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

Mid Term report for the BDM capstone Project

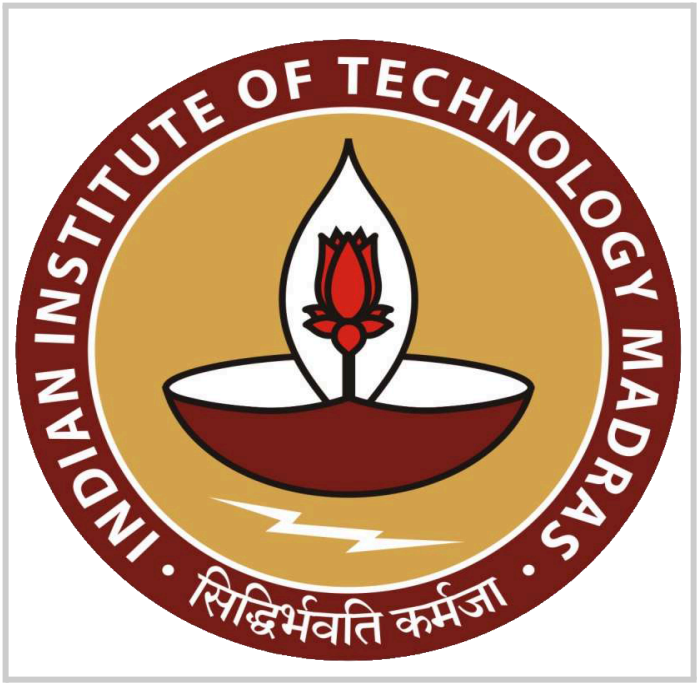



image 2.

**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** |

# INDEX

# 1 Executive Summary :

This project is mid term submission of analysing a hardware store in my town called "Thakur Das and Sons," located in Thunag, District Mandi, Himachal Pradesh (175048). Thakur Das And Sons is dealing with structure material, wall and interior, plumbing, floor design, toiletry etc. category.

As we have seen from the proposal, this hardware store is facing three major problems: poor inventory control, unknown seasonal demand and selling so many unrelated items. To address, I have collected a dataset.It will be organised into sales and inventory sheets.

After cleaning the sales and inventory sheet, various statistical methods, including pivot tables, pie charts, bar charts and time series charts will be deployed to gain insights. Time series sales patterns will also be investigated.Inventory vs sold items and profit vs revenue will be understood graphically. This will help to identify the final approach or method to solve problems.

# 2 Proof of originality of the Data :

To support my claim that data is of primary category and the dataI am sharing video proof of interacting with shop owners, in which I have also shared images of shop and premises.

**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**

**https://drive.google.com/drive/folders/1rv7CN6XqL2S5NZlXRie5vDL0hxp4KSCa**

**Excel Sheet**

**https://docs.google.com/spreadsheets/d/1qfZSq4YQexjo3cVzlP72Zw9RR2T1Ma9L**


## 3 Metadata :

I have continuously gathered data from September 9 to November 8, 2024. This tracks daily sales and inventory, hence 2 sheets were made Sales and Revenue. I choose the best 11 performing(as per owner) items to analyse this project. Here is metainfo.

**Sales Table**

| Time | | | Miscellaneous | | Structure Material | | | Wall and Interior | | | Toiletry | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Date | Month | Day | Screw (kg) | Shovel (piece) | Cement (bag) | Bricks (piece) | Saria (qntl) | Paint (kg) | Glass (sq.m) | Tiles (box) | Tanks (piece) | Pipe {mtr) | Wash Basin (piece) |
| 9/15/2024 | September | Sunday | 1 | 0 | 294 | 5740 | 10 | 22 | 212 | 66 | 3 | 10 | 3 |

