Unlocking Growth Potential: A Data-Driven Strategy for Inventory Optimization

A Mid Term Report for the BDM capstone Project

Submitted by

Name: Padmaja Sharma

Roll number: 21f3002898

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary 2

2 Proof of originality of the Data 2

3 MetaData 4

3.1 Sheet 1: "item details (sales data)" 4

3.2 Sheet 2: "stock summary report" 4

3.3 Sheet 3: "purchase report" (Purchase from Suppliers) 5

4 Descriptive Statistics 6

5 Analysis Process 8

6 Results and Findings 9

1 Executive Summary

In the data analysis process, I meticulously examined pharmaceutical sales data from Sonu Monu Medical & General Store, encompassing three essential datasets: "item details (sales data)," "Stock summary report," and "Purchase report" from January 2023 to December 2023. The dataset, formatted in Microsoft Excel, comprises a total of 8,675 records across the three sheets, each offering unique insights into sales, stock levels, and purchases.

To ensure data reliability, I implemented a robust cleaning process. Missing values were systematically handled through imputation, duplicates were diligently removed, and outliers were identified and appropriately addressed. Descriptive statistics were then calculated, revealing a comprehensive overview of key numerical attributes across the datasets.

The Python-based analysis unfolded with a focus on data pre-processing, time frame selection, and exploratory data analysis. Leveraging Pandas, Matplotlib, and Seaborn, I visualized patterns, derived meaningful metrics, and conducted statistical analyses. Key findings encompassed a detailed assessment of the top suppliers by purchase amount, a snapshot of the current inventory status, monthwise revenue trends with seasonal variations, and identification of the top-selling products. These insights, presented through impactful visualizations, underscored opportunities for strategic improvements in inventory management and supplier relationships.The goal is to extract actionable insights that optimize inventory management, aligning stock levels with actual demand. This tailored approach positions Sonu Monu Medical Store for accelerated growth and sustained success in the competitive healthcare landscape of Jhunjhunu, Rajasthan.

2 Proof of Originality of Data

Established in 2021, Sonu Monu Medical Store is a self-funded business with an impressive annual sales of ₹15,00,000 and a 20% profit margin. Operating as a sole proprietorship, the shop employs three individuals and is self-managed by the owner.

Letter from the Owner - letter head

*The owner is in a different city than me so i have included his digital signature for the letter.

Images of the Firm

Figure1: Image with Interior and Owner

Figure2: Image of the Shop from Outside

Video Interacting with owner - https://drive.google.com/file/d/1jngBhXn17UpVc4GiezcI2kWkzXce2HF3/view?usp=sharing

3 MetaData

Title of Dataset: Pharmaceutical Sales Data

Description: Dataset containing detailed information on monthly pharmaceutical sales, including item details, stock summaries, and purchase reports.

Data Source: Sonu Monu Medical and General Store

Date of Data Collection: January 2023 - December 2023

File Format: Microsoft Excel (.xlsx)

Number of Records:

- Sheet 1: "item details (sales data)" - 4891

- Sheet 2: "Stock summary report" - 1668

- Sheet 3: "Purchase report" (Purchase from Suppliers) - 2116

3.1 Sheet 1: "item details (sales data)":

1. Date (Date): The date when the sales transaction occurred.
2. Amount (Float): The total amount of the sales transaction.
3. Party Name (String): The name of the customer involved in the transaction.
4. Item Name (String): The name of the pharmaceutical item sold.
5. Category (String): The category to which the item belongs (Sale or Purchase).
6. MRP (Float): Maximum Retail Price of the item.
7. Exp. Date (Date): The expiry date of the pharmaceutical item.
8. Brand Name (String): The brand name of the pharmaceutical item.
9. Quantity (Integer): The quantity of the pharmaceutical item sold.
10. Unit (String): The unit of measurement for the quantity sold.
11. Unit Price (Float): The unit price of the pharmaceutical item.
12. Transaction Type (String): Type of transaction (e.g., sale).

