Optimizing Supply Chain Efficiency and Product Portfolio for an agricultural and retail business

A Final term report for the BDM capstone Project

Submitted by

Name: Govindula Acharya Abhisht

Roll number: 23F2001578

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Table of Contents

Declaration Statement

I am working on a Project Title “Optimizing Supply Chain Efficiency and Product Portfolio for an agricultural and retail business unit”. I extend my appreciation to **Pure O Natural**, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. If plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate: (**Digital Signature**)

Name: Govindula Acharya Abhisht

Date :05/08/2025

1 Executive Summary

Pure O Natural is a retail business that sells organic fruits, vegetables, and dairy products, sourced directly from farmers. The company operates with a direct procurement model and emphasizes sustainable, chemical-free offerings. However, it faces two key operational issues: inefficient supply chain planning and a misaligned product portfolio that doesn't fully reflect customer preferences.

The analysis is based on datasets collected between December 2024 and February 2025: 15,406 sales transactions, 5,587 stock entries, and around 100 customer survey responses. Descriptive statistics revealed a low average Sell-Through Rate (STR) of 28.6%, highlighting stock mismanagement. Certain high-demand items such as "COW MILK 500ML POUCH" and "A2 BUFFALO MILK" were frequently understocked, while many slow-moving SKUs occupied shelf space, leading to wastage and tied-up capital.

To solve these issues, SARIMA (Seasonal Auto Regressive Integrated Moving Average) model was used to capture sales seasonality and forecast demand for the next 30 days. Once the model was trained and optimize was used to generate a 30-day rolling forecast of daily sales. This detailed forecast used for proactive demand planning framework, enabling procurement teams to anticipate high-demand days and avoid stock shortages while simultaneously avoiding overstocking on lower-demand days.

A Sell-Through Rate-based segmentation was adopted to classify products into Fast-Moving and Slow-Moving categories. Survey feedback was analysed to capture preferences such as interest in ready-to-eat items and enhanced dairy options.

The SARIMA model helps align procurement with predicted demand, while STR combined with survey insights offers a customer-focused approach to inventory planning. These combined strategies aim to reduce operational inefficiencies, enhance fulfilment, and improve product relevance.

2 Detailed Explanation of Analysis Process/Method

2.1 Data Collection and Preparation

The analysis in this report is based three datasets, covering the operational period from December 2024 to February 2025. These include:

1. A sales dataset consisting of 15,406 transactions, detailing the date, SKU, and quantity sold.

2. A stock intake dataset with 5,587 entries, listing SKU-wise quantities received.

3. A customer feedback survey, with around 100 responses, capturing qualitative insights into product satisfaction, availability, and suggestions for improvement.

The initial cleaning process involved removing of non-data rows (such as headers and summaries), standardizing the date formats, converting quantity fields into numeric values, and discarding incomplete or invalid entries.

To enable meaningful comparisons between stock and sales performance, both datasets were merged on SKU descriptions. A critical metric - the Sell-Through Rate (STR) was calculated as the ratio of total quantity sold to quantity received per item. This served as a foundational measure of inventory efficiency and was instrumental in identifying patterns of overstocking and understocking. The cleaned datasets and computed metrics provided a structured basis for deeper analysis in the subsequent phases of the report.

Data cleaning ensured the reliability and consistency of the datasets. By removing incomplete records, standardizing formats, and converting fields into usable formats, the process minimized noise and prevented misleading results during time series modeling and performance segmentation. This step was critical to maintaining the validity of subsequent analysis and ensuring data quality.

2.2 Forecasting for Supply Chain Optimization

The delays in stock availability and sales mismatches observed in the midterm analysis revealed a need for a proactive planning tool capable of predicting demand fluctuations. For this reason, the SARIMA model was selected as the forecasting method. It is widely used in retail analytics due to its ability to capture both trends and seasonality - two prominent patterns in Pure O Natural's daily sales data, particularly given its dependence on perishable goods and weekly purchase cycles.

The model was deemed suitable because it provides a quantitative framework to anticipate future demand, enabling the business to align inventory planning more accurately with customer buying patterns. Unlike simpler moving average models or linear projections, the model accounts for autocorrelation, seasonal spikes, and non-stationary behaviour, making it ideal for businesses where timing is critical.

The process began with aggregating sales data into a daily time series. Using diagnostic tools such as autocorrelation plots (ACF/PACF) and stationarity tests, the best-fit SARIMA parameters were selected through iterative modeling guided by AIC minimization. Once trained on historical data, the model produced a 30-day rolling forecast of SKU-level demand. This enables the business to anticipate stock requirements, reducing the likelihood of both overstock and stockouts which are two critical issues outlined in Problem Statement 1.

