“Optimizing Inventory Management for a Kitchen Appliances Retailer: A Primary Data Analysis ”

A Mid-Term report for the BDM capstone Project

Submitted by

Name: Yash Karsh

Roll number: 22f3002322

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary

This midterm report presents the progress made on the capstone project titled **"Optimizing Inventory Management for a Kitchen Appliances Retailer: A Primary Data Analysis."** The project focuses on improving inventory practices at *Anand Steel Emporium*, a B2C retail store based in Korba, Chhattisgarh, which deals in home and kitchen appliances.

The business faces ongoing challenges in managing its inventory, frequent stockouts of high-demand products, overstocking of slow-moving items, and capital being locked in unsold inventory. These inefficiencies stem from the absence of structured data practices and demand forecasting. The goal of this project is to introduce a data-driven approach to addressing these issues by collecting primary data directly from the shop.

As of the midterm stage, data collection and cleaning have been completed. Structured transaction data has been analyzed using Python to identify key sales trends, top-performing and underperforming products, and inventory movement patterns. Initial findings confirm significant gaps in stocking practices and the potential to optimize restocking decisions using forecasting models and product segmentation.

The next phase will involve applying ABC analysis, demand forecasting models, and designing an interactive inventory dashboard to enable real-time, data-informed decisions for the retailer.

2 Proof of Originality

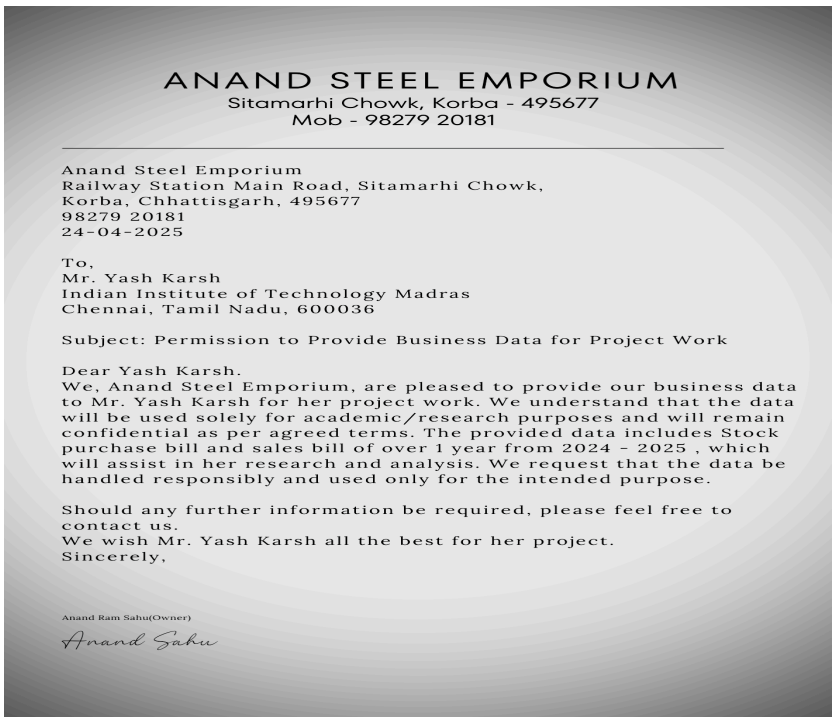
1. Store Front Picture

2. Interaction

   Interaction with the owner

   Video of the store

3. Letter of Permission



ANAND STEEL EMPORIUM
Sitamarhi Chowk, Korba - 495677
Mob - 98279 20181

Anand Steel Emporium
Railway Station Main Road, Sitamarhi Chowk,
Korba, Chhattisgarh, 495677
98279 20181
24-04-2025

To,
Mr. Yash Karsh
Indian Institute of Technology Madras
Chennai, Tamil Nadu, 600036

Subject: Permission to Provide Business Data for Project Work

Dear Yash Karsh.
We, Anand Steel Emporium, are pleased to provide our business data to Mr. Yash Karsh for her project work. We understand that the data will be used solely for academic/research purposes and will remain confidential as per agreed terms. The provided data includes Stock purchase bill and sales bill of over 1 year from 2024 - 2025 , which will assist in her research and analysis. We request that the data be handled responsibly and used only for the intended purpose.

