Optimizing sales and services of a leading Lab equipment firm

A Final Report for the BDM Capstone Project

Submitted by

Name: **V SANJANA**

Roll number: **23f2000029**

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Table of Contents

1 Executive Summary

The final report provides an in depth analysis of Sigma Scientific Products, established in 2015 and headquartered in Chennai which is a leading partnership firm specialized in manufacturing and selling scientific laboratory equipment to engineering colleges and schools for their laboratory needs. Despite its innovative offerings and growing presence, the company faces two major challenges: lack of insights on the product portfolio which is hindering them to rationalize, prioritize and market them effectively; lack of comprehensive understanding of the customer and corresponding sales which leads to ineffective marketing and sales effort.

To address the company's operational challenges, sales data from April 2024 to March 2025 and product data, provided by the business owner, were collected, cleaned and structured into two datasets. Preliminary analysis also confirmed strong information about high variability, with extreme outliers influencing overall revenue. Also, I have made various categorical analyses of data using excel and python libraries and charts to derive insights.The derived insights include information about 'low price high volume' products, 'high price low volume' products etc. As part of customer segmentation advanced techniques like K-Means clustering, RFM (Recency, Frequency, Monetary) analysis and line chart and heatmap based time and geographical analysis were applied.

Key findings revealed that 45% of clients were Mid value, 40% were High and 15% were low value. Cluster analysis proved that low-value customers are large in number but contribute little to the entire revenue.Time series analysis showed that the month of March had the highest sales meaning these months require increased inventory, higher staffing and marketing reinforcement. Product sales were skewed towards a few items like LR-1355, with other categories remaining under-promoted.

Based on these insights, actionable recommendations were proposed.These included flagging long tail SKUs for review, classifying SKUs into price tiers, upsell of medium clusters, win-back & nurturing for low RFM. As a result, Sigma Scientific Products might observe improved engagement rates and higher satisfaction during peak sales by stocking prior. This would also help SIGMA to expand its presence geographically in various neighbouring states. This project successfully transitioned the company from reactive operations to proactive, data informed decision making.

2 Detailed Explanation of Analysis Process

2.1 Detailed cleaning and Preprocessing

A comprehensive data preprocessing was done to prepare the dataset for meaningful analysis. This step was crucial for transforming the raw product and sales data collected from SIGMA Scientific Products Limited, covering April 2024 to March 2025 into a structured format suitable for time series modelling and business insights. The cleaning process addressed several inconsistencies and anomalies, as outlined below:

- **Normalizing column names:** The initial dataset had a lot of extra white spaces which led to the same industry having a different identity. Hence, column names were normalized by trimming whitespace.
- **Datetime Conversion:** Normalized into a datetime column and derived Month for monthly analysis.Certain months were of integer and rest were of words. These were standardised into words.
- **Product key as Foreign Key:** Normalized product key by joining Sales's Product code with Products's Model No.to attach Category, Product Name, Usage, Approx. Price, etc.
- **Cleaned Approx. Price (₹):** The column Approx. Price (₹) into numeric where present by removing symbols like comma and rupees.
- **Customer Name Extraction:** The Sales data had a column called Particulars which had the Customer Name, Location, State for some entries. Therefore, while running the dataset, there were errors as the same customer was treated as different because of this extra information. Hence, a new column called Customer Name was created and used for further analysis.
- **Standardized customer names:** The sales dataset had different representations for the same customer.For example some were entered as “VIT” and some as “VIT University” for accurate customer segmentation the customer names where standardized.
- **Location extraction:** Extracted the location for mapping from the Particulars column in the sales dataset.Also, the state in which the consumers are located where also found out manually using Google Maps and verified with the owner.
- **Merging the datasets:**The product code in Product_data.xlsx were mapped to the Sales_data.xlsx with the Approx price for creating a merged dataset to work with.
- **Missing data:** The dataset was also checked for missing values using python commands.

Data cleaning is important ensuring the integrity of insights and the effectiveness of models. Unclean data leads to misleading patterns and overfitting in machine learning.
By resolving inconsistencies, maintaining uniform data structure and removal of noises enhances the relevance and quality of analysis.Hence, data cleaning helps us model real-world patterns more accurately and draw actionable conclusions that support the company's strategic objectives.

2.2 Feature Engineering & Exploratory Data Analysis (EDA)

Feature engineering was performed on cleaned raw data, to extract relevant variables aligned with the problem statement. Non-essential fields such as "Key Attributes", "Usage" from the Product data.xlsx and "Vch Type" column from Sales data.xlsx were dropped.

Key columns like "Product Code", "Category", "Product Name", "Price" and "Sold By" were retained in the Product dataset and "Date", "Particulars", "Credit" in Sales data were used to form a cleaned, focused dataset.

A new feature called location was added to the cleaned sales dataset for doing geographical analysis by extracting the location details from the particulars column in the sales dataset.

2.3 Product Portfolio Analysis

After cleaning the dataset, I first analyzed Product_data.The column category and price was analyzed well for deriving insights about the product portfolio.

2.3.1 Category distribution analysis

A bar graph *Figure 3.1* was plotted to understand the category distribution in the products present in the portfolio. This is the best way to understand the distribution rather than a pie chart as there are more than 5 numbers of categories. The x axis was labelled as "category" and the y axis had the number of different products in that particular category and was named as "count".

