

Improving Sales and Procurement in the Pharmaceutical Sector through Data Analytics

A Mid-term report for the BDM capstone Project

Submitted by

Name: Shouvik Roy

Roll number: 23f1001348



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

Contents

1. Executive Summary	1
2. Proof of Originality	2-3
3. Meta Data	4-5
4. Descriptive Statistics	5-6
5. Detailed Explanation of Analysis Process	7-9
6. Results and Findings	10-11

Executive Summary

This mid-term report presents the advancements achieved in improving the operational effectiveness and market positioning of a B2C pharmacy located at 50 Millanpolly, JBL Apartment, Airport Gate-2, Kolkata-700079. The pharmacy operates continuously, providing local customers with a comprehensive selection of pharmaceutical products.

The initial focus of the project was on examining sales and inventory data from July 2024 to January 2025, as well as digitizing transaction records to adopt a more organized and data-centric approach. This analysis uncovered significant challenges, such as a decline in sales of Pfizer products attributed to an over-dependence on COVID-related medications, along with Lupin's poor sales performance resulting from an unreliable supply chain and reduced consumer confidence.

Early findings indicate that although the demand for general medications remains stable, brand-specific challenges are considerably affecting revenue. Furthermore, the store faces seasonal variations in certain product categories, highlighting the necessity for a more adaptable inventory and marketing strategy.

To tackle these issues, the pharmacy needs to:

Broaden the range of Pfizer products beyond those related to COVID to align with changing consumer preferences.

Enhance the reliability of the supply chain for Lupin by partnering with more trustworthy suppliers.

Increase customer confidence and brand recognition through targeted marketing initiatives.

Utilize digital tools for data-driven inventory oversight and tailored promotions based on consumer behavior.

The objective is to formulate practical strategies that boost product visibility, customer interaction, and operational processes, thereby ensuring sustainable growth and long-term success.

Proof of Originality

The only source of information for this project was Mr. Suvankar Chakraborty, the founder of Chaya Medical, who provided data from a handwritten diary containing customer details as well as records of sales and services. This data spans the entire period of the shop from July 2024 to January 2025.

Steps Implemented to Ensure Originality

1. Data Gathering

The pharmacy's sales and inventory records were examined over a seven-month duration, concentrating on transaction trends, customer demand, and brand performance. These records were manually converted from physical transaction logs and sales invoices into an organized Excel spreadsheet for analysis. The data incorporated:

- Medicinal brands and types that were sold.
- Pricing, purchase, and sales information to track revenues and profit margins.
- Customer preferences and seasonal demand variations affect inventory levels.

[CLICK HERE](#) for the raw data.

2. Conversion Method

- A uniform format was preserved while digitizing transaction logs to guarantee accuracy.
- The Excel was arranged with designated columns for medicinal brands, sales figures, purchasing costs, and customer insights, reflecting the physical records.
- Data integrity checks were conducted to ensure precision.

3. Verification

- Data was cross-referenced with purchase invoices and supplier records to validate accuracy.
- Confirmed that all sales trends and insights on brand performance corresponded with the pharmacy's actual operations.
- Collaborated with the pharmacy owner to ensure the data recorded was consistent with real-world practices.

4. Supporting Documentation

- Images of physical sales and purchase records.
- Screenshots of the organized Excel spreadsheet containing transaction details.
- [CLICK HERE](#) for access to all pertinent pharmacy images, videos and data.

By employing original transaction data directly from the pharmacy, this project adheres to academic integrity standards, ensuring the authenticity, accuracy, and reliability of all findings and strategic recommendations.

CHHAYA MEDICAL

**50, Milan Pally, Airport Gate No-2,
Kolkata-700079.**

TO WHOM IT MAY CONCERN

It is certified to Sri Shouvik Roy, a student of IITM for B.S. degree . He has collected specific business data for his academic project work named '**BDM CAPSTONE**' from my shop, those should be used solely for academic purpose .

I also confirm that the data supplied is accurate to the best of my knowledge, and that I am the legitimate provider of this data. This information is being supplied to Sri Shouvik Roy for academic purpose.

I wish him a bright future in life.

Date:- 11.01.2025

CHHAYA MEDICAL
Subhankar Chakrabarty
Proprietor

[CLICK HERE](#) to view the Validation letter in pdf format

Metadata

DATA SOURCE:

The information comes from handwritten records of sales and inventory kept by the pharmacy. These documents track customer transactions, medicine brands, pricing, and revenue patterns from the last seven months. The details have been manually entered into a structured Excel file for precise analysis and informed decision-making. [CLICK HERE](#) to access the data.

DATA FORMAT:

The data is organized in a table format, with each row corresponding to a single transaction. It consists of two sheets of sales and purchases.

