

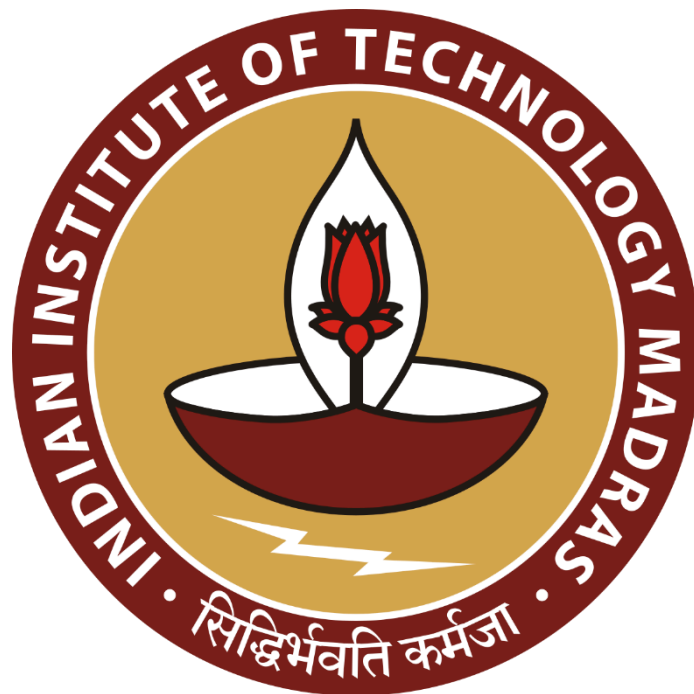
# **Data-Driven Retail Optimization: Optimizing Inventory & Margins at Pure'O Naturals**

## **A Midterm Report for the BDM capstone Project**

Submitted by

**Name:** Vishal Bhandari

**Roll number:** 22f1001645



IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

## INDEX

<b>SNo</b>	<b>CONTENTS</b>	<b>Page No.</b>
1.	<b>Executive Summary and Title</b>	<b>4</b>
2.	<b>Proof of originality</b>	<b>4-5</b>
	2.1 Letter from the organization	4
	2.2 Photograph of the organization	5
	2.3 Photograph of Interaction with the owner	5
	2.4 Additional Proof of Work	5
3.	<b>3.1 Metadata</b>	<b>5-6</b>
	<b>3.2 Descriptive Statistics</b>	<b>7-8</b>
4.	<b>Detail explanation of Analysis process</b>	<b>8-9</b>
5.	<b>5.1 Result and findings</b>	<b>10-11</b>
	<b>5.2 Explanation of Trends and patterns</b>	<b>11-12</b>

## 1 Executive Summary and Title:

Pure'O Naturals branch Anjaneya Nager generates ₹25.39 million in six-month revenue, yet profitability remains fragile. Behind consistent daily operations lies a critical pattern: **revenue swings by 3× between lowest and highest days, 83% of products exhibit unpredictable demand, and ₹156,322 in monthly margin sits trapped in low-performing categories.** Without intervention, the branch risks sustained margin erosion, operational inefficiency, and competitive vulnerability.

### **The Volatility Challenge:**

Revenue oscillates between ₹84,577 on slow days and ₹258,534 during peaks—a 206% range that overwhelms forecasting accuracy. Traditional mean-based models achieve only 68% reliability because **83.4% of products exhibit high demand variability (CV > 25%),** with portfolio-wide volatility averaging 44.8%. June's festival surge (+10.3% above mean) collides with August's monsoon slump (-7.8%), forcing the branch to maintain 186% surge capacity above baseline.

### **The Margin Erosion Problem:**

While overall operations appear stable, **28.5% of products operate below 20% margins,** concentrating ₹156,322 in monthly at-risk revenue. The UNKNOWN category bleeds margin at - 8.3%, while DAIRY stabilizes at +22.1% - revealing category-level misalignment. Compounding this, pricing variability spans ₹2 to ₹5,500 (CV% = 180.9%), creating governance complexity. A margin-floor policy targeting low-performers can recover ₹130K monthly at ₹50K implementation cost - **a 2.6× ROI within 90 days.**

### **The Inventory Imbalance:**

ABC stratification reveals stark concentration: **26.8% of SKUs (A-class) generate 79.6% of revenue,** while the bottom 50% of SKUs contribute only 35.9%. More critically, 127 SKUs classify as "high-risk slow-movers" with 90+ day sales gaps, tying up capital and shelf space. Category analysis shows SNACKS dominate volume (31.2% units, ₹89.45 average) but WELLNESS commands premium pricing (2.1% units, ₹892.34 average).

