

**Payment Pattern Analysis and Optimization of Inventory
for a Packaging Machines supply Company**

Mid Term report for the BDM capstone Project

Submitted by

Name: Suneha Datta

Roll number: 23f1002574



IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

Contents

1. Executive Summary	4
2. Proof of Originality of Data	4
3. Details about the data received from company	
3.1 Metadata	4-5
3.2 Descriptive Statistics	5-6
4. Detailed Explanation of Analysis Process	
4.1 Data Cleaning and Preprocessing	6-7
4.2 Correlation Analysis	7
4.3 Pareto Analysis	7
4.4 Monthly Dues Analysis	8
4.5 Anticipatory Stock Analysis	8
5. Results and Findings	
5.1 To understand the overdue payments	8-10
5.2 To understand the Inventory purchase	10

Declaration Statement

I am working on a Project Title “Payment Pattern Analysis and Optimization of Inventory for a Packaging Machines supply Company”. I extend my appreciation to **S.S. Engineering Services**, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate:



Name: Suneha Datta

Date: 3.3.2025

1 Executive Summary

In this report, the business data collected from S.S. Engineering Services, a packaging-based company located at Mahestala, Kolkata, is analyzed to gain insights about the shortage of money faced by the company to buy products for clients. Two datasets have been procured from the company, as Excel files, where one is related to the sales of product items in the financial year April 2024 to March 2025, and the other is the inventory purchase data in the same year. The data is analyzed as per the problems shared by the company proprietor through a remote conference.

Initially, a descriptive statistics of the data received is shown, which includes the mean, standard deviation, minimum and maximum values, twenty-fifth and seventy-fifth percentiles of all the numerical columns present in the dataset to interpret the ranges as well as the nature of the values that each column can take.

The data is preprocessed and then studied using different Python libraries like Pandas, Matplotlib, and Seaborn to find out details that suggest the reason for the money shortage and the overdue payments in the business. I have performed a Correlation analysis to understand the relationship between the balance amount remaining with the company, the sales revenue, and the overdue payments. A Pareto analysis is done to find out the customers who are the most responsible for the sales, followed by a monthly dues analysis, which shows the customers who are the maximum defaulters over the months. Also, an inventory analysis is done to compute how many units of products are needed to be stocked up and at what time of the year, so that even if there is a shortage of money but the supply workflow can be maintained efficiently.

The findings from these analyses are all plotted as representative graphs, like bar charts, using Python to correlate the results with proper pictorial visuals. This helps the company to make suggestive decisions in order to increase operational productivity.

2 Proof of originality of the Data

The original data, letter from organisation, meeting with the proprietor is present in the link below.

<https://drive.google.com/drive/folders/1dshAHEjAF2rYfG-yPVJ4k9n5zk1ggaRG?usp=sharing>

3 Details about the Data Received from Company

3.1 Metadata:

The organisation has given two excel sheets as a dataset through email(online), where one sheet named inventory purchase data signifies the data about the inventory purchased across 12 months for different product items, and the sheet named outstanding sales data is the data about the already received money for the sell of respective products to customers and the money that is yet to be received from those customers across 12 months data for the financial year 2024-2025(April – March) .

3.1.1 Outstanding sales data:

It has a total of 96 entries with no duplicate contents.

It contains columns such as -

COLUMN NAME	DESCRIPTION	DATATYPE
Invoice No	It is the bill number for each transaction	string
Invoice Date	It contains the issued invoice date	date
Product	It gives details about the sold products.	string
HSN Code	It contains the alphanumeric code for taxation of products.	string
Qtn	It is the number of the product sold for that invoice	integer
Customer	It describes the details of the purchased customer in text format.	string
Party GST No.	It contains the buyer's GST identification number.	string
Amount	It is the numerical price value of the product before applying tax.	float
IGST	It is the applied tax amount for selling in the outer state.	float
CGST	It is the applied tax on the selling product inside the state by Central Government.	float
SGST	It is the applied tax on the product shared with the State Government.	float
Freight	It is a numeric value that deals with the shipping or delivery charges.	float
Total Amount	It is the final value of the product after adding taxation and shipping charges.	float
Due	It gives the details of remaining amount to be paid by the customer.	float
Status	It is a string value containing the current payment situation.	string
Cheque No.	It is the Cheque number used for payment.	string
Cheque Date	It is the date when the cheque was issued for the payment.	date

