Optimizing Inventory and Maximizing the Profit of a Café Business

A Mid-Term report for the BDM capstone project

Submitted by

Name: Sourasish Ghosh

Roll No: 23f3004075

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

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. The restaurant currently has 8 full-time employees with 9 suppliers. This report takes an analytical approach to understand the concurrent problems of the business. The major problem it’s facing are mainly: customer frequency, inventory problems, maintaining profit margins hindering it’s potential and stability.

Sales Data and SKU Data were collected with the collaboration of the restaurant owner and the kitchen staffs. In the Sales data, there is 12 months of data from January, 2024 – December, 2024 like :

Estimated Monthly Sales (₹), Monthly Employee Costs (₹), Monthly Electricity Bill (₹), Monthly Average Raw Goods Costs (₹), Profit Margin of each Month (%), Average Daily Customers.

The SKU dataset consists of 7 different types of SKUs of the restaurant of each month and metadata about price per unit (₹). Descriptive Statistics about the datasets revealed interesting patterns using mean, median, mode, standard deviation etc. For instance, from the SKU data ‘Biriyani’ has highest sales among the other SKUs while ‘Ektu Baithak Special Nawabi Kebab’ has the lowest. Beverages like Coffee, Cappuccino shows higher sales in colder seasons from October – February. The mean sale and profit margin of a month is around ₹274454 and 20% respectively. Although the median Profit Margin is 22.13%, indicating the business suffers loss in some months. Average customers in a month are around 92 and median of 75, shows consistency. Most of the SKUs are positively skewed, meaning the sales are not uniformly distributed. The Pie-Chart identifies the revenue share of each SKUs. Monthly estimated sales with profit margin plotted to identify the seasonal demands. The SKU sales trend plot shows trends of SKUs sold throughout the year. Analysing Break-Even Point gave the minimum number of units to be sold per month (which is around ~3106 units/month) to recover the costs. Assessing T-test and P-value of pairs of SKUs gave similarities of sales. Correlation mapping gave monthly sales spikes similarities of SKUs. Also building linear regression model and calculating R^2 score and Regression Coefficient gave inter relation between Profit Margin and other 5 variables. These methods are used align the restaurant’s problems such as inventory management, seasonal variability, cost effectiveness.

The analysis revealed many significant outcomes. In order to get these findings, methods like time-series analysis, revenue share of the SKUs, seasonal variability impacts etc were very useful. Most of the analysis and plotting were done on Google Sheets and Google Colab using Python and its libraries

(e.g.- Pandas, Matplotlib, Seaborn, Scipy). Analysing the revenue share of the SKUs, 'Biriyani' holds ~35% of the income of the restaurant. During festive seasons mainly October - November, Biriyani sales rise up ~57% , Noodles sales go up to ~68% . The Break-Even Point (BEP) is ~3106 units/month. The Weighted Average Contribution Margin is ~₹64.61/unit. R^2 Score is around 0.964, which is high. These insights helped to understand the inventory changes during seasonal changes, threshold limits of price points, profit margins, SKU inter relation etc.

2. Proof of Originality of the Data

- Business Name: Ektu Baithak
- Adress: Tamlipara bus stand (near STKK road), Jirat, Balagarh, Hooghly
- Owner's Name: Arghya Ghosh

1. Video Interaction with the owner: Video / Drive Link
2. Letter from Organization: Letter / Drive Link
3. Café Performance Report: Google Form Report

Fig-1 : Images of the restaurant, Ektu Baithak

3. Metadata

- Data Format: Excel/Sheets (.xlsx)
- Range: January 2024 – December 2024
- Units of measurement for features involving money: Indian Rupee (₹)

Information about the Sales Data:

Dataset: Sales Data / Link

Metadata refers to the information that provides context about the collected data. The Sales data was collected through the owner of the business and some kitchen staffs of the restaurant. It was mainly collected by verbal conversation and interviews. Each variable below is crucial in terms of analysing business performance. **Month** of the year will organise the data to see the seasonality of the sales. **Estimated Sales (₹)** shows the business performance and growth directly. **Employee Costs (₹)**, **Electricity Bill (₹)**, **Avg. Raw Goods Costs (₹)** tracks expenses and operational costs to understand the cost structures. **Profit Margin (%)** for each month shows the efficiency of the business is running. **Average Daily Customer** for each month indicates the business traffic which is useful to adjust the employees and other costs of the business.

