Data-Driven Insights into Credit Sales and Customer Dynamics in a Jewellery Store

A Final report for the BDM capstone Project

Submitted by

Name : Astitva Agarwal

Roll number : 23f2004131

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary

This study presents a data-driven investigation into the operational challenges faced by Shri Giriraj Ji Jewellers, a rural, unorganized jewellery retailer located in Rajau Paraspur, Bareilly, Uttar Pradesh. Established in 2018, the business primarily serves customers from agrarian and village communities, offering gold and silver ornaments through a traditional business-to-consumer (B2C) model. Despite consistent customer demand, the enterprise operates without digital infrastructure, relying solely on manual billing and bookkeeping, which poses significant limitations in credit management, customer behavior tracking, and inventory planning.

The primary objectives of this research project report were twofold: (1) to analyze credit-based sales and customer purchasing behavior, and (2) to detect seasonal demand patterns for inventory optimization. A total of 985 sales and 297 credit transaction records, spanning from May 2020 to April 2025, were digitized and analyzed. Methodologies employed included descriptive statistics, KPI analysis, customer segmentation via K-Means clustering, and time-series trend analysis.

Key findings reveal that a significant proportion of transactions are installment-based, with repayment delays commonly observed during agriculturally lean months (May–July, December–January). Conversely, repayments peak in post-harvest periods (February–April and October), demonstrating a strong correlation between rural cash flow cycles and credit behavior. Customer profiling identified three distinct segments, including high-value low-frequency buyers and mid-tier loyal customers, offering opportunities for differentiated credit and retention strategies.

Product-level analysis highlighted that lightweight gold items (e.g , earrings, pendants) exhibit peak demand during festive and wedding seasons, whereas heavyweight silver products (e.g., anklets, bangles) align with post-harvest liquidity periods. Inventory inefficiencies, particularly due to overstocking of low-demand heavy items, were also identified.

The study recommends implementing seasonally-aligned credit policies, optimized procurement based on demand cycles, and the adoption of digital tools for billing, inventory, and customer relationship management. These insights provide a robust framework for enhancing profitability and operational resilience in traditional rural retail contexts through data-informed decision-making.

2 Detailed Explanation of Analysis Process/Method

Shri Giriraj Ji Jewellers, operating in a rural and unorganized setup, faces significant challenges due to the lack of digital systems. The key problems include difficulty in tracking delayed payments from credit-based sales, inability to analyze customer purchasing behavior, and lack of insights into seasonal demand trends. These limitations hinder effective decision-making, resulting in revenue losses and inefficient inventory management.

To effectively address the problems, I began by collecting the relevant and available data based on problems. The primary business objective was to analyze the customer purchasing behavior and credit-based sales to make informed decisions for credit sales, leading the business to better profits and bearing fewer losses. There are a variety of products in the jewellery business with variable weights and types, so to predict the future demand, historical sales data was collected.

Data was collected in the form of handwritten bills , a total of 985 sales entries and 297 credit entries which was later digitized and stored in Google Sheets.

23f2004131_BDM_Sales_Data : This dataset contains all the entries of sales between the time period from May-2020 to Nov-2023 . The Dataset contains 13 columns and 986 rows , from which columns refers to the name of customer , date of purchase , category , type of product, its weight , rate per gram making charges , discounts , total amount , payment type , exchange used .

23f2004131_BDM_Credit_Data : This dataset contains records of all credit payments made between May-2020 and April-2025. During this period, entries can be found for customers who purchased products on credit and later paid the remaining amount in installments. The Dataset contains 3 columns and 297 rows . It has the customer's name , date of payment and amount of payment made .

2.1 KPI Analysis :

To understand the operational and financial performance of Shri Giriraj Ji Jewellers, a KPI-based (Key Performance Indicators) analysis was conducted. Key performance indicators such as average transaction value, profit margin trends, and installment frequency patterns were used to evaluate customer contribution, product profitability, and cash flow dynamics. These metrics help the business understand what products drive revenue, how customer behavior differs between payment types, and where financial inefficiencies occur.

Steps to Perform:

Data Preprocessing: Sales and credit data were cleaned to ensure valid numeric entries for rate, weight, and total amount this includes removal of Currency symbols like ₹ and commas columns were converted to float.
Date columns were parsed into datetime format using dayfirst=True.

To handle missing values for *making charges* and *discounts*, a rule-based logic was applied using the formula:

Base Price = Weight (g)* × *Gold Rate (₹/g).

Based on this:

- If **Actual Price > Base Price**, then
  Making Charges = Actual Price − Base Price,
  Discount = 0.

- If **Actual Price < Base Price**, then
  Discount = Base Price − Actual Price,
  Making Charges = 0.

This logic was essential due to the absence of explicit values in handwritten bills and helped normalize the pricing data, ensuring consistency across entries and enabling reliable analysis of cost structures and profitability.

KPI Calculation Formulations :

- **Base Margin (BM):**

$$\mathrm{BM} = \begin{cases} 0.10 \times W \times R, & \text{if Category} = \text{Gold} \\ 0.20 \times W \times R, & \text{if Category} = \text{Silver} \end{cases}$$

Where:

W = Weight (in grams)
R = Rate per gram (₹/g)

- **Final Margin (FM)**

$$\mathbf{FM=BM+MC-D}$$

Where:

BM = Base Margin
MC = Making Charges
D = Discount

- **Average Transaction Value (ATV)**

$$\mathrm{ATV} = \frac{\sum_{i=1}^{n} S_i}{n}$$

Where:

Si = Selling price of transaction
n = Total number of transactions

Visualization: A bar chart of total revenue by product, a summary table by material type, and a dual-axis line chart of monthly average margin vs. total sales were created using Python (Pandas, Seaborn) to visualize product performance, material-wise sales patterns, and profitability trends over time.

Expected Outcome: The visualizations help identify high-performing products, understand sales and track how profit margins evolve relative to overall sales. This supports data-driven monitoring of key performance indicators related to revenue contribution, customer behavior, and profitability.

2.2 Customer Profiling

Customer profiling was conducted to segment buyers based on their purchase frequency and total spend. This analysis helps identify loyal high-value customers, seasonal buyers, and low-engagement clients. It supports personalized credit policies, loyalty offers, and inventory planning tailored to each segment's behavior.

Steps to Perform:

Data Preprocessing:

Name Cleaning and Standardization :

- Inconsistent customer names across the sales and credit datasets were cleaned using fuzzy matching via the *RapidFuzz* library.
- All customer names were: Converted to uppercase and Stripped of leading/trailing whitespace and names were clustered using token_sort_ratio with a 90% similarity threshold.
- A master mapping was generated and applied, resulting in a unified column: Customer Name Clean.
- Currency fields (e.g , Making/Labour charges, Discount, Total Amount(₹)) were cleaned of symbols and converted to float. The Date column was parsed using dayfirst=True and converted to datetime format.

Monthly Aggregation and credit season :

- Both sales and credit data were grouped by month and category to analyze trends.
- From this, the following temporal features were derived:
- Month (numeric) , Year , Payment_Month (e.g., "January", "February") ,Season: Custom-defined as:
    - Rabi Harvest → April to June
    - Zaid → July to September
    - Kharif Harvest → October to November
    - Rabi Sowing/Lean → December to March

Segmentation Algorithm:

To segment customers based on purchasing behavior, we employed unsupervised learning using the K-Means clustering algorithm, which partitions observations into *K* disjoint subsets (clusters) by minimizing intra-cluster variance.

- **Feature Construction**

Each customer i is represented by a feature vector:

$$\mathbf{x}_i = \begin{bmatrix} \text{TotalSpend}_i \\ \text{PurchaseFrequency}_i \end{bmatrix} \in \mathbb{R}^2$$

Where:

Total Spend:

$$\text{TotalSpend}_i = \sum_{j=1}^{n_i} a_{ij}$$

a_{ij} : transaction amount for the j_{th} purchase by customer i

n_i : number of purchases (transactions) made by customer i

$\text{PurchaseFrequency}_i = n_i$

These two continuous variables define the input space for clustering

- **Clustering Algorithm: K-Means**

We aim to partition the dataset into K=3 clusters {C1,C2,C3}by minimizing the within-cluster sum of squares (WCSS):

$$\underset{C_1,\ldots,C_K}{\arg\min} \sum_{k=1}^{K} \sum_{\mathbf{x}_i \in C_k} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|^2$$

Where:

μ_k: centroid of cluster C_k
$\| \cdot \|$: Euclidean norm

Initialize *K* centroids, assign each point to the nearest centroid, then update centroids as the mean of their assigned points. Repeat and update until convergence (no significant centroid movement).