2.4 Comparative Pricing and Sales Intelligence

2.4.1 Price spread analysis

I plotted a box plot *Figure 3.2* for the continuous values in the dataset. Also here, since the prices of the products belonging to same category varies, to find a uniform price for one product, the mean has been calculated using this formula

Approx. Price (₹) =(Sum all products grouped by category)/(No of products in that category)

Q1 = 25th percentile (lower quartile)

Q2 = 50th percentile (median)

Q3 = 75th percentile (upper quartile)

Inter Quartile Range (IQR) = Q3 - Q1

IQR tells the spread of the middle 50% of a dataset showing how clustered or spread out the central values are.

2.4.2 Price vs Quantity Correlation

I counted how many sales rows reference each Product code has which becomes the total_count. The product manager is merged with the cleaned sales data using the Product code which attaches the Category and Price of the product in the Product data for downstream aggregation and plotting.

Category level popularity = sum (total_count across all products in the same category)

Using the above information a scatter plot *Figure 3.3* is plotted between the number of products sold and its approximate price as the company has sold products of the same category at different rates according to the customer's demand.

2.4.3 High-Value vs Low-Value Products

The sales value is calculated by grouping the unique product codes as one and by summing up all the credit values made by those products from the month of May 2024 to March 2025. The values are sorted and the top 10 products which are responsible for high revenue and low revenue are found out using the head and tail functions in Python. A bar chart *Figure 3.4* is plotted to derive insights.

2.5 Customer Segmentation

2.5.1 Customer Segmentation using K-Means on Sales data

The cleaned Sales data was loaded. The Data column was converted into a Date Function to perform further analysis. A separate row called Quantity was created by:

Quantity of a product = Total number of products of that particular Product Code

Similarly,
Num Transactions = sum(Vch No's of grouped by Customer Name)

A column called "Total Sales" was made by aggregating the rows of a Customer Name which was extracted from the row called Particulars in the Sales data while data cleaning.

Total Sales = sum(Credit values of grouped by Customer Name)

A column called "Avg Order Value" was created by dividing the Total Sales and Num Transactions

Avg Order Value of a Customer = Total Sales Value/ Num Transactions (of that Customer)

The selected features, Total Sales, Total Quantity, Num Transactions and Avg Order Value are scaled using the StandardScaler() Function. Scaling is crucial in K-Means clustering because the algorithm is highly sensitive to the magnitude of the features. K-Means relies on distance calculations, usually Euclidean distance to determine similarity and assign data points to clusters.
The Scaled dataset is then fit into the KMeans model with 3 clusters as parameters.

The formulae used by K-Means are:

<u>Euclidean Distance</u>

Distance = $\sqrt{\Sigma}$ **(** x-$\bar{x}$ **)^2**

The Euclidean Distance is used to measure the distance between a data point x and a cluster centroid. This determines which cluster a point is assigned to where $\bar{x}$ is the centroid of the cluster where x belongs .

<u>Centroid Update Formula</u>

The centroid of a cluster is recalculated by finding the mean of all data points assigned to that cluster.

<u>Within-Cluster Sum of Squares</u>

The Within-Cluster Sum of Squares (WCSS) gets minimised. It is the sum of the squared distances between each point and its assigned cluster's centroid.
A plot is displayed, Figure 3.5, to analyze and find results.

2.5.2 Customer Segmentation using RFM Analysis

R - Recency “How recently did a customer buy from you?”
F - Frequency “How often does a customer buy from you?”
M - Monetary “How much money did the customer spend in total?”

RFM is calculated here to identify top customers, customers who may leave, big spenders etc., for marketing segmentation and personalized offers.

Data taken into consideration are quantity and the Total Amount of each customer.

R_Score = Today - Last Purchase Date
F_Score = Count of unique Vch No. grouped by Customer Name
M_Score = sum(Credit grouped by that Customer Name)

The RFM was converted into scores by dividing the values into 4 buckets.
Value = 4 => very recent buyer, buys often, spends the most
Value = 1 => long ago buyer, infrequent buyer, spends the least

RFM_Score = R_Score + F_Score + M_Score   3<= RFM_Score <= 12
A distribution was plotted to find how many customers fall in each RFM Score group.
Figure 3.6 and 3.7 using the above information.

2.5.3 Time series analysis (Monthly sales)

The data had different formats of date in the Date column. For example, some were noted as 14-Apr-2025 and some were noted as 15-04-2024. These dates were standardised for further analysis.

To analyze the monthly income, a line chart, *Figure 3.8* is plotted by extracting the month and year from the Date column in the sales data.

Monthly sales amount was calculated by grouping the credit month wise.

The value was converted into a time stamp. A line chart was plotted using the month-year as x axis and the total sales made in that month in y another axis.Here the peak and low sales cycles are identified for further forecasting.

2.5.4 Geographical Sales Analysis using Heatmap

State names were derived from the Particulars column of the Sales data. The data was grouped by State and the Total Amount column was summed up.

A heatmap *Figure 3.9* was plotted to understand the price distribution among various States.