Features of the dataset:

Name of Field	**Type**	**Description**
Month	Text	Includes all months of a year
Estimated Sales (₹)	Numeric	Estimated sales of each month in INR
Employee Costs (₹)	Numeric	Fixed employee cost of each month
Electricity Bill (₹)	Numeric	Average electricity bill of each month
Average Raw Goods Cost (₹)	Numeric	Estimated raw goods price
Profit Margin (%)	Numeric(percentage)	Includes profit margin for each month
Average Daily Customer	Numeric(frequency)	Avg. daily customers for given months

Information about the SKU Data:

Dataset: SKU Data / Link

Features of the Dataset:

Name of Field	Type	Description
Month	Text	Includes all month of a year
Biriyani (per Plate)	Numeric	Includes both Chicken and Mutton biriyani
Chef Special Tripple Rice (per Plate)	Numeric	Specially made rice in the restaurant
Drums of Heaven (6 Pieces)	Numeric	Chicken lollipop/drumstick specially made in the restaurant
Ektu Baithak Special Nawabi Kebab (6 Pieces)	Numeric	Special kebab dish of the restaurant
Noodles (per Plate)	Numeric	Includes all types of Noodles, e.g.: Hakka Noodles, Chicken Noodles etc.
Beverages (Cappuccino, Coffee etc.) (per Cup)	Numeric	Mostly hot beverages served as per cup
Others (various other SKUs)	Numeric	It includes all other varieties of dishes
Price per unit (₹)	Numeric	Tells the prices of each SKU

The SKU or Stock-Keeping Unit dataset was collected through the owner. The dataset includes all **SKU units sold** throughout the year for each month. It was mostly collected verbally as he had a registry of units sold during the year and I had to manually put down every entry. This gives a great overview about which items are sold more frequently and which of them are affected due to seasonal changes. The dataset is crucial in order to manage and utilise the inventory.

SKU	Total Units Sold	Avg Monthly Sales	Price/Unit (₹)	Estimated Annual Revenue (₹)
Biriyani	7560	630	150	1134000
Chef Special Tripple Rice	2736	228	180	492480
Drums of Heaven	2487	207	160	397920
Nawabi Kebab	1774	148	200	354800
Noodles	4609	384	70	322630
Beverages (Cappuccino/Coffee)	3112	259	60	186720
Others	1996	166	120	239520

Table 1 : Summary of SKU Dataset

4. Descriptive Statistics

Parameters	Monthly Estimated Sales (INR)	Employee costs (INR)	Monthly Electricity Bill (INR)	Average Raw Goods Cost (INR)	Profit Margin (%)	Avg. Daily customers
Mean	274454.55	135000.00	5709.09	60000.00	20.00%	92
Median	244000.00	135000.00	6000.00	60000.00	22.13%	75
Mode	0.00	135000.00	4500.00	60000.00	0.00%	75
Standard Deviation	94144.96	0.00	1292.64	11653.33	0.23	35.99
Sample Variance	8863272727.27	0.00	1670909.09	135800000.00	0.05	1295.25
Maximum	450000.00	135000.00	8000.00	85000.00	49.56%	151
Minimum	160000.00	135000.00	4000.00	45000.00	-16.56%	45
Sum	3019000.00	1485000.00	62800.00	660000.00	220.02%	1016
Skewness	0.50	0.00	0.26	0.76	-0.19	0.54

Table 2 : Descriptive Statistics of Sales Dataset

Parameters	Biriyani (per Plate)	CST Rice (per Plate)	Drums of Heaven (6 Pieces)	EBSNK (Kebab) (6 Pieces)	Noodles (per Plate)	Beverages (Cappuccino/Coffee) (per Cup)	Others
Mean	621.83	223.50	223.92	160.33	380.58	321.50	146.33
Median	570.50	205.50	229.00	148.00	353.00	270.50	123.50
Mode	0.00	0.00	0.00	120.00	286.00	0.00	100.00
Standard Deviation	197.19	69.35	59.53	53.35	124.99	146.12	67.40
Sample Variance	38,883.97	4809.91	3543.90	2846.61	15622.63	21350.82	4543.15
Maximum	980	375	302	270	640	612	310
Minimum	362	150	138	93	221	150	78
Sum	7,462	2,682	2,687	1,924	4,567	3,858	1,756
Skewness	0.58	1.15	-0.14	0.87	0.85	1.04	1.50

