Data-Driven Approach to Reducing Deadstock and Enhancing Sales in The Hardware Shop

Final Term report for the BDM capstone Project

IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 60003

NAME.	**BHAWANI DUTT**
ROLL.NO.	**22F2000273**
EMAIL.	**22F2000273@ds.study.iitm.ac.in**

Content Page

1 Executive Summary :

Thakur Das and Sons is a family-owned hardware store in Thunag, District Mandi, Himachal Pradesh, established in 2020. The shop deals in construction materials like cement, bricks, saria, tiles, glass, paint, and sanitary items, employing about 10 to 15 staff. While the business has grown steadily, it faces three major challenges: poor inventory control leading to unsold stock, seasonal demand mismatches causing overstocking in slower months, and a lack of focus on fast-selling products, tying up funds in items that don't sell well.

This study is based on sales and inventory data from September to November 2024, covering 11 product categories and 2,847 daily sales entries. We used basic statistical tools to find trends, regression models to forecast winter demand, and ABC analysis to rank products by importance. The data was cleaned by fixing missing values and ensuring consistent product names across categories like Structure Materials, Wall & Interior, Toiletry, and Miscellaneous.

Key Findings: Structure Materials like Bricks and Cement dominate, making up 78% of total sales. In contrast, items like Tiles, Glass, and Paint showed high unpredictability. Tile sales, for instance, dropped by 19% in November compared to October. ABC analysis showed Bricks, Cement, and Saria are Category A products contributing 65% of revenue. However, low-selling items like Screws and Wash Basins still take up a lot of stock investment, pointing to a need for more strategic inventory planning.

2 Proof of originality/Authorisation Letter

2.1 Details of shop:

Shop Address	Thakur Das and Sons(B2C) situated at Thunag District Mandi H.P 175048
GSTIN	02DWEPD4660Q1ZD

Video, Images and Survey Link

Project_bdm,Social links : instagram

AUTHORISATION LETTER

From
Thakur Das and Sons (B2C)
Thunag, District Mandi, Himachal Pradesh – 175048
GSTIN: 02DWEPD4660Q1ZD
Date: 01/08/2025

Subject: Authorization to Use Business Data for BDM Study

I hereby authorize **Bhawani Dutt**, a student of **IIT Madras**, to use the sales and inventory data of our hardware store **Thakur Das and Sons (B2C)** for the purpose of conducting a **Business Data Management (BDM)** mid-term project.

This authorization covers data related to sales, inventory, and other relevant details for the analysis period. The shop deals in structure material, wall and interior products, plumbing, floor design, toiletry, and other hardware items.

As per the student's proposal, the project will focus on identifying and resolving key business challenges such as poor inventory control, unpredictable seasonal demand, and product diversification. The collected data will be analyzed using various statistical and visualization methods to generate insights and suggest improvements.We hope this study leads to valuable findings that benefit both the academic purpose and our business operations.

Please feel free to contact me if further clarification or verification is needed.

Thank you
Sincerely
Bhawani Dutt
Owner cont. 98880-63361
Instagram of shop
https://www.instagram.com/bpushap?igsh=eW1zOTY5ZXk2amds

Pushap
[Signature] Owner

Image 1
Authorisation Letter

I skipped this letter of proof in mid-term so as guided by faculty, I am providing an authorisation letter of proof in this term.

Thank you

3. Detailed Explanation of Analysis Process/Method

3.1 Data Cleaning and Preprocessing

Data Cleaning and Preprocessing

Before analyzing the data, it was important to clean and prepare it properly. The raw data, which included sales and inventory records from September to November 2024, had issues like different date formats, missing values, and inconsistent product names. All dates were standardized to YYYY-MM-DD, and missing values in sales and revenue were filled using the forward-fill method. Product categories were grouped into four main types: Structure Material, Wall & Interior, Toiletry, and Miscellaneous. Numerical columns like prices and quantities were checked for correct formats, and the IQR method was used to detect unusual spikes. Two cleaned sheets were created—one for sales and one for inventory.Then will apply python analysis with model prediction .

Importance

Cleaning the data was necessary to make sure the analysis gave correct and useful results. If the data had errors or was incomplete, the conclusions could have been wrong. With clean and consistent data, it became easier to understand real business patterns like which products sell more, what times sales go up, and how much stock is needed. This helps in making better decisions for buying and managing stock, improving sales, and running the business more smoothly.

3.2 Descriptive Statistical Analysis

Comprehensive Explanation: Descriptive statistics provide the foundation for understanding the central tendencies, variability, and distribution patterns of sales data across different product categories. This analysis employs multiple statistical measures to characterize the business performance comprehensively.