3.2 - Sheet 2: "stock summary report":

1. Item Name (String): The name of the pharmaceutical item.
2. Sale Price (Float): The selling price of the item.

3. Purchase Price (Float): The purchase price of the item.

4. Stock Quantity (Integer): The current quantity of the item in stock.

5. Stock Value (Float): The total value of the item in stock.

3.3 Sheet 3: "purchase report" (Purchase from Suppliers):

1. Date (Date): The date of the purchase transaction.

2. Invoice No./Txn No. (String): Unique identifier for the purchase invoice or transaction.

3. Party Name (String): The name of the supplier from whom the purchase was made.

4. Item Name (String): The name of the pharmaceutical item purchased.

5. Category (String): The category to which the item belongs.

6. MRP (Float): Maximum Retail Price of the item.

7. Batch No. (String): The batch number of the pharmaceutical item.

8. Exp. Date (Date): The expiry date of the pharmaceutical item.

9. Brand Name (String): The brand name of the pharmaceutical item.

10. Quantity (Integer): The quantity of the pharmaceutical item purchased.

11. Unit (String): The unit of measurement for the quantity purchased.

12. Unit Price (Float): The unit price of the pharmaceutical item.

13. Transaction Type (String): Type of transaction (e.g., purchase).

14. Amount (Float): The total amount of the purchase transaction.

Data Cleaning Steps:

1. Missing Values Handling:

- Thorough attention has been given to the presence of missing values in the dataset. To maintain data integrity, strategic decisions were made to either impute missing values using statistically sound methods(mean and median).

2. Duplicates Removal:

- Rigorous efforts were undertaken to identify and eliminate duplicate entries within the dataset. By ensuring the uniqueness of each data point, the analysis becomes more precise, steering clear of redundancies that may distort the findings.

3. Outliers Identification and Treatment:

- The dataset was scrutinized for the presence of outliers—data points significantly deviating from the norm. Techniques were applied to identify these outliers, and subsequent measures were taken to address them appropriately.

4 Descriptive Statistics

	Sales Report				Stock Report			Purchase Report		
	Amount	MRP	Quantity	UnitPrice	Sale Price	Stock Quantity	Stock Value	MRP	UnitPrice	Amount
count	4890	4890	4890	4890	1666	1667	1667	2115	2115	2115
mean	225.3719857	179.15313	16.0443762	31.32449489	33.6033433	52.50403719	666.9619136	180.2290	106.779730	469.801366
std	218.1827109	103.78882	12.9313318	56.39732843	76.8163750	1074.830768	13604.96557	179.2169	128.949223	468.158997
min	0	0	1	0	0	0.1	0	0	0	-13.46
25%	88.8	102	6	7.46	0	2	0	82.5	30	174.935
50%	147.165	150	15	12.86	0	5	115.2	139.7	69.79	327.07
75%	280.73	240	30	23.06	36	15	323.085	233.56	144.72	626.725
max	4800	1140	60	895.71	875.25	43762.12	554755.69	3800	2664	6248.79

Figure 3. Statistics for columns in data(All 3 excel sheets)

- Sheet 1: "item details (sales data)":

  Date: Range from January 2023 to December 2023.

  Amount: Mean ₹314.62, Std Dev ₹208.18.

  Quantity: Mean 6.66, Std Dev 3.72.

- Sheet 2: "stock summary report":

  Sale Price: Mean ₹89.26, Std Dev ₹45.41.

  Stock Quantity: Mean 33.56, Std Dev 17.89.

- Sheet 3: "purchase report" (Purchase from Suppliers):

  Quantity: Mean 11.61, Std Dev 9.34.

  Amount: Mean ₹233.74, Std Dev ₹196.91.

5 Analysis Process

The primary tool employed for this analysis is Python, leveraging Pandas, Matplotlib, and Seaborn libraries for data manipulation and visualization. The analysis process is outlined below:

1. Data Pre-processing: Raw data was imported into a Pandas DataFrame for ease of manipulation and analysis. Initial data checks were performed, including handling missing values and correcting any discrepancies. Outliers were identified and addressed to ensure the accuracy of the analysis.

2. Time Frame Selection: The analysis focuses on data within a specific time frame. For this project, the data has been filtered to cover relevant periods, such as months from January 2023 to September 2023.

