Data-Driven Warehouse Inventory Optimization & Stock Movement Analysis

Hemant

Table of contents

Project Overview

Efficient inventory management is a critical factor in maintaining smooth warehouse operations while reducing costs and improving service. This project focuses on analyzing and optimizing inventory performance using real-world style datasets, advanced analytics, and business intelligence tools.

The dataset used in this project simulates a complete warehouse ecosystem, including inventory records, sales transactions, purchase orders, supplier performance metrics, warehouse zones, bin locations, batch numbers, and expiry information. With this data, the project aims to uncover inefficiencies, identify risks, and propose data-driven solutions to improve inventory health.

This project replicates real industry challenges, aligns with supply chain and operations analytics roles, and demonstrates the ability to connect domain expertise with data-driven decision-making.

Business Problem

Warehouse operations often struggle with balancing stock availability against cost efficiency. Excess inventory blocks working capital, increases holding costs, and leads to waste, especially for perishable and slow-moving items. At the same time, inadequate stock levels create stockouts, order delays, and dissatisfied customers.

Unreliable supplier performance, inconsistent lead times, poor demand visibility, and lack of real-time insights further complicate decision making. As a result, organizations face higher operational costs, reduced service levels, and inefficiencies across the supply chain.

This project aims to address these challenges by identifying inventory risks, analyzing stock behavior, evaluating supplier reliability, and optimizing reorder strategies-turning operational data into clear, actionable insights for better warehouse management.

Project Objectives

The objective of this project is to enhance warehouse decision-making by leveraging data analytics to uncover inefficiencies, reduce operational risks, and optimize inventory performance. The project focuses on delivering clear, actionable insights that support strategic and day-to-day operations. Key objectives include:

- Identify stock risks such as slow-moving, dead, and near-expiry items.

- Reduce stockouts by analyzing demand patterns and monitoring reorder levels.

- Optimize inventory levels using data-driven reorder strategies.

- Improve supplier performance through lead-time and reliability analysis.

- Lower operational and holding costs by eliminating excess stock and improving inventory rotation.

- Strengthen warehouse efficiency by evaluating zone utilization and storage distribution.

- Enable informed decision-making through interactive dashboards and analytical reporting.

These objectives collectively demonstrate the impact of analytics in improving warehouse operations and ensuring better control over inventory health.

Dataset Description

This project uses a realistic, multi-table dataset that represents core warehouse operations. It includes:

1. **Inventory Data:** SKU details such as category, stock levels, reorder thresholds, bin locations, batch numbers, and expiry dates.

2. **Sales Data:** Daily item-level sales quantities used for demand analysis and inventory turnover.

3. **Purchase Orders Data:** Replenishment records showing order quantities, dates, and supplier references—supporting lead-time and procurement analysis.

4. **Supplier Data:** Supplier details with lead times and delivery performance metrics.

This dataset provides all essential elements required to evaluate stock health, forecast demand, assess supplier reliability, and optimize inventory decisions.