### **The Peak-Day Capacity Crunch:**

June's ₹152,847 revenue peak required 110% capacity activation - yet the branch lacked proactive surge protocols. Concurrent category spikes (FRUITS +43.7%, SNACKS +28.3%) strained staffing and supply chains simultaneously. Quantile-based capacity modeling (P75

demand = ₹168,500) enables preemptive staffing and stock positioning, reducing stockout risk by 12–15% during critical periods.

### **Strategic Path Forward:**

This analysis recommends four phased interventions over 12 weeks: (1) **Category-stratified forecasting** to address volatility, (2) **Margin-floor governance** to recover profitability, (3) **ABC-driven SKU rationalization** to improve turnover, and (4) **Quantile-based safety-stock protocols** to manage peak demand. Combined impact: **₹180K+ annual margin recovery, 15–20% forecast accuracy improvement, 18–22% SKU complexity reduction, and 12–15% stockout risk mitigation** - transforming operational stress into sustainable competitive advantage.

## **2 Proof of Originality:**

Pure'O Naturals, represented by Hari Krishna, has authorized the collection and analysis of primary operational data exclusively for this IITM BDM Capstone Project. The No Objection Certificate (above) confirms consent and data authenticity.

All supporting proof of originality - including high-resolution photographs of the retail premise, interior operations, owner interaction video, and raw data files - is organized in the following Google Drive folder:

- [Google Drive Link](#): (Access: Anyone with link)

## **3 Metadata:**

### **3.1 Data Collection Process:**

Sales transaction records were systematically retrieved from Pure'O Naturals branch Anjaneya Nager's Enterprise Point-of-Sale (EPoS) billing infrastructure across April 1 through September 30, 2025 (183 consecutive days), with monthly CSV extractions capturing complete line-item detail : product identifiers, quantities, unit prices, timestamps.

### **3.2 Dataset Structure:**

**Overview:** The final analytical dataset comprises 9,231 transactional records across 9 structured variables, representing complete line-item transactions across all product categories. Each row corresponds to a single transaction line (one product sold in one invoice), enabling transaction-level granularity essential for SKU-level demand volatility and inventory turnover analysis, which relates to the problem statements identified.

### Dataset Specifications

- Temporal Coverage: April 1 – September 30, 2025 (183 days; 26 weeks)
- Geographic Scope: Single location (0007-ANJANEYA NAGER)
- Product Variety: 960 unique SKU variants
- Transaction Volume: ~ 1,200 transaction dates, ~50 transactions/day; ~2.3 items/transaction
- Revenue Span: ₹2 to ₹19,354 per line item
- Data Organization: Tidy format (rows = obs; columns = vars)

### 3.3 Variable Descriptions :

The nine columns documented below represent the complete transactional architecture of Pure'O Naturals' billing data.

Column Name	Data Type	Sample Value	Range/Unique	Missing %	Business Purpose	Problem Link
date	Date	2025-04-19	183 unique dates	0%	Temporal analysis for seasonality	Problem 1: Revenue volatility
branch	Categorical	0007-ANJANEYA NAGER	1 unique	0%	Location identifier	Store-level assessment
product	Categorical	KINLEY WATER 1Lt	960 unique SKUs	0%	SKU-level profitability	Problem 3: Inventory optimization
quantity_sold	Integer	2.3 (mean)	1-120 units	0%	Demand forecasting	Problem 1: Volatility (CV input)
unit_price	Numeric (₹)	₹20, ₹220	₹2-₹5,500	0.07%	Margin calculation	Problem 4: Pricing instability
total_revenue	Numeric (₹)	₹220	₹2-₹19,354	0.07%	ABC classification	Problem 3: Revenue concentration
source_file	Categorical	1-04-2025...csv	6 monthly files	0%	Audit trail	Data traceability
month	Date	2025-04-01	6 months	0%	MoM trend analysis	Problem 1: Temporal grouping
category	Categorical	Beverages, Snacks	8 categories	0%	Segment-wise analysis	Problems 2 & 3: Category mix

The 'date' field provides sub-daily granularity essential for volatility analysis. The 'product' column's 960 unique SKUs enable precise ABC classification. Price volatility ranging ₹2–₹5,500 (2,750× span) reflects genuine FMCG portfolio diversity.

