Improving Profitability and Customer Retention for a Rural Latex Business through Data-Driven Strategies

A Mid Term report for the BDM capstone Project

Submitted by

Name: **EMILIN VIJU SAMUEL**

Roll number: **22F2000918**

IITM Online BS Degree Program,

Indian Institute of Technology, Madras,

Chennai Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project Title "**Improving Profitability and Customer Retention for a Rural Latex Business through Data-Driven Strategies**". I extend my appreciation to **Kizhakkekara Latex**, for providing the necessary resources that enabled me to conduct my project.

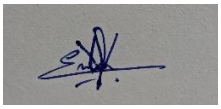
I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate: 

Name: EMILIN VIJU SAMUEL

Date: 08/09/2025

1 Executive Summary and Title

Project Title: "Improving Profitability and Customer Retention for a Rural Latex Business through Data-Driven Strategies"

Kizhakkekara Latex, a rural latex intermediary located in Maruvanchira, Anchal, Kerala, plays a crucial role in connecting small-scale rubber farmers with an export-oriented latex company. Despite operating successfully for over nine years, the business faces several challenges including fluctuating commission-based revenue, dependence on manual record-keeping, lack of permanent staff, and limited customer outreach.

This project aims to leverage Business Data Management (BDM) techniques to address these issues. Specifically, it focuses on three main objectives:

1. Improving profitability through forecasting income and analyzing commission trends.
2. Enhancing customer retention by identifying and prioritizing high-value suppliers.
3. Introducing simple digitalization and workforce strategies to improve operational efficiency.

The dataset consists of transactional records from the past five months, including customer names, sale dates, latex volumes (kg of DRC), pricing, and commission values. Using tools such as Excel pivot tables, descriptive statistics, and forecasting techniques, patterns in revenue, customer engagement, and operational load are being analyzed.

The expected outcome is a set of actionable strategies that will help stabilize revenue, retain key suppliers, and enhance decision-making through digital tools. This report details the originality of the dataset, descriptive statistics, analytical methods, and preliminary findings that form the foundation for the final project outcomes.

2 Proof of Originality of Data

As part of the proof of originality for the collected data, I am providing the following supporting documents and materials: Letter from the Organization, Images of the yard and plantation owned and a video capturing an interaction with the owner. These are securely attached via a Google

Drive link for your reference and review:

https://drive.google.com/drive/folders/1bY3k0CsQ_ObpOzaGGR3aayin3-Cyyvyr?usp=drive_link

(copy and paste the link in browser please)

The mode of the video is in Malayalam Language majorly and an English transcript is provided in the google drive link for your reference and review. Please note that the number of customers mentioned is more than that in the sheet due to decrease in customers because of personal issues experienced by the owners, as the video was recorded in June.

Photos of the yard and plantation have been added in the same drive link for your reference.

3 Metadata and Descriptive Statistics

For this project, data was systematically collected from 1st February 2025 to 29th June 2025, tracking the sales and revenue generated at Kizhakkekara Latex.

The dataset was digitized into structured columns as follows:

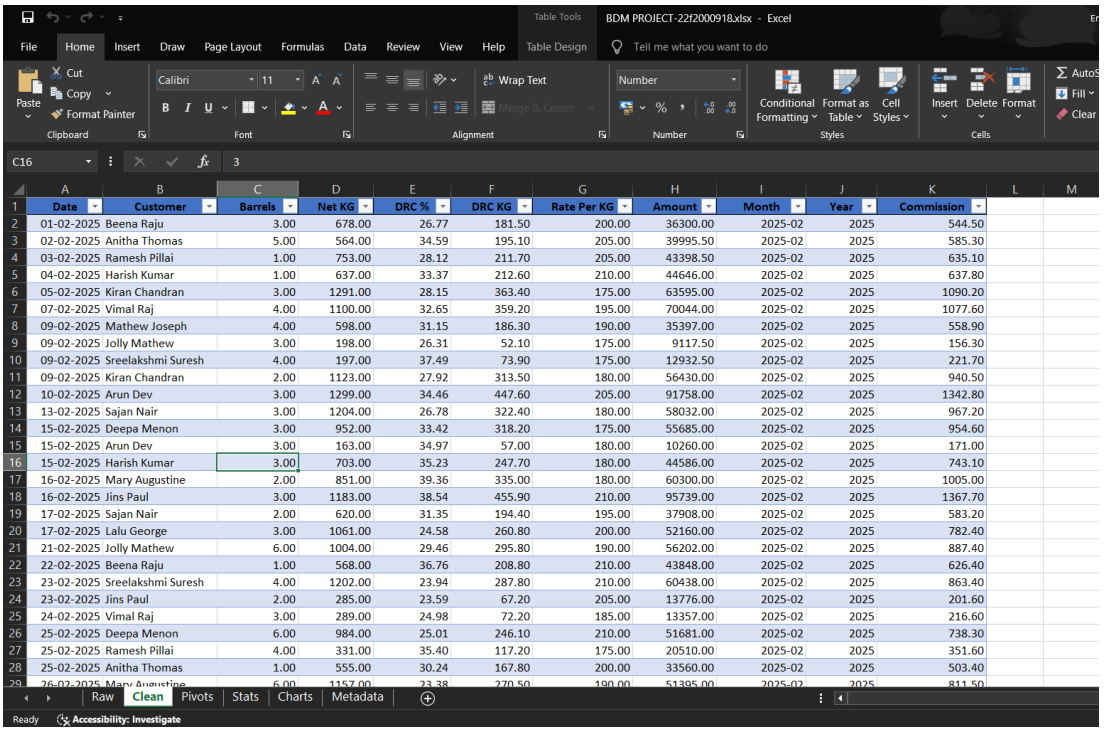
1. **Customer** – Name of supplier/farmer delivering latex.
2. **Date** – Date of sale (DD/MM/YYYY).
3. **DRC KG** – The latex supplied in kilograms of dried rubber content.
4. **Rate per kg** – Price determined by dealer for that period.
5. **Amount** – Quantity × Rate.
6. **Commission (₹3/kg)** – Income earned by the business.

This dataset covers **five months of operations**, capturing customer patterns, revenue flows, and variability in supply.

Once the data was collected, it was transcribed into Excel for detailed analysis. Using Excel , tools such as formulas, pivot tables, and charts were employed to uncover trends in sales patterns, revenue streams, and seasonal fluctuations.

Link to the Data:

https://1drv.ms/x/c/16d5316031157c1e/EZz-z-uFTGVBllE1SeVfMUgBGTH_fZiuTUiYHYg2A_nDig