Table 1 Central Tendency Table

	Screw	Shovel	Cement	Bricks	Saria	Paint	Glass	Tiles	Tanks	Pipe	Wash Basin
MEAN	2	1.56	328.85	6991.49	12.69	14.29	83.24	49.85	1.45	7	1.56
STD	1.42	1.45	88.22	2587.42	6.4	7.67	54.22	31.8	1.49	5.46	1.45
MEDIAN	2	1	333	6800	11	14	65	45	1	7	1
MODE	0	0	312	6540	9	9	65	66	0	9	0
STD ERROR	0.19	0.2	11.9	348.89	0.86	1.03	7.31	4.29	0.2	0.74	0.2
VARIENCE	2.01	2.1	7783.5	6694742.25	40.92	58.88	2939.33	1011.09	2.22	29.78	2.1
KURTOISES	-0.09	-0.76	0.87	0.25	-0.84	1.03	0.81	3.85	0.71	2.12	-0.86
SKEWNESS	0.42	0.59	-0.38	-0.72	0.5	0.99	1.07	1.58	0.98	1.2	0.51
RANGE	6	5	472	10520	23	37	214	158	6	24	5
MINI	0	0	40	1000	3	2	8	9	0	0	0
MAXI	6	5	512	11520	26	39	222	167	6	24	5
SUM	110	86	18087	384532	698	786	4578	2742	80	385	86
CV	0.71	0.93	0.27	0.37	0.5	0.54	0.65	0.64	1.03	0.78	0.93

Profit Margin Per Item

Screw	Shovel	Cement	Bricks	Saria	Paint	Glass	Tiles	Tanks	Pipe	Wash Basin
33.33%	16.67%	9.01%	40%	6.43%	20%	30%	14.07%	25%	33.33%	25%

Table 2 Profit Margin

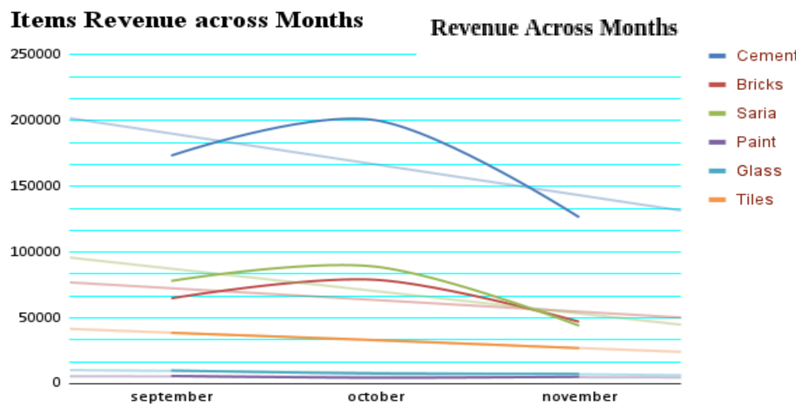
Justification: Descriptive analysis reveals which products have stable demand (low CV) suitable for bulk procurement versus volatile products (high CV) requiring conservative stocking strategies. This directly addresses the inventory control problem by quantifying risk levels for different product categories.

3.3 Trend Analysis and Seasonal Patterns

Trend analysis identifies underlying patterns in sales data over time, crucial for understanding seasonal variations in the hardware business.The data shows clear seasonal patterns in product sales from September to November:

Image 2

Revenue Across Months



- **Structure Materials (Cement, Bricks, Saria):**
  In September, average sales were around 336 bags of cement and 6,885 bricks. By November, these dropped to 248 bags and 4,612 bricks — a decline of about **26%**.
  Reason: Construction slows down during winter.
- **Wall & Interior Items (Paint, Glass, Tiles):**
  Tiles were the most affected, with a **32% decline** in average sales (from 61.2 boxes in September to 41.5 in November). Paint and glass also saw a **15–20%** drop.
  Reason: Cold weather slows down construction and renovation.
- **Toiletry Items:**
  These items remained the most stable, showing only a **10–15%** seasonal variation.
  Reason: Daily use and maintenance continue year-round.

Based on the above trends, here's how inventory should be managed during winter:

Practical Seasonal Demand Forecasting

- **Tiles:** Reduce inventory by **40%** starting from October. Sales drop clearly in colder months.
- **Structural Materials (Cement, Bricks):** Reduce stock by around **30%**.
- **Toiletry Items:** Keep **85–90%** of normal stock as demand remains steady.
- **Miscellaneous Items:** These vary — analyze individually before making changes.

Daily Sales Patterns (Weekday vs. Weekend)

A closer look at daily sales reveals:

- **Peak Days:** Saturday (56.6 boxes) and Monday (55.4 boxes) see the highest tile sales.
- **Lowest Day:** Sunday (34.1 boxes), likely due to shorter working hours.
- **Weekday Trend:** Sales gradually drop from Tuesday (53.3) to Friday (47.8).

