Optimizing Inventory and Maximizing the Profit of a Café Business

An End-Term report for the BDM capstone project

Submitted by

Name: Sourasish Ghosh

Roll No: 23f3004075

Mail: 23f3004075@ds.study.iitm.ac.in

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents Page

1. Executive Summary

‘Ektu Baithak’ is a Café/Restaurant, which has been very popular in our local area, located at Tamlipara Bus Stand, Jirat, Balagarh, near STKK road. Despite the popularity, the business faces some problems that hinders its optimal output. This report takes some analytical, visualization and few forecasting approaches to tackle the problems with proper recommendation and interpretation. The main problems that’s been taken for analysis are: Profit Optimization and Inventory management and wastage spoilage. This report also analyses several other problems also.

To tackle the challenges, we took previous year’s Sales and SKU data in addition with Day-Wise-Sales dataset for May and June and Inventory dataset for May month. The analysis includes Trend and Demand Analysis, Day-Wise-Sales & Inventory Usage Analysis, ABC Analysis, Break-Even Analysis, Regression Analysis. In Trend and Demand Analysis, we have found employee cost is the most (₹135000); Raw Goods Cost vary throughout the months (varies from ₹45000 - ₹85000) with slight cost effect in Electricity Bill (varies from ₹4000 - ₹8000). We have also found that most sales occur during October and from December-to-January and least sales from April-to-August. Analysing the inventory and daily sales, we have found out that most sales in May occurs in weekends and surprisingly during Wednesdays and analysing the wastage flagging, we have found 2 days (05/05/25 and 08/05/25) with inventory item wastages in May. During ABC analysis, we have categorised the SKUs in three segments, where Biriyani and CST Rice are the most valued SKUs (Category A), Beverages, Noodles and Kebab are found to be in Category B, Drums of Heaven (Chicken Drumstick) and All Other SKU items were found to be in Category C. After that, in Break-Even analysis, Break-Even Point gave the minimum number of units to be sold per month (which is around ~3106 units/month) to recover the costs. Although we have found the business currently sells ~2022 units monthly, which is ~1084 units short of the break even. Finally in Regression Analysis, we incorporated Linear Regression which gave a R^2 value of 0.967 or 96.7%, adjusted R^2 value of 0.948, F-statistic of 51.51 and P-value of 0.00000282, indicating a good modelling and that it rejected null hypothesis. We have also analysed some hypothetical ‘What If’ questions that gave a prediction measure that how the business will perform in certain circumstances.

Based on the results and analysis, we recommend to input daily audits to track stocking-restocking, implementing First-In-First-Out in inventory and kitchen storage to manage old stocks. Also, partnering with supplier and SKU prioritization will help to manage inventory and wastage spoilage. Implementing Point-of-Sale to track daily transactions, hiring temporary workers to balance the load of work, doing combo offers and discounts on menu, revising the menu prices according to contribution margin, introducing home delivery services will help to increase the profit margins.

These measures will definitely help the Café to tackle their inventory management issues, wastage spoilage and optimizing their profit margin.

2. Detailed Explanation of Analysis Method

2.1. Data Cleaning and Preprocessing

The data cleaning and preprocessing has been done on multiple datasets. The datasets are:

- The Sales dataset of the Café in year 2024
- The SKU (Stock Keeping Unit) dataset in the same year
- Day-Wise-Sales dataset for May and June in 2025 Link
- Inventory dataset for May 2025 Link

In the Sales dataset, owner gave a rough estimate of the yearly profit (18-22%). I had to work on Google Sheets to calculate monthly profit margin. It helps to track financial performance trends for month-to-month.

In SKU dataset, had to standardise and organise all the menu items. To standardise prices, a mean price per unit (₹) was calculated for each SKU category, such as Biriyani, Noodles, and Hot Beverages, by averaging all their respective variations. It makes grouping and aggregation of sales easier and also gave insights about ingredients demand for SKUs.

The Day-Wise-Sales dataset were put in a standardised form in DD-MM-YYYY format throughout the 2 months (May-June). Along with 'Avg. Sale per Customer' was calculated to check the consistency. This gives customer spending behaviour for days of the week. It makes sure to organise inventory early.

In the Inventory dataset, the inventory items were normalized to facilitate further calculations in the inventory dataset. Some of them are given here:

Items	**Unit**
Rice, Chicken, Mutton, Coffee, Potato, Sugar, Onion , Yogurt , Flour , Butter	Kilogram (Kg)
Milk , Ghee/Oil ,	Litre (L)
Bread	Loaves
Noodles	Packets
Eggs	Units

The standardisation was essential to compute each inventory item consumption throughout the month of May. It is necessary to compare stock-restock usage, prevents mixed unit calculation.

2.2. Comprehensive Explanation of Analysis Method

2.2.1. Trend and Demand Analysis:

We are using trend analysis to understand the behaviour in sales data throughout the year. It is crucial for inventory and waste management of 'Ektu Baithak' Café. The sales trend gets affected during Festival seasons (e.g. Durga Puja, Diwali, Christmas, New Year) and also during Monsoon seasons (May-August). By analysing the trends, we can prepare the inventory items accordingly. We can prevent stockouts in peak seasons and overstocking in Monsoon seasons. With changes of trends, we can identify which types of SKUs are in demand in particular months. It provides an insight about repeating patterns that will help to strategies like combo offers, discounts to improve sales. The analysis methods are given below:

Method 1: Cost Trend Analysis

The analysis method was to observe the month-wise movement of individual cost components using a stacked bar format.

- The Bars in the plot were color-coded to represent different components of the monthly cost structure:
  - – Red for Employee Costs (fixed cost)
  - – Blue for Electricity Bills (variable cost)
  - – Green for Average Raw Goods Cost (variable cost)

By comparing the size and composition of bars across months, we could identify stable and fluctuating cost patterns. The analysis is done using Python using Pandas and Matplotlib.

Method 2: Cost Efficiency Analysis

I calculated the Cost-to-Sales ratio during the months to get higher spendings months and visualised in a line chart. The code-snippets are given below.

```
df['Operating_Cost'] = df['Employee_Cost'] + df['Electricity_Bill'] + df['Raw_Goods_Cost']
df['Gross_Profit'] = df['Estimated_Sales'] - df['Operating_Cost']
df['Computed_Profit_Margin'] = (df['Gross_Profit'] / df['Estimated_Sales']) * 100
df['Cost_to_Sales_Ratio'] = (df['Operating_Cost'] / df['Estimated_Sales']) * 100
```

```
plt.figure(figsize=(12, 6))
plt.plot(df['Month'], df['Estimated_Sales'], label='Estimated Sales (₹)', marker='o')
plt.plot(df['Month'], df['Operating_Cost'], label='Operation Cost (₹)', marker='x', color='red')
plt.title('Monthly Estimated Sales vs Operation Cost', fontsize=20)
plt.xlabel('Months of the Year')
plt.ylabel('INR (₹)')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

print("\n Months with High Cost-to-Sales Ratio:")
print('\n')
print(df[df['Cost_to_Sales_Ratio'] > 90][['Month', 'Estimated_Sales', 'Operating_Cost', 'Profit_Margin']])
```

The above analysis and visualisation have been done on Python using Pandas and Matplotlib library.

2.2.2. Day-Wise-Sales & Inventory Usage Analysis

In order to analyse the inventory and waste management, it is crucial to analyse this section as it provides a direct measure of consumption for each ingredient. By tracking the **'Start'** and **'End'** quantities of key raw materials, we can precisely determine the actual usage for that operational period. Understanding the daily consumption pattern prevents both overstocking and stockouts. So, it's important to analyse this section. Following analysis methods are given below:

Method 1: Inventory Efficiency Ratio Analysis

Analysing Daily Inventory Efficiency Ratio is crucial as it tells percentage of revenue consumed by inventory expenses. To calculate the Ratio, I used Python as the primary tool and to visualize and interpret the data libraries like Pandas and Matplotlib was essential.

