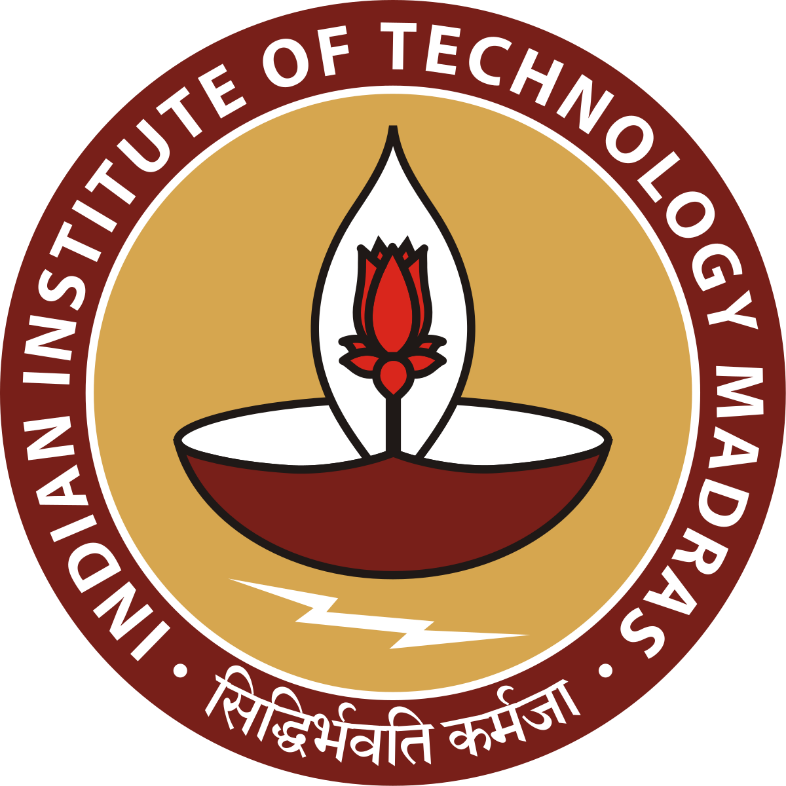
# **ELEVATING SUPERMARKET PROFITABILITY BY OPTIMIZING SALES AND INVENTORY DATA**

**A Final Submission Report for the BDM Capstone Project**

Submitted by

Name: **Preethu Johnson**

Roll number: **22F1001747**



IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

**Table of Contents**

[**Table of Contents** 0](#_Toc185087897)

[Executive Summary and Title 2](#_Toc185087898)

[Detailed Explanation of Analysis Process/Method 2](#_Toc185087899)

[Results and Findings 10](#_Toc185087900)

[Recommendations to Business 18](#_Toc185087901)

# Executive Summary and Title

**Title:** **Elevating Supermarket Profitability by Optimizing Sales and Inventory Data**

The project was conducted on Minnus Super Store, a leading B2C supermarket in Thrissur district, Kerala. The project aims to drive transformation in the store by improving profitability through effective sales strategies and streamlined inventory management.

The analysis began with Exploratory Data Analysis to identify key statistical measures, including descriptive statistics for sales and product quantities. This provided a quantitative context for understanding overall trends and variability in the store's operations.

The advanced data analytics tools, like Microsoft Excel and Microsoft Power BI were used in this study to clean, sort, and analyze the data, gather inferences from it. Subsequently, Explanatory Data Analysis was conducted to identify actionable insights. A combination of line charts, bar charts, and combination charts were employed to visualize sales trends, category-wise performance, and monthly revenue fluctuations. Key findings include the product demand trends, sales patterns, and customer retention. High-performing product categories contributing significantly to revenue were identified. Seasonal and monthly fluctuations were analyzed, showing peak sales during months April and August, and lowest performance in May. Purchase behavior and customer retention patterns were mapped to better understand loyalty and improve targeted sales strategies.

The findings were summarized into clear, practical recommendations aimed at improving operational efficiency and profitability. These recommendations include enhancing customer experience and boosting sales by focusing on locally produced items, seasonal goods, while introducing free delivery, pay-later options, and a small pantry for fresh snacks and optimize operations through daily inventory management, removal of slow-moving items, and strategic customer data collection.

# Detailed Explanation of Analysis Process/Method

Exploratory Data Analysis for Sales and Revenue

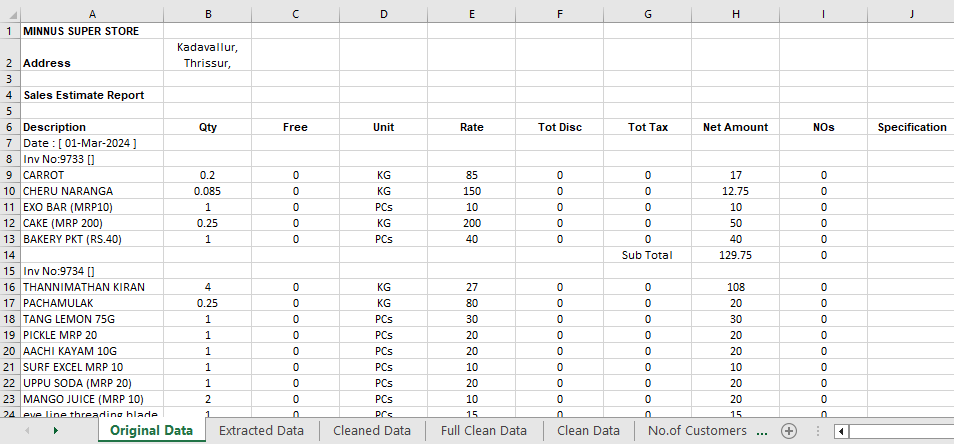
Data Collection

The data received from the Minnus Super Store was their sales data for the past 7 months from March 1 to September 30, 2024. The received data was unorganized, had blank columns and zero values. The data was available as sets of invoice bills issued for the items purchased by each customer on a particular day with the bill number and date.



