Maximising Profitability: A Multi-headed approach for enabling Multi-Channel sales and Inventory Optimisation for T.R Departmental

A Mid-Term report for the BDM capstone Project

Submitted by

Name: **Watham Sabyasachi**

Roll number: **22f2000095**

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The project focuses on analyzing sales and inventory data for T.R. Departmental, a mid-sized convenience store in Hengrabari, Guwahati, to optimize profitability. The business faces challenges such as increasing competition from online retailers and high operational costs, which impact overall profit margins. The goal is to develop a data-driven approach to identify inefficiencies, optimize inventory management, and enhance sales strategies.

To achieve this, raw sales and inventory data were collected and consolidated into a structured format using Python and Google Sheets. The Metadata section provides an overview of the dataset, detailing key attributes like transaction details, product sales, and stock levels. Descriptive statistics were used to break down the data and uncover patterns, allowing for a better understanding of key metrics such as revenue, tax contributions, and sales performance over time.

A detailed analysis was conducted, focusing on various aspects such as monthly sales trends, SKU performance, mode of payment contributions, and customer purchasing behavior. Visual representations like line charts, bar graphs, and pivot tables were used to highlight crucial insights. Additionally, a time-series analysis helped track fluctuations in sales and revenue, identifying potential seasonal effects or external market shifts.

The results and findings from this analysis provide actionable insights that can help improve decision-making in stock management, sales strategy, and cost optimization. The study lays the foundation for further in-depth analysis to maximize profitability by addressing inefficiencies in inventory control and multi-channel sales strategies.

2. PROOF OF ORIGINALITY

Photo with owner

Front of store

TR DEPARTMENTAL

Hengrabari Road, Guwahati

To Whom It May Concern

I, Ohab Ali, the owner of TR Departmental, agree to let Watham Sabyasachi, bearing Roll number 22f2000095 , use my store's data for their project. This data will only be used for research and analysis related to the project.

I understand that Watham Sabyasachi will handle the data carefully and keep it private as needed. If at any time I want to stop sharing this data, I can inform them in writing.

I wish them the best for their future endeavours.

Sincerely,

TR Departmental

Ohab Ali

Proprietor

Ohab Ali,

Owner, TR Departmental

Letter of approval from owner

- Link to more photos here
- Link to video interaction with owner here
- Link to data : Google Sheet and Colab Notebook
- Link to map here

3. METADATA

The excel file contains the data obtained from the store. Link to the excel file containing raw data can be found here. The data was extracted from the store's billing system, the name of which was unfortunately kept private by the owner. It covers the store's sales from the month of September 2024 through December 2024.

➢ **Processed data**:

A	B	C	D	E	F	G	H	I	J	K	L	M	N
InvoiceDate	Month	CustomerName	ItemName	MOP	Qty	CurrStock	MRP	Purchase Rate	SRate	Amount	TaxPer	TaxAmt	NetAmt
2024-09-01	2024-09	CASH CUSTOMER	SUGAR 1 KG	Cash	1	42	55	39.05	52.38	52.38	5	2.62	55
2024-09-01	2024-09	CASH CUSTOMER	SUGAR 1 KG	Cash	1	42	55	39.05	52.38	52.38	5	2.62	55
2024-09-01	2024-09	CASH CUSTOMER	SUGAR 1 KG	Cash	1	42	55	39.05	52.38	52.38	5	2.62	55

Fig: Columns as seen in Data

The column names mentioned here have been retained with the criteria that these are the columns that are useful to the analysis. A short description of the same are given below:

- **CustomerName** : Name of the customer to whom item was sold
- **ItemName** : Name of the SKU
- **InvoiceDate** : Date at which item was sold
- **MOP:** Mode of Payment. Through what mode the payment was made
- **Qty** : Quantity bought by customer
- **CurrStock** : Current stock remaining after purchase
- **MRP** : Maximum Retail Price
- **Purchase Rate** : Rate at which SKU was purchased by the store
- **SRate** : Rate at which SKU was sold by the store to the customer (Selling Rate)
- **Amount** : Same as SRate, but we will be working using this
- **TaxPer** : Tax percentage applied on the item
- **TaxAmt** : What the tax percent amounts to
- **NetAmt** : Total amount payable by the customer (Revenue)

4. DESCRIPTIVE STATISTICS

```
df[desc].describe()
```

	Qty	CurrStock	MRP	Purchase Rate	SRate	Amount	TaxPer	TaxAmt	NetAmt
count	38570.000000	38570.000000	38570.000000	38570.000000	38570.000000	38570.000000	38570.000000	38570.000000	38570.000000
mean	1.450117	100.349780	68.662536	53.548875	63.241488	76.768865	8.719108	6.370932	83.139979
std	1.417149	1261.735872	91.756917	110.183738	85.414095	107.516064	7.779028	12.113196	114.991193
min	-6.000000	-1.000000	0.000000	0.000000	0.000000	-540.000000	0.000000	-58.930000	-550.000000
25%	1.000000	3.000000	25.000000	17.960000	22.320000	30.000000	0.000000	0.000000	33.000000
50%	1.000000	10.000000	46.000000	35.000000	42.370000	50.850000	5.000000	3.050000	55.000000
75%	1.000000	33.000000	80.000000	61.610000	75.000000	85.710000	18.000000	8.040000	95.000000
max	49.000000	48905.000000	2400.000000	7235.000000	2400.000000	4800.000000	18.000000	320.360000	4800.000000

