Business Data Management

A Mid-Term report for the BDM capstone project

TITLE: Precision Enhancement of Customer Care Workflow in Pharmaceutical Setting & Leveraging Data-Driven Strategies.

Preliminary information of the student:

Name: DHULIPALLA UMESH MANIKANTA

Email: 23f2001471@ds.study.iitm.ac.in

Roll No: 23f2001471

Contents

Executive Summary:

➢ *Organization Overview:*

I elected to collect the data from LAXMI GENERIC PHARMACY, located in Guntur, Andhra Pradesh. The store was started by Mr. Balu in the year 2019 with initial investment of 10 Lakhs. As the time moved on, he started to open new branches and hiring people accordingly.

➢ *Brief on Problems Faced by the Business:*

Inventory mismanagement, losses from expired medicines, cash flow blockage, uneven branch performance, and a shift from B2C to B2B are key issues. The focus is on enhancing sales and expanding business reach.

➢ *Synopsis of Pre-Processing:*

Pre – Processing contains the following stages: Data collection, Data Cleaning, Metadata of Data, and Descriptive Statistics.

The owner of the pharmacy provided the data of 9 months which is from June 2023 to Feb 2024.

This segment contains the details of how the data is arranged, cleaned, organized for Analysing and a concise overview of the data's nature and purpose.

Descriptive Statistics of this data contains information like average amount earned per month, average number of sales happens per week/month, variation of profit over time, average loss due to expired medicines. max/avg/min profit earned, etc. Through the above analysis we can come to gain insights into the business's financial performance.

➢ *Overview of Results and findings:*

This above analysis will aim to provide primary solutions to identified problems effectively and help identify precise areas for enhancement. By this analysis Inventory and branches can be managed properly which are few of our key challenges.

Proof Of Originality:

- ***Shop Address are***:

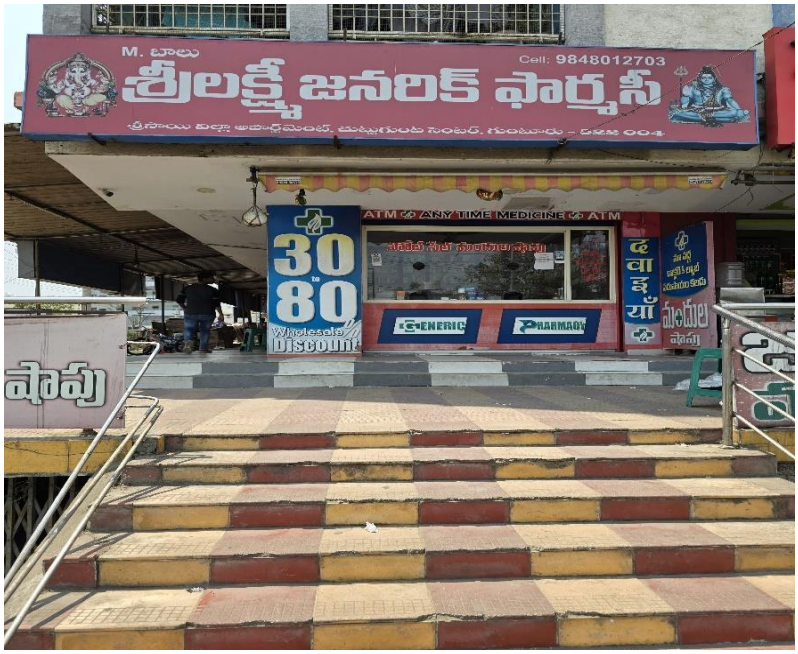
  There are various branches for the shops, the address for few branches are:

  - ✓ 4th Line, Gunturvari Thota, Guntur, Andhra Pradesh – 522001
  - ✓ Chuttugunta Circle, Near SBI ATM, Guntur, Andhra Pradesh-522001