**Figure 1:** The Bill Receipt at Minnus Super Store

The bill generated at the store contains the name and address of the shop, date and time, invoice number, provision for customer name and address, list of items purchased, their code, quantity, rate and amount, total number of items purchased, grand total and net total. This invoice bill data is computerized and was provided in the below format.



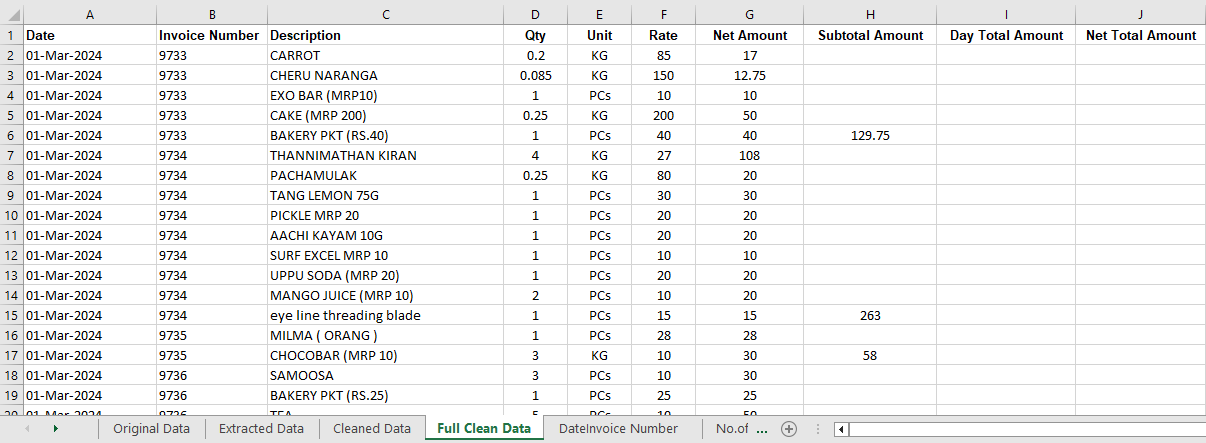
**Figure 2:** The Data Received from the Store

The data obtained had a total of 65257 rows and 9 columns. It had the details of date of purchase, invoice number of the bill, list of items purchased, their quantity, any free items given along, unit in which the item is sold, rate of the item, total discount provided, total tax on the items, net amount of the items, number of orders, and specifications.

Data Cleaning

The first task was to clean this data to an organized structure and do a quality check of data to find out any missing data or outliers that could affect the accurate analysis of data. Microsoft Excel was used for the cleaning process to deduce a time series data showing the sales of items over months.

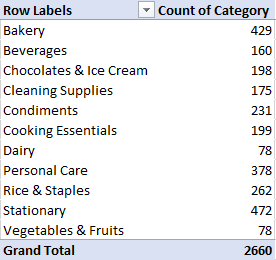
The filtering option in the Excel was used to delete the unwanted rows of invoice number, date, subtotal, day total and net total. For each column where these values were present, filter was applied, selected all and deleted. Then selected the empty rows and deleted them too to get clean data in the tabular form. Created different columns for Date, Invoice Number, Sub Total, Day Total and Net Total so that the analysis is easy by utilizing different excel formulas.



**Figure 3:** Cleaned Data in Structured Form

The next step was to understand the cleaned data and learn what is available from it. Wanted to know the number of items that was sold in the store. The Remove Duplicates option in the Excel was used for that and found that there were 2660 unique items being sold in the store.

The next task was to sort the items to different categories for better analysis. 11 different categories were chosen and manually sorted the data. This was the most time consuming and tiring part of the data cleaning process. But to know how each category contributes to the revenue of the store, this sorting was necessary.



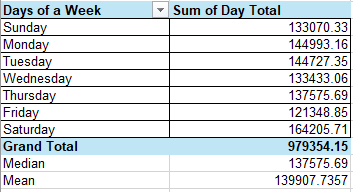
**Figure 4:** Category Sorting of Items

Data Analysis

Once the data is cleaned and sorted, the analysis of data was started. The data had vast information and so the analysis had to be done in various angles.

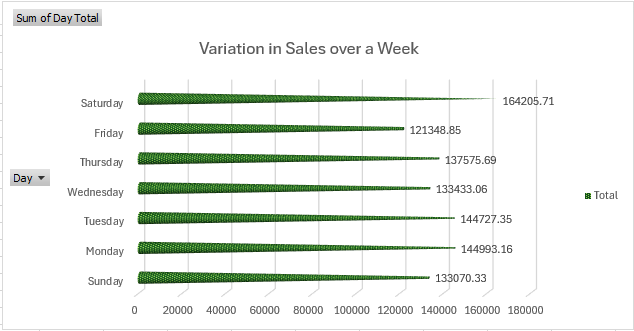
Started the analysis with the problems cited by the store owner. Had to find out statistically how the variations were there in the sales of the items during different days of a week and during different months throughout the selected period.