- The formula that's been used to calculate the Daily Inventory Efficiency is given here:

$$\text{Efficiency Ratio} = \frac{\text{Daily Sales ₹}}{\text{Total Inventory Used (units)}}$$

- The code snippets for calculation and visualisation of Efficiency Ratio are given below:

```
usage_columns = [col for col in inventory.columns if '(Kg) Start' in col or '(Packets) Start' in col or '(Lo
sku_names = [col.replace(' (Kg) Start', '').replace(' (Packets) Start', '').replace(' (Loaves) Start', '').r

# Computing daily usage per item
for sku in sku_names:
    start_col = f'{sku} (Kg) Start' if f'{sku} (Kg) Start' in inventory.columns else f'{sku} (L) Start' if f
    end_col = start_col.replace('Start', 'End')
    inventory[f'{sku}_Used'] = inventory[start_col] - inventory[end_col]

# Total inventory used (sum of all individual SKU usage)
usage_cols = [col for col in inventory.columns if '_Used' in col]
inventory['Total_Inventory_Used'] = inventory[usage_cols].sum(axis=1)

# Date mergeing
df = pd.merge(inventory[['Date', 'Total_Inventory_Used']], sales[['Date', 'Daily Sales INR']], on='Date')

#  Efficiency Ratio
df['Efficiency_Ratio'] = df['Daily Sales INR'] / df['Total_Inventory_Used']
print(f" Daily Inventory Efficiency Ratios for month of May are given below: \n {df['Efficiency_Ratio']}")
```

```
# Plot Efficiency Ratio
plt.figure(figsize=(12, 6))
plt.plot(df['Date'], df['Efficiency_Ratio'], marker='o', label='Inventory Efficiency Ratio')
plt.axhline(df['Efficiency_Ratio'].mean(), color='red', linestyle='--', label='Mean Efficiency')
plt.title('Daily Inventory Efficiency Ratio (Sales / Inventory Used)',fontsize=20)
plt.xlabel('Date')
plt.ylabel('Efficiency Ratio')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

Method 2: Wastage Flagging Analysis

This analysis aims to find out where is the consumption of inventory relative to customer traffic and sales performance to find potential wastage. The 75th percentile of **Inventory per Customer** is chosen (the formula for this: Inventory per Customer = $\frac{\text{Total Inventory Items Used (Daily)}}{No. of\ Daily\ Customers}$). If it exceeds the threshold, it will be considered as high ingredient consumption. And, 25th percentile of **Sales per Customer** is chosen (the formula for this: Sales per Customer = $\frac{Daily\ Sales\ (INR)}{No. of\ Daily\ Customers}$). If it falls below, it will be considered as unusually low revenue generation per customer.

So, the days of the month will be flagged as "Wasteful Days" if both the conditions will meet simultaneously.

$$\text{Waste Flag} = \text{High Inventory Usage AND Low Sales Performance}$$

- The analysis is done on Python using Pandas to manipulate the data.
- The code snippet of this analysis is given below

```
# Step 4: Calculating per-customer metrics
merged['Inventory_per_Customer'] = merged['Total_Inventory_Used'] / merged['Customers']
merged['Sales_per_Customer'] = merged['Daily Sales INR'] / merged['Customers']

# Step 5: Calculating percentiles and flagging wasteful days
inv_75 = merged['Inventory_per_Customer'].quantile(0.75)
sales_25 = merged['Sales_per_Customer'].quantile(0.25)

merged['Wastage_Flag'] = (merged['Inventory_per_Customer'] > inv_75) & \
                         (merged['Sales_per_Customer'] < sales_25)

# Step 6: Display flagged wasteful days
wasteful_days = merged[merged['Wastage_Flag'] == True]
print(" Wasteful Days Identified:")
print(wasteful_days[['Date', 'Daily Sales INR', 'Customers', 'Total_Inventory_Used',
                     'Inventory_per_Customer', 'Sales_per_Customer']])
```

2.2.3. ABC Analysis

ABC analysis is an inventory categorization technique that divides items into 3 categories (A,B and C) based on their relative importance, often measured by contribution to the overall value (e.g.: sales volume, cost, profit etc.). In our analysis, we used SKU-wise profit contribution. So, we applied profitability and the categories indicates as follows:

- Category A is high valued profit items, the most critical SKUs that contribute largest proportion of the total profit.
- Category B is Medium valued profit items. These SKUs contribute a moderate portion of the total profit.
- Category C is Low valued profit items. They contribute to a small percentage of the total profit.

The goal of this analysis is to identify key profit drivers of the business. It will also help managing the operational resources towards the most profitable items. The mathematical formulas that's been used for this analysis is given below:

- Profit per Unit for each SKU: $PPU_i = SP_i - RC_i$
- Monthly Profit for each SKU: $MP_{ij} = S_{ij} \text{ x } PPU_i$
- Total Annual Profit for each SKU: $TAP_i = \sum_{j=1}^{Nm} MP_{ij}$
- Total Overall Profit: $TOP_i = \sum_{i=1}^{Ns} TAP_i$
- Percentage Contribution of each SKU: $PCT\_Contribution_i = \frac{TAP}{TOP} \, x \; 100 \;\%$
- Cumulative Percentage Contribution; as we have to sort the items by their total annual/overall profits in descending order to identify the most important items. So, the mathematical formula will be; $Cum\_PCT_k = \sum_{i=1}^{k} PCT\_Contribution_i$

Where, SP_i is Selling Price per unit SKU i, RC_i is Raw Cost per unit of SKU i, Nm is Total number of months, Ns is Total number of SKUs, S_{ij} Sales units of SKU i in month j .

Here is a code snippet of the analysis which has been done using Python with the help of libraries like; Pandas, Matplotlib, Seaborn.

```
sku_sales_long = sku_sales_df.melt(id_vars='Month', var_name='SKU', value_name='Units Sold')
sku_sales_merged = pd.merge(sku_sales_long, sku_pricing, on='SKU')


sku_sales_merged['Profit per Unit'] = sku_sales_merged['Selling Price'] - sku_sales_merged['Raw Cost']
sku_sales_merged['Monthly Profit'] = sku_sales_merged['Units Sold'] * sku_sales_merged['Profit per Unit']


sku_profit_summary = sku_sales_merged.groupby('SKU')['Monthly Profit'].sum().reset_index()
sku_profit_summary = sku_profit_summary.sort_values('Monthly Profit', ascending=False).reset_index(drop=True)


sku_profit_summary['% Contribution'] = (sku_profit_summary['Monthly Profit'] / sku_profit_summary['Monthly Profit'].sum()) * 100
sku_profit_summary['Cumulative %'] = sku_profit_summary['% Contribution'].cumsum()

def classify_abc(pct):
    if pct <= 70:
        return 'A'
    elif pct <= 90:
        return 'B'
    else:
        return 'C'

sku_profit_summary['ABC Category'] = sku_profit_summary['Cumulative %'].apply(classify_abc)
sku_profit_summary
```

2.2.4. Break-Even Analysis

The Break-Even Point (BEP) is the level of sales (in units or revenue) at which total revenues equal total costs, resulting in zero net income. In other words, it's the point where a business neither makes a profit nor incurs a loss. It helps businesses understand the minimum sales volume required to stay afloat. It highlights the risk associated with different cost structures. Businesses with high fixed costs will have a higher BEP, meaning they need to sell more to break even. The mathematical formula for BEP is given below:

Break-Even Point = $\frac{Total\ Fixed\ Costs}{Weighted\ Average\ Contribution\ Margin}$. Now, Weighted Average Contribution Margin represents the average contribution generated per unit across all SKUs, weighted by the proportion of each product sold. The formula is as following: WACM = $\frac{\sum(CM_i\ x\ Sales\ Volume)}{\sum Sales\ Volume}$. CM refers to Contribution Margin. The CM per unit tells you how much money from each sale is left over to cover fixed costs (like rent, salaries, utilities) after paying for the variable costs (ingredients, packaging). The formula for CM is as following:

CM / Unit = Selling Price / Unit – Variable Cost / Unit .

The calculations are done on Google sheets, which will bw given below:

SKUs	Monthly units sold	% of Total sales	Price per Unit (INR)	Estimated Variable Costs (INR)	Contribution Margin (INR)
Biriyani (per plate)	630	31.15727003%	150.00	70.00	80.00
Chef Special Tripple Rice (per plate)	228	11.27596439%	180.00	60.00	120.00
Drums of Heaven (Chicken Drumstick) (6 pieces)	207	10.23738872%	160.00	100.00	60.00
Ektu Baithak Special Nawabi Kebab(Kebab) (6 pieces)	148	7.319485658%	200.00	100.00	100.00
Noodles (per plate)	384	18.99109792%	70.00	45.00	25.00
Beverages (Cappuccino/Coffee) (per cup)	259	12.8090999%	60.00	30.00	30.00
Others	166	8.209693373%	120.00	70.00	50.00
SUM	2022	100%	940.00	475.00	465.00
Average Fixed Cost	200710				
WACM (Weighted Average Contribution Mragin)	64.61				
BEP(Units)	3106				

Table 1: Break-Even Analysis Calculation Table

Further we visualised BEP and SKU-wise Contribution Margin using Python and its libraries Pandas, Matplotlib. Here is a code snippet of it:

```
# Break-Even Line Chart
units = list(range(0, 6000, 100))
total_costs = [fixed_costs] * len(units)
revenue = [u * wacm for u in units]

# Contribution Margin Bar Chart
fig, axes = plt.subplots(1, 2, figsize=(15, 6))

# Plot 1: Break-Even Point Line Chart
axes[0].plot(units, revenue, label='Revenue', color='green')
axes[0].plot(units, total_costs, label='Fixed Costs', color='red')
axes[0].axvline(bep_units, color='blue', linestyle='--', label=f'BEP: {int(bep_units)} Units')
axes[0].set_title("Break-Even Analysis")
axes[0].set_xlabel("Units Sold")
axes[0].set_ylabel("INR")
axes[0].legend()
axes[0].grid(True)
# Plot 2: Contribution Margin vs WACM
bars = axes[1].bar(df_cm["SKU"], df_cm["Contribution_Margin"], color='skyblue')
axes[1].axhline(wacm, color='orange', linestyle='--', label=f'WACM: ₹{wacm:.2f}')
axes[1].set_title("SKU-wise Contribution Margin")
axes[1].set_ylabel("Contribution Margin (INR)")
axes[1].set_xticklabels(df_cm["SKU"], rotation=45, ha='right')
axes[1].legend()

# Bar highlights
for bar, cm in zip(bars, df_cm["Contribution_Margin"]):
    if cm > wacm:
        bar.set_color('green')
    elif cm < wacm:
        bar.set_color('salmon')
```

2.2.5. Regression Analysis

Regression analysis is a statistical method used to estimate the relationships between a dependent variable and one or more independent variables. Linear regression specifically models the relationship between the dependent variable and the independent variables as a straight line. The goal is to find the "best-fitting" line that minimizes the sum of squared differences between the observed values of the dependent variable and the values predicted by the model. R^2, also known as the coefficient of determination, is a statistical measure that represents the proportion of the variance in the dependent variable that can be predicted from the independent variables. A higher R^2 generally indicates a better fit of the model to the data, implying that the independent variables are good predictors of the dependent variable. The general form of a multiple linear regression model with k independent variables is:

$$\mathrm{Y} = \beta 0 + \beta 1X1 + \beta 2X2 + \cdots + \beta kXk + \epsilon$$

Where:

- Y: The dependent variable (Monthly Profit Margin)
- X1,X2,…,Xk: The independent variables (here it is Estimated Sales, Employee Costs, Electricity Bill, Raw Goods Cost, Average Daily Customers)
- β0: The Y-intercept (the predicted value of Y when all Xi are zero)
- β1,β2,…,βk: The regression coefficients for each independent variable.
- ϵ: The error term (known as Residuals), representing the unexplained variance or random noise.

The predicted value of Y for a given set of X values is denoted by $\hat{Y}$.

$$\hat{Y} = \hat{\beta}0 + \hat{\beta}1X1 + \hat{\beta}2X2 + \cdots + \hat{\beta}kXk$$

The residual for each observation i is, e_i=Yi−$\hat{Y}$i.

The R^2 is calculated as: $R^2 = 1 - \frac{SSR_{error}}{SST}$; where, $SSR_{error} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ and $SST_i = \sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$

After analysing linear regression and determining R^2, we incorporate our analysis into 2 following methods:

Method 1: Variable Inflation Factor (VIF)

Variance Inflation Factor (VIF) measures how much the variance of a regression coefficient is inflated due to multicollinearity among the independent variables. Say if two predictors are highly correlated the model can't distinguish between them even though R^2 is high. The formula for this is given below: For a variable X_i VIF is calculated as, VIF(X_i) = $\frac{1}{1 - R_i^2}$; R_i^2 is R^2 from regression

Method 2: What-If Analysis

What-If analysis tests how much changing input variables affects the outcome. It answers practical questions like: " What if we reduce raw goods price by 10%" or, "What if we increase daily customers by 20%" etc. This type of scenarios provides actionable decisions rather than just a model. The scenarios are chosen based on VIF scores, P-values, Regression Coefficients through the analysis process. The analysis has been done on Python using libraries like Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Statsmodels. The scenarios chosen for this study is given below as a code snippet:

```
scenarios = pd.DataFrame({
    'Scenario': [
        '+15% Customers',           # High footfall
        '-10% Raw Cost',            # Cost saving
        '+10% Sales',               # Marketing push
        '-20% Electricity Bill',    # Energy efficiency
        '-10% Employee Costs'       # Cost-cutting
    ],
```

And the scenarios are chosen based on the following parameter values:

Variable	Regression Coefficient	P-Value	VIF	Interpretation
Estimated Sales	0.0003	0.024	29.30	Significant, Positive Driver of sales
Employee Costs	-0.000005	0.565	74.20	Multicollinearity, not so reliable
Electricity Bill	-0.0037	0.050	1.54	Significant, Negative Effect
Raw Goods Cost	-0.0004	0.374	8.29	Moderate, Negative Effect
Avg. Daily Customers	-0.0182	0.930	20.40	Insignificant, Negative Effect

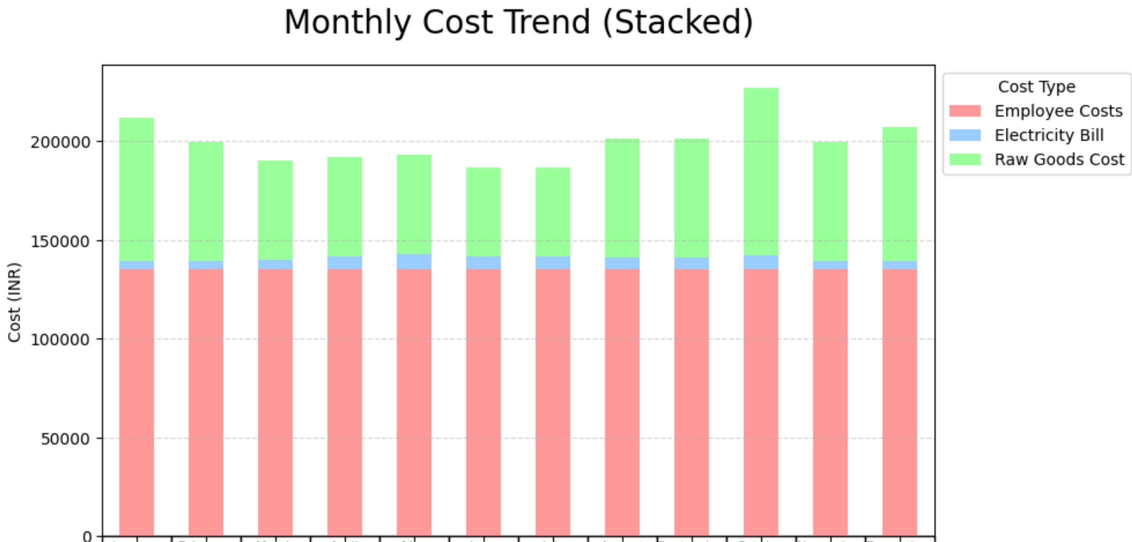
Table 2: Variables with its corresponding parameters and interpretation

Based on those results, we constructed the scenarios.

3. Results and Findings

This section gives a comprehensive analysis of the key business problem of the Café. It is based on two major challenges: Optimizing the inventory management and maximizing the profit of the business. Based on that the results are presented with proper visual representations and calculation to get proper understanding.

The **Cost Trend Analysis** bar chart gives a key insight about the monthly costs in the Café for each month. We can see the Employee costs (in red) are constant for all the months (₹135000). So, the salaries are fixed and it takes most amount of costs. Electricity bills (in blue) take a small component of overall costs. It is slightly higher in between April-to-July (₹6800-8000-6500) and also have higher value in festive month, October (₹7000). Raw Goods cost shows variations across months, indicating sales volume change. Highest costs are at October (~₹85000) and lowest at July (~₹45000). Majority of costs goes towards Employee costs with cost cut in Raw Goods Cost and slight cost cut in Electricity Bill.



	January	February	March	April	May	June	July	August	September	October	November	December
Employee Costs	135000	135000	135000	135000	135000	135000	135000	135000	135000	135000	135000	135000
Electricity Bill	4500	4500	5000	6800	8000	6800	6500	6000	6000	7000	4500	4000
Raw Goods Cost	72000	60000	50000	50000	50000	45000	45000	60000	60000	85000	60000	68000

Fig 1:Stacked Bar Chart of Cost Trends (with dataset table) throughout the year

The Monthly Cost Trend chart shows clear trends in the café's spending. Analysing the plot suggests that current ingredient purchasing may not be perfectly aligned with customer demand. In short, while fixed staff costs limit financial flexibility, the changing cost of raw goods presents a clear chance for savings.

Now by looking at the analysis of **Cost-to-Sales Ratio,** we can identify those months with more operational costs than monthly sales. As per the analysis, from April-to-August the operational costs are greater than estimated monthly sales. The highest difference being in July where sales are around ₹160000 and costs are around ₹186500. The whole trend throughout the year can be seen in the following line plot.

The cause of increase in Cost-to-Sales ratio is mainly due to constant Employee salary bills and Raw Goods

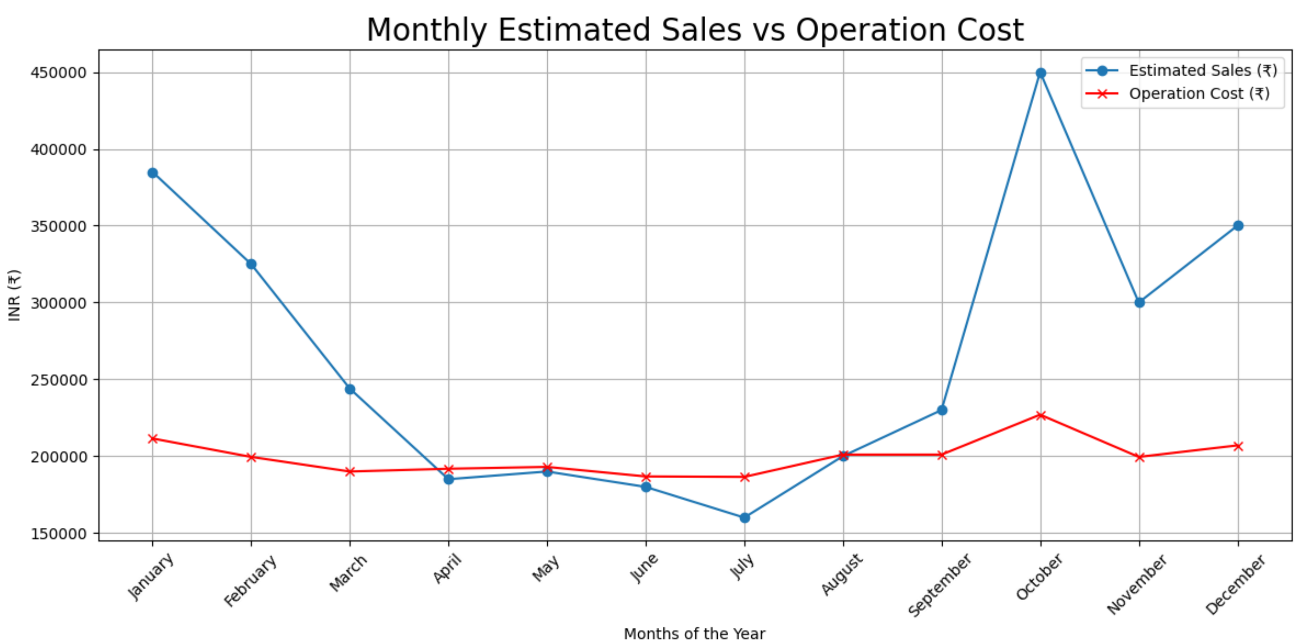


Fig 2: Monthly Estimated Sales vs Operational Costs Line Chart

Cost. The salary of the employees is same for every month, even when the customer frequency is low. And for Raw Goods, even if the demand is low from April-to-August the Raw Goods prices does not fluctuate too much. This constant high cost, regardless of demand or customer numbers, is what's driving the profit loss of the business.

The line plot comparing Monthly Estimated Sales and Operation Costs reveals significant fluctuations in revenue while operational costs remain relatively stable throughout the year. The mismatch highlights the need for demand-driven cost control or promotional strategies during low-sales periods to ensure profitability.

The mismanagement of raw goods causes inventory issues, so the analysis of **Inventory Efficiency** is crucial in this case. The efficiency ratio is being calculated using daily sales and inventory data of month of May. The average daily inventory efficiency ratio for May was **136.90**. This indicates that on average, for every unit of inventory used (in Kg, Litres, Packets, Loaves, or units), approximately ₹136.90 in sales were generated. The minimum and maximum value of Efficiency ratio are **102.34** and **180.11** respectively. So, the daily sales are varied significantly. The Standard Deviation of Efficiency Ratio is around **19.89** suggesting a moderate spread around the mean. The observed variations in Efficiency Ratio highlights the dynamic relationship between daily sales and inventory consumption. Days with higher sales relative to inventory used would result in a higher efficiency ratio, while days with lower sales or higher inventory usage would result in a lower ratio. This can be clearly seen in the following figure.

