Business Data Management

Mid-Term Submission

GrocerAnalytics:

Grocery Insights Deep Dive

Name: Shivani Gupta

Email: 22f2000957@ds.study.iitm.ac.in

Roll No.: 22f2000957

1) **Executive Summary and Title:**

Title: **"GrocerAnalytics: Grocery Insights Deep Dive"** for a unorganized B2C Grocery Store

This report presents an in-depth analysis of Kamal Store, a premier grocery retailer in Narela, Delhi. Specializing in a diverse range of essential household grocery items, Kamal Store has been catering exclusively to business-to-consumer (B2C) markets since its inception in 2014.

The *objective* of this study is to analyze sales trends and optimize inventory management strategies over a two-month period from April 1st to May 31st, 2024. The data utilized in this project was sourced from Kamal Store's informal book-keeping records and sales bills, ensuring authenticity and comprehensiveness.

The *analysis* was conducted using descriptive statistical techniques and data visualization tools. The data was meticulously cleaned, organized, and analyzed using Excel sheets. Various *statistical measures* such as mean, median, mode, and standard deviation were calculated to summarize the data. *Column, bar, line, and pie charts* were created to represent revenue, profit margins, purchase prices, and their respective trends and proportions. Additionally, *Pareto analysis* was employed to identify the top 20% of SKUs that contribute to 80% of the total revenue and profit.

Key findings from the analysis reveal that Chakki Atta, Cooking Oil, and Ghee are the highest revenue-generating items, with a notable upward trend in daily revenue. Cooking Oil, Moong Dal, Urad Dal, and Chole exhibit high-profit margins, while Arhar Dal and Ghee have lower margins. Price stability was observed for items like Rajma, Elaichi and Moong Dal, whereas significant price variations were noted for Arhar Dal and Makhana. The analysis also identified significant markups in items like Ghee, Cashew, and Makhana, indicating effective pricing strategies.

To enhance Kamal Store's performance and profitability, several *strategic recommendations* are proposed. Focusing on top-performing SKUs like Chakki Atta, Cooking Oil, and Ghee can maximize revenue. Improving profit margins by optimizing costs for low-margin items like Arhar Dal and Ghee is essential. Effective inventory management, particularly for items with high price fluctuations such as Arhar Dal and Makhana, is crucial. Furthermore, targeted

promotions for high-margin items can boost sales and overall profitability.

In *conclusion*, this report outlines a data-driven approach to addressing the challenges faced by Kamal Store. Through meticulous analysis and strategic recommendations, this project aims to enhance profitability and streamline inventory management, enabling Kamal Store to thrive in the competitive grocery retail market.

2) Proof of originality of the Data:

- **Details:**

  Name: KAMAL STORE

  Owner: Mr. Kamal Gupta

  Address: Gautam Colony, Safiabad Road, Narela, Delhi-110040

- **About:**

  Kamal Store is a premier retailer exclusively serving B2C clientele with a diverse range of essential household grocery items. Specializing in products such as dal, rice, besan, sugar, nuts, elaichi, masala, and more, this establishment has earned a reputation for offering high-quality goods. Since its inception in early 2014, Kamal Store has become the preferred choice for local residents and individual customers seeking reliable products at competitive prices. With a steadfast commitment to customer satisfaction and product excellence, Kamal Store continues to uphold its reputation as a trusted supplier in the grocery sector.

- **Video:** Due to absence of owner, i create video with owner's son. They both handle this business as discussed in video. Link-1: click here to open
- **Letter from Organization (duly stamped & signed):** Link-2: click here to open
- **Images:**

  To support my statements, I've included visual documentation. This includes pictures of the shop itself, a photo of myself with the owner of the organization, and sample bills and book-keeping records used by the shop to keep track of each sale carefully.

  - Pic for reference of Business Name and Location: Link-3: click here to open
  - Main shop area where all the transactions are done: Link-4: click here to open
  - Pic of Shop: Link-5: click here to open
  - Shop storage area: Link-6: click here to open
  - Pic with the Owner: Link-7: click here to open

Since KAMAL STORE primarily operates at a local level, it only uses informal data storage that makes it impossible to present or obtain all records. Consequently, in order to capture a glimpse of the shop's sales records, some snapshots of the informal bills have been presented. While the absence of official data storage may limit the availability of documentation, these captured snapshots provide valuable insights into the shop's transactional activities and contribute to the overall understanding of its operational practices.

