# Optimizing Inventory Management and Supply Chain Efficiency to Minimize Dead Stock and Reduce Breakage

A Final Report for the BDM Capstone Project


Submitted by:

Name: Aditya Nupani

Email ID: 23f1000873@ds.study.iitm.ac.in

Roll Number: 23f1000873

IITM Online BS Degree Program

Indian Institute of Technology, Madras, Chennai,

Tamil Nadu, India, 600036

# Contents

# 1 Executive Summary

Spllatter is a Kolkata based company which specializes in selling cutlery, tableware, barware and similar products. Its major clients are resorts, hotels and restaurants across the country. The company faces two major challenges: accumulation of dead stock due to ineffective inventory management and losses due to breakage of goods during transportation. The lack of an effective inventory tracking system results in excess stock of low demand items whereas poor packaging and transportation methods contribute to high breakage rates, leading to financial losses.

The data used for the project is from July 2024 to December and the bank statement for 2024. These records were extracted from PDF files, cleaned using Microsoft Excel and analyzed using Excel and Python. Advance time series analysis was used to understand sales trends while demand variability analysis categorized the SKUs based on demand fluctuations. Additionally, geospatial mapping and route wise analysis identified high-breakage regions. Root cause analysis using fishbone (Ishikawa) diagram was used to pinpoint the primary factors contributing to product damage.

On performing the analysis, it was found that the sales peak between the months of July and October driven by wedding seasons ad festivals. Hence, stocking up of inventory in June with popular and revenue-generating SKUs in June is crucial to meeting the demand and preventing accumulation of dead stock. All SKUs were found to have moderate variability indicating an unpredictable sales pattern. Higher breakage rates were observed in northern and western regions for clients in Darjeeling and Maharashtra due to rugged terrain and inadequate packaging.

Some of the recommendations include adjustment of inventory levels based on demand trends to prevent dead stock, and streamlining of SKUs for a stable product lineup. Packaging methods should be improvised and alternative ways of transportation (such as by air, by rail) should be used to reduce the breakage. Lastly, maintaining a breakage log will help to refine the packaging and logistics strategies to enhance profitability. Implementing these strategies is expected to reduce breakage in high-risk regions and improve inventory efficiency. This will contribute to improved profitability and better supply chain control.

# 2 Detailed Explanation of Analysis Methods

## 2.1. Data Cleaning and Preprocessing

The sales and purchase invoices of the company for the months July 2024 to November 2024 were provided to me in the form of PDF documents which were transferred to Microsoft Excel. All duplicates were removed and quantities such as net amount were standardized to Indian Rupees (INR) and number of damaged goods was rounded off. While performing analysis, the taxes and shipping costs were excluded as they were not required for the analysis. Python was also used for data cleaning while geocoding of locations.

```
import pandas as pd
from geopy.geocoders import Nominatim
from geopy.extra.rate_limiter import RateLimiter
import time
df = pd.read_csv('buyers.csv', header=None, names=['company',
'city', 'state'])
location_fixes = {
    'Jaiur': 'Jaipur',
    'Khan Market': 'New Delhi',
    'Bokaro': 'Bokaro Steel City',
    'Marriot': 'Marriott'
}
df['city'] = df['city'].replace(location_fixes).str.strip()
df['state'] = df['state'].str.replace('Telangana',
'Telangana').str.strip()
```

Fig 1: Code for data cleaning before geocoding the locations of the clients

These steps done before starting the analysis ensured that the data was accurate and free of any discrepancy which decreased the probability of erroneous results from the analysis.

## 2.2. Methods for minimizing dead stock through demand driven inventory management

To address the problem of accumulation of dead stock for ‘Spllatter’, a systematic approach was undertaken which included the analysis of sales and purchase invoices with Excel and Python to gain insights into the problem.

### 2.2.1 Advanced Time Series Analysis

Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time. In this analysis, data points are recorded at certain intervals over a given period of time which are used to study how the variables change over time which is very useful in detecting certain trends. Time series analysis has been used to identify how the sales patterns vary which will be useful in determining the inventory levels at the different times of year.

To analyze the sales trends, the day-wise sales were extracted from the sales invoices and ban statements into Excel spreadsheets and the SUM () function was used to compute the monthly sales from July 2024 to December 2024. This spreadsheet was stored as a CSV file which was used as a dataset for analysis which has been done in Python (Google Colab notebook). The **pandas** library was used to load the dataset and **matplotlib** and **seaborn** were used for visualizing the results. For the time series analysis, the **statsmodels** library's seasonal_decompose () was used to decompose the time series data into the following constituent components: trend, seasonality and residuals. Each of the components is useful in the following ways:

Trend refers to the long-term direction of the data which means it specifies if it is increasing, decreasing or stable over a given time period. This is useful in identifying the nature of sales of the company from July to December.

Seasonality represents repeating patterns within a specific time period such as seasonal fluctuations which has been observed from the preliminary analysis mentioned in the mid-term report. This is useful for identifying if there is a particular period during which the company's sales varies significantly.

Residuals are the random variation left over after removing the trend and seasonality, and is essentially considered as noise in the data. In simple words, it is the difference between the actual data points and the values predicted by the model and are useful in determining if the model chosen (additive or multiplicative) is correct or not. Ideally, if the residuals behave like white noise, then the model is making accurate predictions.