Fig 1. Details about columns present in Outstanding Sales data

3.1.2 Inventory purchase data:

It has a total of 44 entries with no duplicate contents.

It contains columns such as – (After Cleaning)

COLUMN NAME	DESCRIPTION	DATATYPE
Invo Date	It contains the issued invoice date.	date
Item Code	It contains the alphanumeric code of products purchased.	string
Quantity	It is the number of the product sold for that invoice.	integer
Basic Rate	It is the cost price value of the product before applying tax.	float
Sub Total	It is the product of Basic Rate and Quantity.	float
IGST	It is the applied tax amount for selling in the outer state.	float
Total Amt in Rs	It is the final value of the product after adding taxation of IGST.	float
Product	The category of the product.	string

Fig 2. Details about columns present in Inventory Purchase data

3.2 Descriptive Statistics:

3.2.1 Outstanding sales data:

	Qtn	IGST	Total Amount	Due
mean	61.604167	826.159787	74117.04323	14391.48958
std	62.681396	2401.469995	73904.12186	35399.03268
min	1	0	0	0
25 percentile	8.75	0	9398.11	0
50 percentile	44	0	44025.8	0
75 percentile	100	0	127484.25	46
max	200	13500	330400	171100

Fig 3. Descriptive statistics of Outstanding Sales data

The maximum quantities of any product ordered by any client is 200 and the minimum number of products ordered is 1 which was a fully automatic Machine order. The mean ordered quantity is around 61. The maximum bill amount among clients in the year FY 2024-2025 is Rs 3,30,400. The order amount typically is around Rs 73,856 with a deviation of about Rs 72,090. The due amount of bill money ranges from Rs 0 to Rs1,71,100. The average due bill amount is of Rs 14,697, with a deviation of Rs 35,714.

3.2.2 Inventory purchase data:

	Quantity	Basic Rate	Sub Total	IGST	Total Amt in Rs
mean	118.090909	5883.329755	77774.50909	13999.41182	91773.91636
std	66.674432	28501.29601	39486.6592	7107.5983	46594.25197
min	1	340	6490	1168.2	7658.2
25 percentile	70	444	47603.5	8568.63	56172.13
50 percentile	100	651.134615	73675	13261.5	86936.5
75 percentile	200	735	105332.5	18959.85	124292.35
max	230	185000	185000	33300	218300

Fig 4. Descriptive statistics of Inventory Purchase data

The number of items purchased is about 118 on an average. The overall purchased items' quantity has been in the range of 1-230. The mean basic rate of products is Rs. 5883.33 with a deviation of Rs 28,501, that is the basic rate is widely spread distribution. Although the maximum base rate of products is Rs 1,85,000 and the minimum base rate being Rs 340 which is mainly for purchase of spare parts. The final purchase price for an invoice range from Rs 7658 - Rs 2,18,300. The mean cost price being Rs 91773.92 with a standard deviation of Rs. 46,5964.