Data Analysis

Q1. Which items are at risk of stockout based on current stock and demand trends?

```
WITH
    last_month_qty_sold AS (
        SELECT
            item_id,
            SUM(sold_qty) AS qty_sold_last_month
        FROM
            sales
        WHERE
```

```sql
            sale_date BETWEEN '2025-06-01' AND '2025-06-30'
        GROUP BY
            item_id
    ),
    current_stock_demand AS (
        SELECT
            inv.item_id,
            inv.item_name,
            inv.current_stock,
            lmqs.qty_sold_last_month,
            ROUND(lmqs.qty_sold_last_month / 30::NUMERIC, 1) AS avg_daily_demand
        FROM
            inventory inv
            JOIN last_month_qty_sold lmqs ON lmqs.item_id = inv.item_id
    )
SELECT
    item_id,
    item_name,
    current_stock,
    qty_sold_last_month,
    avg_daily_demand,
    CASE
        WHEN avg_daily_demand > 0 THEN ROUND(current_stock / avg_daily_demand, 1)
        ELSE NULL
    END AS days_of_stock_left,
    CASE
        WHEN avg_daily_demand = 0 THEN 'No Demand'
        WHEN current_stock / avg_daily_demand <= 10 THEN 'High Risk'
        WHEN current_stock / avg_daily_demand <= 30 THEN 'Medium Risk'
        ELSE 'Low Risk'
    END AS stockout_risk
FROM
    current_stock_demand
ORDER BY
    days_of_stock_left;
```
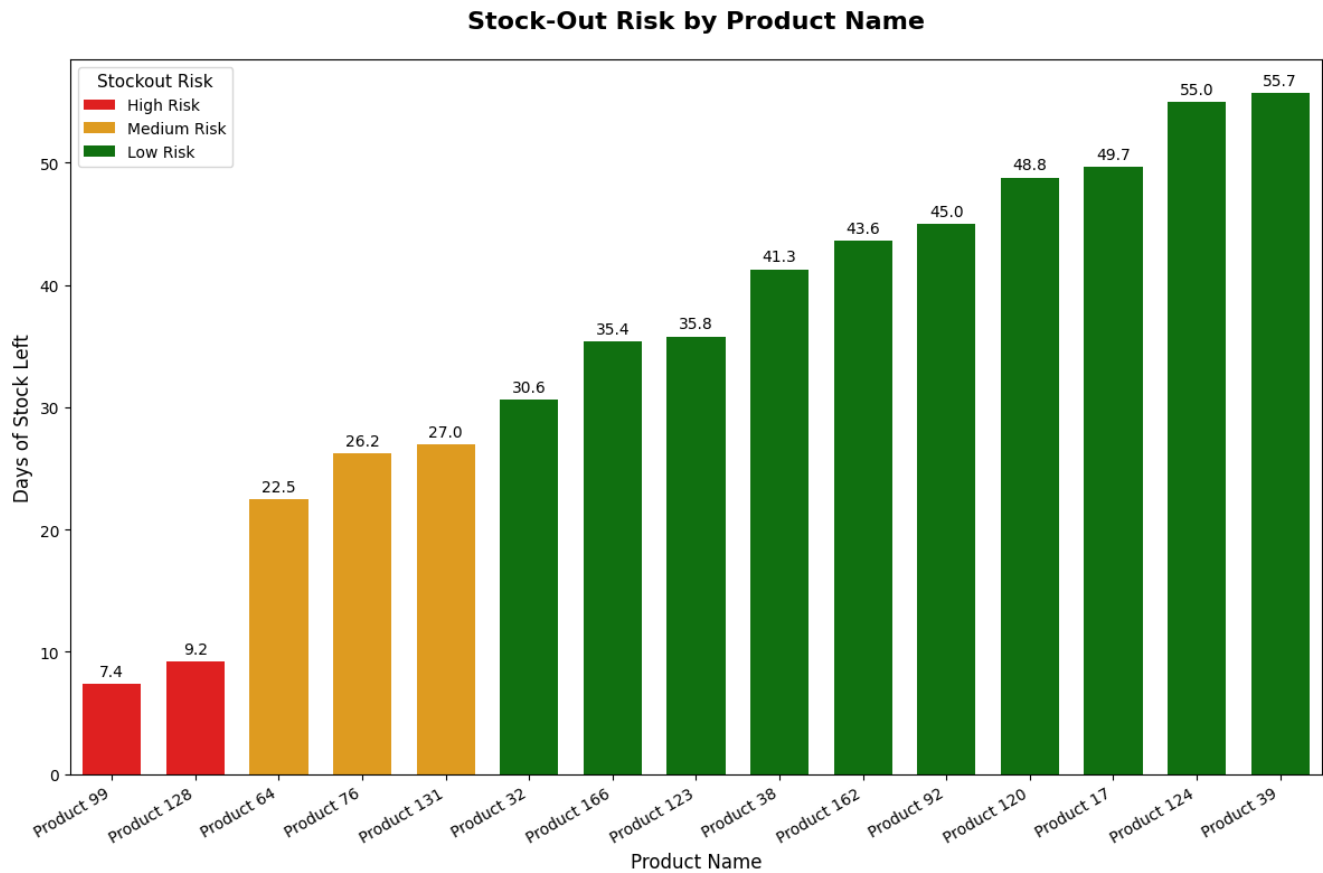
Figure 1: Stock-Out Risk by Product