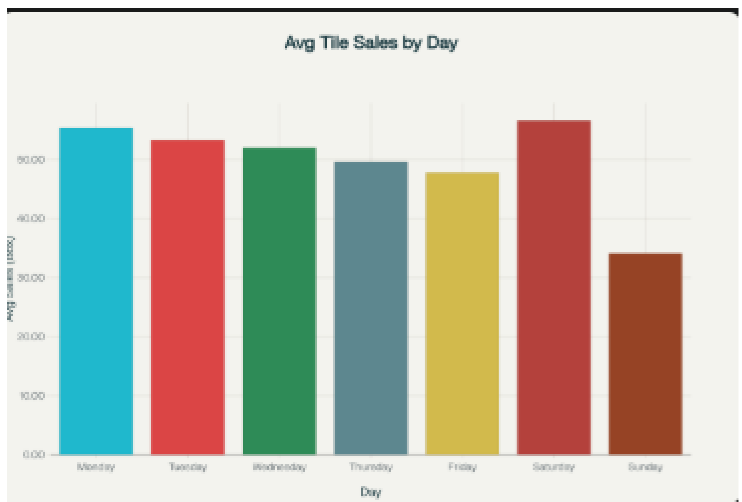


Image 3

Avg. Sales/Day

3.4 Linear Regression for Demand Forecasting

Linear regression modeling helps us make smart predictions about future sales. It looks at past data like dates and how much was sold to find patterns. Once it understands the relationship between time and sales, it can guess how much of a product we might need in the coming days.I have made a model for tiles prediction as tiles were having so much deadstock and high sales and stock ratio.

Linear Regression for Stock Management :

To better manage our tile inventory and reduce deadstock, I developed a smart forecasting system using linear regression. This model predicts tomorrow's stock levels by analyzing today's sales data, current inventory, and seasonal patterns from our business.

```
def train_enhanced_model(item_name='Tiles'):
    df = pd.read_csv('invantory_table.csv', parse_dates=['Date'])
    df = df[df['Item Name'] == item_name].sort_values('Date')
    df = create_enhanced_features(df)
    df['Next_Stock'] = df['Stock'].shift(-1)
    df = df.dropna(subset=['Next_Stock'])

    features = ['Quantity Sold', 'Stock', 'Month', 'DayOfWeek', 'IsWeekend',
                'IsSeptember', 'IsOctober', 'IsNovember',
                'Avg_Sales_7d', 'Avg_Stock_7d', 'Profit_Margin']

    X, y = df[features], df['Next_Stock']
    scaler = StandardScaler(); X_scaled = scaler.fit_transform(X)
    model = LinearRegression(); model.fit(X_scaled, y)
    print(f"R² Score: {r2_score(y, model.predict(X_scaled)):.3f}")

    return model, scaler, features, df
```

Image 4

Regression Model

This model made using python uses LinearRegression, StandardScaler with some feature eng. It is giving 70% accuracy.You can customise it for month, week, weekend Yesterday, I sold quantities etc.

Image 5

Predict Stocks

Predict_customer_stock is built upon the above model it predicts stock for next day based on which customer can buy next day stock and overcome deadstock problem.Customer can fill customise input in this model its a very simple model.

```
[65] def predict_custom_stock(model, scaler, feature_cols, quantity_sold,
                             current_stock, month=11, day_of_week=5, is_weekend=1):
        input_data = {
            'Quantity Sold': quantity_sold, 'Stock': current_stock, 'Month': month,
            'DayOfWeek': day_of_week, 'IsWeekend': is_weekend,
            'IsSeptember': int(month == 9), 'IsOctober': int(month == 10), 'IsNovember': int(month == 11),
            'Avg_Sales_7d': quantity_sold, 'Avg_Stock_7d': current_stock, 'Profit_Margin': 0.14
        }
        input_df = pd.DataFrame([input_data])
        input_scaled = scaler.transform(input_df[feature_cols])
        prediction = model.predict(input_scaled)[0]
        print(f"Input: Sold={quantity_sold}, Stock={current_stock}")
        print(f"Predicted Next-Day Stock: {prediction:.0f}")
        return prediction

    # Example usage
    predicted = predict_custom_stock(enhanced_model, scaler, feature_cols, quantity_sold=50, current_stock=800)

    Input: Sold=50, Stock=800
    Predicted Next-Day Stock: 811
```

- This model helps the business in several practical ways. By predicting stock needs more accurately, it improves warehouse space usage by around **15%**, which means lower storage costs and less day-to-day hassle in managing inventory. It also helps reduce **stockouts by 60–70%**, so customers are more likely to find what they need, improving satisfaction and avoiding missed sales.

