Intelligent Inventory Optimization Solution for a Retail fashion brand of India

by Praveen Shahani

Contents

1 The Problem Statement

This problem statement is for the retail fashion indian brand.

This mismatch manifests in two primary symptoms:

- **Broken Size Runs:** High-demand sizes of a product sell out quickly while other sizes become dead stock.
- **Seasonal Overstocking:** A large surplus of trend-based merchandise remains unsold after its peak selling period.

These symptoms create critical business bottlenecks that directly hinder growth and profitability by:

- **Trapping Capital:** Tying up significant working capital in unproductive, slow-moving inventory.
- **Diluting Margins:** Forcing deep, margin-eroding markdowns to clear the leftover stock.
- **Creating Lost Sales:** Resulting in stockouts of popular items, directly sacrificing revenue and disappointing customers.

1.1 Key Factors Affecting the Problem

The solution is built upon a comprehensive understanding of four key areas. These factors are not just inputs; they represent the complex, interconnected dynamics of the retail environment that the model must navigate.

- **Core Performance & Product Data**
  This is the foundation, focusing on the products themselves.
  - **Historical Sales Performance:** What has sold well in the past? This is our primary measure of demand.
  - **Product Attributes:** The specific characteristics of each item (category, style, color, fabric, price point) that allow us to understand *why* it sells.
  - **Profitability Metrics:** The unit cost, selling price, and anticipated end-of-season holding/markdown cost of every item to ensure decisions maximize profit, not just sales.
- **Inferred Customer Demand Patterns**
  This is how we understand customer behavior using only sales data.
  - **Seasonality & Trends:** Identifying predictable sales peaks (like festive seasons or winter) and emerging fashion trends.
  - **Size Preference (The "Size Curve"):** Calculating the historical sales percentage for each size within a product category to create an ideal size assortment.

  - **Geographic Variation & Store Clustering:** Recognizing that buying behavior differs by location. We group stores into clusters (e.g., "Coastal & Resort," "Urban Metro") based on their sales patterns to create tailored local assortments for places like Goa, Bangalore, and Delhi.

- **Product Portfolio Interactions**
  This layer analyzes how products affect each other's sales.
  - **Cannibalization:** The risk that a new product will simply steal sales from a similar existing item, leading to no net growth and excess stock.
  - **The Halo Effect:** Identifying "hero" products that drive store traffic and increase the sales of other items in a customer's shopping basket.

- **Operational & Business Constraints**
  These are the real-world rules that ensure the model's output is practical.
  - **Physical Store Capacity:** The solution must fit within the physical stockroom and shelf space available at each store.
  - **Overall Procurement Budget:** The total cost of inventory must not exceed the seasonal budget.
  - **Supplier Terms:** The plan must adhere to conditions like Minimum Order Quantities (MOQs).
  - **Brand & Merchandising Strategy:** The model incorporates rules to protect brand-defining "must-stock" items and respect visual merchandising limits on the number of unique products that can be effectively displayed.

1.2 Data Requirements

To address all the specified factors, we need four primary data tables and two key business parameters. This data forms the essential fuel for the entire analytical and modeling process.

1.2.1 Transaction Data (`Transaction_Data`)

This table is the most critical asset. It provides the raw material for understanding historical performance, inferring all customer demand patterns, and analyzing product interactions.

1.2.2 Product Master Data (`Product_Master`)

This table defines "what" we sell. It covers the core product attributes and profitability metrics.

Column Name	Data Type	Description	Example
transaction_id	Text	Unique ID for the sale/basket.	TRN-20251013-001
timestamp	Datetime	Exact date and time of the sale.	2025-10-13 14:35:10
store_id	Text	Unique ID for the store.	DEL-CP-01
sku_id	Text	Unique ID for the product.	WM-TSH-BL-GRPH-042
size	Text	The specific size of the item sold.	M
quantity_sold	Integer	Number of units sold.	1
price_per_unit	Decimal	Actual price the item sold for.	799.50

Column Name	Data Type	Description	Example
sku_id	Text	Primary Key to link to sales.	WM-TSH-BL-GRPH-042
product_name	Text	Human-readable name.	Men's Blue Graphic Tee
category	Text	Main product grouping.	T-Shirt
style	Text	Sub-grouping (e.g., Fit, Neck).	Slim Fit, Round Neck
color	Text	Primary color of the item.	Blue
fabric	Text	Material of the item.	Cotton
unit_cost	Decimal	Cost to acquire one unit.	250.00
mrp	Decimal	Manufacturer's Retail Price.	999.00

1.2.3 Store Master Data (Store_Master)

This table defines "where" we sell. It's essential for the geographic analysis and understanding physical constraints.