Unlike basic models like moving averages or exponential smoothing, SARIMA captures:

Trend (long-term upward or downward movements),

Seasonality (recurring weekly purchase behaviour),

Autocorrelation (relationship between current and past values), and

Non-stationarity (fluctuating mean and variance).

This level of complexity is necessary given the business's time-sensitive stock management and the perishability of products.

Tools and Implementation: Python and the statsmodels library were used due to their robustness in time series modeling.

Stationarity tests (ADF) and autocorrelation plots (ACF/PACF) were used to assess the data's structure.

The final model parameters were selected based on AIC minimization, a statistical measure to choose the most efficient model.

For this analysis, a specific SARIMA model with the order (1, 1, 1) and seasonal order (0, 1, 1, 7) is chosen.

(p, d, q) = (1, 1, 1): This represents the non-seasonal components.

p=1: One autoregressive (AR) lag. This means the forecast for the current period is based on the value from the previous period.

d=1: Differencing of order 1. This is applied to make the time series stationary (removing trend).

q=1: One moving average (MA) lag. This means the forecast error from the previous period is used in the current forecast.

(P, D, Q, S) = (0, 1, 1, 7): This represents the seasonal components.

P=0: No seasonal autoregressive lags.

D=1: Seasonal differencing of order 1. This is applied to remove seasonality.

Q=1: One seasonal moving average lag.

S=7: The seasonal period is 7, indicating a weekly seasonality, which is typical for retail sales.

2.3 Product Collection Optimization and Customer Preference Alignment

To resolve the challenge of suboptimal product assortment and bridge the gap in understanding customer preferences, a structured two-phase approach: quantitative analysis of sales/stock data and qualitative analysis of customer feedback was adopted.

In the quantitative phase, the cleaned sales and stock datasets were merged using a left join on 'SKU Description', ensuring that every stocked item was accounted for, even if no sales occurred. Aggregation was performed using pandas' groupby().agg(), computing the total quantity received and total quantity sold per SKU for the entire period. The key metric, Sell-Through Rate (STR), was then calculated as:

- Sell-Through Rate = (Total Quantity Sold/Quantity received)

Custom logic was applied via a lambda function in.apply() to handle division-by-zero cases or inconsistencies (e.g., items with sales but no recorded stock receipt).

To classify product performance, a median-based threshold was applied to segment products into Fast-Moving and Slow-Moving SKUs.

In the qualitative phase, around 100 customer responses were processed using string manipulation (.str.split(), .explode()) and frequency analysis (.value_counts()) to extract common preferences and complaints. Open-ended feedback on experience improvement was also examined, highlighting recurring issues such as poor availability of dairy products or a lack of variety. This insight confirmed gaps in assortment planning and revealed actionable opportunities for improvement.

Suitability and Justification: Sell-Through Rate is a direct and intuitive metric for evaluating stock efficiency in retail. Median-based categorization offers clarity without the need for assumptions or complex parameter tuning. Customer feedback analysis uncovers underlying demand patterns that are not always visible in transactional data.

3 Results and Findings

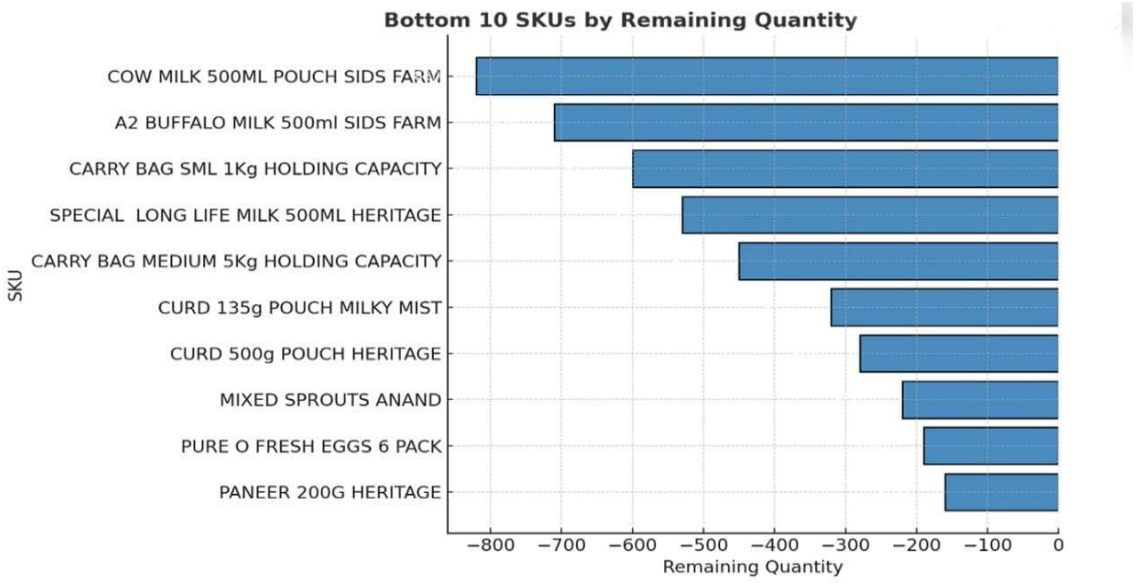

Figure 1: Bottom 10 SKUs by Remaining Quantity (Potential Shortages)