To know that, the Day Total amounts and the Net Total (Monthly Sum of Amount) amounts were analyzed. Sorted the day total data and took the cumulative sum of sales for each day in a week. Did the same for month total to get the cumulative sum of sales for each month in the selected period of March to September. Also, the mean and median of the sales over the week and through months were found.

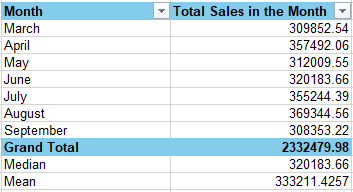


**Figure 5:** Mean and Median of the Sales Over a Week

The following bar chart represent the variation in sales over weeks during the 7 months.

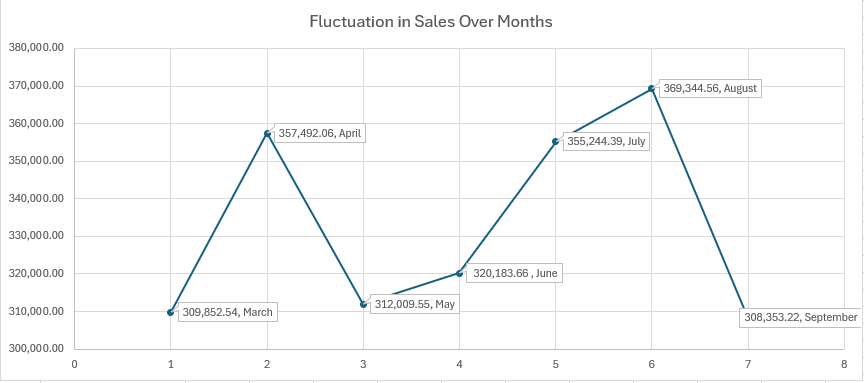


**Figure 6:** Sales in Different Days of a Week



**Figure 7:** Mean and Median of the Sales Over the Months

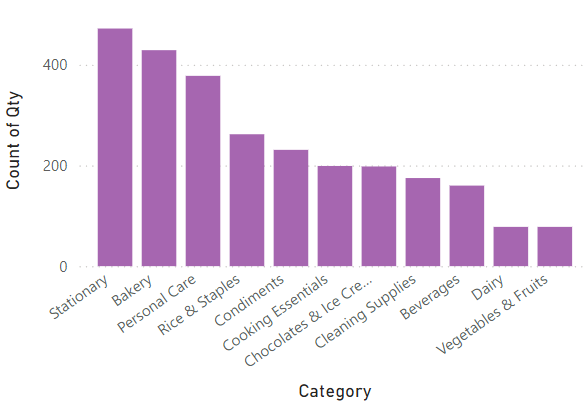
The following line chart represent the variation in sales over months in the selected period.



**Figure 8:** Fluctuation in Sales Over Months

The charts clearly show the problem raised by the owner, the drastic fluctuation in sales over months. The demand was to reduce the fluctuations so that there is a constant average revenue generated every month.

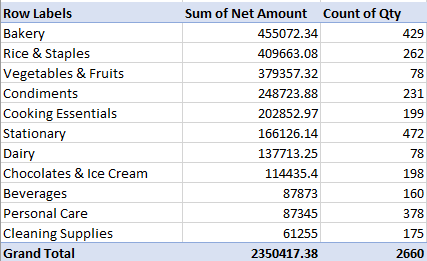
Next, focus was to know how the categories of items contribute to the revenue generated at the store. So, did the category wise analysis of the data. Microsoft Excel was used to sort the data and Microsoft Power BI to generate the charts for a better visual appeal.



**Figure 9:** Count of Items in Each Category

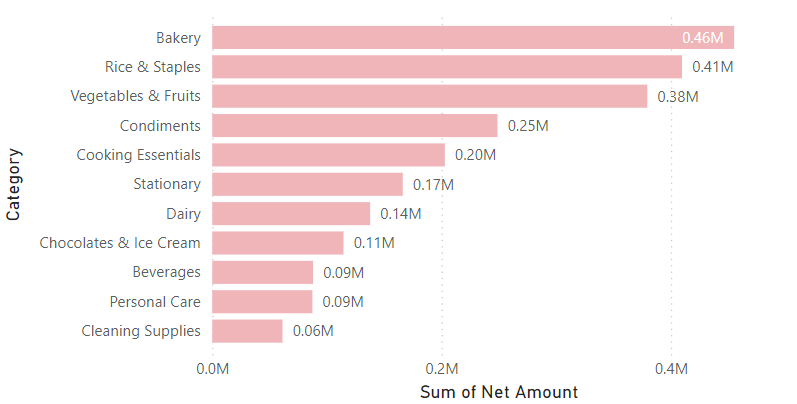
As seen from the bar chart, the store has more of stationary items which are non-perishable and have a higher shelf life. But the second highest is Bakery items and they can be easily perishable and might have less shelf life as well. So, the quick sales of those items are of high importance. The least number of products are in Dairy and fresh produce, that is Vegetables & Fruits. Both the category contains easily perishable items and their daily intake and sales is to be done with utmost care.