- Pic of Sales record (Informal): Link-8: click here to open

3) Metadata and Descriptive Statistics:

I diligently gathered an extensive dataset covering the period from April 1st to May 31st, 2024. Throughout these two months, I collected data using a combination of previously issued bills and book-keeping records stored in the bill book. This approach ensured that I compiled a comprehensive dataset for the entire timeframe, which is essential for analyzing and gaining insights for my project.

Metadata:

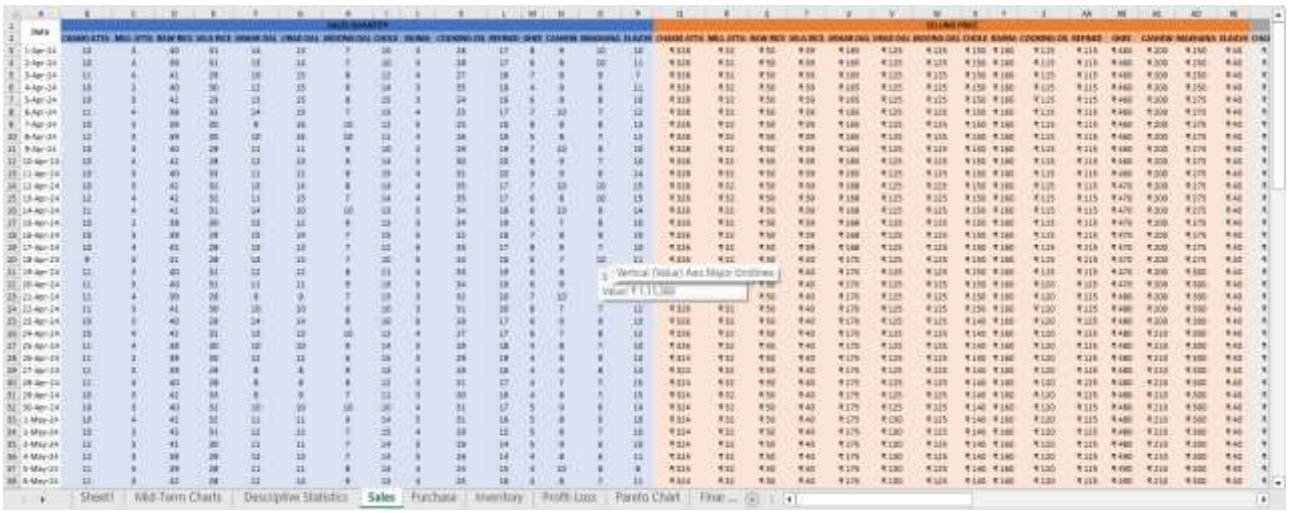

Figure 1: Snapshot of Data

Column Name	Description
Date	Date of the transaction in the format “DD-Month-YY”. This column indicates the specific day when the transaction occurred.
Item Name	Name or description of the item being sold or purchased in the transaction. This helps in identifying the specific product involved in each transaction.

Quantity	Quantity of the item sold or purchased, measured in kg, ltr & packets. This column provides the amount of the product involved in each transaction.
Rate	Price per unit or rate of the item sold or purchased. This column indicates the cost or selling price per unit of the product.
Amount	Total amount for the item sold or purchased, calculated by multiplying quantity by rate. This provides the monetary value of each transaction.
Transaction Type ✓ Sales Sheet ✓ Purchase Sheet ✓ Inventory Sheet	This helps in distinguishing b/w different types of transactions through the organization of data across different sheets. • Contains sales transactions • Contains purchase transactions • Contains inventory details

Table 1: Metadata

Descriptive Statistics:

BASIS	SALES	PURCHASE	INVENTORY
Mean	₹ 25,846.08	₹ 30,781.07	₹ 2,564.93
Standard Error	₹ 170.00	₹ 3,746.26	₹ 109.72
Median	₹ 25,822.00	₹ 27,000.00	₹ 2,337.00
Mode	₹ 24,897.00	₹ -	₹ -
Standard Deviation	₹ 1,327.74	₹ 29,259.22	₹ 856.97
Sample Variance	₹ 17,62,882.64	₹ 85,61,02,130.10	₹ 7,34,392.30
Kurtosis	-0.31	17.71	-0.65
Skewness	0.42	3.80	-0.05
Range	₹ 5,575.00	₹ 1,77,505.00	₹ 3,239.00
Minimum	₹ 23,443.00	₹ -	₹ 702.00
Maximum	₹ 29,018.00	₹ 1,77,505.00	₹ 3,941.00
Sum	₹ 15,76,611.00	₹ 18,77,645.00	₹ 1,56,461.00
Count	61	61	61

Table 2: Descriptive Statistics

➢ Sales:

- Mean: ₹25,846.08 - This is the average sales amount. =AVERAGE(range)
- Standard Deviation: ₹1,327.74 - Indicates the variability in sales; lower variability means more consistent sales. =STDEV.S(range)

- Median: ₹25,822.00 - The middle value, which is very close to the mean, suggests a relatively symmetrical distribution. =MEDIAN(range)
- Mode: ₹24,897.00 - The most frequently occurring sales value. =MODE.SNGL(range)
- Kurtosis: -0.31 - Indicates a slightly flatter distribution than a normal distribution. =KURT(range)
- Skewness: 0.42 - Suggests a slight right skew, meaning there are some higher sales values. =SKEW(range)
- Range: ₹5,575.00 - Difference between the highest and lowest sales values. =MAX(range)-MIN(range)
- Minimum: ₹23,443.00 - The lowest sales value. =MIN(range)
- Maximum: ₹29,018.00 - The highest sales value. =MAX(range)
- Sum: ₹15,76,611.00 - Total sales over the period. =SUM(range)
- Count: 61 - Number of sales observations. =COUNT(range)

✓ In above formulaes, "***range***" refers to the column of data to which descriptive statistics are applied.

✓ The formulas are applied to Sales, Purchase, and Inventory data by substituting the respective column references.

➢ <u>Purchase</u>:

- Mean: ₹16,977.54 - The average purchase amount.
- Standard Deviation: ₹15,530.42 - Indicates high variability in purchases, suggesting inconsistency.
- Median: ₹17,400.00 - The middle value, indicating a distribution with some large deviations.
- Kurtosis: -1.62 - Indicates a significantly flatter distribution than normal.
- Skewness: 0.07 - Nearly symmetrical distribution.
- Range: ₹40,875.00 - Wide range indicating significant differences between the highest and lowest purchase values.
- Minimum: ₹0.00 - The lowest purchase value.
- Maximum: ₹40,875.00 - The highest purchase value.
- Sum: ₹10,35,630.00 - Total purchases over the period.
- Count: 61 - Number of purchase observations.

➢ <u>Inventory</u>:

- Mean: ₹2,564.93 - The average inventory amount.

- Standard Deviation: ₹856.97 - Indicates variability in inventory levels.
- Median: ₹2,337.00 - The middle value, slightly below the mean.
- Kurtosis: -0.65 - Slightly flatter distribution than normal.
- Skewness: -0.05 - Almost symmetrical distribution.
- Range: ₹3,239.00 - Difference between the highest and lowest inventory values.
- Minimum: ₹702.00 - The lowest inventory value.
- Maximum: ₹3,941.00 - The highest inventory value.
- Sum: ₹1,56,461.00 - Total inventory over the period.
- Count: 61 - Number of inventory observations.

❖ General Insights:

- ***Sales*** are relatively consistent, with a slight right skew and a small range compared to the mean.
- ***Purchases*** show high variability with a wide range and flatter distribution, indicating occasional large purchases.
- ***Inventory levels*** are quite stable with low variability and almost symmetrical distribution.

The variability and distribution in purchases suggest that while sales and inventory are relatively stable, purchasing decisions may involve larger fluctuations, potentially due to bulk buying or varying supply needs.

