Navigating B2C Expansion and Optimizing Warehouse Efficiency Amid Labour Shortage

Mid-term report for the BDM capstone Project

Submitted by

Siddharth Pandey

22F2001147

IITM BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1. Executive Summary

SNP is a B2B manufacturing firm based in Maharashtra. The firm engages in the production of industrial equipment such as ladder type cable trays, APFC panel, MCC panel, etc. The demand of the products manufactured by SNP is majorly from electrical firms, hotels, malls, factories and builders, etc. The company employees a total of 12 permanent workers along with some others hired on a daily wage depending on seasonal trends in demand for the products.

Though currently profitable, the company has been facing the challenge of being able to continue to produce and sell at its current profitability amidst labour shortage in the manufacturing industry. This labour shortage has been impeding the firm's growth, thereby hindering its aspiration to expand into sectors like B2C.

This project aims to provide potential solutions to alleviate the problems of the firm with the help of extensive data analysis methods. For this purpose, the data pertaining to the sales and purchase such as the company name, product, quantity and price were collected for the years 2022-24 along with the labour attendance data for the same period. This report entails an extensive analysis of the aforementioned data using visualisation techniques such as bar charts, line charts, heatmaps, etc. These visualisations help to understand the underlying patterns in the data such as inefficient use of labours, frequently purchased raw materials and relation between labour availability and sales. Insights obtained from the same will be used to conduct regression analysis using software like sci-kit learn and Excel *Forecast.Ets* which will help in forecasting the demand for various products during different periods in the year. The results will further be used to plan the labour requirement and warehousing of products accordingly.

2. Proof Of Originality

Document	**Description**	**Link**
Letter	The pdf file is a scanned image of the letter issued by the company stating the originality of data and has been provided by them for educational purposes.	Link
Interview Video	This is a mp4 file of around 4.5 minutes of my interview with the company owner.	Link
Data	This folder contains the google sheets with the processed data and also snapshots of the raw data provided by the organisation	Link

Table 1. Proof Of Originality

3. Meta Data and Descriptive Statistics

3.1. Meta Data

For the period July-2022 to June-2024, comprehensive data pertaining to both inputs and outputs was collected to analyse the productivity of the firm. The data includes, sales data, purchases data and labour data. The company manages their data in Tally software for sales, bills for purchases and a daily register-book for labour data.

3.1.1. Sales and Purchase Data

Column	Data Type	Description
Date	Datetime	Represents the date of sale or purchase of a product.
Company	String	Name of company from where the product is purchased or to whom the product is sold.
Product	String	Name of the product sold or purchased. E.g. Cable tray, Raceway, Nuts, Washer, etc.
Quantity	Float	The quantity of product purchased or sold.
Units	String	The unit of measurement of product. E.g. inch, mtr, pcs, nos, etc.
Rate	Float	The price per unit of the product. E.g. ₹100 for 1 mtr of cable tray.
Value	Float	The total price of the product, usually the product of Quantity and Rate.

Table 2. Sale and Purchase Meta Data

3.1.2. Labour Data

Column	Data Type	Description
Date	Datetime	Represents the date of the year
Labour Count	Integer	The count of labours present on that date.

Table 3. Labour Meta Data

3.2. Descriptive Statistics

3.2.1. Sales Data

The sales data is available for two fiscal years, 2022-23 and 2023-24. The data for both years have approximately 2100 samples representing the product sold by the business to other business across the year. The mean of product rate for the year 2022-23 is approximately ₹467.01 representing the average rates of the product sold. The same for the year 2023-24 is approximately ₹774.38, this gives the idea that the price of individual products have increased

over the year mainly due to the market price fluctuations. Further the standard deviation for the product rate in 22-23 is approximately ₹3,779.14 and the same for year 23-24 is ₹9,081.10, the high standard deviation indicates the high variability in the price of products which is further corroborated by the range of product rates which is ₹1,28,999 and ₹3,74,999 respectively. On the other hand, the sale value, which is the product of rate and quantity, has higher mean and standard deviation. Even though the standard deviation for product rate in 2022-23 is lower than that of in 2023-24, the sale value standard deviation shows an inverse trend, the reason for the same is the increase in the quantity of products demanded. The range of the sale value follows that of product rate indicating the large range of prices offered by the business for its products.