Table 3 : Descriptive Statistics of SKU Dataset

The above results are produced using Excel formulas e.g.: AVERAGE,MODE,STDEV etc. to calculate parameters from the Sales data-wise and SKUs-wise descriptive statistics. Some of them includes:

- The maximum sales happened during a month is about **₹450000** and the month is **October.** Lowest sale is about **₹160000** in **July** .
- Average daily customers at the restaurant : **89**.
- Maximum profit margin is **50%** and minimum is about **-17%** .
- Highest number of sold dish in a month is **Biriyani (980 plates)** and following **Noodles (640 plates)** and **Cappuccino/Coffee (612 cup servings)** .

The calculated descriptive statistics data is directly linked to the problem statement.

- Sales Data: The data on monthly sales, costs surrounding the business, profit margin, average daily customers give a financial overview of the business. Analysing and setting up margins, by cost cutting, the desired income can be achieved.
- SKU Data: The amounts of dishes being served every month can be used to utilise the inventory and build an inventory management system around it.

5. Detailed Explanation of Analysis Process/Method

5.1. Data Cleaning and Preprocessing

Most of the data is being collected verbally from the owner. So, there were multiple issues like: typographical issues in the sales, inappropriate profit margins, Non-Standardisation in SKU units, issues calculating Descriptive Statistics etc. After cleaning the data, it gives a standardised way of operating throughout the dataset.

While collecting Sales dataset, the owner gave approx. profit margin of the business of year 2024 (about 18-22%), which is not sufficient. I had to work on Google Sheets calculating profit margin from monthly sales to monthly costs for each month [e.g.: Profit Margin =(B2-(C2+D2+E2))/B2*100 , where column-B: Estimated Monthly Sales, column-C: Employee Costs, column-D: Monthly Electricity Bills, column-E: Average Raw Goods Cost ; same formula for other rows].

In SKU dataset, the SKUs were not standardised. As an example, Biriyani includes all the categories of Biriyani served by the restaurant (e.g. Chicken Biriyani, Mutton Biriyani etc.). 8 types of Noodles are on the menu and 5 types of hot serving beverages. Had to take **mean** value of each SKU to make a standard price(₹)/unit .

Most of it's being done by Excel/Google Sheets using proper data entry and also using commands like **IFERROR** . As an example, in Sales data, all the monthly sales and profit margin are unique valued so MODE values were showing ERROR in Sheets. To avoid that using IFERROR omits the error [e.g.: =IFERROR(MODE(B2:B12),0) ; where B2:B12 is monthly estimated sales from Jan-Dec]. Similar method applied in SKU data also. Can be seen in *Table 2* and *Table 3*.

5.2. Analysis Process/Method

1. SKU Performance and Revenue Impact

Descriptive Statistics Analysis : By looking and comparing the statistics of the SKU dataset we can identify which SKUs are performing poorly or inconsistently. *Table2* and *Table 3* will be useful.

Other than that, comparing SKU revenue share with **Pie-Chart** will give clear picture of high-demand SKUs, thus managing the inventory accordingly.

2. Trends and Demand Forecasting

Time-Series analysis: It is very suitable for financial data as it gives out the trends, patterns, and variations of parameters of parameters over time. In *Fig-3* and *Fig-4*, it's already been shown and it can be clearly seen the trends in the course of a year. As an example, during the rainy seasons, to maintain the margin cost-cutting can be done in various ways.

3. Cost Recovery Threshold

Break-Even Analysis: This analysis gives the point where Total Revenue = Total Costs. Also calculating Break-Even Points in Units or **BEP (Units)** gives how much units of each SKUs are needed to Break Even the sales.