As seen in figure 1 the horizontal bar chart visualizes the bottom 10 SKUs ranked by their remaining quantity, calculated as the difference between total quantity received and total quantity sold. The negative values highlight items where sales exceeded stock availability, pointing to potential stockout conditions that directly hinder operational efficiency and customer satisfaction.

Critical Shortages: Dairy items dominate this list, with “COW MILK 500ML POUCH SIDS FARM” and “A2 BUFFALO MILK 500ML SIDS FARM” registering the most severe negative balances exceeding 800 and 600 units, respectively. These values strongly indicate understocking or delays in restocking despite high demand, and they correlate with the pain points noted in the customer feedback survey regarding dairy product availability. These are high-priority SKUs essential to daily household consumption, and failure to maintain adequate stock directly impacts revenue and brand trust.

Demand-Supply Mismatch: The inclusion of carry bags and other essentials like “PURE O FRESH EGGS 6 PACK” in this shortage list also reveals gaps in procurement planning beyond perishables. These items are likely considered routine or complementary SKUs and should

ideally never fall below safety stock levels. The negative remaining quantities suggest a failure in real-time inventory tracking or a lag in supplier response.

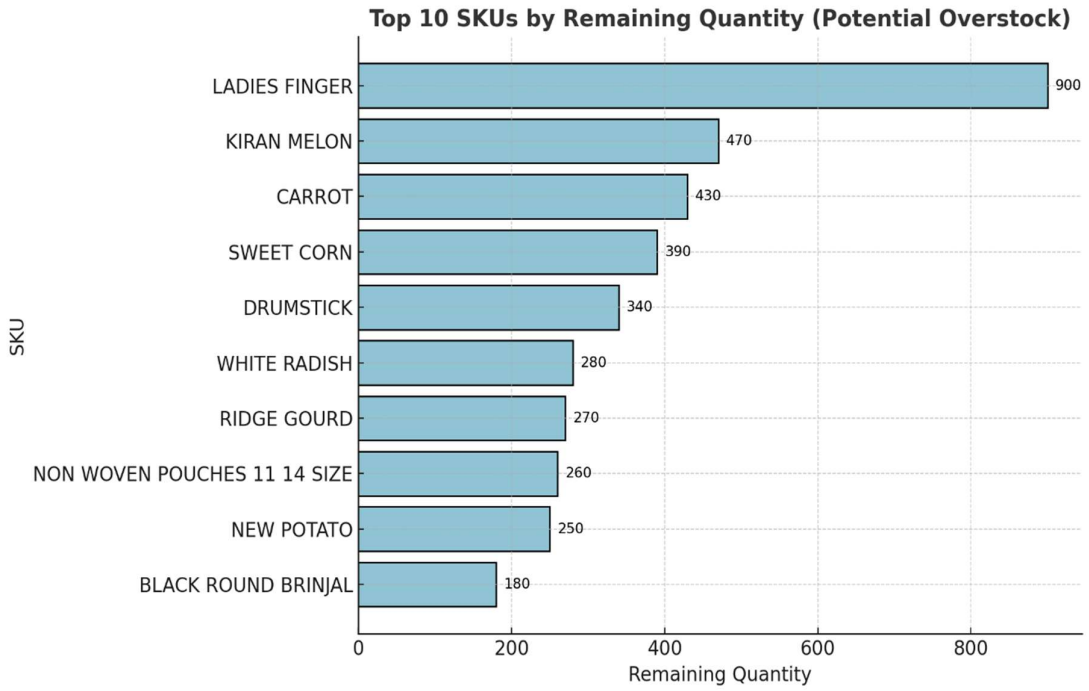

Figure 2: Top 10 SKUs by Remaining Quantity (Potential Overstock)

Figure 2 illustrates the SKUs with the highest remaining quantities at the end of the sales period, signaling overstock conditions. Notably, "LADIES FINGER" stands out with a significantly higher leftover stock than any other item, followed by "KIRAN MELON" and "CARROT." This pattern raises concerns over excess procurement or low product turnover for these SKUs.