The count of items really does not matter much. What matters is the revenue generated out of it. So, checked the revenue generated out of each category. The order of items seemed bit different when it came to the revenue generation.



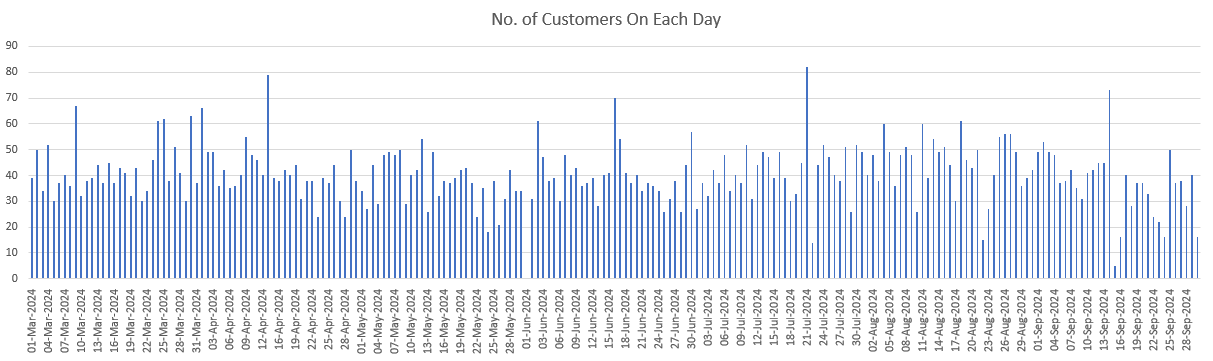
**Figure 10:** Revenue Generated by Each Category

Even though stationary items were in highest count, they did not generate the highest revenue in the store. The highest revenue was generated by Bakery, then Rice & Staples and Vegetables & Fruits, which had the least count.



**Figure 11:** Revenue Generated by Each Category

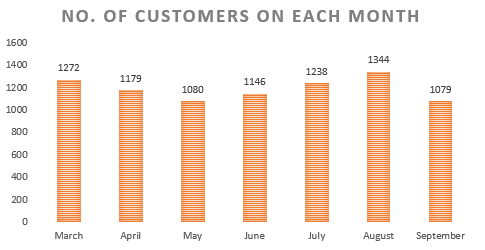
Considering each invoice bill was created for a unique customer, analyzed the number of customers visited every day during the selected time and the number of customers visited in every month. This was needed to know about the customer retention.



**Figure 12:** Number of Customers Visited Each Day During the Selected 7 months

The average of customer visited in a day

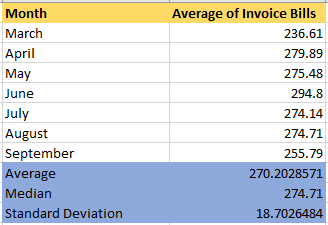
The standard deviation of customer visited in a day in selected period



**Figure 13:** Number of Customers Visited Each Month During the Selected 7 months

The average of customer visited in a month = =

The standard deviation of customer visited in a month in the period



**Figure 14:** Number of Customers Visited Each Month During the Selected 7 months

From the customer visit charts, it is visible that even though there is variation in the number of customers visit every day, when we see over months, the customer visits are almost stabilized with a standard deviation of just 99.34. That is a good sign, but from the point of sales and profitability, the customer visits should increase and there should not be any decrease in the customer count moving forward. That increases the profitability of the store.

# Results and Findings

Explanatory Data Analysis

The primary analysis done by creating tables and generating charts unveiled various insights. The mean (₹139,907.73) and median (₹137,575.69) values of the sales over a week are close, suggesting that sales are evenly distributed across the week. Saturday has the highest sales with a total of ₹164,205.71. This indicates higher purchases during weekends, possibly due to customers stocking up for the week.

Friday has the lowest sales with ₹121,348.85. This could imply fewer customers or reduced spending, perhaps due to people waiting for the weekend for bulk purchases. Sunday, however, has below-average sales, indicating that customer activity drops slightly after the Saturday rush. Monday has above-average sales, suggesting a spike in activity, possibly from customers completing weekend errands or starting their week with necessary purchases.

Both Saturday and Monday have above-average sales, suggesting increased activity around weekends. Tuesday, Wednesday, and Thursday remain close to the mean, indicating consistent but not peak sales during midweek.

The average number of customer visits per day is approximately 40.43, indicating a moderate daily footfall. The standard deviation of 11.47 suggests that daily customer visits vary moderately around the average. The weekend and week day trend caused this variation.

The mean (₹333,211.43) and median (₹320,183.66) values of the sales over months are relatively close, suggesting a consistent sales trend across months with no major outliers. April, July, and August consistently performed above average. March, May, and September are below the mean, indicating potentially quieter months in terms of sales. June has total sales of ₹320,183.66, which is exactly equal to the median and slightly below the mean (₹333,211.43). The sales in June are steady but not exceptional, as they are closer to the median value.

