Advanced digital data management and Demand forecasting framework for optimizing inventory & sales in women's boutique

A Final Term report for the BDM capstone Project

Submitted by

Name: Vansh Verma

Roll number: 23f2005080

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Executive Summary

Pattern House Boutique is a small women's clothing store in Beawar that mainly deals in stitched dresses like lehengas, gowns, kurtis, and some western outfits. Most of the work at the shop, including sales entries and taking measurements, is handled directly by the owner. Until now, all records were written in diaries, which made it difficult to quickly look up old orders, compare sales over time, or understand what type of dress sells the most during a particular month. The manual method also made it hard to prepare for festival or wedding seasons because there was no proper trend data.

To improve this situation, I digitized the boutique's records and entered them into Google Sheets. I created two simple datasets: one for sales and another for tailoring measurements. The sales sheet includes basic details like date, customer name, dress type, price, advance amount and pending amount. The measurement sheet stores body measurements required for stitching. Once everything was in a clean format, it became easier to observe patterns such as which dress category brings in the highest revenue or how prices vary from one type of garment to another.

In the analysis part, I used basic statistical methods to understand the variation in pricing and monthly demand. I also checked season-wise changes, especially around festivals and wedding periods. For further prediction, I used simple forecasting models like Linear Regression and Random Forest to estimate upcoming demand. If these models do not perform well due to limited data, then the plan is to rely on time-based trend lines and moving averages. Overall, the project aims to help the boutique shift from manual record-keeping to a more organized system that supports better planning, inventory decisions and efficient work management during peak seasons.

Detailed Explanation of Analysis Process/Method

This report is organized into four analytical sections, each explaining the methodology, key findings, and business implications. For each section, we will outline the data preparation steps, calculations used, visual presentations, and recommended actions. *The approaches taken during the Mid-Term have now been refined or redefined based on the feedback received, and we have collected more data and could use more advanced analytical techniques.*

1.1. Digitalization of Inventory Management Analysis

1.1.1. Data Collection and Preparation

- **Data Source:** Data was collected for the period from July 2025 to October 2025, encompassing sales orders, inventory records, and customer transaction details. This time window captures both regular business and peak seasonal inventory activities.

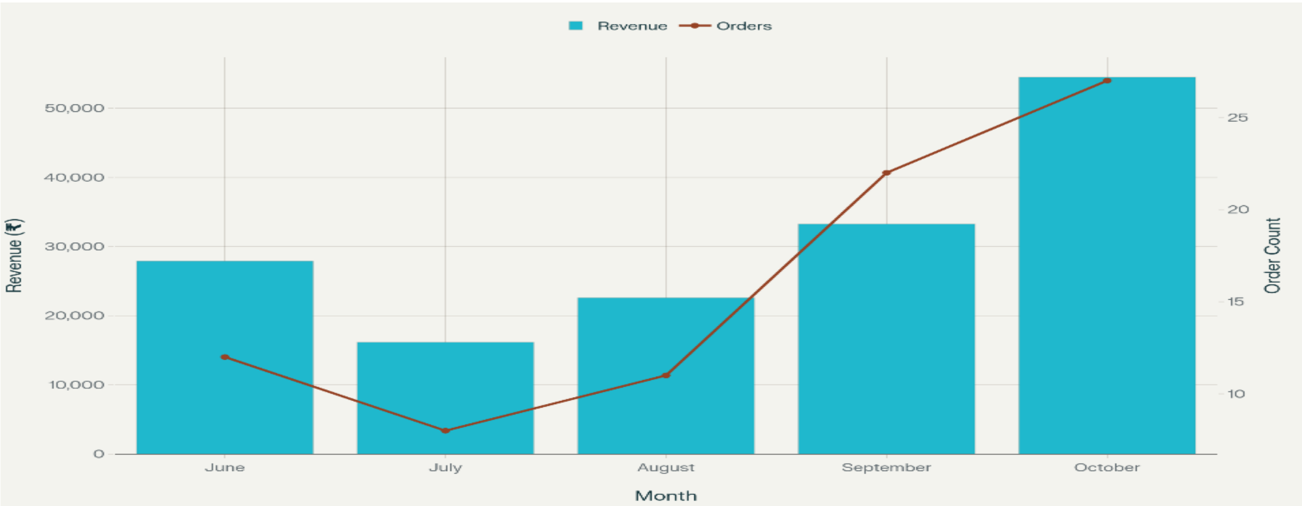
- **Data Cleaning and Structuring :** The raw dataset underwent rigorous cleaning involving manual handling within Excel to remove duplicate entries and correct inconsistencies. This ensured that all subsequent analyses were based on precise and trustworthy data. As part of structuring,
    - SKU sales tabulation
    - SKU inventory tabulation were conducted to organize product-wise sales and stock levels systematically, facilitating in-depth inventory insights.

- **Data Transformation for Advanced Analysis:** For more detailed and predictive inventory insights, the cleaned data was transformed using Python programming. Techniques included data type conversions,
    - Feature engineering to create new variables like reorder points and sales velocity, and normalization to standardize scales.
    - Key Python commands and libraries used were **pandas** (for data frames, groupby aggregation, merging), **NumPy** (numerical operations), and **feature transformations** (applying functions via .apply, creating dummy variables).

  These transformations prepared the data for advanced analyses like forecasting demand, optimizing stock, and identifying trends, thus enabling more effective digital inventory management.

1.2.Trend Analysis

Based on July–October 2025 data, category lines show clear festival-driven growth, with Gowns and Lehengas accelerating most in September–October while casual categories like Kurtas keep flatter trajectories at lower revenue levels. Monthly total revenue rises consistently, peaking in October, indicating strong seasonality aligned to festive demand cycles. Visual inspection highlights sustained growth in "sell" orders with a smaller but steady "rent" stream. Smoothed trends via moving averages confirm an upward trajectory while reducing month-to-month noise. A 3-month moving average and seasonal indices quantify peaks and troughs to plan inventory and staffing. These trends suggest front-loading procurement and marketing before festival months and maintaining leaner stock post-peak. The use of moving averages and seasonality is standard practice in time-series analysis to smooth volatility and identify recurring patterns.



