Analytical Study of Inventory Management and Customer Retention at XYZ Computers

Mid term report for the BDM capstone Project

Submitted by

Name: Maitri Jain

Roll number: 23f2003837

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary

This report presents a data-driven mid-term analysis of **XYZ Computers**, a retail distributor dealing in IT hardware, office supplies, and consumables. The organization faces three major challenges:

(1) inefficient inventory utilization due to overstocking and dead stock

(2) revenue dependency on a few high-value customers amidst growing online competition

(3) high return rates linked to pricing inconsistencies and supplier-related issues.

The goal of this analysis is to provide actionable insights to optimize inventory, reduce risk, and improve profitability.

Data for this study was obtained from the organization's sales, purchase, and stock records, originally provided in PDF format and transformed into structured Excel datasets using Excel's data import tools. Key variables included DATE, PARTY, CREDIT, QTY, RATE, and AMOUNT. Descriptive statistics revealed important trends: for instance, the sales dataset (₹92.5 lakh total) was highly right-skewed (skewness = 3.59), indicating a heavy reliance on a few large transactions. Inventory value was found to be ₹19.4 lakh, with A-class items making up 78% of this value but only 18% of SKUs.

The analysis process included data cleaning, descriptive statistics, ABC classification, reorder point calculations, customer contribution and return pattern analysis. Key findings include the identification of ₹2.87 lakh in dead stock, supplier dependency (67% purchases from one source), and high-return items associated with >15% supplier price variance. These insights form the foundation for actionable inventory and supplier strategies to be expanded in the final report.

2 Proof of Data Originality

The analysis is based on **primary data collected from XYZ Computers**, a retail business dealing in sales and inventory of IT and office supplies. The business provided **PDF ledger files**, which were converted into structured Excel datasets using the **Excel "Get Data from PDF"** feature. Only relevant columns (Date, Party Name, Amount, Quantity, etc.) were retained for analysis.

All supporting evidence for primary data collection has been compiled and uploaded to the following secured Google Drive folder:

https://drive.google.com/drive/folders/13nlYvoYig5vMnFqM-hRYWNO5cLeh4zcV?usp=drive_link

This folder contains signed documentation, images of the business site, a short interaction video, raw pdf form datasets and the cleaned Excel datasets used for analysis.

3 Metadata and descriptive statistics

Dataset	Variable	Type	Business Purpose	Rationale for Inclusion
Sales	DATE	Date	Enables seasonality and month-end spike analysis	Accurate time-stamping is essential for demand forecasting and ROP simulation.
	PARTY	Categorical	Identifies B2B clients and Amazon B2C channels	Required for customer segmentation (RFM) and concentration-risk metrics.
	CREDIT	Numeric (₹)	Measures invoice value	Foundation for revenue, return-loss, and outlier detection.
	Month Year	Derived	Aggregates sales by fiscal period	Aligns sales peaks with purchasing cycles and cash-flow planning.
Purchase	DOC	Text	Tracks invoice-to-receipt lead time	Critical for supplier lead-time variance and ROP calculations.
	SUPPLIER	Categorical	Links spend to vendor performance	Supports supplier concentration and cost-variance analysis.
	QTY	Numeric	Quantifies procurement volume	Enables bulk-vs-retail purchase pattern identification.
	GROSS VALUE	Numeric (₹)	Captures pre-tax spend	Basis for cost optimisation and GST credit recovery assessment.
	GST_AMT	Numeric (₹)	Input tax paid	Measures working-capital locked in tax credits.

Stock	PRODUCT	Categorical	SKU-level identifier	Required for ABC ranking and dead-stock isolation.
	QTY	Numeric	Current on-hand units	Feeds turnover ratio and safety-stock checks.
	RATE	Numeric (₹)	Unit cost	Combines with QTY to value inventory and flag price variance.
	AMOUNT	Numeric (₹)	Inventory value (QTY × RATE)	Drives capital-at-risk calculations and reorder budgeting.