**Table 1**

**Inventory Table**

| Date | Category | Item Name | Unit | Quantity Sold | Sale Price (per unit) | Purchase Price(unit) | Total Revenue | Stock |
|---|---|---|---|---|---|---|---|---|
| 9/15/2024 | Structure Material | Bricks | Per Item | 5740 | 10 | 6 | 57400 | 7254 |

**Table 2**

**Quantity Sold:** Total quantity sold that day of a specific item

**Sale Price Per Unit:** Selling price of item

**Purchase Price:** Price at which item was purchased

**Total Revenue:** Selling Price * Quantity Sold

**Stock:** Stock Left

### 3.1 Descriptive statistics :

**Measure of central tendency**

| | 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 4

- Highest Margin = Pipe and Screw (Bricks have much transportation involved in it),
- We can see structural items have lowest margin (Cement,Bricks,Saria)
- **inconsistent sales vs steadier demand.**High CV ($>1$): Tanks ($\approx 1.03$), Glass ($\approx 0.65$), Screw(0.71) and Wash Basin ($\approx 0.93$) **vs** Low CV ($<0.5$): Cement, Bricks, Paint, Tiles shows 3rd business problem of unrelated items in the shop.
- For most categories (e.g. Cement, Bricks, Paint, Tiles), the mean and median are quite close, suggesting symmetric distributions without extreme outliers.

# 4 Detailed Explanation of Analysis Process and Methods :

## 4.1 Data Collection and Preprocessing

The main sources of data are bills, invoices, and data collected by myself. which provide comprehensive sales information for products sold. On the basis of this data 2 sheets were made, Sales and Inventory. For the preprocessing part i have cleaned data like duplicate date and ensure dates are Date type, numeric fields like price/quantity are Number.

After making sheets I have categorised items in categories like structure material (saria,cement,bricks), wall and interior (tiles, glass, paint), toiletry (wash basin, pipe, tank) and miscellaneous (screw, shovel).

**Justification :** For accurate and correct results in analysis a good data cleaning is necessary. I have made sure not to remove any column to maximize data points.

**Descriptive Statistics Analysis** : For the feature engineering on Sales **Revenue** (Quantity × Sale Price), **Profit** (Sale Price - Purchase Price * Quantity Sold) and **Coast** (Revenue - Coast) for **per day** are calculated. For Inventory sheet total Total **Revenue** and **Stock left** for **per day/per items** are calculated also daily profit and cost are kept in track by adding extra columns. Central tendencies are calculated to know the overall picture of data as shown in **table 3** Daily

data processing like daily sales, purchase, profit, cost. This gives us an overall picture and pattern in data.

**Justification** : From table 3 we conclude Pipe and Screw have higher margin because bricks have higher transportation cost. Structure Material have higher daily demand while category in miscellaneous and wall and interior have lower daily demand. We also see differences between coefficient variance (CV) among unrelated or different Niche Items which validate 3rd business problem

## 4.2 Trend Analysis

**Purpose: Understand how sales or stock change over time**

1. Daily/weekly/monthly sales trends per item or category
2. Seasonal peaks (e.g., Tiles selling more in summer)
3. Identify growth/decline in sales volume over time

**Pivot tables:**Pivot tables are the most powerful tools available to summarizing and analyzing large datasets. They give users the ability to segment data into components and perform multiple calculations, such as sums, counts, averages, and percentages. I have used it for month wise, and categories wise in sheet 7 at above provided sheet link. It was very helpful to compare across different categories

**Why Pivot Tables?**

Pivot tables are favored over traditional data analysis methods due to their flexibility and efficiency. Notable advantages are:

Quick Summarization, Quick Adjustments, Focused Analysis, Efficiency.

**Graphical Tools: Different Charts for Different Analyses:**

**Pie Chart:** I have used pie charts to show category wise revenue and profit in **sheet 1**. It tells us how much each category is participating in the total portion.

- **Why Pie Charts?** Pie charts are generally used for proportion analysis. The objective, of course, is how immediately apparent it is visually that a particular portion represents a part of the whole. It is also easy to view the percentage of each category.