Table 1: Descriptive statistics for numerical columns

- The dataset reveals key insights into sales, pricing, and taxation trends. The quantity sold (Qty) averages 1.45 per transaction, but negative values suggest returns or data errors. Stock levels are highly imbalanced, ranging from -1 to 48,905 units.
- Pricing data shows wide variation, with an average MRP of ₹68.66 and a mean selling rate of ₹63.24. Some items have ₹0 prices, possibly due to discounts or misentries. The purchase rate varies significantly, with certain high-value items (max ₹7,235).
- Revenue (Amount) fluctuates between -₹540 and ₹4,800, hinting at refunds or adjustments. Tax rates range from 0% to 18%, with negative values possibly indicating exemptions or reversals. The data contains outliers and inconsistencies, requiring further cleaning and validation for accurate analysis.

Fig 1: Total Monthly Sales

- Through a quick visual representation in the attached chart above, we can observe fluctuations in monthly revenue, with **October** showing the highest sales, while **December** and **November** experienced a slight decline compared to the previous months.

5. DETAILED EXPLANATION OF ANALYSIS PROCESS AND METHODS

5.1. Data Extraction

The dataset was initially stored in multiple month-wise Excel files. To streamline the analysis, Python was used to consolidate these files into a single dataset. Pandas was instrumental in handling the merging process efficiently, ensuring that all data points were accurately combined. Once consolidated, the dataset was imported into Google Sheets for further manipulation and ease of accessibility.

5.2. Data Cleaning and Preprocessing

The raw data contained redundant and empty columns that provided no meaningful insights. After a thorough manual inspection, these unnecessary columns were removed using Pandas. The InvoiceDate column was identified as a key candidate for indexing, which facilitated time-series analysis. Additionally, missing values were handled appropriately to maintain data integrity, and numerical values were converted into appropriate formats for accurate calculations.

5.3. Exploratory Data Analysis (EDA)

Several exploratory analyses were conducted to gain a preliminary understanding of the dataset:

- **Descriptive Statistics:** Using df.describe(), key statistical measures such as mean, standard deviation, and percentiles were computed for numerical columns like Qty, MRP, Purchase Rate, SRate, and NetAmt. This helped identify trends, anomalies, and possible outliers.
- **Identifying Top Performing SKUs:** Pivot tables were created in Google Sheets to determine the top 10 SKUs by total revenue and total number of units sold, helping to highlight the best-selling products.
- **Payment Mode Contributions:** The distribution of revenue across different modes of payment (Cash, Card, Credit) was analyzed, revealing their percentage contributions to the total sales. This could serve as a preliminary form of customer segmentation based on payment preferences.

- **Word Cloud Analysis:** A word cloud was generated using Python's WordCloud library to visualize the most frequently sold items, aiding in the identification of popular SKUs.

5.4. Time-Series Analysis

Time-series analysis was a major focus of this study to understand sales trends over time. Several key analyses were conducted:

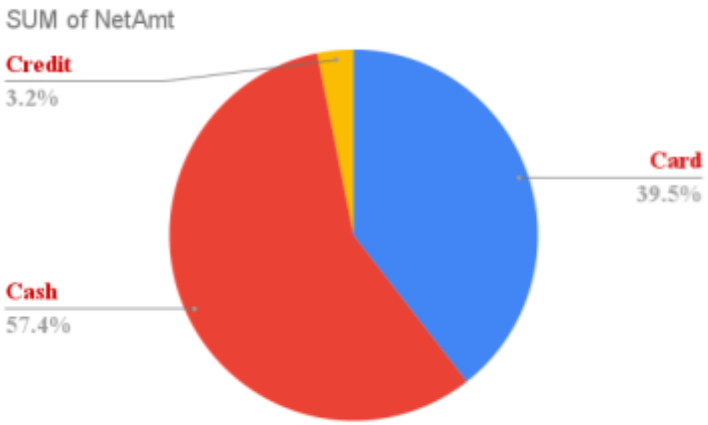
- **Monthly Revenue Trends:** A pivot table was used to calculate the total NetAmt for each month, and the results were plotted as a line graph in Google Sheets. This revealed fluctuations in sales performance over different months.
- **Weekly Sales Trends:** Weeks were extracted from InvoiceDate using Google Sheets formulas, and a line graph was created to track sales on a weekly basis.
- **Seasonality and Market Disruptions:** The analysis indicated a stable sales trend in October, a volatile November with peaks and declines, and a sharp drop in revenue towards late December and early January, possibly due to seasonal factors.

5.5. Visualization and Reporting Tools Used

- **Python (Pandas, Matplotlib, WordCloud):** Used for data cleaning, transformation, visualization, and generating word clouds.
- **Google Sheets:** Used for pivot tables, additional manipulation, and visualization of trends using charts.