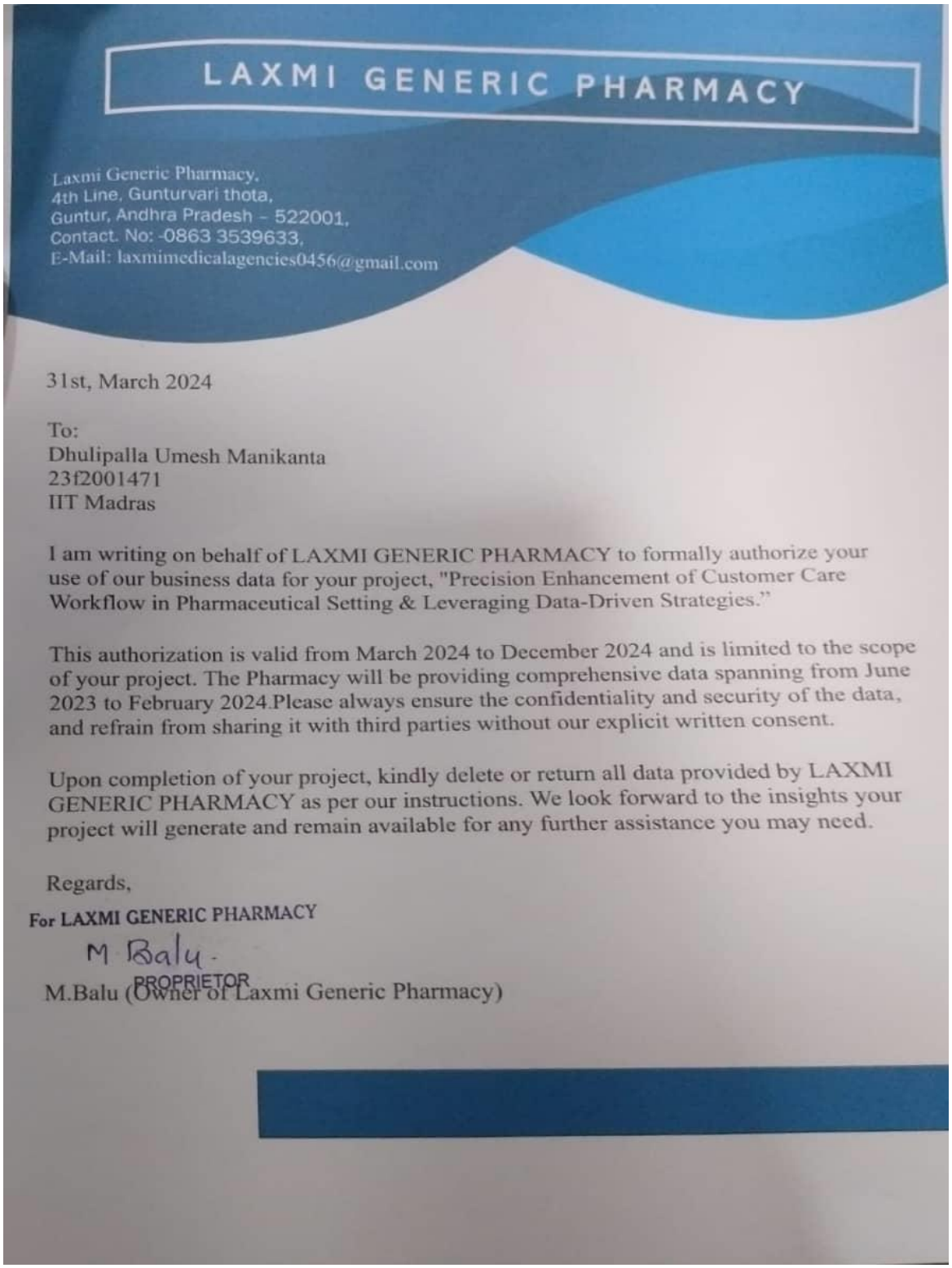
➢ ***Recorded interaction with the shop owner (Link):***

https://drive.google.com/drive/folders/118ABh83LQCKFCY_9fzCNOqM49VGF-89w?usp=drive_link

➢ ***Images Of the Shop are:***



➢ *Authorization Letter from the owner:*



LAXMI GENERIC PHARMACY

Laxmi Generic Pharmacy,
4th Line, Gunturvari thota,
Guntur, Andhra Pradesh – 522001,
Contact. No: -0863 3539633,
E-Mail: laxmimedicalagencies0456@gmail.com

31st, March 2024

To:
Dhulipalla Umesh Manikanta
23f2001471
IIT Madras

I am writing on behalf of LAXMI GENERIC PHARMACY to formally authorize your use of our business data for your project, "Precision Enhancement of Customer Care Workflow in Pharmaceutical Setting & Leveraging Data-Driven Strategies."