Insights:

1. Product 99 and Product 128 has less than 10 days of stock. Putting them in high stock out risk.

2. Product 64, Product 131 and Product 76 fall under medium risk, with stock ranging 22–27 days.

3. These SKU's are likely fast-moving or inaccurately forecasted, needs immidiate procurement.

4. Wide variation from 8 to 66 days showing inefficient stock planning and inconsistent stock planing.

Q2. How much inventory value is tied up in slow-moving or dead stock?

```
WITH
    inventory_value AS (
        SELECT
            inv.item_id,
            inv.current_stock * inv.unit_cost AS stock_value,
            MAX(sls.sale_date) AS sale_date
        FROM
            sales sls
```

```
            JOIN inventory inv ON inv.item_id = sls.item_id
        GROUP BY
            inv.item_id
    ),
    inventory_value_status AS (
        SELECT
            item_id,
            stock_value,
            sale_date,
            CASE
                WHEN sale_date <= '2025-03-31' THEN 'slow-stock'
                ELSE 'fast-stock'
            END AS stock_status
        FROM
            inventory_value
    )
SELECT
    stock_status,
    ROUND(SUM(stock_value) / 1000000, 2) AS stock_value_mn
FROM
    inventory_value_status
GROUP BY
    stock_status;
```
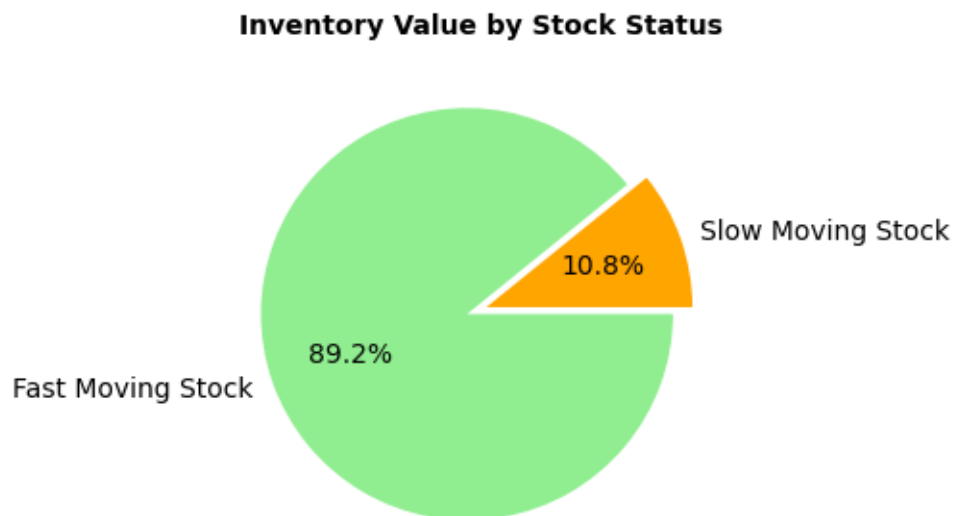

Figure 2: Inventory Value by Stock Status

Insights :

1. 89.2% of stock which is fast moving, could be due to low price, essential products which is used in day to day life or popularity of the brands among the customers.

2. 10.8% stock is identified as slow moving, In this category the products are more likely to be luxury goods or seasonal.