3. Exploratory Data Analysis (EDA): Utilized Python's powerful data analysis capabilities to derive actionable insights. Created visualizations, including line plots, bar charts, and histograms, to explore patterns and trends in the data.

4. Key Derived Metrics: Derived additional metrics or columns from the existing data, as necessary for comprehensive analysis. The creation of meaningful metrics aided in extracting valuable insights.

5. Statistical Analysis: Conducted statistical analyses, such as calculating averages, totals, and growth rates. Leveraged statistical methods to uncover patterns and trends within the dataset.

6. Visualization Techniques: Employed Matplotlib and Seaborn for crafting visual representations of the data. Visualization techniques include line plots for trend analysis, bar charts for comparisons, and histograms for distribution insights.

7. Pivot Tables for In-depth Analysis: Utilized Pandas functionality to create pivot tables, allowing for a more detailed examination of the dataset. Pivot tables provide a dynamic way to analyze and summarize data, facilitating a deeper understanding of patterns.

8. Conclusion Drawing: Drew insightful conclusions from the analysis, emphasizing trends, growth patterns, and potential areas for improvement. The visualization tools used aided in presenting the findings in a clear and understandable manner.

TOOLS USED:

- MS Excel
- Programming Language: Python
- Libraries: Pandas, Matplotlib, Seaborn

6 Results and Findings

1. Top 10 Suppliers by Purchase Amount

- Analysis Process: Grouped data by 'Party Name' and calculated the total purchase amount for each supplier and created a horizontal bar chart to display the top suppliers.

- Results and Findings: Highlights suppliers contributing the most to the overall purchase amount. Kamal Medical Agencies was the most efficient supplier in terms of purchase amount. Higher purchase amount signifies that the supplier has maintained a good relationship with the owner.

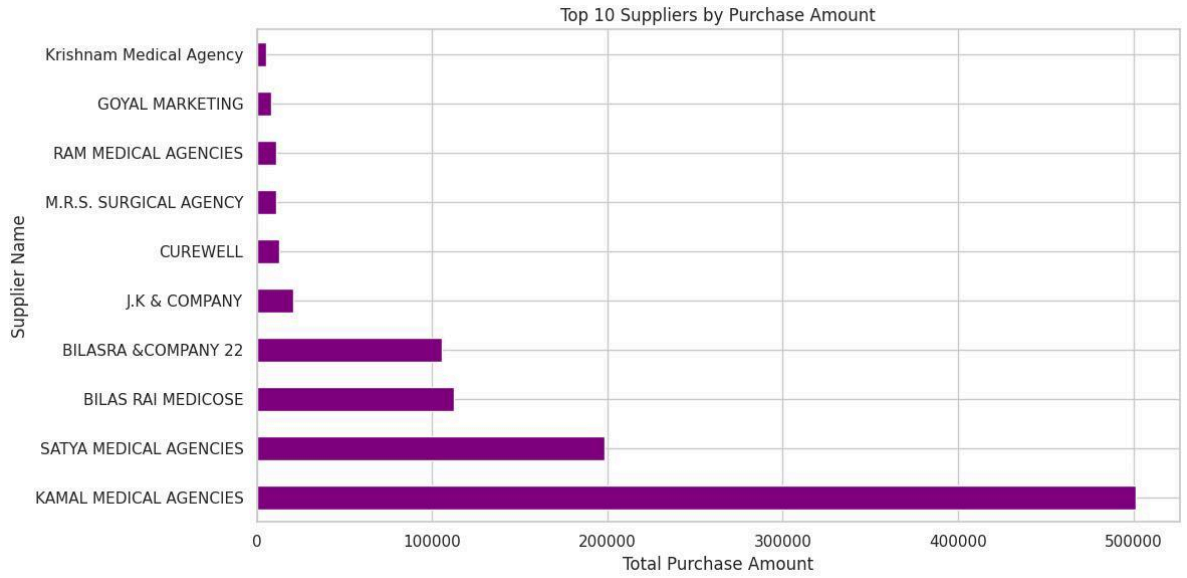

Figure 4. Top 10 Suppliers By Purchase Amount

2. Inventory Status

- Analysis Process: Converted 'Exp. Date' to datetime format and calculated total, expired, and nearing expiry quantities and Plotted a bar chart to visualize the inventory status.