6. RESULTS AND FINDINGS

6.1. Mode of Payment



MOP	NetAmt
Cash	1839636.76
Card	1265234.23
Credit	101838.00

Fig 2: MOP breakdown

From this breakdown, it is evident that cash transactions dominate with over half of the total payments, suggesting that a large portion of customers prefer immediate payments. Card payments account for a relatively smaller portion, indicating that a substantial number of customers opt for digital transactions, possibly for convenience or reward points. The credit-based transactions contribute only 3.2%, implying that very few customers rely on deferred payment methods.

NOTE: Card payments conveniently mean UPI payments and not through Debit or Credit cards.

6.2. Weekly Sales

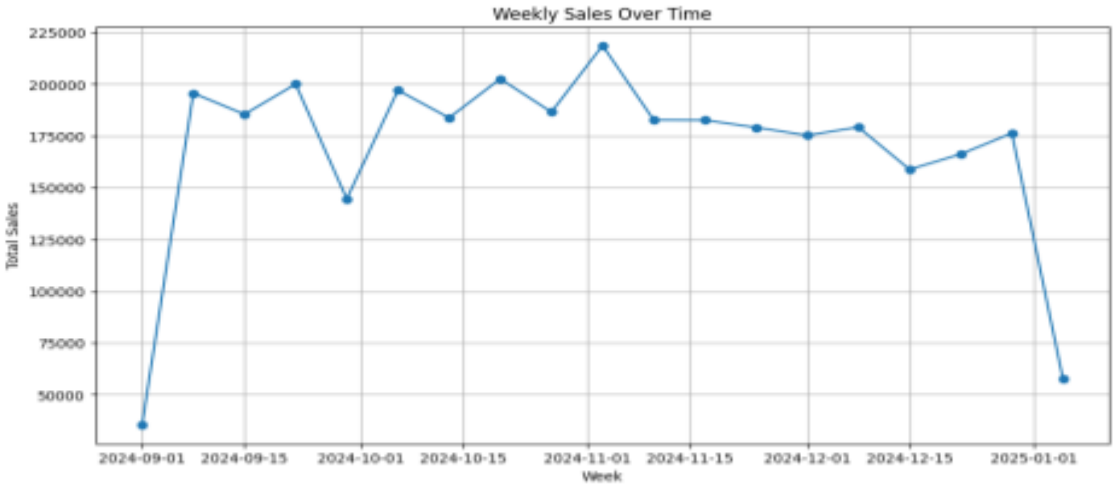


Fig 3: Weekly Sales Trend over time

- Looking at the weekly sales trend, the graph shows that October had steady sales, while November was more inconsistent.
- November had a peak week of around ₹220,000, but sales dropped to about ₹180,000 in the following weeks, which may have reduced the overall performance. This shows that even though there was one strong week, November's sales might not be as high as October's more stable sales.
- This then continues through December, staying stagnant and possibly being attributed to seasonal similarities.

6.3. **Top performing SKUs**

ItemName	SUM of NetAmt
DHARA KACHI GHANI MASTERD OIL 1L	57334.45
KAMAKHYA PANEER 200G	53128.85
PUROBI STANDARD MILK 500ML	39933.28
POTATO 1KG	39376.85
PORI MSOOR DAL 1KG	33816.5
EGG PLATE	33111.9
AMUL TAZZA 1LTR	32553.34
ONION 1KG	31672.4
DHARA KGMOB 1LITRE	27219.2
BPT RICE LUNIA 1KG	26564

Table 2: Top performing SKU by revenue

- The above table shows the top 10 SKUs by revenue generated.
- Generated using a pivot table, this analysis helps identify the highest-grossing products, providing insights into which items contribute the most to overall sales. By organizing the data in this way, it becomes easier to recognize trends, optimize inventory, and make informed business decisions based on product performance.

ItemName	SUM of Qty
EGG	2780
MAGGI 52GM	1401
AMUL TAZZA 200ML	1300
PUROBI STANDARD MILK 500ML	1175
POTATO 1KG	893
AMUL TAZZA 500ML	625
KAMAKHYA PANEER 200G	556
ONION 1KG	464
AMUL TAZZA 1LTR	440
SUGAR 1 KG	405

Table 3: Top performing SKU by quantity sold

- The above table shows a list of the top 10 SKUs with respect to quantities sold in total.
- This analysis highlights the best-selling products in terms of quantity, helping to identify which items are in the highest demand. Understanding these trends can assist in optimizing stock levels, improving supply chain decisions, and ensuring that high-demand products are always available to meet customer needs.
- SKU such as Purobi Standard Milk 500ml and Onion 1kg were common to both the tables indicating strong candidacy of being high revenue and high demand SKU. As the analysis progresses more such SKUs can be identified.

Fig: Wordcloud showing most popular items

This is just a preliminary analysis aimed at identifying key trends and patterns in the data. While it provides valuable insights into sales performance, customer preferences, and revenue distribution, a more in-depth analysis shall be conducted for the final report. The upcoming stages will include deeper statistical evaluations, segmentation studies, and trend forecasting to gain a comprehensive understanding of the business dynamics. The problem of online delivery will also be tackled using appropriate research.