3 Results and Findings

3.1 Product Portfolio Analysis

3.1.1 Category distribution analysis

The below bar chart reveals how products are distributed across major categories. A bar chart is chosen over a pie chart to show the category distribution as there are more than 5 categories in the current product portfolio.

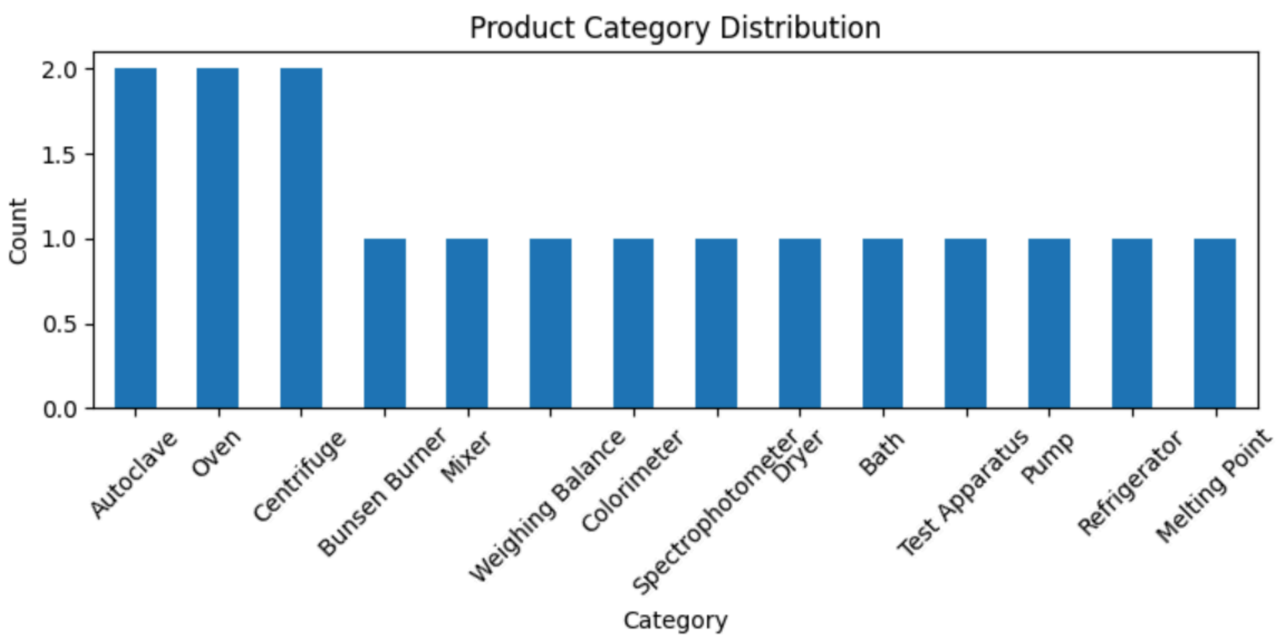


Figure 3.1:Category distribution analysis using bar chart

The top 3 categories which includes Autoclave, Ovens and Centrifuge form the bulk of the product portfolio, while all remaining categories contribute marginally.Category overcrowding comes into play if one or two categories dominate.This indicates potential overlap within that category learning to possible SKU duplication.There is a long Tail of products as the like Bunsen Burner, Mixer and so on which includes categories with very few products.These may not contribute meaningfully to sales and can be reviewed for discontinuation.A healthy product portfolio should show reasonably distributed categories.
Too many products in a single category leads to complexity in marketing, stocking and pricing which is not prevalent in this product portfolio.

3.2 Comparative Pricing and Sales Intelligence

3.2.1 Price spread analysis

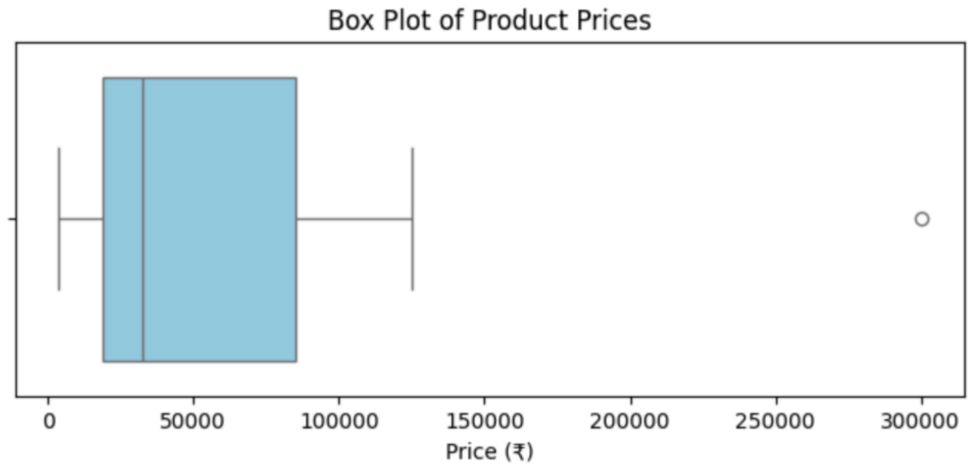


Figure 3.2:Price spread analysis using box plot

The box plot reveals price dispersion across all products including:

Variable	Minimum price	Q1	Median (Q2)	Q3	Maximum price	Mean	Standard deviation
Value (₹)	3,600	18,900	32,500	85,000	3,00,000	63,111	73,191.00

IQR = Q3-Q1 = ₹66,100

Large Price Variability (High IQR) indicates high inconsistency in pricing, even possibly within the same product category. The presence of extreme outliers are because some products are priced significantly higher, such as the LR-1335 which costs around ₹3,00,000. This is highly specialized lab equipment which is overpriced slow-sellers. These products require individual profitability evaluation.