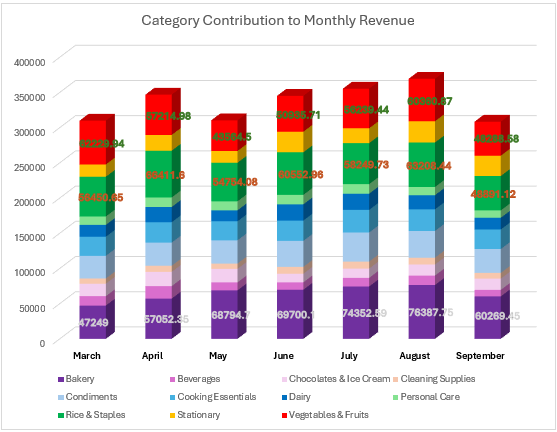
April and August had higher performance due to the seasonal festivals Vishu in April and Onam in August. March and May showed a dip because of the summer season. The sales started to increase from June as the rainy season started. It started to decrease after the rainy season in September.

The average number of customer visits per month is approximately 1191.14, which indicates a consistent flow of customers when aggregated over a month. The standard deviation of 99.34 indicates some fluctuation in the monthly customer visits, but this variability is relatively low compared to the monthly average. This could point to a stable customer base with occasional spikes or drops due to external factors like festivals or season changes.

June has the highest average invoice bill (294.8) but does not correspond to the highest revenue. This suggests that although customers spent more per bill, fewer customers visited compared to peak months. August generated the highest total revenue (369,344.56) even though the average invoice bill was moderate (274.71). This is due to the higher number of customers (1344) visiting during the month. March had the lowest average invoice bill (236.61) but still managed moderate revenue (309,852.54) because of a relatively high customer count (1272). August saw the highest footfall (1344 customers), directly contributing to the highest revenue despite no significant spike in the average invoice bill. May and September had the lowest customer counts (1080 and 1079), which reflects in their lower total revenue.

The daily variability is proportionally higher relative to the daily average compared to the monthly variability relative to the monthly average. This suggests that while daily customer visits can fluctuate noticeably, these fluctuations tend to stabilize when viewed over a month.

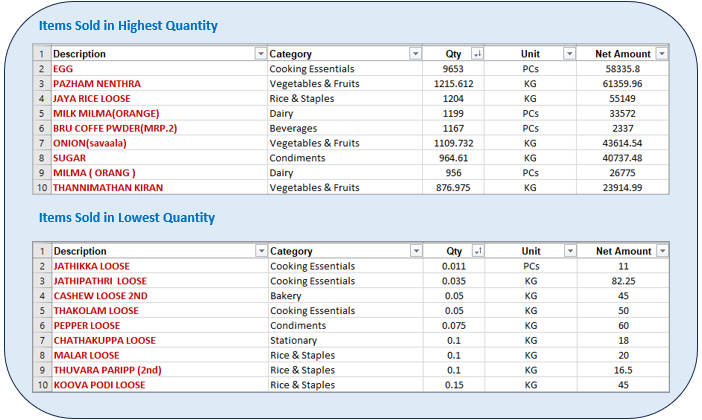
Once the primary analysis of the data was done, wanted to know how much each category is contributing to the monthly sales revenue. Stacked bar chart was used to analyze that.



**Figure 15:** Contribution of Each Category to the Monthly Revenue

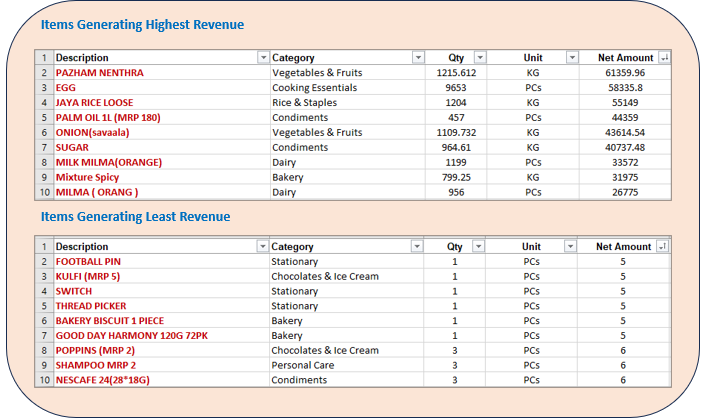
Every month, Rice & Staples, Vegetables & Fruits, and Bakery categories contribute maximum revenue. These categories consistently contributing the maximum revenue indicates that these are the core demand drivers for the store. This suggests that customers primarily rely on the store for essential, everyday food items and fresh produce, reflecting their staple dietary needs. The demand for staples and fresh produce suggests most customers likely come for routine household needs rather than luxury or specialty items.

After categories, next search was for the items that had the maximum and minimum sales in these 7 months. To know what products, generate greatest revenue and which least, the products that were sold the most and sold the least in every month were to be identified. Also wanted to know to which category they belong to for obtaining a clear picture. Needed to know whether the products belong to easily perishable goods or goods with large shelf life. If the least revenue generating goods are perishable, they create loss due to less sales and easy contamination. Avoiding that is very important.



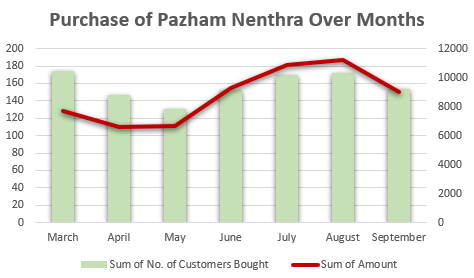
**Figure 16:** Products with Maximum and Minimum Quantity of Sales

But the highest count did not reflect in the highest revenue generated. The highest and lowest revenue generating items seemed bit different.