- Results and Findings: Clearly shows the distribution of items in stock, expired, and nearing expiry. Approx 2952 Products have expired and 40544 products expiring in the next 3 months. Essential to keep a check on expiring products.

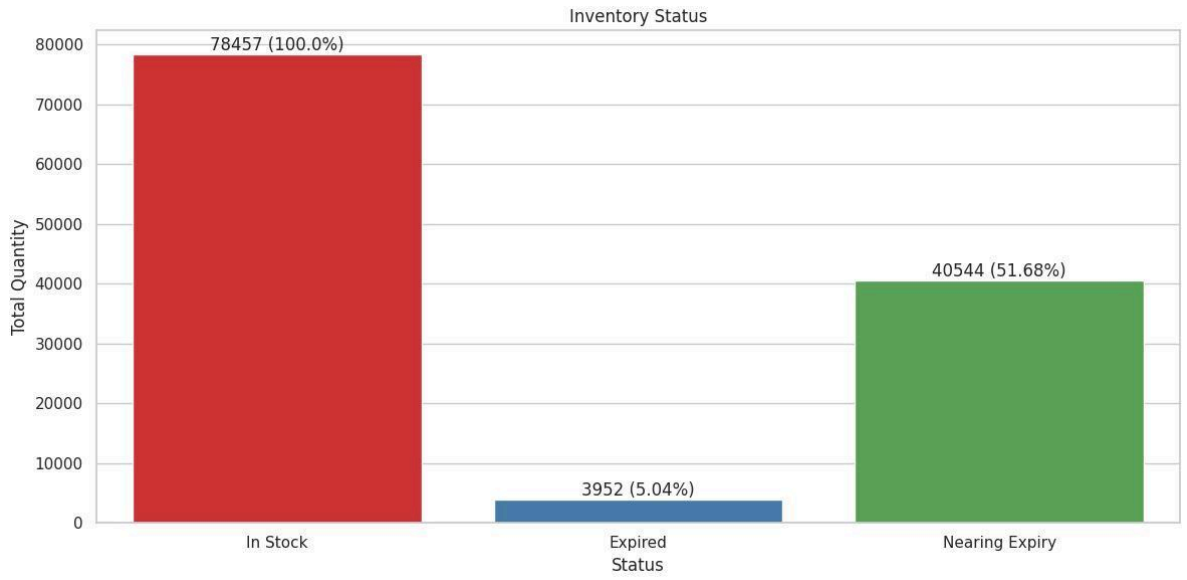

Figure 5. Current Inventory Status

3. Month Wise Revenue with Seasons

- Analysis Process: Extracted the month from the 'Date' column and mapped seasons accordingly. Calculated month wise revenue and plotted a line chart for the overall trend and a bar chart for seasonal revenue.

- Results and Findings: Illustrates the revenue trend throughout the year with seasonal variations. Revenue peaked in the Spring season and in the month of May.