In the time series analysis, the data is broken down into three components: -

$Y_{t\,=}$ Observed sales at time t

$T_{t\,=}$ Trend component at time t

$S_{t\,=}$ Seasonality component at time t

$R_{t\,=}$ Residual component at time t

The additive model can be expressed as:

$Y_t = T_t + S_t + R_t$

The multiplicative model can be expressed as:

$Y_t = T_t * S_t * R_t$

```
import pandas as pd
from statsmodels.tsa.seasonal import seasonal_decompose
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('darkgrid')
df2 = pd.read_excel("Analysis.xlsx",
sheet_name="advtimeseries")
forecast = seasonal_decompose(df2['sales'],
model='multiplicative', period=4)
fig, (ax1, ax2, ax3, ax4) = plt.subplots(4, 1, figsize=(12,
15))
forecast.observed.plot(ax=ax1, title="observed")
forecast.trend.plot(ax=ax2, title='trend')
forecast.seasonal.plot(ax=ax3, title='seasonality')
forecast.resid.plot(ax=ax4, title='residuals')
```

Fig 2: Code used for performing advanced time series analysis

Both these models have been used to capture seasonal variation of sales.

Advanced time series has been chosen because of the following reasons:

- It will help to identify the periods during which their sales are high/low which will help them to stock up their inventory by deciding the right order time and quantity of a certain product thus preventing the accumulation of dead stock.
- The company does not keep track of their inventory which makes it difficult to record the products available at any given time. Time series analysis works on the available data only thus eliminating the need of inventory data.
- Unlike other methods, it quantifies seasonal variation which may otherwise rely on subjective memory or experience thus providing more accurate results. Moreover, it also helps to separate noise from data in the form of residuals using seasonal decomposition.

2.2.2 Demand Variability Analysis

This method has been used to monitor how the demand for the different products varies with time. It is an extension from the ABC analysis conducted which has been included in the mid-term report. In this method, the monthly sales data as well as the data from ABC analysis were used to find the demand variability of each product. The mean sales and standard deviation in sales were computed and used to find the variability of every SKU (Stock Keeping Unit). Based on the variability, the SKUs were divided into the following categories:

1. Stable Demand (variability < 0.3)
2. Moderate Demand (variability >= 0.3 and variability < 1.0)
3. Low Demand (variability >= 1.0)

This helps in understand how the demand for each product varies which will also help in identifying if the company is stocking up more products with low demand / high variability. The coefficient of variability is calculated as:

$CV = \sigma / \mu$ where

$\sigma$ = the standard deviation in sales

$\mu$ = the mean sales

```
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import pandas as pd

data = pd.read_excel('/content/Analysis.xlsx',
sheet_name='rough')
# data.head()
route_an = data.groupby(['buyer']).agg({
    'quantity' : 'sum',
    'breakage count' : 'sum',
    'loss' : 'sum'
}).reset_index()
route_an['breakage_rate'] = route_an['breakage count'] /
route_an['quantity']
```

```
route_an['breakage_cost'] = route_an['breakage count'] *
data['rate'].mean()
```

Fig 3: Code for conducting demand variability analysis

This method has been chosen because of the following reasons:

- CV helps to identify the variation in demand which helps to determine if a particular SKU is potential dead stock or not. This can help in inventory management by stocking up the inventory with products which have moderate or low variability.
- Regression analysis was deemed unsuitable for Spllatter's inventory challenges due to transient product relevance (201 SKUs with shifting industry trends) and limited explanatory power (40% unexplained variance in residuals from times series decomposition, Fig 9). Regression requires stable predictor relationships such as fixed seasonal demands which clashes with Spllatter's 5- month dataset and dynamic SKU lineup.
- Monte Carlo simulations were not chosen because of insufficient historical data for new SKUs and urgent June stock-up deadlines (Fig 7) requiring immediate action. The breakage of goods during transportation requires real-time decisions which cannot be simulated using the Monte Carlo simulations.

## 2.3. Methods for reducing breakage of goods while transportation

The following approaches were taken to analyze the issue of breakage of goods while transportation:

### 2.3.1 Geospatial Mapping

The breakage of goods which was computed for every client in the period July 2024 – November 2024 was plotted using python libraries like **plotly_express** and **folium**. This helped to visualize the regions with high number of clients and more importantly, the regions where the breakage rate was high. The coordinates of the clients' location were plotted on a map and breakage per region was used as a metric to be visualized. This method was used because:

- Visualizing the breakage data on a map helps to identify patterns in specific regions where breakage is frequent.
- It also helps to optimize logistics performance by pinpointing high breakage zones and selecting delivery routes to minimize the breakage while transportation.

### 2.3.2 Route-wise Analysis

The route wise analysis was conducted to understand how different transportation routes impact product breakages for each client. The goal was to identify the patterns and trends in breakage incidents across various delivery paths. In this method, the data was aggregated by buyers and the order quantity, breakage count for that client and loss due to breakage were computed using Python libraries such as pandas.

```
import matplotlib.pyplot as plt
import seaborn as sns

route_an = data.groupby(['buyer']).agg({
    'quantity' : 'sum',
    'breakage count' : 'sum',
    'loss' : 'sum'
}).reset_index()
route_an['breakage_rate'] = route_an['breakage count'] /
route_an['quantity']
route_an['breakage_cost'] = route_an['breakage count'] *
data['rate'].mean()
```

Fig 4: Code to perform route wise analysis in python using pandas

The computed quantities were plotted using visualization libraries such as matplotlib and seaborn.

```
plt.figure(figsize=(30,10))
top_routes = route_an.nlargest(5, 'breakage_rate')
sns.barplot(x='buyer', y='breakage_cost', data=top_routes)
plt.title("Top 3 routes with the highest breakage rates")
plt.savefig('route_breakage_analysis.png')
```

Fig 5: Code to visualize the breakage for different clients

### 2.3.3 Root Cause Analysis using Fishbone Diagram

A fishbone diagram (also known as an Ishikawa diagram) is a visual method for root cause analysis which organizes the cause-and-effect relationships into categories. The mouth of the fishbone diagram denotes the problem which is being solved. Each of the bones feeding into the spine of the diagram represents a specific category of the potential contributors to the problem. The lines joining the problem causing factors to the spine also contain sub-branches which denote the causes of the corresponding factor. This method has been chosen because of the following reasons:

- This problem is caused due to several factors and using a fishbone diagram ensures that all possible contributors to this problem are taken into account.
- It will help to chart out a comprehensive strategy on steps which are to be taken in order to minimize the losses due to breakage of goods.
- The company does not keep track of data related to breakage such as how many items were broken which made it cumbersome to perform other kinds of analysis.