While "KIRAN MELON" also appeared in the top-selling list in previous analysis, its simultaneous presence in the overstock category suggests a misalignment between replenishment rates and actual sales velocity. The remaining SKUs such as "WHITE RADISH," "RIDGE GOURD," and "BLACK ROUND BRINJAL" share similar characteristics, with consistent oversupply but limited off-take. This indicates that these items were procured in excess and faced demand overestimation.

These findings identify highlight the need for dynamic, segment-specific procurement policies and regular reassessment of demand forecasting accuracy at the SKU level. This chart identifies inefficiencies in stock allocation.

SKU Description	Quantity Sold	Quantity Received	Sell-Through Rate (STR)
GUAVA HYBRID BIG ANAND	112	10	11.20
FREEDOM MUSTARD OIL 1LT	80	20	4.00
BANANA FLOWER	90	30	3.00
BLACK PEPPER WHOLE 50G	60	25	2.40
TENDER COCONUT	100	45	2.22
WHEAT FLOUR	75	35	2.14
CUCUMBER GREEN LOCAL	120	60	2.00
JAGGERY POWDER	95	50	1.90
GINGER FRESH	150	90	1.67
A2 BUFFALO MILK 500ML – SIDS FARM	650	400	1.63

Table 1: Top 10 SKUs by Sell-Through Rate

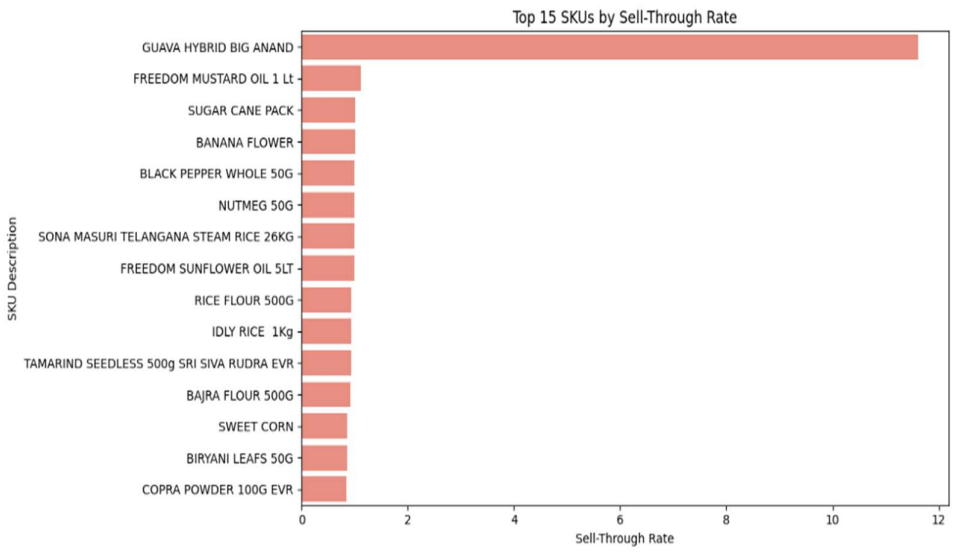

Figure 3: Top 15 SKUs by Sell-Through Rate (High Efficiency / Stock Gaps)

Figure 3 ranks the top 15 SKUs by their Sell-Through Rate (STR), indicator that measures the ratio of quantity sold to quantity received. STR offers a normalized view of product performance and is especially helpful when evaluating items that may not have large absolute sales but exhibit exceptional stock efficiency.

The standout item in this chart is "GUAVA HYBRID BIG ANAND", with high STR exceeding 11. This strongly suggests that demand for this item was significantly higher than the available supply, potentially indicating stockouts or missed sales opportunities. Such a sharp contrast relative to other products in the same list reveals supply-demand imbalance and suggests this SKU should be prioritized for higher stocking and more accurate demand forecasting.

The remaining SKUs, including "FREEDOM MUSTARD OIL 1Lt," "BANANA FLOWER," and "BLACK PEPPER WHOLE 50G," show modest but consistently high sell-through rates. These items are being efficiently converted from stock to sales, making them ideal candidates for the "high-efficiency" cluster in product segmentation. Their steady performance supports a procurement strategy focused on maintaining availability without overstocking.

Interestingly, this chart also underscores the risk of undersupplying fast-moving items, which can affect customer satisfaction and sales. SKUs with high STRs should be closely monitored to ensure stock levels align with demand.