**Line Chart:** I have used a line chart to show sold vs stocked tiles, sold vs stocked tiles over the time to compare selling vs stocking speed of any item. That give good overview of overall

picture

- **Why Line Charts?** Line charts are generally used for individual measures with respect to trends over time, thus making it much easier to analyze patterns, fluctuations.

## 4.3 Seasonality Analysis

To identify seasonal trends in the data seasonal analysis was used. **Most items show a decline in demand during the winter months**, indicating a **seasonal demand pattern**..This can be seen in **figure 13.** This confirmed that all items in stock follow seasonal demand patterns.

- **Why Time Series Charts?** Time series charts are ideal for trend analysis because they display data points in chronological order, allowing easy observation of patterns, fluctuations, and long-term trends over time.

## 4.4 Deadstock & Niche Item Detection

**Goal:Find items with very low movement or sporadic demand.**

To identify deadstocks, High Coefficient of Variation (CV) and Sold vs Stock left for each day were drawn using a line chart. It can be seen in figure 15 below or in sheet.I have done same comparison analysis across all items but tiles have most difference between Sold VS Stocked

## 4.5 Customer Demand Pattern Analysis

**Goal:Understand how customers behave with what they buy together, frequently.**

I Identify Consistent vs Inconsistent Items by Cross-Item Correlation CV < 0.5 → **stable demand and CV > 0.8 → variable or unpredictable.**

Cement ↔ Bricks (0.73) → Likely bought together.

Tiles ↔ Paint/Glass → Mild correlation → Often used for finishing/interior.

Screw, Wash Basin, Tanks → Weakly correlated with others → Independent demand.

## 4.6 Correlation Analysis

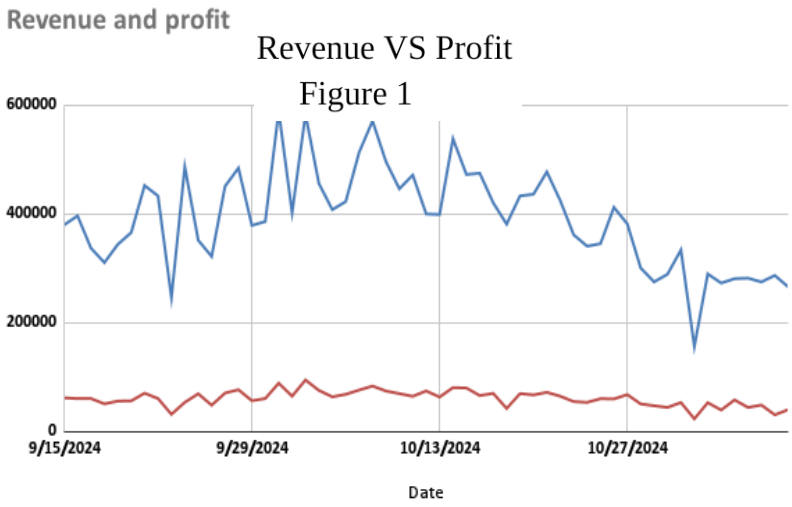
**By finding correlation between items these were the results.**

Strong positive correlation: Cement & Bricks → 0.73 (Likely sold together frequently)

Low/No correlation:Paint & Glass → –0.08 (Sold independently**)**

Shovel,Screw,Paint,Pipe Wash Basin have very low corr across all items. They seem unrelated.

## 5 Results and Findings :




Revenue VS Profit

Figure 1

**6.1 To Study variation of average revenue and profit across the time:** I plotted Average revenue And Average profit for a day wise over a period of time.

**What?**This represents the cyclical variations of the Total average Revenue and Profit across the time period of two months. As it shows that there is a **cyclical nature t**o the revenue and profit.