Microsoft Excel was used for creating the fishbone diagram.

# 3 Results and Findings

## 3.1 Minimizing dead stock through demand driven inventory management

### 3.1.1 Advanced Time Series Analysis

Time series analysis was performed on the monthly sales data for the year 2024 using Python. The entire data was decomposed into 3 parts namely trend, seasonality and residuals.

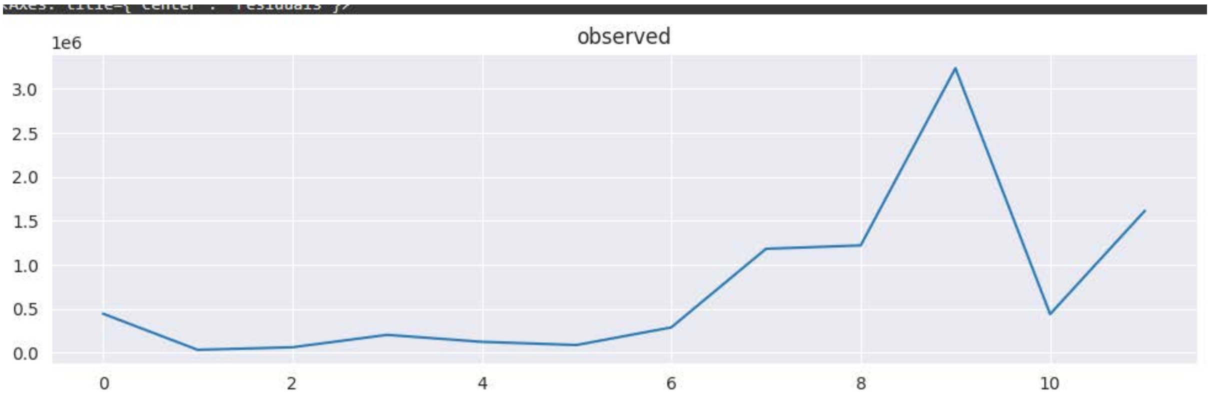



Fig 6: Plot for the observed sales throughout 2024

In the above graph (Fig 6), the x-axis denotes the month and the y-axis denotes the sales per month (in millions). It can be seen that the sales drops after January and increase slowly till July where there is a spike. The company records its highest sales in October which is because of festivals like Diwali and Durga Puja. Following this the sales drop in November. Since a major portion of the company's sales happen in July-October, the company should stock up their inventory in June to cater to the demands of clients and also prevent dead stock accumulation.

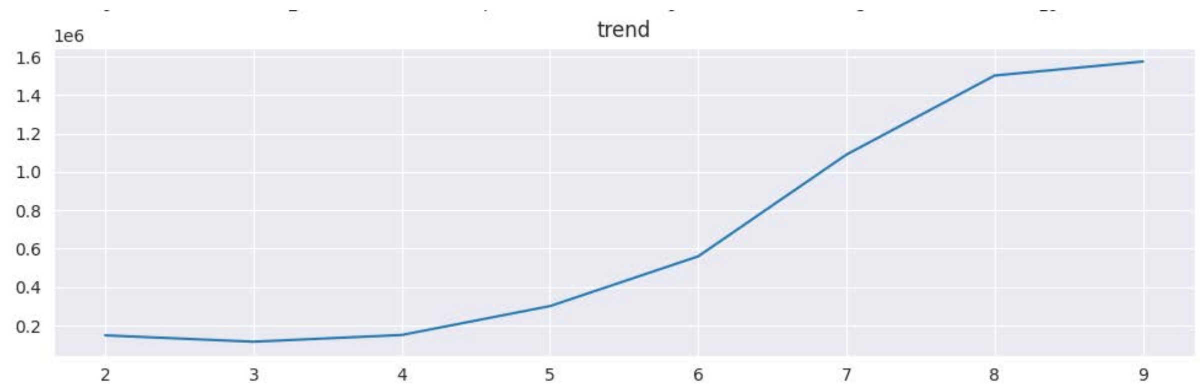



Fig 7: Observed trend after removing residuals and seasonality

The trend plot (Fig 7) shows the long-term pattern of sales after removing the residuals and seasonality components. The plot shows that the sales are very low in the first half of the year as compared to the second half of the year. This is again due to the fact that the second half (July – December) has a lot of festivals as well as wedding season which requires many hotels and resorts to buy cutlery, tableware and other such products from sellers like Spllatter.