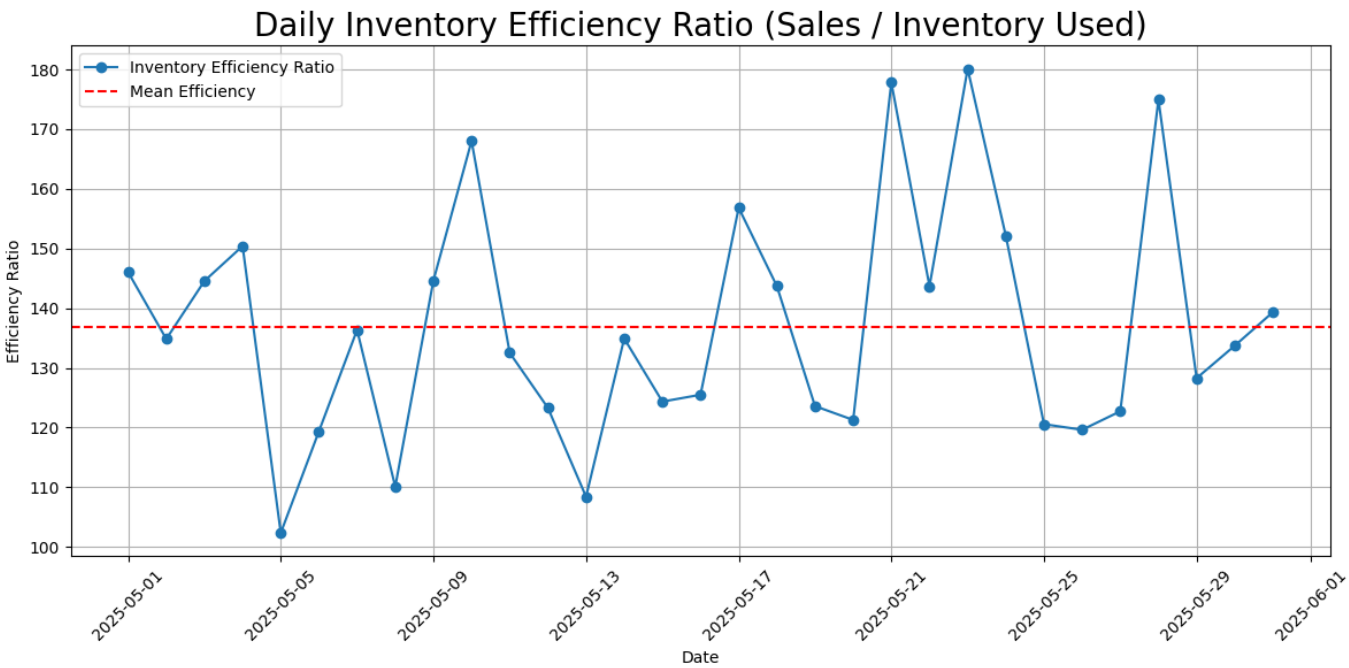


Fig 3: Inventory Efficiency Ratio Daily Trend Plot for May

The plot illustrates that while the average inventory efficiency for May was around the mean line, there were significant daily fluctuations. The peaks are mainly in Saturdays and Sundays. In general, the Weekdays suffer in sales as demand decreases. Although the maximum sales occur during Friday notably. Also, sales during Wednesday have a peak which is very much noticeable. The Efficiency Ratios for the month of May can be seen in the following table.

These fluctuations are mainly caused due to customer traffic difference in Weekdays vs. Weekends, weather conditions. Retail businesses often see higher customer traffic and sales on weekends compared to weekdays. And, weather conditions in May due to extreme heats and thunderstorms in evening period are probably the main causes of loss in sales.

Date	Efficiency_Ratio	Day of the Week
01-05-2025	146.039604	Thursday
02-05-2025	134.844869	Friday
03-05-2025	144.565217	Saturday
04-05-2025	150.414938	Sunday
05-05-2025	102.345416	Monday
06-05-2025	119.375	Tuesday
07-05-2025	136.363636	Wednesday
08-05-2025	110.096154	Thursday
09-05-2025	144.628099	Friday
10-05-2025	168.017058	Saturday
11-05-2025	132.519531	Sunday
12-05-2025	123.304158	Monday
13-05-2025	108.411215	Tuesday
14-05-2025	134.920635	Wednesday
15-05-2025	124.349882	Thursday
16-05-2025	125.501114	Friday
17-05-2025	156.822811	Saturday
18-05-2025	143.712575	Sunday
19-05-2025	123.595506	Monday
20-05-2025	121.281465	Tuesday
21-05-2025	177.809798	Wednesday
22-05-2025	143.686869	Thursday
23-05-2025	180.107527	Friday
24-05-2025	152.129817	Saturday
25-05-2025	120.567376	Sunday
26-05-2025	119.628099	Monday
27-05-2025	122.727273	Tuesday
28-05-2025	174.925373	Wednesday
29-05-2025	128.275862	Thursday
30-05-2025	133.726415	Friday
31-05-2025	139.344262	Saturday

Table 3: Efficiency Ratio Table for May

The inventory inefficiency causes a lot of inventory wastage. To identify those wasteful days, we need to flag them, known as **Wastage Flagging**. Days were flagged wasteful if inventory used per customer is above 75th percentile and sales per customer is below 25th percentile (High Consumption – Low Revenue). By analysing May month, two dates were found with unusually high inventory but poor in sales:

- **05-05-2025**: 49.6 units of inventory for 58 customers; sales per customer ₹82.76
- **08-05-2025**: 44.9 units for 51 customers; sales per customer ₹89.80

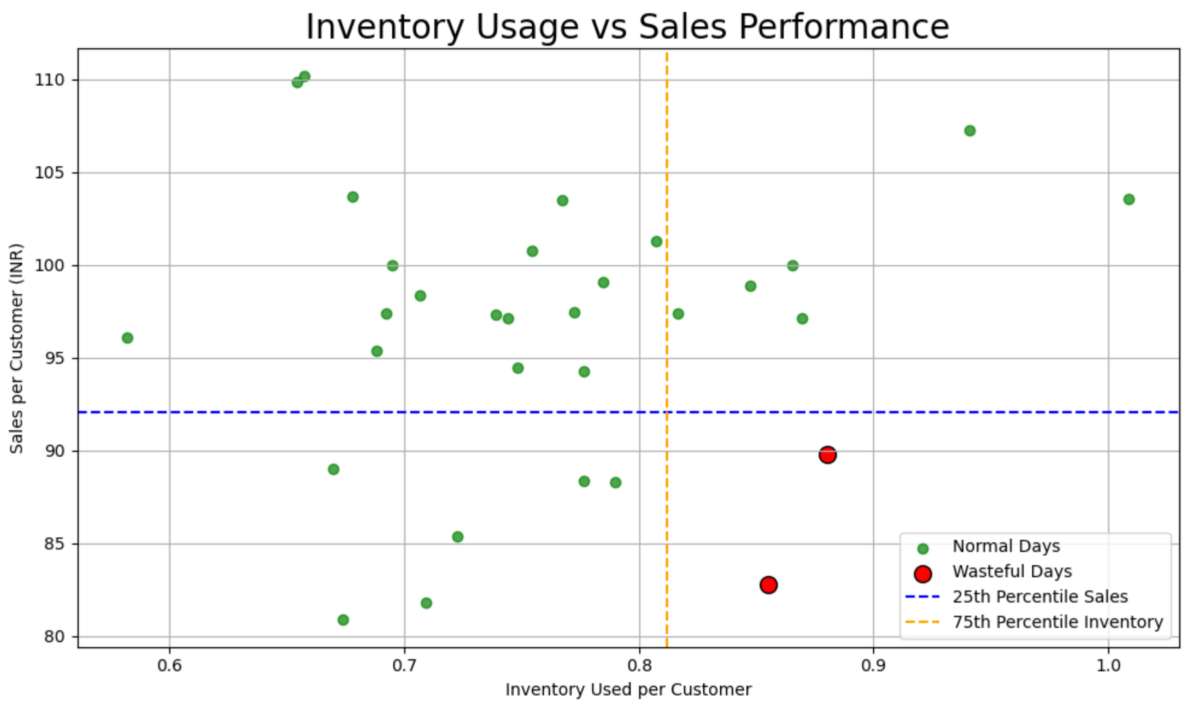


Fig 4: Wastage Flagging Scatter Plot: Inventory Usage vs. Sales per Customer

These days indicate high possibility of over-production, item spoilage. 05/05/2025 and 08/05/2025 are both weekdays, Monday and Thursday respectively. The kitchen might have prepared more food than usual. Bad weather could be a possible reason for low customer frequency. The wastage days are visualised in the above scatter plot.

Now to allocate more storage, prioritizing SKU management, enhancing inventory cost control **ABC analysis** is vital for it. Its primary purpose is to help businesses focus their resources effectively. For this business, the analysis method was to classify SKUs into categories **A**, **B**, and **C** based on their cumulative contribution to total annual profit. Top 70% profit contributors are in Category **A**, next 20% contributors are in Category **B** and the rest 10% are in Category **C**. The analysis identified top-performing SKUs based on their annual contribution to overall profits:

- **Category A** SKUs — **Biriyani** and **CST Rice** — contributed approximately 63% of total profits, marking them as primary revenue generators. **Biriyani** solely contributed around 41% of the total sales profit. **CST Rice** contributes around 22% of the total profit. SKUs in this section should be given high attention in terms of stocking levels, supply chain management and offers/promotions due to its significant impact on sales.

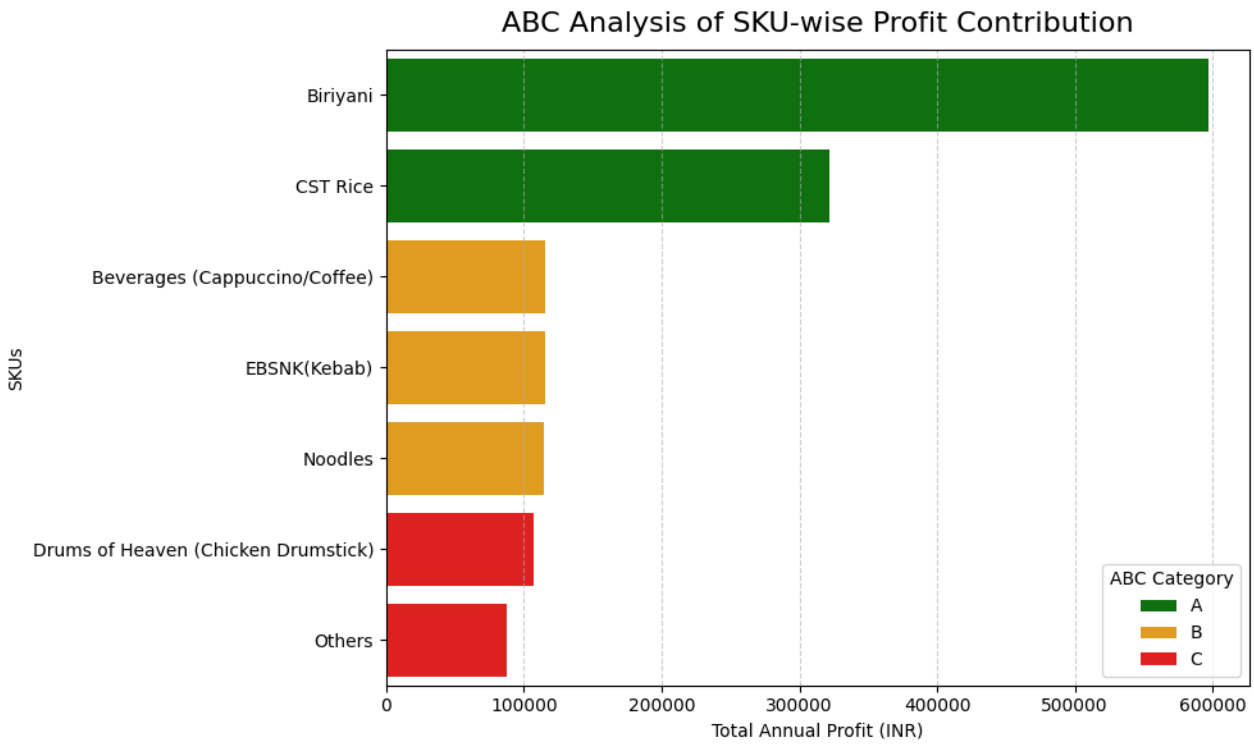


Fig 5: ABC Analysis of SKU-wise Profit Contribution using Bar-Chart

- **Category B** SKUs — including **Beverages, Kebab**, and **Noodles** — collectively accounted for about 24%, indicating moderate profitability with potential for upselling or bundling strategies. Three of the SKUs similarly contribute around (7.8-7.9%) of the profit. These SKUs are not as much important from Category A, but regular demand pattern can ensure limitation of wastage.

- **Category C** SKUs — **Drums of Heaven** and **Others** menus items— contributed the least (13%) and may require menu optimization, cost reduction, or strategic repositioning. Using promotions and discounts might help move the excess stock. The whole ABC analysis is visualised in the above bar chart.