4. Detailed Explanation of Analysis Process

4.1 Data Cleaning and Preprocessing

At first, I loaded the two data files into Excel and then I have removed columns from the Outstanding sales data which are not relevant to the analysis that includes HSCN code, Party GST No., Freight, Status Cheque No., Cheque Date, Bank. In Inventory purchase data I have collabed all three parties' Signor, SIA pack, Quality Tech, and merged columns like Invo Date, Item Code, Quantity, Basic Rate, Sub Total, IGST, Total Amt in Rs, Added New column 'Product' which is to denote the 'Product' column of the outstanding sales data. This helped me to structure the data in a proper format in order to do further analysis. After which I checked for all mismatched cell values and removed blank rows from both the sales and inventory sheets and corrected spellings across every column and converted all cell values to their respective type or cell data type. Then I converted the files into csv for further analysis. In google colab, I used the pandas library of python to load these csv files as pandas dataframe. Afterwhich, I checked for any missing values as well as NAN values present in the dataset.

Purchase 2024-2025														
Sl. No.	Invoice No.	Invo Date	Partys Name	Item Code	Quantity	Basic Rate	Sub Total CF	IGST	CGST	SGST	Total Amt in Rs	Paid Amt	Cheque No.	Date
	pp/212	20/4/2024	WRELT INDIA PVT LTD	1206 fa kdm	200	623.70	124740.00	22453.20			394297.00			
				960 FA V	100	1050.00	105000.00	18900.00			147193.20	200000.00	348077	26/4/2024
											123900.00	80535.00	348081	6/5/2024
	P/317	2/5/2024	WRELT INDIA PVT LTD	1260 ER	72	661.50	47628.00	8573.04			56201.04			
				5mm	40	790.00	31600.00	5688.00			37288.00			
												113762.00	348085	14/5/2024
	P/334	3/5/2024	WRELT INDIA PVT LTD	1206sa kdm	150	444.00	66600.00	11988.00			78588.00	261093.00	348094	6/6/2024
					0		0.00	0.00			0.00	100000.00	381796	13/6/2024
							0.00	0.00			0.00	82077.00	381798	18/6/2024
							0.00	0.00			0.00	100000.00	381803	3/7/2024

Fig 5. Inventory Purchase data before data preprocessing

Invo Date	Item Code	Quantity	Basic Rate	Sub Total	IGST	Total Amt in Rs	Product
20/04/2024	KDM FA	200	623.70	124740.00	22453.20	147193.20	PP Strap
20/04/2024	960 FA V	100	1050.00	105000.00	18900.00	123900.00	PP Strap
02/05/2024	1260 ER	72	661.50	47628.00	8573.04	56201.04	PP Strap
02/05/2024	5 mm	40	790.00	31600.00	5688.00	37288.00	PP Strap
03/05/2024	KDM SA	150	444.00	66600.00	11988.00	78588.00	PP Strap
27/05/2024	1206 FA	200	682.5	136500.00	24570.00	161070.00	PP Strap
20/06/2024	KDM FA	200	623.7	124740.00	22453.20	147193.20	PP Strap
20/06/2024	KDM SA	200	444	88800.00	15984.00	104784.00	PP Strap
20/06/2024	960 FA V	70	1063.12	74418.40	13395.31	87813.71	PP Strap
02/07/2024	Stout 1260	100	735	73500.00	13230.00	86730.00	PP Strap
02/07/2024	KDM SA	200	444	88800.00	15984.00	104784.00	PP Strap
08/07/2024	KDM SA	150	444	66600.00	11988.00	78588.00	PP Strap
08/07/2024	Bosch 1260	50	509.2	25460.00	4582.80	30042.80	PP Strap
14/07/2024	KDM SA	200	444	88800.00	15984.00	104784.00	PP Strap

Fig 6. Inventory Purchase data after data preprocessing

4.2 Corelation Analysis

I have calculated the monthly balances and the monthly purchased amount, where the monthly balance is computed as Outstanding Sales - Due amount - Purchased Price. I have modified the balance to add money from the capital whenever there is a purchase price shortage, only when there is enough capital available. Then I computed the correlation using python Pandas library's 'corr' function between balance due and purchased price to check whether it is the monthly dues for which the purchasing power is decreasing even if the year starts with a capital. I have used the Pearson correlation formula where the correlation between two continuous variables x and y is calculated as $r = \frac{\sum(x_i - x_m)^2 \cdot \sum(y_i - y_m)^2}{\sum(x_i - x_m)(y_i - y_m)}$ where y_m and x_m are the corresponding mean values.