The median helps determine the core pricing zone.
Products deviating significantly which includes the LR-1355, VO-300, TD-12 etc., may need repricing or better positioning to ensure steady marketing and income.

3.2.2 Price vs Quantity Correlation

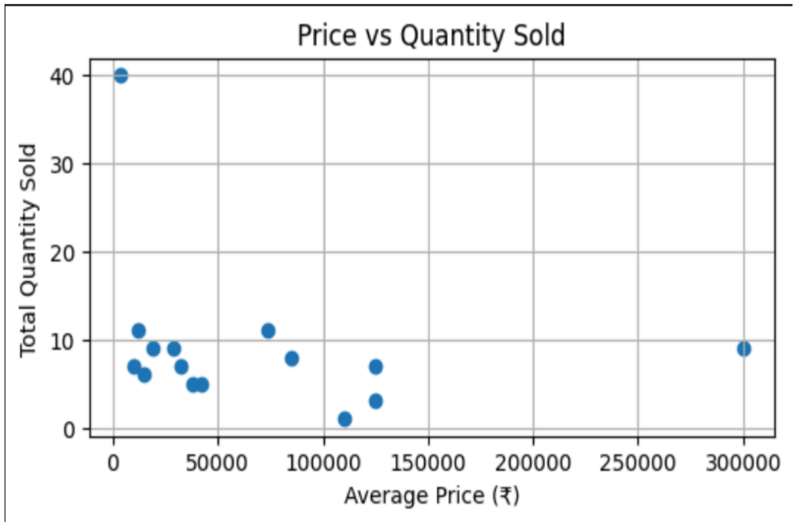


Figure 3.3:Price vs Quantity Correlation using scatter plot

The scatter plot comparing product average price vs. quantity sold. Overall there is a negative relationship as the bulk of high-volume sales come from lower-priced items while higher-priced items typically sell fewer units. The plot also shows distinct sub-groups that require different business actions. A low-price outlier with very high quantity indicates a volume-driven bestseller . Several mid-price items cluster around moderate sales roughly around 3 to 11 units, representing steady performers. There are also a few high-price, low-quantity products around ₹3,00,000 with approximately 9 units and others between ₹80,000 to ₹1,30,000 with very low volumes which appear niche or underperforming relative to price.There are low-price, low-quantity points that indicate items that are not moving despite being inexpensive.

Low-price, high-volume items such as BB-E01 are the workhorses of revenue through volume. Maintaining supply, ensuring availability, and considering incremental margin improvements like optimizing the cost or increase in cost if demand is inelastic.These products are ideal for promotional bundles and cross-sell placements because they drive transactions.

Low-price, low-volume items are poor performers despite low price. Evaluating them for discontinuation or replacement is important.

High-price, low-volume items might be legitimate niche SKUs, mispriced or suffering from poor marketing/awareness. Here, LR-1355. Though the revenue is mainly because of this product, the volume of this product being sold is very less.

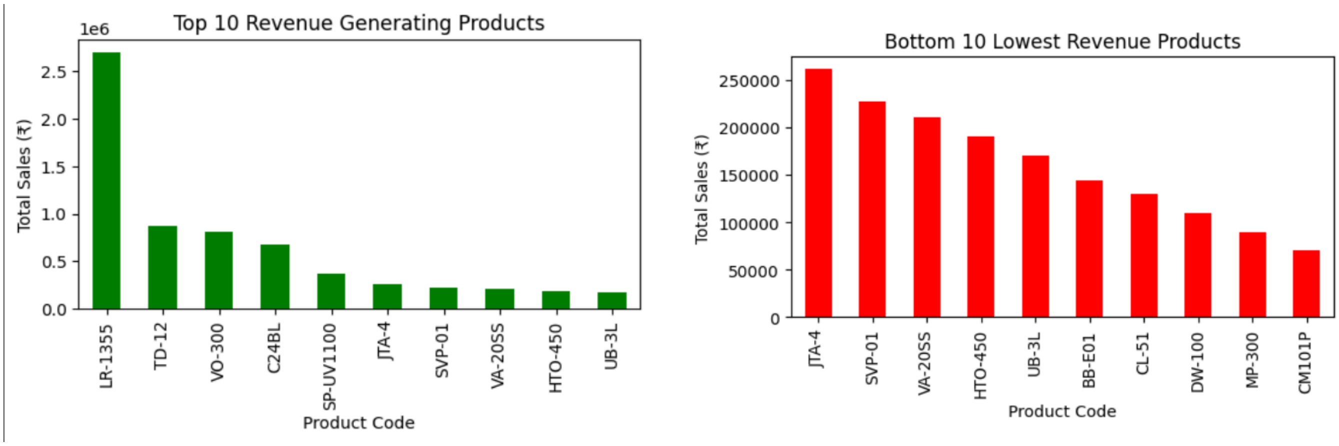


Figure 3.4:High-Value vs Low-Value Products

The first chart highlights the highest earning products in the portfolio. The standout findings are:Extreme revenue concentration by LR-1355 alone generates over ₹26 lakhs far higher than any other product.This product is a high value, low volume scientific instrument, indicating it is a core revenue driver and must be protected in the portfolio.