Visualization: Customer segments and behavior patterns were visualized using scatter plots, bar charts, time series plots, and heatmaps to uncover spending, repayment, and seasonal trends. Tools used include Python libraries: Pandas (data handling), RapidFuzz (name clustering), Scikit-learn (K-Means), Matplotlib & Seaborn (static visualizations), and Plotly (interactive plots).

Expected Outcome: This segmentation enables the business to identify high-value customers, optimize credit terms, and tailor loyalty schemes. Insights on repayment cycles aligned with crop seasons help improve collection timing and reduce credit risk through data-driven strategy.

2.3 Product and Inventory Forecasting & Optimization

This analysis aims to uncover cyclical patterns in product sales and category-wise performance, particularly influenced by festive seasons, regional buying behavior, and agricultural cycles. Instead of relying solely on model-based forecasting, a combination of trend analysis and pattern-based time series insights was applied to guide inventory planning and optimize product assortment.

Steps to Perform:

Data Preprocessing:

Cleaned and filtered transaction-level data by date, product type, and material (Gold/Silver).
Segregated products into lightweight (<20 gm) and heavyweight (≥20 gm) for more targeted insight.
Handling missing values and Margin formulations provided in KPI analysis's data preprocessing for better context .

Trend Analysis:

To uncover meaningful sales patterns and seasonal effects in the jewelry business, the following statistical and analytical methods were applied:

- ***Monthly Aggregation of Sales:***

  Sales data was resampled to monthly frequency using Pandas' .resample('M') method, enabling smoothed visualization of time-series trends and eliminating daily noise.

- ***Trend Smoothing:***

  For each product weight category (lightweight <20g, heavyweight ≥20g), a monthly moving average of revenue was plotted. This smoothing helped highlight broader sales patterns while reducing the impact of transaction-level fluctuations.

$$\mathrm{MA}_t = \frac{1}{k} \sum_{i=t-k+1}^{t} x_i$$

  where MA_t is the moving average at time t over window k (in this case, months), and x_i represents total sales in each month.

Forecasting Method(Non-Model-Based): Although SARIMA/SARIMAX models were explored, they proved ineffective due to data variability and inconsistent patterns.
Instead, the analysis focused on trend-based forecasting and visual seasonality recognition to

identify high-demand periods and category-level sales consistency.
Insights were derived through pattern interpretation rather than statistical prediction.

Visualization: Line Charts showing revenue trends segmented by product weight and material (Gold/Silver).
Bubble Charts to analyze product performance based on transaction volume, profit margins, physical weight, and revenue impact.

Expected Outcome: Understand sales dynamics between gold and silver, Forecast high-demand months (e.g., October, March, November) based on historical trends.Use visual and multi-dimensional product performance insights to inform storage planning, promotional targeting, and capital deployment.

3 Results and Findings

After completing the analysis , several meaningful insights were revealed from credit and sales data . All visualizations were generated using the Matplotlib module of python on google colab . ***BDM_Project_Analysis_Code.ipynb***

3.1 KPI Analysis

This bar chart visualizes the best-selling products by total revenue, providing a clear snapshot of the top-performing items over the 42-month analysis period. Such visual representation is essential for guiding decisions related to inventory planning, promotional focus, and product lifecycle management.

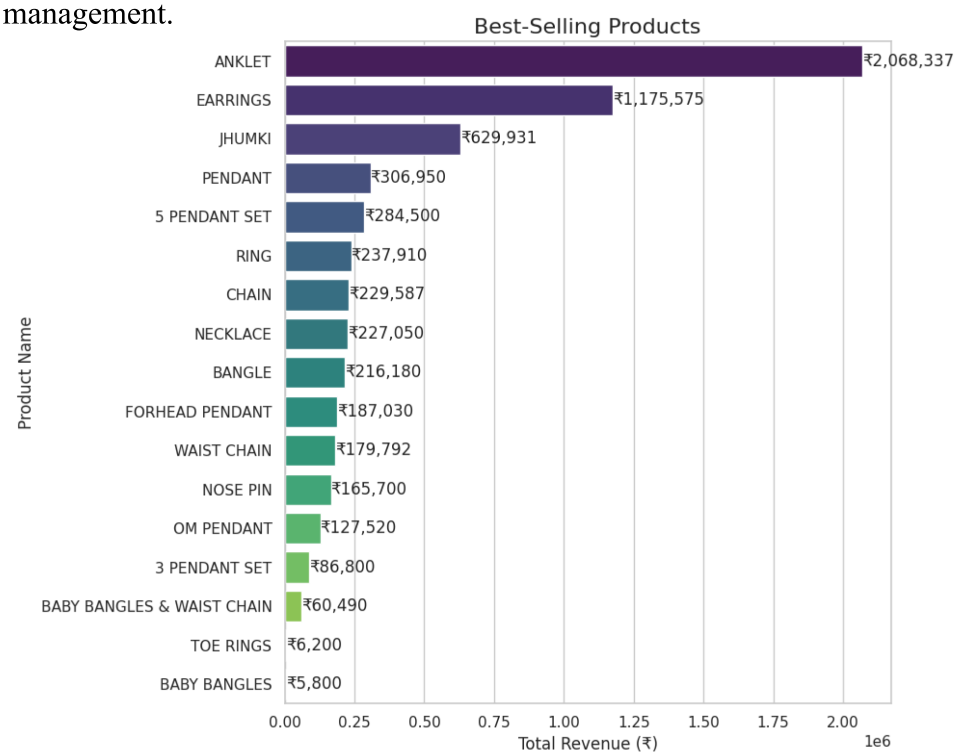

Figure 1 Bar Chart Showing Best-Selling Products by Total Revenue

It can be clearly observed that Anklets and Earrings are the most significant contributors to overall revenue, with Anklets alone generating over ₹20 lakhs, making it the leading product in terms of sales. These two product categories form the backbone of revenue, indicating strong and consistent customer demand.

The following table presents descriptive statistics of the sales data across material categories (Gold and Silver). Notably, Gold items account for approximately 40% of total transactions by volume, yet exhibit a much higher mean transaction value (~₹9,500) which is 2.25 times greater than that of Silver. Silver, while contributing 60% of transaction count, maintains a lower average transaction value of approximately ₹4,200.

Table Showing Transaction Metrics by Product Category (Gold and Silver)

Category	Transaction_Count	Mean_Value	Median_Value	Value_Range
GOLD	391	9548.54	7400.00	80050.00
SILVER	594	4232.6	3600.00	21200.00

The Average Transaction Value (ATV) is a crucial metric that represents the average amount spent per transaction. For the total dataset:

ATV = ₹6,342 (₹6,247,566 / 985 transactions)

This implies that on average, a customer spends between ₹6,000–₹7,000 per purchase.

Depth and Insight

To gain deeper insights, ATV was calculated for the top five revenue-generating products.

Anklets, contributing ₹20.68 lakhs from 444 units, have an ATV of ₹4,659. Based on the average silver rate of ₹63.5/gm, this corresponds to an approximate average weight of ~73.4g per unit, indicating high-volume, silver-based sales.

Earrings, made of gold, generated ₹11.75 lakhs from 131 units, resulting in an ATV of ₹8,973. Using the average gold rate of ₹5,053.4/gm, this translates to an average transaction weight of ~1.78g, typical for lightweight gold jewelry.

Premium items like Jhumkis and 5 Pendant Sets recorded high ATVs of ₹24,229 and ₹21,885, with estimated weights of ~4.8g and ~4.3g, respectively consistent with their positioning as lower-volume, high-value gold products.
Pendants, with an ATV of ₹9,029, reflect a mid-range offering, averaging ~1.79g per transaction.

Overall, these figures reveal a balanced sales structure: silver drives volume through heavier, lower-value units, while gold delivers value through lightweight, high-ATV transactions.

This dual-axis line (Figure 2)chart illustrates the relationship between Monthly Average Margin (in ₹) and Total Sales (in ₹) from January 2021 to October 2023. It is used to analyze how profitability per unit (margin) moves in relation to overall sales revenue over time. This view is essential for understanding whether growth is being driven by high-value transactions or simply higher volume, and how product mix and pricing strategies impact margin performance month-to-month.