4. Impact Assessment

T-test Analysis: Doing T-test will compare the means of two groups of data. Grouping the SKUs will give information if their sales are related or not. Comparing **P-values** gave the clear picture of Sale-similarities of SKUs. It will improve managing inventory.

Correlation Mapping: Doing Correlation analysis of the SKUs gave insights how strong they are inter-related. Applying heatmaps also made sure more cohesiveness of the data.

5. Monthly Profit Drivers

Regression Analysis: Doing regression analysis and interpreting **R^2-Score** and comparing **Regression Coefficients** we have information about which of them in Sales dataset are influencing the profit margins.

Also having a conversation with the owner provided a qualitative insight about the business. It gave a clear understanding about the customer behaviour and current trends that repeat and new customers are adopting. As an example, customers in the rural area are much lesser prone to go have an outing in rainy days comparing to urban areas.

These analysis methods have been useful to find fruitful insights of the business. Most of the analysis and plotting were done in Google Sheets and Google Colab. All the links are provided at the end of the document.

6. Results and Findings

6.1. SKU Performance and Revenue Impact

Evaluating *Table-2* and *Table-3*, we get concise overview of **descriptive statistics** of datasets. The mean profit margin is about **20%** and mean monthly sales is around **₹274454**. Maximum and Minimum customers visited during a month are **151** and **45**, respectively. Skewness of Profit Margin is about **-0.19** (Negative Skewness) tells few of the month's sales are unusually low which drags the Profit Margin overall. Monthly Raw Good Costs have skewness of **0.76** (Positive Skewness) tells some months had significantly higher costs than usual.

Some of the key findings of sales are:

- Mean : ₹266583
- Maximum : ₹450000
- Minimum : ₹160000
- Standard Deviation : ₹93814

The revenue share can be sought out with the help of **Pie Chart** given below at *Fig-2*. It can be clearly seen that **Biriyani** holds the most of the revenue share (35.2%) following **Chef Special Tripple Rice** (15.2%), **Drums of Heaven** (Chicken Drumstick) (13.5%), **Ektu Baithak Special Nawabi Kebab** (12.1%) and rest can be seen in the pie chart.

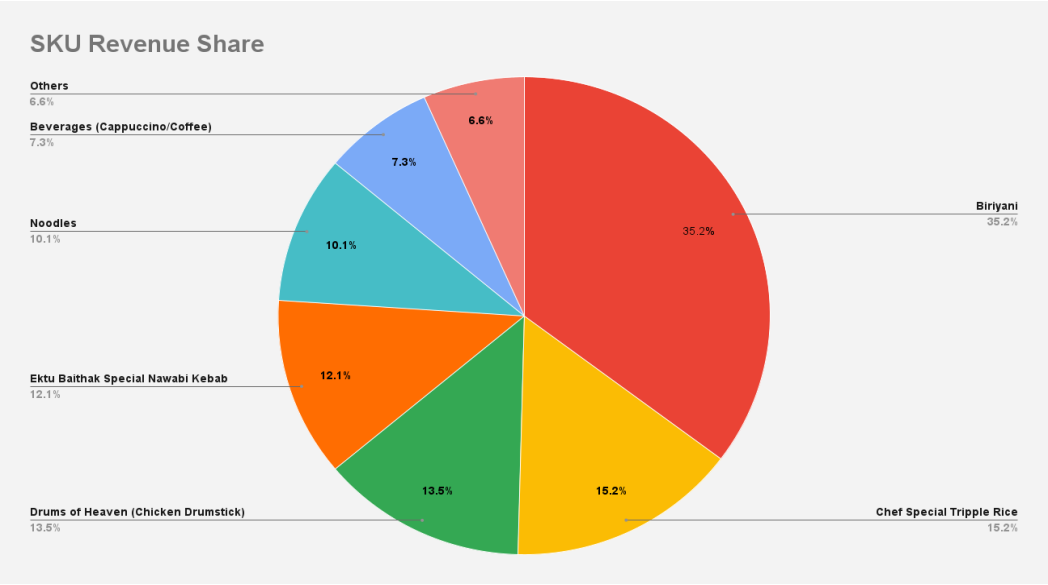

Fig-2: Pie Chart of SKUs Depicting the Revenue Share (%)

