Optimizing Inventory, Data Management and Branding Strategy at Music & Music

A Mid-term report for the BDM capstone Project

Submitted by

Name: Kitiksha Chakrawarti

Roll number: 23f3001245

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary

Music & Music, founded in 1990 by Mr. Rajendra Prajapati, has evolved from a cassette duplication service into a retail shop selling packaged food, cold drinks, and digital services. Despite stocking popular items and staying open for over 12 hours daily, the shop faces persistent challenges of overstocking and underperforming items, which directly affect sales, revenue, and profitability. Additional operational hurdles include difficulties complying with Indore's plastic ban and the absence of promotion for lesser-known services like recharge and D2H.

In the first phase, the project focused on digitizing incoming bills and handwritten sales records to understand inventory dynamics. Analyses included supplier-wise purchase patterns, daily restocking and selling trends, and exploration of top-performing items like cold drinks and ice creams that contribute most to revenue and profit. These findings were shared with the shop owner, who found visual insights like trend plots particularly useful for planning.

A key challenge identified was the mismatch between incoming and outgoing data, as outgoing records capture only sales amounts without item-level detail. To address this, the idea of a lightweight token-based system is being explored to log item names quickly without adding manual burden.

Next steps will extend data collection to May and June to capture seasonal demand shifts, apply regression models to forecast inventory needs more accurately, and explore targeted branding strategies for underperforming categories. The project will also work toward building a predictive alert system to notify the shop owner before key items run low.

2 Proof of originality of the Data

- Name of shop- Music & Music
- Address- 52, Vrindavan Colony, Banganga Main Road, Indore (M.P)
- Owner/ Founder- Mr. Rajendra Prajapati

- *Letter from Organization link*- Letter from organization
- *Interaction video with shop owner link*- Interaction video
- *Raw data link*- Music & Music data
- *Images with the founder/ service scape link*- Shop images

3 Meta-Data

The data for this project was collected in two parts- incoming data and outgoing data. The incoming data was collected from physical purchase bills sourced from various suppliers from **1 April, 2025** to **30 April, 2025.** The outgoing data was collected from handwritten daily sales record maintained by the shop owner from **1 April, 2025** to **15 April, 2025**.

To digitize these records, the data was manually entered into the spreadsheet, resulting in two structured datasets :-

3.1 Incoming dataset

- Data format - CSV (Comma Separated values) and XLSX (Excel open XML Spreadsheet)
- Collection Period - 1 April, 2025 to 30 April, 2025
- Source - Physical purchase bills from various suppliers
- Raw data link - Purchase data'25

Original Structure (Wide Format)

The raw incoming dataset was entered manually in a spreadsheet with one row per item and thirty daily quantity columns. The table below describes each column of incoming data as it was recorded initially.

Column	Description
Item	Name of the product (e.g., "Mini Matka").
Category	High-level grouping (e.g., "Ice cream & frozen")
Supplier	Supplier name
Per unit price (MRP)	Maximum Retail Price per unit (in rupees)
Per unit wholesale rate	Cost price per unit (in rupees)
Profit per unit item	MRP minus wholesale rate
1, 2, …, 30	Quantity restocked each day (columns labeled by day of April)
Total Incoming Qty	Sum of daily quantities
Total Cost	(Total Incoming Qty) × (Per unit wholesale rate)
Potential revenue	(Total Incoming Qty) × (Per unit price (MRP))
Potential total profit	(Total Incoming Qty) × (Profit per unit item)

Spreadsheet link (Horizontal format): Inventory'25

Transformation to Long (Tidy) Format

To enable time-series analysis and category level analysis, the wide table was converted to a vertical (CSV) format with one row per item. The table below describes each column.