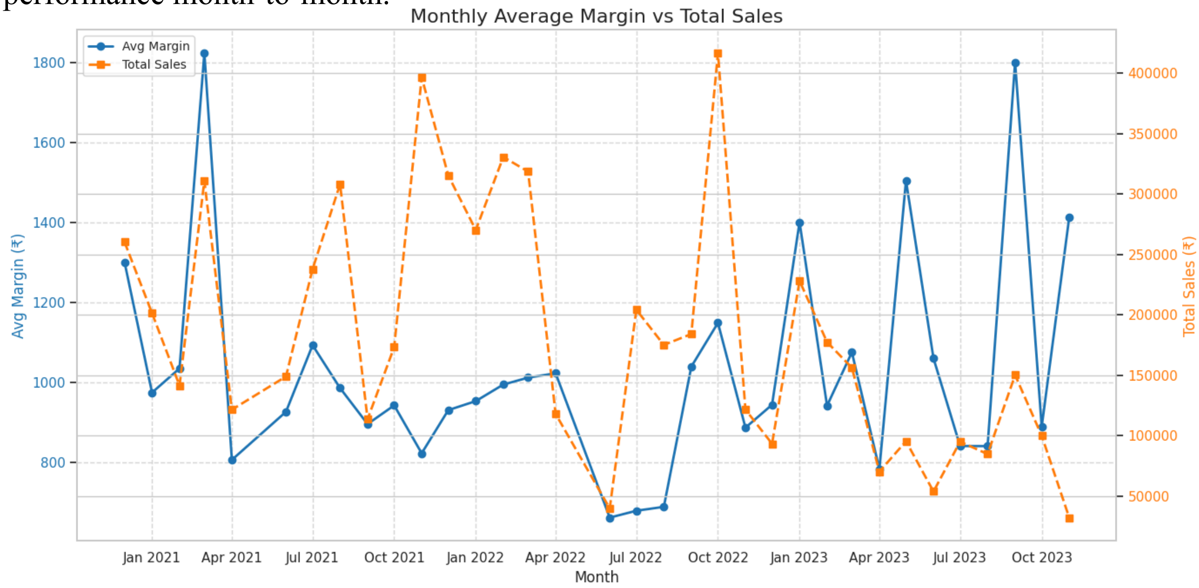

Figure 2 Dual-Axis Line Chart Showing Monthly Average Margin vs. Total Sales

Depth and Insight

Across the timeline, the relationship between sales and margins appears inconsistent, revealing underlying shifts in product strategy and transaction types. Months like October 2022, which recorded the highest sales (₹4.1 lakhs), coincide with only moderate margins, indicating high-volume movement likely supported by more flexible transaction terms or lower-margin products. In contrast, March 2021 and September 2023 exhibit elevated average margins (₹1,800) despite lower overall sales, hinting at concentrated, high-value purchases.

Stable dual-performance periods such as January 2023 and May 2023 suggest a healthy blend of volume and margin, whereas dips seen in April 2021 and July 2022 reflect market slowdowns or off-season activity. On closer inspection, months where profitability rises as volume drops may reflect more upfront, value-focused sales, while the reverse spikes in sales with softened margins may point toward deferred or diluted revenue collection mechanisms.

Together, these patterns highlight the importance of balancing sales growth with transactional quality and profitability consistency.

3.2 Customer profiling

Now below are the insights from Problem statement 1 for credit sales and customer dynamics This scatter plot segments customers based on two key behavioral variables: total spend (in ₹) and purchase frequency. Customers are clustered using an unsupervised machine learning technique K-Means , and each cluster is labeled as a distinct "Customer Segment."

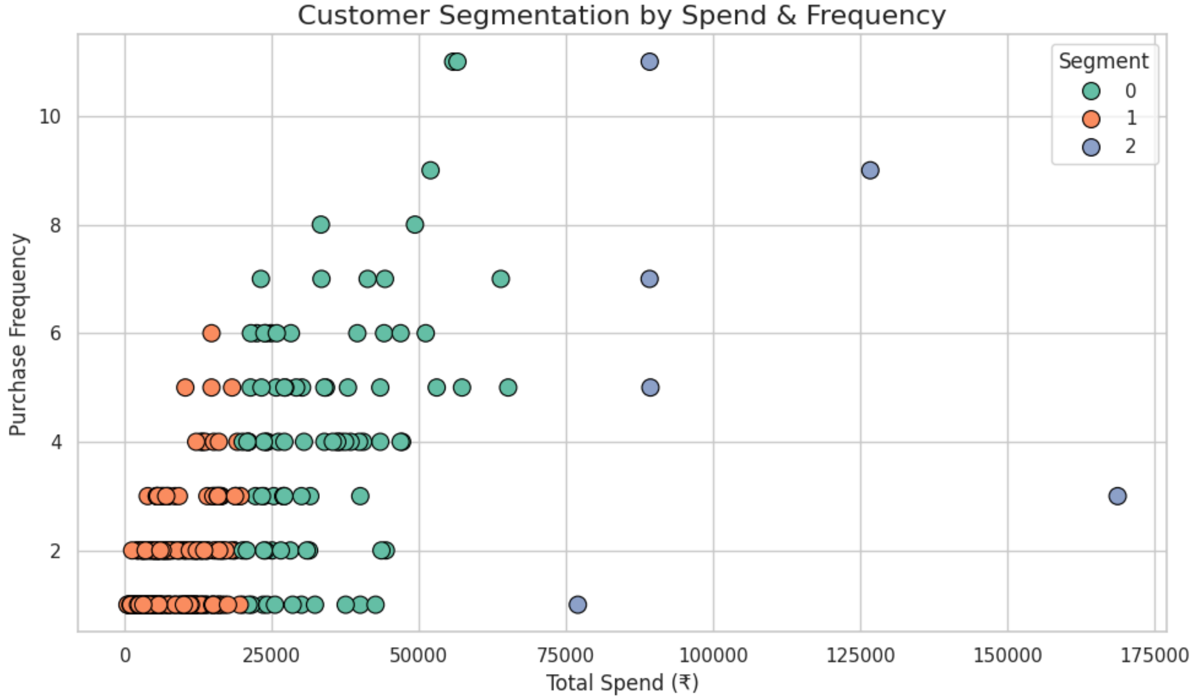

Figure 3 Scatter Plot Showing Customer Segmentation by Total Spend and Purchase Frequency

Segment 0 (green): This group has moderate to high frequency but varying levels of spend (up to ₹75,000), suggesting they are regular but mid-value customers.

Segment 1 (orange): This cluster comprises low spend and low frequency customers, possibly one-time or occasional low-value buyers.

Segment 2 (blue): Small in number but significantly high in total spend. These are high-value customers with potentially lower frequency but larger ticket sizes.

Depth and Insight

The presence of high-value, low-frequency customers (Segment 2) highlights a niche segment with substantial purchasing power ideal for targeted premium services or credit offers.
Segment 0's pattern shows potential for repeat business, indicating this cohort might be more responsive to loyalty or retention strategies.
Segment 1 may represent first-time users or financially constrained customers, useful for assessing creditworthiness or onboarding interventions.

This bar chart (Figure 4) visualizes how many customers made installment payments and the average amount per transaction within each group, excluding full (one-time) payments.

The majority of customers (140) made exactly 2 installment payments, with an average transaction of ₹7,873.

A gradual drop-off is seen as the number of installments increases, but the average transaction value increases sharply with each additional installment. E.g., 5+ installments had a transaction value of ₹52,500 much higher than the rest.