This authorization is valid from March 2024 to December 2024 and is limited to the scope of your project. The Pharmacy will be providing comprehensive data spanning from June 2023 to February 2024.Please always ensure the confidentiality and security of the data, and refrain from sharing it with third parties without our explicit written consent.

Upon completion of your project, kindly delete or return all data provided by LAXMI GENERIC PHARMACY as per our instructions. We look forward to the insights your project will generate and remain available for any further assistance you may need.

Regards,

For LAXMI GENERIC PHARMACY

M Balu.

PROPRIETOR

M.Balu (Owner of Laxmi Generic Pharmacy)

Metadata:

Metadata implies the brief description about the data which means data about data. Metadata of this pharmacy data contains the following information: Establishment details, Business details, financial profile like average net worth average profit, revenue, purchase amount, etc., Operation management details, etc.

❖ *Metadata of Business Organization:*

S.No	Key Metric	Value
1	Organization Name	LAXMI GENERIC PHARMACY
2	Location	Guntur, Andhra Pradesh
3	Founder	Mr. Balu
4	Year of Establishment	2019
5	Initial Investment	10 Lakhs
6	Organization Type	B2C
7	Number of Branches	5
8	Number of Employess	10
9	Key Issues Addressed	Inventory mismanagement, expired medicines,cash flow, branch performance

❖ *Metadata of Data Used in Analysis:*

Owner of the Organization had provided data of 9 months from June 2023 to Feb 2024.

There are three datasheets for each month. They are:

1. Sales of the Month.
2. Purchases of the Month.
3. Stock maintenance of the month.

For the sake of Analysis, I have prepared the 2 types of data from the data provided: one is by merging all sales data and purchases data separately over all months for getting a brief insight. For getting a detail insights, I have combined sales and purchase data separately, aggregating them into three-month intervals.

The Important Columns in datasets are:

Sales

1.Product Name
2.Company
3.HSN Code
4.Invoice Date
5.Branch
6.Batch Number
7.Expiry
8.MRP
9. Cost price
10. Selling Price
11. Quantity Sold
12. Subtotal
13. CGST
14. SGST
15. Amount
16. Profit

Purchase Rate

1.Product Name
2.Company
3.HSN Code
4.Invoice Date
5.Distributer
6.Batch Number
7.Expiry
8.MRP
9.Item Discount
10.Bill Discount
11. Cost price
12. Selling Price
13. Quantity Sold
14. Subtotal
15. CGST
16. SGST
17. Amount
18. Net PurchaseRate
19.Landing Cost
20.Margin

Stock

1.Product Name
2.Company
3.Opening Stock qty
4.Opening Stock Value
5.Sale_qty
6.Sale_Value
7. Adjusted Stock qty.
8. Adjusted Stock Value
9.Closing Stock qty
14.Closing Stock Value

Other than this information we also have the information about the type of Drug that medicine belong to. There are 4 major variety of drugs, they are schedule type-G, type-H, type-H1, type-X.

1. Sales:

- Sales data, contains the information of medicine sold, its company name, Expiry, MRP, etc.
- Also contain, from which Branch the sale happened.
- Selling and cost price of medicine and quantity sold.
- Total amount, profit from that medicine.

2. Purchase:

- Purchase data contains information of purchased medicine, company, Distributer, Expiry.
- Also contain, Item Discount on that medicine, net purchase, Landing cost.
- CGST, SGST and margin on that medicine.