3. In slow moving stock we have to identify why these products stuck at the warehouse and plan inventory accordingly.

Q3. Which SKUs are approaching expiry and require priority action?

```sql
WITH
    product_expiry AS (
        SELECT
            item_id,
            item_name,
            (current_stock * unit_cost) AS inventory_value_m,
            date_expiry,
            (date_expiry - '2025-07-16') AS days_in_expiry
        FROM
            inventory
    )
SELECT
    item_id,
    item_name,
    inventory_value_m,
    date_expiry,
    days_in_expiry,
    CASE
        WHEN days_in_expiry < 0 THEN 'Expired'
        WHEN days_in_expiry <= 30 THEN 'High Risk'
        WHEN days_in_expiry <= 90 THEN 'Medium Risk'
        WHEN days_in_expiry <= 180 THEN 'Low Risk'
        ELSE 'Safe'
    END AS expiry_status
FROM
    product_expiry
ORDER BY
    days_in_expiry;
```
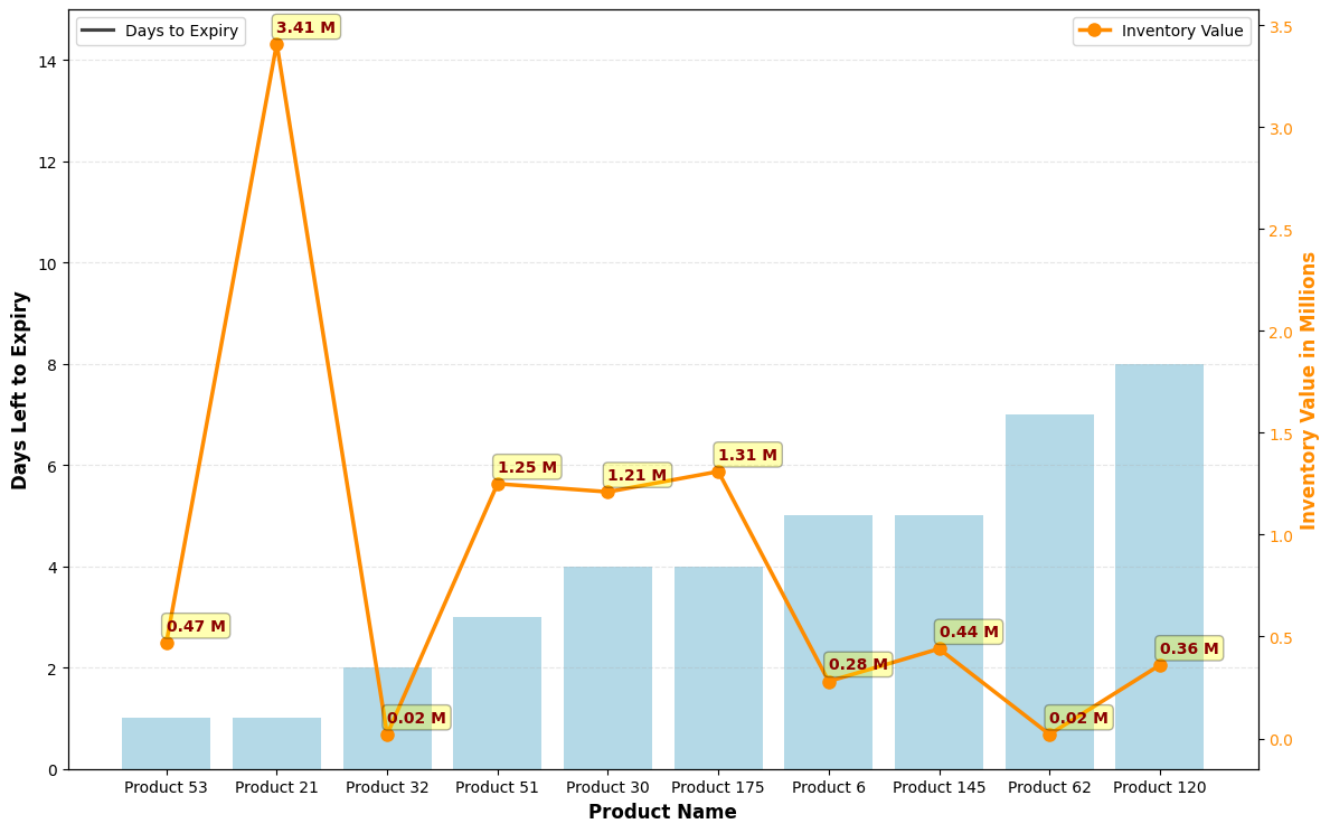
Figure 3: Product Expiry Status

Insights:

1. Product 21, Product 175, Product 30 and Product 51 has very high value 3.41M, 1.31M, 1.25M and 1.21M but less than 4 days left to expiry date, making them highest financial risk.