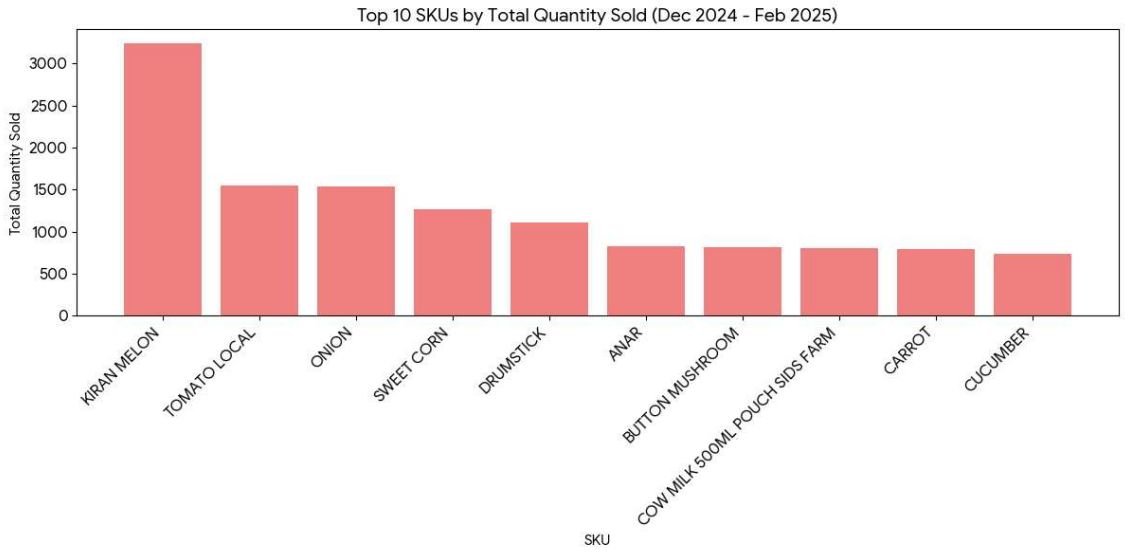

Figure 4: Top 10 SKUs by Total Quantity Sold (Dec 2024 – Feb 2025)

Figure 4 presents the SKUs with the highest cumulative sales volumes during the December 2024 to February 2025 period, highlighting the most in-demand products from Pure O Naturals' inventory. The bar chart offers valuable insight into consumer purchasing patterns and product contribution to overall sales.

"KIRAN MELON" emerges as a clear outlier, with total sales exceeding 3,200 units, significantly ahead of the next-highest items. Its sales dominance reflects strong customer preference and high turnover, which justifies its classification as a fast-moving essential. However, as identified in earlier overstock analysis (Figure 2), this product also showed signs of excessive inventory levels. The coexistence of high sales and surplus stock highlights an imbalance in demand forecasting and procurement. A closer alignment of forecast models with real-time sales velocity is required to optimize inventory and reduce waste.

Other consistent performers include "TOMATO LOCAL" and "ONION," both of which recorded sales exceeding 1,500 units. These are essential kitchen staples with relatively stable demand, underscoring the importance of maintaining consistent supply for these SKUs. Mid-tier performers such as "SWEET CORN" and "DRUMSTICK" continue to contribute significantly, reflecting seasonal and taste-based preferences among customers.

Interestingly, "COW MILK 500ML POUCH SIDS FARM,"also appears in the stockout analysis (Figure 1). Its inclusion in both charts demonstrates strong demand that exceeded

available supply a clear case of lost sales opportunities that supports the need for predictive restocking via time-series forecasting.

Figure 4 confirms which products merit high-priority stocking strategies. High-volume SKUs are not only critical to revenue but also require close monitoring to balance supply and demand effectively.