6.2. Trends and Demand Forecasting

Analysing the **Time-Series** data following from *Fig-3* and *Fig-4* there are some key insights that can be seen . **October** had the highest sales of about **₹450000** and profit margin of ~**50%**. April **(-4%)**,

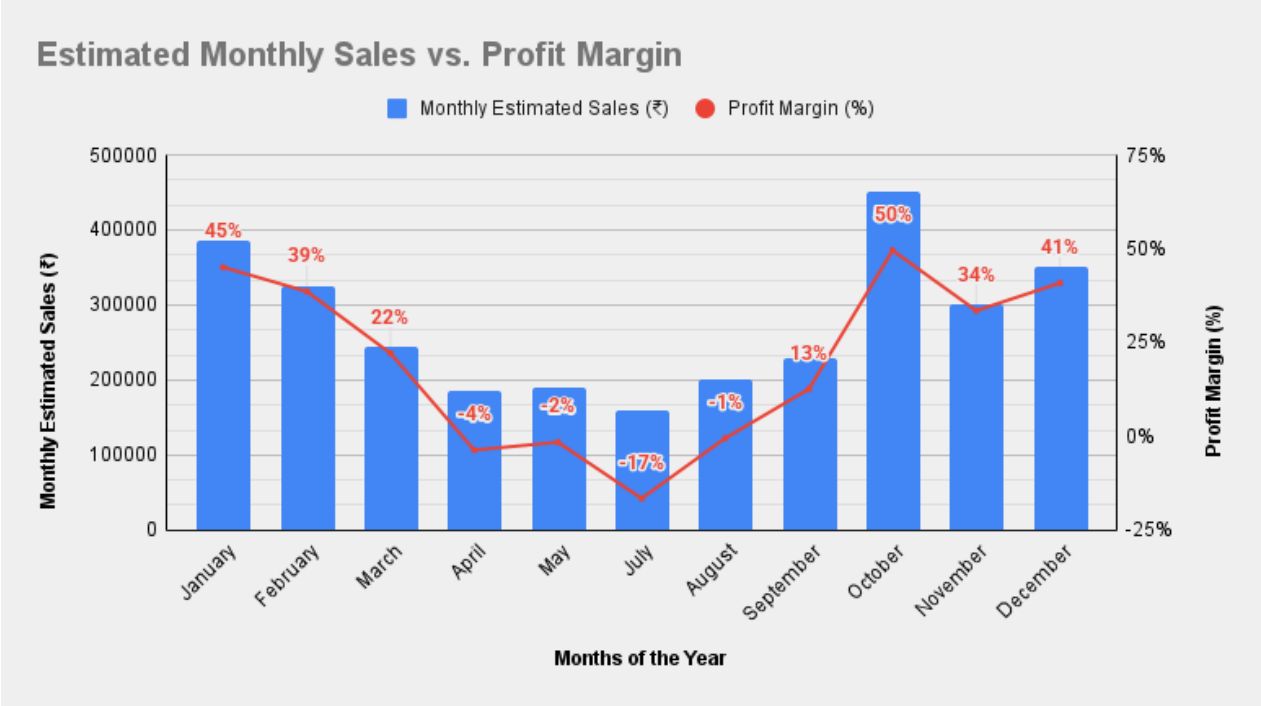

Fig-3: Combo Chart between Monthly Estimated Sales (₹) and Profit Margin (%)

