Business Data Management

Capstone Project Final

Data-Driven Demand and Inventory Planning for a Rural Grocery Shop

Submitted by

Name: Sk Noor Alam

Roll No.: 23f1000205

IITM Online BS Degree Program,

Indian Institute of Technology, Madras,

Chennai Tamil Nadu, India, 600036

Contents

1. Executive Summary:

Mallick Bhandar is a family-owned B2C grocery store located in Khoriop, Howrah, West Bengal. While serving its local community for years, the shop faced two major business challenges: lack of visibility into seasonal demand patterns, particularly during local festivals (like Durga Puja, Eid, and Kali Puja), and inefficient inventory planning resulting in either overstocking or stockouts. These issues led to poor space utilization, unsold stock accumulation, and lost sales opportunities.

To investigate and address these problems, one year of primary data was collected by digitizing handwritten bills, covering 18 essential product categories including rice, wheat flour, tea, oil, snacks, and toiletries. Descriptive statistics such as mean, median, standard deviation, skewness, and kurtosis were computed. Rice (Local), Pulses have shown high and stable demand, where Rice (Premium), Spices and Snacks have shown festival sensitivity. To dive into deeper, analysis like ABC analysis, correlational analysis and festival demand comparison have done.

The analysis revealed key sales patterns across both time and category dimensions. Rolling average trends highlighted consistent spikes in demand before major festivals, especially for essentials like Rice, Pulses, Egg, and Potatoes & Onions. Festival demand comparison confirmed that items such as Cooking Oils, Spices, and Snacks experienced a significant uplift of 40–90% during event weeks. ABC analysis identified that staples are dominating the sales, highlighting inventory priorities. Correlation mapping uncovered natural product groupings (e.g., Snacks with Soft Drinks), while the Demand Variation Index classified products based on demand stability, guiding risk-adjusted restocking strategies.

The insights derived from the data provided clear direction for action. Festival-driven stocking calendars, rolling average-based forecasting, and ABC-based inventory prioritization were recommended to shift operations from reactive to proactive. High-DVI items like Spices and Coconut Oil were suggested for buffer stocking, while stable-demand items were ideal for fixed-cycle replenishment. Correlated product pairs can be bundled or co-placed to optimize shelf layout and improve cross-selling. Collectively, these interventions aim to reduce stockouts, minimize overstock, and improve customer

satisfaction through consistent availability of essential products, especially during seasonal peaks.

2. Detailed Explanation of Analysis Process/Method:

2.1 Data Collection

The data used in this project was collected primarily from Mallick Bhandar, a rural grocery shop located in Khoriop, Howrah, West Bengal. The store maintained daily bills of sales which were digitized manually in Excel. The dataset contains 365 rows (one in each day from June 2024 to May 2025) and 19 columns, representing date and quantity sold for 18 essential grocery products.

- **Justification:** This daily-granular dataset allowed for the analysis of seasonality, volatility, and festival-driven patterns, which are central to the inventory and demand planning challenges the store faces.

2.2 Data Cleaning & Preprocessing

The data preprocessing steps involved in the following steps:

- **Handling of Zeros**: A value of '0' (zero) was retained to indicate days with no sales for a product, distinguishing between true zero demand and missing data.
- **Outlier Detection**: Used box plots to identify extreme values. Outliers were cross-verified with original bills or confirmed with the store representative. If invalid, they were corrected using median imputation from nearby days.
- **Data Type Formatting:** All 18 products' quantity converted into integers to avoid floating point-precision errors and generalization of analysis. While ensure that the "Date" Column is in datetime object.
- **Justification:**

  Clean and consistent data is critical for time series and statistical analysis, and mis-entries can lead to misleading insights. Preprocessing ensured high-quality inputs for reliable findings.

2.3 Trend Analysis

2.3.1 Purpose

The demand trend has captured through rolling average and the primary objective of conducting a rolling average analysis was to uncover subtle, underlying demand trends that are often masked by daily fluctuations in sales data. This approach is especially useful for identifying seasonal uplift ahead of major festivals and for distinguishing between normal volatility and meaningful patterns. By smoothing the data, it becomes easier to detect gradual increases or decreases in product demand that can inform strategic stocking decisions and marketing initiatives.

2.3.2 Implementation

A 30-day rolling average was computed for the top 10 high-performing products based on annual sales volume. This smoothing technique helped reveal long-term consumption patterns. For example:

- Rice (Local) and Pulses show upward spikes around mid-October and March—aligning with Durga Puja and Shab-e-Barat.

These trends would have been obscured in a raw daily sales view but were clearly visualized through rolling averages, making the data more actionable and easier to interpret.