Column Name	Data Type	Description	Example
store_id	Text	Primary Key to link to sales.	DEL-CP-01
store_name	Text	Human-readable name.	Westside - Connaught Place
city	Text	City where the store is located.	Delhi
state	Text	State where the store is located.	Delhi
inventory_capacity	Integer	Total inventory capacity (units).	50000

1.2.4 Supplier & Business Rules Data

This information covers the operational and business constraints.

Column Name	Data Type	Description	Example
sku_id	Text	The product this rule applies to.	WM-TSH-BL-GRPH-042
moq	Integer	Minimum units required per order.	200

Table: Supplier_Data

Single Input Parameters:

Total_Seasonal_Budget: A single decimal value (e.g., 50,00,00,000).

List_of_Must_Stock_SKUs: A simple list of sku_ids provided by the brand team.

---

2 The "Crawl, Walk, Run" Implementation Framework

This framework is a structured journey from raw data to a fully intelligent, profit-maximizing engine.

2.1 Phase 1: Crawl - Foundational Insights (Months 1-3)

Goal: Establish a single source of truth and identify immediate, low-risk opportunities.

2.1.1 Diagnostic Analysis (ABC)

We rank every product by its actual profit contribution to identify heroes and laggards.

Mathematical Justification: We calculate the total profit for each SKU i.

$$\text{Total Profit}_i = \sum_{\text{transactions}} (\text{price\_per\_unit} - \text{unit\_cost}_i) \times \text{quantity\_sold} \tag{1}$$

2.1.2 National Size Curve

We calculate the historical demand for each size to create a data-driven ratio.

Mathematical Justification: For any product category C, we calculate the sales share for each size j.

$$\text{Size Share}_{j,C} = \frac{\sum \text{quantity\_sold for size } j \text{ in } C}{\sum \text{quantity\_sold for all sizes in } C} \tag{2}$$

2.2 Phase 2: Walk - Predictive Pilot (Months 4-9)

Goal: To build the first version of the predictive model and prove its value in a controlled pilot.

2.2.1 Store Clustering (K-Means)

We group stores into "personas" based on their sales DNA.

Mathematical Justification: The algorithm's objective is to minimize the within-cluster sum of squares (WCSS), grouping store sales vectors ($\mathbf{x}$) into K clusters (S_k).

$$\underset{S}{\operatorname{argmin}} \sum_{k=1}^{K} \sum_{\mathbf{x} \in S_k} \|\mathbf{x} - \boldsymbol{\mu}_k\|^2 \tag{3}$$

2.2.2 Market Basket Analysis (Halo & Cannibalization)

We analyze transaction data to find product relationships.

Definition (Halo Effect): The **Halo Effect** occurs when the sale of one product (the "hero") significantly increases the likelihood or value of other products being sold in the same transaction.

Mathematical Justification (Halo Effect): We identify "hero" products by calculating the **Lift** of one product on another. A Lift value greater than 1 indicates a positive association.

$$\text{Lift}(A \to B) = \frac{\text{P}(A \cap B)}{\text{P}(A) \times \text{P}(B)} > 1 \tag{4}$$

Definition (Cannibalization): **Cannibalization** occurs when the introduction of a new product reduces the sales of another, existing product.

Mathematical Justification (Cannibalization): We identify substitutes by looking at the **Pearson correlation coefficient** (ρ) of their sales time series. A strong negative correlation ($\rho \approx -1$) signals cannibalization.

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{5}$$

2.2.3 Demand Planning & Forecasting

Logic: We build a predictive model to forecast sales for the upcoming season.