4) Detailed Explanation of Analysis Process/Method:

Over a period of two months, i meticulously collected, cleaned, organized, and analyzed data using Excel Sheets.

i) Data Cleaning and Organization:

- I carefully examined the data to ensure its structure and format were clean and organized for effective analysis.
- This step involved checking for inconsistencies, missing values, and outliers that could impact the results.

ii) Descriptive Statistical Analysis:

- I applied various descriptive statistical techniques to summarize and explore the data.
- Measures such as mean, average, profit, and inventory stocks were calculated to gain insights into the overall dynamics of the shop.

iii) Data Visualization:

- I used column, bar, line, and pie charts to visually represent revenue, profit of SKUs, purchase price, and their respective proportions.
- These charts provide a clear and concise way to understand the performance and distribution of various metrics within the business.
- They help identify trends, patterns, and areas of focus for decision-making.

iv) Pareto Chart for Revenue or Profit:

- I used a combined line and bar chart to show the Pareto Chart of revenue or profit for all SKUs.
- The bar chart component displays individual revenue or profit values for each SKU, sorted in descending order from highest to lowest.
- The line chart component overlays the cumulative percentage of revenue or profit on the same graph.
- This combination allows for a clear and effective representation of the cumulative contribution of each SKU to the overall revenue or profit, highlighting the significant contributors.

v) Pie Chart for Profit Proportions:

- I utilized a pie chart to display the proportions of profit for different SKUs.
- The pie chart effectively represents the relative contribution of each SKU to the total profit by dividing a circular graph into sectors or slices.
- Each slice corresponds to a specific SKU, and its size is proportional to the profit generated by that SKU, with larger slices indicating greater profit contributions.

vi) Line Chart for Trends:

- I used a line chart to display trends in revenue and purchasing price over time.
- This chart provides a visual representation of how these variables change with respect to time, helping to identify trends and inform strategic decisions.

By utilizing these descriptive statistics and visualizations, I gained a comprehensive understanding of the data and its implications for the business. This analysis will enable us to make informed decisions and optimize business strategies for improved performance and profitability.

5) **Results and Findings**:

Here is the visualization and data analysis conducted based on the collected data:

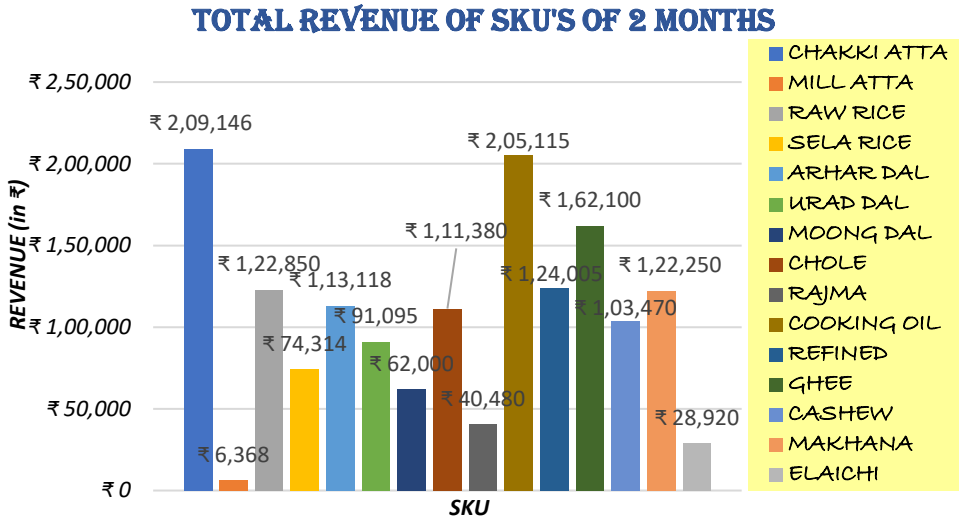

Chart 1: Total Revenue of all the SKU's over 2 months

- Top Revenue Generating Items: Chakki Atta, Cooking Oil, Ghee, Moong Dal, Raw Rice and Makhana are the highest revenue generators.
- Revenue Contributions: Chakki Atta contributed the most with ₹209,146, followed by Cooking Oil with ₹205,115, and Ghee with ₹162,100.