4.3 Pareto Analysis

To carry out the Pareto analysis for finding out the top customers who are actually contributing to the sales, at first, the outstanding sales data is grouped by customers in order to find the total sales per customer using the Python pandas library's groupby function on the outstanding sales data frame. Then the grouped data is sorted in descending order. Thereafter, the cumulative percentage of the total sales for each customer is calculated by dividing the individual sales by the total sales amount to get the relative contribution of each of them. Then the top 20 % of the customers are selected, contributing to 80% of the sales. To get a visual representation of this Python matplotlib library's bar plot is used to plot the cumulative sales as bars representing the values on the y axis and the names of the respective customers on the x axis.

4.4 Monthly Dues Analysis

For calculating the monthly dues to check which customers are likely to make dues in any month. At first, the column 'Invoice Date' of the outstanding sales data is converted to a proper date-time format and any rows which has missing dates are removed. Then, a new column is created that has the extracted month and year for the respective invoice dates for each row. After which, using the Python pandas groupby function, the due amounts have been grouped per month and then again by customer, following which the due values are summed up for every month. Then the months where there have been no due payments or the customers who have no dues are both removed from the grouped dataframe. Next, using Python's Seaborn library, this grouped data is being plotted as a horizontally stacked bar graph, where the y axis depicts the monthly due amounts in rupees and the x axis represents the months. Using a legend, the stacked bars for a particular month are differentiated for various customers present.

4.5 Anticipatory Stock Analysis

At first, the two datasets as the Outstanding Sales data and the Inventory purchase data, have been modified together. The columns like 'Invoice Date', 'Qtn' in the first dataset, and the 'Invo Date', 'Quantity' are renamed to 'Date' and 'Quantity' simultaneously. Then the dues column present in the first dataset is grouped by month using Python's pandas library, and alongside the latter dataset's Quantity column is grouped by Product and month. After that, the date in the first grouped data is shifted to one month ahead and labeled as next month due. Now, based on the months where there are dues greater than Rs 50,000 as well as there has been sales of any products, are segregated. The quantity of units that is hence sold in the next months of the segregated months is collated in the second dataset. This filtered second dataset is then plotted as a horizontally stacked bar graph using Python Seaborn library, where the x axis corresponds to the month and the y axis represents the units of product that is needed to be stocked up in that month.

5. Results and Findings

5.1 To Understand the overdue payments:

5.1.1 Correlation Analysis

	purchase	balance	dues
purchase	1.000000	-0.177518	-0.147191
balance	-0.177518	1.000000	-0.654679
dues	-0.147191	-0.654679	1.000000

Fig 7. Correlation Matrix

From the previous realisation that there must be potential correlation between the purchase price, balance and due amounts as seen in the graph I have calculated the correlation between these variables grouped by months. The purchase amount and the dues are not related so closely whereas the balance and the dues are related with correlation >0.5 but they are related inversely that is if the balance increases the dues decreases and vice versa, so it can be stated that it is

because of the high uncollected due payments which is causing the balance to be so low in some months.