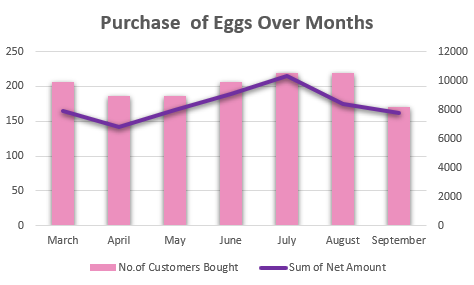
**Figure 17:** Products that Generated Maximum and Minimum Revenue

Next, interest was in knowing the month in which the highest revenue items were sold the most. Whether there is any pattern in that data. Analyzed the purchase pattern of Pazham Nenthra, the highest revenue generated item for this. The highest sales of this item were in the August month pertaining to the seasonal festival, Onam.



**Figure 18:** Purchase Pattern of Pazham Nenthra

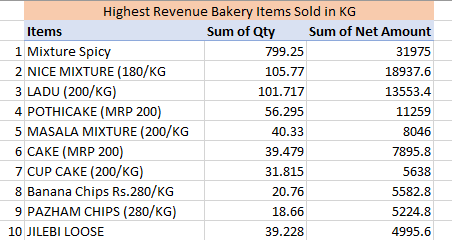
The purchase pattern of Egg was analyzed the next, the second highest revenue generated item. Egg was sold most in July, indicating a higher demand for protein during the rainy season.



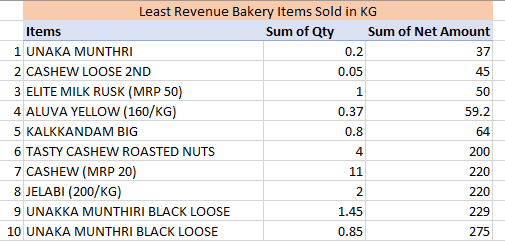
**Figure 19:** Purchase Pattern of Egg

After the overall analysis of items, next wanted to know the items in each category that was sold the most and the items that generated the highest revenue. The products sold in the store was in three different units. In kilogram, marked as KG in the data sheet, in packets, marked as PCs and in numbers marked as Nos. So, moved ahead with unit wise analysis, rather than an overall view, to get more insights about the sales. The filter option in the Pivot Table in Excel was used for analyzing the unit wise sales of the items.

Started with the highest revenue generating category, Bakery. The 429 items in the store generate a total revenue of ₹4,55,072.34. In that, 65 items were sold in KG in Bakery. A total of 1740.092 kg of bakery items were sold at the store and a revenue of ₹2,01,587.4 was generated. This contributes to 44% of the revenue generated by the whole bakery category.



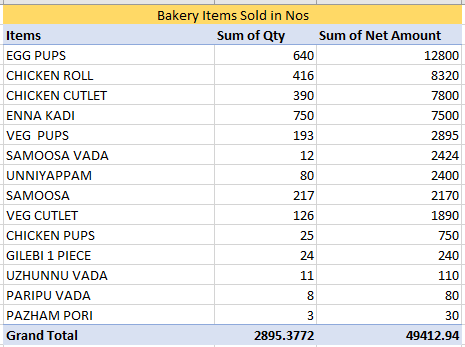
**Figure 20:** Top 10 Revenue Generating Items in Bakery Sold in KG



**Figure 21:** Least 10 Revenue Generating Items in Bakery Sold in KG.

These items with least sales just contribute 1% altogether to the total sales of items in Bakery in KG. The list shows that dry fruits do not have much market in the locality.

Only 14 items were sold in numbers and they generate a revenue of ₹49412.94. This cumulates to around 11% of the revenue generated by the whole bakery category.