**Why The Peaks and Troughs?** The Peaks and Troughs or cyclic nature in chart is due to **weekly cycle,** on sunday shop open for half Figure day sometine. So they can be seen on both revenue and profit.The second thing we noticed is that lines are going down as we are approaching winter.That means this line chart is **showing seasonal demand** affecting revenue and profit.

**6.2 To study segment wise contribution to revenue and profit:**

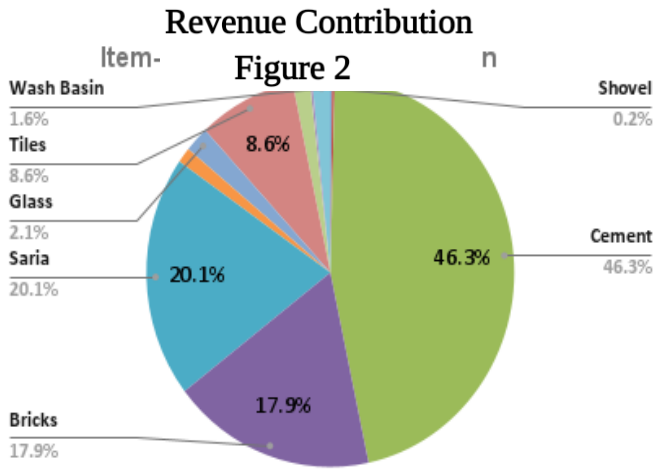


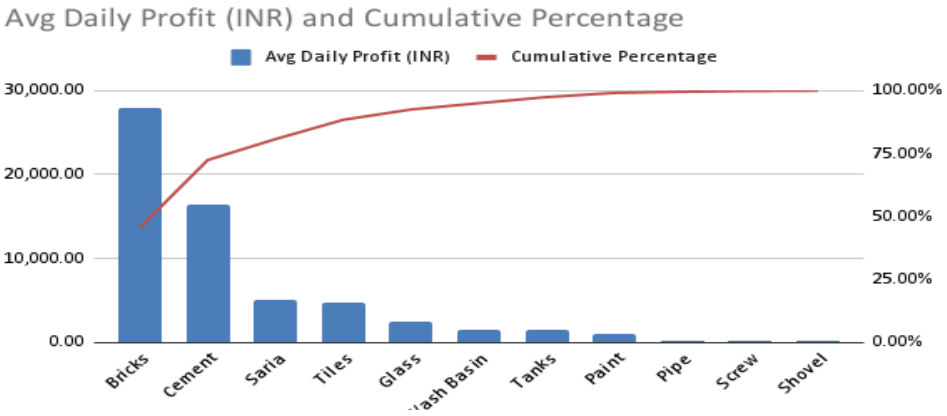

Figure 2

Pie charts are drawn for revenue to see how much each item is contributing to total revenue.

**What ?** Here we can see that **46 %** of contribution comes from **cement**, followed by saria with 20% and bricks with 18%. The lowest contribution comes from **shovel 0.2%,** wash basin with 1.6% We can clearly see on both chart **structural material** category is dominating, it has **84%** share followed by Wall and interior 12%

**We** can see structural material have a bigger piece yet lower margin from (table 4)

**Why?** This happens because structural sections have bulk demand also,higher demand for construction. But this category has a lower margin because of its high transport cost.



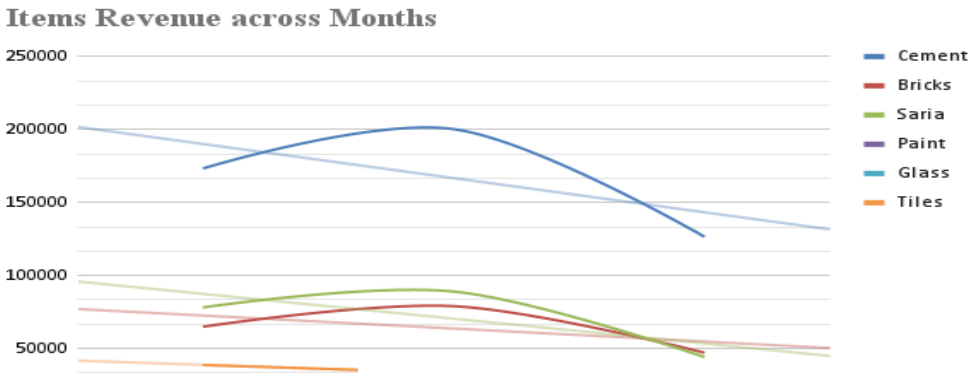

Avg. Daily Profit (Cum.)