Sales				
	Rate (₹) (2022-23)	**Value (₹) (2022-23)**	**Rate (₹) (2023-24)**	**Value (₹) (2023-24)**
count	2057	2057	2139	2139
mean	467.014939	17919.023749	774.382352	16586.047789
std	3779.14419	41526.849359	9081.096395	36870.302382
min	1	20	1	14
25%	55	2066.25	53.5	2250
50%	159	6565	145	6720
75%	330	16610	310	16500
max	129000	707256	375000	531360
range	128999	707236.0	374999	531346

Table 4. Sales Data Descriptive Statistics

3.2.2. Purchase Data

The purchase data follows a similar trend to that of sales data. The purchase data is available for the year 2023-24 and has almost 900 entries. The mean of rate of products is approximately ₹2,892.87 and that of for sales is ₹35,092.36. Similar to that of sales data the standard deviation for both the rate and value are very high, at ₹12,311.78 and ₹35,092.36 respectively.

Purchase		
	Rate (₹) (2023-24)	**Value (₹) (2023-24)**
count	898	898
mean	2892.865924	35092.35827
std	12311.77742	87096.55939
min	0.07	3.6
25%	24.5	580
50%	66.4	3000
75%	130	19381.5
max	78500	845142.5
range	78499.93	845138.9

Table 5. Purchase Data Descriptive Statistics

3.2.3. Labour Data

Similar to that of sales data, the labour data is present for two fiscal years, 2022-23 and 2023-24. There are a total of 306 entries in both of the years (excluding the holidays and Sundays). The mean labour presence for both years is not too far apart, being 6.3 for 22-23 and 7.2 for 23-24. The standard deviation for both the years, though not very high, is moderately high given the range of data being between 2 and 12 for both the years indicating a decent variation in the presence of labours over the months.

Labour		
	Labour Count (2022-23)	**Labour Count (2022-23)**
count	306	306
mean	6.300654	7.238562
std	1.551667	2.308633
min	2	2
25%	5	5
50%	6	7
75%	8	9
max	10	12
range	8	10

Table 6. Labour Data Descriptive Statistics

4. Detailed Explanation of Analysis Process

The problem identified after the interview with the owner of SNP, was that the business faces acute labour shortage in the months of May-June and also during the festival of Diwali i.e. in Oct-Nov. Secondly, the firm is based on the idea of demand-based production, hence it has not explored the idea of warehousing it's products, mostly due to being unaware about the benefits of the same. A detailed analysis of the labour, sales and purchase data is performed to identify the trends in the same which in turn will lead to identification of potential solutions.

The data required for the analysis and finding solutions to the problems was collected from the company's database. The sales data provided for the period July 2022 to June 2024 was in the form of a pdf file which was extracted from the firm's Tally database. The data was then imported into excel.

The purchases made by the firm are not correctly stored in the Tally database so to get the required data the pictures of the firm's purchase bills were taken which were then converted into a tabular format manually.

Lastly, since there is no software used for storing the labour data, the firm uses a traditional register to note down the presence and absence of its employees. The same was again converted into tabular form manually which made it feasible for further analysis.

This process of data extraction was followed by cleaning and pre-processing the data. The sales and purchase data had multiple unwanted columns like CGST/SGST values, GSTIN number of customers, etc. removal of these and selecting the relevant was essential which was performed using Pandas' data-frames and other data manipulations techniques offered by the same. Lastly, the product and customer names in the sales and purchase data had a huge number of errors like spelling mistakes, unnecessary punctuations, etc. which was cleaned using Open Refine, a data wrangling software.