3. Stock:

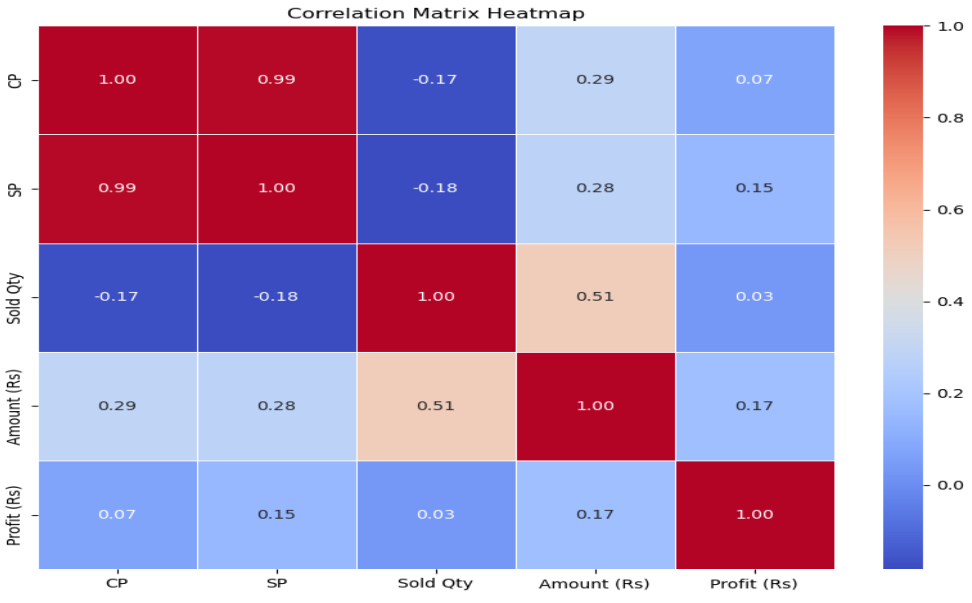
- This Stock dataset contain the Opening stock on 1st of the particular month and Closing stock on particular month.
- This data is internally uses above 2 datasets for calculating Closing stock.

Descriptive Statistics:

- Descriptive Statistics involves the summarizing of the data and explaining the Characteristics of the variables involved in the data for this data descriptive Statistics contain information like average amount earned per month, average sales month, variation of profit over time, max/avg/min profit earned, etc.
- One the good methods of Descriptive Statistics is ***five-number summary,*** it contains ***Minimum, 1st Quartile, Median, 3rd Quartile, Maximum***. For the sake of our analysis, we will add 3 more important metrics to it, they are ***mean and variance*** which comes under under Measures of central Tendency and variance, ***Skewness*** which describes the distribution of data.

S.No	Metric	No of Sales	Revenue	Profit	No of Purchases	Purchase_value
1	Min	98	₹ 3,69,535.59	₹ 7,044.04	29	₹ 4,79,286.33
2	1st Quartile	454	₹ 5,57,457.83	₹ 8,492.51	34	₹ 5,24,370.83
3	Median	895	₹ 5,75,287.92	₹ 9,684.73	39	₹ 5,61,617.85
4	3rd Quartile	1093	₹ 5,97,361.10	₹ 12,445.78	41	₹ 6,42,881.46
5	Max	1244	₹ 6,20,091.16	₹ 15,737.22	45	₹ 7,33,204.48
6	mean	737	₹ 5,54,194.55	₹ 10,582.12	37.67	₹ 5,82,384.35
7	Variance	170109.75	6111391343	8910100.77	25.5	6776478681
8	Std.Dev	412.44363	78175.38835	2984.97919	5.049752469	82319.37002
9	Skew	-0.47021	-1.945542109	0.53561105	-0.497569942	0.675971978

- In this Analysis we also need to consider correlation between variables for understanding the patterns. The metric which is used to measure the correlation between the variable is "***Correlation Coefficient***" which is value between ***-1 to 1,*** – for inverse relation, + for direct relation.



If we observe the Correlation matrix there is strong positive relation between Cost Price and Selling Price of each medicine. Moderate positive Relation between Amount and Quantity Sold, surprisingly there is no strong correlation between profit and other variables.