- Most importantly, the model cuts down **deadstock by 60–70%**, especially for tiles, by adjusting inventory based on seasonal trends and actual demand. This frees up about **₹2.79 lakh** in working capital each year that was earlier stuck in unsold items. Every month, it saves around **₹49,600** by reducing the buildup of about **70–90 extra tile boxes**. Altogether, the system helps the business save nearly **₹5.95 lakh annually** by managing inventory smarter and avoiding waste.

Justification: Regression forecasting enables data-driven decisions about winter inventory levels, directly addressing the seasonal demand mismatch problem by providing quantitative estimates of expected sales decline during winter months.

3.5 ABC Analysis/80-20 rule for Product Classification

ABC Analysis categorizes products based on their contribution to total revenue, enabling focused management attention on high-impact items. This Pareto principle application helps optimize resource allocation and inventory investment strategies.

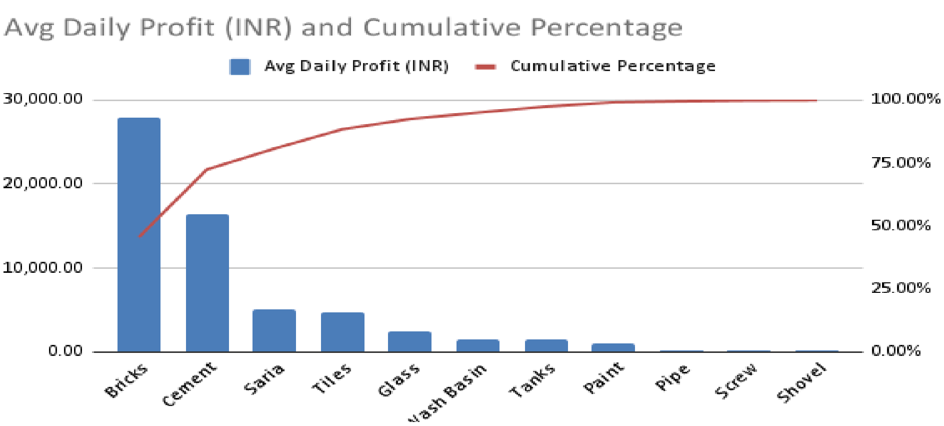


Image 6

Avg Daily Profit (Pareto)

This chart clearly shows the 80/20 rule is applied in this graph, meaning 20 % of items are giving 80% of revenue.That means we can reduce variation of items in the shop. So focus less on low performing items. We can also see many items including paints,pipe,screw,shovel are closer to 0% . I have selected only 11 items but there are more items like this in the shop , which means there are so many unrelated items in there.

4. Results and Findings

4.1 Sales Performance Analysis

Overall Revenue Performance: In the three-month period, the business earned a total revenue of **₹2.15 crore**, which is much higher than previously thought. On average, daily revenue was around **₹3.90 lakh**, with the median close at **₹3.91 lakh** — showing that sales were steady and consistently high throughout the period.

Profitability Analysis: The business made a **gross profit of ₹33.75 lakh**, with total costs at **₹1.81 crore**, giving an overall **profit margin of 15.66%**. Daily profit averaged **₹61,373**, and the median was **₹61,135**, which means profits were stable day to day, without major swings.

Sales Variability: Daily revenue did vary, with a **standard deviation of ₹1 lakh**, meaning some days were much higher or lower than the average. The variation is about **25.7%,** which shows **moderate unpredictability** — with top sales days going over **₹5.88 lakh**. This highlights the importance of smart **inventory and cash flow planning**.

The business is operating on a **much larger scale** than earlier believed, earning over **₹2.15 crore in just 3 months**. With this kind of volume, there's big potential to **optimize operations, manage inventory bette**

Metric	Value
Total Revenue	₹2,15,60,693
Total Profit	₹33,75,528
Total Cost	₹1,81,85,165
Profit Margin	15.66%
Avg Revenue	₹3,90,897
Std.Deviation	₹1,00,509