Should any further information be required, please feel free to contact us.
We wish Mr. Yash Karsh all the best for her project.
Sincerely,

Anand Ram Sahu(Owner)
Anand Sahu

3 Metadata and Descriptive Statistics

3.1 Metadata

Source of Data:

Data was collected from physical purchase records maintained by the shopkeeper, which include product restocking entries for two years. The records are noted only at the time of placing a restock order with the supplier and do not include daily opening and closing stock levels.

Files Used:

1. cleaned_Restocking_Records_Data.csv

Size - 20 kb

Rows - 300

Columns - 6

2. cleaned_Extended_Sales_Data.csv

Size - 45 kb

Rows - 801

Columns - 5

Total Records:

• Restocking Records: 300 entries (10 products, 5 suppliers)

• Sales Records: 1,000 entries (includes product, date, quantity sold, price)

Description of Variables:

Restocking records data

Variable Name	Type	Source	Description	Relevance to Problem

Restock Date	Date	Purchase records	Date of restocking from the supplier	Helps determine restocking frequency and seasonality
Product Name	Categorical	Purchase/Sales records	Name of the item	Used to identify fast-moving items
Supplier	Categorical	Purchase records	Name of the vendor/supplier	Helps assess vendor reliability and cost patterns
Quantity Purchased	Numeric (Integer)	Purchase records	Units bought in each restocking event	Used to estimate replenishment needs
Purchase Price per Unit	Numeric (Decimal)	Purchase records	Cost per unit at time of purchase	Used to compute margins
Total Purchase Value	Numeric (Decimal)	Derived	Quantity × Purchase Price	Used for value-based inventory segmentation

Sales Data

Variable Name	Type	Source	Description	Relevance to Problem
Date	Date	Sales records	Date on which the product was sold	Helps identify demand frequency and seasonality
Product Name	Categorical	Purchase/Sales records	Name of the item	Used to identify fast-moving or frequently sold items
Quantity Sold	Numeric (Integer)	Sales records	Units sold per transaction	Used to assess product-level demand
Sale Price	Numeric (Decimal)	Sales records	Selling price per unit	Useful for revenue and profitability analysis
Total Sale Value	Numeric (Decimal)	Derived	Quantity Sold × Sale Price	Determines revenue contribution of each product

3.2 DESCRIPTIVE STATISTICS

Using Excel, the following descriptive statistics were calculated over the dataset of restocking and sales records:

Sales Data

Descriptive Statistics (Sales Data)					
Quantity Sold		**Sale Price**		**Total Sale Value**	
Mean	4.027465668	**Mean**	1402.483533	**Mean**	5646.700936
Standard Error	0.110549895	**Standard Error**	22.60513584	**Standard Error**	193.5836274
Median	3	**Median**	1381.17	**Median**	3615.09
Mode	3	**Mode**	2238.88	**Mode**	9594.1
Standard Deviation	3.128776869	**Standard Deviation**	639.7692749	**Standard Deviation**	5478.792865
Sample Variance	9.789244694	**Sample Variance**	409304.7251	**Sample Variance**	30017171.25
Kurtosis	-0.307513454	**Kurtosis**	-1.191350858	**Kurtosis**	2.260789278
Skewness	1.038671516	**Skewness**	-0.009292303	**Skewness**	1.681108269
Range	9	**Range**	2199.08	**Range**	24675.79
Minimum	1	**Minimum**	300.41	**Minimum**	300.41
Maximum	10	**Maximum**	2499.49	**Maximum**	24976.2
Sum	3226	**Sum**	1123389.31	**Sum**	4523007.45
Count	801	**Count**	801	**Count**	801

Interpretation :

Quantity Sold

- Mean > Median implies **positive skew** (few high sale quantities).
- Skewness +1.04 confirms mild right skew.
- Most orders are **of low quantity (with a median** of 3 units), accompanied by a few larger orders.

Sale Price

- Mean ≈ Median, negligible skewness → roughly symmetric price distribution.
- Std dev ₹639 is high relative to the mean (₹1402), suggesting a wide price range across products or variants.

Total Sale Value

- The mean (₹5646.7) is higher than the median (₹3615.09); skewness of +1.68 indicates a **strong positive skew**, suggesting that most sale transactions are small, but a few large-value orders inflate the mean.

✔ Kurtosis

- Positive kurtosis (+2.26) in Total Sale Value suggests **heavy tails/outliers** (few very large sale transactions).