Detailed Explanation of Analysis of Process:

The steps contain the information of how the data is cleaned, analysed and why that analysis is important.

Cleaning Data:

This part includes checking the format in which the data is being provide and identifying the appropriate columns and checking for unwanted columns, checking for missing values. The data contain some duplicated columns and some columns which are not necessary for our analysis. Some such columns are:

- "Bill_with_GST" this column contains "Yes" for all bills in sales data.
- "Created_by" this column contains "Balu (Owner)" for all bills.
- "IGST" percentage and amount are 0 for all the medicines.
- "Bill Discount" is zero for all the bills.

These columns are not providing any kind of useful info related to analysis.

There is a column which stores from which branch the sale happened. There are some missing values.

As the Column is having Categorical Data, Mode is preferred for filling missing values. So, I used MODE function in excel to fill missing values.

Analysis on Data:

- ***On Inventory & Cash Flow:***

As we are analysing on huge amount of data it is better to work on small chunks by dividing the data. So, month wise is preferred. we have Invoice data for sale as well as purchase, with the help of text function I extracted month and year of the invoice.

I did all the analysis with the help of ***"PIVOT TABLE"***. This one of the easiest ways of analysing data. with this we can directly find all kinds of information like mean, sum, count, min, max, etc.

Mean is one the important metric for all numerical variables. With this metric we can get a proxy for how much revenue, profit, we are getting, same for purchase value. With the help of variance, we will come to know how spread is a particular, this is help you to manage Inventory, how to place the orders for purchases.

With this we can try to figure out from which Distributors we are good profits. By figuring on which tablets we are getting loss, we can find appropriate ways like change distributor or company or increase to resolve it. This analysis will help in managing inventory upto a short extent. This one of the problem statements.

- ***Branch Analysis:***

Another problem statement is Branch Analysis i.e. which branch is performing good, from which branch the business is getting good Profits. As discussed above first we filled missing values and used Pivot Table is for this analysis.

Results and Findings:

Visualizations helps us understand better. Charts like Bubble, Staked Bar Chart, line chart, etc are used.