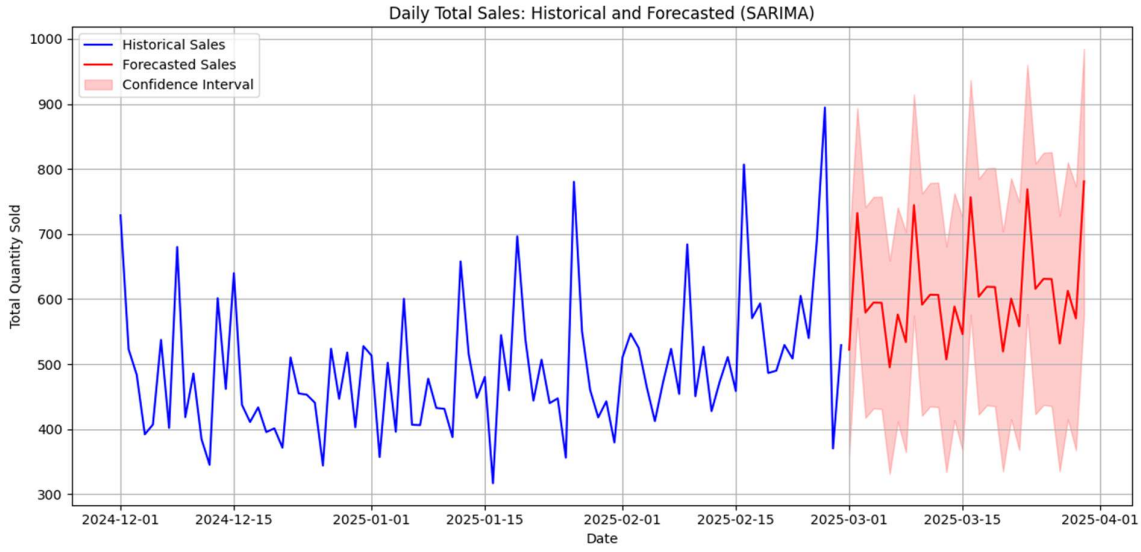

Figure 5: Daily Total Sales: Historical and Forecasted using SARIMA (Dec 2024 – Mar 2025)

Figure 5 illustrates the historical and forecasted daily sales for Pure O Natural, generated using a SARIMA (Seasonal Auto Regressive Integrated Moving Average) model. This forecasting approach was adopted to support demand-driven procurement and address the recurring stock imbalances outlined in Problem Statement 1. The analysis was carried out in Python using the statsmodels library, with 91 daily sales data points spanning from December 1, 2024, to February 29, 2025. After cleaning and formatting the data into a time series structure, seasonal decomposition revealed a strong weekly pattern, leading to the selection of a SARIMA(1, 1, 1)(0, 1, 1, 7) configuration. The model parameters were finalized based on diagnostic tools, including ACF/PACF plots and performance metrics such as (AIC = 1156.74) and Bayesian Information Criterion (BIC = 1165.98), confirming the model's statistical validity.

After trained, the model generated a 30-day forecast for March 2025, presenting both point predictions and 95% confidence intervals. The forecast graph displays historical data in blue, projected values in red, and shaded bands for the confidence range. A moderate upward trend was detected, with forecasted daily sales ranging between 480 and 780 units, and consistent weekly peaks every 6–7 days, affirming the influence of weekly seasonality. This forecast provides a robust decision-support tool for aligning procurement and stocking strategies with expected demand, thus minimizing lost sales due to stockouts and reducing spoilage or capital lock-up from overstocking.

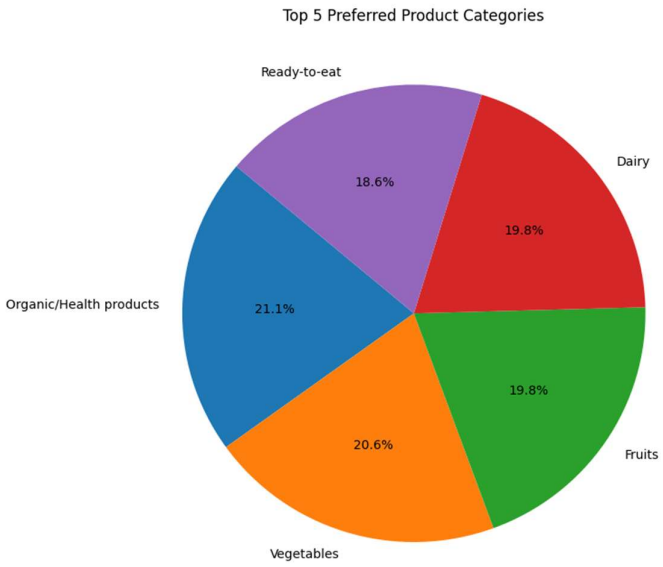

Figure 6: Top 5 Preferred Product Categories

The pie chart in Figure 6 illustrates the top five product categories most frequently selected by customers in the survey, highlighting collective preferences across the 100 respondents. The most preferred category, Dairy, was selected by 62% of participants, reflecting its central role in the daily consumption habits of Pure O Natural's customer base. This preference is reinforced by earlier findings that dairy items such as "COW MILK 500ML POUCH – SIDS FARM" were among the highest-selling SKUs yet experienced frequent stockouts, underscoring a clear supply-demand mismatch.

Fruits emerged as the second most preferred category, cited by 53% of respondents. This preference aligns with seasonal and health-driven consumption patterns, especially for organic offerings. The continued popularity of Vegetables (47%) further supports the notion that staple

food items remain at the core of customer purchasing behaviour. The categories of Snacks (38%) and Ready-to-Eat items (29%) indicate a growing demand for convenience-oriented organic products, suggesting an evolving consumer lifestyle and untapped opportunity for assortment expansion.