Sales sheet columns: Date, Name of customer, Brand of Medicine, Medicine type, Selling Price, Quantity sold, Total selling

Purchase sheet columns: Aate, Brand of medicine, Medicine type, Price, Quantity, Total purchase.

COLUMNS EXPLANATION:

The below table provides you with the Purchase sheet structure:

Column Name	Description	Example Value	Attribute Type
Date	The date of the transaction in DD-MM-YYYY format.	13-07-2024	Date/Time
Brand of Medicine	The pharmaceutical brand name of the medicine.	Cipla, Pfizer, Lupin	Categorical
Medicine Type	The type and name of the medicine, including dosage details.	Capsule (Metrol XR 50mg)	Categorical
Price (Per Piece)	The cost of one unit of the medicine in Indian Rupees (₹).	₹12.35, ₹95.98	Numerical
Quantity	The number of units purchased in the transaction.	10, 50, 30	Numerical
Total Purchase	The total cost for the given quantity (Price × Quantity).	₹1,213.50, ₹687.00	Numerical

The below table provides you with the Sales sheet structure:

Column Name	Description	Example Value	Attribute Type
Date	The date of the transaction in DD-MM-YYYY format.	13-07-2024	Date/Time
Name of Customer	The name of the customer who purchased the medicine.	Sunetra Ghosh, -	Categorical
Brand of Medicine	The pharmaceutical brand name of the medicine.	Cipla, Pfizer, Lupin	Categorical
Medicine Type	The type and name of the medicine, including dosage details.	Capsule (Becosules Z)	Categorical
Selling Price	The price at which one unit of medicine was sold (₹).	₹54.94, ₹300.00	Numerical
Quantity Sold	The number of units sold in the transaction.	1, 2, 10	Numerical
Total Selling	The total revenue for the given quantity (Selling Price × Quantity).	₹109.88, ₹1,214.44	Numerical

TIMESPAN:

The time span of the data is from June 2024 to January 2025, offering perspectives on historical patterns as well as current advancements.

INSIGHTS:

This dataset is extensive and encompasses essential business indicators such as:

- Revenue fluctuations across several months.
- The demand for particular medications among various brands.
- Customer buying behaviors, including commonly purchased types of medicines and their quantities.

SCOPE OF THE DATASET:

The dataset total has 1199 rows and 13 columns from the combined Sales and Purchase dataset, giving useful insights:

- Total revenue generated after selling is ₹2,11,719.81 in the past seven months from July 2024 to January 2025.
- Total purchasing cost of medicine from July 2024 to January 2025 is ₹2,30,869.63.
- Total customer involvement is 73%.
- The highest-selling brand is Dr. Reddy's with 21% and lowest-selling brand is Lupin with 1%.

Descriptive statistics

This dataset comprises 1199 entries and 13 columns, created by merging Sales and Purchase data. The time range covered by the dataset spans from July 2024 to January 2025. The dataset organizes data patterns across 13 columns, including date, brand of medicine, name of medicine, type of medicine, and more. Based on the data, the shop has generated approximately ₹2,11,719.81 in revenue over the past seven months (from July 2024 to January 2025). Dr. Reddy's brand has recorded the highest sales, while Lupin's brand has seen the lowest. Customer engagement frequency has reached as high as 73%.

Overview of Dataset

- Total Records: 1,197 (Purchase: 598, Sales: 601)
- Timeframe: July 2024 to January 2025
- Total Revenue: ₹2,11,719.81

- Medicine Categories Analyzed:
 - Tablets & Capsules: 1780
 - Syrups: 656
 - Injections: 1211

Key Statistics

- Amount Spent & Earned (INR):
 - Average Purchase Price: ₹454.214
 - Maximum Purchase Price: ₹10,406
 - Average Selling Price: ₹483.379
 - Maximum Selling Price: ₹8,005

Customer, Purchase & Sales Insights

- Repeat Customers: Bashkar Chakraborty (29 counts, 6.62%).
- Most Purchased Medicine Brand: Cipla (27.503%).
- Most Sold Medicine Type: Capsule(Becosules Z), 9.36%.

Sales & Purchase Trends

1. Monthly Purchase & Sales Trends

- November 2024 has the highest amount of sales, and July 2024 has the least amount of sales.
- October 2024 has the highest amount purchased, and November 2024 has the least amount purchased.

2. Top-Selling & Least-Selling Medicines

- Cipla is the highest-selling brand, and Lupin is the least-selling brand.

Detailed Explanation of Analysis Process

At first, I collected the raw data from the shop and structured it digitally through Excel. I cleaned the data using Pivot table, which helps me to find out the data inconsistencies, data anomalies, blank data, etc.