The above cleaning and pre-processing resulted in data which was feasible to analyse and draw conclusions from, to find potential solutions. The below analysis makes use of the following graphs for the mentioned reasons:

1. Line Chart: Line charts are ideal for visualizing time series data, offering a clear and effective way to compare trends across multiple periods. By connecting data points with lines, they make it easy to observe changes over time and identify patterns or fluctuations.

2. Treemap: Treemaps are highly effective for visualizing hierarchical data, particularly when representing parts of a whole with multiple components. They allow for a clear comparison of proportions, making it easy to see the relative sizes of each part within the overall structure.
3. Calendar-heatmap: For a voluminous dataset such as data for each day of the year, heatmaps provide a way to summarize the data in a single glance.
4. Bar chart: Bar charts are an effective tool for visualizing and comparing the volume of a quantitative variable across various categories, allowing for clear identification of patterns and trends.

4.1 Sales Data

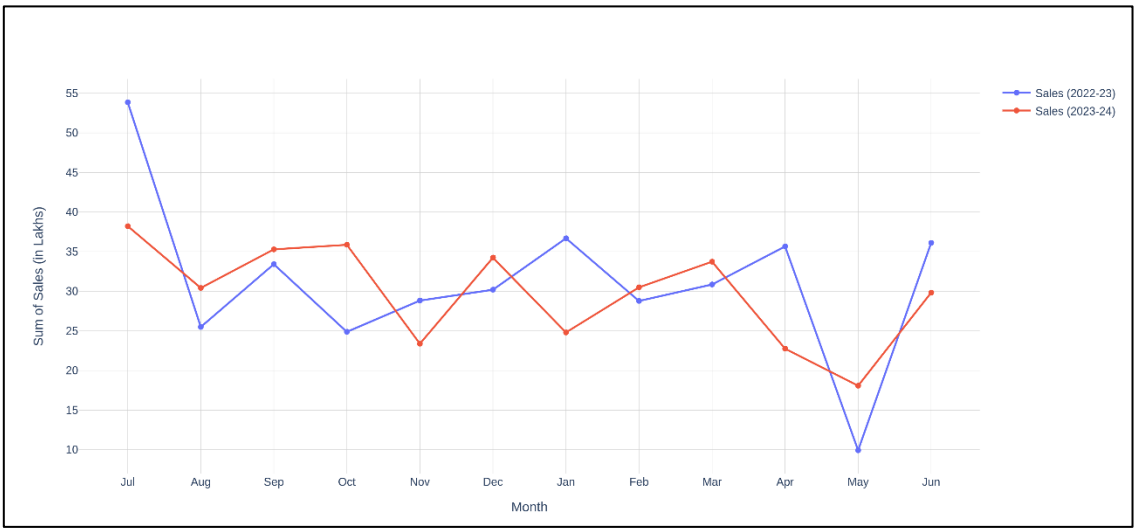

Figure 1. Sales Value Trend

Figure 1 shows the sum of sales value over the months for the year 2022-23 (red line) and 2023-24 (blue line). It helps in discovering the pattern in the sales value over the months and also compare the same across both financial years. The sales value seems to follow a similar trend for both the years, it can be seen that the sales in the month of May for both the years is at its lowest, approximately ₹10 lakhs for 22-23 and just over ₹18 lakhs for 23-24. It is also visible that sales drop in the month of November for the year 22-23 and in the month of October for the year 23-24.