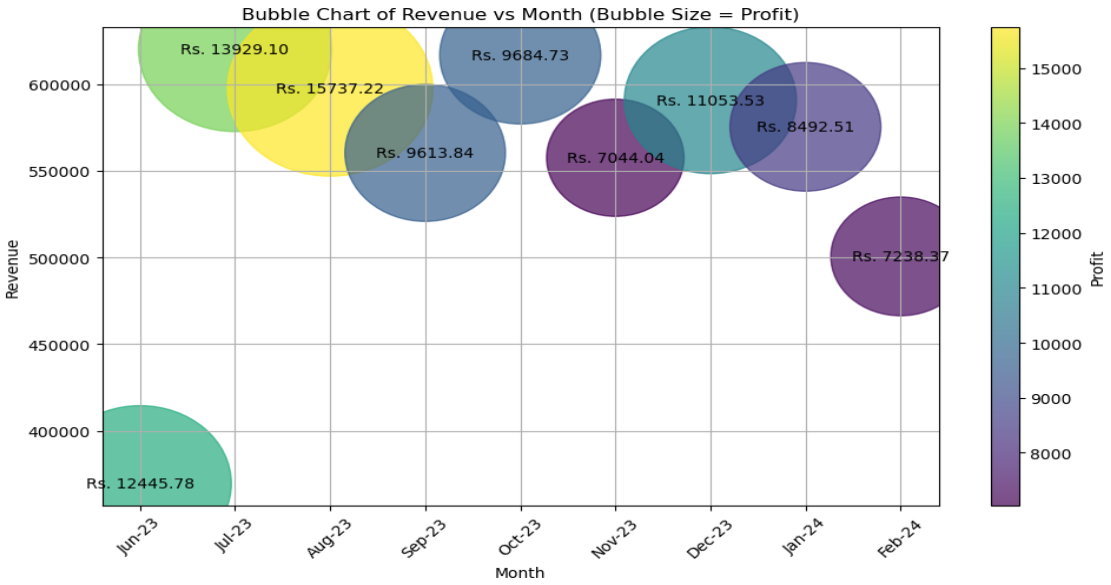
- ***Link to Data & Analysis Done:***

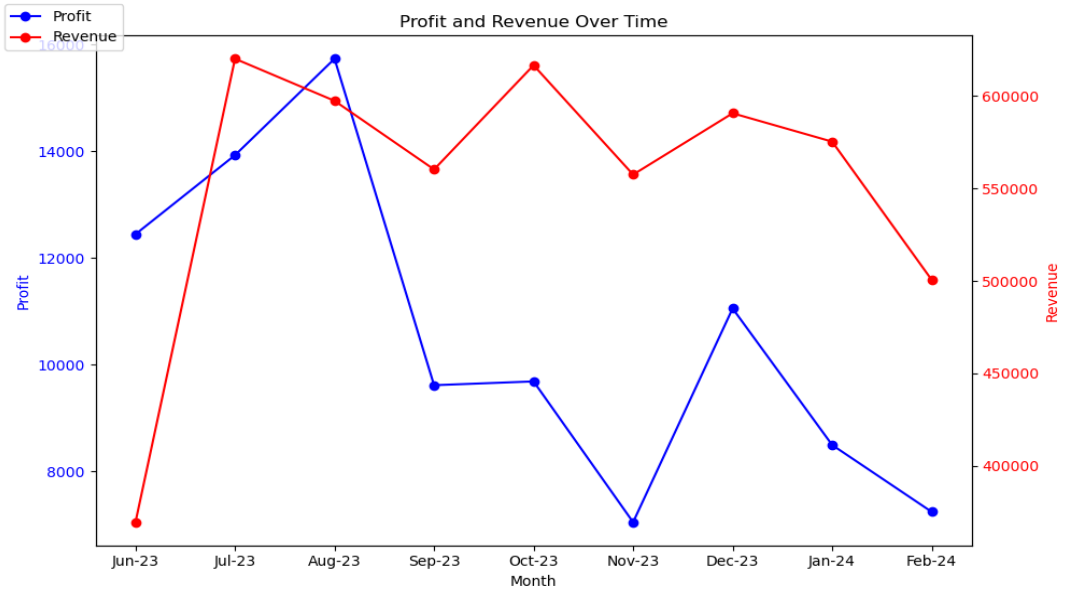
https://drive.google.com/drive/folders/1ywOzNio9QKoI2LTSOCXk9Sn391N1WtSZ?usp=drive_link

** Because of Space constraint Some of the important visualisations are shown here remaining are provided along with data. Python Code for the analysis is also attached.

- **Revenue and Profit:**

Bubble chart is Used for showing Revenue and profit is changing over past 9 months. In the above chart bubble size and colour of the bubble indicate the profit earned. Findings from the above chart are business had Earned **Good profit in the month of August 2023** and **low profit in the month of November 2023.**