Column	Description
Item	Item name
Category	Grouping into 5 categories
Supplier	Supplier name
Per unit price (MRP)	MRP per unit (in rupees)
Per unit wholesale rate	Cost price per unit (in rupees)
Profit per unit item	Unit-level profit (in rupees)
Day	Day of April (1–30)
Quantity	Quantity restocked on that day

CSV file link (Vertical format): vertical_data_incoming.csv

3.2 Outgoing dataset

- Data format - CSV (Comma Separated values) and XLSX (Excel open XML Spreadsheet)
- Collection Period - 1 April, 2025 to 15 April, 2025
- Source - Handwritten daily sales record maintained by shop owner in physical registers
- Raw data link - Sales data

Original Structure (Wide Format)

The raw outgoing dataset was entered manually in a spreadsheet in a wide format. For every category, there were separate columns capturing Total sale, profit and quantity sold on a particular day. The table below describes the columns

Column examples	Description
Pan Total sale, Pan Profit, Pan Quantity sold	Total daily sales in ₹, total profit, and quantity sold for Pan on that day
Cold drink Total sale, Cold drink Profit, Cold drink Quan	Same for Cold drink
...	and so on for all other categories

Spreadsheet link (Horizontal format): Outgoing data'25

Transformation to Long (Tidy) Format

To enable daily and category-level analysis, the data was transformed into a long format (CSV), where each row corresponds to a specific day and category, with these columns:

Column	Description
Day	Day of the month (1–15)
Total sale	Total sales amount (₹) for that category on that day
Profit	Total profit (₹) for that category on that day
Quantity sold	Number of units sold (available for some categories)
Category	Name of the category (e.g., Pan, Cold drink, Ice cream, etc.)

CSV file link (Vertical format):outgoing_data_filtered.csv

4 Descriptive Statistics

Descriptive statistics- Incoming data

	Per unit price (MRP)	Per unit wholesale rate	Profit per unit item	Day	Quantity	Total Cost
count	288.000000	288.000000	288.000000	288.000000	288.000000	288.000000
mean	44.621094	31.738017	12.883077	15.739583	39.677083	462.964972
std	60.929292	43.059739	18.127133	8.830078	50.532529	565.008847
min	5.000000	3.622500	0.250000	1.000000	1.000000	38.900000
25%	10.000000	7.200000	2.542500	8.000000	10.000000	199.320000
50%	20.000000	13.890000	5.762857	17.000000	24.000000	333.560000
75%	40.000000	26.370000	10.677500	22.000000	48.000000	547.990000
max	300.000000	208.470000	91.530000	30.000000	420.000000	7001.640000

- The average MRP is ₹44.6 (range: ₹5–₹300) and average wholesale rate is ₹43, both showing high variation. Premium-priced items like special family packs of ice-creams and large cold-drink packs account for the upper extremes.
- Most products yield low profits (₹0.25–₹10), with a few niche items earning up to ₹90, indicating a focus on low-margin daily-use items.
- Items are restocked at an average of 40 units/day (median 24), with ice-creams restocked most heavily (up to 420 units) due to perishability and seasonal demand. Other items show frequent but smaller restocks.
- The average daily cost per item is around ₹463, with a few high-cost items pushing the maximum to nearly ₹7,000. While most items stay under ₹550 (75th percentile: ₹547.9), bulk restocking of multiple ice-cream varieties from the main supplier, GJ's Scoopy Spoon, accounts for the higher costs.

Descriptive statistics by category- Quantity Restocked (per item × day)

Category	count	mean	std	min	25%	50%	75%	max
Biscuits and chocolates	24.0	39.25	20.86	10.0	27.0	33.0	43.5	84.0
Cold Beverages and dairy	65.0	32.38	31.37	1.0	15.0	28.0	41.0	240.0
Ice cream & frozen	175.0	41.65	60.60	1.0	8.0	24.0	48.0	420.0
Instant and pantry items	7.0	54.29	31.23	16.0	36.0	40.0	78.0	96.0
Namkeen & snacks	17.0	41.88	24.12	10.0	24.0	36.0	60.0	90.0

- **Instant and Pantry items**, though having least variety, are consistently restocked and have highest average restocking per day(₹ 52.29). This might be because of 1-2 widely popular items like Maggi 2 min noodles and coffee sachets.
- **Ice cream and frozen** items have the maximum number of unique items and a huge difference between 75th percentile and maximum of 420 item restock per day suggests major restocking days within an interval. The same conclusion goes for **Cold beverages and dairy** category.
- **Biscuits and chocolates** and **Namkeen & snacks** have lower average quantities but are consistently restocked (median ~33–36).