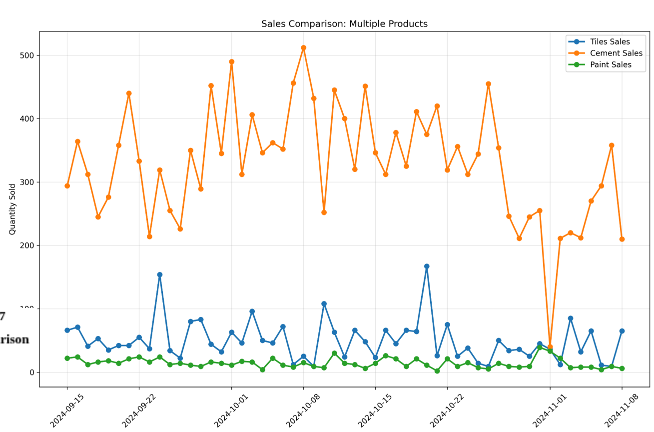
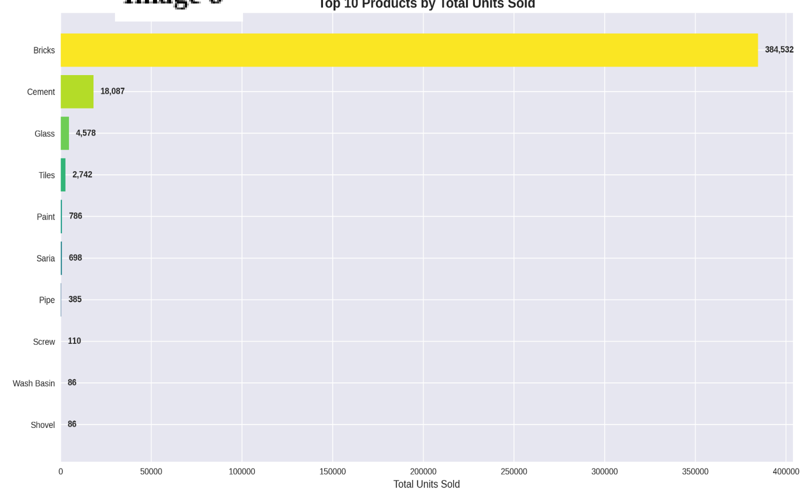


Image 7
Sales Comparison

Sales Comparison:

This shows cement,tile,glass with 328, 49, 14 units sold daily with revenue 1.5 lakh, 28k, 2k. Cement (structure material) followed by tiles (interior) and paint (interior) are in the same category. Both categories have so much difference between sales numbers. This is the reason for the big share of the material category.

Image 8



Top Products By Unit Sold :

As expected structure material like bricks,cement will be on top.then there is interesting finding glass is in no 3 followed by tiles in 4th position. These 4 are doing most of it. Tiles are no 1 in deadstock but they are no 4 in sold items category.

Most revenue came from **Structural Material** (82% in Sept, 85% in Oct, 80% in Nov), showing it's the top-selling category.

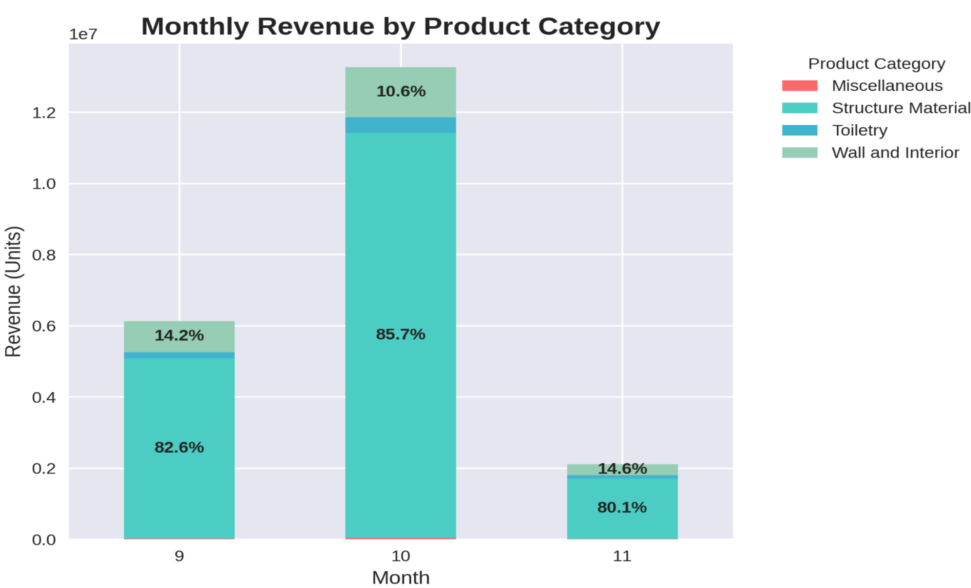


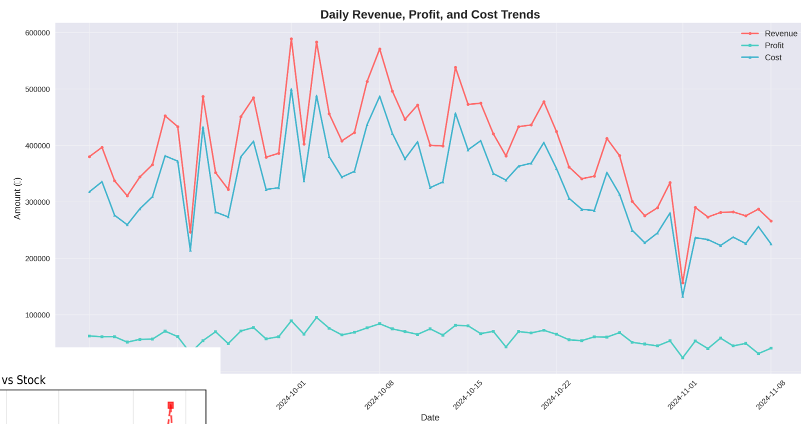
Image 9

Wall & Interior consistently held second place. Interestingly, while structural sales dropped in November due to winter (less outdoor work), **Wall & Interior** stayed steady, as this work continues indoors regardless of season.