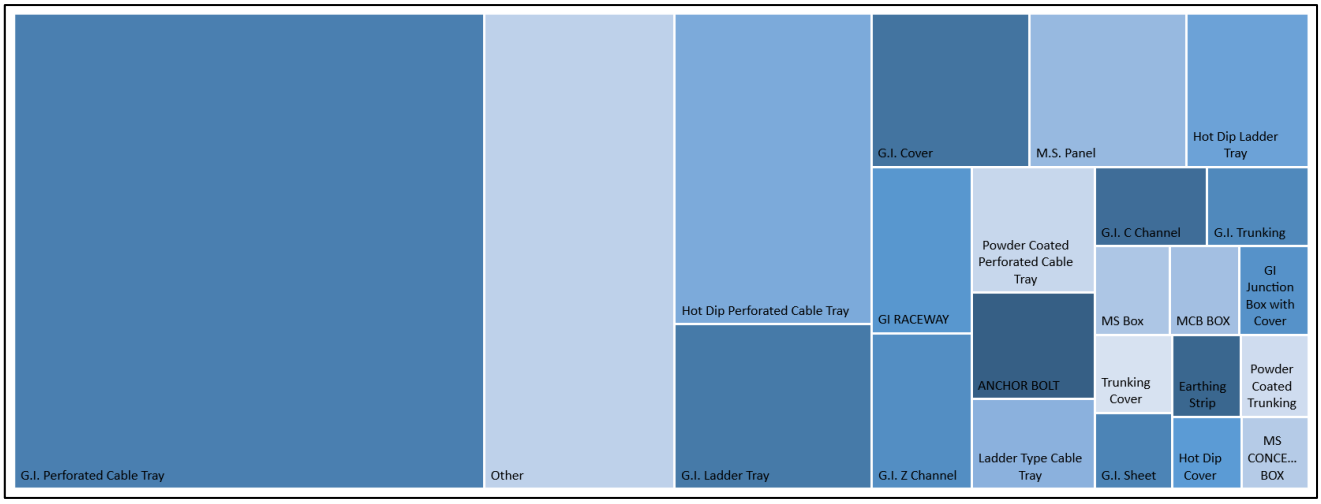

Figure 2. Sales Value (Product-wise)

Figure 2 represents different SKUs of the company with respect to the sales value. A tree-map is more favourable in representing this data due the large number of unique products supplied by the company, which when represented using a bar chart looks cluttered. The SKU G.I. Perforated Cable tray contributes to more than 35% of the total sales. The product 'Other' consists of all the products which do not contribute more than ₹5 lakhs to the total sales value. The second highest product in terms of sales value is Hot Dip Perforated Cable Tray followed by G.I. Ladder Tray contributing 9% and 5% to the total sales value respectively.

4.2 Labour Data

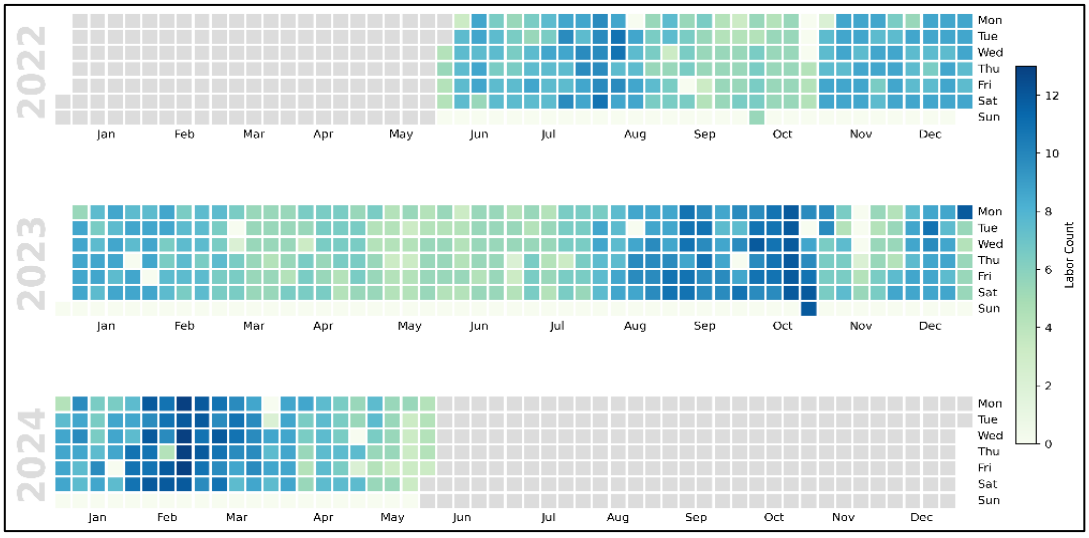

Figure 3. Calendar Heatmap of Labour Data

Figure 3 represents the monthly labour count. A calendar heatmap is the most appropriate visualization for this particular data as the data spans each day of two fiscal years. Figure 3 helps us identify the pattern in the labour count for the available data i.e. from June 2022 to