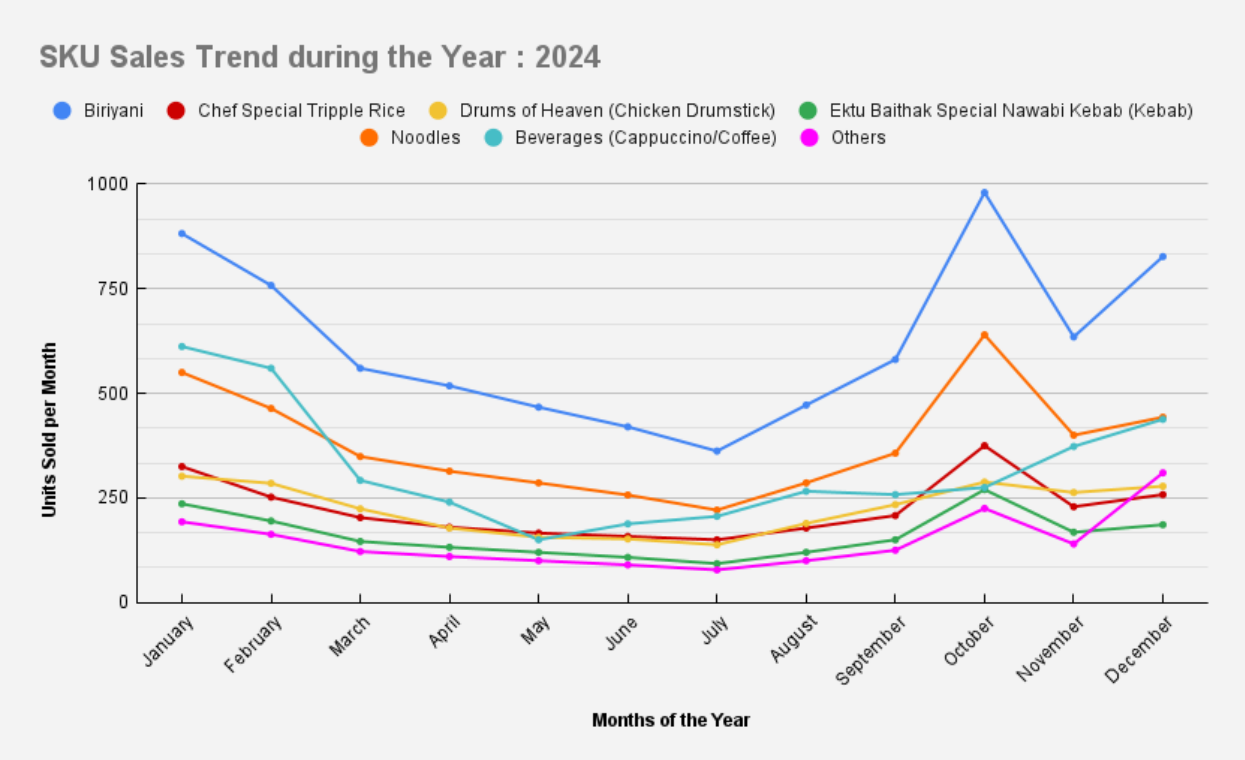

Fig-4: SKU Sales Trend During year 2024

May **(-2%)**, July **(-17%)**, and August **(-1%)** show **negative profit margins**, indicating **losses** throughout these months. **July** had the lowest sales of about **₹160000** and profit margin of about **-17%** indicating steady and higher cost spendings. From September the sales start to have a steady growth till March. Mostly the business gets affected in Rainy seasons. Analysing through ***Fig-4***, it shows the SKU selling trends throughout the year. **Biriyani** being the highest sold SKU in all months, highest at October (**980** units or plates) following **Noodles** (**640** units or plates) and **Beverages** (**612** units or cups). Most SKUs have high sales in October except Beverages. They have much higher sales in the interval November to February, max in January due to winter seasons.

6.3. Cost Recovery Threshold

To recover the fixed costs (including raw goods cost) the business needs about **₹200710/month** in sales. The Break-Even Point (**BEP**) Units is ~**3106** units/month, it means the business have to sell these amounts of SKUs (across all SKUs) to recover the fixed costs. **Chef Special Tripple Rice** has the highest **Contribution Margin** (CM) per unit : **₹120/unit** and **Noodles** have the lowest about **₹25/unit**. The **Weighted Average Contribution Margin (WACM)** is **₹64.61/unit**. 3 SKUs have more than WACM (Biriyani, Tripple Rice & Special Kebab).