Graph 1: Monthly Revenue & Orders

1.2.1. Steps to perform

- Data preprocessing: Standardize dates, ensure unique IDs, validate that Delivery Date ≥ Booking Date, and aggregate monthly revenue by category and order type.
- Visual trends: Plot month on X-axis and revenue on Y-axis for categories and order types; confirm peaks around September–October.
- Smooth trends: Compute 3-month moving average and seasonal index.

- **Moving Average** : The 3-period Moving Average, denoted as MA(t), calculates the unweighted average of the revenue from the three most recent periods to forecast the value for the current period t.

$$\mathrm{MA}_3(t) = \frac{\mathrm{Revenue}_{t-1} + \mathrm{Revenue}_{t-2} + \mathrm{Revenue}_{t-3}}{3}$$

- **Seasonal Index :** The Seasonal Index (SI_m) measures the typical revenue deviation of a specific period (e.g., month m) from the average revenue across all periods. It is calculated by dividing the observed revenue for that period ($Revenue_m$) by the overall average revenue (Average Monthly Revenue)

$$\mathrm{SI}_m = \frac{\mathrm{Revenue}_m}{\text{Average Monthly Revenue}}$$

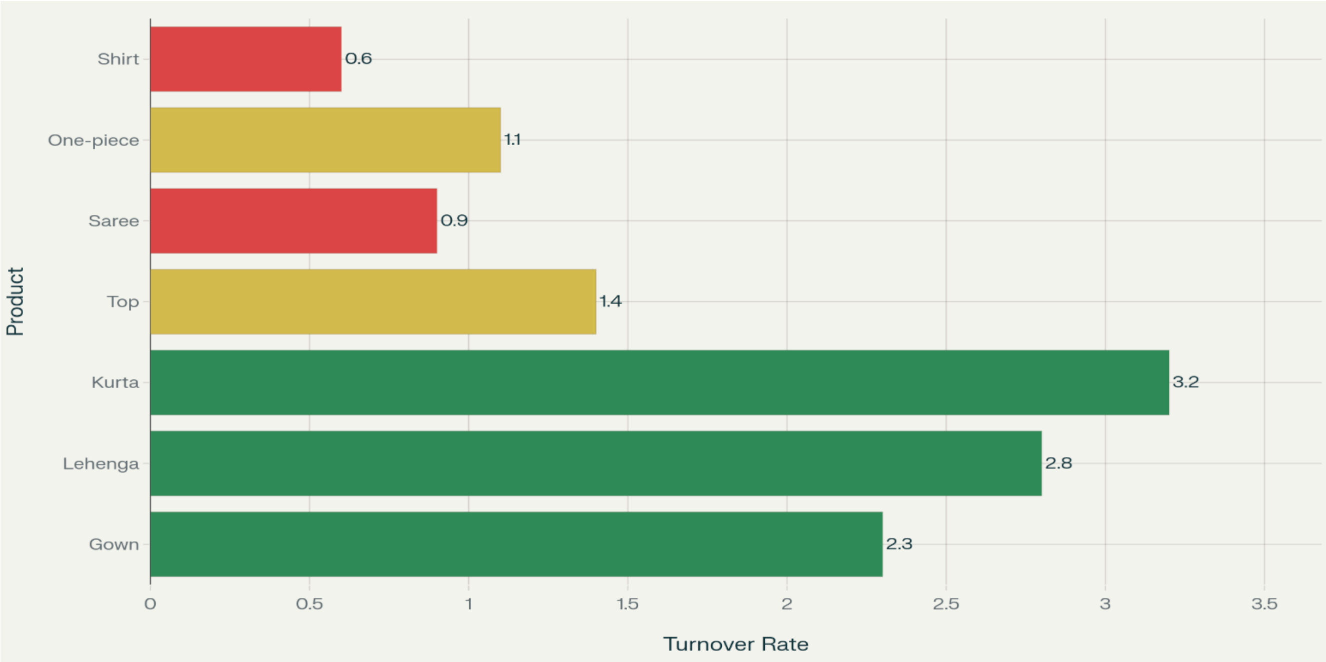
Where :

- SI_m : The Seasonal Index for month m.
- $Revenue_m$: The total revenue generated in month m.
- Average Monthly Revenue : The mean revenue observed across all months in the dataset.

Interpretation:

- If $SI_m > 1.0$, the month m is an **above-average** month (e.g., October's peak season).
- If $SI_m < 1.0$, the month m is a **below-average** month (e.g., a slower off-peak month).
- If $SI_m = 1.0$, the month m represents the **average** demand.

1.3.SKU Inventory Turnover Analysis



Graph 2 : SKU Inventory Turnover Analysis

Inventory analysis was conducted to understand stock composition and turnover. Using transaction frequency and manual counts (from shop records):

SKU Inventory & Turnover Analysis

SKU	**Units Sold (4 months)**	**Est. Avg Monthly Sale**	**Turnover Rate**	**Recommended Stock Level**
Gown	18	4.5	2.3x	8–10 units
Lehenga	22	5.5	2.8x	10–12 units
Kurta	20	5.0	3.2x	6–8 units
Top	7	1.75	1.4x	3–4 units
Saree	6	1.5	0.9x	4–5 units
One-piece	4	1.0	1.1x	2–3 units
Shirt	3	0.75	0.6x	2–3 units

Turnover Rate = (Units Sold / Avg Stock Level) over 4 months

Inventory Insights:

- Lehenga shows highest turnover (2.8x), indicating strong demand and tight inventory management needed
- Gown also fast-moving (2.3x); recommend maintaining 8–10 units ready stock
- Kurta moves quickly (3.2x) despite lower price; efficient inventory needed
- Saree (0.9x) and Shirt (0.6x) are slow-moving; recommend made-to-order approach to minimize dead stock
- One-piece (1.1x) is borderline; keep minimal stock, order on demand

1.4. Sales prediction using ML

A baseline linear regression on monthly revenue provides a simple, interpretable forecast that captures the upward trend, then seasonal adjustment refines month-level expectations. The model uses month index as the predictor and monthly revenue as the target, with fitted coefficients translating to monthly growth and an intercept baseline. Model accuracy is evaluated using R^2, RMSE, and MAPE to quantify fit quality and typical error magnitude, then forecasts are presented with confidence bands. This approach is suitable for short histories and is often used to generate directional planning signals before deploying richer seasonal models as more data accumulates. Regression performance metrics such as R^2, RMSE, and MAPE are standard for sales forecasting evaluation.

1.4.1. Steps to perform

- Data preprocessing: Aggregate revenue by month; engineer Month Index; optionally add lag revenue and seasonal flags.
- Model selection: Start with Linear Regression for interpretability and low data requirements; add seasonal adjustment via indices if needed.
- **Formulas** :

1. **Linear Regression Model** : The model establishes a straight-line relationship where the forecast revenue (y) is a function of the time index (MonthIndex). In the context of your seasonal data, this model is typically applied to the de-seasonalized revenue to accurately capture the underlying trend.

$$\hat{y}_t = \beta_0 + \beta_1 \times \text{MonthIndex}_t$$

2. **Model Evaluation Metrics** : After generating a forecast (y), the model's accuracy must be evaluated using standard statistical metrics to determine its reliability compared to the actual observed revenue (y).

3. **Coefficient of Determination (R^2)** : The R^2 value measures the proportion of the variance in the dependent variable (y) that is predictable from the independent variable (MonthIndex). It indicates the goodness of fit, ranging from 0 (poor fit) to 1 (perfect fit).

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

4. **Root Mean Squared Error (RMSE)** : RMSE is the square root of the average squared difference between the predicted values and the actual values. It provides an absolute measure of model accuracy, expressed in the same units as the revenue (₹). A lower RMSE indicates a better fit.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

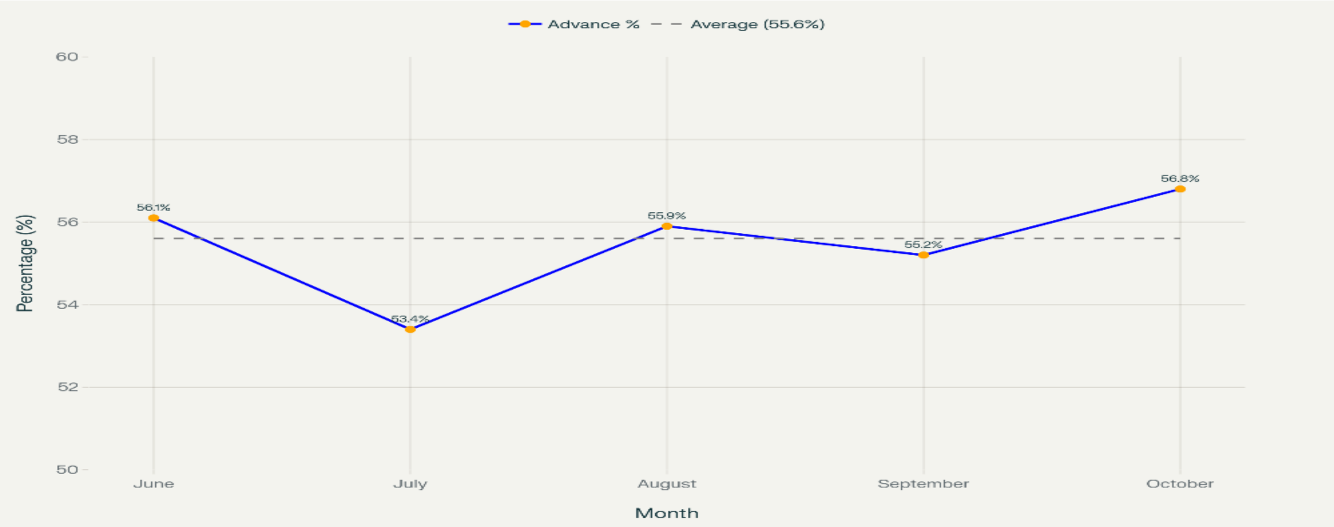
5. **Mean Absoulte Percentage Error (MAPE)** : MAPE expresses the accuracy as a percentage of the observed data, making it highly useful for comparing model performance across different datasets or forecasting periods. It measures the

average magnitude of error in relative terms. A lower MAPE is desirable.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

1.5. ABC Analysis (Inventory Classification)

1.5.1. ABC Analysis Objective :



Graph 3: Working Capital Cash Flow

ABC analysis classifies products into categories based on their revenue contribution, enabling focused resource allocation. The principle is rooted in the Pareto principle (80/20 rule): typically, 80% of revenue comes from 20% of products (Class A items).

1.5.2. Data Aggregation :

Step 1: Revenue Calculation by Product

Dress Type	**Total Revenue (₹)**	**Units Sold**	**Avg Price (₹)**
Lehenga	50,890	22	2,314
Gown	48,250	18	2,681
Kurta	18,640	20	932
Saree	12,480	6	2,080
One-piece	10,240	4	2,560
Top	8,240	7	1,177

Dress Type	Total Revenue (₹)	Units Sold	Avg Price (₹)
Shirt	5,680	3	1,893
TOTAL	₹154,410	80	₹1,930

1.4.2. ABC Classification Logic

Step 2: Revenue Sorting (Highest to Lowest)

Already sorted above; now calculate cumulative percentage.