Date	Customer	Barrels	Net KG	DRC %	DRC KG	Rate Per KG	Amount	Month	Year	Commission
01-02-2025	Beena Raju	3.00	678.00	26.77	181.50	200.00	36300.00	2025-02	2025	544.50
02-02-2025	Anitha Thomas	5.00	564.00	34.59	195.10	205.00	39995.50	2025-02	2025	585.30
03-02-2025	Ramesh Pillai	1.00	753.00	28.12	211.70	205.00	43398.50	2025-02	2025	635.10
04-02-2025	Harish Kumar	1.00	637.00	33.37	212.60	210.00	44646.00	2025-02	2025	637.80
05-02-2025	Kiran Chandran	3.00	1291.00	28.15	363.40	175.00	63595.00	2025-02	2025	1090.20
07-02-2025	Vimal Raj	4.00	1100.00	32.65	359.20	195.00	70044.00	2025-02	2025	1077.60
09-02-2025	Mathew Joseph	4.00	598.00	31.15	186.30	190.00	35397.00	2025-02	2025	558.90
09-02-2025	Jolly Mathew	3.00	198.00	26.31	52.10	175.00	9117.50	2025-02	2025	156.30
09-02-2025	Sreelakshmi Suresh	4.00	197.00	37.49	73.90	175.00	12932.50	2025-02	2025	221.70
09-02-2025	Kiran Chandran	2.00	1123.00	27.92	313.50	180.00	56430.00	2025-02	2025	940.50
10-02-2025	Arun Dev	3.00	1299.00	34.46	447.60	205.00	91758.00	2025-02	2025	1342.80
13-02-2025	Sajan Nair	3.00	1204.00	26.78	322.40	180.00	58032.00	2025-02	2025	967.20
15-02-2025	Deepa Menon	3.00	952.00	33.42	318.20	175.00	55685.00	2025-02	2025	954.60
15-02-2025	Arun Dev	3.00	163.00	34.97	57.00	180.00	10260.00	2025-02	2025	171.00
15-02-2025	Harish Kumar	3.00	703.00	35.23	247.70	180.00	44586.00	2025-02	2025	743.10
16-02-2025	Mary Augustine	2.00	851.00	39.36	335.00	180.00	60300.00	2025-02	2025	1005.00
16-02-2025	Jins Paul	3.00	1183.00	38.54	455.90	210.00	95739.00	2025-02	2025	1367.70
17-02-2025	Sajan Nair	2.00	620.00	31.35	194.40	195.00	37908.00	2025-02	2025	583.20
17-02-2025	Lalu George	3.00	1061.00	24.58	260.80	200.00	52160.00	2025-02	2025	782.40
21-02-2025	Jolly Mathew	6.00	1004.00	29.46	295.80	190.00	56202.00	2025-02	2025	887.40
22-02-2025	Beena Raju	1.00	568.00	36.76	208.80	210.00	43848.00	2025-02	2025	626.40
23-02-2025	Sreelakshmi Suresh	4.00	1202.00	23.94	287.80	210.00	60438.00	2025-02	2025	863.40
23-02-2025	Jins Paul	2.00	285.00	23.59	67.20	205.00	13776.00	2025-02	2025	201.60
24-02-2025	Vimal Raj	3.00	289.00	24.98	72.20	185.00	13357.00	2025-02	2025	216.60
25-02-2025	Deepa Menon	6.00	984.00	25.01	246.10	210.00	51681.00	2025-02	2025	738.30
25-02-2025	Ramesh Pillai	4.00	331.00	35.40	117.20	175.00	20510.00	2025-02	2025	351.60
25-02-2025	Anitha Thomas	1.00	555.00	30.24	167.80	200.00	33560.00	2025-02	2025	503.40
26-02-2025	Mary Augustine	6.00	1157.00	23.38	270.50	190.00	51395.00	2025-02	2025	811.50

Descriptive Statistics

Descriptive statistics were calculated for the four main variables: **DRC KG, Rate per KG, Amount, and Commission.** The results are shown below (see **Stats sheet**).

Key Findings:

- **DRC KG:**
  - Mean = 214.35 kg per transaction, Median = 210.35 kg.
  - Mode = 73.9 kg, suggesting a recurring transaction size at this level.
  - Std Dev = 114.27 kg and Range = 453.1 kg, indicating wide variability in supplier contribution.
- **Rate per KG (₹):**
  - Mean = ₹192.13, Median = ₹190, Mode = ₹180.
  - Std Dev = 11.82, Range = 35.0 → showing that prices remained relatively stable.
- **Amount (₹):**
  - Mean = ₹41,350 per transaction, Median = ₹40,634.50.
  - Std Dev = ₹22,348 and Range = ₹88,820, showing a large spread between small and bulk transactions.
- **Commission (₹):**
  - Mean = ₹643.05, Median = ₹631.05, Mode = ₹221.7.
  - Std Dev = ₹342.83, Range = ₹1,359.3 → demonstrating high fluctuation in the business per-transaction income.