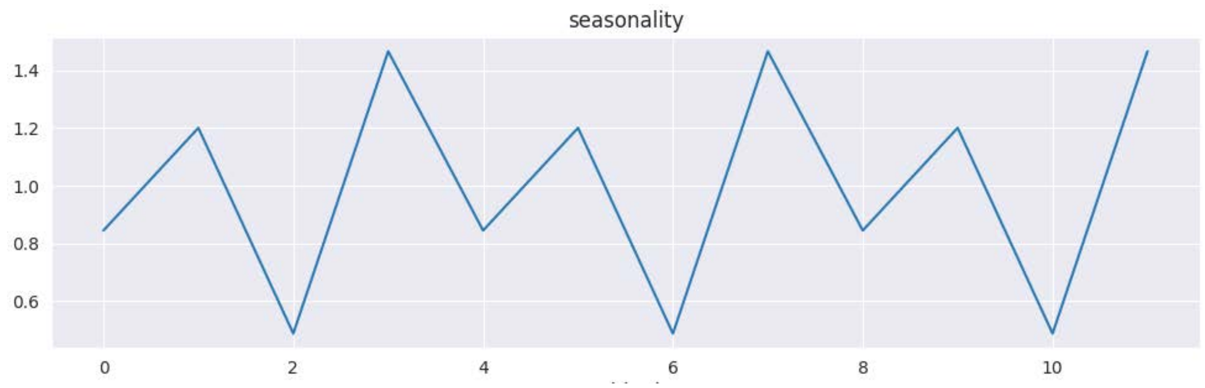



Fig 8: Seasonality plot

The above plot (Fig 8) exhibits a clear repeating pattern which suggests that the sales are affected by seasonal factors quarterly (period=4). The seasonal index ranges from 0.6 to 1.4 meaning that the sales fluctuate by ~40% due to seasonality. When the seasonal index > 1, sales are higher than average during that period and similarly, sales are below average when seasonal index < 1. This indicates that the company's sales are highly seasonal.

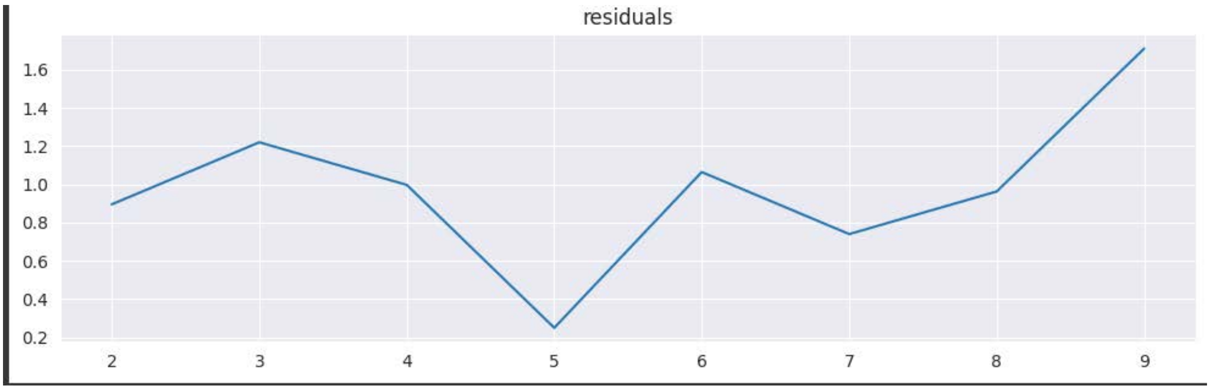



Fig 9: Residual plot

In time series analysis, residuals represent the unexplained variation in the sales after removing trend and seasonality (Fig 9). Here, the x-axis represents the months whereas the y-axis represents the deviation between the actual sales and the sales predicted using the seasonal decomposition model. Large spikes in the plot suggest that the model failed to consider key demand factors such as festive seasons. This indicates a risk of stockouts as since the model underestimates the sales during the peak sales period.

### 3.1.2 Demand Variability Analysis

The variability in demand for every SKU was computed after ABC analysis was performed. The objective of using ABC analysis was to identify which products contributed most to the revenue and demand variability was computed to see how the demand for a particular SKU varies.

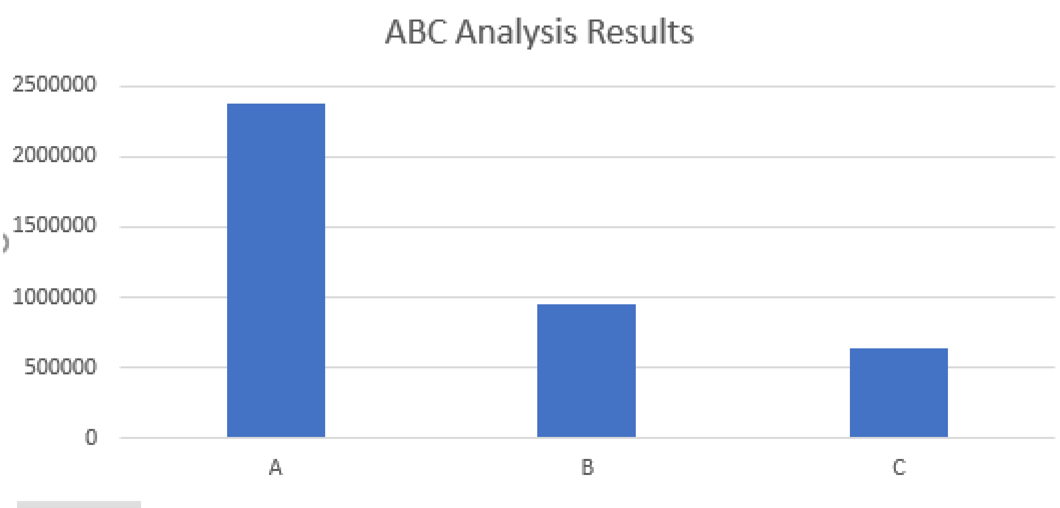



Fig 10: Results of ABC analysis

ABC analysis was performed for 201 different SKUs and the categorization was done as follows:

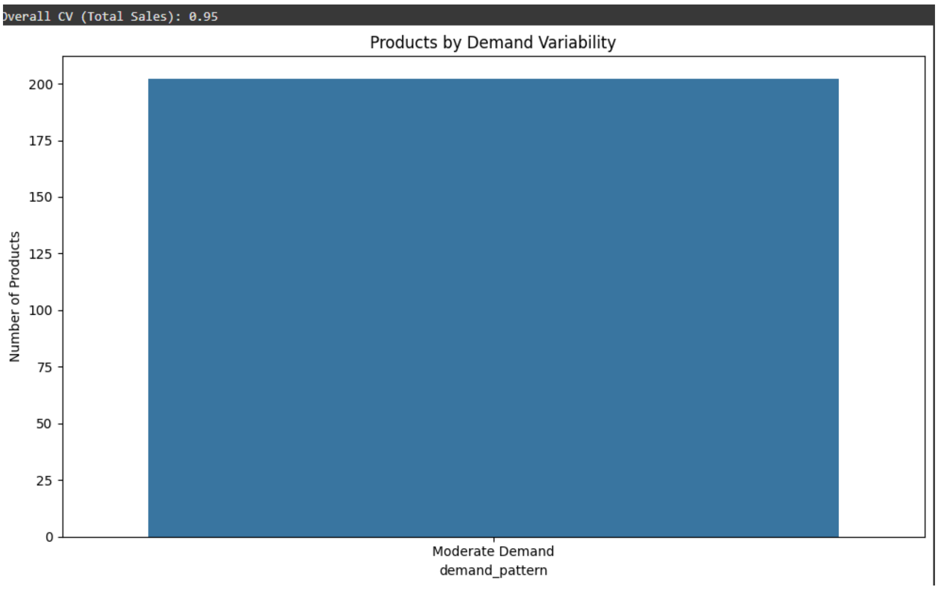
Top 60% of sales – Category A

60% - 85% of sales – Category B

80% - 100% of sales – Category C

From the ABC analysis (Fig 10), it was found that 86 items belonged to group A, 60 items belonged to group B and the remaining belonged to group C. The graph (Fig 5) shows that items in group A contributed the most to the revenue. Hence their variability should be low.

To confirm the ABC categorization of the SKUs the demand variability was computed and the results of the same were as follows:




| demand_pattern | count |
|---|---|
| Moderate Demand | 202 |

dtype: int64

Fig 11: Demand variability analysis results

The demand variability (CV) was computed with the help of the mean sales and the standard deviation in sales. If CV < 0.3, then the SKU is said to have stable demand, if CV between 0.3 and 1.0 then the SKU is said to have moderate demand otherwise is it said to have high variability (or consequently low demand). It was found that all the products have a demand variability of 0.95 which places them in the category of moderate demand (Fig 11). This indicates that the since the company has a large number of SKUs, there is no such SKU which is very popular or unpopular. Thus, stocking these SKUs in large quantities may lead to accumulation of dead stock.

### 3.2 Minimizing losses due to breakage of goods during transportation

3.2.1 Geospatial Mapping

The coordinates of all the clients were geocoded using Nominatim and were plotted on a geospatial map using plotly_express and folium.

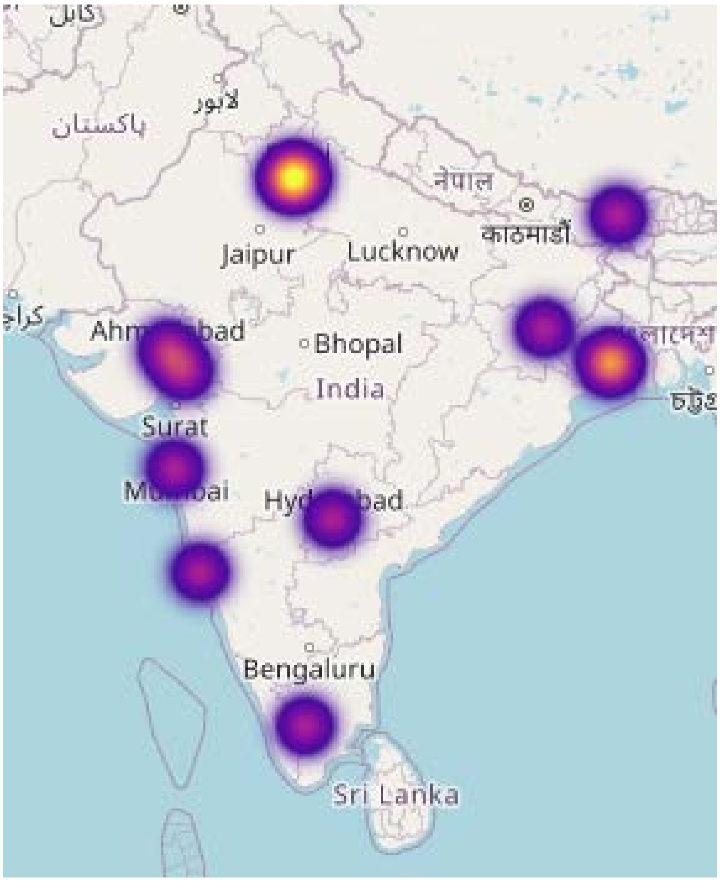



Fig 12: Heatmap denoting the client distribution across the country

The above map (Fig 12) was constructed using plotly_express to visualize the distribution of the company's clients across the country. It was found that a major part of the clients are from northern regions like Delhi and Haryana followed by western regions like Mumbai and Gujarat. Using this information, the breakage rates were also plotted on a geospatial map for each client to see how the breakage varies with client locations and identify trends is loss incurred by the company due to breakage while transportation.