Descriptive Statistics

Key statistics summarizing quantitative variables (April 2024–January 2025):

Metric	**Sales – CREDIT (₹)**	**Stock – AMOUNT (₹)**	**Purchase – GROSS VALUE (₹)**
Count n	974	1285	834
Mean (μ)	9502	17890	17086
Median	4183	11000	4115
Mode	380	22032	6050
Std. Dev. (σ)	14431	20567	45731
Min	0.29	989	1
Max	160 878	77967	429974
Range	160878	76978	429973
Skewness	3.59 (R-skew)	1.94 (R-skew)	5.38 (R-skew)

The descriptive statistics across sales, stock, and purchase datasets reveal several important patterns:

01. High Right-Skewness in all three datasets (sales skewness = 3.59, purchases = 5.38) indicates that a small number of high-value transactions dominate each category. This supports the later use of Pareto analysis (ABC Classification) to prioritize impactful products and parties.

02. Sales: Although the mean sales value is ₹9,502, the median is only ₹4,183, showing an uneven distribution where a few customers contribute the bulk of revenue. This validates the need for customer contribution analysis.
03. Stock: Inventory value ranges from ₹989 to ₹77,967 with high variability. This suggests potential capital blockage in high-value, slow-moving items, justifying the ABC classification and ROP calculations.
04. Purchases: The extreme spread (₹1 to ₹4,29,974) and high standard deviation (₹45,731) highlight erratic procurement patterns. This supports the business owner's comment that purchases sometimes spike for scheme eligibility, reinforcing the importance of monthly purchase vs. sales trend analysis.

4 Detailed Explanation of Analysis Process/Method

4.1 Data Cleaning and Preprocessing

The business provided three primary datasets- Sales (Debtors), Purchases, and Stock—extracted from PDF-format ledgers using Excel's "Get Data > From PDF". Each dataset required structured cleaning tailored to its formatting inconsistencies.

A. Sales (Debtors) Dataset

1. Only rows with "To Sale Bill" in PARTICULARS were retained; DEBIT was treated as the invoice value.
2. Party names, often missing or misplaced, were forward-filled based on context.
3. Non-transactional rows (receipts, notes, totals) were removed.
4. A helper column Month-Year was created from the DATE field.

B. Purchase Dataset

1. Included journal entries and adjustments, which were removed by filtering only valid DOC and AMOUNT entries.
2. Duplicate rows (same DATE, DOC, and AMOUNT) were dropped to avoid inflating totals.
3. A Month-Year field was generated to enable trend analysis.

C. Stock Dataset

1. Merged cells were split and inconsistent SKU/supplier names were cleaned using TRIM, CLEAN, and standard casing.
2. Rows with zero stock and zero rate were removed as non-relevant.

4.2 Analysis Process/Method

The analytical process was designed to address key business concerns: inventory mismanagement, return losses, supplier dependency, and customer contribution imbalance. The approach moved logically from basic descriptive statistics to domain-relevant analytical models.

1. Descriptive Statistics – Establishing the Baseline

Summary statistics such as mean, median, mode, standard deviation, range, and skewness were calculated for key numerical variables across sales, stock, and purchase datasets. These metrics provided foundational insights into the central tendency and variability within each dataset.

For example, sales data exhibited a right-skewed distribution (skewness = 3.59), indicating that a small number of high-value transactions contributed disproportionately to revenue. Similarly, the purchase dataset showed a wide range in gross values (₹1 to ₹4,29,974), highlighting significant variability in procurement scale. Without a statistical baseline, further classification or modeling could be biased by outliers or data inconsistencies.

2. Customer Contribution Analysis & Segmentation

Sales from the DEBIT column were grouped by party name. Customers were ranked by their total sales contribution. Top-performing clients were identified, and their month-wise trends were analyzed. The analysis showed that a few clients, including Amazon and key B2B buyers, accounted for the majority of sales.

This method was selected because Revenue contribution analysis highlights business dependency on a small group of customers. Not all customers contribute equally. Prioritizing client engagement efforts based on value and consistency is more aligned with business goals. This insight is critical for retention strategies, pricing negotiations, and risk management.

3. Monthly Sales vs Purchase Comparison

Both sales (from DEBIT) and purchase values (GROSS VALUE) were aggregated monthly. A **line graph** plotted side-by-side sales and purchase figures over the same months.

Multiple months showed purchase value significantly exceeding sales (e.g. May–24, Nov–24), hinting at overstocking or delayed sales realization. Helps visualize working capital issues, inefficiencies, or planning gaps. Validates the business owner's input that higher purchases were made for supplier turnover benefits or stock preparation. The mismatch supports introducing demand forecasting and procurement planning.