Descriptive statistics by category- Profit per Unit Item (₹)

Category	count	mean	std	min	25%	50%	75%	max
Biscuits and chocolates	24.0	2.19	2.83	0.45	0.54	1.23	2.30	12.17
Cold Beverages and dairy	65.0	12.65	13.27	0.25	1.50	6.99	17.29	39.11
Ice cream & frozen	175.0	15.99	20.81	1.50	3.05	6.10	12.20	91.53
Instant and pantry items	7.0	2.20	0.72	1.11	1.74	2.71	2.71	2.71
Namkeen & snacks	17.0	1.30	0.52	0.84	1.11	1.11	1.21	2.35

- Categories other than **Ice cream** and **Cold drink** have thinner profit margins of ₹(1-2). These categories also don't contribute much to sales. Hence, there is a need to ensure items belonging to these categories are not overstocked.

Descriptive Statistics- Outgoing data

	Category	total_sales	total_profit	avg_daily_sales	avg_profit_margin(%)
7	Others	58965.0	10787.40	3931.000000	18.29
4	Ice cream	45372.0	7278.27	3024.800000	16.04
1	Cold drink	25765.0	4685.35	1717.666667	18.18
8	Pan	12448.0	1800.80	829.866667	14.47
5	MB	11745.0	331.29	783.000000	2.82
3	Electronic	7867.0	21.00	524.466667	0.27
6	MR	4163.0	111.16	277.533333	2.67
0	Aim	3400.0	40.00	226.666667	1.18
2	D2H	664.0	20.92	44.266667	3.15
10	Roll	550.0	150.00	36.666667	27.27

- **Ice cream** and **Cold drink** dominate both sales and profit. Their average daily sales are significantly higher than other categories. Profit margins are also healthy (~16-18%).
- **D2H, MR, MB** (Mobile recharge & Mobile balance) and other services are occasional sellers. Hence, they underperform both in sales and profit.

Descriptive statistics by category-Total sales(1-15 April)

Category	count	mean	std	min	25%	50%	75%	max
Aim	15.0	226.67	559.93	0.0	0.0	0.0	0.0	2000.0
Cold drink	15.0	1717.67	570.57	915.0	1242.5	1680.0	2060.0	2855.0
D2H	15.0	44.27	117.44	0.0	0.0	0.0	0.0	364.0
Electronic	15.0	524.47	1773.41	0.0	0.0	0.0	0.0	6867.0
Ice cream	15.0	3024.80	498.94	2175.0	2622.5	3070.0	3430.0	3690.0
MB	15.0	783.00	479.74	0.0	376.5	841.0	1166.5	1719.0
MR	15.0	277.53	386.89	0.0	0.0	91.0	463.5	1345.0
Others	15.0	3931.00	836.70	2538.0	3422.5	3668.0	4502.0	5606.0
Pan	15.0	829.87	180.97	380.0	732.5	851.0	932.5	1082.0
RMC	15.0	0.00	0.00	0.0	0.0	0.0	0.0	0.0
Roll	15.0	36.67	115.12	0.0	0.0	0.0	0.0	440.0

- Services such as **mobile recharge**, **Airtel recharge** are still stable despite being in less quantity.
- However services such as **D2H, Roll(Camera reel), RMC(Microchip/ pen drive/ CD) recording, Aim(Airtel money transfer through google)** show almost no sales on most days (median=0) but very high max values.
- Pan has moderate daily sales with less variation, suggesting it is a small but steady contributor.