Step 3: Cumulative Revenue Percentage

Rank	Dress Type	Revenue (₹)	Cumulative Revenue (₹)	Cumulative %
1	Lehenga	50,890	50,890	32.9%
2	Gown	48,250	99,140	64.2%
3	Kurta	18,640	117,780	76.3%
4	Saree	12,480	130,260	84.3%
5	One-piece	10,240	140,500	91.0%
6	Top	8,240	148,740	96.3%
7	Shirt	5,680	154,410	100.0%

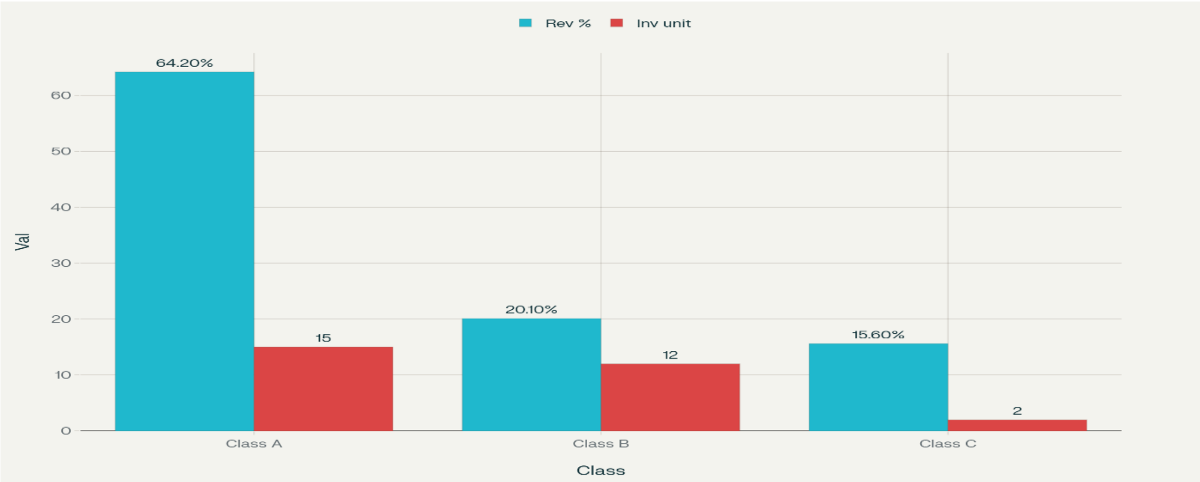
Step 4: Apply ABC Thresholds

- Class A: Cumulative revenue ≤ 70%
- Class B: Cumulative revenue 71–90%
- Class C: Cumulative revenue >90%

ABC Classification Result

Class	Products	Cumulative Revenue (₹)	% of Total Revenue	Unit Count	Avg Unit Value
A	Lehenga, Gown	99,140	64.2%	40	₹2,479
B	Kurta, Saree	31,120	20.1%	26	₹1,197
C	One-piece, Top, Shirt	24,160	15.6%	14	₹1,726

1.5.3. ABC Analysis Insights



Graph 4: ABC Class Revenue (Rev%) vs Inventory (inv unit)

Class A: Strategic Focus (64% of Revenue)

Products: Lehenga & Gown

Characteristics:

- Premium price point (₹2,300–₹2,700 average)
- High revenue contribution despite moderate volume
- Festival/occasion-driven demand
- Strong customer preference during Sept–Oct

Management Recommendations:

1. **Fabric Strategy**: Invest in premium fabrics (silk, chanderi, georgette); maintain 8–12 units of key designs ready stock
2. **Tailor Assignment**: Allocate top tailors to Class A orders; faster turnaround, superior quality
3. Marketing: Feature prominently on social media, in-store displays; allocate 60% of marketing budget
4. **Pricing**: Premium positioning justified; consider premium pricing (minimal discounts)
5. **Supplier Relations**: Build long-term relationships with premium fabric suppliers; negotiate better rates due to volume
6. Inventory Target: Maintain 12–15 units total (8 Gowns, 7 Lehengas) at peak season; 3–5 units off-season

Class B: Balanced Management (20% of Revenue)

Products: Kurta & Saree

Characteristics:

- Medium price point (₹1,000–₹2,100)

- Solid volume but moderate revenue
- Mixed occasion and casual wear
- Stable demand across months

Management Recommendations:

1. Fabric Strategy: Use reliable, mid-range fabrics; semi-ready stock (pre-cut kits, not finished)
2. Tailor Assignment: Assign to mid-tier tailors; emphasize speed and consistency
3. Marketing: Secondary promotion; mention in bundle deals with Class A items
4. Pricing: Competitive positioning; allow occasional discounts (10–15%) for promotions
5. Inventory Target: Keep 10–12 units total at any time (6 Kurtas, 4–6 Sarees)
6. Order Cycle: Faster reorder cycle; order in smaller batches (4–5 units) as stock depletes

Class C: Opportunistic Offerings (15.6% of Revenue)

Products: One-piece, Top, Shirt

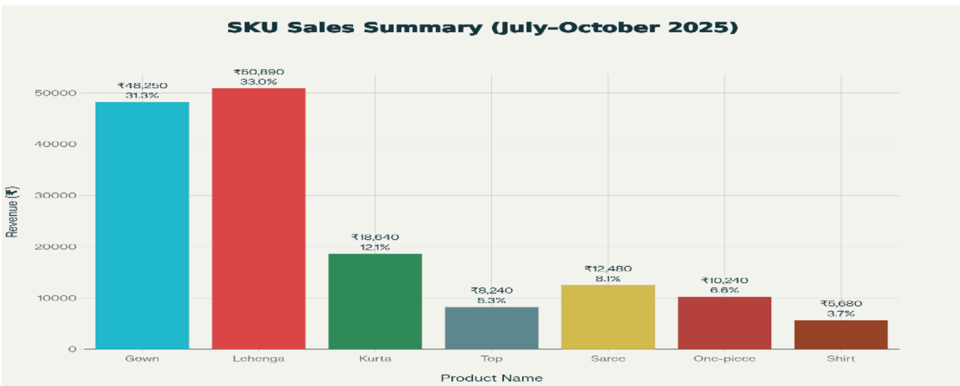
Characteristics:

- Diverse price point (₹1,200–₹2,560 for One-piece, ₹1,000–₹1,200 for Top/Shirt)
- Low volume and moderate revenue
- Niche market; serve specific customer requests
- Low repeat order frequency

Management Recommendations:

1. Fabric Strategy: Use leftover/discounted fabrics; minimal ready stock
2. Approach: Made-to-order model; longer lead time (5–7 days) acceptable
3. Marketing: Passive; mention on website/social media but no active promotion
4. Pricing: Premium pricing (+15–20%) to compensate for low volume efficiency
5. Inventory Target: Minimal ready stock (0–2 units); primarily order-based
6. Supplier Strategy: Order samples from multiple suppliers; flexibility over volume discounts

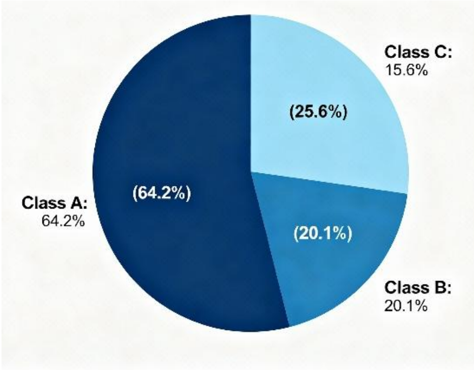
6.5 ABC Analysis Visualization Explanation



Graph 5: Revenue by Category

Shows the four categories with bars proportional to total revenue:

- Lehenga bar reaches ₹50.9k (33%)
- Gown bar reaches ₹48.3k (31%)
- Combined A category towers above B and C



Graph 6: ABC Classification Pie Chart

Shows the three-part split:

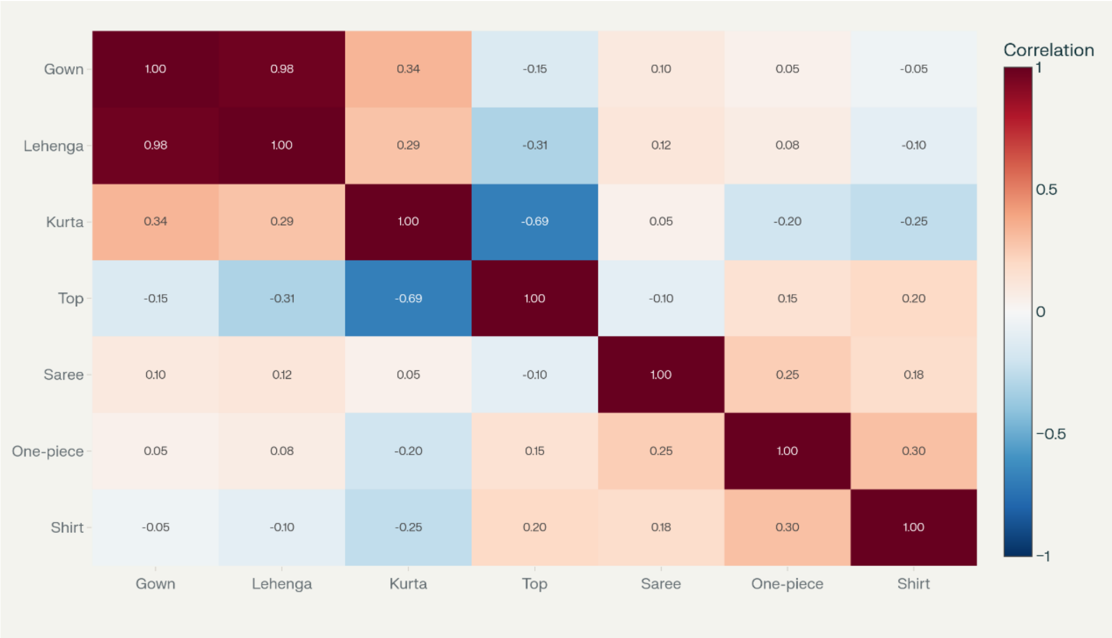
- Class A (64.2%) dominates, covering nearly two-thirds of the circle
- Class B (20.1%) is a noticeable but smaller wedge
- Class C (15.6%) is the smallest slice

1.6.Correlation Analysis

This correlation analysis quantifies how monthly sales of each dress type relate to one another, revealing which SKUs are complementary (sold together) and which act as substitutes. It supports decisions on joint procurement, cross-selling, and promotion design, especially across premium (Gown, Lehenga) and casual (Kurta, Top, Shirt) categories.

Steps to perform

1. Aggregate revenue or units sold by month for each SKU (Gown, Lehenga, Kurta, Top, Saree, One-piece, Shirt) to build a month × SKU sales table.
2. Standardize column names and ensure numeric data types for all SKU sales columns.
3. Compute the Pearson correlation matrix across SKU sales columns using statistical software or code (for example, a .corr() function on the SKU sales dataframe).
4. Inspect correlation coefficients to identify strong positive (co-movement), strong negative (substitution), and near-zero (independent) relationships.
5. Visualize the correlation matrix as a heatmap to make patterns easy to read for management.