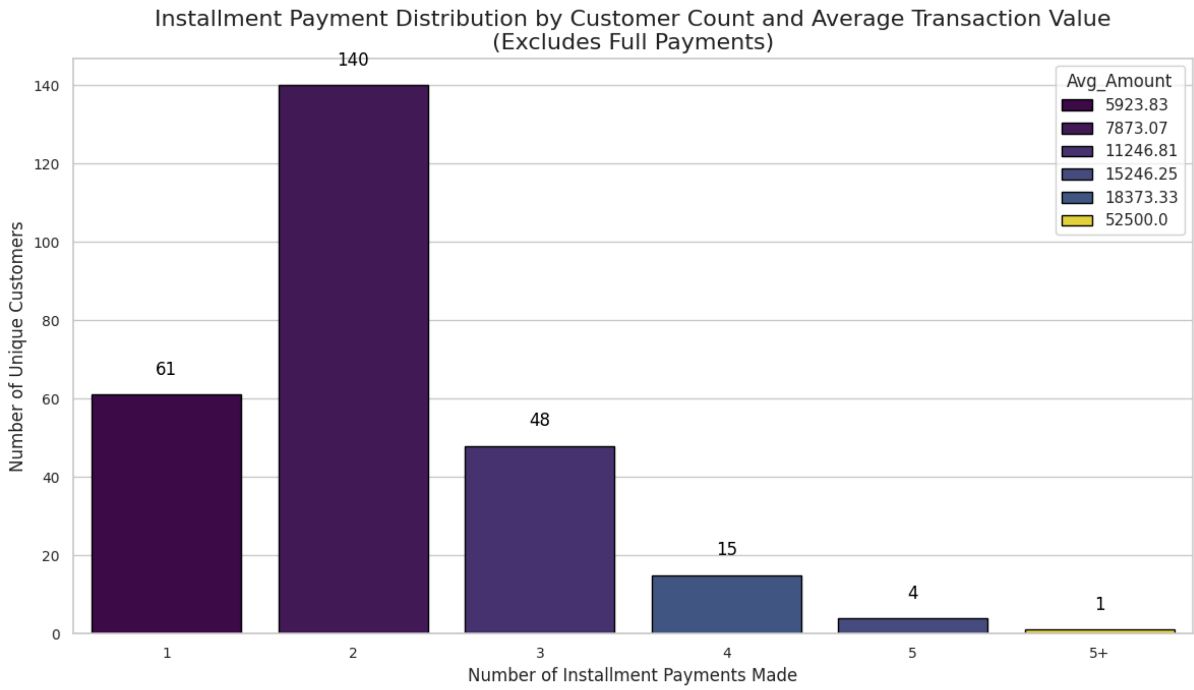

Figure 4 Bar Chart Showing Installment Payment Distribution by Customer Count and Average Transaction Value (Excluding Full Payments)

Depth and Insight

Higher installment counts correlate with larger purchase values, suggesting installment flexibility is essential for enabling high-value sales.

Most customers prefer short installment plans, for purchase between 7000 to 8000 the customer opts for 2 installment but tend to delay the payment for months possibly due to limited cash flow in rural and farmer community

The skew toward low installment counts can inform default risk assessments and guide payment plan design tailored to transaction size.

This time-series plot (Figure 5)tracks delayed payments (payment lag >28 days) month-wise and categorizes them by severity of delay giving visibility into payment discipline and customer liquidity stress over time.

A large portion of 2021 shows severe delays (red markers, average lag >75 days).

There's a gradual decline in the number of delayed payments from late 2021 through 2023, indicating the credit policy changes implemented by the business owner and limiting the installment plans .

Green zones are rare, indicating timely payments were infrequent overall.

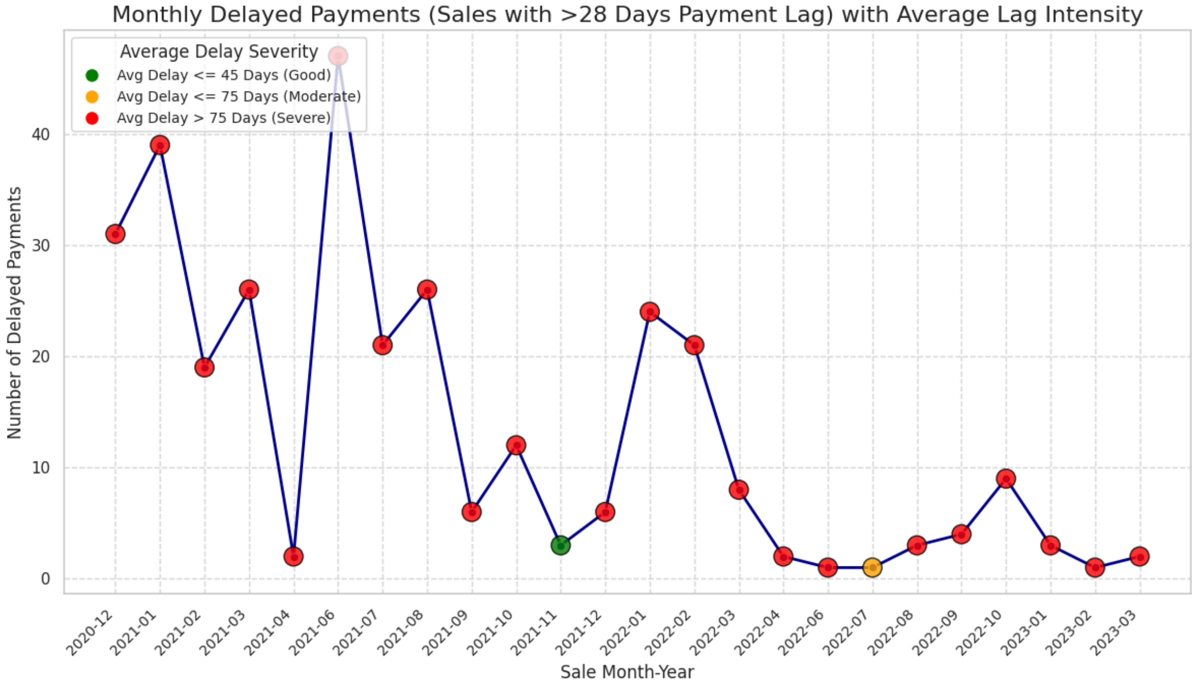

Figure 5 Line Chart Showing Monthly Delayed Payments (Sales with >28 Days Payment Lag) with Average Lag Intensity

Depth and Insight

It can be interpreted that November and December have the least amount of delayed payments in 2021 but peaked during July and August . A good amount of delayed payments can also be seen in January to late April in both 2020 and 2021 years .

2021 may represent COVID after-effects or post-harvest uncertainty, causing severe repayment stress. The falling trend in 2022-23 could indicate:

Adjusted credit vetting criteria , Improved cash flows due to good harvests , Better collection strategies

Seasonal peaks in delayed payments could be cross-validated with Figure 2 for causal analysis between low harvest periods and repayment failures.

The analysis of credit repayment data from December 2020 to March 2023 reveals a distinctly seasonal repayment behavior, closely tied to agricultural income cycles rather than retail events or sales promotions. While jewellery purchases spike around weddings and festivals, and overall sales trends (as shown in Figure 2) follow festive and wedding seasons, credit repayments follow a different rhythm driven primarily by when rural customers, especially from farming households, have actual cash flow post-harvest.