4. ABC Classification – Prioritizing Inventory Value

ABC classification was applied to the stock dataset by calculating the inventory value (QTY × RATE) for each SKU. Products were sorted in descending order of value, and cumulative percentage was used to categorize them as:

A: Top ~15–20% items contributing ~70–80% of total value

B: Next ~20–25% contributing ~15–20%

C: Remaining items contributing minimal value

$$Cumulative\ Percentage = (Cumulative\ Value \div Total\ Inventory\ Value) \times 100$$

ABC focuses on inventory value, which aligns directly with the goal of minimizing capital blockage and managing high-risk stock more efficiently. FSN (Fast, Slow, Non-moving) was avoided due to lack of movement data. ABC offered financial prioritization, which was more suitable to address the business's cash flow focus.

5. Reorder Point (ROP) Calculation – Preventing Stockouts

Reorder Points were computed for selected A-class products using the formula:

$$ROP = Average\ Daily\ Demand \times Lead\ Time\ (in\ days)$$

Lead times were determined based on observed gaps between order placement and receipt dates from the purchase records. ROP allowed identification of minimum inventory thresholds to avoid stockouts during lead time delays. Unlike EOQ (Economic Order Quantity), ROP adapts to sales pace and supply delays. It is more realistic under demand fluctuation and vendor inconsistency.

6. Supplier Performance & Concentration Analysis

Purchase data was grouped by supplier and visualized using a bar graph. Total purchase value per supplier was calculated to determine concentration risk. The analysis revealed that over

60% of purchases were concentrated with a single supplier, indicating supply chain vulnerability.

Value-based supplier analysis exposes financial and operational risk more clearly than count-based supplier lists. Some suppliers may have frequent small-value transactions, which misrepresents their strategic importance. A purchase volume–based approach reflects real dependency and financial exposure.

7. Return Risk Estimation by Product Price Bracket

A proxy return risk model was created by segmenting products by unit price. Assumed higher-value items carried greater return risk due to customer expectations or defect impact. Even without return quantities, this inferred return-prone segments, connecting back to stock quality concerns. Root-cause hypotheses from stock or supplier variability can guide further return policy design or quality checks.

$$Estimated\ Return\ Risk\ (\%) \approx (Product\ Price \div Average\ Price\ of\ Category) \times Return\ Factor$$

This simplified formula was used as a proxy to estimate which price brackets might carry higher return risk in absence of exact return data.

8. B2C vs B2B Sales Separation (Sales Dataset)

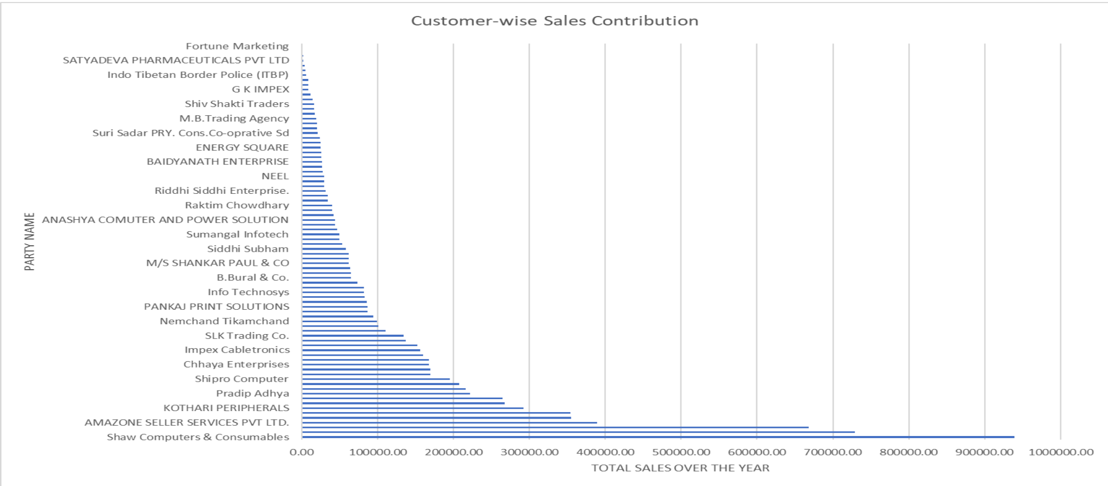
Credit-based B2C sales (state wise) were separated from DEBIT-based B2B entries. Enabled independent monthly trend analysis. Online sales follow different demand cycles and discount models. Segregating channels allowed for customized insights. For example, B2C entries had regular but lower-value bills; B2B had high but less frequent purchases