4 Detailed Explanation of Analysis Process/Method

The analysis for this project was carried out in a structured and systematic way to ensure that meaningful insights could be derived from the raw transaction data of Business A. The process broadly consisted of four stages: **data preparation, descriptive analysis, customer segmentation, and forecasting/insight generation.**

4.1 Data Preparation and Cleaning

- The original records of latex transactions were digitized into Excel (see **Raw sheet**).
- Errors such as missing values, duplicates, or misrecorded weights were corrected in the **Clean sheet**.
- Commission values were validated against the formula: **Commission = Quantity (kg) × ₹3**, ensuring consistency across all records.

4.2 Descriptive Analysis

- Using Excel formulas and pivot tables, descriptive statistics were generated (see **Stats sheet**).
- Metrics such as **Mean, Median, Mode, Standard Deviation, and Range** were calculated for key variables: *Quantity (kg)*, *Rate per kg*, *Amount (₹)*, and *Commission (₹)*.
- This provided an initial understanding of revenue variability, customer patterns, and seasonality.

	A	B	C	D	E	F
1	Column	Mean	Median	Mode	Std Dev	Range
2	DRC KG	214.3486667	210.35	73.9	114.27687	453.10
3	Rate Per KG	192.1333333	190	180	11.820053	35.00
4	Amount	41350.04667	40634.5	No Mode	22348.131	88820.00
5	Commission	643.046	631.05	221.7	342.83062	1359.3

4.3 Customer Segmentation

- A pivot table was created in the **Pivots sheet**, summarizing latex supplied and commission earned by each customer.
- Based on their contribution, customers were classified into three categories:
    - **High-volume suppliers (≥ 2,500 kg)**: These are the most critical contributors, accounting for the bulk of latex supply and commission income.
    - **Medium-volume suppliers (2,000–2,499 kg)**: These suppliers form the majority group, contributing consistently and significantly to overall revenue..
    - **Low-volume suppliers (< 2,000 kg)**: Smaller contributors who add marginally to total volume but still represent potential for growth if retained and nurtured.
- This segmentation is critical for retention strategies, as high-volume suppliers account for the majority of revenue.

4.4 Forecasting and Trend Analysis

- Monthly totals were aggregated to create trend graphs (see **Charts sheet**).
- A **bar chart of Monthly Latex Supply** highlighted seasonal fluctuations.
- A **clustered column chart of Supply vs Commission** revealed the correlation between supply and earnings.
- A **line chart of Commission Trend** was used to visualize volatility and identify peak and lean months.
- Simple **moving averages** were also tested to forecast income for upcoming months, which will guide workforce allocation and financial planning.

4.5 Tools and Rationale

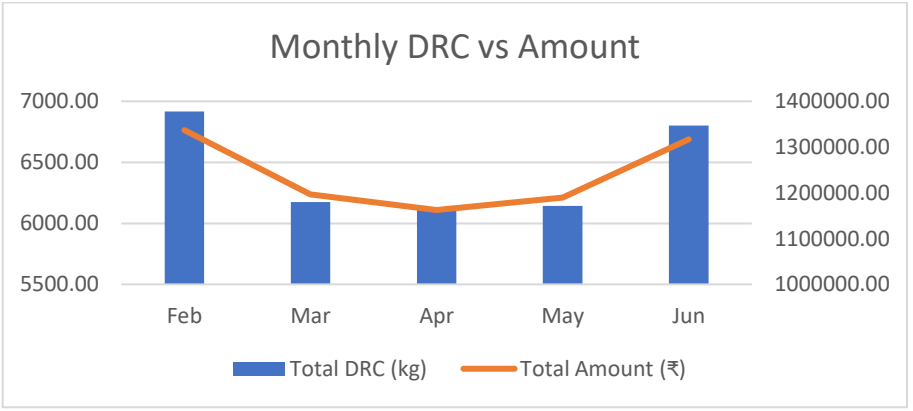
- **Excel** was chosen as the primary tool due to its accessibility, ability to handle small-to-medium datasets, and built-in visualization and pivot functionalities.
- The structured approach (Raw → Clean → Stats → Pivots → Charts) ensured that the analysis remained transparent and reproducible.
- This method provides insights into both historical performance (descriptive) and forward-looking patterns (forecasting), making it well-suited for small rural businesses with limited digital infrastructure.