5.1.2 Pareto Analysis for finding most contributing customers

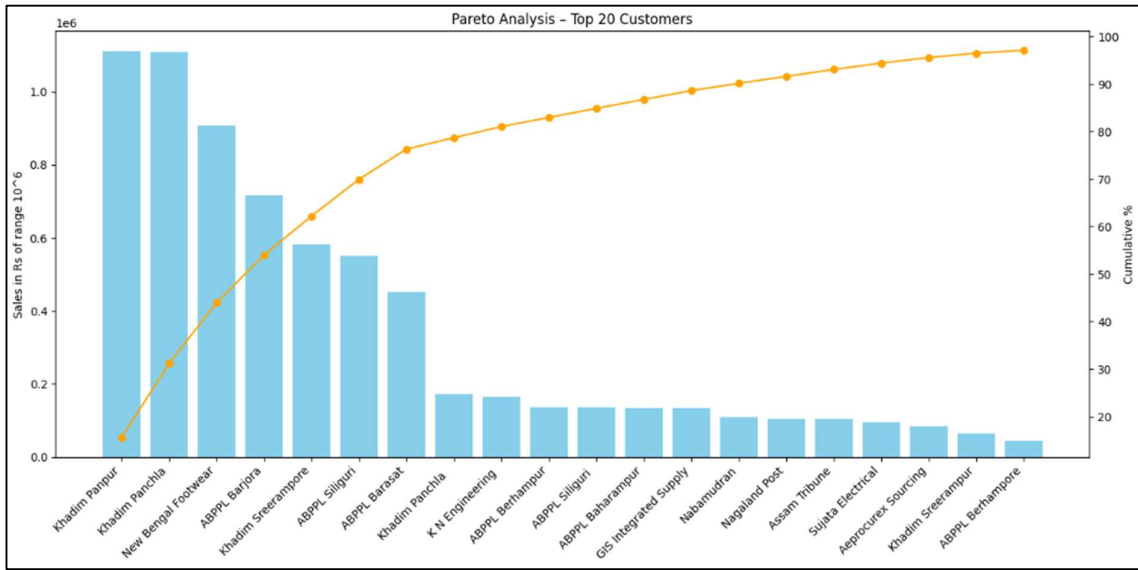


Fig 8. Top 20 customers contributing to 80% of sales

In the chart above, the blue bars correspond to the total sales made by each of the customers, while the orange line corresponds to the relative percentage of the total revenue. As evident, only a small group of customers like the 'Khadim Panpur', 'Khadim Panchla', 'New Bengal Footwear', and 'ABPL Barjora' produce the maximum sales uniquely. The initial sharp rise in the trend line depicts the dominance of these high-value customers, while the flat tail shows the small returns made by the remaining customers among the top 20 of them. Therefore, these customers should always be given at least a small favour over others and also some acceptance on keeping outstanding dues, whereas from the rest customers, the company should ask for advanced payments in case of dues, since their contribution to the sales is disproportionately lower.