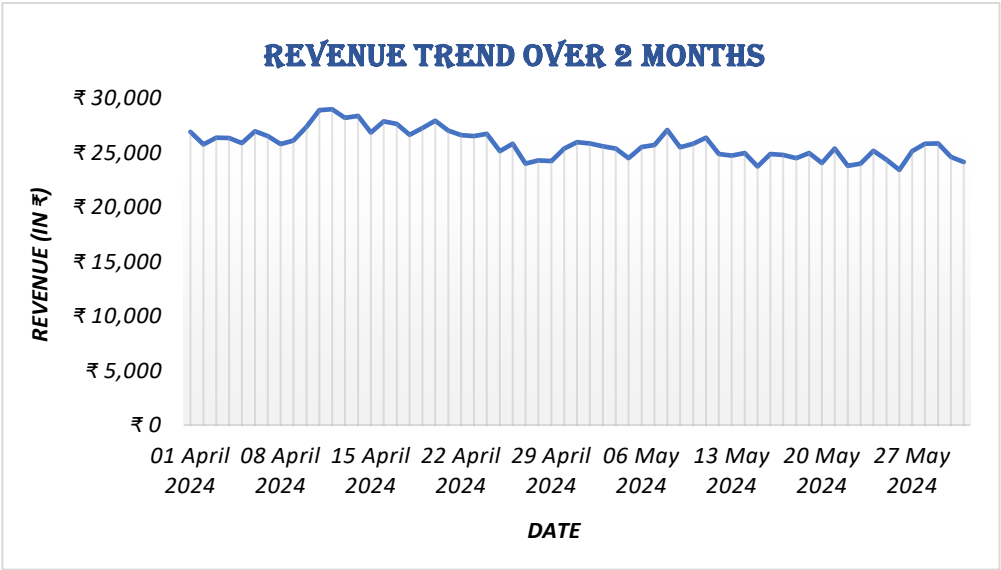

Chart 2: Total Revenue Trend observed over 2 months

- Daily Revenue Variation: The total daily revenue shows fluctuations, with a general upward trend towards the end of the period.
- Highest Revenue Day: The peak revenue day recorded a total of ₹29,018.

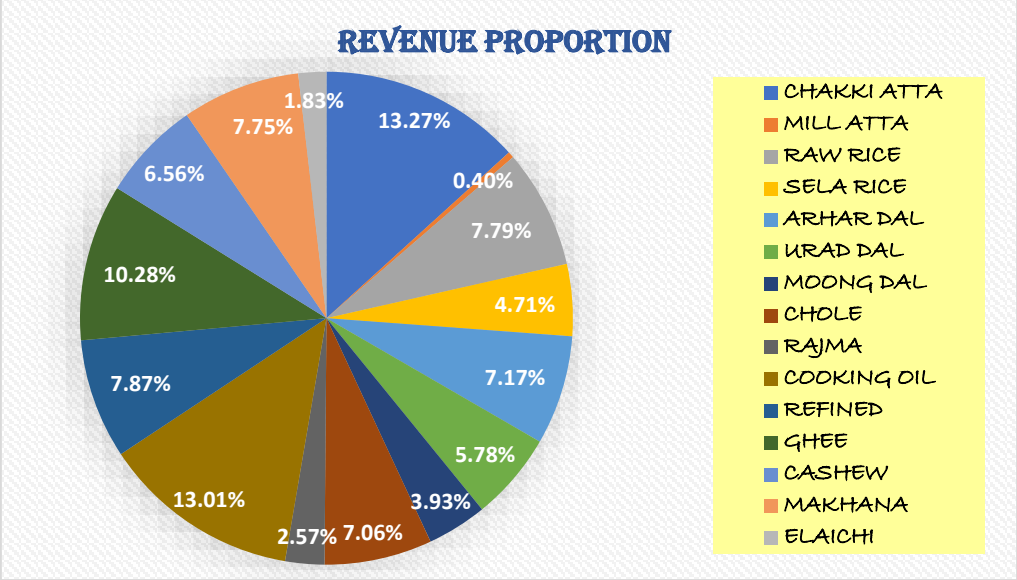

Chart 3: SKU's proportion to Total Revenue

- Major Contributors: Chakki Atta, Cooking Oil, and Ghee together contribute a significant portion of the total revenue.
- Least Contribution: Mill Atta contributes the least to the total revenue.