Restocking Records Data

Descriptive Statistics (Restocking Records Data)					
Quantity Purchased		**Purchase Price per Unit**		**Total Purchase Value**	
Mean	55.86	**Mean**	1159.4752	**Mean**	65990.9881
Standard Error	1.526861384	**Standard Error**	29.34618471	**Standard Error**	2620.615744
Median	55	**Median**	1185.775	**Median**	57759.265
Mode	91	**Mode**	#N/A	**Mode**	#N/A
Standard Deviation	26.44601493	**Standard Deviation**	508.2908292	**Standard Deviation**	45390.39615
Sample Variance	699.3917057	**Sample Variance**	258359.5671	**Sample Variance**	2060288063
Kurtosis	-1.24972085	**Kurtosis**	-1.154364765	**Kurtosis**	-0.485811943
Skewness	-0.050443065	**Skewness**	-0.120146766	**Skewness**	0.660203348
Range	90	**Range**	1784.48	**Range**	188849.44
Minimum	10	**Minimum**	203.04	**Minimum**	3857.76
Maximum	100	**Maximum**	1987.52	**Maximum**	192707.2
Sum	16758	**Sum**	347842.56	**Sum**	19797296.43
Count	300	**Count**	300	**Count**	300

Interpretation :

Quantity Purchased

- Mean ~ Median, negligible skewness (−0.05) indicates **symmetrical purchase quantity distribution**.
- Std dev (26.45) shows moderate variability.

Purchase Price per Unit

- Mean (₹1159.48) < Median (₹1185.78), slight negative skew (−0.12), prices are slightly concentrated towards higher end.

Total Purchase Value

- Mean (₹65990.99) > Median (₹57759.27) → positive skew (+0.66), with some high-value purchase orders inflating the average.
- High Std dev (₹45390) and range (₹188k) suggest wide variability in purchase order values depending on product type or supplier deals.

✔ Kurtosis

- Negative values indicate **light tails / flatter distributions** compared to normal, except Total Purchase Value (−0.49), which is closer to normal.

<u>Strategic Insights</u>

Sales Data

- **Low daily sales quantity (mean ~4 units)** → small transaction sizes, suggesting focus on individual customers rather than bulk orders.
- **High variability in sale prices** → diversified product pricing or premium variants.
- **Right-skewed total sale values** → Few large transactions drive revenue.

Restocking Data

- **Consistent purchase quantities (mean ~56 units)** → bulk buying per restocking event.
- **High purchase cost variability** → multiple product categories or suppliers.

Business Implications

✔ Inventory Management

- Restocking quantities (mean 56 units) far exceed daily sales (mean 4 units) → potential overstocking risk.

✔ **Product Strategy**

- Wide sale price range and positive skew in total sales suggest focusing on **high-value products or upselling premium variants**.

4. Analysis Process and Methods

The analysis methodology for this project is structured around a primary data-driven approach, designed to extract actionable insights from unstructured sales and inventory data of a kitchen appliances retail store. The objective is to identify fast- and slow-moving products, uncover stock inefficiencies, and build a foundation for inventory forecasting and optimization.

4.1 Data Cleaning and Preparation

Given the unstructured and inconsistent nature of the collected data, thorough data cleaning was a critical step. This included:

- Parsing mixed date formats and standardizing them to ISO format.
- Handling missing values using imputation (e.g., estimating totals when price and quantity were available).
- Removing duplicates and rows with invalid values (like negative quantities).
- Standardizing product categories (e.g., grouping "P Cooker", "Pressure Cooker", and "Pressure Cooker Deluxe").

These steps were essential to prepare the dataset for meaningful segmentation and forecasting.

4.2 Exploratory Data Analysis (EDA)

This phase involved exploratory data analysis to understand the structure, range, and trends within the dataset:

1. Descriptive statistics were computed to summarize key variables such as quantity sold, price ranges, and total sale value.
2. Sales were aggregated product-wise and monthly to identify top performers and seasonal trends.
3. Visual tools like bar plots and line charts were used to present sales concentration and time-based movement.
4. For Sales and Revenue Analysis, I have used a Pareto chart to show which product contributes the maximum in sales and revenue. This analysis is done by finding the percentage of each product's contribution in sales and revenue, and I have also found the cumulative percentage of sales and revenue to compare with the Pareto principle. This method helps me to find those products which is contributing the most.
5. To analyze the trend of the sales and the revenue over the fifteen days and weekdays, I used a line graph to depict the trend.
6. To find the number of days each product will last, and to find the average inventory.