Break-Even Point (BEP) Formula

$$\text{BEP} = \frac{\text{Fixed Costs}}{\text{Weighted Average Contribution Margin per unit (WACM)}}$$

$$\text{WACM} = \sum_{i=1}^{n} (\text{CM per unit}_i \times \text{Sales Mix Percentage}_i)$$

Where:

- n = total number of products
- CM per unit_i = Contribution Margin per unit for product i
- $\text{Sales Mix Percentage}_i$ = Sales Mix Percentage for product i

Formulas of BEP and WACM

Where, **Contribution Margin (CM)** is the amount of revenue remaining from sales after covering variable costs.

6.4. Impact Assessment

Assessing the **T-test** gave the pairs of SKUs with similar sales, such as; Chef Special Tripple Rice and Drums of Heaven (Chicken Drumstick) have similar average monthly sales. Their T-statistics is about **-0.0158** which is very small and P-value of ~**0.98**, indicates they have very high similarities statistically. Following Noodles and Beverages also have similar pattern. Their T-statistics is about **1.064** and P-value of ~**0.3** . Other than that, all other pairs are more or less statistically different. The overall sales difference between SKUs are very high. It can be clearly seen in **Correlation Heatmap** in the following *Fig-5.* In the Correlation Matrix it can be seen that correlation between some SKUs

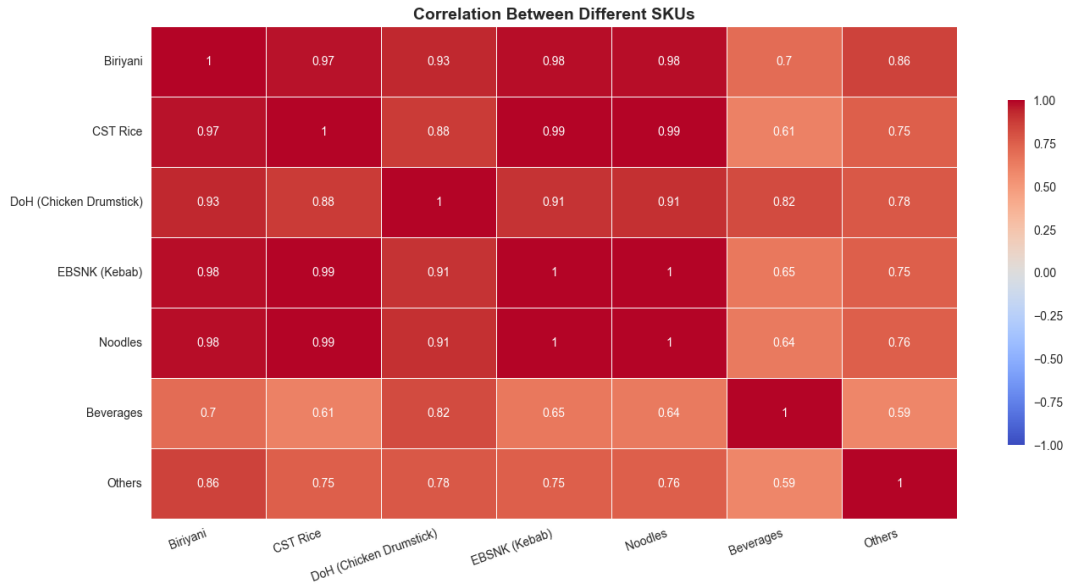

Fig-5: Correlation Heatmap between SKUs

is very high, as an example; Biriyani and Noodles have correlation of **0.98**, Biriyani and Tripple Rice have **0.97** correlation, Tripple Rice and Special Kebab and also Tripple Rice and Noodles have **0.99** correlation. These high correlations indicates that their monthly sales spikes are same but the actual sales difference and volume is much higher. As an example, T-statistics of Biriyani and Noodles is about **0.00167** which is way lower comparing their correlation, which indicates the sales are different.

6.5. Monthly Profit Drivers

Building a Linear Regression Model to predict the monthly profit margin gave a very high R^2-Score of **0.964**, meaning the model explains 96.4% variation of the profit margin. Profit Margin was taken as dependent variable and other 5 parameters (Estimated Sales, Employee Costs, Electricity Bill, Raw Goods Cost, Average Daily Customers) were taken as independent variable. Calculating **Regression Coefficients** gave the inter relation between the 5 variables and Profit Margin. It's 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

All datasets and additional information : Mid-Term Project