Figure 3

**6.3 To study Pareto principle in the sales data:**

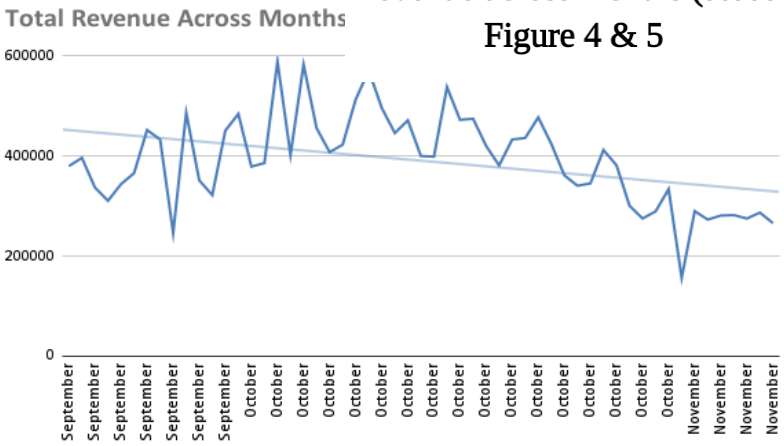
This chart clearly shows the 80/20 rule is applied in this graph, meaning 20 % of items are giving 80% of revenue.That means we can reduce variation of items in the shop. So focus less on low performing items.

**6.4 Plotting Seasonal Trend:**We will see effect of seasonal trend on each item and total revenue




Revenue across months (seasonal)

Figure 4 & 5

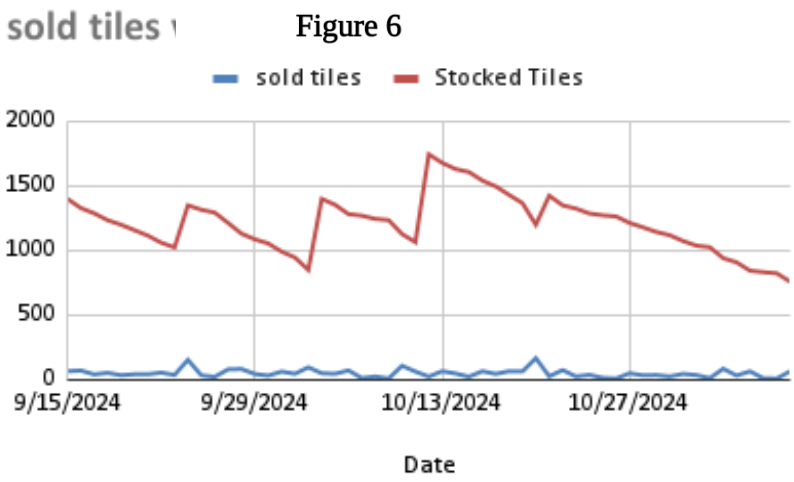



**What?** Fig.13 and 14 tells us there is a seasonal trend.We can see the graph of smooth line chart is tilted towards November, also trend lines are going down as approaches to winter that clearly shows that there is seasonal demand

**Why?**This seasonal trend is seen because of the shop's geographical location.Due snowy atmosphere and extreme cold, construction works just stop during winter so people does not purchase items