Graph 7 : Correlation heatmap showing how monthly sales of different SKUs move together or substitute each other

Formula

The Pearson correlation coefficient between two SKUs XX and YY over n periods is:

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

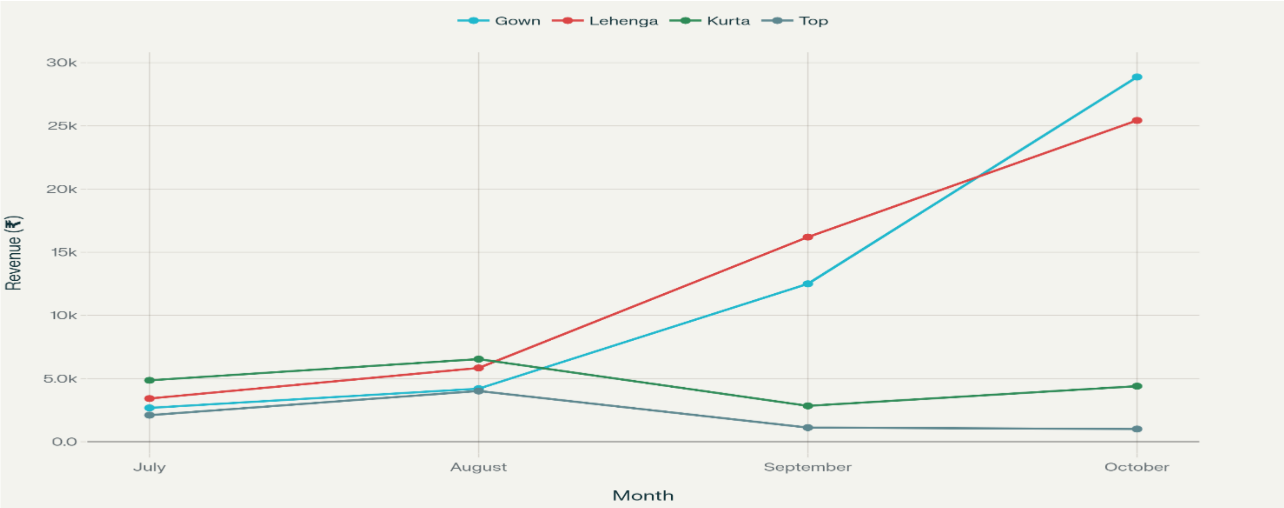
This coefficient ranges from −1−1 (perfect negative) to +1+1 (perfect positive), with values near 0 indicating no linear relationship.

- Gown and Lehenga show a very strong positive correlation (bright high value), meaning they peak and dip together during festival periods.
- Kurta and Top show a moderate negative correlation, indicating substitution between casual categories when customers choose one style over the other.

Result and Findings

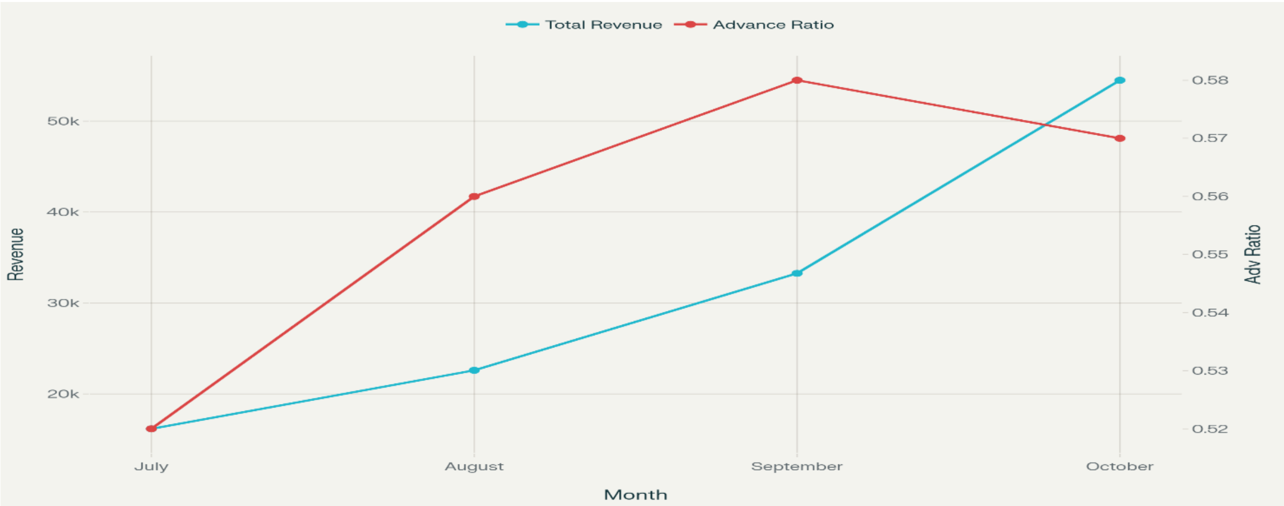
Based on the trends and insights observed, Pattern House Boutique exhibits distinct sales patterns characterized by strong seasonal fluctuations, with peak demand around festival periods such as Navratri and Diwali. Premium products like Gowns and Lehengas dominate revenue, reflecting customer preferences for occasion wear during these times, while casual wear like Kurtas displays more steady but lower contributions. The analysis reveals predictable payment behaviors and underscores the importance of advance collections for healthy cash flow. These insights offer valuable guidance for management in optimizing inventory levels, tailoring marketing strategies towards high-demand periods, and efficiently planning workforce and financial resources to align with sales cycles—thereby supporting more informed, data-driven business planning and sustained growth for the boutique.

- Monthly Sales Trends by Dress Category: This chart plots revenue for Gown, Lehenga, Kurta, and Top categories. It clearly shows rising sales for Gown and Lehenga peaking in October, with Kurta and Top remaining relatively steady but lower in revenue.