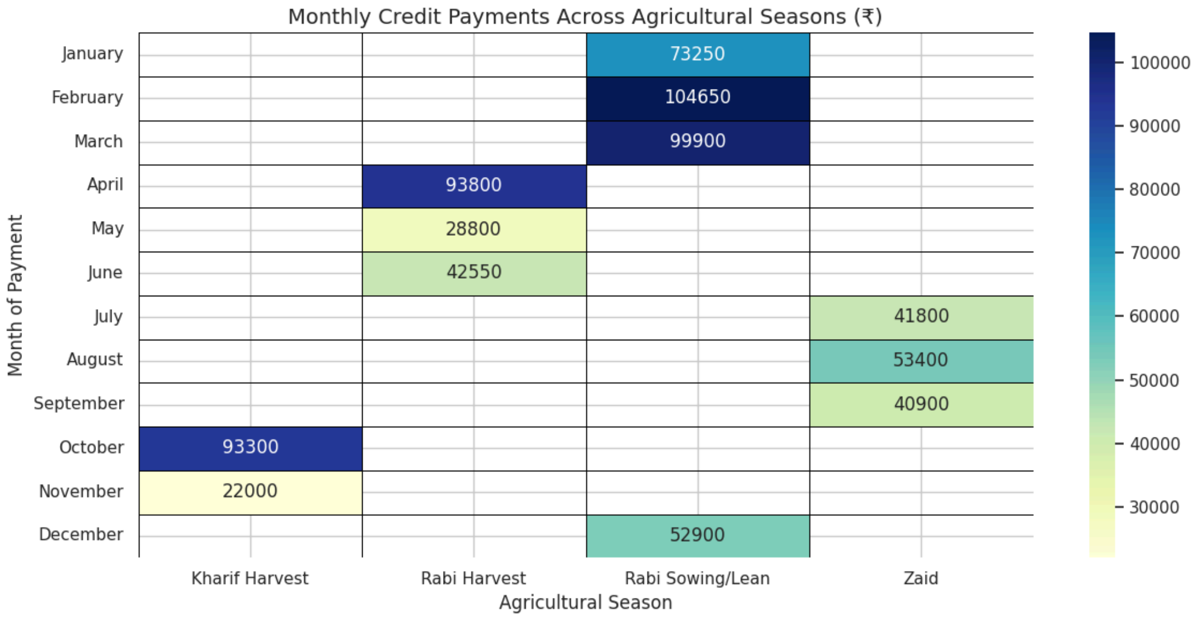

Figure 6 Heatmap Showing Monthly Credit Payments Across Agricultural Seasons

The seasonality in repayment behavior is clearly illustrated in Figure 7, which shows total credit payments aggregated by agricultural season. Repayment activity is concentrated in a few key phases February (₹104,650), March (₹99,900), and April (₹93,800) which align with the end of the Rabi crop cycle. This period marks one of the most stable income phases for farmers due to wheat harvests, and the repayment intensity suggests that customers clear outstanding dues after liquidating produce or receiving payments from agricultural markets. Another notable spike occurs in October (₹93,300), following the Kharif harvest (rice), though this is shorter-lived, as repayments drop sharply in November (₹22,000), indicating a narrow liquidity window.

These observations are further reinforced in Figure 6, which presents a heatmap of monthly credit payments across different agricultural seasons. It shows concentrated repayment intensity in the Rabi harvest and immediate post-harvest months, while payments remain consistently low during the Zaid and sowing periods particularly May–June and July–September, which align with lower or delayed agricultural income. Even though jewellery purchases rise during these months (as shown in Image 2), repayment behavior remains weak, confirming that credit usage peaks during lean periods but repayment is deferred until post-harvest months.

Figure 8, which tracks the monthly credit payment trend over the entire time period, highlights this cyclical pattern in more detail. The graph consistently shows a dip in repayments every December–January, followed by a sharp rise in March–April. This reflects the typical lag between sowing-related expenditure and the realization of income from Rabi crop sales. The visual also captures a gradual plateauing and drop in repayments after mid-2022, which corresponds to the company's strategic shift toward limiting credit disbursement due to growing mismatch between credit usage and repayment capacity during non-harvest months.

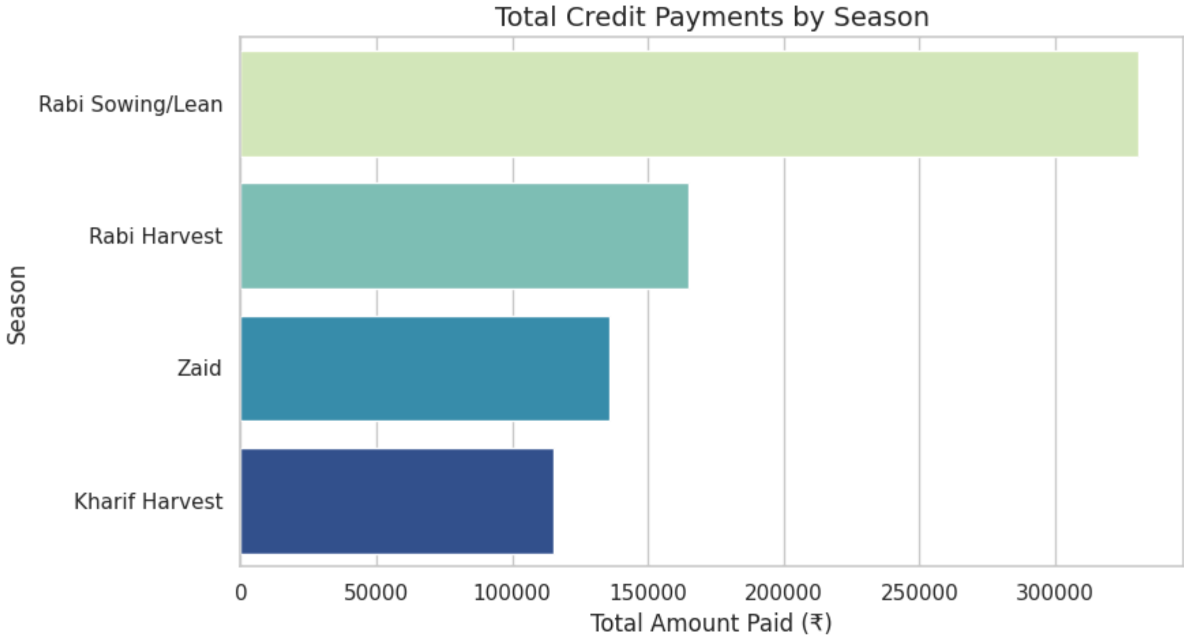

Figure 7 Horizontal Bar Chart Showing Total Credit Payments by Season

Depth and Insight

Seasonally aggregated data confirms that the Rabi sowing and early-year phase (Jan–Mar) has the highest total repayments (₹330,000+), even though this is traditionally considered a lean agricultural period. This may be due to early crop sales, repayment before wedding seasons, or stored liquidity from Kharif earnings. The Rabi harvest phase (Apr–May) also shows strong repayment totals (₹170,000), while the Zaid and Kharif harvest seasons are lower at ~₹140,000 and ~₹120,000 respectively. These repayment cycles are driven less by retail purchasing timelines and more by when agricultural liquidity becomes available.

These trends indicate that repayment is less a function of sales timing and more a reflection of agricultural liquidity. Customers tend to repay shortly after harvests and delay during sowing or non-income phases. The business's evolving credit strategy must continue aligning with these agricultural patterns. Focused follow-ups in February–April and October, and deferred or flexible collection in low-income months, can significantly improve repayment reliability in rural, agriculture-dependent markets.

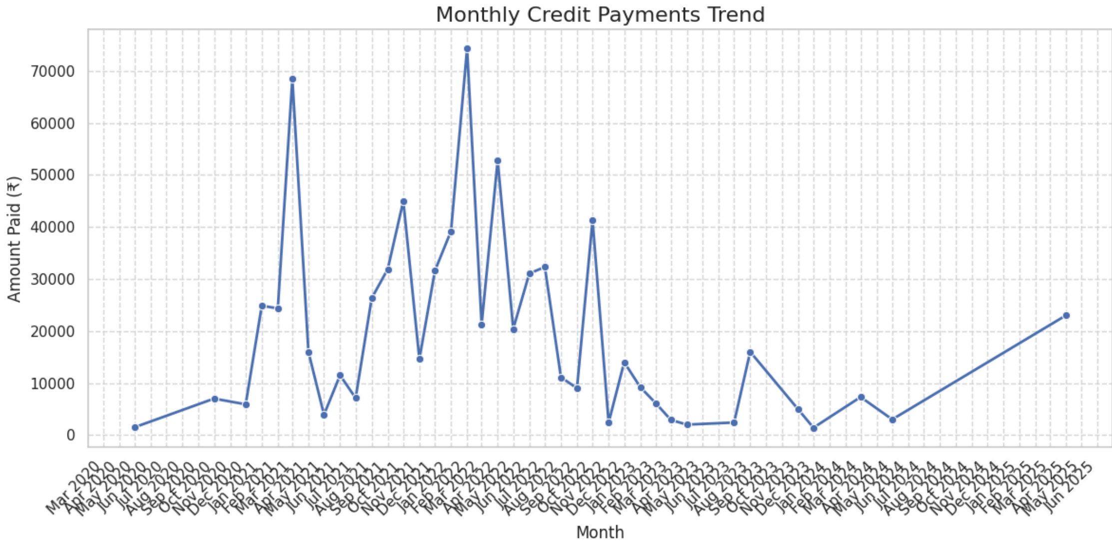

Figure 8 Line Chart Showing Monthly Credit Payments Trend

This repayment pattern supports the need for credit structuring and recovery strategies that align with local crop calendars. Timely follow-ups during February–April and October, and flexible or deferred collection during Zaid and sowing months, can significantly improve repayment reliability in rural, agriculture-dependent markets.