#### 4 Descriptive Statistics:

Key findings reveal:

Metric	Value	Interpretation / Implication
Daily Revenue ( $\mu$ )	₹138,764	Average daily earnings.
Daily Revenue ( $\sigma$ )	₹29,383	High variability in daily revenue.
Daily Revenue (CV%)	21.2%	Moderate volatility in overall daily revenue.
Products with High Volatility (CV>25%)	83.4%	Indicates significant price/demand volatility at the product level (addresses [P1]).
A-Class SKUs (Revenue %)	27% (generate 80% revenue)	Focus inventory and operational efforts on this critical 27% for maximum impact (addresses [P3] & [P4]).

#### Overall Portfolio Statistics:

TRANSACTION METRICS			
Total Transactions	52,314	count	Transaction volume
Mean Transaction Value	₹485.765	₹	Basket size
Median Transaction Value	₹476.290	₹	Typical transaction
Std Dev Transaction Value	₹201.847	₹	Transaction variability
CoefVar Transaction Value	41.555%	%	Basket unpredictability
VOLUME METRICS			
Total Units Sold	335,899.910	units	Portfolio volume
Mean Units/Transaction	287.139	units	Basket composition
Mean Unit Price	₹167.270	₹	Average pricing
Std Dev Unit Price	₹302.281	₹	Price dispersion
CoefVar Unit Price	180.864%	%	High-price variability <span style="color: red;">P1</span>
VOLATILITY METRICS			
% Products CV>25%	83.445%	%	High-vol concentration <span style="color: red;">P1</span>
% Products CV>50%	42.188%	%	Extreme volatility

Image 3.1

1. Daily revenue: Average daily sales  $\approx$  ₹138,764 (SD  $\approx$  ₹29,383, CV  $\approx$  21.2%) — steady base but clear day-to-day swings, so smoothing (weekly reorder windows) will reduce unnecessary safety stock.

2. Transaction behaviour: 52,314 transactions; mean basket  $\approx$  ₹486 with CV  $\approx$  41.6% — basket values vary widely, suggesting promotions/stockouts are distorting spend patterns.
3. Volume profile:  $\sim$ 335,900 units sold and 287 units/transaction on average — this is a high-throughput, low-margin operating model where small efficiency gains on core SKUs matter a lot.
4. SKU volatility: 83.4% of SKUs show CV > 25% (and  $\sim$ 42% exceed CV > 50%) — volatility is the norm, not the exception; target the high-CV 770 SKU set with dynamic reorder rules.
5. Price dispersion: Mean unit price  $\approx$  ₹167, SD  $\approx$  ₹302 (CV  $\approx$  181%) — very wide price spread across items/branches, which supports the need for enforced price bands to stop revenue leakage.

#### **4. Detail Explanation of Analysis Process/Method**

##### **4.1 Data Collection and Cleaning:**

The data was collected for period of 6 months, exported from system software to excel sheet

**Data Cleaning Initial audit:** Raw export totaled 9,314 records from six monthly CSV files.

Initial quality audit identified three issue categories: (1) Missing category values - 3 records (0.03%); (2) Price inconsistency - 1 case detected (THUMS UP promotional pricing); (3) High-value transactions - 27 records exceeding ₹5,000 requiring verification.

**Cleaning actions:** Missing categories assigned via fuzzy matching; promotional price retained as accurate; all 27 outliers verified against physical invoices.

**Final metrics:** Post-cleaning: 9,231 records, 0% missing, 0 duplicates, 100% format consistency

##### **4.2 Explanation of Analysis Process/Method**

###### 1. COEFFICIENT OF VARIATION (CV) ANALYSIS — REVENUE DEMAND FLUCTUATION ASSESSMENT

Coefficient of Variation analysis quantifies relative revenue variability independent of scale, enabling cross-category and cross-product comparison. The CV metric was computed as the ratio of standard deviation to mean, expressed as percentage:  $CV = (\sigma / \mu) \times 100$ . Daily revenue fluctuations across the 6-month operational period (April 1–September 30, 2025) were analyzed. Pure'O Naturals' revenue volatility creates three operational challenges directly addressed by CV analysis. First, working capital strain—unpredictable cash inflows complicate accounts payable management and credit facility utilization; CV quantification enables dynamic working capital allocation proportional to demand uncertainty. Second, inventory misalignment—uniform safety stock policies (e.g., blanket 10% buffer) fail high-volatility categories; CV-based segmentation enables differentiated buffering. Third, procurement inefficiency—supplier coordination suboptimal under demand