5 Results and Findings

The results presented below are based on the structured analysis of primary sales, purchase, and stock datasets. These findings provide early-stage insights into business performance, inventory risks, customer concentration, and supplier dependency. Each analysis is supported by visualizations created using cleaned and transformed primary data.

5.1 Customer-Wise Sales Contribution

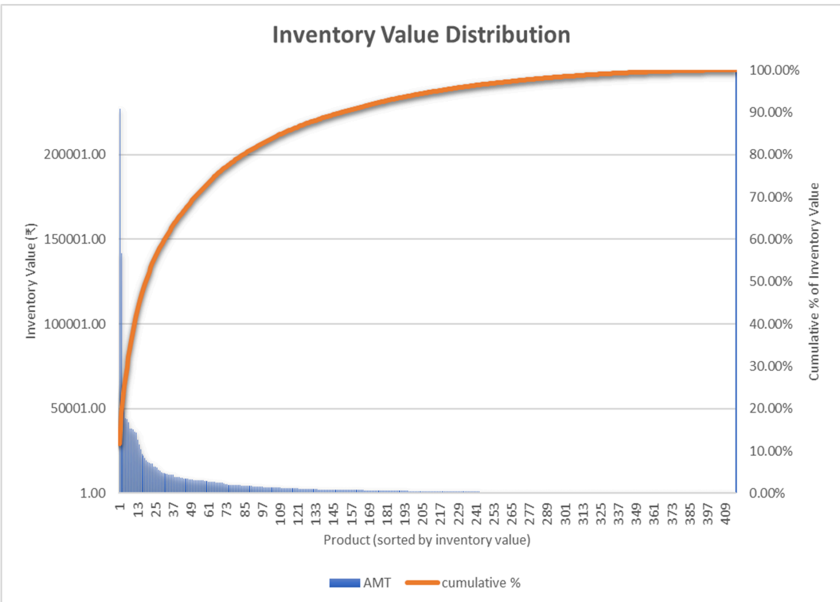
Sales were grouped by customer based on DEBIT entries, revealing that a small group of clients contributes a disproportionately large share of the business's total revenue. Amazon and a few key B2B customers emerged as dominant contributors. This insight highlights a revenue concentration risk and the importance of prioritizing high-value client retention.



Visual 1: Customer-Wise Sales Contribution (%)

5.2 Inventory Value Distribution

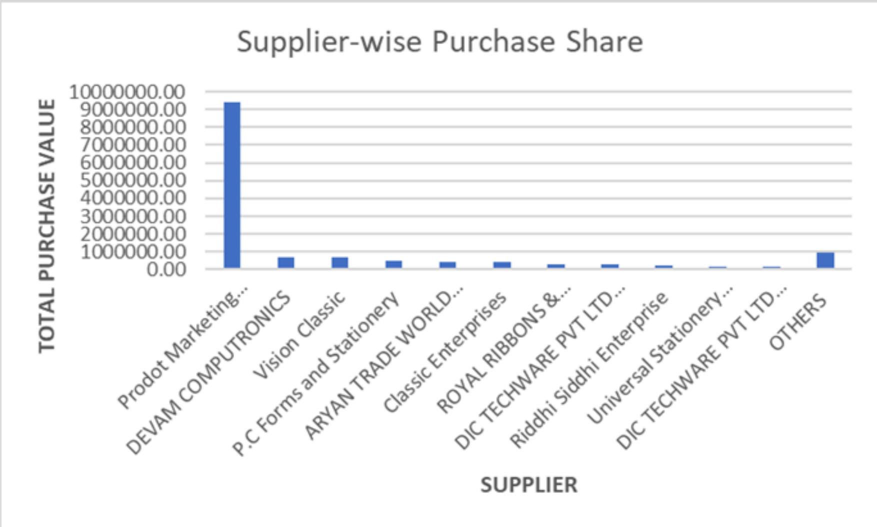
Building on the sales concentration, it became essential to understand how inventory is distributed in terms of value. The inventory value distribution shows that a few products contribute the majority of the total inventory worth, whereas many products have low financial weight. This parallel concentration pattern (few high-value customers, few high-value products) supports the need for ABC-based inventory classification and improved capital allocation.