Mathematical Justification (Existing Products): We use a Time-Series Forecasting model (e.g., SARIMA).

$$\text{Sales}_t = f(\text{Trend}_t, \text{Seasonality}_t, \text{Past Sales}_{t-n}) + \text{Error}_t \tag{6}$$

Mathematical Justification (New Products): For new items, we use an attribute-based model.

$$\text{Forecast}_{i_{\text{new}}} = f(\beta_{A_1}, \beta_{A_2}, \ldots, \beta_{A_p}) \tag{7}$$

2.3 Phase 3: Run - Scale & Automate (Months 10-18)

Goal: To expand the proven model across the business and embed it as a core operational engine.

2.3.1 Full-Scale Optimization Model

We deploy the complete mathematical model, a Mixed-Integer Linear Program (MILP), to find the most profitable inventory plan.

Notations Definitions (The Model's Legend)

Sets Indices

$i \in I$ The set of all unique **SKUs** (products).

$I_{core} \subset I$ The subset of SKUs classified as **core/evergreen**.

$j \in J$ The set of all **sizes**.

$k \in K$ The set of all **store clusters**.

Parameters (Data Inputs)

p_{ij} The retail **price** of SKU i in size j.

c_{ij} The unit **cost** to procure one new unit of SKU i in size j.

h_{ij} The **holding/markdown cost** for one unsold unit of SKU i.

D_{ijk} The forecasted **demand** for SKU i, size j, at cluster k.

β_i The **Halo Effect bonus** from selling one unit of hero SKU i.

I^0_{ijk} The **initial on-hand inventory** of SKU i, size j, at cluster k.

C_k The physical inventory **capacity** of an average store in cluster k.

B The total seasonal **procurement budget**.

v_{ij} The physical **volume** of one unit of SKU i.

M_i The **Minimum Order Quantity (MOQ)** for SKU i.

σ_{ijk} The **standard deviation of forecast error** for SKU i, size j, at cluster k.

Z_i The **service level factor** for SKU i (e.g., 1.65 for a 95% service level).

Decision Variables (The Answers We Seek)

x_{ijk} The **quantity to purchase** of SKU i, size j, for cluster k.

y_{ijk} A **binary variable** (1 if we purchase, 0 if not) for the decision to purchase new stock.

s_{ijk} The **expected units sold** of SKU i, size j, at cluster k.

Objective Function: The goal is to maximize the total, fully-adjusted profit across the entire business.

$$\text{Maximize Z} = \sum_{i,j,k} (p_{ij}s_{ijk} - c_{ij}x_{ijk} - h_{ij}((I^0_{ijk} + x_{ijk}) - s_{ijk}) + \beta_i s_{ijk}) \tag{8}$$

2.3.2 Interpretation of the Objective Function

The objective function is the "brain" of the model, perfectly aligning its mathematical goal with our business strategy. It balances four key financial levers:

Term 1: $p_{ij}s_{ijk}$ **(Revenue)** This term seeks to maximize total revenue by rewarding the model for every unit it expects to sell.

Term 2: $-c_{ij}x_{ijk}$ **(Cost of Goods)** This term subtracts the cost of all new inventory purchased, ensuring the model is conscious of procurement expenses.

Term 3: $-h_{ij}((I^0_{ijk} + x_{ijk}) - s_{ijk})$ **(Penalty for Leftovers)** This is the critical component that solves our overstocking problem. It applies a direct financial penalty for every single unit of inventory (both initial and newly purchased) that is projected to be left unsold at the end of the season.

Term 4: $+\beta_i s_{ijk}$ **(Strategic Bonus for Hero Products)** This term adds a strategic bonus for selling "hero" products, telling the model that their value extends beyond their own price tag by driving larger baskets and store traffic.

Key Constraints: The solution must obey all real-world limitations.