   Days of Inventory = Average inventory / Cost of Goods Sold per day.

This phase provided an overview of demand behavior, product concentration, and periods of high transaction volume.

Method	Why It Was Chosen	Why Better Than Others
Pareto Chart	To identify top-selling products easily	Shows both absolute and cumulative impact
Line Graph	To reveal sales/revenue trends over time	Best for visualizing continuous data like dates
Scatter Plot	To compare category performance using dual metrics	Handles multiple variables in one chart
Days of Inventory	To estimate how long stock will last	Simple and meaningful for small businesses

4.3 Planned Analysis Methods

The following analytical techniques are planned for the next phase:

- ABC Classification: To segment inventory into A (high value), B (moderate), and C (low value) categories using cumulative sales value.
- Time-Series Forecasting: Demand prediction using methods like Moving Averages or ARIMA for high-selling items, enabling proactive restocking.

These methods are chosen for their suitability in small retail setups and their ability to provide interpretable insights that the business owner can implement directly.

5. Results and Findings :

Based on the cleaned and structured sales data collected over two years, several key insights have emerged that reflect current inventory patterns, product performance, and operational inefficiencies at the business.

5.1 Year-wise Summary of Sales and Purchases

- Total Sales Value increased significantly in 2024 (₹22 lakhs), showing peak business activity compared to 2023 and 2025.
- Total Quantity Purchased followed a similar trend, highest in 2024 (~9,000 units), indicating alignment between purchase planning and sales volume.

- A notable decline in both sales and purchases in 2025 may indicate either:
    - Over-purchasing in 2024, leading to excess inventory in 2025
    - Seasonal slowdown or reduced demand

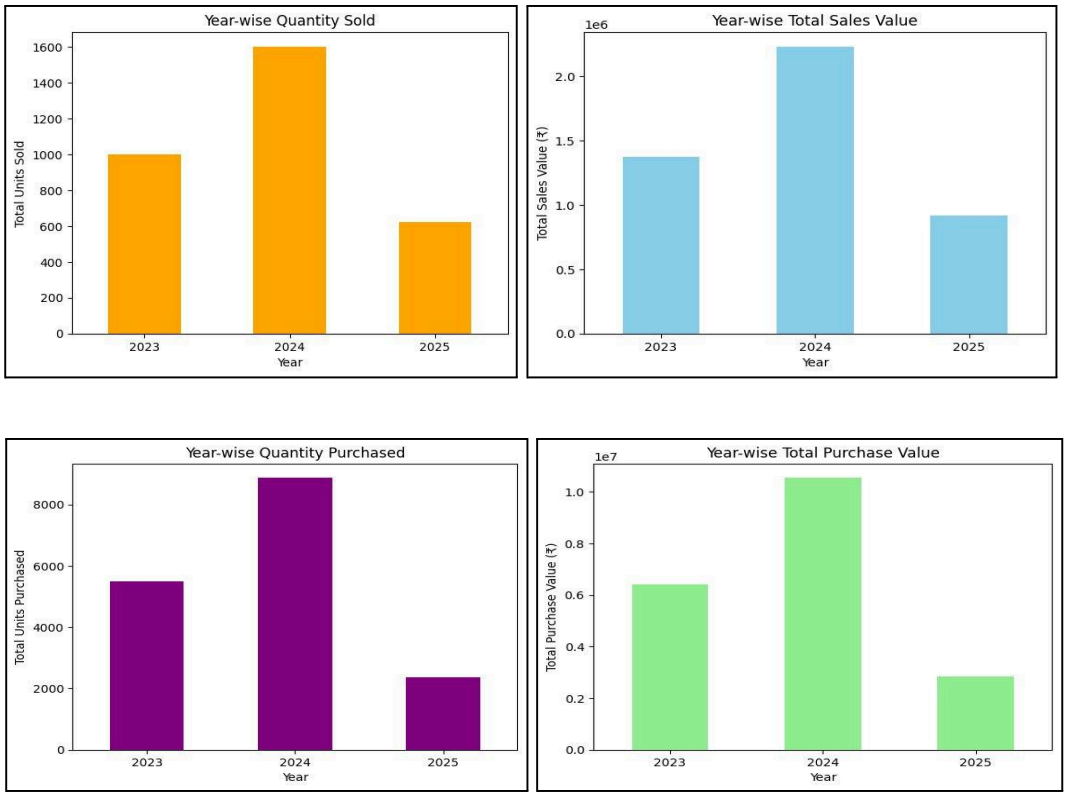


Fig 1. Year-wise Total Sales Value, Year-wise Quantity Purchased]

5.2 Top Performing Products

- Best-selling items included:
    - Utensil set with induction-compatible cookware
    - Non-stick tawas and electric kettles
    - Toaster and gas stove
- These accounted for higher quantities sold, suggesting popularity, affordability, or seasonal relevance.