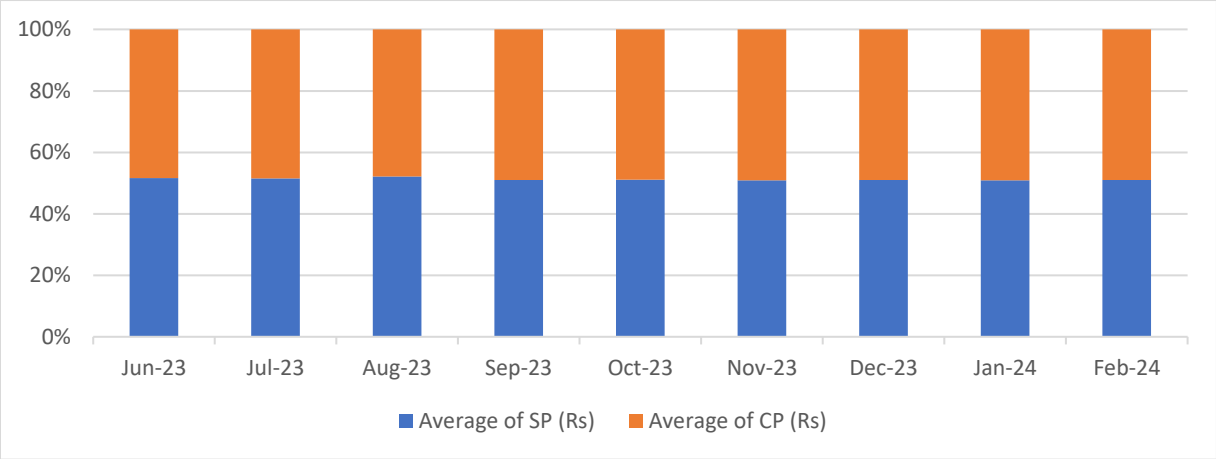


For Observing the trend of revenue and profit we used scatter plot with Multi axis, since the range of both features are different, we used multi axis. Left vertical axis is for profit and right is for Revenue.

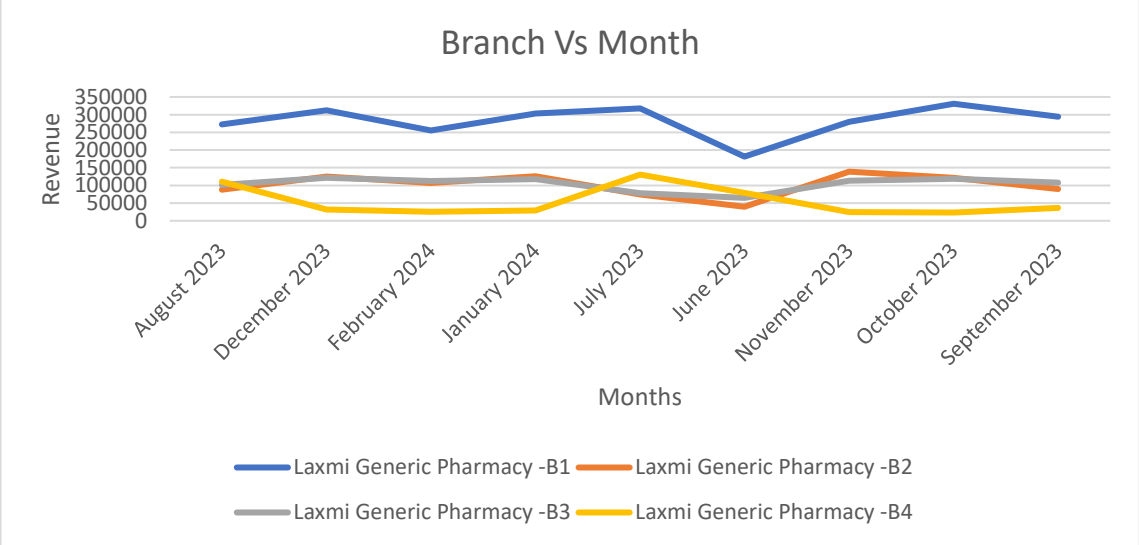
Findings are the following:

- The is huge jump in revenue from June to July 2023. I tried to find the reason for this, the reason for this is the sales data is available from 9th June. Data don't have first 8 days of June in data. Approx. Conclusion we can draw is revenue varies between 50 Lakh to 60 Lakh.
- Coming to Profit there is an abnormal down fall of profit in the months Sep, Oct, Nov of 2023 even though the sale and Revenue are high.

- One more important thing to Notice is revenue is in lakhs (50L – 60L) and profits are in thousands (7k- 15k), These might be the reasons for this abnormality: improper way of deciding Selling price, Cost fluctuations of medicines [cost of goods sold (COGS)].
- With this *100% staked Bar Graph* we can observe *SP is slightly greater than CP*, we can find *Avg CP and Avg SP are roughly same*, This *Might be the reason for low profits.*



- **Branch Analysis:** Out of 5 branches one is supplier to other 4 branches, very low retail sales happen, so 4 branches are considered for analysis.



From this Chart we can Observe the Branch-1 is performing very well, Branch-2,3 are having similar trends and are performing moderately. Branch – 4's performance is somewhat low when compared with other branches.