2.3.3 Link to Problem

This analysis directly addresses Problem 1: Lack of Visibility into Seasonal Patterns. By leveraging rolling averages, the business can now clearly observe recurring festival-related demand surges and seasonal cycles, which were previously difficult to quantify. This insight enables proactive planning of inventory, promotions, and supplier coordination, reducing the risk of stockouts or overstocking during critical periods.

2.4 Demand Stability Analysis

2.4.1 Purpose

The Demand Variation Index (DVI) was introduced to quantitatively measure the demand stability or volatility of individual SKUs over time. While average sales

values provide insight into product popularity, they fail to capture how consistent or erratic the demand is. The DVI enables direct comparison of demand volatility across products, helping the business:

- Identify which products have reliable, steady demand.
- Flag those that exhibit erratic or seasonal buying behaviour.
- Make informed decisions about inventory risk management, procurement strategy, and shelf placement.

2.4.2 Implementation

Using the formula: $\text{DVI} = \left(\frac{\sigma}{\mu}\right) X\ 100$

Where:

- σ = Standard deviation of daily sales (a measure of spread or variability)
- μ = Mean daily sales (average demand)

This metric reflects how much a product's demand fluctuates relative to its average. A lower DVI indicates stable, predictable demand, while a higher DVI suggests irregular or seasonal consumption patterns.

Products like Spices (DVI ≈ 62%), and Snacks (DVI ≈ 39%) showed high variability—demand was less predictable. In contrast, Rice (Local) and Wheat Flour had DVI around 30%, indicating stable demand.

2.8.3 Link to Problem

This analysis helps address the broader challenge of inventory uncertainty and inefficient shelf allocation.

2.5 Festival Demand Analysis

2.5.1 Purpose

The purpose of the Festival Demand Analysis was to understand the influence of culturally significant events on consumer purchasing behaviour. While rolling averages and DVI metrics provide general trends and volatility levels, they may miss short-term spikes tied to specific festival periods.

This analysis aims to:

- Isolate and quantify temporary demand surges that occur during regional festivals

- Reveal hidden or short-duration seasonal patterns not visible in broader statistical trends

Help anticipate SKU-specific festival uplift, enabling more precise inventory, marketing, and logistics planning during high-demand windows.
Understanding these demand peaks is essential for businesses operating in culturally diverse markets, where religious and regional events significantly impact consumption patterns.

2.5.2 Implementation

Dates around Eid, Durga Puja, Kali Puja, and Shab-e-Barat were tagged. Average sales during these periods were compared against off-season data. For example, Egg and Spices exhibited over 1.5x spike in festival weeks.

2.5.3 Link to Problem

Addresses Problem 1 (Lack of Visibility into Seasonal Patterns) by quantifying sales uplift due to festivals, validating the need for seasonal forecasting.

2.6 ABC Analysis (Volume-Based Pareto Inventory Classification)

2.4.1 Purpose

The purpose of conducting an ABC Analysis was to segment inventory based on the relative contribution of each product to total annual sales volume. This classic Pareto-based classification helps distinguish between high-impact and low-impact products, allowing the business targeted inventory control.

2.4.2 Implementation

Each product's annual sales were calculated using total units sold across 365 days. Products were then sorted in descending order and their cumulative percentage contribution to total sales was computed. Based on the Pareto principle:

- *"A" items (top ~70%)* include high-frequency essentials like Rice (Local), Pulses, and Potatoes & Onions.
- *"B" items (next 20%)* include moderately sold products like Tea and Dairy Product.
- *"C" items (bottom 10%)* include infrequently sold goods such as Coconut Oil

and Toiletries.

2.4.3 Link to Problem

Addresses Problem 2 (Data-Driven Inventory Planning) by focusing restocking efforts on critical items to avoid shortages.

2.7 Correlation Analysis

2.7.1 Purpose

To identify complementary products that are frequently purchased together. Understanding these relationships enables the business to:

- Design cross-selling strategies and product bundles.
- Improve shelf layout and in-store navigation.
- Enable coordinate replenishment for interrelated items.

This analysis is especially valuable in identifying complementary goods, where changes in the demand of one product can indicate or even drive changes in the demand of another.

2.7.2 Implementation

A Pearson correlation matrix was computed across all product pairs using their daily sales data over the full time period. The Pearson correlation coefficient (r) measures the linear relationship between two variables, ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation).

A Pearson correlation matrix revealed strong correlations between Rice (Local) and Pulses ($r \approx 0.68$), and between Snacks and Soft Drinks ($r \approx 0.61$). These suggest bundling or co-placement strategies.