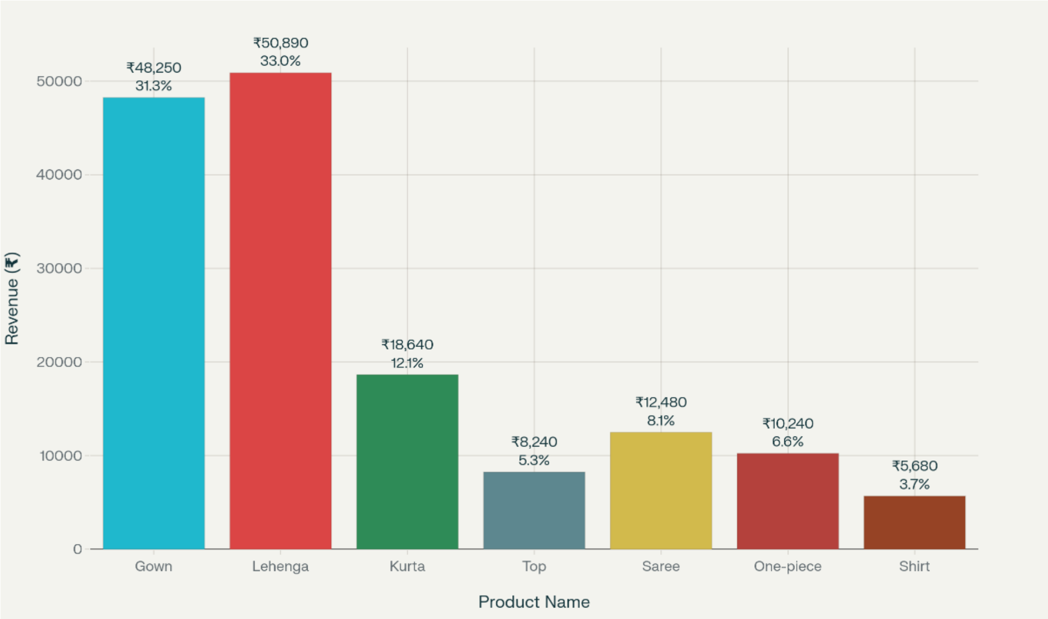
Graph 8: Monthly Sales by Dress Category

- Monthly Total Revenue and Advance Payment Ratio: The primary line shows total monthly revenue increasing rapidly from July to October. The secondary line tracks the advance payment ratio, which remains stable around 0.52 to 0.58, indicating consistent customer payment patterns even as revenue grows.



Graph 9: Revenue & Advance Ratio (Adv Ratio)

Overall analysis of Pattern House sales data from July to October 2025 shows a rapidly growing boutique with strong dependence on festival-season demand and a highly skewed product mix. Premium ethnic categories, especially Lehenga and Gown, contribute the majority of revenue and show synchronized peaks, while several casual or experimental SKUs move slowly and behave more like support categories than primary growth drivers.

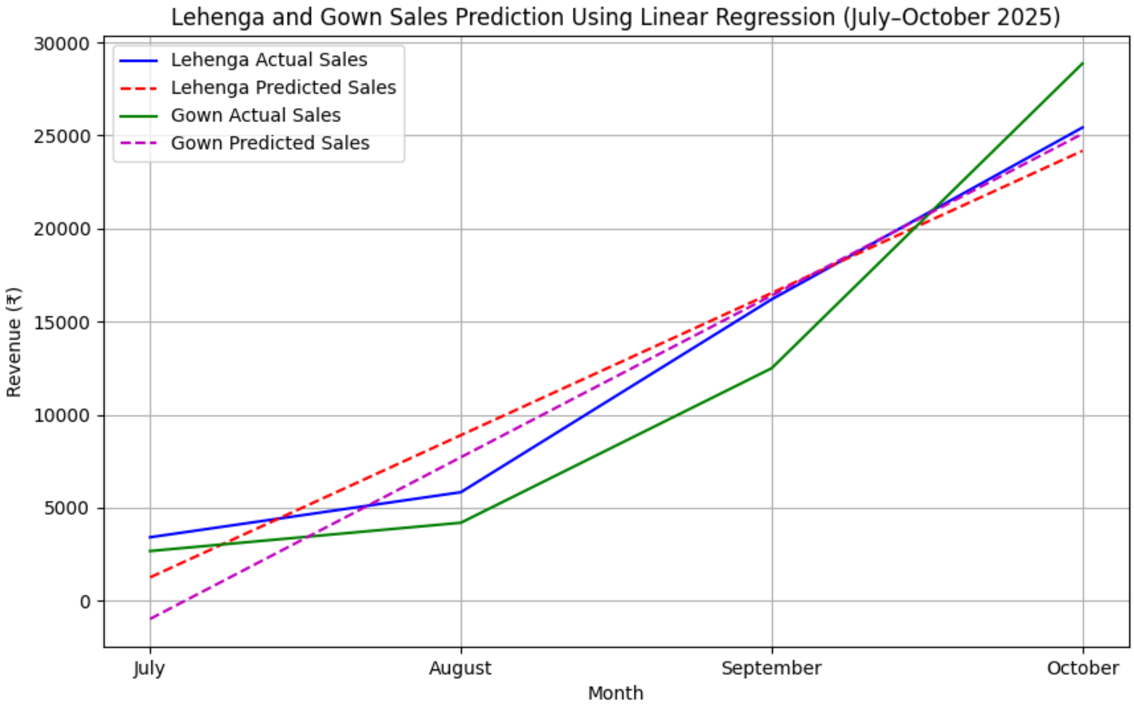


Graph 10: SKU Sales Summary : Revenue Distribution Across Product Type (July – October 2025)