5 Detailed explanation of Analysis Process with Justification

Choice of tools with justification

After initial cleaning and manual entry in Google Sheets, the data was exported as CSV files for structured analysis.
Google Colab was chosen as the primary analysis environment because it supports combining code, explanations and outputs in one place
Key Python libraries used:

- **pandas**: for data cleaning, reshaping from wide to long format, aggregation, and summary statistics
- **matplotlib** & **plotly**: for interactive visualizations suitable for any number of items, including bar plots, pie charts, and line charts that help explain trends clearly
- **numpy**: for numerical operations and efficient data manipulation
- **dataprep**: an advanced EDA library known for its fast, automated data profiling and visualization — used to quickly identify missing values, outliers, and distributions

In early stages, **Google sheets** was used for:

- Filtering and sorting
- Built-in functions (e.g., SUM, AVERAGE, SUMPRODUCT, VLOOKUP) for quick calculations
- Pivot tables for supplier-wise, item-wise, and category-wise summaries
- Preliminary visualizations (bar plots, pie charts, line charts)

Google sheets was ideal for manual data entry whereas python provided a means for automated and advanced analysis.

Analysis process in detail

1. **Google Sheets Preliminary Analysis**

- Used filtering to identify top performing items in terms of: Incoming stock quantity, Total cost, Potential revenue, Potential profit
- Plotted quick visual summaries: Pie chart showing distribution of items by category → highlighted that Ice cream & frozen and Cold Beverages & dairy dominated stock and Bar plot for top 10 most restocked items.

Built pivot tables to analyze: Total cost, potential revenue, and potential profit by category and supplier
Drew a line plot of total restocking volume per day (1–30 April) which helped to identify whether restocking was evenly spread or clustered.

2. In python (deeper analysis of incoming data)

Plotted:

- Bar chart: top 10 items by total restocking quantity → to spot high-volume drivers
- Bar chart: supplier-wise share in total inventory cost → to identify dependence on suppliers
- Bar chart: top 10 items by unit profit margin → to see which items are high-profit
- Clustered column chart: restocking trend of top 5 items → to check restocking trend of top 5 items and identify patterns
- Scatter plot (with quadrants): fast vs slow-moving items based on restocking qty & profit per unit.
- Computed descriptive statistics across items, categories & profit.

The above analysis helped visualize not just which items sell more, but which are profitable. Scatter plot helped identify overstocking and understocking of items.

3. Outgoing data analysis (first 15 days):

Plotted:

- Bar charts: total sales & profit by category → rank categories by revenue & profitability
- Line plots: daily total sales & profit → visualize day-to-day fluctuations
- Pie chart: category share of total sales.

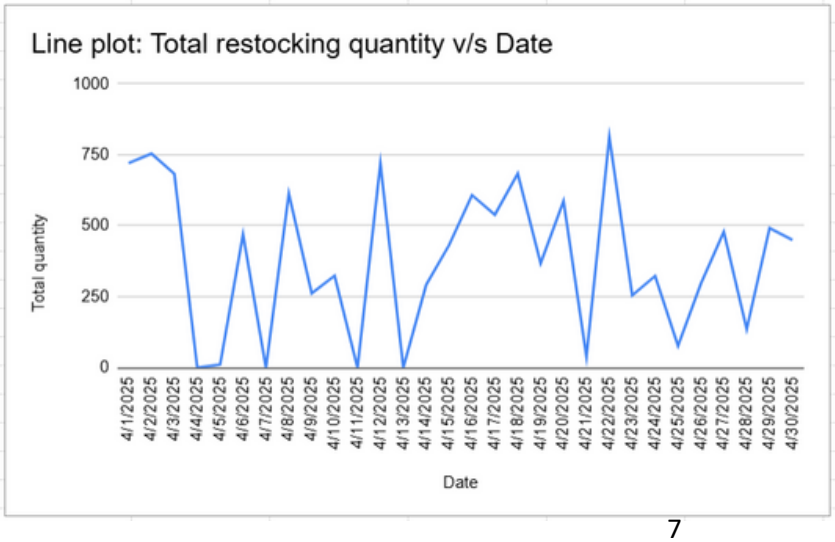
Computed summary stats per category: Total sales, total profit, average daily sales, average profit margin.
Used GROUPBY to compute descriptive statistics by category.