2.7.3 Link to Problem

This analysis supports the resolution of Problem 2: Lack of Data-Driven Inventory Management. By highlighting product interdependencies.

3. Results & Findings:

This section presents key findings derived from various quantitative analyzes performed on the sales data of 18 grocery products over 365 days. Each analysis is aligned with the two central business problems: lack of visibility into seasonal demand and absence of

data-driven inventory planning. Insights are supported by numerical indicators and linked back to business implications for Mallick Bhandar.

3.1 Trend Analysis: Uncovering Smoothed Demand

The main focus of this trend analysis is on time-based seasonal spikes and sales cycles.

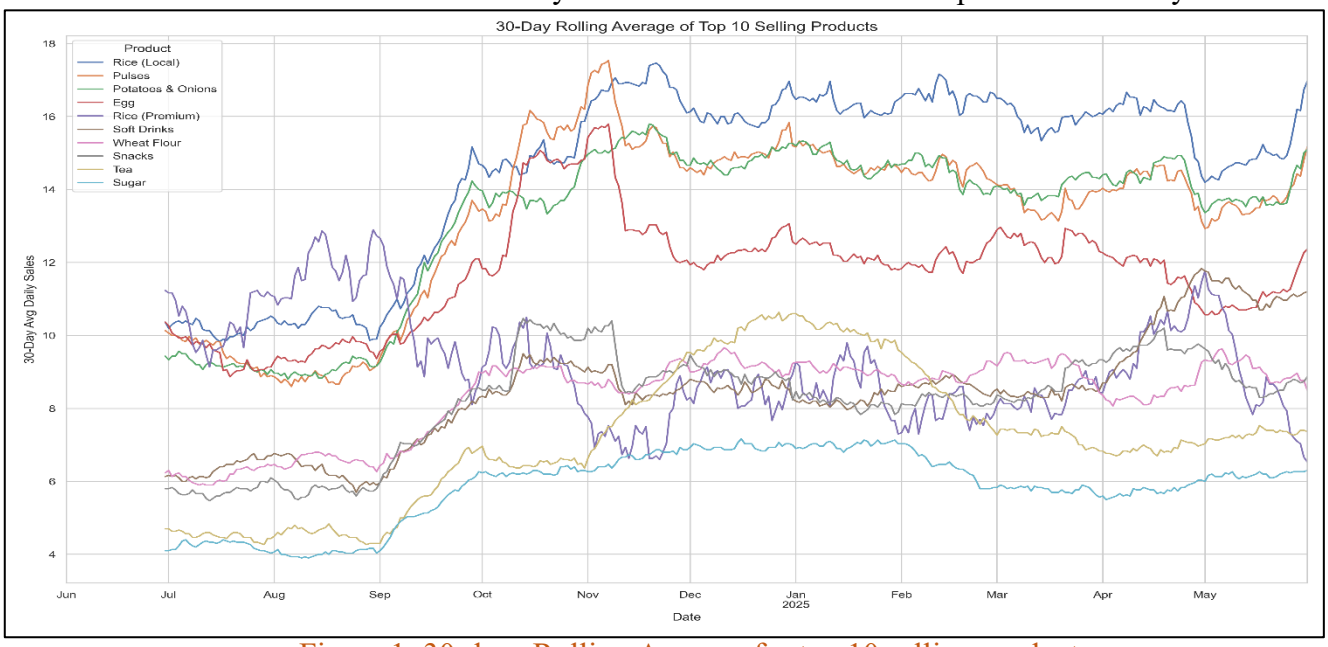

Figure 1: 30-days Rolling Average for top 10 selling products

The above figure of top 10 selling products' rolling average (in Figure 1) clearly state the below facts about these products:

3.1.1 Festival/Seasonal Spikes

- All top 10 products show sharp demand increases between September and November, consistent with festivals.
- Especially prominent in:
  - Pulses
  - Egg
  - Potatoes & Onions
  - Rice (Local)

3.1.2 Post-Festival Normalization

- Most products stabilize post-December, indicating a return to baseline demand.

- Exception: Rice (Premium) and Soft Drinks show erratic behavior even after peak months**.**

3.1.3 Off-Season Growth

- Products like Snacks and Soft Drinks show a rise in late Feb–May, possibly linked to summer or school holidays.

3.1.4 Persistent High Demand

- Rice (Local), Potatoes & Onions, and Pulses maintain consistently high rolling averages throughout the year => staples that require sustained stock planning.