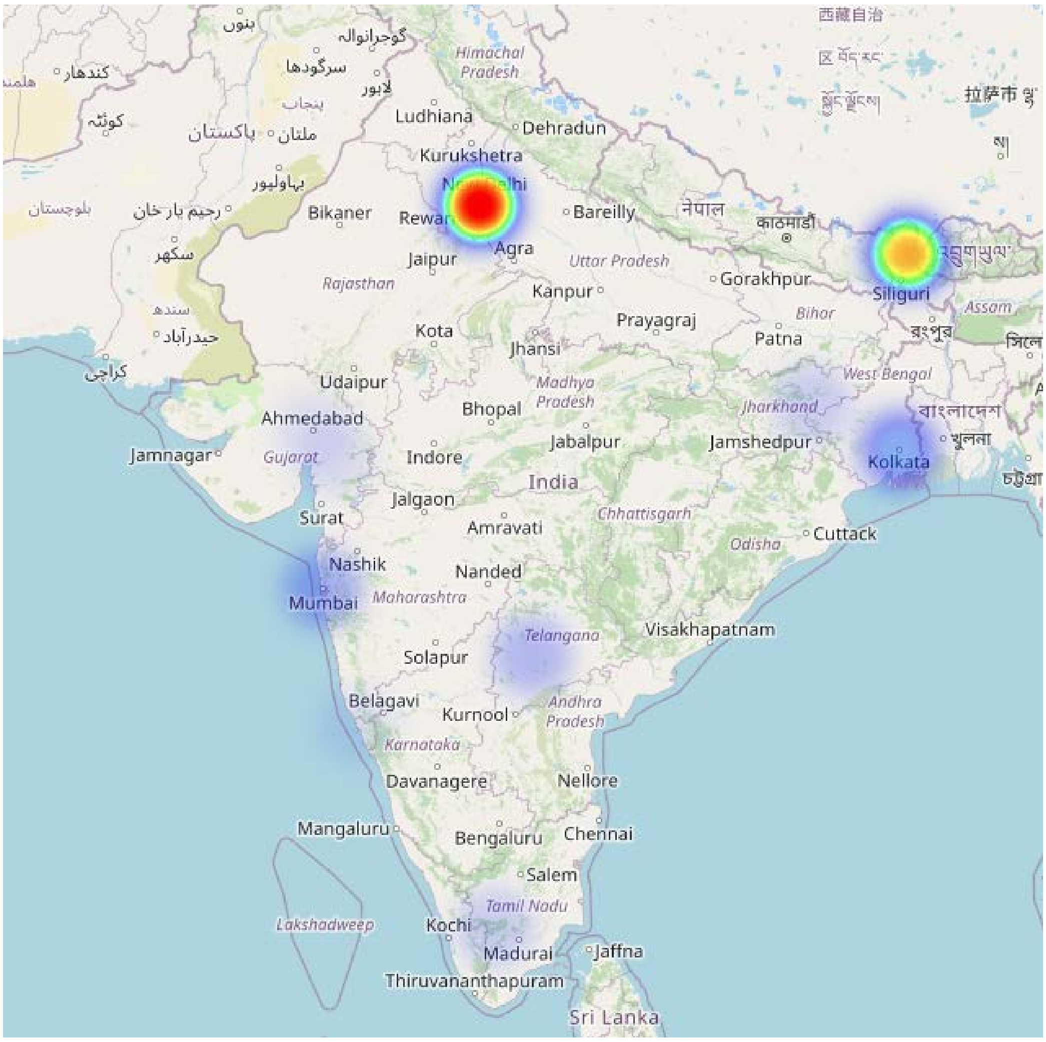



Fig 13: Geospatial mapping of clients to analyze breakage

In the above map (Fig 13), the highlighted regions show the breakage for clients at that location. Lighter colors (purple, blue) denote less breakage whereas darker colors (green, orange, yellow) denote high breakage. It can be observed that the breakage rate is comparatively very high in the northern and western parts of the country as compared to the southern and western parts. This can be because of factors such as rugged terrain in the north-eastern regions, improper packaging for highly fragile items and extreme climatic conditions (heavy rainfall). This observation is confirmed from the previous analysis in the midterm report where the highest breakage was found to be in the northern and western regions of the country.

### 3.2.2 Route-wise Analysis

In this method, the loss due to breakage was computed for different clients in order to get an idea about high-risk regions.