These mid range products such as TD-12, VO-300 help balance the portfolio by contributing reliable income even if the top SKU is volatile.

The bottom chart lists products with very low total sales revenue even though some have been sold.Several products show consistent low revenue.A few products such as VA-205S, HTO-450 appear in both charts which suggests inconsistent performance across time periods or customer segments.These items may be dependent on a few institutional buyers or seasonal cycles.

Low price, low volume products like MP-300 and CM101P show poor turnover making them candidates for repricing or discontinuation.

3.3 Customer Segmentation

3.3.1 Customer Segmentation using K-Means on Sales data

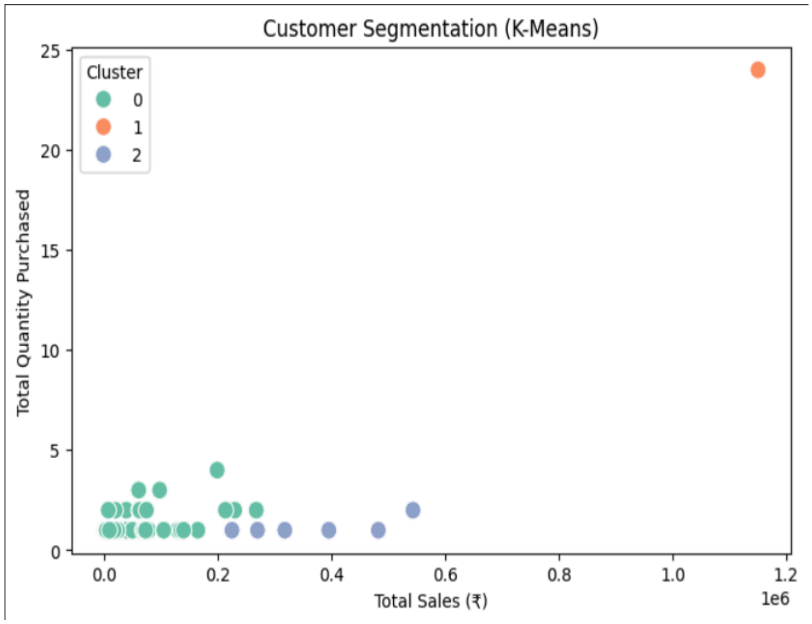


Figure 3.5: Customer Segmentation using K-Means

Around 15% of customers contribute over 60% of revenue.Low-value customers are large in number but contribute little (cluster 2)

Cluster 0: High Value Customers

- Highest total purchase value
- Large and frequent orders
- Buy across multiple product categories.
- Should be target for premium offerings, loyalty plans

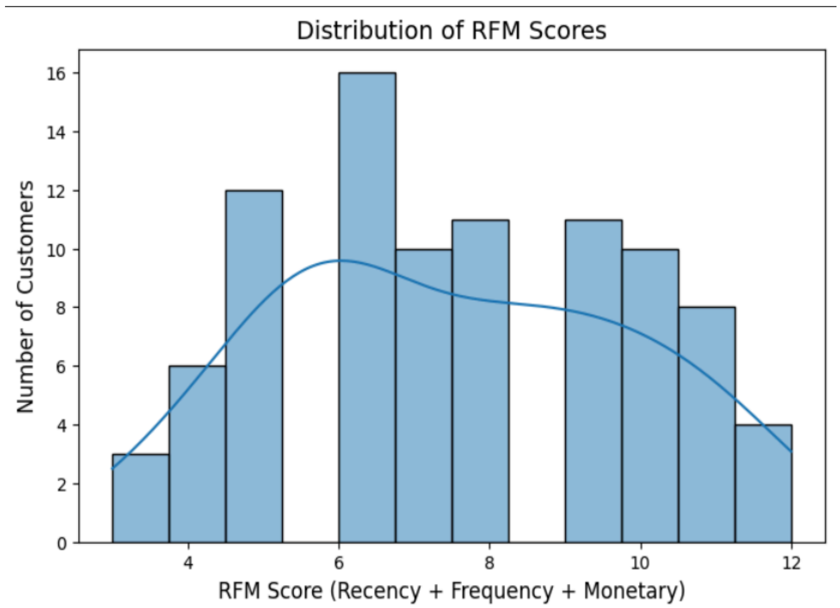
Cluster 1: Medium Value Customers

- Moderate sales contribution
- Regular but smaller purchase
- Good candidates for upselling

Cluster 2: Low Value or One-time Buyers

- Low frequency
- Very small order sizes
- Should receive promotional offers, beginner bundles, price incentives