3.2 Demand Stability Analysis: Figuring Consistency in Sales

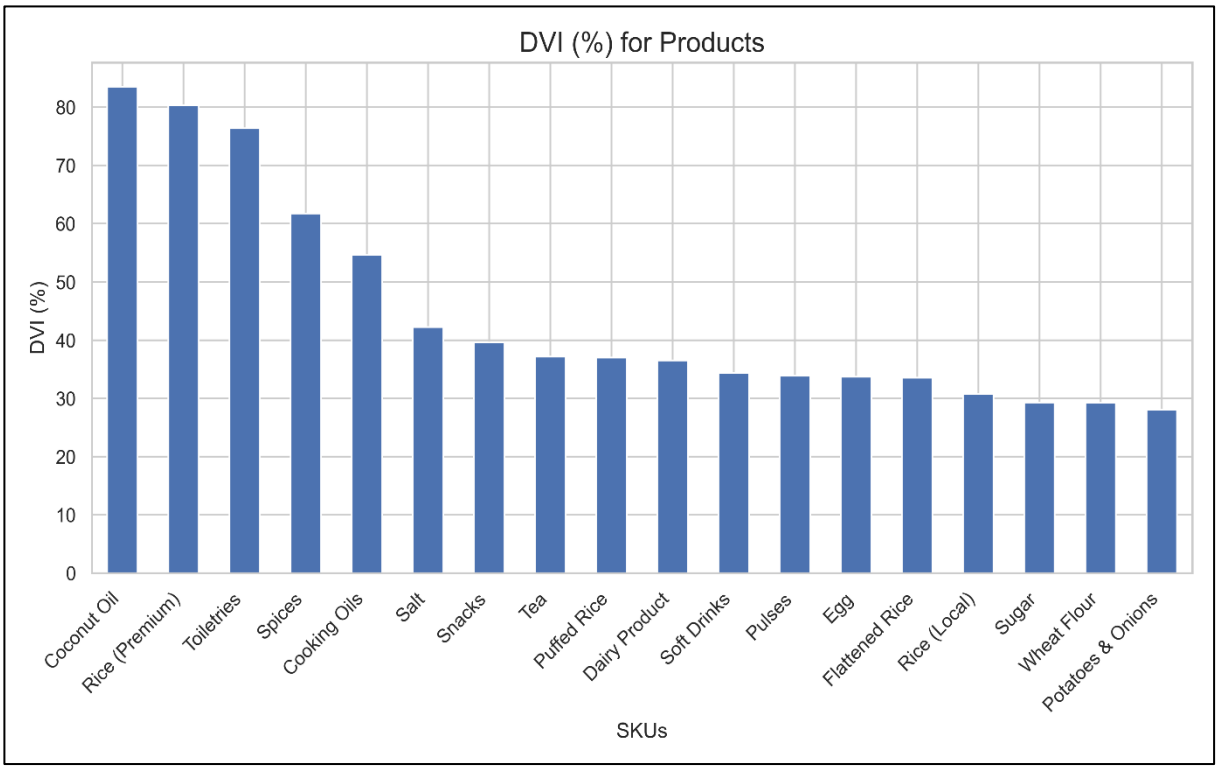

Figure 2: Bar Plot of Percentage DVI

Using DVI, we could track product predictability and volatility.

Based on the bar plot in Figure 2 the all 18 SKUs have shown the below natures**: -**

3.2.1 High-Variability Products (Less-predictable)

- *Top 5:* Coconut Oil, Rice (Premium), Toiletries, Spices, Cooking Oils
- These products exhibit high swings in daily demand, requiring buffer stock and cautious procurement.

- Likely influenced by price sensitivity, holiday promotions, or discretionary purchase patterns.

3.2.2 Low-Variability Products (Stable/Predictable Demand)

- Bottom 5: Sugar, Wheat Flour, Potatoes & Onions, Rice (Local), Flattened Rice
- These are predictable staples, ideal for automated replenishment, low safety stock, and forecast-based ordering.

3.2.3 Strategic Inventory Actions

- High DVI => Flexible supply chain, promotion tracking
- Low DVI => Lean inventory, lower safety stock, regular ordering cycles