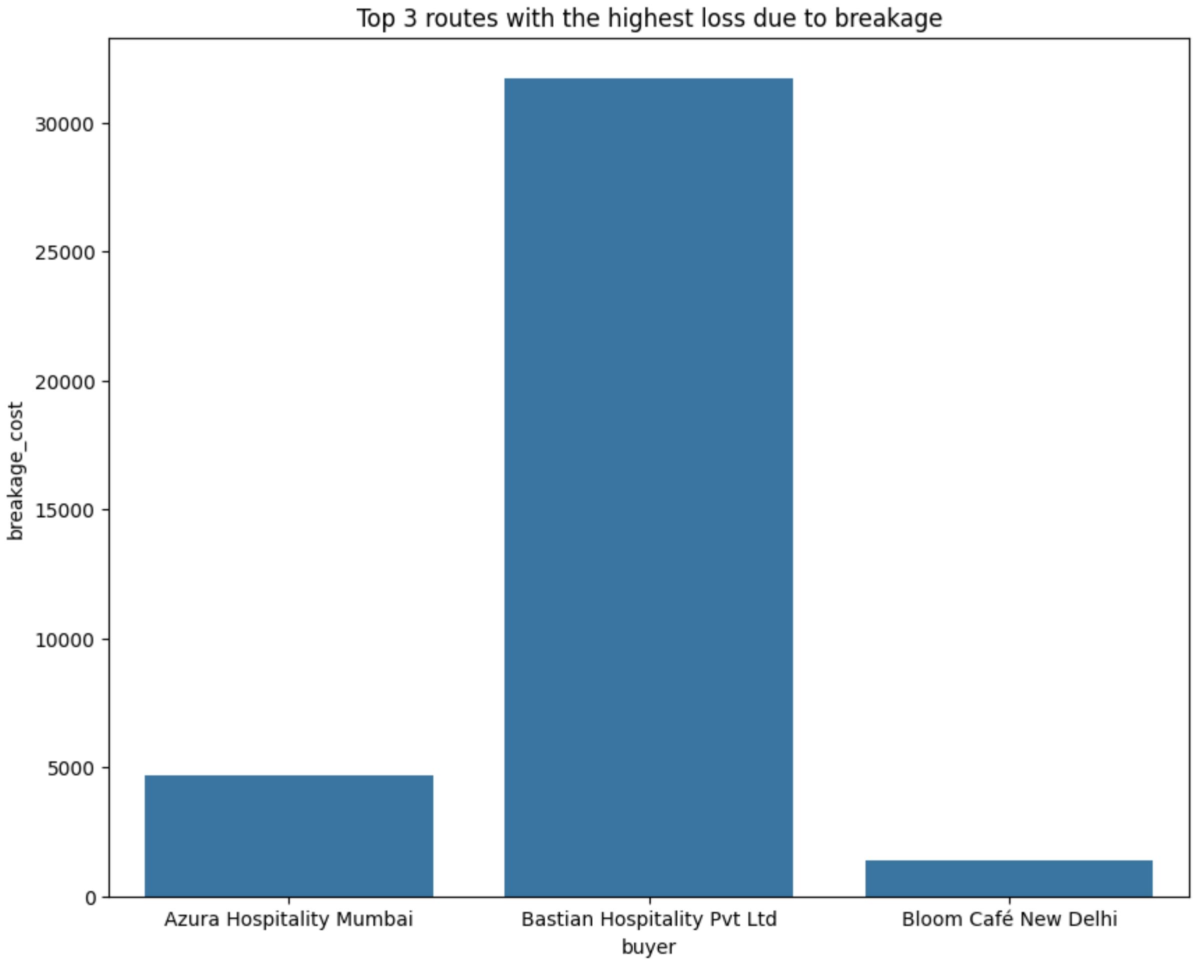



Fig 14: Bar chart representing the top 3 clients with highest loss due to breakage

The above graph (Fig 14) shows the top 3 clients for whom the loss incurred by the company due to breakage of goods while transportation is the highest. An important fact is that the above clients are located in the western region where the breakage is lower as compared to the norther regions but higher when comparted to the southern and eastern regions. It can be concluded that these clients order items which are expensive as compared to the items ordered in other regions of the country or their order size is higher than other clients owing to which the loss due to breakage is very high. Some factors contributing to this are improper packaging of expensive items or manhandling during transportation. Since this is a central part of the country, the roads are well developed so rugged terrain and extreme climatic factors are ruled out in this case.


Fig 15: Regression between quantity of item ordered and loss due to breakage

Another important observation brought forth by this observation (Fig 15) was that there is a positive correlation between the ordered quantity of an SKU and the loss due to breakage. This suggests that products which are ordered in large quantities are more prone to damage because of improper packaging and hurried transportation.

### 3.2.3 Root Cause Analysis using Fishbone Diagram

The fishbone analysis was used to systematically identify and categorize the root causes of problem of breakage of goods while transportation. The primary focus was on understanding breakage trends, packaging failures and inefficiencies in transportation routes. The fishbone diagram was constructed with the help of findings from the aforementioned methods as well as the information gathered during in-person discussions with the company. There were 4 key areas which were identified to be responsible for breakage of goods.