Descriptive statistics were utilized to assess trends in pricing, sales behaviors, and customer purchasing patterns. Monthly trends were analyzed to identify revenue increases and variations in seasonal demand.

The tools I have utilized for data framing include Excel. I have examined analytics trends in graphs using Python (pandas, matplotlib, chart.js) as well as some free online platforms like Plotly.

METHODS OF DATA PREPARATION:

1. Validation of Source

- Information was gathered from unprocessed handwritten data that includes two categories: Purchase and Sales.
- Important columns in the Purchase dataset include Date, Brand of Medicine, Medicine Type, Price (Per piece), Quantity, and Total Purchase.
- Important columns in the Sales dataset consist of Date, Customer Name, Brand of Medicine, Medicine Type, Selling Price, Quantity Sold, and Total Selling.

2. Data Cleaning

- Transformed all numeric columns (Price, Selling Price, Total Purchase, Total Selling, Quantity) into a numerical format.
- Standardized names and categories of medicines to prevent duplication.
- Corrected date formats to ensure accurate time-based analysis.
- Removed missing and inconsistent data, guaranteeing precision in the final analysis.

ANALYSIS STRUCTURE:

1. Descriptive Statistics

- Calculated key metrics, such as:
 - Median Purchase Price: ₹40.24
 - Median Selling Price: ₹105.55
 - Profit per unit: Selling Price – Purchase Price
 - Total Revenue and Profit Analysis
- Customer Purchase Patterns:
 - Identified repeat customers and their contribution to total sales.
 - Analyzed customer spending behavior.

2. Trend Analysis

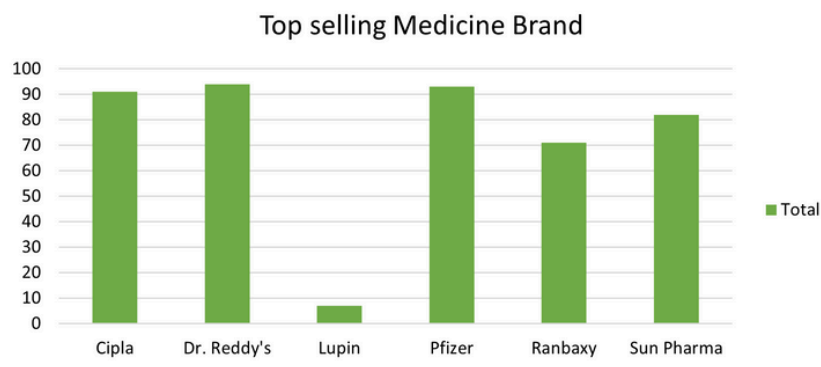
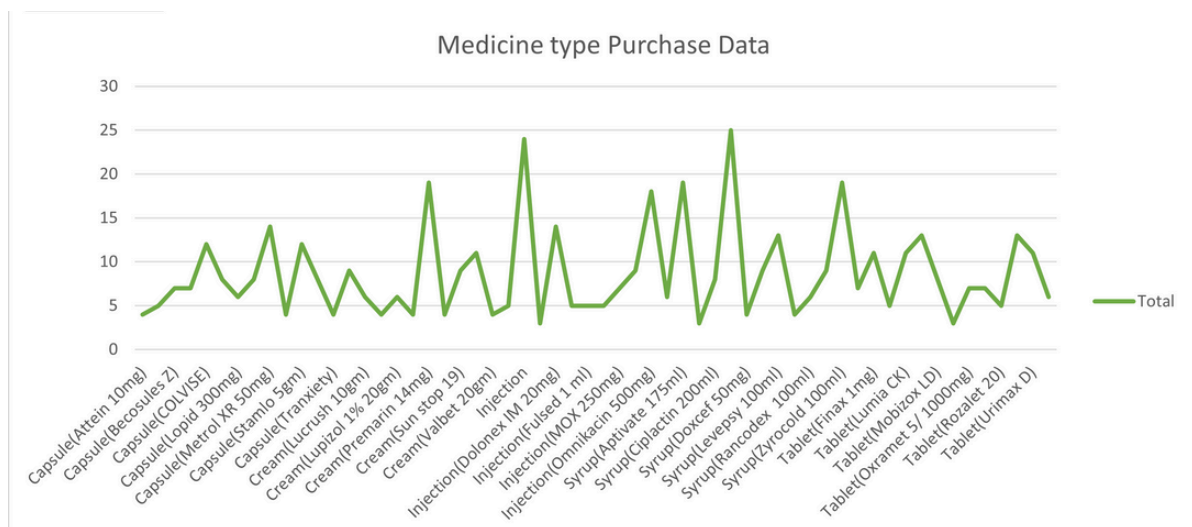
- Monthly Sales Trends:
 - Analyzed variations in revenue across various months.
 - Recognized the months with the highest sales and shifts in seasonal demand.
- Profit/Loss Trends:
 - Identified the medicines with the highest profitability.
 - Evaluated the products that are currently unprofitable and considered possible price changes.