3.3 Festival Demand Analysis: Demand Comparison

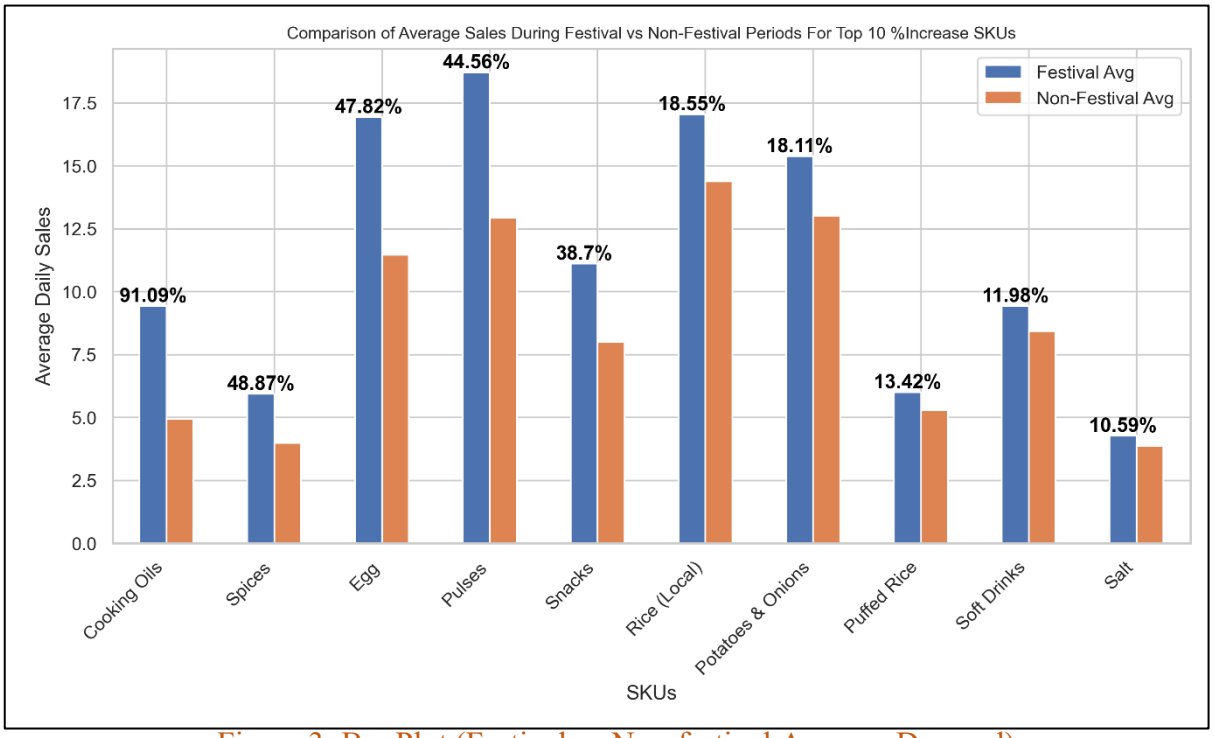

Figure 3: Bar Plot (Festival vs Non-festival Average Demand)

Now the attention is on the seasonal pattern of the products specially in festival seasons. From the above bar plot (Figure 3): -

3.3.1 *Cooking Oils* saw the highest spike in sales during festivals:

- **+91.09%** increase compared to non-festival periods.
- Indicates it's a major ingredient in festival-related cooking.

3.3.2 *Spices and Egg* also had large jumps:

- *Spices:* **+48.87%**
- *Egg:* **+47.82%**
- Suggests increased cooking activity and possibly baking or protein-rich meals.

3.3.3 *Pulses and Snacks* showed strong increases:

- *Pulses:* **+44.56%**
- *Snacks:* **+38.7%**
- Likely due to traditional dishes and increased snacking during celebrations.

3.3.4 *Staples like Rice (Local) and Potatoes & Onions* also rose notably:

- *Rice:* **+18.55%**
- *Potatoes & Onions:* **+18.11%**
- Shows a consistent uplift in basic food items.

3.3.5 Smaller but still notable increases in:

- *Puffed Rice:* **+13.42%**
- *Soft Drinks:* **+11.98%**
- *Salt:* **+10.59%**
- These items are either complements or part of festive meals/snacks.