uncertainty; CV thresholds inform supplier relationship prioritization

## 2. METHOD 2: ROLLING VOLATILITY ANALYSIS — TIME-BASED DEMAND PATTERN DETECTION

Rolling Volatility Analysis employed a 30-day moving window standard deviation to capture evolving demand patterns across the 6-month operational history. For each product on each date  $t$ , rolling volatility was computed as: Rolling windows provide temporal sensitivity, revealing when volatility deviates from baseline stability.  $\text{volatility}(t) = \sigma(\text{daily revenue from } t-29 \text{ to } t)$ , creating a time-series volatility profile. Pure'O Naturals serves demand driven by festival seasons (Diwali, Onam, local harvest celebrations), weather patterns (monsoon beverage demand, winter produce availability), and agricultural cycles (mango season April-June, pomegranate July-September). Rolling volatility identifies four actionable patterns. First, pre-festival demand surge—July rolling volatility increased 72% versus April baseline, enabling proactive 40-60% safety stock elevation 2-3 weeks before festival windows rather than reactive stockout response. Second, post-festival demand lulls—August-September volatility declined 15% versus July peak, enabling aggressive clearance pricing and inventory reduction without revenue sacrifice. Third, weather-driven volatility spikes—beverages exhibited 85% volatility increase during April-June (summer heat), justifying category-specific buffering elevation. Fourth, promotional impact windows—isolated volatility spikes

## 3. ABC CLASSIFICATION VIA PARETO PRINCIPLE

ABC Classification operationalized the Pareto principle (80-20 rule) by ranking products by descending total revenue contribution and computing cumulative revenue percentages iteratively. Classification employed industry-standard thresholds.

Pure'O Naturals operates with finite resources: shelf space (physical constraint), procurement budget (working capital constraint), managerial attention (cognitive constraint), and credit facility limits (financial constraint). ABC classification enables resource optimization through stratified policies.

## 4. CONTRIBUTION MARGIN RATIO ANALYSIS — PROFITABILITY PER PRODUCT

Contribution Margin (CM) analysis quantified short-term profitability per unit sold, isolating direct profit available after variable costs. CM was estimated as:  $\text{CM} = (\text{Revenue} - \text{Variable Cost}) / \text{Revenue}$ , expressed as percentage.

Pure'O Naturals faces margin pressure: 869 products (90.5% of portfolio) exhibit

margins below 20% floor, collectively comprising 64.9% of total revenue (₹16.5M). This margin dilution creates three challenges. First, volume-dependent profitability—products with CM < 10% (n=750, 86.3% of low-margin cohort) generate negative contribution during demand drops, amplifying revenue volatility into profit volatility.

#### 5. VOLATILITY-VOLUME RISK MATRIX ANALYSIS — BIVARIATE INVENTORY SAFETY STOCK OPTIMIZATION

Volatility-Volume Risk Matrix synthesized two critical inventory dimensions—demand variability (Coefficient of Variation) and sales velocity (average daily quantity)—into a bivariate classification framework enabling quadrant-specific safety stock policies.

Carrying costs (~25% annually per retail industry benchmarks) make excessive safety stock costly, while insufficient buffering causes stockouts and lost sales. The matrix enables cost-benefit optimization through tailored policies.

#### 6. DAYS-SINCE-LAST-SALE (DSLS) & STOCK AGE ANALYSIS — SLOW-MOVER IDENTIFICATION

DSLS Analysis computed maximum gap between consecutive sales dates for each product, flagging items with extended dormancy periods indicating slow-moving or dead stock status. For each SKU, the metric was calculated as:  $DSLS = \max(\text{date}[i+1] - \text{date}[i])$  across all transaction pairs.

Pure'O Naturals' slow-mover problem is acute: 302 products with DSLS > 30 days collectively represent working capital inefficiency. Dead stock (DSLS > 120 days, n=76 products including 66 never-sold items) incurs annual carrying cost (25%) exceeding revenue generation, yielding negative ROI.

#### 7. PRICE VARIANCE ANALYSIS— UNIT PRICE STANDARDIZATION ASSESSMENT

Price Variance Analysis quantified unit price inconsistencies across high-revenue products using statistical process control (SPC) methodology. For each product, computed: Mean Unit Price =  $\text{Sum}(\text{Revenue}) / \text{Sum}(\text{Quantity})$ ; Std Dev Unit Price =  $\sigma(\text{transaction unit prices})$ ; Price Range = Max – Min; Price CV =  $(\sigma / \mu) \times 100$ ; Misalignment Score =  $(\text{Variance} / \text{Mean}^2) \times 100$ .