3.3.2 Customer Segmentation using RFM Analysis



Customer Name	Recency	Frequency	Monetary	R_Score	F_Score	M_Score	RFM_Score	Customer_Value_Segment
AIC Specialities	278	1	13500.0	2	1	2	5	Low Value
Accuracy Analabs Pvt Ltd	345	1	55000.0	1	1	3	5	Low Value
Aic Enterprises	312	1	2840.0	1	1	1	3	Low Value
Alfa Rubber & Springs Pvt Ltd	173	1	3975.0	2	1	1	4	Low Value
Altra Chemtech Pvt Limited	287	3	97800.0	1	4	4	9	High Value
...	...	...	...	...	...	...	...	...
V.B Ceramic Consultants	81	1	140000.0	4	4	4	12	High Value
VIJAYANI NUTRACEUTICALS PRIVATE LTD	51	1	17875.0	4	4	2	10	High Value
VIT University	106	4	199000.0	3	4	4	11	High Value
Veltech Rangarajan Dr.Sanguthala	99	1	9500.0	3	4	1	8	Mid Value
Yogivemmana University+	57	2	7500.0	4	4	1	9	High Value

Figure 3.6: Customer Segmentation using RFM Analysis

The RFM score ranges from 3 to 12, where:
Higher score = better customer (recent, frequent, high spending)
Lower score = weaker customer (old purchase, infrequent, low spending)
Most customers have scores between 6 and 10, which forms a balanced bell-curve-like distribution.

High RFM Scores (8 -12) says they are the loyal and high-value customers as they have recently purchased or they order frequently or produce the highest monetary value or all. Here, Harikrish Technologies, The Precision Scientific Co etc., have high RFM Scores. These companies should receive priority service, premium catalog and exclusive discounts.

Medium RFM Scores (5-7) should be targeted with regular engagement marketing. These companies are stable, but could churn. For example., Godrej Consumer Products Ltd.
Low RFM Scores (3-5) are comparatively inactive, low frequency product buyers. They might requires reactivation campaigns, email reminders or discount offers.

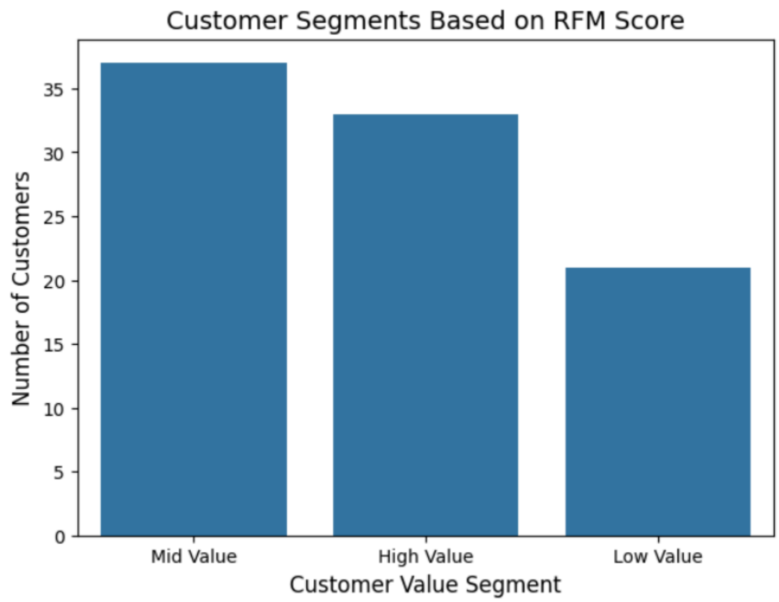


Figure 3.7: Number of customer in various segments

Majority of customers fall in the mid-range which suggests most of the companies customers buy somewhat frequently but aren’t extremely loyal or extremely inactive.Smaller groups at the extreme ends of the graph show that very low score customers have stopped buying or buy very rarely and very high scores implies strong loyal customers who purchase recently and frequently.
The distribution is relatively smooth. Here there are no sharp peaks or drops, meaning the customer behavior is well-distributed and there are no extreme abnormalities.

3.3.3 Time Series Analysis - Monthly

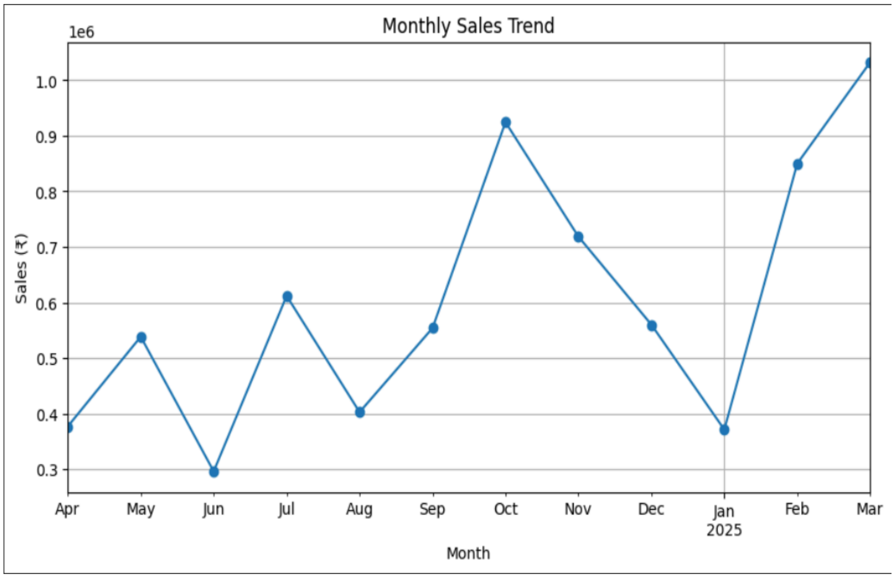


Figure 3.8: Time series analysis month wise

Peak sales months are identified when months consistently show high sales.The month of March shows peak sales which means these months require increased inventory, higher staffing and marketing reinforcement.Low Sales Months includes the month of June and January.These periods show dips. Hence, the company has to optimize inventory and reduce stock holding cost.