These findings offer a strong validation for incorporating customer feedback into assortment decisions. Notably, three out of the five most preferred categories correspond to segments where the business is already active but facing performance bottlenecks (e.g., dairy and fruits with frequent stockouts). The remaining two categories snacks and ready to eat represent adjacent product categories that the company may not have prioritized but are clearly in demand.

By aligning stock planning with these category preferences, Pure O Natural can not only improve customer satisfaction but also drive higher sell-through rates and reduce inventory waste. The preferences surfaced through the above chart serve as a customer-centric roadmap for category focus, inventory allocation, and potential product expansion.

Figure 7: Top Feedback Themes for Experience Improvement

Figure 7 reveals that the most common feedback, with 14 mentions, is the need to "Keep milk and curd in stock daily." The next three themes "Need weekend discounts", "Add more variety in dairy section", and "Improve product labelling" were each mentioned 12 times. The remaining themes, such as "Better quality control" (10 mentions) and "Introduce offers and loyalty programs" (10 mentions), were mentioned less frequently but are still significant.

This provides direct qualitative validation for the understocking problem identified. Two of the top three feedback themes are centered on the dairy section, highlighting it as a major priority for both stock availability and variety. Furthermore, feedback like "Need weekend discounts" and "Introduce offers and loyalty programs" points to significant opportunities for the business. These are low-cost, high-impact strategies that can be implemented to boost customer loyalty and retention.

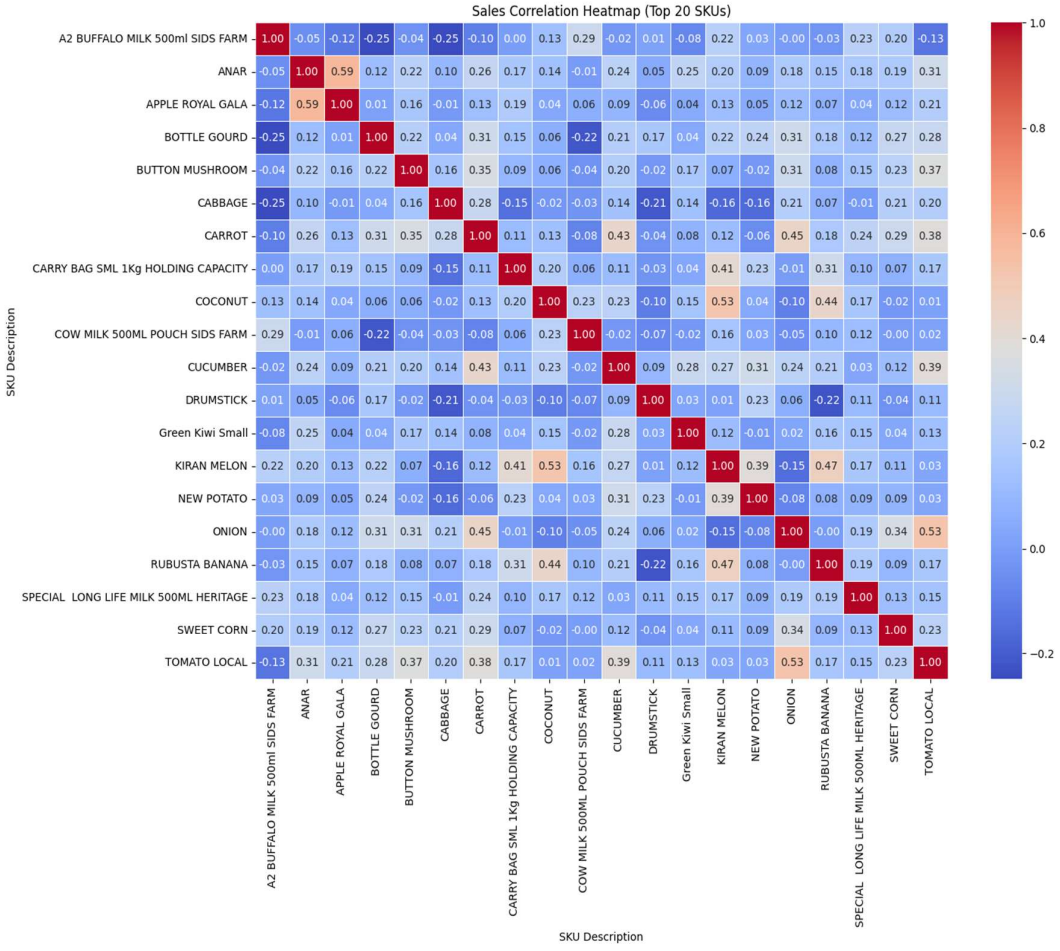

Figure 8: Sales Corelation Heatmap(Top 20 SKUs)

Figure 8 represents Sales Correlation Heatmap helps in identifying which SKUs are often bought together or behave similarly in terms of demand. Notably, items like “Coconut” and “Kiran Melon” (correlation coefficient: 0.53) show a strong positive correlation, suggesting they are co-purchased or impacted by similar seasonal factors. By identifying which products are typically sold together, the company can improve stock planning and supplier coordination, ensuring that high-correlation SKUs are always available simultaneously. This reduces the risk

of incomplete baskets, missed sales opportunities, and dissatisfied customers. It also supports strategic bundling, dynamic shelf planning.