3.3 Product and Inventory Forecasting & Optimization

Understanding which products deserve priority in inventory and marketing is not simply a matter of sales count or profit margin; it requires a balance of multiple variables. The bubble chart (Figure 9) used in this analysis was chosen to present that balance visually and intuitively. Without overtly segmenting the products, it reveals key strategic insights by allowing us to view frequency of sale, profit margins, product weight, and comparative performance in one comprehensive glance. In this chart:

- The X-axis maps the number of transactions (sales frequency) per product type.
- The Y-axis shows the average profit margin per product.
- Bubble size corresponds to average product weight   vital for understanding storage and logistics implications.
- A color gradient represents the intensity of profit margin , warmer colors signify higher profitability.

By plotting these factors simultaneously, the chart helps identify not only the best performers but also products with hidden potential or hidden costs.

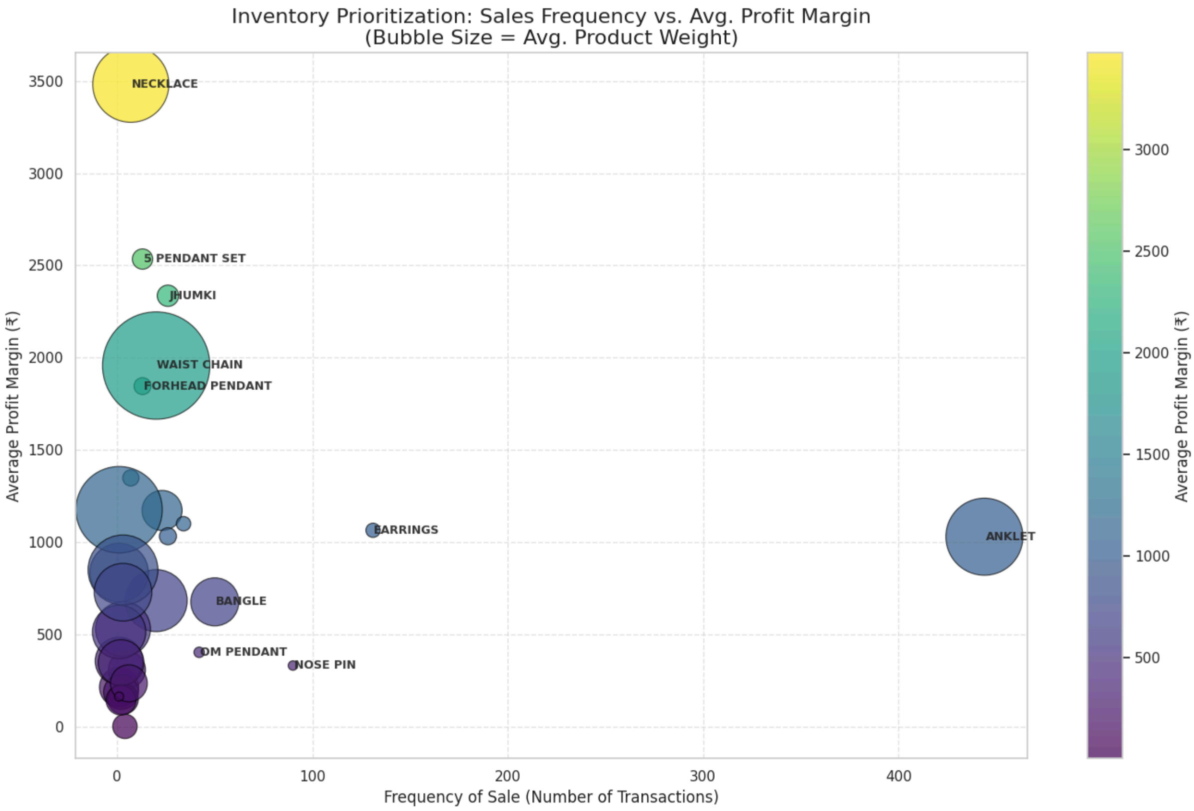

Figure 9 Bubble Chart Showing Inventory Prioritization: Sales Frequency vs. Average Profit Margin (Bubble Size Represents Average Product Weight)

Depth and Insight

The chart naturally divides products into four broad behavioral quadrants:

High Frequency + High Margin (Top-Right):

This is the ideal zone for inventory products here delivering both volume and value. While no product lands squarely in this quadrant, Anklet comes closest. With the highest frequency (444 transactions) and a moderate profit margin (~₹1000), it's a clear leader in terms of turnover. However, its significantly large weight means storage and handling efficiency must be closely managed.

High Frequency + Low Margin (Bottom-Right):

Products like Nose Pin, Earrings, and Bangles sell well but bring in relatively low margins. Nose Pin has a modest margin (~₹300) but high frequency, making it ideal as a promotional or entry-level item. Earrings stand out slightly with a higher margin (~₹1050) and 130+

transactions, positioning them as balanced performers.Bangles, with moderate sales and mid-level margins, may benefit from bundling or margin optimization strategies.

Low Frequency + High Margin (Top-Left):

These are luxury or occasion-driven products low in volume but high in profitability. Necklace leads here with the highest margin (~₹3500), though its frequency is low and weight high. It likely fits well in boutique or event-driven inventory strategies.
Other products like Pendant Set, Waist Chain, Jhumki, and Forehead Pendant follow a similar trend, offering good margins but infrequent demand. Targeted marketing and lean inventory can help balance risk and reward here.

Low Frequency + Low Margin (Bottom-Left):

This quadrant includes several smaller items with low sales frequency, modest margins, and minimal weight. Most product names in this section were intentionally left unlabeled in the chart to avoid visual overcrowding, as this area is densely populated. Still, these remaining items hold relevance often serving occasional, seasonal, or rural demand, where specific preferences and affordability drive purchasing behavior.

Among them, the ***Om Pendant*** stands out as a balanced performer. With a light weight typically under 1 gram, it maintains a healthy average profit margin of around ₹400 and a frequency close to 50 transactions. Positioned near the intersection of low and moderate ranges on both axes, it effectively bridges the gap between lower-value and higher-performing segments. Its consistent turnover, low logistical cost, and cultural appeal make it a strategically valuable product within this category.

While the initial bubble chart (Figure 9) provided critical insight into average profit margins, weight efficiency, and sale frequency guiding assortment planning and promotional design it did not fully account for total business impact. Notably, its focus on average product weight limited its utility for storage, replenishment, and supply chain considerations.

To address this, a second bubble chart (Figure 10)was introduced, shifting the analytical lens to total revenue generation against total physical weight, with bubble size representing transaction volume and color distinguishing material category (Gold or Silver). This view enables an operationally grounded assessment of each product's contribution to revenue, space usage, and sales activity essential for inventory forecasting, warehousing optimization, and capital deployment.

By integrating these three dimensions revenue, physical footprint, and transaction count the chart bridges product-level performance with logistical and financial efficiency. The distribution across the chart can be interpreted through four spatial zones, each revealing distinct patterns in product behavior.