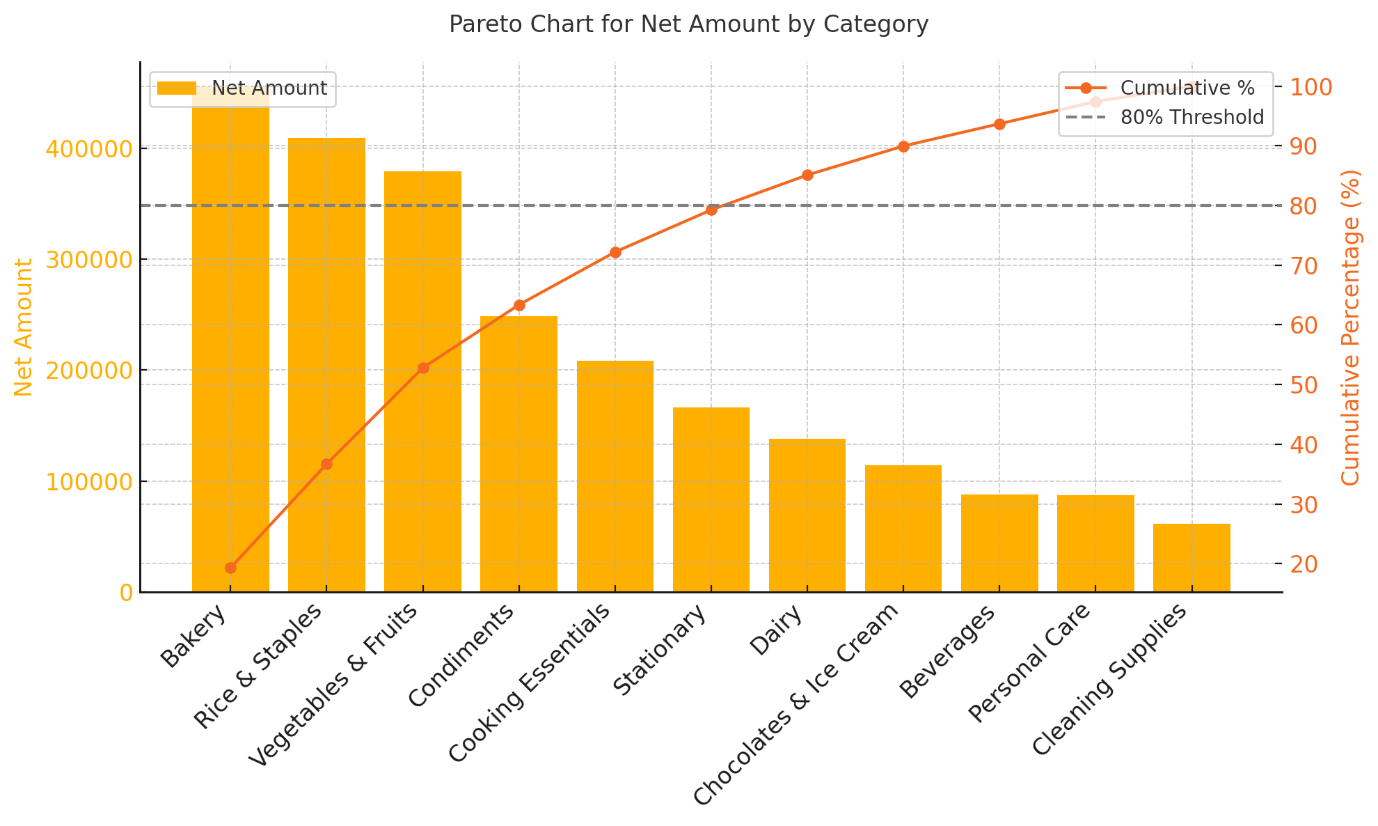
**Figure 22:** Top 10 Revenue Generating Items in Bakery Sold in Nos

A total of 350 items were sold in packets in the bakery category with a total revenue of ₹204072. This cumulates to around 45% of the revenue generated by the whole bakery category.

A similar analysis was done for all the items in all the categories to draw insights about them.

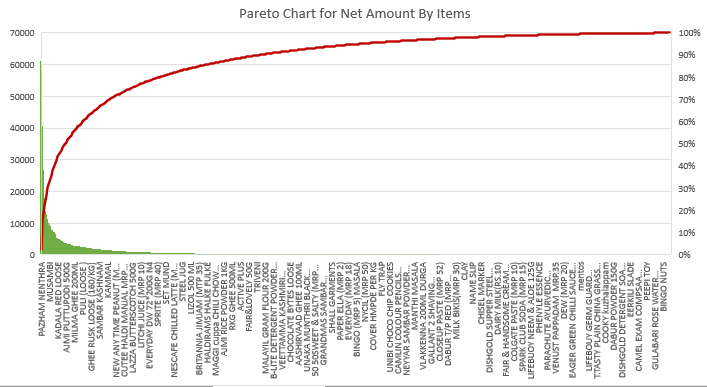
Then finally checked whether the sales follow the pareto principle. The Pareto Principle, also known as the 80/20 rule, is a concept that suggests that 80% of effects come from 20% of causes. It highlights an imbalance in input and output. If approximately 80% of the total sales comes from 20% of the categories or products, then pareto principle is valid.

The Pareto charts was generated for categories and items in the Minnus Super Store. The charts are given below.



**Figure 23:** Pareto Chart for Category

The Pareto Principle is not strictly valid with categories because it takes more than 20-30% of the categories (6 out of 11, ~55%) to achieve approximately 80% of the revenue. The revenue is more evenly distributed across categories compared to a typical Pareto distribution. This indicates that while some categories are strong contributors, the revenue spread is broader, and no small group of categories overwhelmingly dominates the total revenue.



**Figure 24:** Pareto Chart for Items

The Pareto principle is valid for the sales of items at the store. The analysis confirms that a small subset of products contributes disproportionately to the total sales.

All the calculation and analysis can be found in the Excel file at the following link:

[BDM Project Final Submission Excel Report](https://docs.google.com/spreadsheets/d/1bz0cp66VMdwyVzAGXjeAubEKnaJjdMnu/edit?usp=sharing&ouid=106796579137076179815&rtpof=true&sd=true)

# Recommendations to Business

**Recommendation 1: Introduce and Increase Locally Produced Items**

The items that had maximum sales was mostly freshly produced local items like Pazham Nenthra(Kerala Banana), Egg, Jaya Rice, Milma milk. Being in a village area, the demand for locally procured items are high. Increasing the stock of this kind of items will increase the sales. But all the local produced items need not have the same market value because the vegetables and fruits that are grown naturally in every household will be of very less demand, like Pazham Poovan, Cheru Pazham, Pacha Mango, unlike Pazham Nenthra, which are normally cultivated in the household. So, the items should be selected such that they are locally available, but not naturally grown in the households.