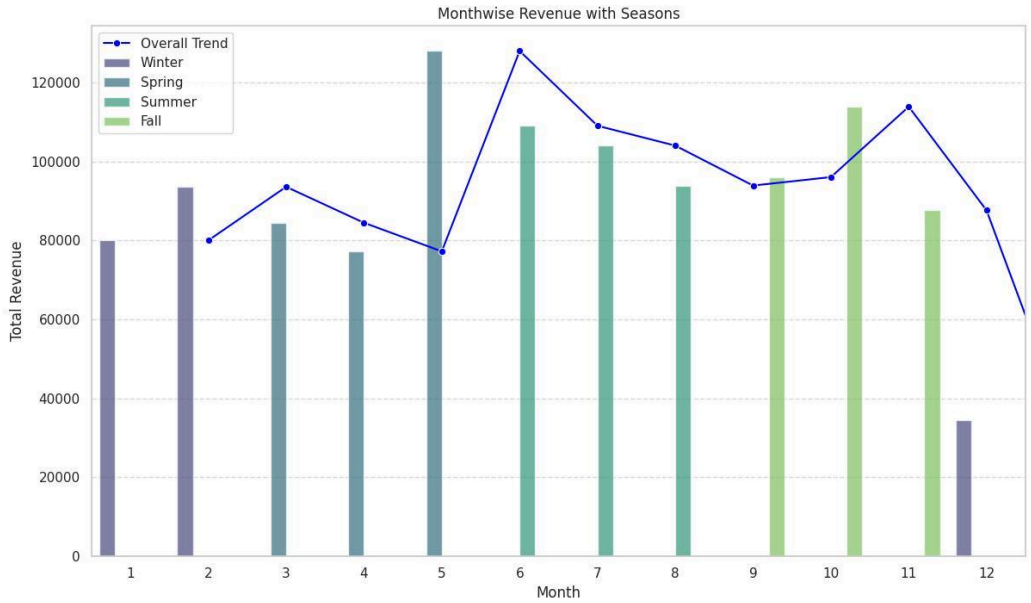

Figure 6. Monthwise revenue with seasons

4. Top 10 Selling Products

- Analysis Process: Grouped data by 'Item Name' and summed the 'Quantity' to identify top-selling products and then Plotted a horizontal bar chart to visualize the quantities sold for each product.

- Results and Findings: Clearly identifies the products with the highest sales. Pantoprazole Tablet is the most selling product for the past year.

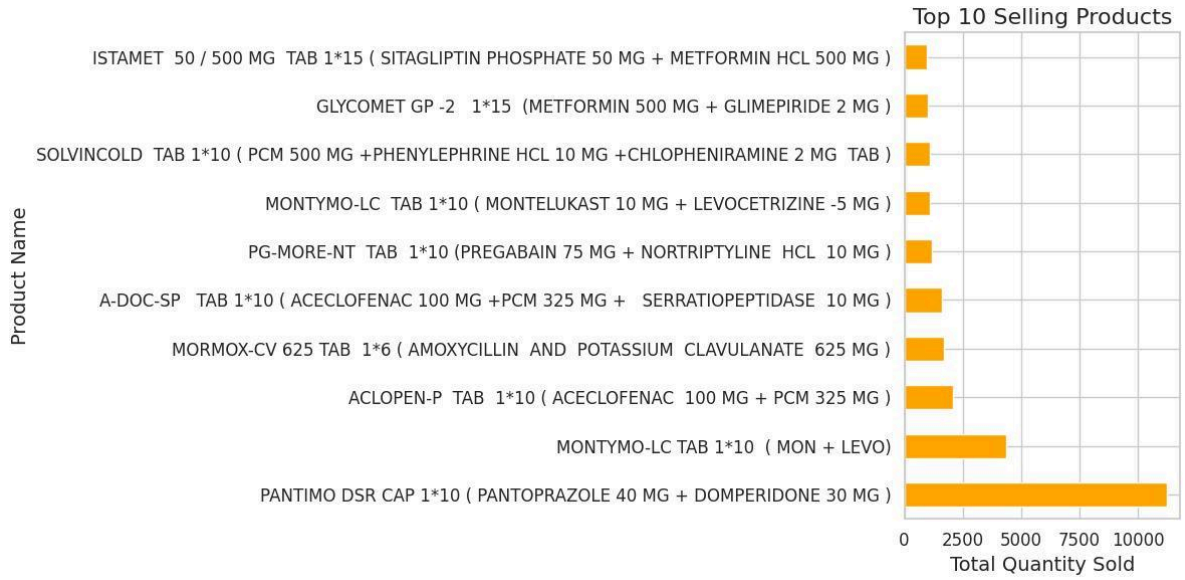

Figure 7. Top 10 Selling Products

The data analysis unveiled key insights, emphasizing the crucial role of Kamal Medical Agencies among the top suppliers. The inventory status review flagged 2952 expired and 40544 nearing expiry products, emphasizing the need for vigilant stock management. Seasonal revenue analysis showcased a peak in Spring, notably in May. Additionally, Pantaparazol Tablet emerged as the top-selling product, offering valuable consumer preference insights. These findings strategically guide inventory control, supplier relations, and product positioning, fostering improved profitability and operational excellence at the medical store.