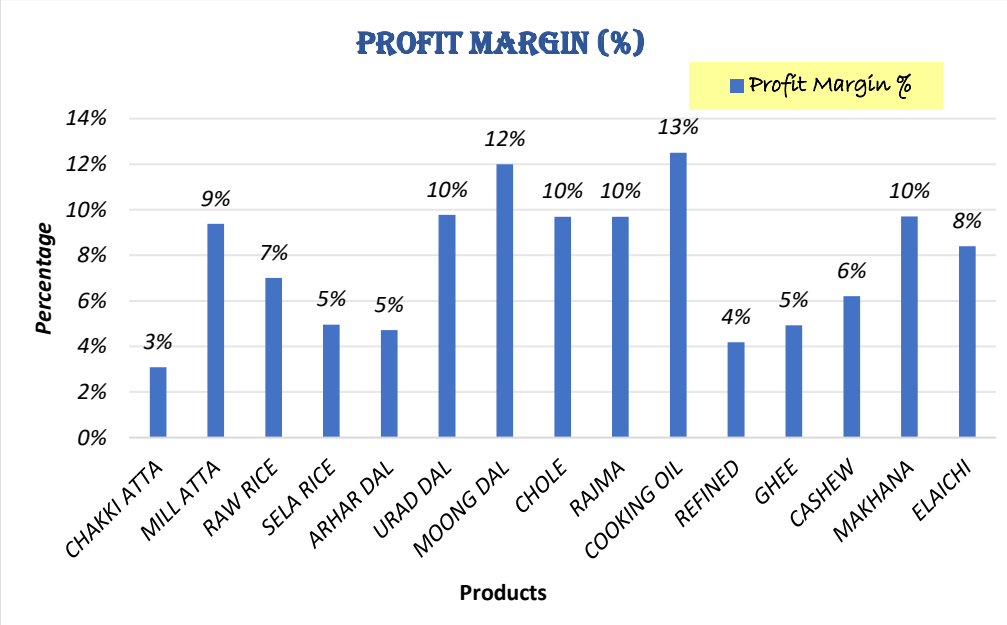

Chart 4: SKU's Profit Margin percentage

- High Margin Items: Cooking Oil, Moong Dal, Urad Dal, Chole, Rajma and Makhana have high-profit margins.
- Low Margin Items: Chakki Atta, Refined, Sela Rice, Arhar Dal and Ghee have the lowest profit margins.

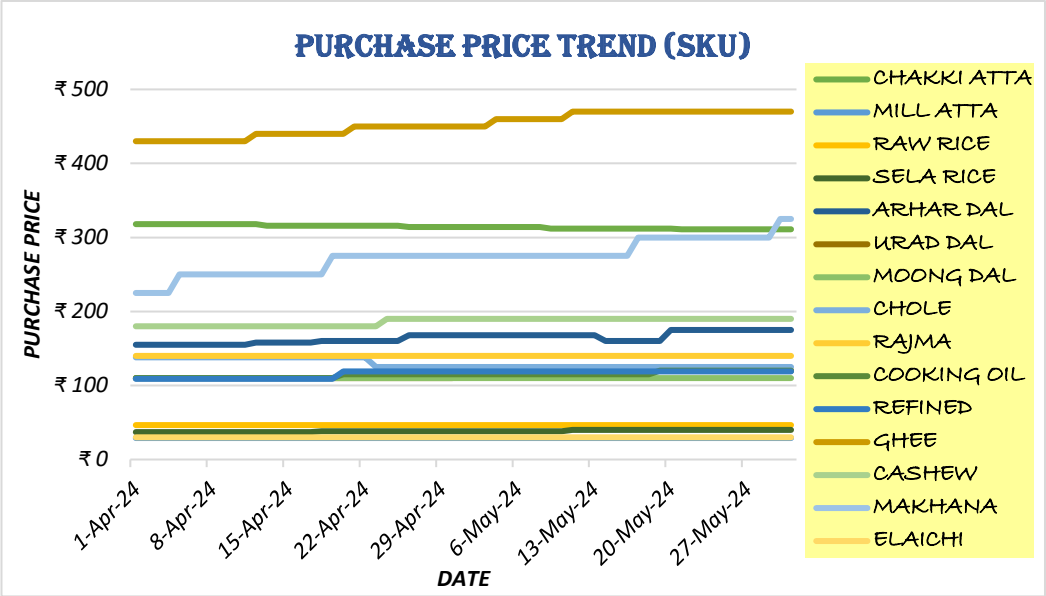

Chart 5: Purchase Price of all the SKU's over a period of 2 months (for shop owner)

- Price Stability: Items like Rajma, Elaichi, Moong Dal and Raw Rice showed stable prices.
- Price Variations: Items such as Makhana, Urad Dal and Arhar Dal showed significant price fluctuations.