4. Incoming vs Outgoing alignment & comparison:

- Aligned common categories from incoming and outgoing datasets
- Plotted line plots of incoming quantity vs quantity sold for 15 days → to see if stock matches demand.
- Computed: Expected vs actual revenue (based on potential revenue from incoming data vs actual sales), Expected vs actual profit, Profit realization ratio, revenue realization ratio. These were calculated to understand performance of stock.

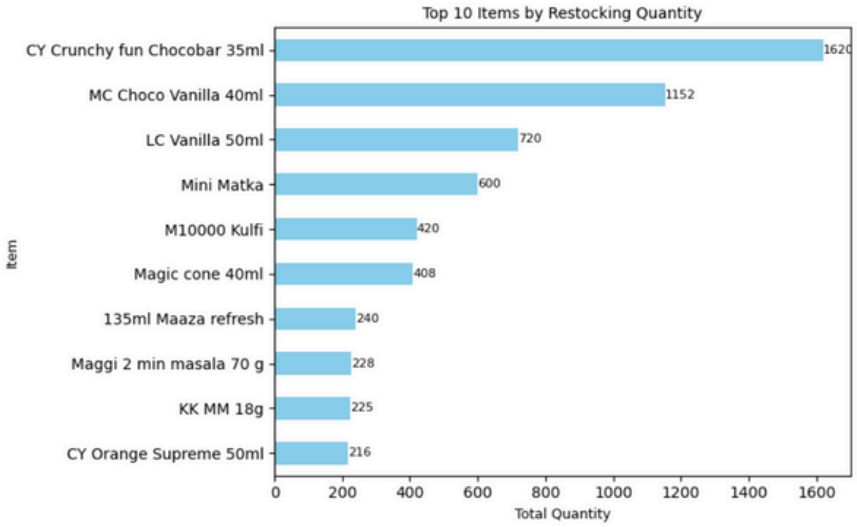
Note: The incoming data of *Pan Masala*, which dominates the outgoing data is not available in any form. The *Others* category in Outgoing data includes many things like- biscuits, namkeen, ice, plastic glasses, stationary chips etc. Also, item-wise outgoing data is not available. Only the category and MRP of selling item is known.

