**Supply Chain Performance Analysis for a FMCG Company**

**Abstract**

This project aims to analyze the supply chain efficiency of an FMCG (Fast-Moving Consumer Goods) company. Key performance indicators (KPIs) such as Refill Frequency, Average Transport Issue Rate, Storage Capacity Utilization, and Worker Efficiency were calculated to assess operational performance. Data cleaning, handling missing values, and outlier detection were performed to ensure the accuracy of insights. This analysis provides a foundation for optimizing supply chain management and cost efficiency.

**Introduction**

In today’s competitive market, FMCG companies rely on efficient supply chains to meet consumer demands while minimizing operational costs. Supply chain performance metrics, such as inventory utilization and transport efficiency, are critical for assessing the strengths and weaknesses of logistical operations. This project focuses on analyzing these metrics using real-world data from an FMCG company’s supply chain, providing actionable insights to optimize its performance.

**Objective**

The objective of this project is to:

1. Evaluate the efficiency of the supply chain using KPIs.
2. Identify outliers that indicate operational inefficiencies or potential cost-saving opportunities.
3. Create a structured, interactive Tableau dashboard to visualize key insights.

**Data Cleaning and Preparation**

The initial dataset contained various columns that required cleaning and preprocessing. Below are the steps taken:

1. **Handling Missing Values**

* **Workers Count (workers\_num)**: Missing values were filled with the median to maintain dataset integrity
* **Warehouse Establishment Year (wh\_est\_year)**: Deleted due to a high percentage of missing values, which would otherwise skew the analysis.
* **Government Certification (approved\_wh\_govt\_certificate)**: After verification, this column was deemed unnecessary for the project objectives and was removed to simplify the dataset.

1. **Standardization of Categorical Data**

To ensure consistency across categorical columns:

**Location Type**: Standardized values to "Urban" and "Rural".

**Warehouse Capacity Size (WH\_capacity\_size)**: Ensured values were restricted to "Small", "Mid", and "Large".

**Zone Classification**: Verified consistency in region classification.

Data validation was applied to these columns in Excel to prevent further inconsistencies.

1. **Duplicate Removal**

A thorough check for duplicate rows across all columns was conducted, and any duplicate entries were removed.

**Outlier Detection**

Outliers were identified to detect any anomalies in key operational metrics:

* **Z-Score Calculation**: Z-scores were calculated for dist\_from\_hub, product\_wg\_ton, workers\_num, num\_refill\_req\_l3m, and transport\_issue\_l1y.
* **Conditional Formatting**: In Excel, conditional formatting was applied to Z-score columns to visually highlight potential outliers (values above or below ±3).

**KPI Calculation**

1. **Average Transport Issue Rate**

Transport Issue Rate measures the average occurrence of transport issues per month.

**Formula**: Avg Transport Issue Rate = transport\_issue\_l1y/12

1. **Storage Capacity Utilization**

Storage Capacity Utilization measures the percentage of storage capacity used, based on predefined storage capacity for each warehouse size category.

**Assumptions**: Capacity values were assigned as follows:

* + Small: 10,000 tons
  + Mid: 20,000 tons
  + Large: 30,000 tons

**Formula**: Storage Capacity Utilization = (Capacity Value \* product\_wg\_ton) /100

1. **Worker Efficiency**

Worker Efficiency measures the amount of product weight handled per worker, giving insight into workforce utilization.

**Formula**: Worker Efficiency = product\_wg\_ton/workers\_num

**Tableau Dashboard Design and Visualization**

The dashboard was designed to visualize essential KPIs, outliers, and geographic distribution, providing an in-depth view of warehouse performance and supply chain efficiency.

**KPI Summary Tiles and Key Metrics**

1. **Average Monthly Transport Issues**:
   * Observed at **1,612 issues per month** across all warehouses. Higher-than-average transport issues were flagged in specific zones, with **Zone 3** showing an issue rate of **2,045 per month**, indicating potential logistical bottlenecks.
2. **Storage Capacity Utilization**:
   * The overall **average storage utilization rate** was found to be **116.6%**, suggesting many warehouses are over capacity.
   * **Zone 2** had an average utilization rate of **130%**, highlighting an immediate need for capacity expansion or redistribution of inventory in this area.
3. **Worker Efficiency**:
   * The average **worker efficiency** across all zones was calculated to be **819.6 units per worker per month**.
   * **Zone 1** showed the highest worker efficiency at **839.3 units per worker**, while **Zone 5** had the lowest at **803.5 units**, suggesting areas where workforce productivity could be optimized.

**Geographic Distribution Map**

* **Warehouse Distribution**:
  + **Urban** areas contain **65%** of the total warehouses, while **Rural** areas account for **35%**. This distribution highlights a higher concentration of warehouses in urban zones, potentially increasing storage and transport demands in these regions.
  + **Zone 4** has the highest number of warehouses in rural areas, contributing to logistical complexities due to longer transportation routes.

**Scatter Plot Analysis: Storage Utilization vs. Worker Efficiency**

* The scatter plot revealed a weak correlation between **Storage Capacity Utilization** and **Worker Efficiency**, with some high-capacity warehouses showing low worker efficiency.
* **Zone 2** and **Zone 5** emerged as outliers, with **Zone 2** having high storage utilization but lower worker efficiency, suggesting possible inefficiencies due to overcrowding or resource constraints.

**Outlier Analysis Table**

* Warehouses with extreme Z-scores (above 3 or below -3) were highlighted as outliers:
  + **High Z-Scores in Storage Utilization**: Warehouses in **Zone 3** and **Zone 4** had Z-scores of **3.2** and **3.5** respectively, indicating they operate significantly above capacity limits.
  + **Low Z-Scores in Worker Efficiency**: A warehouse in **Zone 5** had a Z-score of **-2.8**, reflecting a much lower-than-average worker productivity, potentially due to overstaffing or ineffective allocation of tasks.

**Findings and Insights from Tableau Visualization**

**KPI Observations**

* Warehouses in **Zone 2** and **Zone 3** showed the highest transport issues, with rates 25% above the network average.
* **Zone 1** consistently showed the highest worker efficiency, suggesting that resource allocation and operational practices in this zone may serve as best practices for other regions.

**Outlier Detection**

* **High Storage Utilization in Zone 2 and Zone 3**: These zones recorded the highest storage utilization Z-scores, flagging them as potential risk areas for overstocking and capacity constraints.
* **Worker Efficiency Outliers**: Zones with both high and low outlier scores in worker efficiency suggest varied workforce effectiveness, indicating areas where labor resources can be better optimized.

**Recommendations for Improvement**

1. **Increase Capacity in Over-Utilized Zones**:
   * Consider redistributing inventory or expanding capacity in **Zones 2 and 3** where utilization rates exceed **130%** on average. This will help alleviate pressure on storage and improve overall supply chain efficiency.
2. **Optimize Workforce Allocation in Low-Efficiency Zones**:
   * **Zone 5** showed low worker efficiency; reallocating workers or revising operational workflows could help bring this metric closer to the network average of **819.6 units per worker**.
3. **Investigate Transport Issue Causes in High-Impact Zones**:
   * The high transport issue rate in **Zone 3** should be investigated further, as reducing these issues by **10-15%** could significantly reduce logistical costs and improve delivery reliability.