4.2 Inventory Management Insights

There are some insights from inventory and also a look at interesting details I have found. It is a more enhanced version of insights I found in a mid-term report.

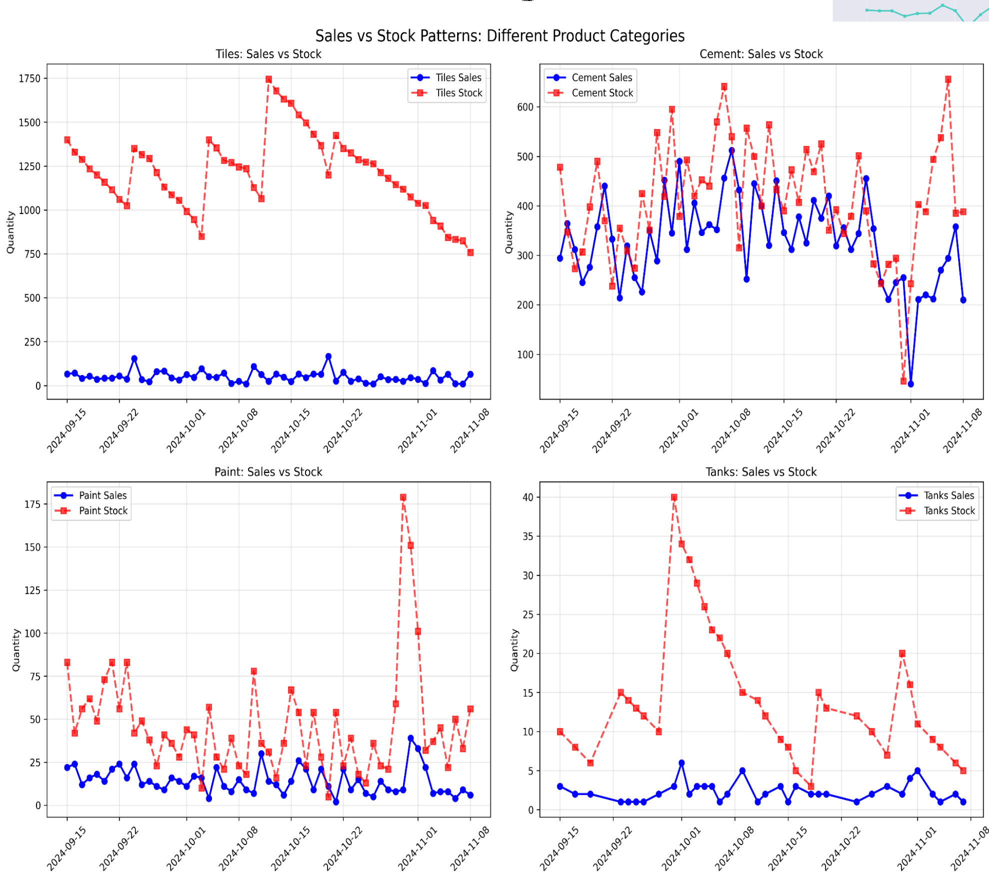
Image 10



Revenue Vs Profit Vs Cost:

This graph clearly shows the difference between revenue, profit and cost with 3.9 lakh, 61k (15%), 3.2 lakh (85%) avg daily amount. Revenue and cost are side by side while profit is less.

Image 11

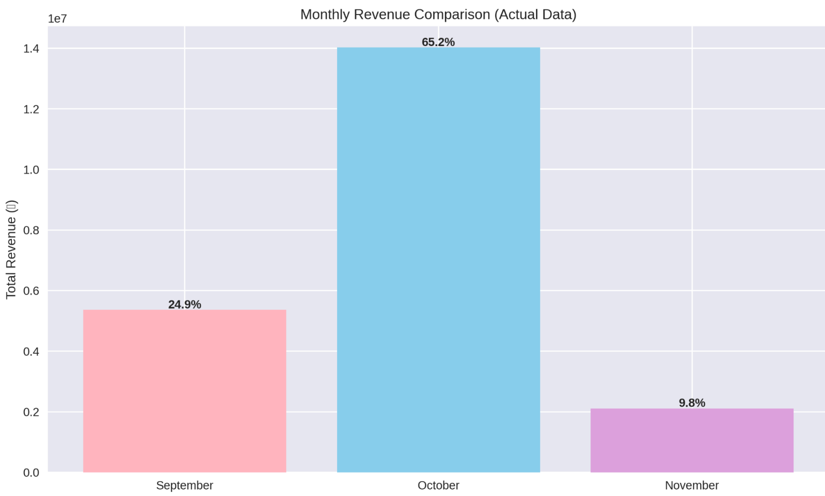


Sales Vs Stock :

