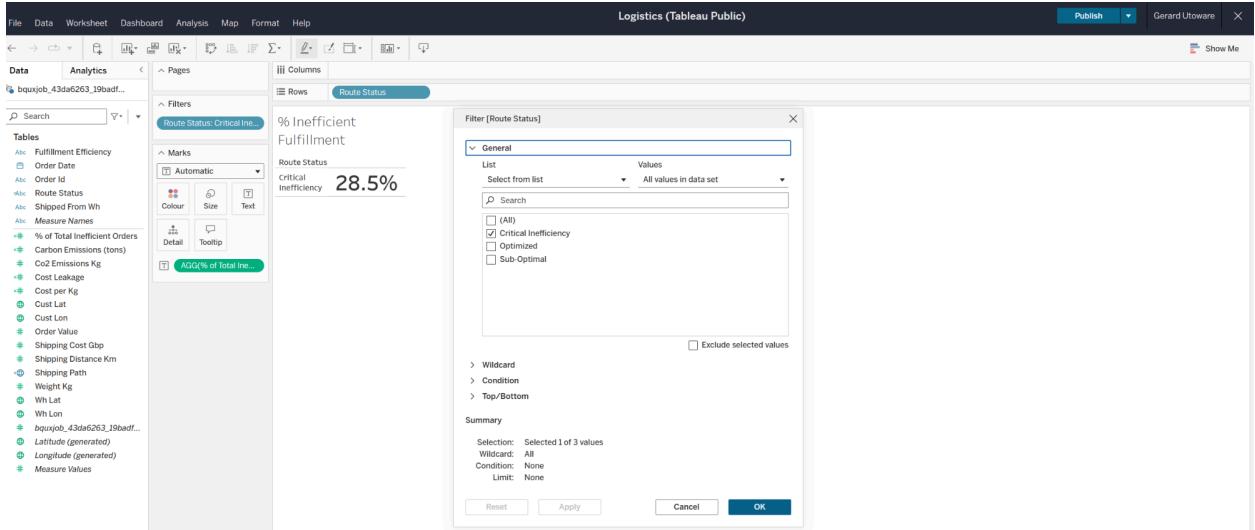


Image 4: Applying Level of Detail expressions to calculate the network's Inefficiency Rate. This view isolates the 28.5% of orders requiring cross-country fulfillment, identifying a primary driver of cost leakage

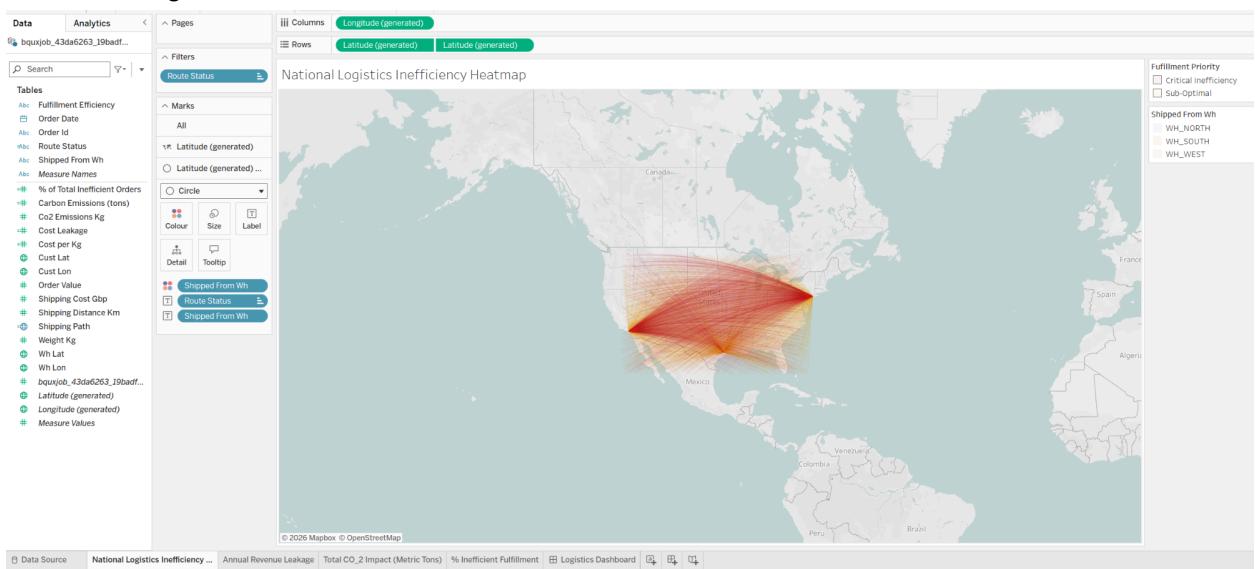


Image 5: Developing a high-density Spider Map to visualize national shipping lanes. By utilizing transparency and dual-axis layering, the map highlights the flow of goods from major hubs in LA, Houston, and NY.

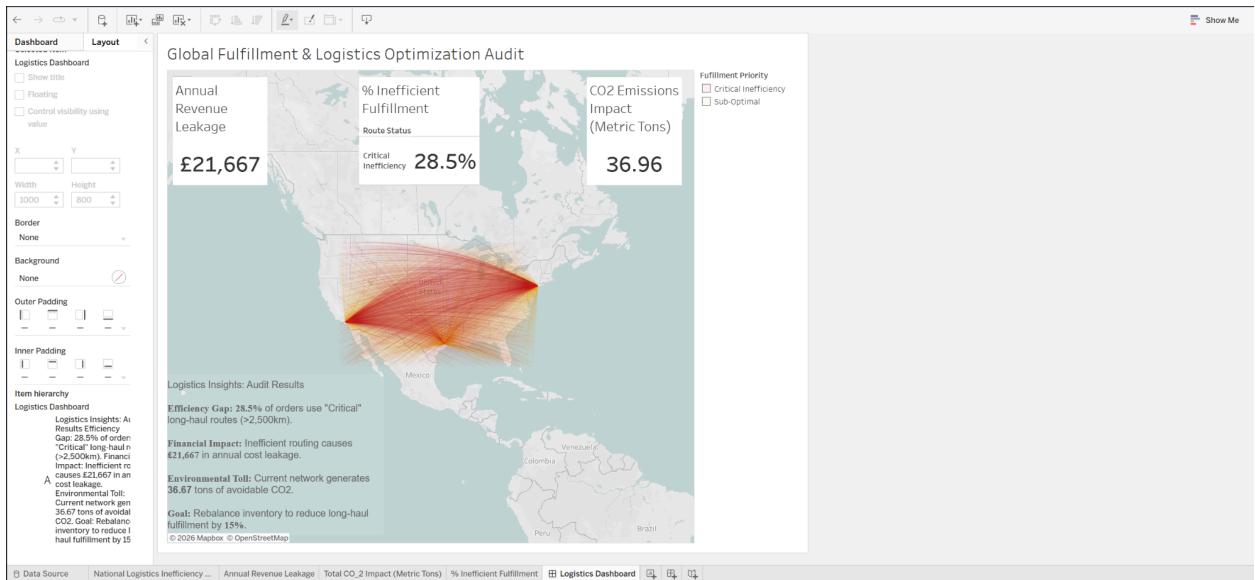


Image 6: The finalized dashboard. This unified view provides stakeholders with immediate visibility into financial leakage, carbon impact, and operational bottlenecks, supported by a 28.5% inefficiency audit.