2. These products have very low remaining shelf life with high cost. Urgent action required like discount offers, bundle promotions, clearance offers or return to vendor.

3. There are some products with low value and about to expire soon. These products do not cause financial risk but shrinking warehouse storage capacity.

Q4. What is the current inventory turnover rate, and which categories perform the best or worst?

```
WITH
    sales_cogs AS (
        SELECT
            inv.item_id,
            inv.category,
```

```sql
            SUM(inv.current_stock * inv.unit_cost) AS inventory_value,
            SUM(sls.sold_qty * inv.unit_cost) AS cogs
        FROM
            inventory inv
            JOIN sales sls ON (sls.item_id = inv.item_id)
        GROUP BY
            inv.item_id,
            inv.category
    ),
    inventory_turnover_per_item AS (
        SELECT
            item_id,
            category,
            inventory_value,
            cogs,
            ROUND(cogs / inventory_value, 2) AS inventory_turnover_rate
        FROM
            sales_cogs
    )
SELECT
    category,
    ROUND(AVG(inventory_turnover_rate), 2) AS inventory_turnover_rate
FROM
    inventory_turnover_per_item
GROUP BY
    category;
```
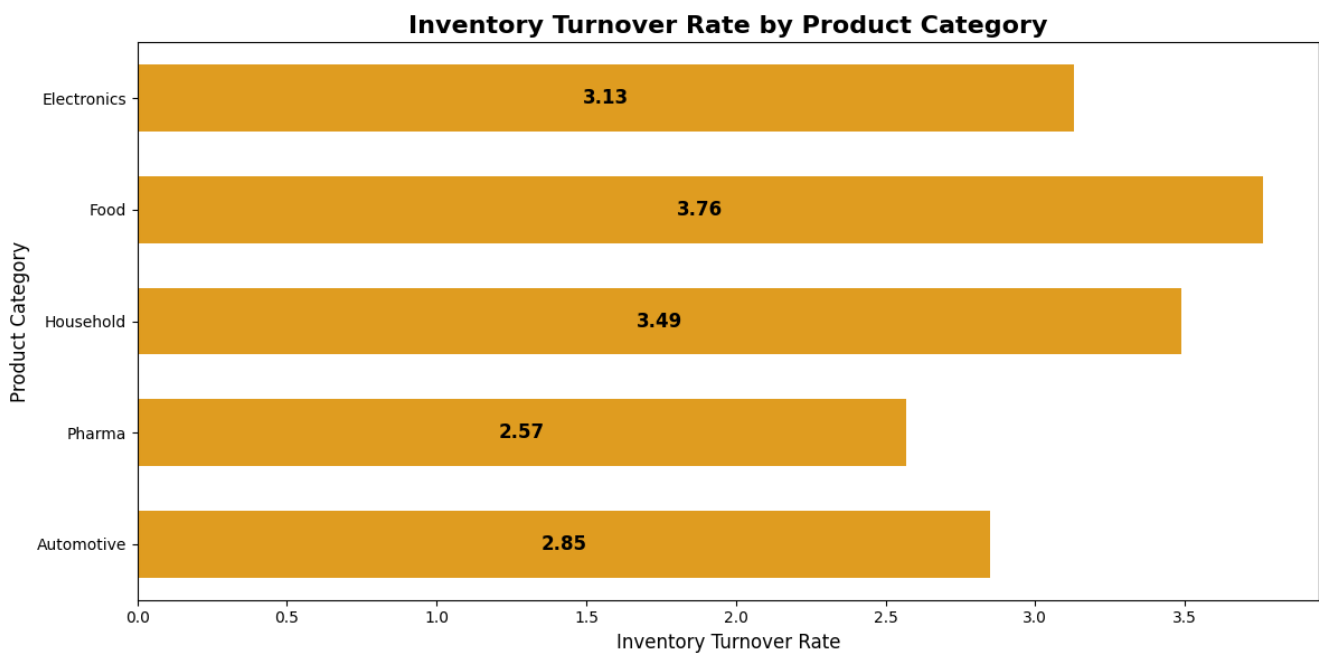