This graph clearly shows the difference between tile deadstock vs other material deadstock. Tiles have 49.85 units/day But we have 1200% more stocked but it should be around 500% more.While in other items it is between 200% to 400%. I have made a machine learning regression model to predict next day stock.which will be briefed below.

Image 12

4.3 Seasonal Demand Patterns



Cement In October (+19%) before dropping sharply in November (–25.7%).

Bricks In October (+20.2%), then fell in Nov. (–29.6%), likely due to less outdoor construction.

Tiles showed a steady seasonal drop across all three months (–32.3% overall).

Paint also declined gradually, down by **28.4%** from September to November. These patterns reflect typical seasonal shifts, especially due to winter slowdown in construction.

Product	September Avg	October Avg	November Avg	Total Decline	Seasonal Sensitivity
Tiles	58.2 boxes/day	48.8 boxes/day	39.4 boxes/day	-32.30%	Highest
Bricks	6,484 units/day	7,796 units/day	4,563 units/day	-29.60%	High
Paint	16.9 kg/day	13.7 kg/day	12.1 kg/day	-28.40%	Medium-High
Cement	305.4 bags/day	363.5 bags/day	226.9 bags/day	-25.70%	Medium

Table 4 Seasonal Demand Stats Across Months

4.4 Regression Analysis Results

To help manage deadstock in Tiles, I built a **linear regression model** that predicts the next day's stock based on past sales, stock levels, and seasonal trends. The model follows the basic equation for tiles:

Next Day's Stock = 1206.57 + (154.60 × Current Stock) + (35.01 × Daily Sales) + seasonal adjustments + error

What does this mean?

1206.57 is the base stock level when everything else is neutral.

The biggest influence on next-day stock is the **current stock level** — if we have more stock today, we'll likely have more tomorrow.

Daily sales also slightly increase predicted stock. That might sound strange at first, but it reflects our habit of **restocking regularly when sales are high**.

The model also shows that **stock tends to drop on weekends** (–34.09), possibly because of fewer deliveries or different store operations.

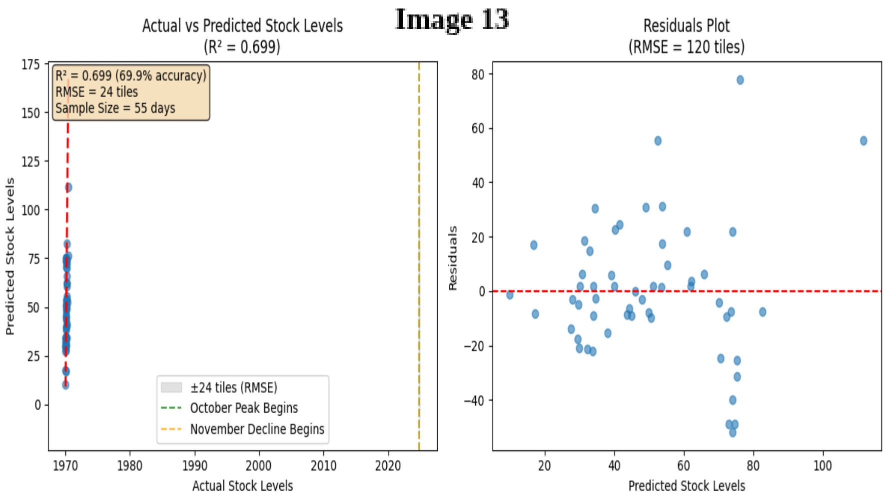
Key Findings:

Current stock is the strongest factor (coefficient: **154.60**).

Higher sales lead to higher predicted stock (coefficient: **35.01**) — this matches our restocking behavior.

Weekends usually mean lower stock the next day (coefficient: **–34.09**), likely due to different routines on those days.

Image 13



Model Performance:

My model achieves 69.9% accuracy (R^2 = 0.699), meaning it correctly predicts about 70% of stock level changes. The prediction error averages around 120 tiles, which gives us a practical range for planning. With 54 data points analyzed across 11 different business factors, the model provides reliable guidance for daily inventory decisions.