**Recommendation 2: Attach a Small Pantry in Evenings**

The bakery items which are the local snacks sold in numbers contribute the highest to the revenue generated for bakery category like puffs, cutlets, rolls, etc. These items are currently procured from outside. Starting a new section where these are freshly made there itself can increase the demand for these items. Many shops are attaching a small pantry alongside in the evening to attract crowd. This not only increase the sales of bakery items but also the whole sales of store. If tea and coffee is served along with snacks, that will be highly crowd pulling as evening tea or coffee is consumed by a huge majority in the village irrespective of the age.

**Recommendation 3: Free Home Delivery Over a Certain Amount**

Introduce a free delivery system for purchase above a certain value. Keep the delivery charge minimal like ₹10 or ₹20. But a natural tendency will be there to place orders to avoid this delivery fee. This is a common consumer behavior known as the free shipping threshold effect. Platforms like Zepto, BigBasket, and Swiggy Instamart strategically use this concept to increase order sizes. The average invoice bill value over the selected 7 months is ₹270.33. Setting the free threshold slightly above this value incentivizes customers to add more to their cart to save on delivery. For example, someone with a cart value of ₹250 might add ₹50 worth of items to unlock free delivery at ₹300. As the order size increases, the delivery cost becomes a smaller percentage of the total revenue, making it more economical for the store. Free delivery creates a positive experience, increasing the likelihood of repeat orders which increase customer retention.

**Recommendation 4: Cooldrinks in Summer**

The sales during the summer season are less because of the scorching sun and heat which discourages customers from stepping out to shop. As the day progresses, the heat increase and customers will not be shopping. To avoid that time and revenue loss, start the shop earlier than other months and close late at night encouraging night shopping. The shop can leave closed during the peak hot hours so that the staff can rest to make up these extra extended hours. Introduce more drinks during this period, focus more on locally procured items like tender coconut, tender palm fruit, buttermilk, sipups, etc. Focus to arrange drinks from local groups like Kudumbasree or Swayam Seva Sangham, women empowerment groups in Kerala to encourage the small units and attract the local people in more sales. Homely items always do have higher demand.

**Recommendation 5: Add Seasonal Goods to Cater Festival Sales**

The higher sales during April and August due to the festivals, Vishu and Onam can still be increased by introducing festival special items around that time. These cultural festivals have strong traditions of purchasing specific items for celebrations, rituals, and gifting. Introducing additional seasonal goods tailored to these festivals can further maximize sales during these periods. Include traditional items like Athapoo (flower garlands for Pookalam), Thrikkakarappan idols, Valkannadi (traditional mirrors), and decorative lamps that are highly sought-after during Vishu and Onam. Stock up on essential ingredients for festival recipes such as banana chips, jaggery, coconut products, pickles, payasam mixes, and spices. Introduce pre-orders for items like fresh flowers, Onakodi (new clothes), and specialty ingredients, ensuring customers can secure high-demand goods without last-minute hassles.

**Recommendation 6: Pay Later for Regular Customers**

Being in a village side with more known people as customers, a Pay Later option can be given to known regular and consistent customers. This is a very popular option given to customers by bigshots like Amazon and Flipkart. Introducing this option offers several advantages to customers, especially those with fixed incomes or limited immediate cash flow. Salaried customers often face cash shortages toward the end of the month. The Pay Later option helps bridge this gap, enabling them to manage essential purchases without stress. And for common man, it offers relief during financial emergencies, allowing access to goods without requiring upfront payments. This way more trust can be build and increase customer retention.

**Recommendation 7: Introduce Daily Inventory Management**

Currently, the store has monthly record of inventory data. They were not willing to share the data as they were stating it is not recorded accurately. The store is incapable of managing the stock inventory well. Introducing a Daily Inventory Record system can optimize the inventory recordings and watch the sales and profit. To initiate this, store can dedicate a specific personnel and time to record inventory details daily. This practice will provide real-time and accurate stock information, helping to minimize stock shortages or overstocking, identify fast-moving and slow-moving products, ensure efficient inventory replenishment and better planning for seasonal demand variations. As seen in the data, the rate of some items keep varying daily especially for vegetables and fruits. So, a daily record of that clearly points the profit generated out of it. Store can utilize the college student community in the locality for this as an internship to them. Store will get the work done at a lower cost and the students will get a good exposure, win-win for both the parties.

**Recommendation 8: Remove or Reduce Least Sold Products**

When the per unit sales of items in each category is inspected (analysis attached in the excel file), there are few items which was sold very less. Kulfi, Doublemint, Milon Semiya Bar, Ariel Mala, Laxmanrekha, Vim Liquid, Sugar Free Sugar, Pickles, Butter milk, Milky mist curd, Maggi, etc. These items were sold less than 3 quantities in these 7 months and clearly state the least demand for them. In this there are items which have low shelf life and are perishable. Removing them from the stocks can reduce the loss caused by expiring. If removed, these products can save space for other products which are moving fast.

**Recommendation 9: Customer Data Collection**

Collecting customer data during billing is a valuable practice that can help understand customer behavior, build relationships, and offer personalized services. However, it should be implemented carefully to ensure privacy and transparency. Analyze the products which are most frequently purchased, preferred brands, and shopping intervals to identify trends and optimize inventory. Use data to create targeted promotions such as discounts on preferred products or festival-specific items. This data can be used to sort customers for Pay Later option as well.