Seasonality Patterns are there in the months of February and March. As discussed with the owner the main reason behind this is that colleges and industries try to exhaust their funds by the end of March.With this, the company can estimate next month's demand based on trend direction.

3.3.4 Geographical Sales Analysis

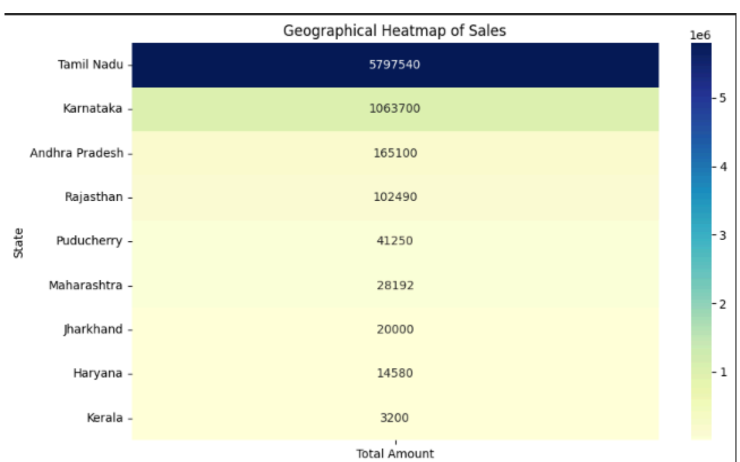


Figure 3.9:Geographical Sales Analysis using Heatmap

High performing states include Tamil Nadu, Karnataka and Andhra Pradesh as they contribute disproportionately to total revenue.These states should be the focus of sales representatives, dealer support inventory placement.

Medium performing states are Andhra Pradesh and Rajasthan. They have potential for growth. They should be targeted with promotions or educational outreach.

Low performing states are Kerala, Haryana, Jharkhand and Maharasthra. These could be due to lack of distributors, weak brand visibility and competitor dominance.

4 Interpretation of Results and Recommendations

4.1 Product Portfolio Analysis

Category distribution analysis

Interpretation of results

Autoclave, Ovens and Centrifuge are the top 3 categories and form the bulk of the product portfolio whereas the rest of the categories are long-tail items with very few SKUs each.
Top categories concentrate product complexity and potentially sales effort. If many SKUs in a category are very similar, internal cannibalization and unnecessary inventory cost might happen.
Long-tail categories, that is single SKU categories may not contribute materially to revenue but increases catalog maintenance and SKU management overhead.
The Product Portfolio is wide and partially skewed.A few categories hold most SKUs.This raises the risk of duplication.
The long tail suggests many niche or legacy SKUs that may be underperforming.

Actionable Recommendations

1. **SKU feature matrix**: For Autoclave, Oven, Centrifuge producing a side by side matrix of key attributes such as capacity, temp range, material, warranty. Label SKUs as “distinct tier” vs “overlap candidate”.
2. **Flag long-tail SKUs for review**: A method can be created such that if SKU annual sales < X and no orders in the last 12 months the company can mark for "catalog on demand" or retirement.

4.2 Comparative Pricing and Sales Intelligence

Price spread analysis

Interpretation of results

The box plot shows large price dispersion like min ₹3,600, median ₹32,500, mean ₹63,111, max ₹300,000, IQR ₹66,100 and clear outliers (LR-1355). High IQR and heavy right tail show two distinct business models: Low cost fast movers vs High cost equipment.
High priced outliers drive mean and may tie up capital that is inventory risk if they sell slowly.
There are distinct clusters which include fast moving low price items which are volume drivers and high value low volume items which may be for project or institution sales.
Some high-price SKUs such as LR-1355 are revenue critical but risky as the company is heavily dependent on that product.

Actionable Recommendations

Classifying SKUs into price tiers: The tires might include Low / Mid / High using quartiles.Use these tiers to set inventory.

Core SKU protection: For top revenue SKUs like LR-1355 create supplier redundancy and safety stock policy.

Repricing and repositioning experiments: For mid/high items that underperform, small price tests or targeted marketing can be experimented.

Price vs Quantity Correlation

Interpretation of results

A Negative relationship exists as low priced items sell more units while high priced items sell fewer units. This can be separated into distinct sub groups:

- Low price, high volume : SKUs are revenue drivers via transactions and they need high availability and efficient replenishment.
- Low price, low volume : SKUs are poor performers and candidates for discontinuation.
- High price, low volume : SKUs might be specialized and require targeted sales and service.

This shows the need for differentiated inventory and marketing strategies by SKU class.