Now the **Break-Even Analysis** was conducted to evaluate the financial sustainability of the SKUs by comparing fixed costs, contribution margins, and sales performance. The calculated table has been given earlier in *Table 1*. Doing the analysis of the table we can see that;

- The **Total Fixed Cost** for the business was estimated at **₹200,710**.
- The **Weighted Average Contribution Margin (WACM)** across all SKUs was calculated as **₹64.61**.
- And after calculation, the **BEP** (**Break-Even Point**) is around **3106 units**. That means the business needs to sell approximately 3106 units per month to cover all fixed and variable costs and start generating profit.
- The **Contribution Margin** (= Selling Price – Variable Price) varied significantly in all across the SKUs:
    - Highest CM : CST Rice -- ₹120
    - Lowes CM : Noodles -- ₹25
- SKUs like **CST Rice (₹120)**, **EBSNK Kebab (₹100)**, and **Biriyani (₹80)** performed **above the WACM threshold** and they contribute more toward profitability. In contrast, **Noodles (₹25)** and **Beverages (₹35)**, despite having higher unit sales, **fall below the WACM**, indicating lower profitability per unit.

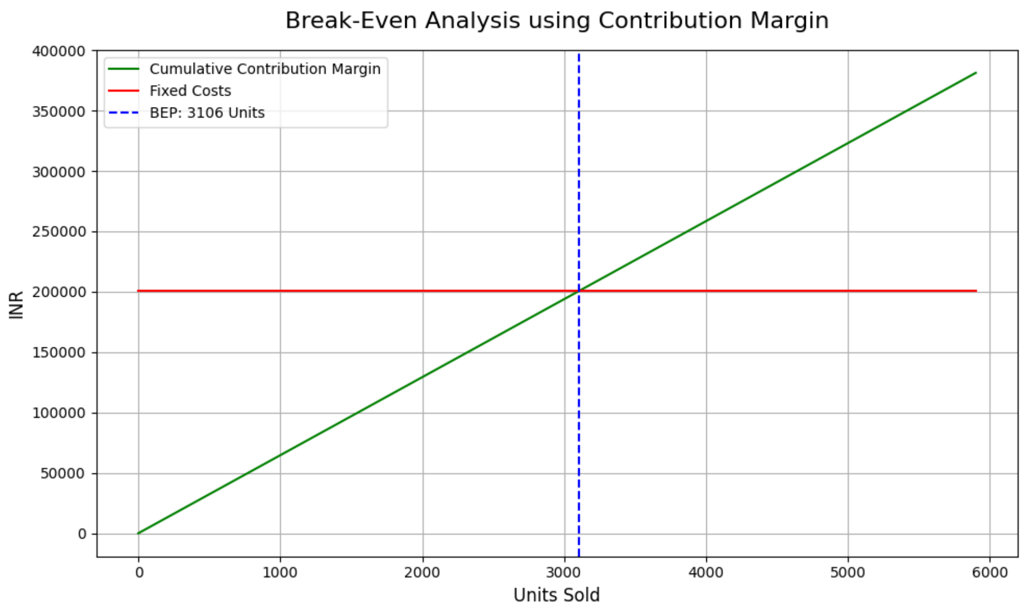


Fig 6: Break-Even Point Analysis: Contribution Margin vs. Fixed Costs

Above all the SKUs, **CST Rice** has the highest CM but the sale is around **11%**. That means they are being **under-leveraged**. The business is not utilizing its full potential. Noodles contribute around **19%** of the business but the margins are too low. Thorough check up on **cost reduction** or **price revision** needs to be considered. The business currently sells ~**2022** units monthly, which is ~**1084** units short of the break even. It indicates loss making situation in the business. The Break-Even Point can be seen in the above figure.

In the plot, the Green Line indicates Cumulative Contribution Margin, which is the total profit generated over fixed costs. The Red Line is the Fixed Cost (~₹200,710). The intersection of Green and Red lines is the BEP (~3106 units). So beyond selling more than that, the business starts making profit.

Next there is the visualisation of **SKU-wise Contribution Margin** bar plot that shows CM of each SKU relating to WACM. The dotted line indicates the WACM (~₹64.61). SKUs that cross the line are Biriyani, CST Rice and Kebab. SKUs below the line are Drums of Heaven (Chicken Drumstick), Noodles, Beverages, and Other SKU items. The plot can be seen in the following figure:

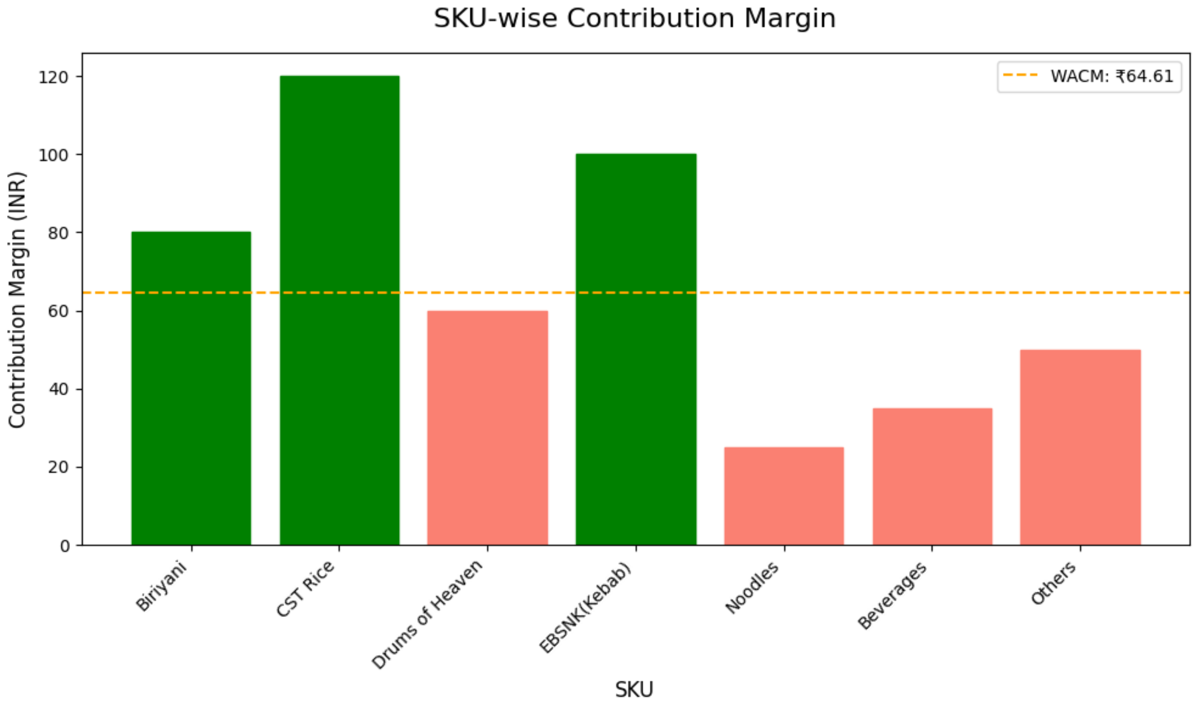


Fig 7: Contribution Margin by SKUs Using Bar Plot

Further a multiple **Linear Regression** was conducted to understand how various operational factors influence the monthly profit margin. So, Profit Margin (%) was taken as dependent variable. And all other 5 variables were taken as independent variable i.e., Estimated Sales, Employee Costs, Electricity Bill, Raw Goods Cost, Average Daily Customers. We have found $\mathbf{R^2}$ value to be **0.967** or in other words, the model explains **96.7%** of the variability in profit margin. This indicates an excellent fit, which means all the predictors are highly effective at explaining the changes in profit margin. The **Adjusted** $\mathbf{R^2}$ value is **0.948** which is good considering that it penalizes the model for adding predictors that do not significantly improve the model's explanatory power. **F-statistic** and **p-value** came around **51.51** and **0.00000282** respectively. It indicates that null hypothesis can be rejected. Further calculating the regression coefficients gave the inter relation between the 5 variables and Profit Margin. It is shown in the following table:

Variable	Coefficient	Interpretation
Estimated Sales	2.70 * 10^-04	For every ₹1 increase in sales, profit margin increases by ~0.00027%.
Employee Costs	1.25 * 10^-16	No measurable impact (As the value is close to zero).
Electricity Bill	-3.59 * 10^-03	For every ₹1 increase in electricity cost, profit margin decreases by ~0.0036%.
Raw Goods Cost	-4.01 * 10^-04	Every ₹1 increase in raw goods reduces profit margin by ~0.0004%.
Avg Daily Customers	-2.40 * 10^-02	Every additional customer reduces profit margin by ~0.024% on average.

Table 4: Variables and their Corresponding Regression Coefficient and Brief Interpretation

Now to see how well the model's predictions compared to the actual profit margin data, the regression plot is given below:

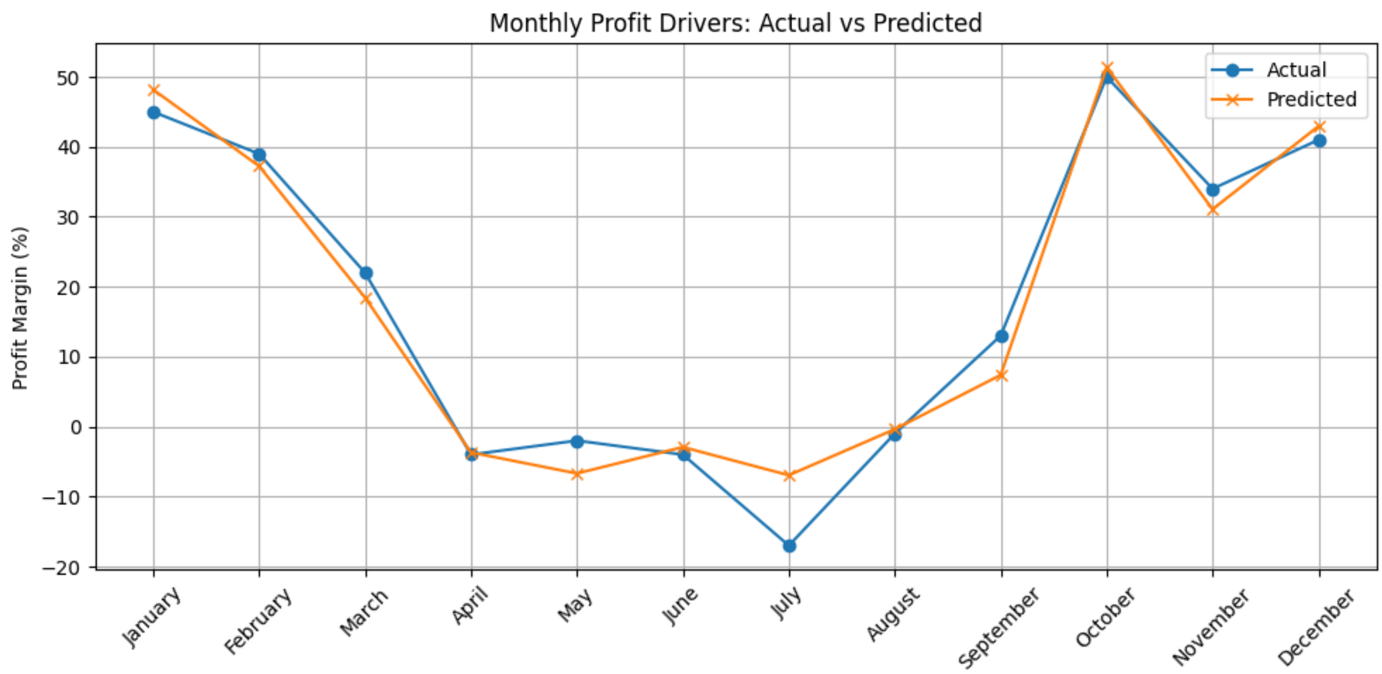


Fig 8: Monthly Profit Margin: Actual vs. Predicted