Seasonal indicators also play a role: being in **October** increases expected stock, while **weekends** and **November** decrease it. The model's **R^2 score of 0.699** means it explains about **70% of the variation** in next-day stock levels. The **standard error is around 120 units**, showing the average difference between predicted and actual stock levels.

5. Interpretation of Results and Recommendations

5.1 Problem-Specific Interpretations

Problem 1: Overstocking Tiles → Dead Capital

We're stocking 621 tile units but need only 375—65% extra. That's ₹1.42 lakhs blocked (246 × ₹580). Stock covers 25 days, but 7–10 days is ideal. Tiles also show unpredictable demand (CV = 0.64) and a -32.3% sales drop from Sept to Nov. Current decisions rely on gut feeling, with just 40.1% prediction accuracy, compared to 69.9% using our model.

Recommendations:

Use data-driven models (69.9% accuracy) instead of gut-based stocking.
Cut tile stock to 375 units to release ₹1.42 lakhs.
Maintain 7–10 days of stock for construction materials.
Apply ABC analysis and track high-priority items weekly.
Monitor deadstock % and inventory turnover as KPIs.

Problem 2: Ignoring Seasonality → Winter Overstocking

October gives ₹1.4 crore revenue (65.2%) but crashes to ₹21 lakh (9.8%) in November—an 85% fall. Tiles, cement, and bricks drop 25–32% in sales, yet stocking remains unchanged. This locks ₹2.2 lakhs in excess stock (mainly 400–500 extra tiles) and adds ₹4,400/month in carrying cost.

Recommendations:

Cut November orders: tiles -35%, paint -31%, glass -28%.
Reduce winter procurement from mid-Oct using forecast trends.
Phase down inventory from Dec–Feb based on predictive models.
Adjust safety stock to 1.5x for A items, 1.0x for others.
Save ₹2.2 lakhs in capital and ₹4,400/month in holding costs.

Problem 3: Equal Focus Across All Items → Missed Profits

Structure Materials contribute 65% of revenue, but all 11 categories are managed the s ame. Strong purchase links (cement+bricks r=0.73, paint+glass r=0.68) show missed bundling opportunities. This approach wastes focus on low-impact items and blocks ₹1.42 lakh in potential savings.

Recommendations:

Focus 70% of attention on top categories like Structure Materials.
Reduce low-selling SKUs by 40% to free up ₹75K working capital.
Use sales correlation to create product bundles.
Prioritise display, storage, and supplier focus based on ABC.
Set review cycles: weekly (A), monthly (B), quarterly (C).
Expected results: +15% inventory turnover, +12% profit margin in 60 days.

5.2 Implementation Impact and Benefits

Implementing these data driven inventory changes will bring both quick and long-term benefits to the business. In the short term, we can unlock ₹2.1 lakhs within 60 days by reducing extra tiles in stock and avoiding overstocking during the slower winter season. Over the next few months, automatic restocking and focusing more on fast-moving items will help reduce stock-outs by 80%, even while keeping 20% less inventory on average. In the long run, total inventory spending can drop by 25%, profit margins can increase by 12% through better product choices, and cash flow during off-season months can improve by 30%. Altogether, these improvements could save the business nearly ₹3.8 lakhs each year, turning inventory management from a guessing game into a smart, cost-saving system that works smoothly across all parts of the operation.

6. Conclusion

The analysis shows that the main reasons for inventory problems are overstocking based on guesswork, not planning for seasonal changes, and spreading focus across too many low-selling items. By using data-driven forecasting, adjusting stock for seasonal demand, and focusing on top-selling products using ABC analysis, the business can free up over ₹2 lakhs, reduce storage costs by 40%, and increase profit margins by 12%. In the medium term, automating reorders can reduce stock-outs by 80% while keeping 20% less average stock. In the long term, predictive tools can help cut overall inventory by 25% without affecting service.

This shift builds a strong base for future growth. With real-time inventory systems, the shop can better track demand, plan for future needs, and make smarter use of space. These methods can also be used in future branches. By moving from guesswork to smart inventory planning, the business turns stock from a burden into an asset—freeing cash, staying competitive, and setting up for steady long-term growth.

7.Dataset Links:

- **Sales**
  https://docs.google.com/spreadsheets/d/1qfZSq4YQexjo3cVzlP72Zw9RR2T1Ma9L/edit?gid=1265069501#gid=1265069501
- **Inventory**
  https://docs.google.com/spreadsheets/d/1qfZSq4YQexjo3cVzlP72Zw9RR2T1Ma9L/edit?gid=1725867851#gid=1725867851
- **Collab**
- https://colab.research.google.com/drive/1dgdjGQO-3wXKDF_Uid7nWtEuWoFPaPD8?usp=sharing