5.1.3 Monthly Dues Analysis

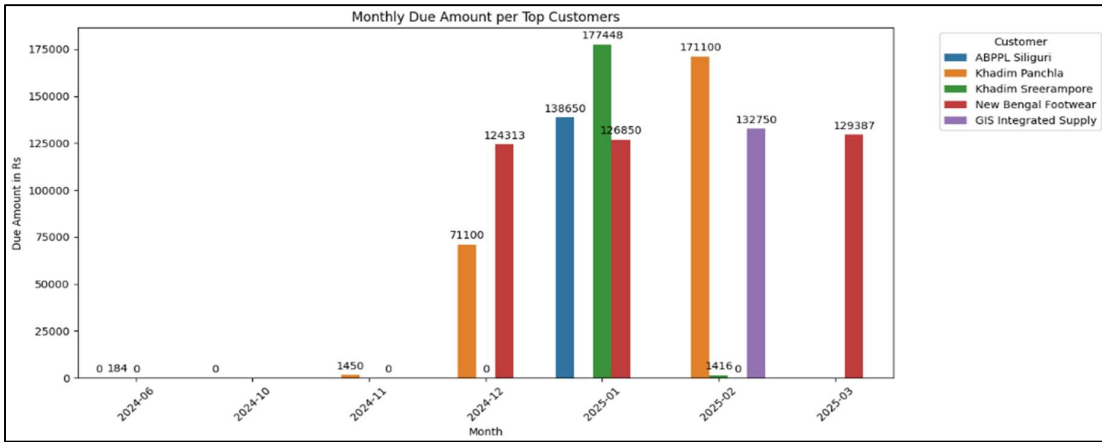


Fig 9. Monthly Dues per customer chart

From the chart above, it can be seen that during the months of December to March, the due payment amounts rise quite as high as Rs 1,77,000 approximately, which is alarming. The major dues were done by customers like 'Khadim Sreerampore', 'Khadim Panchla', and 'ABPPL Siliguri' of about Rs 1,77,448, Rs 171,100, and Rs 1,38,650 respectively. Also, it can be observed that the customer 'New Bengal Footwear' consistently keeps a large amount as dues during the months of December, January, and March. The customer like 'GIS Integrated Supply' has also kept up in February. It can be stated that these small revenue contributing customers, as seen in the previous charts, like 'GIS Integrated Supply', 'Khadim Sreerampore' and 'ABPPL Siliguri' should be persuaded for advanced payments since they are more likely to make high dues, though their purchase value over the year seems low. On the contrary, customers like 'New Bengal Footwear' and 'Khadim Panchla' can be considered a little since even if they keep dues, they are the most revenue-generating clients. Since it's evident that during the second half of the financial year, there is the highest probability of high dues, the company should be prepared beforehand to stock up on inventory enough in the first half of the year.

5.2 To Understand the Inventory purchase:

5.2.1 Anticipatory Stock Analysis

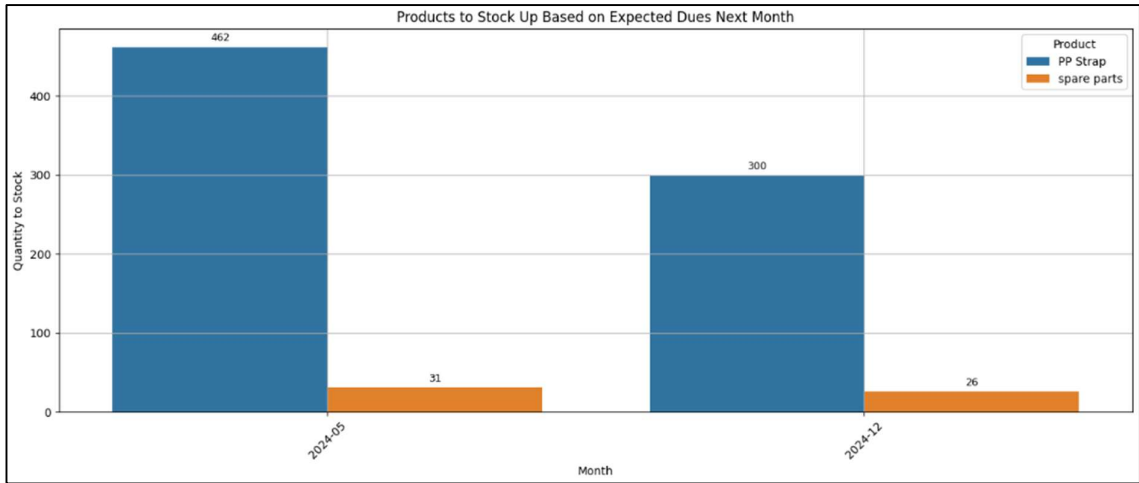


Fig 10. Suggestive units of inventory to stock up

The chart above shows that the months of May and December are the months in which the company should stock up its inventory, since it's likely that, due to a shortage of money in the upcoming month, there will be a constraint on money to buy new items. The need for PP straps is far more than the demand for spare parts. If the company stocks up around 462 units of PP straps and 31 units of spare parts in the month of May, followed by 300 units of PP straps and 26 units of spare parts in the month of December, then in the next month, all orders can be satisfied efficiently. It is evident that PP straps drive most of the business, and hence, to overcome any shortage of it and also to avoid monetary issues for its purchase, it's better for the company to buy the stated units of PP straps in the months shown for a smooth flow of supply.