This **Regression Plot** visually validates the statistical results, showing that the model is an excellent fit for the data. The model performs exceptionally well in predicting the highest profit margin of the year in **October**. Both the actual and predicted values align almost perfectly. The noticeable deviation occurs during the low-profit period from **May to July**. In actual case, the monthly profit during July is around -17%, but the prediction is at somewhere around 8-9%. Overall, the model captures all the trends.

The model's performance looks good. After that first, we will use **VIF** values to check its multicollinearity. And after that we will explore some '**What If**' scenarios for more practical forecasting and decision making.

Multicollinearity makes it hard to isolate the individual effect of a variable on the dependent variable. This leads to: Unstable coefficient estimates, Inflated standard errors, Difficulty in model interpretation etc.

For this model the following VIF table with interpretation is given below:

Feature	VIF	Interpretation
Estimated Sales	29.296952	Extremely high. Strongly correlated with other variables, possibly Estimated Sales or Raw Goods Cost.
Employee Costs	74.204526	Very high. Likely because many other costs scale with sales.
Electricity Bill	1.544886	High value, likely to overlap with Estimated Sales.
Raw Goods Cost	8.297495	Moderate value, less concerning.
Average Daily Customers	20.403768	Moderate value, less concerning.

Table 5: VIF Table of the Model with Interpretation

It's useful to visualize how the variables relate to each other using a **correlation heatmap**. This helps identify strongly co-moving variables and decide which ones might be redundant, merged, or prioritized in the model.

The correlation heatmap of business metrics are given below:

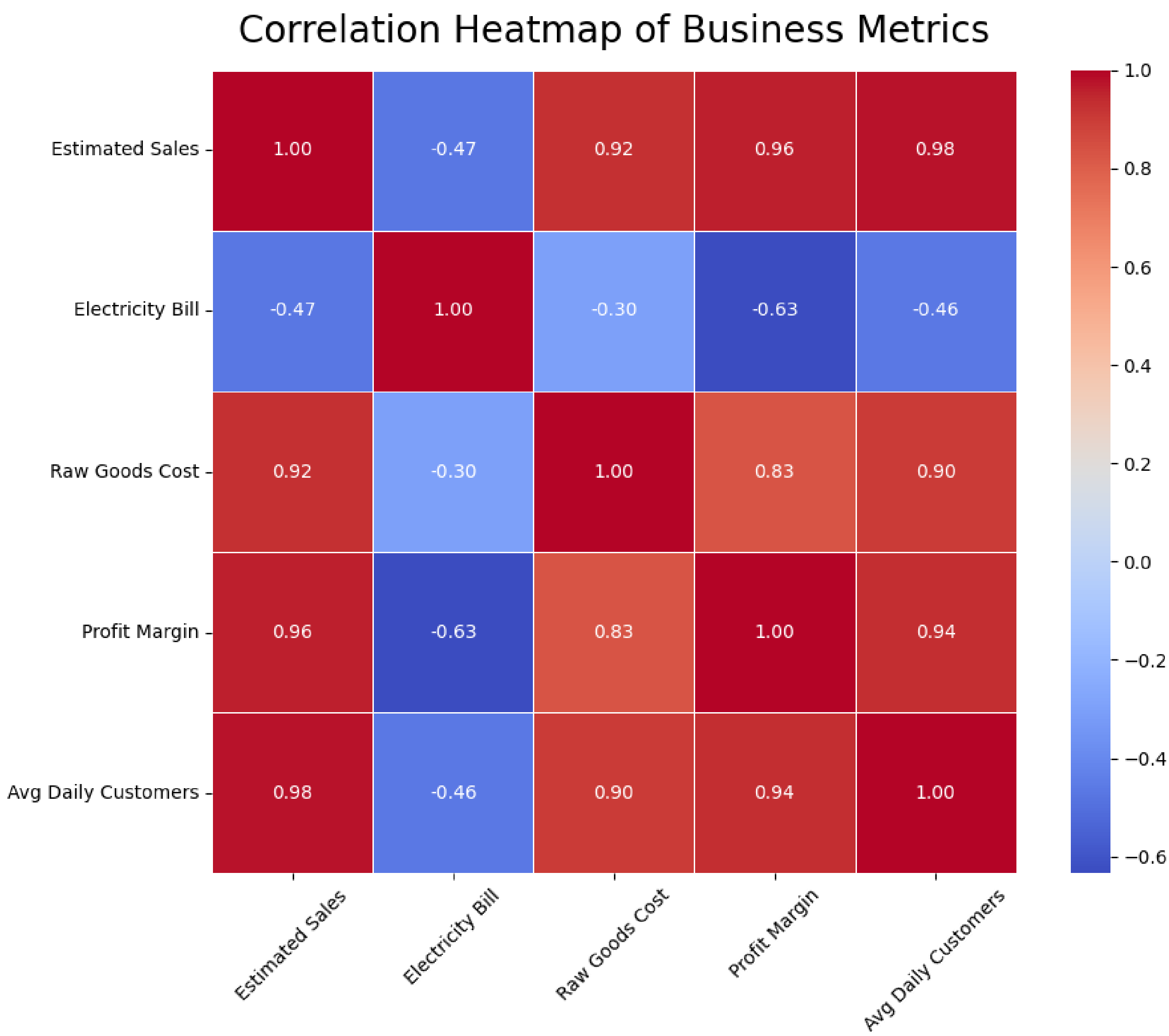


Fig 9: Correlation Heatmap of Business Metrics (Note: Employee Costs, which remains constant across all months, is excluded from this heatmap since their correlation is undefined.)

The interpretation of this heatmap can be seen in the following table given below:

Variable Pair	Correlation	Insights
Estimated Sales – Avg Daily Customers	0.98	More daily customers strongly drive sales.
Estimated Sales – Profit Margin	0.96	High sales almost always result in higher profit margins – due to efficient cost absorption.
Estimated Sales – Raw Goods Cost	0.92	Sales and raw goods cost increase together. Indicates that ingredient usage scales directly with demand.
Avg Daily Customers – Profit Margin	0.94	Suggests that higher footfall also improves margin, possibly due to economies of scale.
Raw Goods Cost – Profit Margin	0.83	Implies optimized raw material usage is connected to profitability.
Raw Goods Cost – Avg Daily Customers	0.90	More customers equals more raw material usage.
Profit Margin – Electricity Bill	-0.63	High electricity bills cut into profit margins, suggesting poor energy efficiency or high usage in low-profit months.
Estimated Sales – Electricity Bill	-0.47	Higher electricity bills don't lead to more sales, possibly due to off-season cooling or kitchen inefficiencies.

Table 6: Correlation Heatmap Interpretation of Business Metrics

Next, we will look into some '**What If**' scenarios based on the business metrics. This technique helps simulate how changes in key variables could affect profitability, cost structure, and overall business performance. To see how specific operational changes might influence overall profitability, we simulated five What-If scenarios and predicted their corresponding profit margins. The results are given in the following table:

Scenarios	Predicted Profit Margin (%)	Interpretation
+15% Customers	**17.76%**	Slight increase in profitability due to higher sales volume., showing the start of moderate economies of scale.
–10% Raw Cost	**20.15%**	Significant gain in margin. Raw materials form a major cost driver, and any efficiency of inventory management or better negotiation with suppliers greatly improves profit.
+10% Sales	**25.05%**	Best-case scenario. Direct revenue growth outcomes cost increases. Indicates that boosting demand or increasing prices has strong potential.
–20% Electricity Bill	**22.29%**	Operational savings help margins. Energy efficiency investments or off-peak operations can contribute to better profitability.
–10% Employee Costs	**18.00%**	There is slight improvement. Labor optimization affects margin positively, but not as strongly as raw materials or sales strategies.

Table 7: 'What If' Scenario Analysis with Predicted Profit Margins

4. Interpretation of Results and Recommendations

Problem 1: Inventory Management and Wastage Issues

Interpretation:

- Most of high sales in May happened during weekends (Saturday and Sunday) with a surprising sales peak in Wednesday. Keeping the stock prepared for those days is important to manage inventory and wastage reduction.

- Overstock vs. Demand Mismatch: Wasteful-day analysis flagged 05-May and 08-May as having high inventory per customer but low sales, confirming over-preparation on slow days.