3.4 ABC Analysis: Product Prioritization

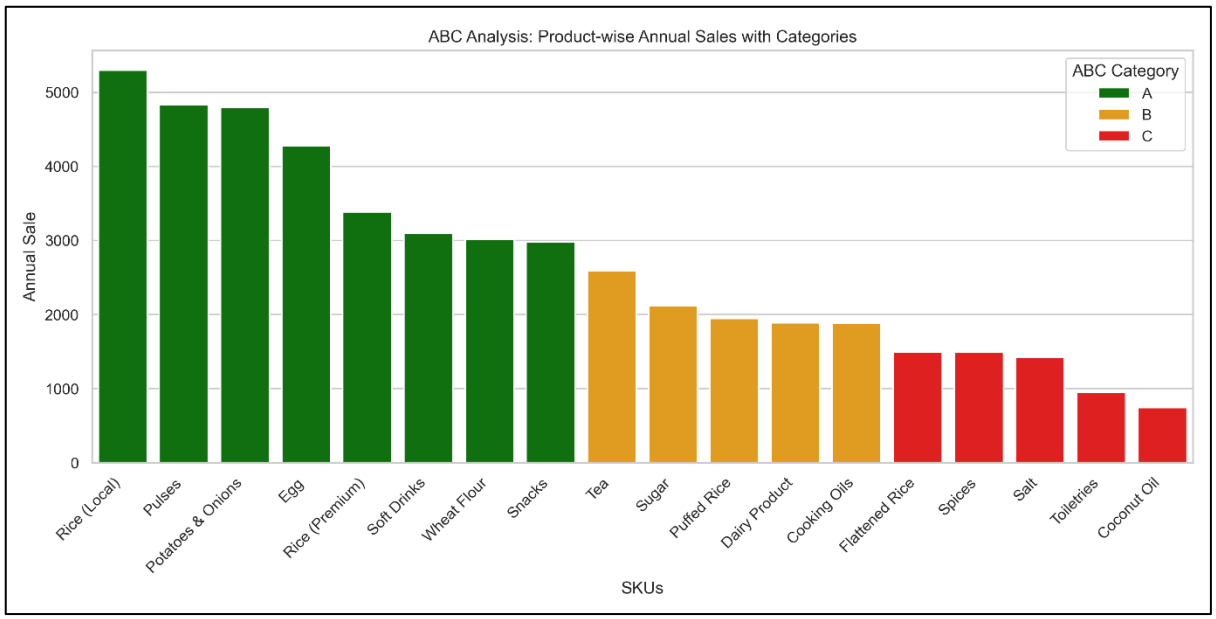

Figure 4: Bar Plot (ABC Classification)

To make inventory planning more data-driven and efficient, the 18 core products sold by Mallick Bhandar have been categorized into three strategic groups—Category A, B, and C—based on their total annual sales volume. This type of volume-based classification helps determine which products need the most attention and tighter control, and which can be managed with simpler strategies.

3.4.1 **Category A** – High-Priority Products (Green Bars in Figure 4)

These are the top-performing 8 products, which collectively account for over 70% of total annual sales.

Products Included:

- Rice (Local), Pulses, Potatoes & Onions, Egg, Rice (Premium), Soft Drinks, Wheat Flour, and Snacks.

Key Takeaways:

- These products are the backbone of overall revenue and are sold in high volumes year-round.
- They require frequent demand forecasting, tight inventory control, and prioritized restocking to avoid stockouts.
- Because of their contribution to sales, automation and real-time monitoring should be considered for these items.

- Ensuring high service levels and availability for these items is essential for maintaining customer satisfaction.

3.4.2 **Category B** – Medium-Priority Products (Orange Bars in Figure 4)

This category includes 5 products that contribute moderately (around 15–20%) to total yearly sales.

Products Included:

- Tea, Sugar, Puffed Rice, Dairy Product, and Cooking Oils.

Key Takeaways:

- These products are important but do not require as intense a focus as Category A items.
- They should be reviewed periodically—perhaps biweekly or monthly—depending on demand cycles.
- Inventory decisions (like safety stock or reorder points) should be based on historical trends and supplier lead times.
- Some products in this category may temporarily shift to Category A during festive seasons, such as Cooking Oils or Tea.

3.4.3 **Category C** – Low-Priority Products (Red Bars in Figure 4)

This group also includes 5 products, which together make up less than 10% of total sales.

Products Included:

- Flattened Rice, Spices, Salt, Toiletries, and Coconut Oil.

Key Takeaways:

- These are low-volume, low-contribution products, even though they occupy shelf space and resources.
- They can be managed using simpler inventory techniques, like periodic stock checks or fixed review intervals.
- May be candidates for bundling with high-demand items, or discontinuation if they don't add strategic value.
- Optimizing stock levels here can reduce holding costs without significantly impacting revenue.