$$s_{ijk} \leq (I^0_{ijk} + x_{ijk}) \quad \forall i,j,k \quad \text{(Sales Unit} \leq \text{Total Stock)} \tag{9}$$

$$s_{ijk} \leq D_{ijk} \quad \forall i,j,k \quad \text{(Sales} \leq \text{Demand)} \tag{10}$$

$$\sum_{i,j}(v_{ij} \cdot (I^0_{ijk} + x_{ijk})) \leq C_k \quad \forall k \quad \text{(Store Capacity)} \tag{11}$$

$$\sum_{i,j,k}(c_{ij} \cdot x_{ijk}) \leq B \quad \text{(Total Budget)} \tag{12}$$

$$x_{ijk} \geq M_i \cdot y_{ijk} \quad \forall i,j,k \quad \text{(Supplier MOQs)} \tag{13}$$

$$x_{ijk} \leq M \cdot y_{ijk} \quad \forall i,j,k \quad \text{(Logical Gate for Purchase)} \tag{14}$$

$$(I^0_{ijk} + x_{ijk}) \geq D_{ijk} + Z_i \cdot \sigma_{ijk} \quad \forall i \in I_{core}, \forall j,k \quad \text{(Service Level Constraint)} \tag{15}$$

Constraint Explanations: Each constraint translates a critical business rule into a mathematical statement, ensuring the model's output is realistic, actionable, and aligned with our strategy.

- **Sales Unit $\leq$ Total Stock:**
  - **Business Logic:** The number of units sold (s_{ijk}) cannot exceed the total available stock. Total stock is the sum of what you already have (initial inventory, I^0_{ijk}) and the new inventory you are purchasing (x_{ijk}).
  - **Simple Analogy:** The "Empty Shelf Rule." If you start the day with 5 shirts and get a delivery of 10 more, you have 15 in total. You can sell a maximum of 15 shirts.
  - **Implication:** This grounds the model's revenue projections in physical reality.
- **Sales $\leq$ Demand:**
  - **Business Logic:** Expected sales (s_{ijk}) cannot exceed the forecasted demand (D_{ijk}). Even with ample stock, you can only sell what customers actually want to buy.
  - **Simple Analogy:** The "Movie Ticket Rule." A theater has 200 seats (stock), but if only 80 people want to see the movie (demand), you will only sell 80 tickets.
  - **Implication:** This forces the model to respect the data-driven forecast and prevents overly optimistic planning.
- **Store Capacity:**
  - **Business Logic:** The total physical volume of all products in a store must fit within the store's capacity (C_k). This includes both the initial inventory and the newly arriving stock.
  - **Simple Analogy:** The "Suitcase Rule." The total volume of clothes already in your suitcase plus the new clothes you want to add cannot exceed the suitcase's total capacity.

- **Implication:** This ensures the allocation plan is physically executable and prevents the logistical error of sending inventory to a store with no space.

- **Total Budget:**
  - **Business Logic:** The total cost of all the *new* inventory purchased (x_{ijk}) cannot exceed the seasonal procurement budget (B). Note that this constraint rightly applies only to new purchases, not existing stock.
  - **Simple Analogy:** The "Shopping Budget Rule." You have a 50,000 budget for new clothes. The total cost of everything you decide to buy cannot be more than 50,000.
  - **Implication:** This ensures the final plan is financially sound and aligns with the company's financial strategy.

- **Supplier MOQs & Logical Gate:**
  - **Business Logic:** These two constraints work together. They state that *if* you decide to purchase an item ($y_{ijk} = 1$), the quantity you buy (x_{ijk}) must be at least the supplier's Minimum Order Quantity (M_i). If you don't purchase it ($y_{ijk} = 0$), the quantity must be zero.
  - **Simple Analogy:** The "Wholesale Club Rule." You can't buy just one can of soda; you must buy the whole case of 24 (the MOQ). You either buy the case or you buy nothing.
  - **Implication:** This ensures the purchase orders generated by the model are realistic and actionable for the procurement team.

- **Service Level Constraint (for Core Items):**
  - **Business Logic:** This is a sophisticated rule applied only to essential, low-risk core items ($i \in I_{core}$). It ensures that the total stock is high enough to cover not just the average expected demand (D_{ijk}), but also a **safety stock buffer** to account for demand uncertainty. The size of this buffer is determined by a service level factor (Z_i, e.g., 95 percent service level) and the demand's variability (σ_{ijk}).
  - **Simple Analogy:** The "Milk in the Fridge Rule." For an essential item like milk, you don't just buy enough for your average daily use. You buy a little extra (safety stock) to make sure you don't run out if you have guests. The more important it is not to run out (a higher Z_i), the more extra milk you buy.
  - **Implication:** This strategically protects the business from stockouts on its most important, high-demand products, safeguarding revenue and customer satisfaction.