3. Demand & Market Trends

- Most Purchased Medicine Types: Identified the most frequently bought medicine categories.
- Top-Selling Brands: Identified the brands that experienced the greatest level of demand.
- Inventory Management Insights: Proposed strategies for stocking based on observed demand patterns.

USED TOOLS:

- Excel: Initial data structuring, data cleaning, calculation, graph.



JUSTIFICATION FOR THE ANALYSIS METHODS USED:

1. Descriptive Statistics

Why Used?

- Provides a quick overview of important financial indicators such as average selling price, median acquisition cost, and profit margins.
- Helps identify outliers, such as unusually high or low transaction values.
- Provides insights into customer behavior by analyzing repeat purchases.

Why Not Other Methods?

- Machine Learning Models (Regression, Clustering): Not required for simple summary statistics.
- Deep Learning: Overkill for straightforward trend analysis.

2. Trend Analysis (Time-Series Analysis)

Why Used?

- Sales and purchase records are time-dependent, so analyzing monthly and yearly trends helps in:
 - Predicting future revenue.
 - Identifying seasonal demand patterns.
 - Detecting growth opportunities.
- Allows visualization of revenue trends and their impact on business strategies.

Why Not Other Methods?

- Random Sampling or A/B Testing: Unnecessary since we are analyzing complete historical data, not conducting an experiment.
- Simple Averages: Would overlook fluctuations in demand across different months.

3. Demand & Market Trend Analysis

Why Used?

- Helps forecast inventory needs by analyzing frequently purchased medicines.
- Identifies underperforming medicines that may need price adjustments or promotions.
- Ensures optimal stock levels based on past sales patterns.

Why Not Other Methods?

- Neural Networks for Forecasting: Unnecessary for structured, tabular data with clear seasonal trends.
- Sentiment Analysis: Irrelevant since we are analyzing sales data, not customer reviews.

Results and Findings

1. Monthly Sales Trend

The analysis of monthly sales trends revealed fluctuations in revenue over different months. The data shows a pattern where sales exhibit seasonal variations, with peaks occurring in certain months. This could be due to increased demand for specific medicines during particular seasons, such as flu season or allergy seasons.

◆ Key Observations:

- A steady increase in sales revenue was observed in certain months, indicating growing demand.
- Some months showed a decline, which could be attributed to reduced customer visits, stock shortages, or external factors like holidays.
- The sales trend highlights a potential opportunity to optimize inventory and marketing efforts based on seasonal demand.

Graph Representation:

- A line graph was plotted to visualize the monthly sales trend, showing the variations in total revenue over time.
- The graph clearly illustrates high sales periods, which can be leveraged for promotional campaigns.

2. Profit and Loss

A comparison of the purchase and selling prices was conducted to determine profit margins across different brands of medicines.

◆ Key Observations:

- Profitable Medicines: Certain brands had consistently high profit margins, indicating a strong demand and optimal pricing.
- Loss-Making Medicines: Some medicines had little to no profit, suggesting either mispricing or low market demand.
- Break-Even Items: A few items were sold at nearly the same price as purchased, implying negligible profits.

Graph Representation:

- A bar chart was used to highlight the top 10 most profitable medicines.
- Another bar graph showed loss-making medicines, helping identify areas for price adjustments.

3. Top-Selling Medicines

The analysis of sales frequency by medicine brand revealed which products were in the highest demand.

◆ Key Observations:

- Some brands, like Cipla and Pfizer, consistently ranked among the best-selling medicines.
- The most popular medicines were those that catered to common ailments, indicating high consumer demand.
- Low-selling items might require targeted promotions or reconsideration of stock levels.

Graph Representation:

- A horizontal bar chart was used to display the top 10 best-selling medicines based on the quantity sold.
- The visualization helped in identifying customer preferences and aligning stock levels accordingly.

4. Demand Trends and Customer Preferences

An analysis of demand trends over time showed shifts in customer buying behavior.

◆ Key Observations:

- Medicines related to chronic illnesses had consistent demand throughout the year.
- Seasonal medications, such as those for flu or allergies, showed spikes during specific periods.
- Customer loyalty was observed for certain brands, suggesting a preference for trusted manufacturers.

Graph Representation:

- A trendline graph was used to show fluctuations in demand over time.
- The insights from this trend analysis can be used to adjust supply chain decisions and marketing strategies.

TOOLS USED:

- Python: I have used matplotlib, seaborn and pandas.

[CLICK HERE](#) for the code.