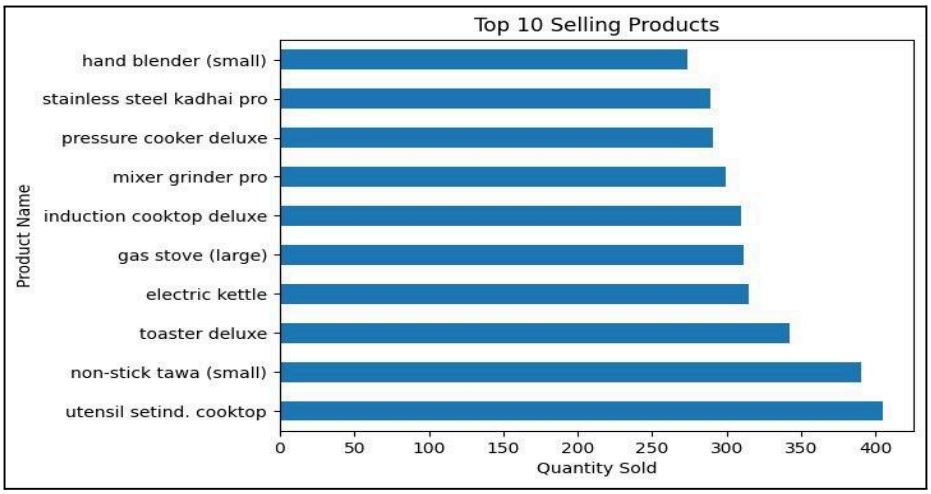


Fig 2. Top 10 Selling Products

- In terms of purchases, the most frequently restocked products were:
    - Electric kettle, non-stick tawa, induction cooktop, hand blender
- The close alignment of purchased vs sold quantities indicates effective restocking decisions.

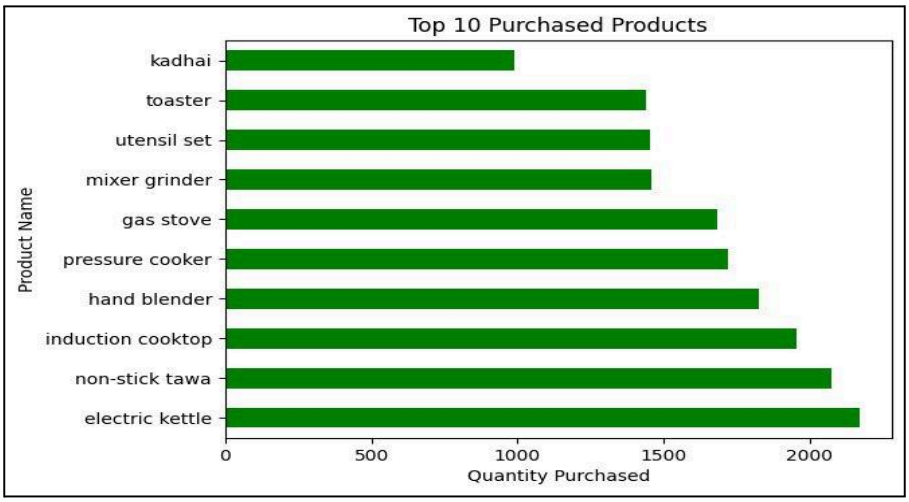


Fig 3. Top 10 Purchased Products

5.3 Daily Purchase and Sales Pattern (Temporal Trends)

Restocking Patterns:

- The restocking cost varied significantly day to day, showing:
    - Bulk purchases at the start of quarters (e.g., April, July)
    - Spikes near festival seasons (e.g., Diwali – October/November)
- In 2024, more frequent and higher-value purchases indicate an aggressive procurement strategy, possibly in response to increased sales.