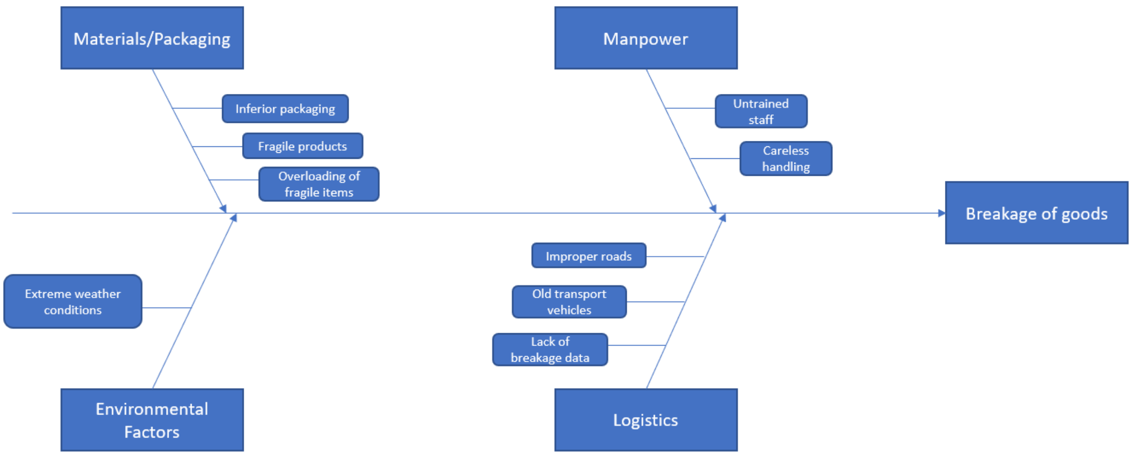



Fig 16: Root Cause analysis using fishbone (Ishikawa) Diagram

- Materials and Packaging: As seen from the above analysis methods, the highest breakage loss is due to inadequate packaging for large orders and on orders with expensive products. Overloading of items especially very fragile goods puts them at a high risk of breakage.
- Manpower
  Another reason for breakage of goods during transportation is careless handling of goods. This included throwing of cartons containing the products on the ground and the delivery being transferred through many hands which significantly increases the risk of breakage. It was also identified that the company uses some local transporters for their delivery purposes who are not experienced in such transportations which also contributes to the breakage.
- Environmental Factors
  Around monsoons (during July-August), the demand in sales increases (as seen in time series analysis) and thus the products being transported are at a risk of breakage due to environmental factors such as torrential rains and slippery roads. Improper packaging for goods in this time also incurs heavy loss due to damages.
- Logistics
  Another important factor contributing to breakage is faulty logistics practices such as making deliveries through routes which have uneven terrain, using old vehicles for transportation and most importantly not maintaining breakage data which can help the company in identifying such issues and taking appropriate measures to improve them.

# 4 Interpretation of Results and Recommendations

**Interpretation of Advanced Time Series Analysis:**

From the time series analysis conducted, it was found that the sales are low from February to May and they start increasing from June. From the trend plot, it is observed that the sales continuously increase throughout the year. These observations indicate that majority of the sales start increasing from June and hence the company should stock up their inventory accordingly.

**Recommendations:**

The company should stock up their inventory with the SKUs having moderate demand at the beginning of the year. SKUs which are expensive and ordered in large quantities should be stocked up in May-June as their demand is during the period of July-December. During this period, the items with moderate variability should not be kept in inventory as they will become dead stock.

The company should make use of an inventory management system as the above analysis is purely based on historical sales data and the same patterns may or may not followed in coming years. Moreover, the sales trend is not very reliable as the company was started in 2022 and is relatively new so the trend cannot be assumed to repeat every year.

**Interpretation of Demand Variability Analysis:**

This is a continuation of the ABC analysis conducted to identify the high value, mid value and low value SKUs of the company. The ABC analysis was conducted as described earlier because the company has a large number of SKUs and they keep changing with time according to the general trend in the industry. Although 86 out of 201 items contributed to 60% of the total sales, all the products were found to have moderate demand which means that there is no guarantee that they will sell in very high or very low quantities. This is an important observation as it suggests that the company's SKUs are very diverse due to which there are no items which are expected to generate high revenues every year. The company orders products based on the industry trends which keep changing with time and hence have no fixed products. It is very much possible that the SKUs which generated the most revenue in 2024 may not sell at all or sell in very less quantities in 2025 or any of the following years and thus become dead stock.

**Recommendations:**

The company should limit their SKUs as this will help them have a fixed product base and also help them generate a certain revenue every year. These SKUs should be chosen carefully after thorough research and trials.

The company should work on their research methodology and techniques to identify which product is to be ordered next. This will help them in accurately identifying the potential revenue generating SKUs and also prevent accumulation of dead stock.

**Interpretation of Route-Wise Analysis, Geospatial Mapping and Fishbone Diagram**

Route – Wise analysis, geospatial mapping and root cause analysis through fishbone diagram were used collectively to gain a deeper insight into the problem of breakage of goods while transportation. The above methods pointed out several factors which lead to breakage. The primary contributor to breakage while transportation is improper packaging of products. While the company has taken steps to reduce their breakage in the past, these steps are not very efficient as the breakages still occur and imposes heavy losses on the company. Some discrepancies in packaging include not using adequate cushioning materials like polystyrene, thin bubble wraps. Another important contributor to breakage is the choice of delivery through risky routes such as rugged terrain. All the deliveries are made by road only. Other factors leading to breakage include manhandling of goods and delivery through small local agencies where the staff is not trained to do the deliveries. Lastly, no breakage data is maintained which can prove very useful in identifying which delivery agencies are suitable for deliveries and also help in improving packaging for the damaged products in upcoming deliveries.

**Recommendations:**

The company should use high quality and adequate amount of packaging material to prevent breakage of goods.

To address high breakage rates, alternative transport methods like airways or railways should be considered. While they are more expensive as compared to roadways, they offer significantly lower breakage risks. These options should be considered based on the shipment size, product fragility and client urgency to maintain a balance between cost and safety.

To reduce breakage more effectively, a dedicated logging system should be implemented which tracks SKU type, packaging details, delivery route and breakage. This data will help in identification of patterns thus enabling targeted and efficient future improvements.

# 5 Additional

Dataset link:
https://drive.google.com/drive/folders/1x_KrXi5O_MzlBqnUPGaSKFivGWDYA6ES?usp=sharing

Analysis link:

https://docs.google.com/spreadsheets/d/1rVJBCeJG3Ydzf4g423COnRVWDMklCk8-/edit?usp=sharing&ouid=111809023545961421534&rtpof=true&sd=true

Google Colab Link:

https://colab.research.google.com/drive/1oOrKyo3fiRkh9YjH4sgyixIeTvYGOPoM?usp=sharing