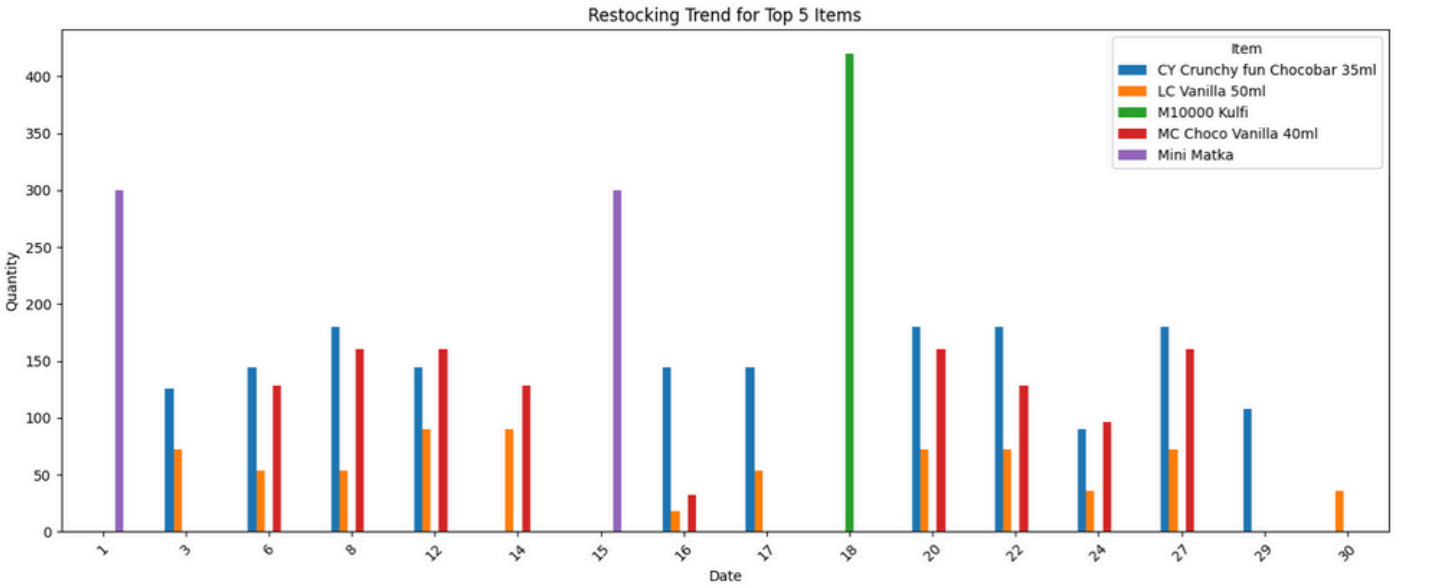
6 Results and Findings



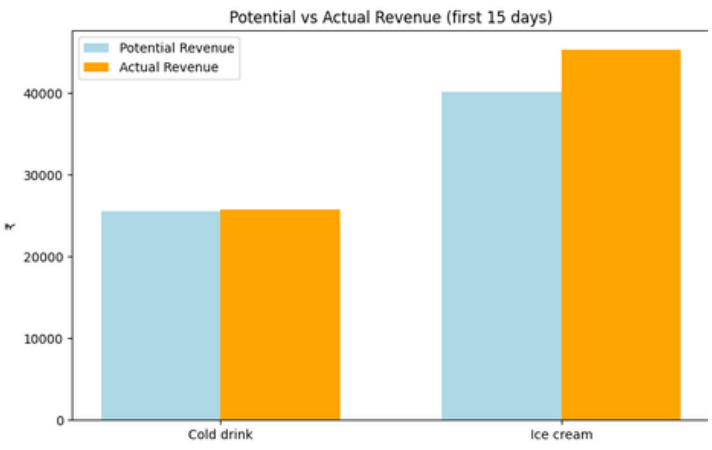
- This line chart reveals a major stocking every 10 days when more than 750 items are stocked.
- This will be helpful to predict restocking requirements in advance in future.

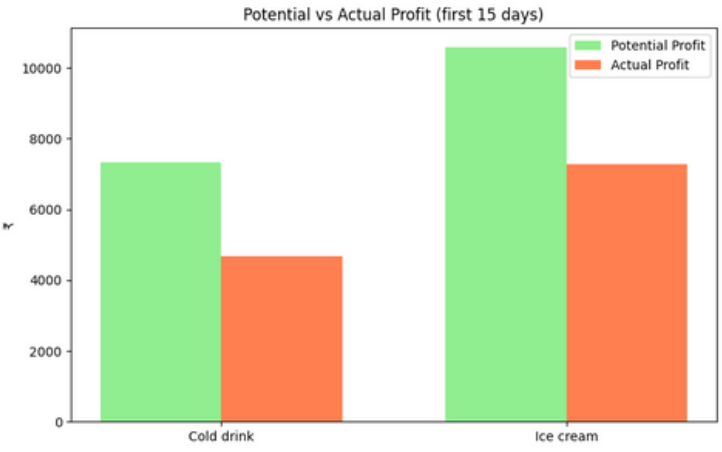
- The shop heavily restocks items like *CY Crunchy fun Chocobar* (1620 units) and *MC Choco Vanilla 40ml* (1152 units) and *LC Vanilla* (720 units). These 3 products come from the same supplier- *GJ's Scoopy Spoon*.
- Other popular items are *Matka kulfi* and small packs of *Kurkure*, *Maggi noodles* and *Maaza*.