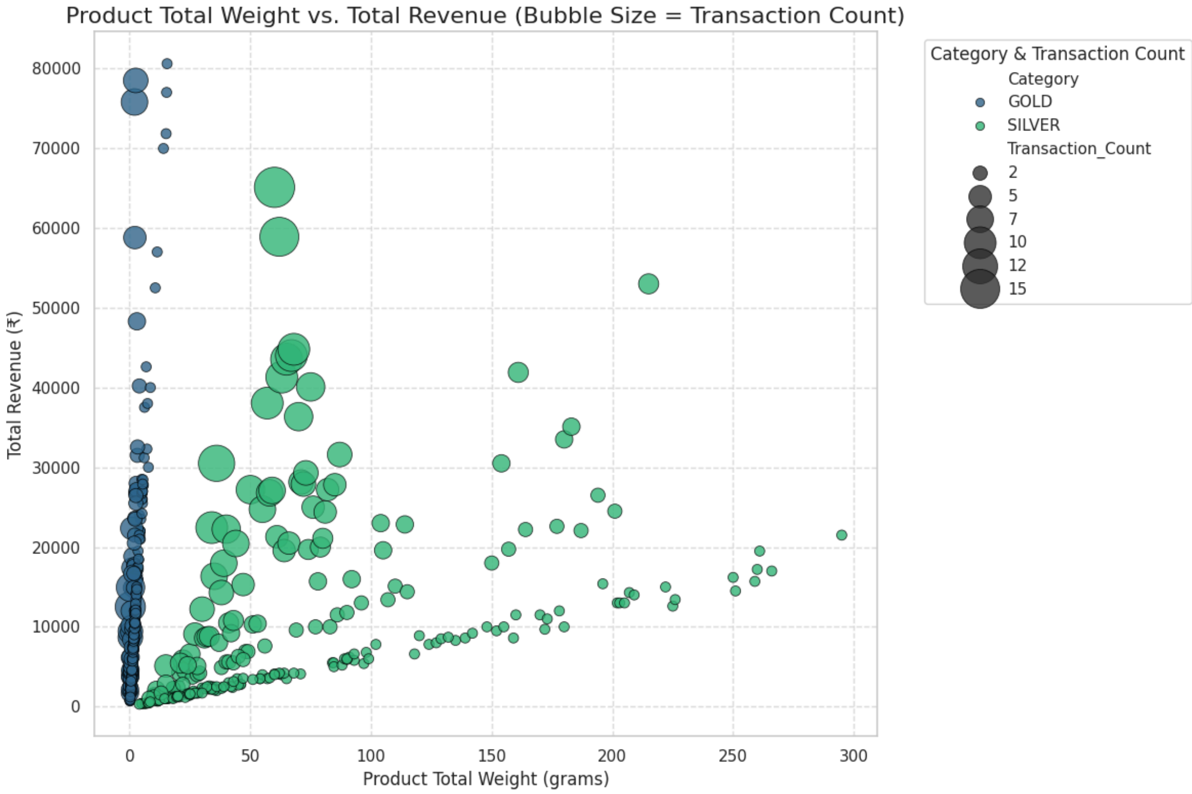

Figure 10 Bubble Chart Showing Product Total Weight vs. Total Revenue (Bubble Size Represents Transaction Count)

Depth and Insight

Top-Left Quadrant – High Revenue, Low Weight: Efficient Value Generators

This region is densely populated by Gold items, characterized by:

- Revenue levels often exceed ₹30,000, even when product weights remain under 25–30 grams.
- Numerous small bubbles not overlapping each other, signifying a high variable weight and small transaction frequency.
- A tight vertical clustering near the Y-axis, showing consistent high value from minimal weight contribution.
- These products deliver high financial returns with medium storage cost, making them ideal for inventory strategies high-margin turnover.

Top-Right Quadrant – High Revenue, High Weight: Revenue-Heavy Logistics Loads

- This zone displays a wider spread of Silver products, often with:

- Revenues above ₹30,000, sometimes reaching ₹60,000–₹80,000.
- Total weights typically in the 50–150g range, with some items extending above 200g.
- Medium to large bubble sizes, indicating moderate to high sales volumes.
- These are high-revenue drivers, but their bulkier physical footprint imposes greater demands.
- Heavier silver jewelry such as anklets, bracelets, and waist chains are often linked to traditional, ceremonial, or regional preferences where value is associated with physical weight.

Bottom-Left Quadrant – Low Revenue, Low Weight:

- This area contains a large number of both Gold and Silver items, characterized by:
- Revenues typically below ₹10,000, with many under ₹5,000.
- Product weights under 20g, often in the single digits.
- Small and medium bubble sizes, reflecting low to medium transaction counts and it can be clearly visualized that they are overlapping each other which means there are weight similarities among the items .
- These items have limited financial impact individually but may carry strategic or cultural value. Their small size minimizes storage burden but also limits revenue contribution.
- Low-cost accessories such as toe rings, rings , baby bangles , chains for silver and om pendant , nose pin , earrings for gold .

Bottom-Right Quadrant – Low Revenue, High Weight:

- This is the least efficient zone, dominated by:
- Low-revenue products (often < ₹20,000) combined with high total weights above 200g or even 250g.
- Scattered bubble sizes, indicating low to moderate transaction activity, but not enough to offset physical cost.
- Primarily Silver products, many of which may be ceremonial or outdated in design.
- These items tie up physical space and capital without proportional return. They may persist but not overstock, good as limited seasonal demand.

Following the multi-dimensional insights gained from the prior bubble charts(Figures 9 & 10), which established the foundational relationship between product revenue, weight, and transaction volume, we now delve into a time-series perspective(Figure 11). This phase evaluates demand volatility and weight-based sales seasonality to aid in optimizing inventory stocking strategies for lightweight (<20 gm) and heavyweight (≥20 gm) product segments.

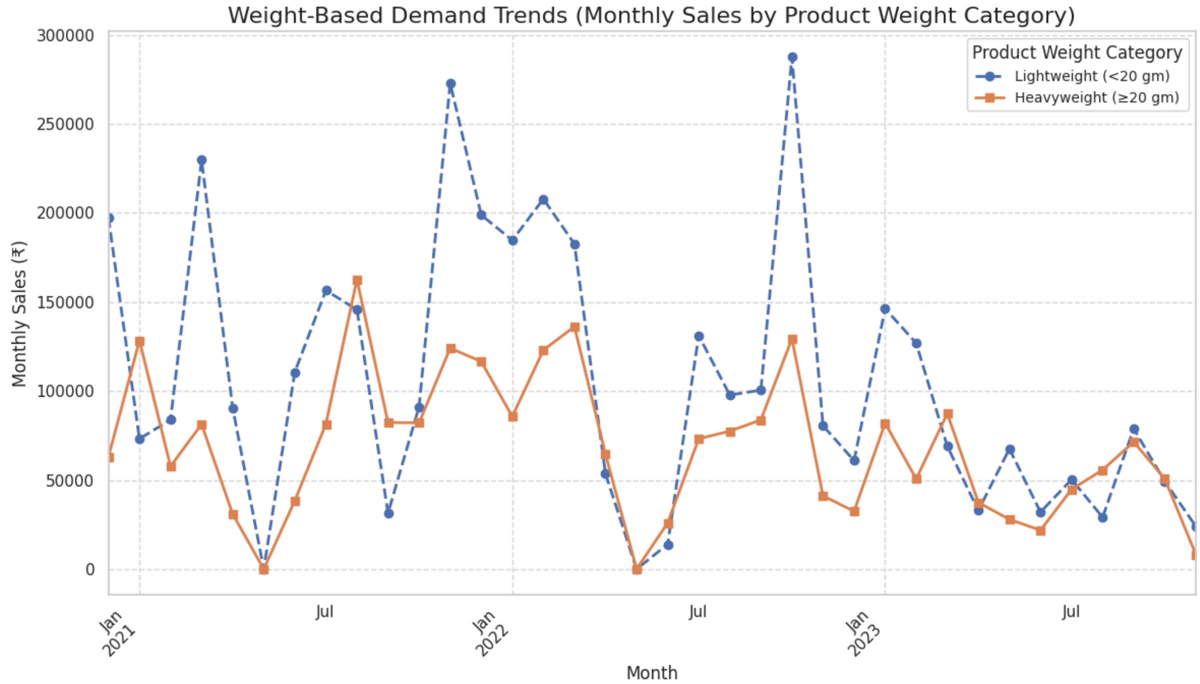

Figure 11 : Line Chart Showing Weight-Based Demand Trends (Monthly Sales by Product Weight Category)

Depth and Insight

The weight-based time-series trend (Figure 11) reveals cyclical demand patterns, clearly distinguishing when to prioritize lightweight versus heavyweight stockkeeping:

Lightweight Products (<20 gm):

- Show consistent spikes in revenue during festive or high-gifting months, such as December 2021, March 2022, and November 2022. These are likely periods of jewelry gifting for events like weddings, Diwali, and Post harvest events .
- Lightweight items predominantly consist of high-margin, low-footprint Gold products (as seen in the top-left quadrant of the bubble chart), which make them prime candidates for lean inventory strategies with high rotation frequency.
- Sales for lightweight items drop sharply during mid-year periods (May–August), suggesting low purchase urgency in off-season months.

Heavyweight Products (≥20 gm):

- Exhibit a steadier demand curve but with less frequent peaks. Spikes occur around November 2021, February 2022, and November 2022, aligning with specific ceremonial or cultural triggers where bulkier jewelry (especially Silver anklets or waist chains) is favored.
- These products are more often Silver-dominated, contributing to heavier physical footprint and slower turnover corroborating the inefficiencies visualized in the bottom-right quadrant of the second bubble chart.