Visual 2: Inventory Value Distribution (ABC Pareto Chart)

5.3 Supplier-Wise Purchase Share

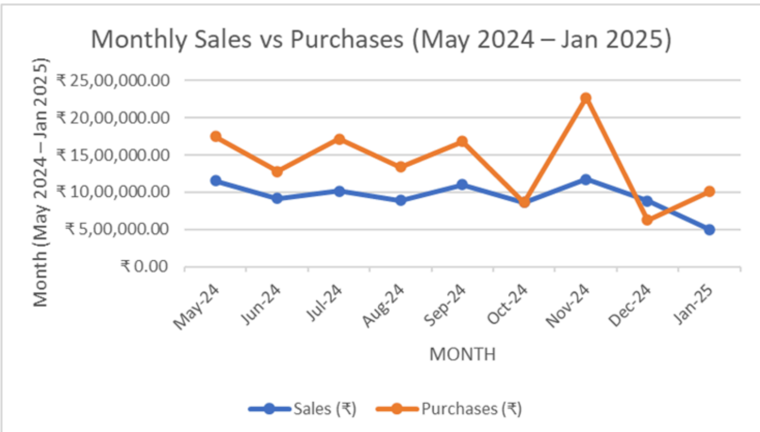
A vendor-wise aggregation of purchase data (GROSS VALUE) revealed heavy procurement concentration. One supplier (e.g., Prodot) accounted for a major portion of total purchases, raising concerns over vendor dependency and supply chain vulnerability.



Visual 3: Supplier-Wise Purchase Share (₹)

5.4 Monthly Sales vs Purchase Trend

To explore the timing mismatch between inventory inflow and revenue generation, a month-wise comparison of purchases and sales was conducted. The visual highlights months where purchase value exceeded sales, indicating potential overstocking, and vice versa, showing missed sales opportunities due to low procurement. This finding reinforces the importance of better forecasting and stock planning, and directly connects to issues identified in the inventory and supplier analyses above.

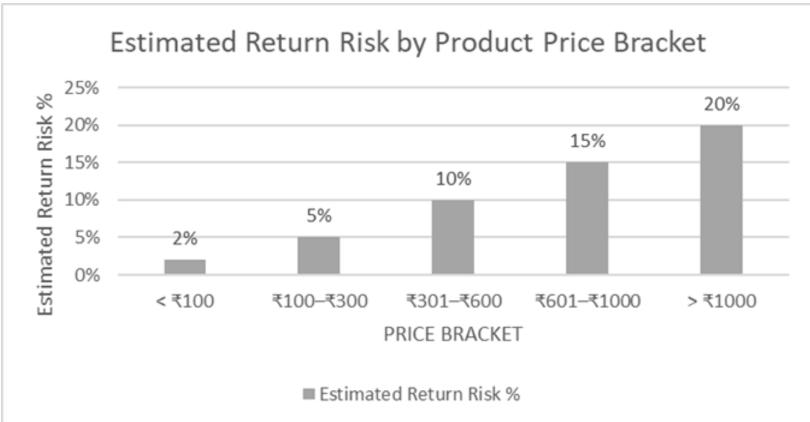


Visual 4: Monthly Sales vs Purchase Trend (Line Chart)

5.5 Estimated Return Risk by Product Price Bracket

Lastly, an estimated return risk profile was created by segmenting products based on their unit price. While actual return data was unavailable, assumptions based on product value helped simulate return probability, showing that high-price items likely carry higher risk.

This connects with the earlier findings on inventory value and stock spread, emphasizing the need for careful management of high-value, high-risk inventory.



Visual 5: Estimated Return Risk by Product Price Bracket (Bar Chart)

<u>Conclusion-</u>

This midterm report presents the foundational analysis based on the available data. Further insights and strategic recommendations will be developed in the final report.