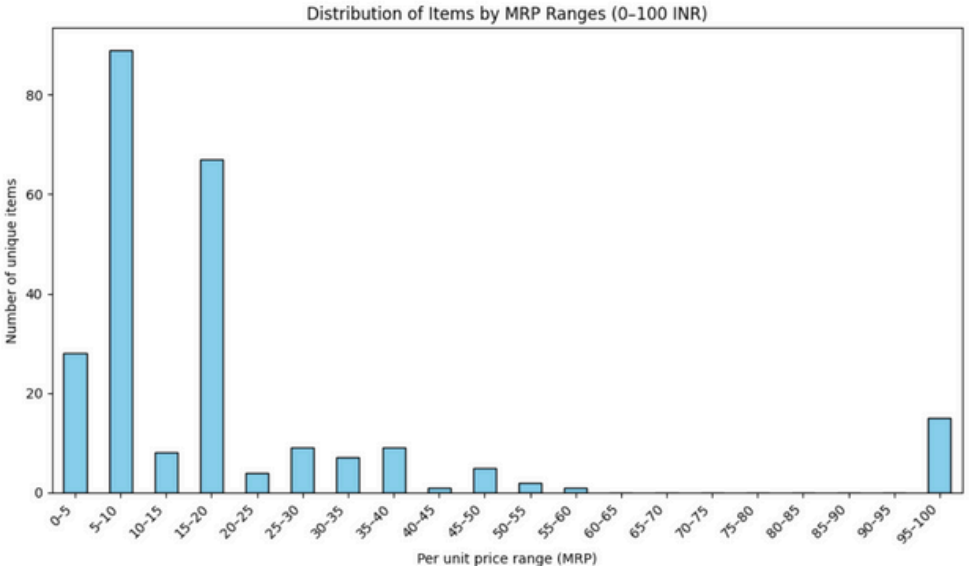


- *Matka kulfi* follows a 15 day restocking trend in bulk.
- The restocking of 3 items- *CY Crunchy fun chocobar, LC Vanilla* and *MC Choco vanilla* are restocked every 3-4 days in almost same proportion each day, this can be explained as they come from same supplier.