May 2024. The heatmap shows that the labour strength in the Oct 2022 is moderately low mainly due to that being a festival season. In the year 2023, low labour strength is noticeable in the months of April to June and then again in November, the trend follows the information obtained from the interview with the owner that the labour tends to leave for their native towns in April-May and also during the Diwali season in the month of Oct-Nov. The month of May 2024 follows a similar trend as that of May 2023 pertaining to similar reasons.

4.3 Purchase Data

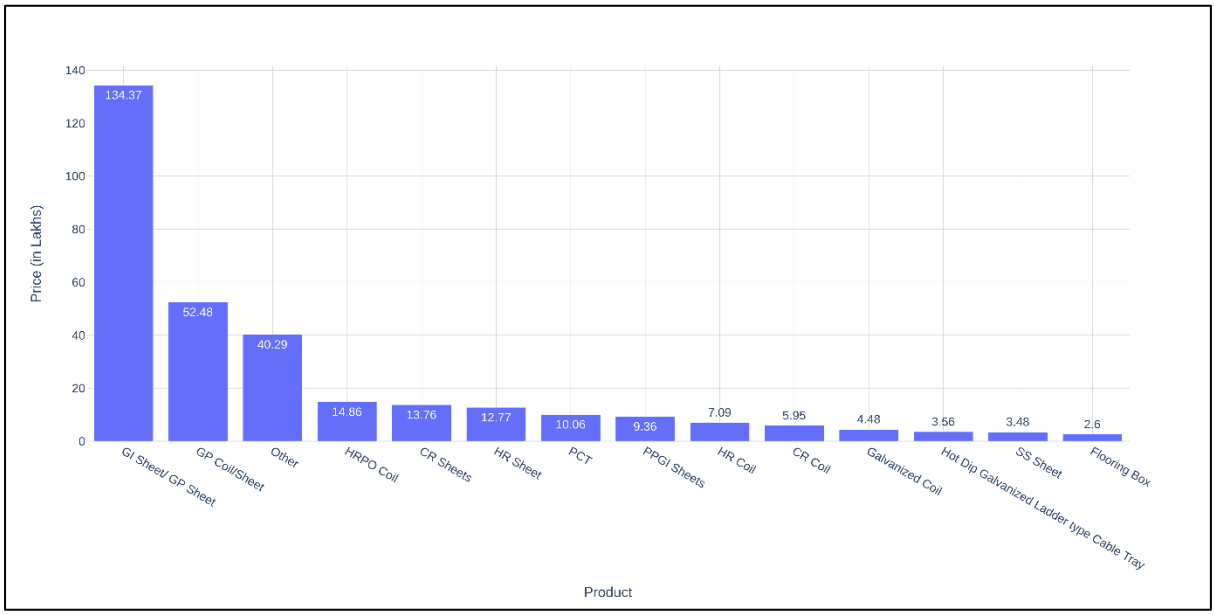

Figure 4. Purchase Trend (Product-Wise)

Figure 4 represents the purchases made by SNP in the fiscal year 2022-23. The figure represents the frequent purchases, in terms of purchase value, which were made by the company. It can be seen that GI Sheet/ GP Sheet is the product which was most purchased followed by GP Coil/Sheet. The 'Other' category represents the products which have a cumulative purchase value of less than ₹2.5 lakhs and hence do not contribute much to the purchases made by the company. The figure helps in identifying the primary raw materials required for the manufacturing of goods offered by the company.