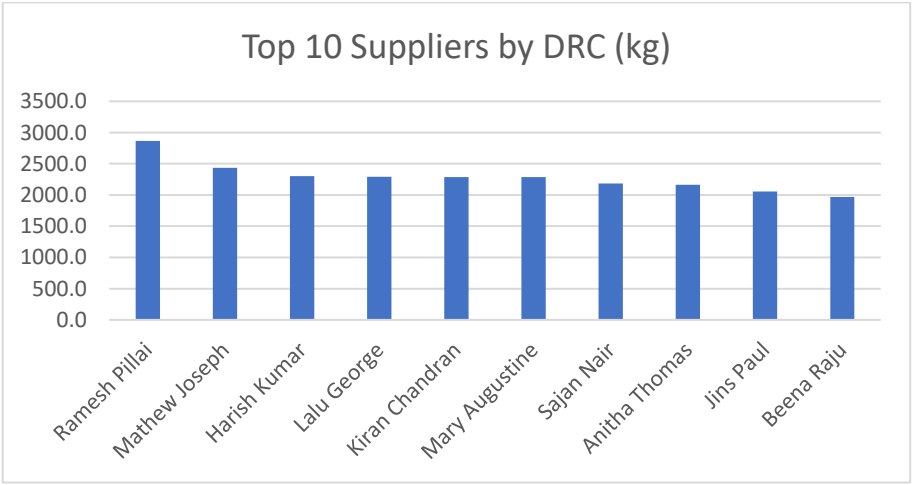
5 Results and Findings

The analysis of the latex transaction data provides several insights into revenue patterns, supplier contributions, and overall business performance. These findings are based on descriptive statistics, pivot tables, and visualizations created in Excel.

5.1 Revenue Trends

- The monthly trend reveals **February as the most profitable month**, with a total commission exceeding ₹20,000, while **April showed the lowest earnings** at approximately ₹18,345.
- These fluctuations are linked to both **seasonal latex production cycles** and **global pricing decisions by dealer**.
- Despite fixed operating costs, revenue is highly volatile, emphasizing the need for forecasting models.