3.5 Correlation Analysis: Product Basketing

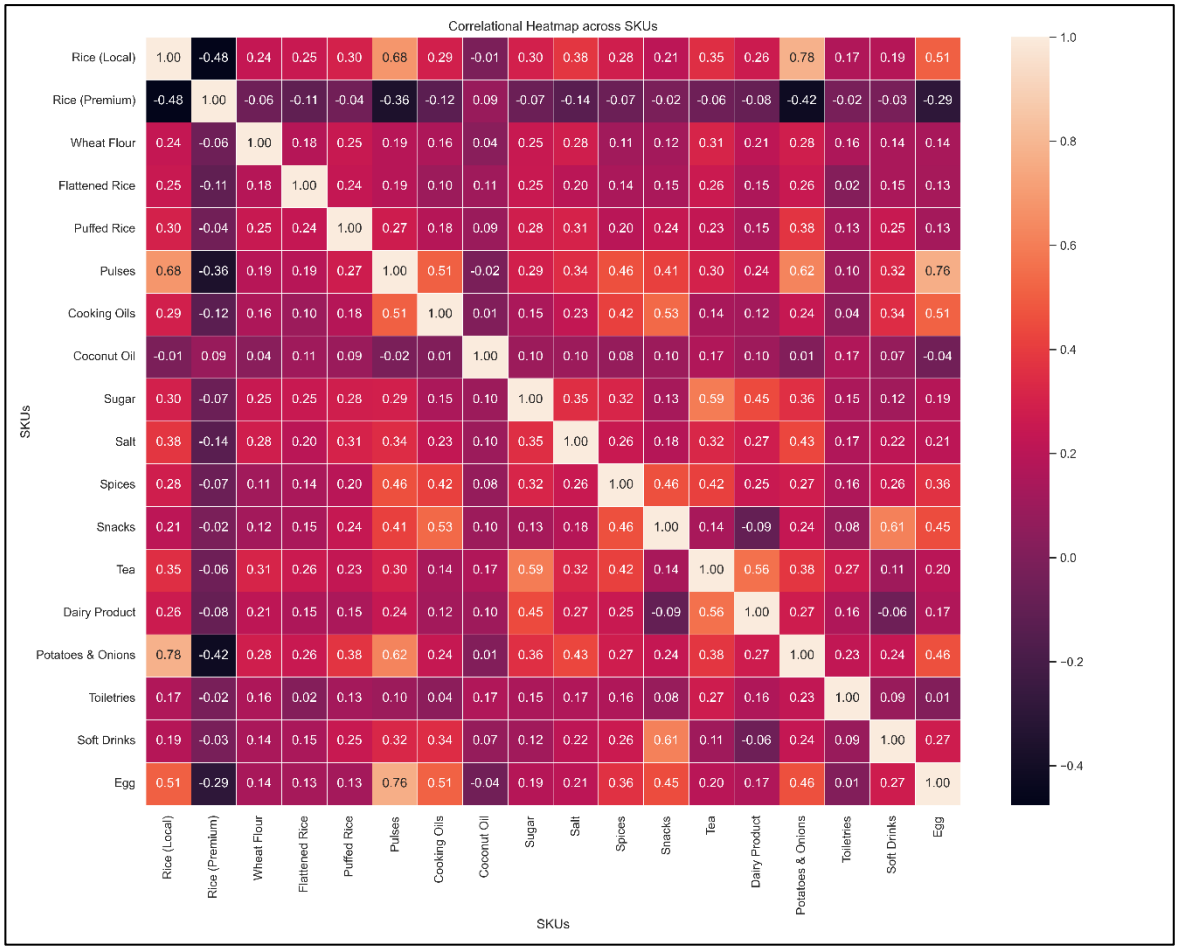

Figure 5: Correlational Heatmap

The Pearson correlation heatmap offers valuable insights into inter-product purchase behaviour, revealing patterns that can guide bundling, placement, and joint inventory strategies:

- Staples Often Bought Together
    - Rice (Local) has strong positive correlations with:
        - Pulses (r = 0.68)
        - Potatoes & Onions (r = 0.78)

    These relationships indicate that customers frequently purchase these staples together, reflecting core dietary habits.

- Snacks and Beverages Alignment
    - Snacks and Soft Drinks show a significant correlation of r = 0.61, suggesting they are commonly bought together—particularly relevant during festive or

leisure periods.

- Cooking Essentials as Cohesive Groups
    - Cooking oils are moderately correlated with:
        - Pulses (r = 0.51)
        - Eggs (r = 0.51)

  These correlations imply a natural bundling of ingredients used in cooking, especially during meal preparation spikes.

- Beverage Pairings
    - A moderate correlation between Tea and Sugar (r = 0.59) points to their typical usage together, validating opportunities for co-marketing and joint inventory controls.

- Category Overlaps
    - Several food essentials such as Salt, Cooking Oils, and Pulses show overlapping correlations in the 0.4–0.5 range with other staples, indicating shared consumption contexts.

- Negative Correlation Examples
    - A negative correlation exists between Rice (Local) and Rice (Premium) (r = -0.48), suggesting customers tend to choose one type over the other rather than both—highlighting a substitute relationship.