Actionable Recommendations

1. **Inventory policy by SKU class**: Fast movers will need multiple replenishment sources. Long tail, the products which do not provide much income can be made to order or limited stock.
2. **Bundle strategy**: Bundling up fast movers with mid range SKUs might increase attach rate.
3. **Validate niche SKUs:** The niche SKU products' rates can be negotiable by interviewing top account representatives and checking competitor offerings before deciding.
4. The Centrifuge currently sold by SIGMA is a white-label product manufactured by REMI. Despite being externally produced, it consistently performs well in sales and shows strong customer acceptance even though it belongs to the long tail of products being sold.Given this high demand, SIGMA should evaluate the feasibility of transitioning from white label sourcing to in house manufacturing. This shift could help SIGMA capture higher profit margins, improve control over product quality and reduce reliance on third party suppliers.

Interpretation of results

Revenue is heavily concentrated in a small number of SKUs like LR-1355 as they contribute more to the company's revenue. As discussed there is a long tail of underperformers.
Core SKUs should be protected and prioritized in sales & inventory. Long tails should be rationalized with clear rules as catalogue on demand or retire those products but the company is dependent on these very small sets of hero SKUs.

Supply chain protection for these items is critical because any disruption directly hits revenue. Marketing and sales campaigns should heavily support the top 3-5 products.

Actionable Recommendations

1. **Core SKU program**: Top 10 SKUs might get expedited procurement, premium listing, dedicated account managers.
2. **Reporting cadence**: Monthly review of top SKUs and long tail with actionable changes might be important.

4.3 Customer Segmentation

Customer Segmentation using K-Means

Interpretation of results
Three clusters:

- High-value -> few customers, large purchases
- Medium-value -> steady
- Low-value / one-timers

This shows where to allocate personalized sales effort vs automated marketing.

A small set of customers which is the top cluster provides the majority of revenue; these are the core accounts.
A large number of low value customers create noise but represent an opportunity for scale via automated channels.

Actionable Recommendations

1. **Account based plays for high-value clusters**: Ensuring prioritized outreach, service contracts, customized bundles will always ensure good probability of having them intact.
2. **Upsell programs for medium cluster**: Targeted campaigns with mid-tier bundles is important to gather their interests in buying more or additional products.
3. **Automate for low-value cluster**: Constant mail nurtures, low cost remarketing would attract the low value cluster of people to get more products from the company.

Customer Segmentation using RFM

Interpretation of results

Mid Value is the largest customer base in this company.These customers are active enough but have a high chance of becoming high value if they are nurtured properly or Low value if ignored.

A strong high value customer base shows customers who are consistent and repeat buying, customer loyalty and higher lifetime value.

Low Value Segment is significant because around 20+ customers have not purchased recently, low monetary contribution and low engagement
These customers are at high risk of churn.

RFM segments directly inform which customers to retain, upsell or re-engage.

- High RFM customers: High RFM customers are the most valuable buyers because they purchase frequently, spend more and buy recently. These customers should receive priority service to maintain their loyalty and maximize long term revenue. Investment can be in the forms of retention by providing services or exclusive offers.
- Low RFM customers: Low RFM customers include buyers who have not purchased in a long time, buy infrequently or spend very little.Executing run win-back campaigns to bring back customers who stopped buying from the company.

Actionable Recommendations

1. **Priority service for High-RFM**: This may include offering strict Service Level Agreements (SLA) for faster delivery or support, sending proactive reorder reminders before they run out of stock and providing special contract pricing or exclusive deals. By giving personalized attention to high-value customers, businesses can increase retention and strengthen long-term relationships.
2. **Win-back & nurturing for Low-RFM**: These customers might require re engagement strategies to bring them back. Win-back campaigns can include targeted discounts,

limited-time promotions or personalized messages reminding them of product value. Nurturing strategies such as product education, usage tips and scheduled follow up campaigns help rebuild interest and encourage them to return. The goal is to re activate these inactive customers before they are permanently lost.

Time Series Analysis - Monthly

Interpretation of results

The peak is found to be around March which seems to be the end of fiscal or academic spend. It dips again in January & June. Patterns likely linked to institutional purchasing cycles.

Understanding this enables planning inventory, staffing and marketing campaigns to match demand.High seasons require increased inventory and staff while slow months require cost control.

Actionable Recommendations

1. **Seasonal planning**: This might increase procurement lead times ahead of peak months and align promotions to slow months.
2. **Safety stock calibration**: Setting higher safety stock for SKUs with peak demand in March will be very useful in generating revenue.

Geographical Sales Analysis

Interpretation of results

Tamil Nadu, Karnataka and Andhra Pradesh are top performing states while other states underperform comparatively.Regional focus directs local salesforce placement, dealer appointment and distribution centers which plays a major role in the revenue generation.
Regions with high sales should get localized investment by sales representatives and stocks.
Low performing regions may need distributor recruitment or paid marketing.

Actionable Recommendations

1. **Regional GTM plan**:
    - High performing: Investing in the field of sales and local spare parts warehouses might make them a frequent buyer.
    - Medium: Targeted promotions and distributor outreach might let them buy more products.
    - Low: evaluate distributor channels or pilot campaigns to test demand.

2. **Geographic bundling**: Stocking of fast movers in local warehouses will reduce lead times.