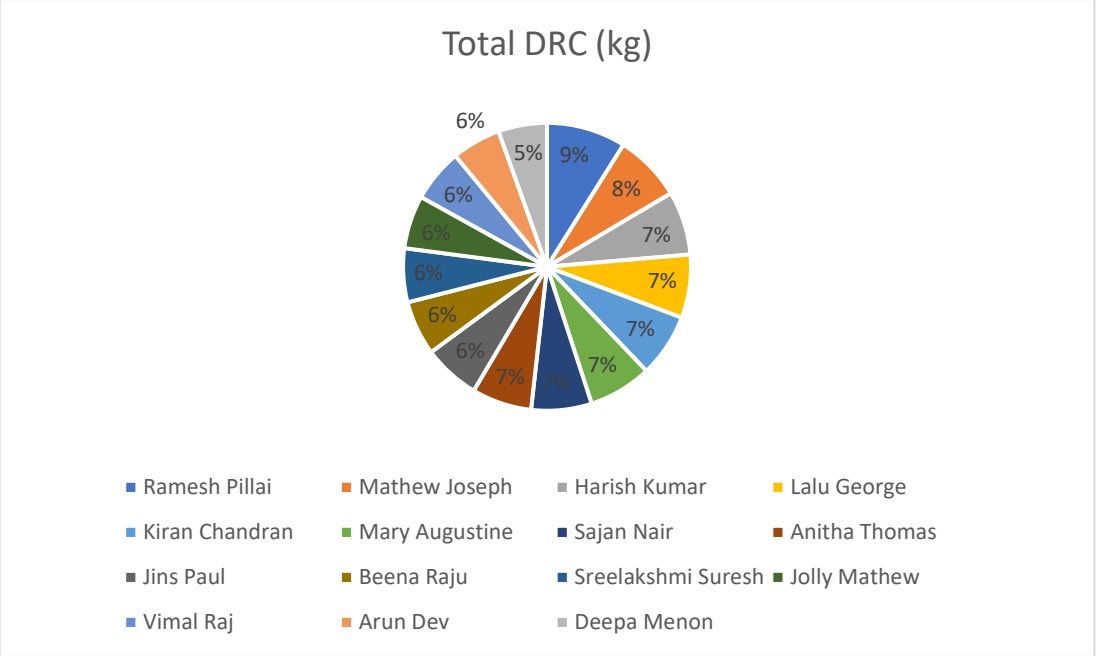




5.2 Customer Contribution and Segmentation

- Customer analysis indicates that Business A relies heavily on a handful of key suppliers.

- A **pie chart from the Pivots sheet** shows that the **top five customers contribute the majority of total latex supplied**, highlighting the concentrated nature of the supplier base.



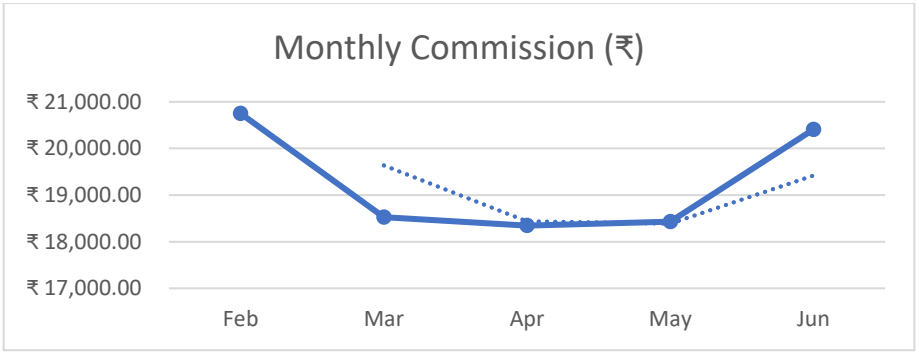
- Based on total DRC supplied, customers were segmented into three categories:
  - **High-volume suppliers (≥ 2,500 kg):** These suppliers are critical to sustaining the business.
  - **Medium-volume suppliers (2,000–2,499 kg):** They form the backbone of the supplier base.
  - **Low-volume suppliers (< 2,000 kg):** Smaller contributors who add marginal value but represent an opportunity to expand sourcing if better engaged.

	A	B	C	D
1	**Customer**	**Total DRC (kg)**	**Total Amount (₹)**	**Total Commission (₹)**
2	Ramesh Pillai	2863.6	₹ 5,47,021.00	₹ 8,590.80
3	Mathew Joseph	2433.1	₹ 4,56,658.00	₹ 7,299.30
4	Harish Kumar	2303.1	₹ 4,47,459.00	₹ 6,909.30
5	Lalu George	2289.3	₹ 4,40,726.50	₹ 6,867.90
6	Kiran Chandran	2287.6	₹ 4,39,212.00	₹ 6,862.80
7	Mary Augustine	2285.9	₹ 4,44,329.50	₹ 6,857.70
8	Sajan Nair	2185.2	₹ 4,16,727.00	₹ 6,555.60
9	Anitha Thomas	2163.4	₹ 4,14,356.00	₹ 6,490.20
10	Jins Paul	2055.5	₹ 4,04,558.00	₹ 6,166.50
11	Beena Raju	1970.7	₹ 3,86,894.50	₹ 5,912.10
12	Sreelakshmi Suresh	1944.5	₹ 3,75,658.00	₹ 5,833.50
13	Jolly Mathew	1921.0	₹ 3,54,839.50	₹ 5,763.00
14	Vimal Raj	1919.4	₹ 3,65,701.00	₹ 5,758.20
15	Arun Dev	1770.4	₹ 3,58,633.00	₹ 5,311.20
16	Deepa Menon	1759.6	₹ 3,49,734.00	₹ 5,278.80
17	**Grand Total**	**32152.3**	**₹ 62,02,507.00**	**₹ 96,456.90**