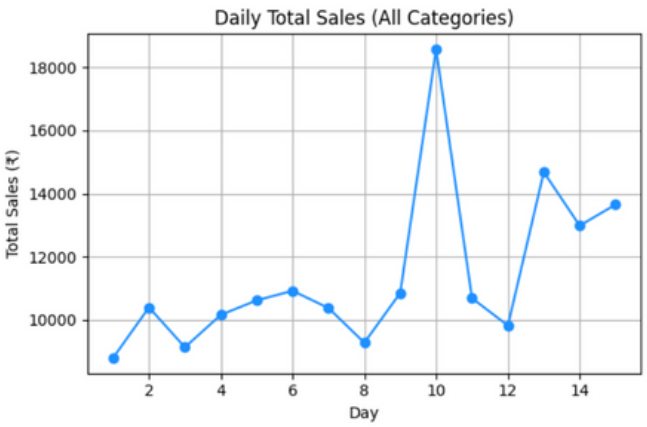


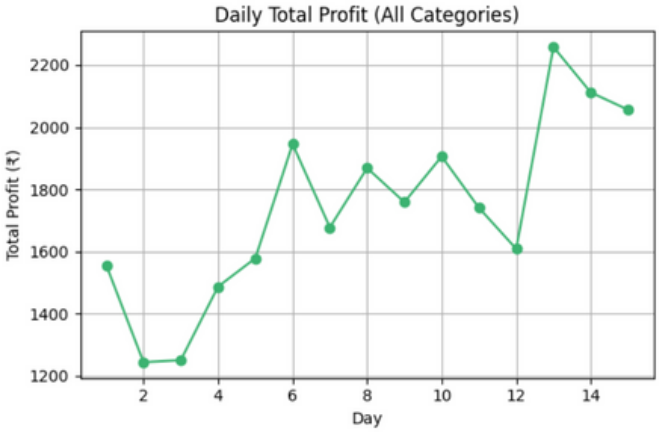


- Both categories underperformed in terms of profit. Despite the sales of ice cream exceeding the Potential revenue, it still has a large profit gap. This again indicates overstocking.
- Revenue realization ratio for ice- cream and cold drink- 113% and 101% respectively
- Revenue realization ratio for ice- cream and cold drink- 69% and 64% respectively
- Profit can significantly improve for the shop if it's top 2 categories are managed efficiently.

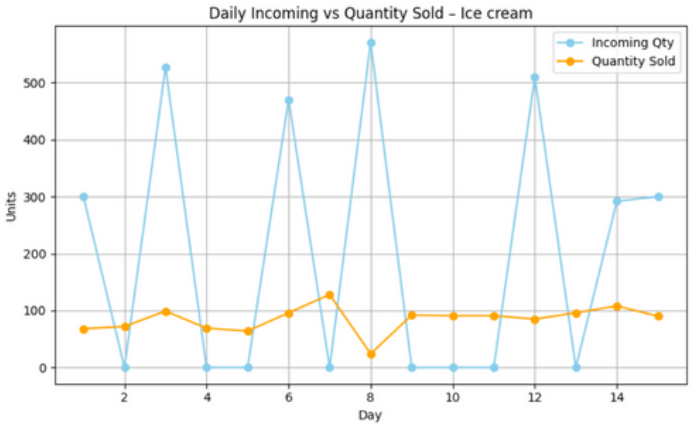


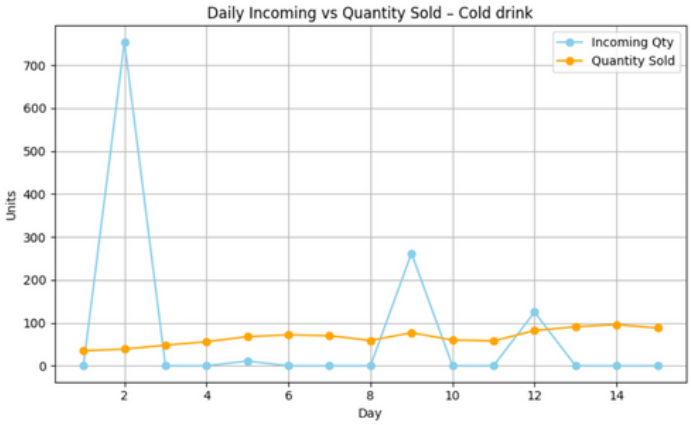
- Most items fall in the affordable ₹5–20 MRP range, which aligns well with the local customer base's price sensitivity and helps maintain high daily turnover.





- On Day 10, sales peaked (₹18,500), but profit was only moderate (₹1,920), indicating low-margin or discounted sales. This anomaly needs to be further investigated.
- On Day 13, profit was the highest (₹2,250) despite only moderate sales (₹14,800), suggesting sale of high-margin products.
- Profit is More Stable Than Sales. While sales fluctuated significantly, profit remained relatively steady, this might indicate some strategy followed by shop owner to maintain consistent profit.





- *Ice-creams* are heavily restocked every 2-3 days (~500) units but the quantity sold remains steady between 60-130 units. There is a possibility of overstocking since ice-creams have 75+ unique items.
- Restocking of *cold drinks* seems more planned. It's sales are steadier than that of ice-creams despite few restocking days.

END OF REPORT