Figure 4: Inventory Turnover Rate by Product Category

Insights :

1. Food & Household performing best when compare to rest of the other categories. It's high turnover rate 3.76 & 3.49 indicates strong demand and efficient stock movement.

2. Worst Performing category is Pharma in all the categories with turnover rate 2.57. Capital is tied up in these category with high carrying cost and risk of expiry.

3. Low turnover rate suggest that these product more likely to sit in warehouse for longer time compared to other categories.

4. Overall Inventory Turnover Rate is 3.16.

Q5. Which suppliers have the most inconsistent lead times or poor delivery performance?

```sql
WITH
    supplier_performance AS (
        SELECT
            spl.supplier_id,
            spl.supplier_name,
            COUNT(pod.*) AS no_of_purchase_order,
            ROUND(AVG(pod.received_date - pod.order_date), 2) AS avg_lead_time_days
        FROM
            purchase_orders pod
            JOIN suppliers spl ON (pod.supplier_id = spl.supplier_id)
        GROUP BY
            spl.supplier_id,
            spl.supplier_name
    )
SELECT
    supplier_name,
    no_of_purchase_order,
    avg_lead_time_days,
    CASE
        WHEN avg_lead_time_days <= 5 THEN 'Good'
        WHEN avg_lead_time_days <= 10 THEN 'Moderate'
        ELSE 'Poor'
    END AS Performance_status
FROM
    supplier_performance
ORDER BY
    avg_lead_time_days;
```
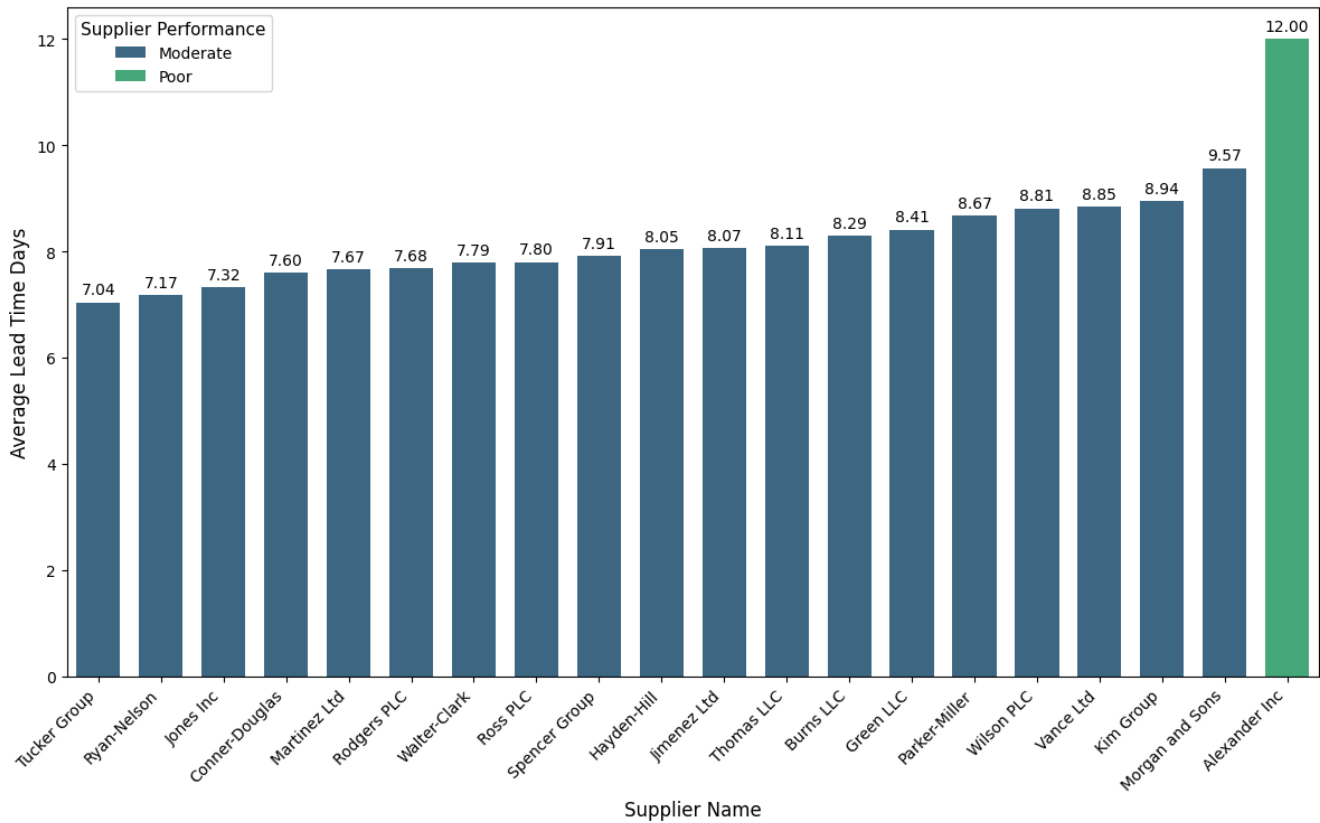
Figure 5: Supplier Delivery Performance