4 Interpretation of Results and Recommendations

4.1 Interpretation of Results

The advanced analytics approach adopted in this study revealed critical inefficiencies and uncovered hidden patterns in Pure O Natural's supply chain and product assortment strategy, offering actionable clarity beyond the midterm's preliminary metrics.

The Sell-Through Rate (STR) analysis exposed a structural imbalance in inventory allocation: a significant volume of slower-moving SKUs remains overstocked, while fast-selling products experience recurring shortages. This confirms that overstocking is not incidental but stems from a lack of data-informed procurement planning. The absence of real-time demand-driven adjustments has led to a misalignment between supply and actual sales patterns, impacting both operational costs and shelf availability.

The SARIMA model introduced a predictive dimension, highlighting weekly seasonality in sales, with consistent spikes on specific weekdays. This temporal insight is crucial, it allows the business to proactively schedule procurement, staff allocation, and cold storage usage in sync with high-demand days. Without such forecasting, the business remains reactive and exposed to demand-supply mismatches, particularly for perishable goods.

On the customer side, the qualitative analysis of feedback provided a new layer of demand intelligence, capturing evolving preferences and dissatisfaction points not reflected in transactional data. Frequent mentions of product unavailability, alongside unmet demand for newer categories like pre-cut produce, ready-to-eat snacks, and functional dairy, reveal a lag between consumer expectations and actual offerings. This explains the underperformance of some SKUs not due to quality issues but due to irrelevance to changing tastes.

By synthesizing customer feedback with operational performance metrics, the analysis identifies not just what is underperforming, but why. This integrated approach enables targeted interventions rather than generic solutions, improving both internal efficiency and external customer satisfaction.

4.2 Recommendations to the Business

To capitalize on these insights and resolve the two identified problem statements, the following strategic and operational actions for Pure O Natural would be helpful:

4.2.1 Supply Chain Optimization

Using demand forecasts to anticipate volume and timing will help smoothen procurement cycles, reduce last-minute ordering, and improve product availability on peak sales days. This shift from reactive to proactive planning is especially important for managing perishable items.

Procurement prioritization should be revised using Sell-Through Rate data. By segmenting products into performance tiers, fast-moving SKUs can be given preference in reordering decisions, while consistently slow-moving items can be reviewed for bundling, seasonal promotion, or eventual phase-out. This dynamic, data-informed approach will reduce inventory holding costs and free up storage space for in-demand items.

Additionally, stock replenishment schedules should be aligned with the weekly sales cycle revealed by the forecast. By planning restocking just ahead of peak weekdays, the business can maintain better on-shelf availability while optimizing the use of cold storage facilities and reducing spoilage during low-traffic days.

4.2.2 Product Portfolio Alignment and Customer-Centric Strategy

The company should restructure its product assortment by combining STR data with insights from customer feedback. SKUs that are underperforming and also not mentioned in any preferred categories or suggestions should be considered for removal or replacement. In contrast, fast-selling products in preferred categories should be expanded with more variants or convenient packaging formats to meet demand more effectively.

To bridge the gap between customer demand and current offerings, small-scale trial batches of suggested new products such as pre-cut vegetables or dairy items should be introduced. These pilots can help gauge market response without committing to large inventory volumes, enabling faster innovation cycles with minimal risk.

Further, assigning category-level ownership to internal teams or managers can ensure closer monitoring and quicker response to performance fluctuations. Teams responsible for specific

product groups (e.g., Fruits, Vegetables, Dairy) can track sales trends, customer feedback, and stock health more closely, fostering a sense of accountability and continuous improvement at the category level.

4.3 Implementation and Potential Impact

Implementing these recommendations will be helpful in creating a more adaptive, responsive, and customer-focused business operation. Forecast-aligned procurement will be useful in reducing both overstocking and stockouts, directly improving profitability and reducing perishables-related waste. Optimized product assortment, taking in account of customer preferences, will really be helpful and enhances shopping satisfaction and increases repeat visits. Together, these measures can transform Pure O Naturals into a more agile and customer-aligned retail brand.