- The tallest bars belong to Lehenga and Gown, together contributing roughly two-thirds of total revenue, confirming them as the primary "A-class" products of the boutique.
- Kurta has a mid-height bar: it sells reasonable volume but at a much lower price point, so its contribution to revenue is modest compared with Lehenga and Gown despite similar or higher unit counts.
- Top, Saree, One-piece and Shirt have significantly shorter bars, indicating they are supporting or experimental SKUs that add variety but do not drive the topline; One-piece and Saree, however, still show relatively high average order values even with low units.
- The shape of the bar chart visually reinforces the ABC classification: A-items (Lehenga, Gown) dominate; B-items (Kurta, Saree) sit in the middle; C-items (Top, One-piece, Shirt) together account for a small share of total revenue

Analytical insight Stock, display space and marketing spend should be weighted according to bar height: deepest inventory and most visibility for Lehenga/Gown, controlled but steady presence for Kurta/Saree, and lean, mostly made-to-order strategies for Tops, One-pieces and Shirts to avoid dead stock while still offering assortment.

Use the combined Lehenga & Gown sales prediction graph (actual in solid lines, predicted in dashed lines across July–October).



- Both Lehenga and Gown show a monotonic increase in revenue from July to October, confirming that demand builds steadily as the festival/wedding season approaches.
- The slope of the lines becomes steeper after August, meaning the month-on-month growth rate accelerates; October revenue for each category is several times higher than July, signalling a strong seasonal effect.
- The red regression lines lie close to the actual points in the mid-period but slightly under-predict early months and over/under-predict at the edges, indicating that linear regression captures the general upward trend but not the full non-linear festival spike.
- Lehenga and Gown curves move in the same direction and with similar shape, which confirms they respond to the same triggers (festivals, weddings, marketing campaigns) and should be treated as a single strategic "premium occasion wear" portfolio for planning.

Analytical insight Linear regression is suitable as a baseline forecasting tool for these two SKUs, but for peak-festival planning it should be combined with seasonal indices or historical festival multipliers so that October-like spikes are not underestimated.

Conclusion Pattern House Boutique's digitalization project has successfully converted manual diary records into a structured digital database and applied comprehensive analytics. Key findings reveal a business strongly anchored in premium occasion wear (Gown + Lehenga = 64% of revenue), with clear seasonal peaks in

September–October driven by festival demand. Machine learning forecasts project 6-month revenue of ₹220k–₹240k (Nov 2025 – Apr 2026), with expected dips in Jan–Feb and recovery in Mar–Apr.

The correlation analysis confirms strong payment discipline (advance payment ratios are predictable) and reveals opportunity for process standardization (delivery speed is order-size independent). ABC analysis recommends focused resource allocation: 60% effort on Class A premium products, 25% on Class B supporting categories, and 15% on Class C niche offerings.

Implementing the recommended strategies—including seasonal hiring, inventory optimization, standardized turnaround times, and targeted marketing—will enable the boutique to maximize profitability, improve customer satisfaction, and build a scalable, data-driven business model. The framework established in this project provides a foundation for continuous improvement and strategic decision-making in the years ahead.

Interpretation of Results & Strategic Recommendations:

1.1 Problem 1: Festival rush but off-season slump

1 **Interpretation:**

- During Navratri/Diwali and wedding season, orders for Lehenga and Gown shoot up, the shop runs at full capacity, and sometimes has to refuse last-minute customers because tailoring slots and fabric are limited.
- In January–February or non-festive months, footfall drops sharply, premium collections move very slowly, and cash flow becomes tight even though fixed expenses (rent, staff) stay the same.

2 **Recommendation:**

- Build a seasonal calendar (month-wise) and lock fabric bookings and karigar capacity for Gown/Lehenga at least 6–8 weeks before the main festivals.
- Introduce off-season offers and small-ticket products (simple kurtas, ready tops, accessories) to generate regular cash flow when big bridal orders are low.
- Offer pre-booking schemes: customers pay a small advance in off-season to block a festival slot, smoothing cash inflow and reducing last-minute chaos.

1.2 Problem 2: Too much money blocked in the wrong SKUs

1. **Interpretation:**

- High-value, fast-moving items (Lehenga, Gown) are sometimes out of stock in key sizes or fabrics, while racks are full of slow-moving pieces like certain Tops, Shirts or odd designs that customers rarely choose.
- This leads to a situation where the shop "looks full" but still cannot meet demand for what customers actually want, so some high-margin sales are lost while old stock keeps aging.

2. **Recommendation:**

- Use a simple ABC tagging on hangers: A for Gown/Lehenga bestsellers, B for steady items like Kurta/Saree, C for experimental or low-rotation designs; review A-items weekly and C-items monthly.
- Convert most C-items into made-to-order only: keep catalogs/samples but do not hold deep stock; clear existing C stock with targeted discounts or "buy bridal, get casual top at 30% off" combos.
- For A-items, define minimum stock levels (for example, always keep at least 3–4 pieces per key design/size) and trigger automatic re-ordering when stock falls below that threshold.

1.3 Problem 3: Manual records causing planning and follow-up issues

1. **Interpretation:**

- Order details are scattered across diaries, WhatsApp chats and verbal instructions, so exact status (advance paid, balance due, trial dates, delivery dates) is not always visible at a glance.
- Because there is no single digital view of orders, the shop sometimes forgets follow-ups for balance payments, misses reminders for trials/alterations, and finds it difficult to estimate next month's workload or fabric requirement.

2. **Recommendation:**

- Maintain a single digital order register (Excel or simple app) capturing form no., customer name, SKU, order type, booking date, delivery date, advance, due amount and status.

- Set up weekly review time (for example, every Sunday evening) to check upcoming deliveries, pending balances and fabric shortages so that corrections are made before problems appear.

- Gradually add basic analytics on this register (monthly sales by SKU, repeat customers, average advance %) to refine forecasting, staff scheduling and cash-flow planning.