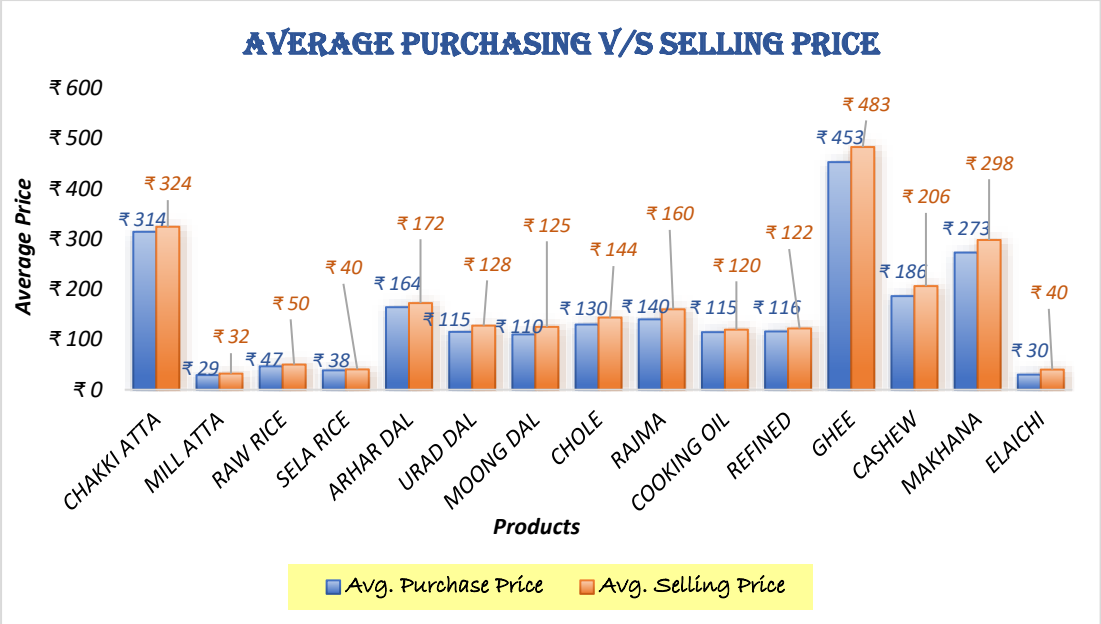

Chart 6: Avg. Purchase Price vs Avg. Selling Price

- Markups: Significant markups are seen in items like Ghee, Cashew, Makhana and Rajma.
- Competitive Pricing: Minimal price differences are observed in commodities like Mill Atta, Raw Rice, Sela Rice and Cooking Oil.