- ABC analysis identified Drums of Heaven and Other Items as low-margin yet inventory-heavy, meaning stocking these items amplifies wastage risk. Their contribution margin significantly lower than others. On the other hand, Biriyani, CST Rice (Category A) contributes around 63% of total profit, indicating these should be prioritized for stock optimization.
- In a simulated 'What If' scenario, 10% reduction in raw cost improves profit margin from ~17% to ~22.4%, illustrating large upside to cutting waste. Proper management of inventory and wastage reduction are crucial for profit-margin improvement.

Recommendations:

1. **Daily Inventory Audits**
    - Records of opening and closing stock levels and compare against sales data to identify wasteful days. Flag days with exceeding 75th percentile in inventory per customer ratio. With daily audits the flagged days could be reduced.
2. **FIFO Implementation with Real-time Dashboard**
    - Older stock gets wasted whenever newer stock comes through. To tackle this enforcing **First-In-First-Out (FIFO)** in kitchen storage and ingredient bins using shelf labels or colour codes is crucial in this case.
    - And, to sufficiently monitor the inventory, applying Real-Time inventory dashboard is beneficial to show daily stock, usage, reorder alerts, and waste flags.
3. **Coordination with Supplier**
    - Building good connections with the suppliers to build dynamic and more fast supply of the goods will ensure to maintain a good stock level.
4. **SKU Prioritization**
    - Using Combo offers and discounts on low margin and lower category (Category C) SKUs can clear out the inventory to reduce wastage.
    - Also doing partnerships with local NGOs to donate unsold items, perishables to reduce wastage and supporting the local community.

By reducing wasteful days, the business can preserve cash and significantly cut spoilage expenses. Adding FIFO and implementing dashboard will reduce the wastage problem more significantly. Also, SKU bundling can shift demand toward higher-margin items, which lowers inventory costs and accelerates turnover.

Problem 2: Profit Optimization

Interpretation:

- Most of the sales during the year occurred during October and from December-to-January during festive seasons. These months are critical for business performance. Least amount of sales happened during April-to-August in the rainy seasons which affected the overall profit (~18%) in a year.
- Break-even requires ~3,106 units/month but current sales are ~2,022 units. Around **1,084**-unit gap, implying consistent monthly losses. This significantly affects the overall profit margin.
- Apart from Raw Goods Costs, Electricity Bills also erode profits (regression coefficient = -0.0036) as a hidden cost. Mostly it affects during rainy seasons.
- 'What If' Scenarios shows:
    - **+10% sales** → margin ↑ to **27.4%**
    - **–20% electricity** → margin ↑ to **24.1%**
    - **–10% raw cost** → margin ↑ to **22.4%**

  These simulations show how certain scenarios can affect the profit margin.

Recommendations:

1. **Installation of POS**
    - To track daily sales for more efficient sales tracking, applying **Point-of-Sale** (POS) system is very much beneficial. Integrating it with inventory dashboard will give an advanced view of supply chain management.
2. **Hiring Temporary Workers**
    - In the peak seasons during October and from December-to-January, hiring temporary workers for inventory work, stocking-restocking, customer handling could increase the sales drive and increase the operational efficiency.
3. **Sales Growth Initiative**
    - Advertising digital campaign and loyalty discounts will boost average ticket size and footfall. Since sales volume is the strongest profit driver, running customer retention campaigns (combo offers, loyalty points, etc.) will boost profit margins.
4. **Menu Optimization Based on Contribution Margin**
    - Using the SKU-wise contribution margin data to promote high-margin items (e.g., combo meals, Biryani SKUs, CST Rice SKUs etc.) will increase average order value without increasing cost base.
5. **Introducing Home Delivery as a Service**
    - As there is no such Delivery Partner in a more Remote/Local area, introducing home delivery

services will enhance customer satisfaction and convenience.

- Incorporating local people with efficient ways of transportation (e.g.: Cycle, Bike, Scooter etc.) will also manage delivery costs.

Applying POS will track daily transactions to optimize the operations. In peak seasons, incorporating part-time workers will increase the speed of operation. Prioritizing high valued SKUs will boost the sales and will increase the cost base. Increasing prices of those SKUs would not affect the demand also. So, a total make over of menu will definitely boost the margins.

All datasets and additional information: Link