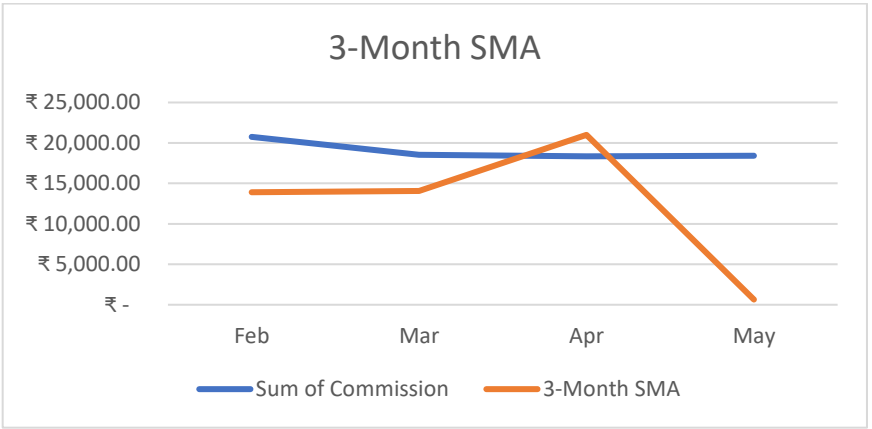
This segmentation shows that Business A's operations are **highly dependent on medium-to-high contributors**. Losing even one of them could significantly impact revenue, reinforcing the importance of **retention strategies**.

5.3 Seasonality and Income Volatility

- The **commission trend line chart** shows noticeable peaks and troughs, with sharp increases in May - June and dips in Feb - March.
- Such volatility makes **income planning difficult** and puts pressure on cash flows during lean months.
- A moving average applied to the dataset suggests that the cycle is **seasonal**, and income dips are likely to recur annually.



- To further smooth out short-term fluctuations, a **Simple Moving Average (SMA)** was applied. This revealed the underlying direction of earnings, independent of sudden monthly spikes or dips.
- The SMA curve indicates a **gradual upward trend overall**, even though monthly values fluctuate. This can help **anticipate future patterns** and plan accordingly.



5.4 Operational Implications

- **Fixed Costs:** Rent, transport, and casual labor expenses remain constant, but revenue fluctuations reduce profit margins during lean months.

- **Workforce Planning:** High-volume months (e.g., May–June) may justify hiring permanent or semi-permanent staff, while lean months can be managed with temporary labor.
- **Digitalization:** Manual data entry was slow and error-prone. Transitioning to digital billing systems would save time, reduce errors, and provide real-time insights.

5.5 Strategic Implications

The findings lead to several actionable insights:

1. **Customer Retention is Crucial** – High and medium suppliers should be prioritized with personalized engagement, timely payments, and trust-building measures.
2. **Revenue Forecasting Needed** – Time series forecasting (even simple moving averages) can help anticipate income dips and allow the owners to plan expenses better.
3. **Expand Low-Volume Suppliers** – Although currently contributing little, these suppliers represent growth opportunities if properly incentivized.
4. **Seasonal Strategy** – Prepare for Feb–March downturn by planning cash reserves or diversifying into related services during that period.
5. **Digital Tools** – Adoption of simple Excel-based dashboards or low-cost mobile applications will improve transparency and support data-driven decisions.

Summary of Results

- February is the best-performing month; April is the weakest.
- A small number of suppliers dominate latex contributions.
- Segmentation shows clear High, Medium, and Low groups, with no extreme small-scale suppliers.
- Seasonal fluctuations create income volatility.
- Operational improvements in workforce and digitalization are essential for stability.