Variable Definitions:

$$
\begin{aligned}
&x_{ijk} \in \mathbb{Z}^+ && \forall i, j, k && \text{(Purchase quantity is a non-negative integer)} \\
&y_{ijk} \in \{0, 1\} && \forall i, j, k && \text{(Purchase decision is binary)} \\
&s_{ijk} \geq 0 && \forall i, j, k && \text{(Sales is a non-negative number)}
\end{aligned}
$$

2.3.3 Interpreting the Supplier MOQ Logic Gate

The two constraints governing supplier minimums work together as a single "on/off" switch. This logic forces the model to make a realistic, strategic decision about whether to commit to a full MOQ for a product.

Table 1: The MOQ Logic Gate in Action (Example MOQ = 200)

Scenario	Constraint 1: $x \geq M_i \cdot y$	Constraint 2: $x \leq M \cdot y$	The Resulting Rule
Decision: BUY The model sets $y = 1$.	$x \geq 200 \times 1$ (Purchase must be at least 200 units).	$x \leq M \times 1$ (Purchase is allowed up to a large limit).	The quantity to purchase, x, must be 200 or more.
Decision: DO NOT BUY The model sets $y = 0$.	$x \geq 200 \times 0$ (Purchase must be at least 0).	$x \leq M \times 0$ (Purchase must be less than or equal to 0).	The only number satisfying both rules is $x = 0$. No purchase is made.

2.3.4 Risk-Adjusted Portfolio Strategy

We apply different rules for different product types to manage the risk of the long lead time.

Mathematical Justification: For high-risk fashion items, we artificially increase their **holding cost parameter** (h_{ij}) in the model. This makes the model inherently more conservative, as the "Penalty for Leftovers" in the objective function becomes much higher for these specific items, discouraging overstocking.

---

3 Actionable Outputs of the Optimization Model

The output of the model is not just a number; it is a set of concrete, actionable plans designed to be directly implemented by key business teams.

3.1 The Master Purchase Plan

This is the aggregate, business-wide procurement plan, serving as the master "shopping list" for the season.

Intended For: Procurement and Buying Teams.

Content: A detailed list of every SKU and size to be procured from suppliers, along with the total quantity needed across all stores. It is optimized to respect supplier MOQs and the overall budget.

Business Value: It provides a single, data-driven directive for all purchasing, ensuring capital is allocated to the most profitable mix of products.

Table 2: Example: Master Purchase Plan

SKU ID	**Size**	**Total Units to Procure**
WM-TSH-BL-GRPH-042	M	12,500
WM-TSH-BL-GRPH-042	L	11,200
AP-DRS-FLRL-MAXI-007	S	5,000

3.2 The Store Allocation Plan

This plan details the "last-mile" distribution of inventory, specifying where each procured item should be sent.

Intended For: Logistics and Supply Chain Teams.

Content: A granular breakdown of how many units of each SKU and size should be allocated to each specific store cluster.

Business Value: This executes our tailored assortment strategy, ensuring that the right products in the right sizes arrive at the right local markets to meet specific customer demand.

Table 3: Example: Store Allocation Plan

SKU ID	**Size**	**Store Cluster**	**Units to Allocate**
WM-TSH-BL-GRPH-042	M	Urban Metro	72
WM-TSH-BL-GRPH-042	M	Coastal & Resort	35
AP-DRS-FLRL-MAXI-007	S	Urban Metro	50
AP-DRS-FLRL-MAXI-007	S	Coastal & Resort	15

3.3 The Projected Performance Report

This is the strategic financial forecast of the model's recommended plan.

Intended For: Management, Finance, and Merchandising Teams.

Content: A projected Profit & Loss (P&L) statement for the recommended assortment, including forecasted sales, revenue, cost of goods, and expected end-of-season leftover inventory.

Business Value: It sets realistic, data-backed financial targets for the season and provides a clear benchmark to measure actual performance against, quantifying the expected ROI of the plan.

Table 4: Example: Projected Performance Report

Metric	**Projected Value (INR)**
Total Revenue	126 Cr
Cost of Goods Sold	59 Cr
Gross Margin %	53.2%
Projected Leftover Stock (Value)	12.7 Cr