4. Interpretation of Results & Recommendations:

This section draws meaningful conclusions from the data analysis and proposes actionable solutions for Mallick Bhandar's two primary challenges:

1. Lack of Visibility into Seasonal Demand Patterns
2. Absence of Data-Driven Inventory Planning

Recommendations are structured as **SMART** (Specific, Measurable, Achievable,

Relevant, and Time-bound) actions to ensure practical implementation.

4.1 Problem 1: Lack of Visibility into Seasonal Demand Patterns

4.1.1 Interpretation

Multiple analyses confirm that Mallick Bhandar's sales are significantly influenced by seasonal and festival-driven demand surges, which were previously untracked.

- **Rolling Average Analysis** revealed visible demand spikes before October and March, especially for staples like Rice, Pulses, Egg, and Potatoes & Onions.
- **Festival Demand Comparison** showed dramatic sales increases during festivals:
    - Cooking Oils: +91%
    - Spices: +49%
    - Eggs: +48%
- **DVI** further reinforced that several products (e.g., Spices, Snacks) have inconsistent but high festival-linked demand patterns.

These seasonal behaviours were not visible in day-to-day operations and had led to stockouts during peak periods or over-purchasing post-festival.

4.1.2 Recommendations (Problem 1)

- Create a Festival-Specific Procurement Calendar:
    - Maintain a calendar of local events and pre-define a demand multiplier for each product category (e.g., Spices = 1.5x in October).
    - Action Timeline: Implement by next quarter.
- Implement 30-Day Rolling Average Monitoring:
    - Use Excel to monitor demand of 5–10 high-selling SKUs.
    - Adjust procurement volumes weekly during pre-festival buildup.
- Introduce Pre-Festival Stock Reviews:
    - Conduct manual reviews of high-variance items (like Snacks and Cooking Oils) 2–3 weeks before major festivals.
    - Target Products: Spices, Eggs, Oils, and Soft Drinks.

- Align Shelf and Storage Based on Seasonality:
  - Reserve flexible shelf space for products with known seasonal surges.
  - Relocate festival-relevant items to front racks during the season.
- Use Printed Demand Reports Before Events:
  - Prepare handouts for owners/staff showing top-10 demand increases before festivals to drive manual preparedness.

4.2 Problem 2: Absence of Data-Driven Inventory Planning

4.2.1 Interpretation

Analysis revealed clear inefficiencies in restocking and inventory control:

- **ABC Analysis** showed 70% of revenue comes from 8 essential products, but the shop treats all SKUs equally.
- **DVI** exposed high volatility in key SKUs like Spices (DVI ≈ 62%), Coconut Oil, and Toiletries, which need careful buffer stock planning.
- **Correlation Analysis** found strong purchase links (e.g., Rice ↔ Pulses, Snacks ↔ Soft Drinks) that are not yet leveraged in ordering or shelf organization.

Overall, the absence of SKU-wise planning has resulted in overstocking of low-impact products and reactive replenishment of high-impact ones.

4.2.2 Recommendations (Problem 2)

- Adopt ABC-Based Inventory Prioritization:
  - Focus daily checks and weekly restocking on **Category A** products: Rice, Pulses, Eggs, etc.
  - For **Category C** items (e.g., Coconut Oil, Toiletries), adopt monthly or bi-monthly restock strategy.
- Use DVI to Define Restock Frequency:
  - High DVI items → weekly monitored restocking + 8buffer (e.g., Spices).
  - Low DVI items → fixed reorder cycle (e.g., Sugar, Wheat Flour).
- Bundle Products Based on Correlation:

- Snacks + Soft Drinks, Tea + Sugar can be bundled in offers or co-located for joint purchase.

➢ Build a Reorder Dashboard in Excel:

- Track ABC category, DVI, and last stock date for each product.
- Helps set up reorder thresholds based on demand type (predictable vs. unpredictable).

➢ Begin Supplier Negotiations for Category A SKUs:

- Use annual sales data to negotiate discounts or priority delivery for essential high-volume products.

4.3 *Impact of Project Implementation*

➢ **Improved Forecasting Accuracy**:
By analyzing historical sales patterns and seasonal peaks, the business can now better predict demand 2–4 weeks ahead of key events.

➢ **Reduced Inventory Waste**:
Data-driven restocking strategies will help avoid overstocking of volatile items and reduce storage costs.

➢ **Enhanced Customer Satisfaction**:
Ensuring product availability during peak times leads to fewer lost sales and more satisfied customers.

➢ **Informed Decision-Making**:
The transition from manual intuition to data-backed insights empowers the store to operate proactively rather than reactively.