### 6.5 Analysis of the inventory

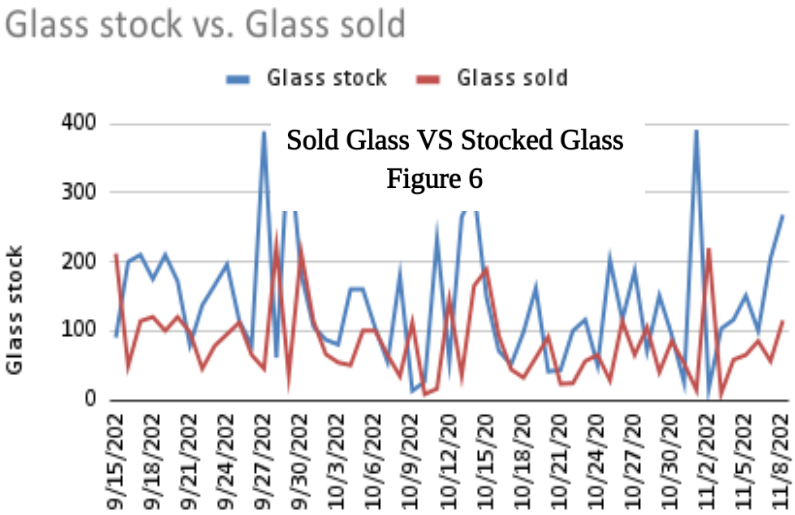
Sold Tiles VS Stocked Tiles
Figure 6




**what?**These charts show sold **tiles vs stocked tiles**. The difference between the two red and blue tells us the presence of deadstock.That means items are stocked frequently but are not sold at that pace.The vertical jump on red line shows stock is filled that day.

**Why ?**Tiles are sold in bulk, if the customer did not find them. If customers don't find the desired quantity or preferred color/design, they often choose not to buy at all. Additionally, many tile designs become outdated after a year, leading to a decline in customer interest. As a result, unsold stock accumulates, increasing deadstock in this category. But shops have to fill more tiles because some have left less design and some go outdated.

Sold Glass VS Stocked Glass
Figure 6

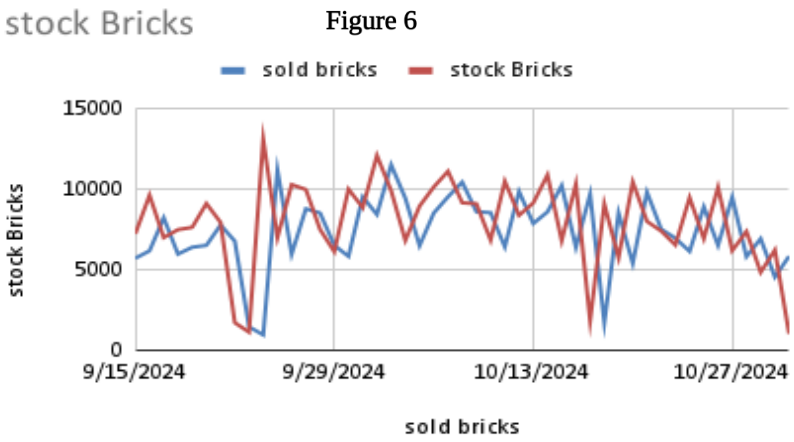



**We** can not see so much of a difference on glass and bricks charts. This tells us that these items are sold frequently, while **Tiles** have so much **deadstock**

**Why?** Glasses don't have that much variety and have good demand so we can see less difference between stock glasses vs sold glasses. They don't have customer segments like tiles have. We can say tiles have different customer types than other items.

Sold Bricks VS Stocked Bricks

Figure 6




**What?** The same pattern is seen in bricks. Bricks are higher revenue item.All structure Material have same graph also they save strong correlation between them

Why? Structure Materials are always purchased in bulk and they don't have variation like tiles have so they have not shown much difference between Sold vs Stocked graph. Also Structural Material have strong correlation between them.