5. Results and Findings

The problems identified about the business after the interview with the owner are visible in the above analysis. The labour shortage in the month of April-May and Oct-Nov affects the sales of the firm in the respective months. The average of labour available per day in the month of May-2022 was 4.33 and the same for May-2023 was 4.5, the lowest for the respective years. The average daily sales in the month of May-2023 and May-2024 were ₹8,705.32 and ₹15,216.58, lowest for the respective years, which is a direct effect of the low labour strength

in the respective months. The effect of labour on sales is quite obvious and can also be corroborated by figure 6, showing a positive relationship between the two with a correlation coefficient of +0.36.

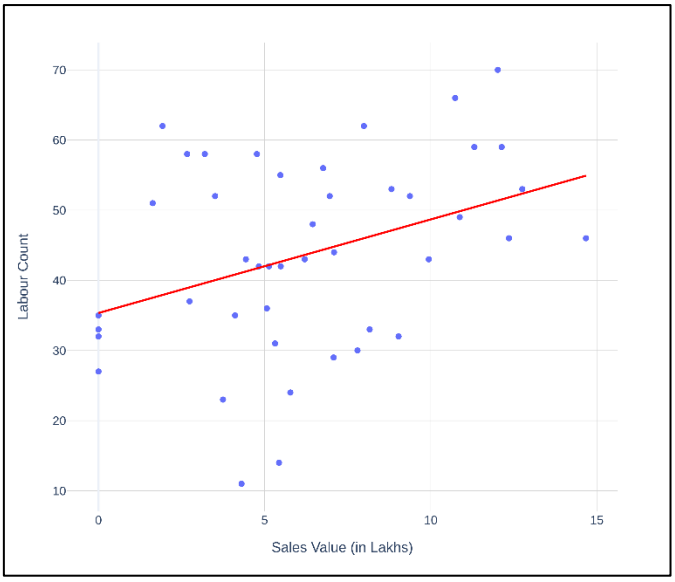

Figure 5. Sales Value v/s Labour Count

The sales value analysis revealed the key performing SKUs, this analysis can be beneficial when exploring the idea of warehousing of products. The products such as various types of cable trays, ladder trays, covers, etc. can be manufactured well in advanced and can cater to the need of the company's business partners as these products are high in demand throughout the year.

On the other hand, figure 4 represents the frequent purchases made by the firm, these products are the essential raw materials required for manufacturing the in-demand products. The trend in the figure can help the firm to order the required products in time so that the production process is uninterrupted making the warehousing and inventory management smoother.

The interview with the owner revealed that the firm is profitable on a yearly basis but when conducting a comparative analysis between the sum of labour and profits for particular months, it is inferred that the labour is not utilized to its full potential. To justify the same, figure 6 plots the normalized values of profit and sum of labour per month. The below formula was used to calculate the profit.

$$Profit = Sales - (Purchase + Labour\ Cost + Factory\ Bills + Miscellaneous\ Costs)$$

In order to scale the values to make them comparable, the below formula was used.

$$x_i' = \frac{x_i}{max\ (|x|)}$$

The figure shows that for almost all the months the normalized value of labour count is more than the normalized value of profit. The only month where profit exceeds labour count is the month of March, depicting the true labour potential. An efficient plan to cut down on labour in the months with low demand or increase production in the months with high labour strength can help optimise the costs and profits of the firm.

Figure 6. Labour count and Profit (monthly)

6. Goals for the Final report

- The insights gained from the above sales analysis will be used for demand forecasting, helping the firm to estimate the demand in advance and hence warehouse the products accordingly.
- Conducting a regression analysis to predict need of labour followed by a cost benefit analysis between the actual values and the predicted values will help in deciding between hiring labour at a higher cost or sticking to the approach currently followed by the firm.
- Although specific data for B2C sector demand is unavailable, obtaining an estimate from the owner regarding the products and quantities in demand will enable a cost comparison between running a website and the potential profit from meeting B2C sector needs.