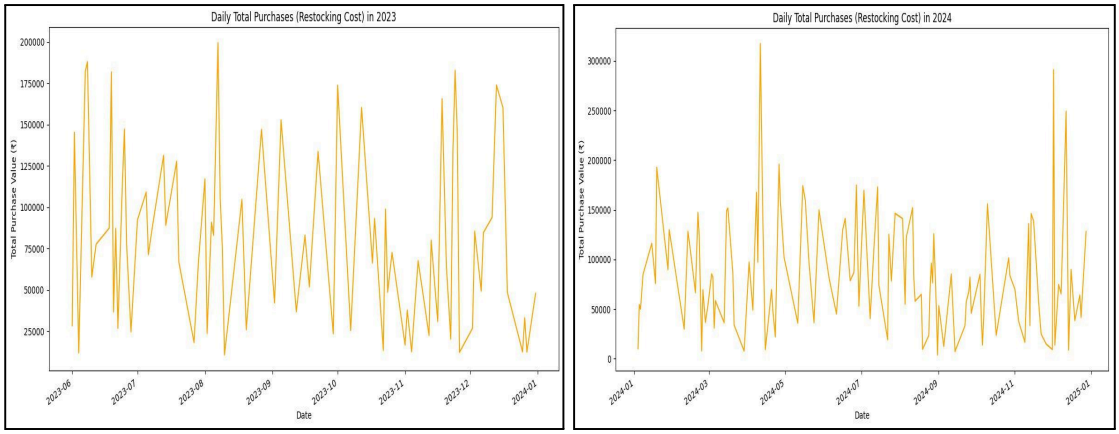


Fig. Daily Purchase Graphs for 2023 and 2024

Sales Patterns:

- Sales followed a more scattered daily trend but had noticeable peaks:
    - In 2023 and 2024, spikes occurred around May, August, and November.

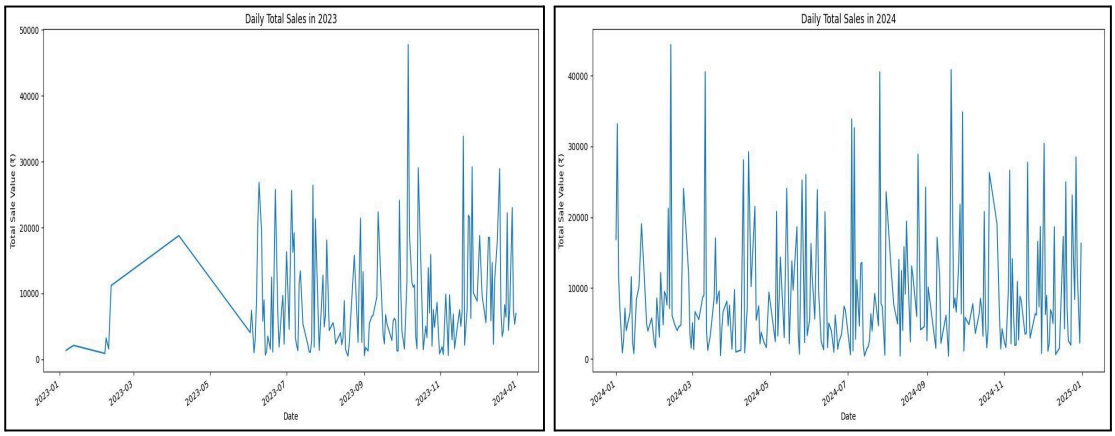


Fig. Daily Sales Graphs for 2023 and 2024

5.4 Insights Derived:

- The business is heavily seasonal, with evident peaks around mid and late year.
- Inventory procurement mostly aligns with demand, but some over-purchasing (2024) and under-selling (2025) hint at inefficiencies.
- Top products contribute disproportionately to revenue — ideal candidates for ABC classification and reorder automation.
- Vendor and item-level trends could be used to build purchase planning templates, reducing manual guesswork.

5.5 Implications for Inventory Strategy

The current stocking pattern appears to be **reactive rather than data-driven**. Frequently sold products are not consistently available, while several low-demand items occupy shelf space for extended periods. This creates a mismatch between demand and stock allocation.

The findings highlight the opportunity to:

- Implement the **ABC classification** for inventory focus
- Use **forecasting models** to predict reorder points
- Eliminate **dead stock** to free up working capital