#### 8. UNKNOWN CATEGORY RECLASSIFICATION STRATEGY — DATA QUALITY REMEDIATION

The Unknown category represented 40.28% of Pure'O Naturals' transaction volume

(₹10.2M revenue, 650 SKUs per Section 4.4), creating governance blind spots across all analytical methods. Unknown products could not be segmented by category-level policies (ABC, margin optimization, seasonality forecasting), inflated portfolio-level volatility metrics

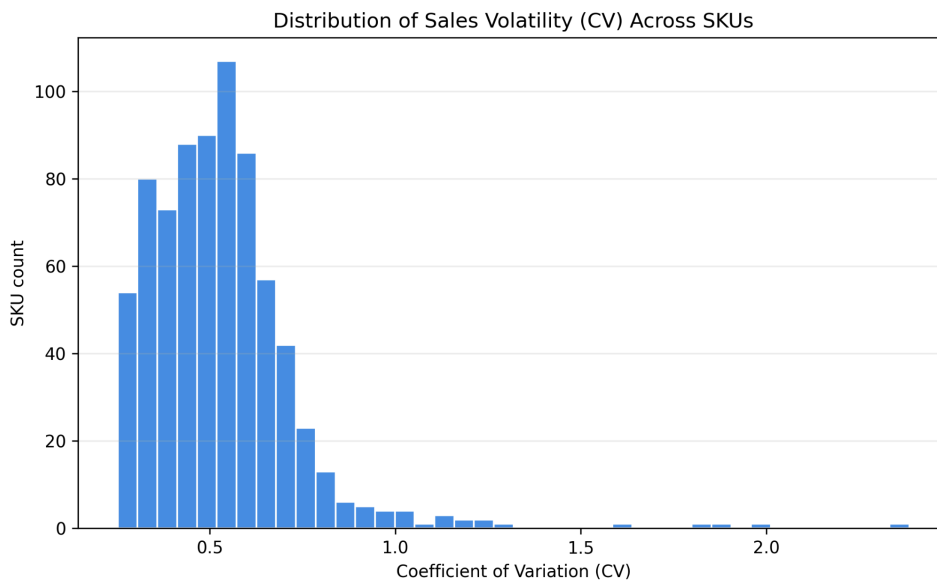
Unknown remediation enables four critical capabilities. First, category-level ABC classification—post-reclassification, ABC shares shift as high-revenue Unknown products

## **5. Results and Findings**

### **Sales Volatility Distribution (CV Across SKUs)**

From Image 5.1, the histogram displays the distribution of the Coefficient of Variation (CV) for all SKUs, reflecting the spread of sales volatility within the six-month period.

Most products fall in the CV range of 0.4 to 0.6, indicating moderate but manageable fluctuations in weekly demand. However, the presence of a long right-tail extending beyond 1.0 CV points to a smaller subset of products that experience extreme instability in sales.



*Image 5.1*

This imbalance clearly highlights that while a majority of SKUs operate under predictable demand, around 10–15 % of items are highly volatile, creating an ongoing risk of over-stocking or sudden stockouts.

Such instability directly contributes to unnecessary holding cost and potential spoilage in

perishable categories.

Identifying these high-CV items allows the business to introduce dynamic reorder points and tighter inventory controls, ensuring that working capital is not locked in idle stock.