Insights :

1. Tucker Group has the lowest average delivery time 7.04 days. Ideal for urgent procurement, fast moving goods and stock out prevention strategy.

2. Alexander Inc has the highest average delivery time 12 days. Indicating high risk of stock outs, delayed orders and customer dessatisfaction.

3. Rest of the supplier has the average delivery time ranging between 7-9 days. These suppliers are good for reguler replenishment, but not for high demand or urgent products.

Key Insights

1. 4% of inventory have high to moderate stock-out risk.

2. 10.8% inventory is tied up in slow moving stock or dead stock.

3. Some products are going to expire within 10 days with value 8.77M rupee.

4. Food category has best 3.76 inventory turnover rate while Pharma and Automotive has worst 2.57.

5. Supplier Alexander Inc has un-reliable lead time, indicating risk of delayed replenishment.

Recommendations

1. maintain stable stock levels for high demand and fast-moving SKU's to reduce stock out risk and maintain healthy stock levels.

2. liquidate slow moving or dead stock through bundling, discounts and clearance sales and stop routine procurement of these type of stocks.

3. Promote discount schemes for near expiry products to reduce financial loss. Apply FIFO or FEFO method to sell older product before new one.

4. Increase procurement frequency for high inventory turnover categories to maintain product availablity.

5. Avoid overstocking for low inventory turnover categories to save storage space and storage cost.

6. Engage with underperforming suppliers with long or inconsistant lead time for escalation and contact negotiation and shift critical products to suppliers with low or stable lead time.

Conclusion

This analysis shows us the inventory health, stock risk and operation efficiency. By identifying stockout prone items, dead or slow moving stock, near expiry products and unreliable suppliers. The study helps drive smarter replenishment planning and better financial control.

By implementing data-driven decision-making. The orgnizaion can move from reactive inventory management to proactive, efficient and cost-focused approach. At the end improving profitablity and service levels.

Tools & Technologies

Category	Tools Used
Database & Storage	PostgreSQL
Data Extraction & Analysis	SQL (PostgreSQL), Python (Pandas, NumPy)
Visualization	Power BI, Python (Matplotlib, Seaborn)
Database Connector	psycopg2, sqlalchemy
Reporting & Documentation	Markdown, PDF
Version Control & Sharing	Git, GitHub