● Due to their higher carrying cost and lower agility in movement, heavyweight products are better managed with event-driven or regional forecasting models, not continuous stocking.

Building upon the weight-based findings, the seasonal sales trend (Figure 12) compares revenue patterns of Gold and Silver jewelry on a month-wise basis. Interestingly, this chart mirrors several patterns from the weight category trend, yet nuanced differences offer a deeper segmentation for inventory strategy.

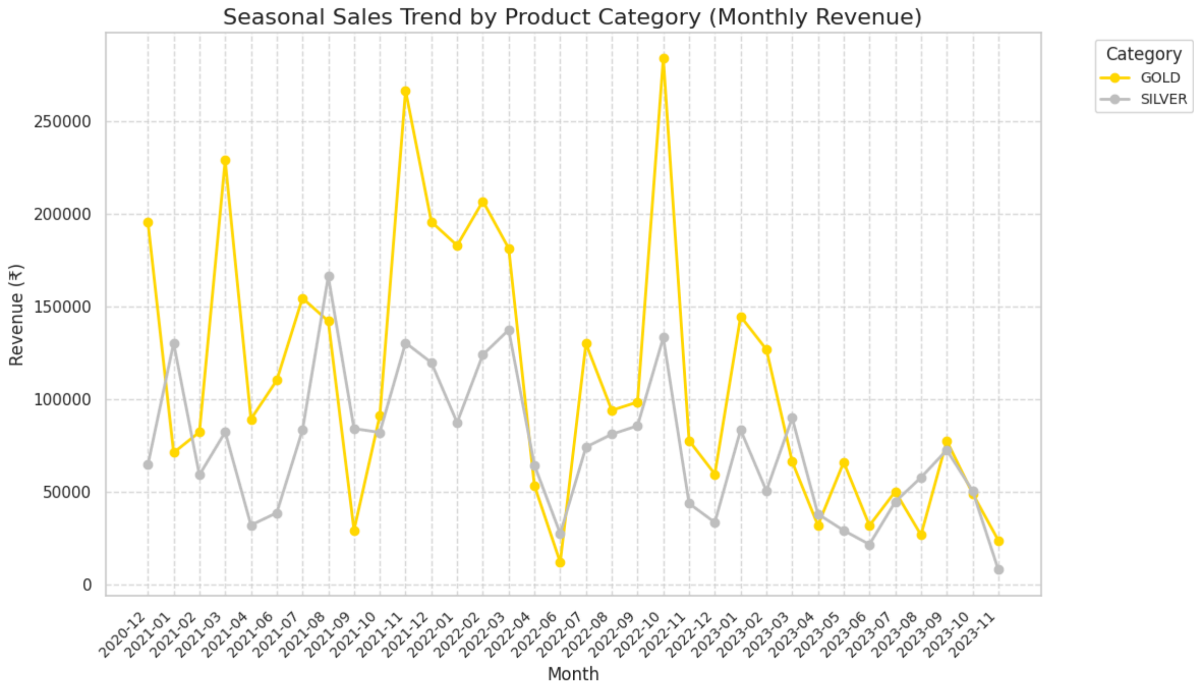

Figure 12 Line Chart Showing Seasonal Sales Trend by Product Category (Monthly Revenue)

Differentiating Factors & Cross-Insights

- ***Gold Revenue Curve (Typically Lightweight):***
  - Closely aligns with the Lightweight product trend in the previous chart.
  - Exhibits sharp and frequent spikes (e.g., Dec 2020, Apr 2021, Nov 2021, Oct 2022), reaffirming Gold's position as a preferred high-value, light-weight purchase.
  - Drops in off-peak seasons mirror lightweight category drops, suggesting that Gold inventory strategies can lean on lightweight sales forecasts for timing.
- ***Silver Revenue Curve (Often Heavyweight):***
  - Demonstrates a more stable, moderate trajectory, matching the heavyweight product trends.
  - Peaks are broader and less abrupt (e.g., Feb 2021, Apr 2022, Mar 2023), indicating steady but less impulsive purchase behavior.

   - While not as volatile as Gold, Silver's heavier average weight and lower revenue-to-weight ratio indicate it should be stocked conservatively and tactically preferably around known demand periods.

While both charts share similar shapes, the amplitude of spikes in Gold/Lightweight sales is notably more dramatic than in Silver/Heavyweight categories.Synchronize Gold Inventory with Lightweight Demand Peaks .Leverage time-series peaks to ensure agile restocking of fast-moving Gold items, optimizing warehouse space and improving cash flows.

4 Interpretation of Results and Recommendation

4.1 Problem Statement 1 : Delayed Credit Repayment and Unstructured Customer Behavior Impacting Profitability

Interpretation: Customer transactions are heavily installment-based, with most buyers opting for two-part payments and an average ticket size of ₹7,873. However, repayment delays are common, with an average lag exceeding 75 days, especially during May–July and December–January, aligning with agricultural lean periods. Conversely, repayment intensity is highest in February, March, April, and October, which coincide with the Rabi and Kharif harvest seasons.

Customer profiling via K-Means clustering reveals three segments:

- Segment 0 (Green): Most valuable cohort 2–8 purchases, spending between ₹21,000 and ₹70,000, often using credit.
- Segment 1 (Orange): Low-frequency (1–2 purchases), low-spend (<₹21,000) typically new, unprofitable customers.
- Segment 2 (Blue): High-value, infrequent buyers spending ₹75,000–₹1.6L on average.

Notably, even customers in Segment 1 with low-value transactions frequently delay payments, indicating credit risk is behavioral, not just financial. July 2022 marked a case of high sales and deep margin compression likely due to aggressive credit without adequate repayment control.

Recommendation:

- Establish a uniform 30–60 day installment schedule, with interest applied on delayed balances to preserve pricing power.
- Restrict credit disbursal during months with historically weak repayment (May–July, December–January), and enable credit in January, February, August, and September so repayment falls during strong harvest-income months.
- Design segment-based policies:
   - Offer credit and retention programs to Segment 0.
   - Move Segment 1 customers to Segment 0 through first-time buyer incentives with capped credit.

    - Engage Segment 2 using personalized high-ticket offers with flexible, premium repayment options.
- Introduce an internal behavioral credit score that tracks past delay duration and frequency across all purchase sizes, helping vet future credit eligibility more effectively.

4.2 Problem Statement 2 : Lack of Seasonal Demand Visibility and Inventory Inefficiencies Across Product Lines

Interpretation: Analysis of 985 sales transactions over 42 months revealed pronounced seasonal variation in product demand and material preferences.

- Lightweight gold items such as earrings (~1.78g, ₹8,973 ATV) and Om Pendants (<2g) peak during March, October, and December, coinciding with weddings and festivals.
- Heavyweight silver items like Anklets (~73.4g, ₹4,659 ATV) and Bangles see higher demand in February, April, and November, post-harvest periods when rural liquidity improves.

Gold items dominate in revenue efficiency   clustering in low-weight, high-margin zones while silver is scattered across higher weights, with revenue driven by volume.

Top-performing items include Anklets, Earrings, Om Pendants, Jhumkis (~4.8g, ₹24,229), and 5-Pendant Sets (~4.3g, ₹21,885).
A distinct pattern emerged in July 2022, where sales surged but margins dropped, confirming that relaxed credit terms without profitability checks can erode business health.

Meanwhile, overstocking of low-frequency, heavy silver items is visible during off-peak months (May–August), leading to poor inventory rotation and capital lock-up.

Recommendation:

- Stock lightweight gold products (1.5g–5g) like Earrings and Om Pendants in preparation for March, October, and December, driven by gifting and wedding demand.
- Procure 60–80g silver items like Anklets and Bangles ahead of February, April, and November, reflecting post-harvest rural buying patterns.
- Avoid building inventory in May–August beyond daily wear essentials   restrict procurement of slow-moving items and repurpose capital for fast-moving stock.
- Balance high-volume sales promotions with margin strategy   especially avoid July-like cycles where sales volume increases at the expense of profitability.
- Optimize warehouse allocation and product lifecycle planning using weight-to-revenue ratios and seasonal turnover to improve space efficiency and ROI.

All the Data and relevant Graphs are present in Drive Folder :

BDM Capstone Project