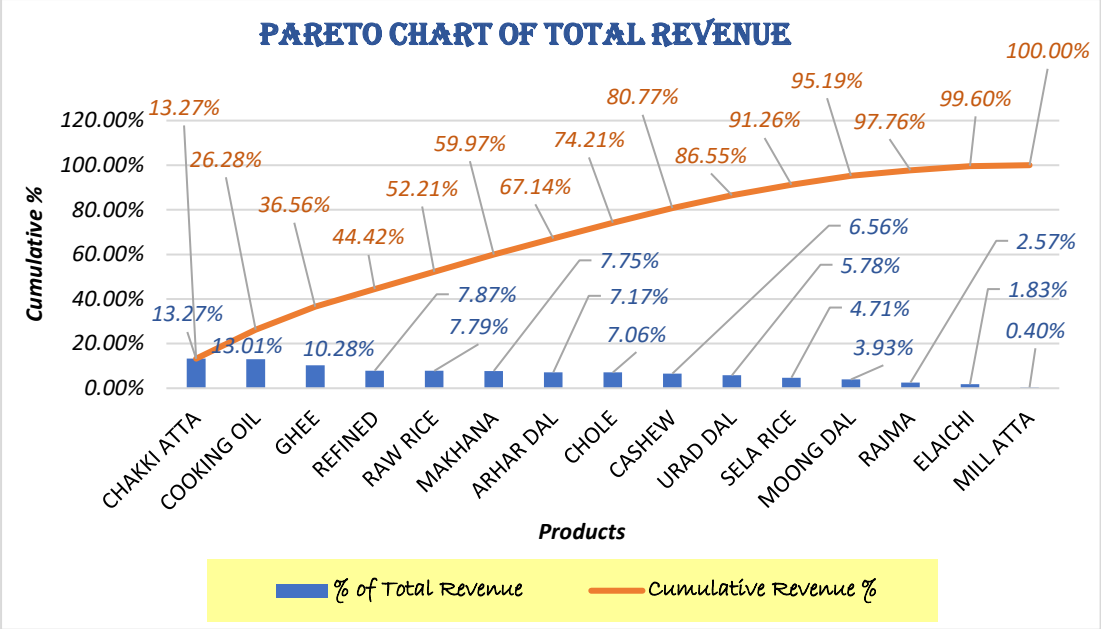

Chart 7: Pareto Chart of Total Revenue

- Key Revenue Drivers: The 9 SKUs (Chakki Atta, Cooking Oil, Ghee, Refined, Raw Rice, Makhana, Arhar Dal, Chole, Cashew) contribute to approximately 80% of the total revenue.

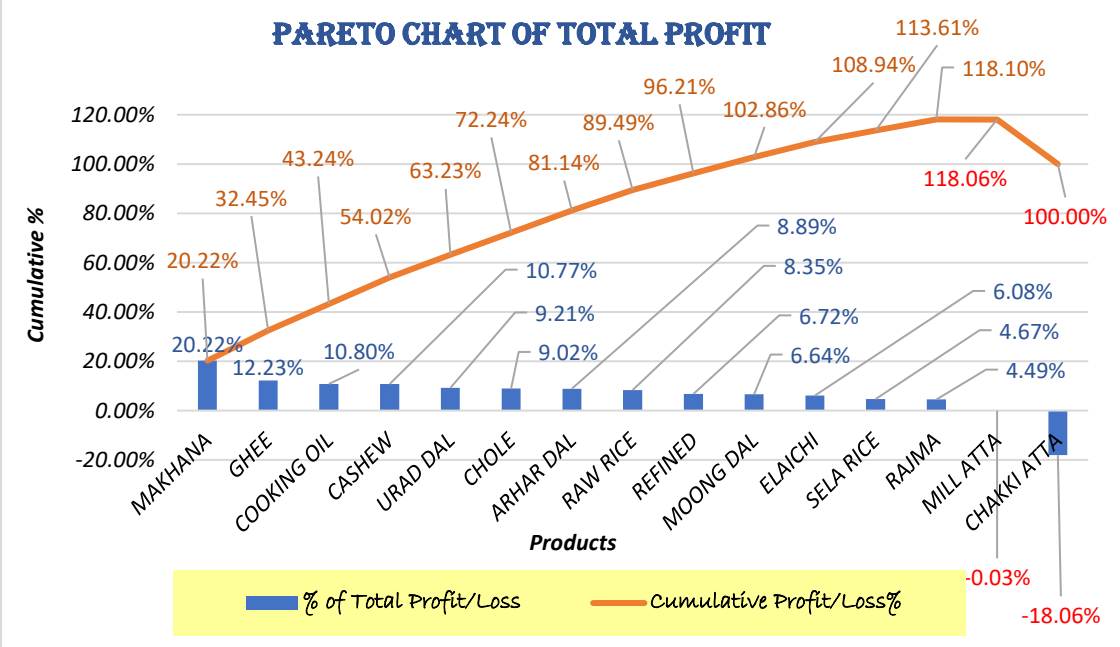

Chart 8: Pareto Chart of Total Profit

- Key Profit Drivers: Items such as Makhana, Ghee, Cooking Oil, Cashew, Urad Dal, Chole, Arhar Dal are significant contributors to the total profit.