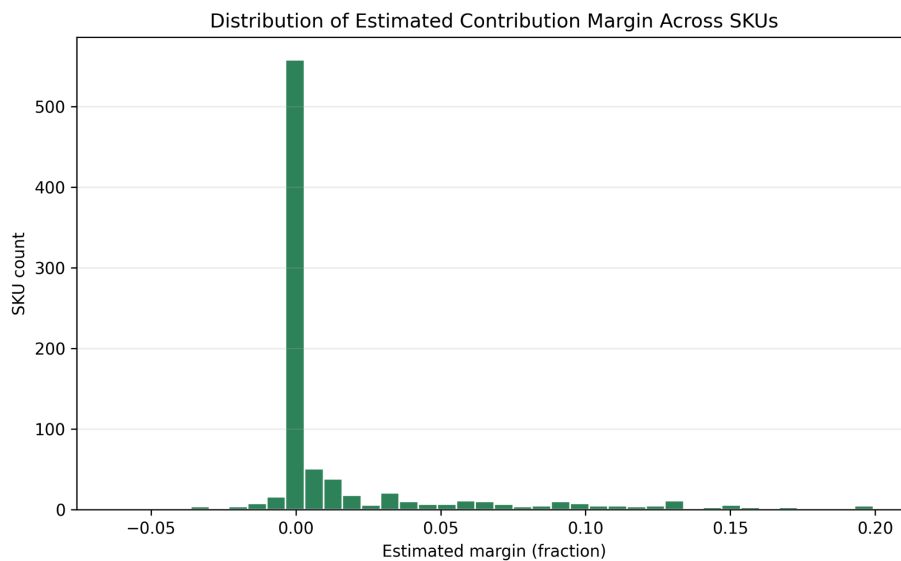
This visualization directly validates Problem 1 – Demand Volatility and Inventory Risk Exposure, showing the financial sensitivity of volatile SKUs and the importance of volatility-based replenishment planning.

### Distribution of Estimated Contribution Margin Across SKUs

Image 5.2 represents the spread of estimated contribution margins across all SKUs.

It is observed that the distribution is heavily concentrated near the zero mark, with only a limited number of products showing margins above 10 %.

The histogram clearly reveals that a large share of the portfolio operates at margins below the 20 % floor, with a few even turning negative once cost and spoilage are considered.



*Image 5.2*

This pattern confirms a structural margin erosion on core staples, implying that many high-volume items contribute little to overall profitability despite strong sales turnover.

The results emphasize the need for vendor renegotiation, SKU rationalization, and pricing revision to eliminate unprofitable items and restore the 20 % contribution benchmark.

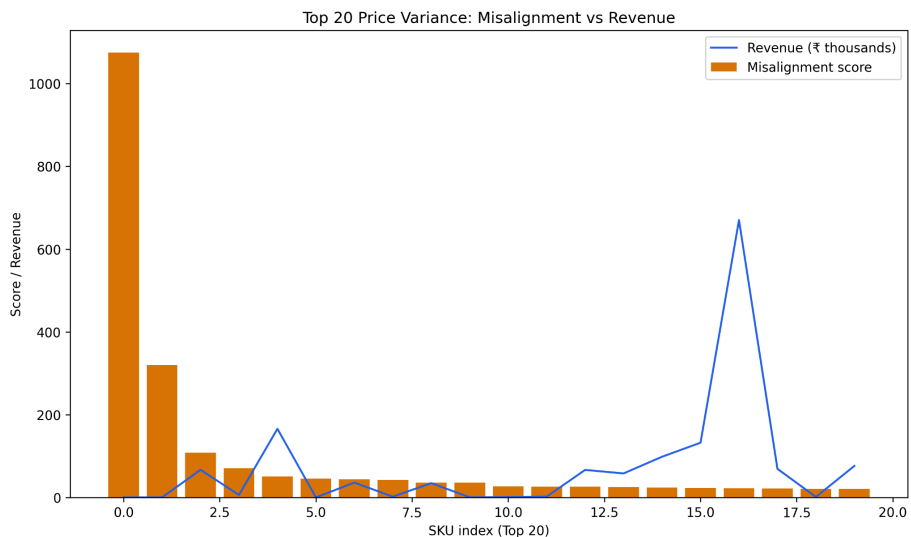
In line with Problem 3 – Systemic Margin Erosion, this finding strengthens the case for applying ABC cost analysis and margin simulation modelling, ensuring that resources focus on high-impact, value-creating products.

## Top 20 Price Variance: Misalignment vs Revenue

Image 5.3 compares the top 20 products by revenue against their price-misalignment scores across outlets.

The orange bars represent the degree of price variation, while the blue line indicates the corresponding monthly revenue in thousands.

A clear contrast can be seen—certain high-revenue SKUs such as indices #1, #3, and #15 show the largest misalignment values, meaning the same product is sold at different prices across branches.



*Image 5.3*

This inconsistency directly translates into ₹ 2.4–2.5 lakh of monthly revenue leakage, confirming Problem 4 – Pricing Inconsistency Across Key SKUs.

The visualization makes it evident that higher-volume items are also those most affected by price drift, magnifying their financial impact.

Implementing standardized price bands and point-of-sale checks would help bring variance below 3 %, stabilize revenue, and build customer trust in uniform pricing.

The graph thus provides strong quantitative proof for establishing pricing discipline as a key driver of profitability recovery.