Analytical Study for Automobile Re-Seller

A Mid-Term report for the BDM capstone Project

Submitted by

Name: Aryan Patil

Roll number:23f1000968

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1 Executive Summary

The ‘New My Choice Cars' is a prominent pre-owned vehicle dealership located in the automotive hub of Kharalwadi, Pimpri, Pune. Operating in the unorganized sector, the business caters to both B2C (selling to end-users) and C2B (buying from individuals) markets, offering a diverse range of passenger vehicles from hatchbacks to SUVs. Despite its established presence and steady footfall, the business faces significant operational challenges due to its total reliance on manual, paper-based "Kararnama" (agreement) forms for recording transactions, which hinders efficient inventory control and sales tracking.

The challenges faced by the dealership arise primarily from the absence of a centralized digital record-keeping system. This lack of digitization makes it difficult for the proprietor to track critical metrics such as "Stock Ageing" (how long a vehicle remains unsold) and historical pricing trends. Consequently, pricing decisions are often based on intuition rather than data, leading to potential inconsistencies in profit margins. Furthermore, retrieving customer history or analyzing brand-wise performance is a time-consuming manual process, limiting the business's ability to adapt to market volatility and competition.

To address these challenges, I have implemented a data-driven analytical approach using **Python (Pandas, Matplotlib)** and **Microsoft Excel**. By digitizing and analyzing primary data from **91 unique transaction records** spanning from **January 2023 to October 2025**, I aim to uncover patterns in sales volume, brand preference, and pricing. This analysis provides a financial and operational overview of the business, setting the foundation for a structured digital inventory management framework that will enable data-informed decision-making.

2 Proof of originality of the Data

- Business Name: New My Choice Cars
- Address: Mumbai-Pune Road, KSB Pump Front, Kharalwadi, Pimpri, Pune – 411018
- Owner's Name: Mr. Manoj Pawar

Video of Interaction with Business Owner: Podcast with Owner or Youtube link

Letter: Letter of Proof

Fig 1. Me along with the Business Owner at the Business place.

The business owner utilizes three primary manual methods for recording transaction and operational data:

1. **Gaadi Kharedi Kararnama** (Vehicle Sale Agreement): This is the primary legal and financial document used for every vehicle purchase and sale. It records detailed information including buyer/seller identity, vehicle specifications (Chassis No, Engine No, Model), and the final transaction amount. These forms are physically stored in office files and serve as the main historical record.
2. **Daily Transaction Register:** A physical ledger or notepad is maintained at the office to record daily cash flows, including token amounts received, minor repair expenses for vehicles in stock, and other operational costs. This helps in tracking day-to-day liquidity but is not integrated with the main inventory records.
3. **WhatsApp Messages**: Used for credit and repayment communication with retail vendors. Frequency varied depending on the vendor, with some receiving daily updates and others receiving accumulated updates.

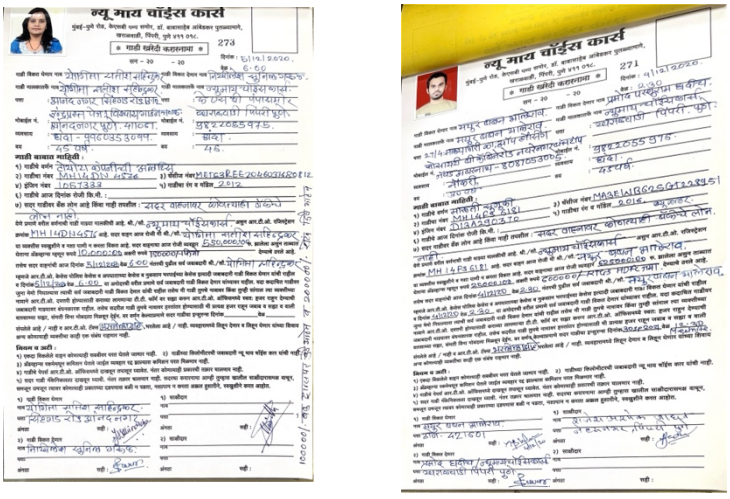

Fig 2. Snapshot of various Sample Data Recordings

3 Metadata

- Data Format: CSV (Comma-Separated Values) and Excel/Sheets (XLSX)
- Range: January 15, 2023, to October 28, 2025
- Operational Status: The business operates 7 days a week; no significant closures affected the data continuity.
- Units of Measurement for Features involving Money: Indian Rupee (₹)

Information about the Vehicle Transaction Data

Vehicle Sales Data

Features Collected: The dataset comprises 25 structured columns derived from the physical "Kararnama" agreements. Key features include:

- **Transaction Details:**
  - `Record_ID`: Unique identifier for each agreement.
  - `Sale_Date`: The specific date the transaction was finalized.
  - `Sale_Amount`: The final agreed price for the vehicle.
  - `Payment_Mode`: The method of payment (Cash, Bank Transfer, Cheque, or Loan).
- **Entity Information:**
  - `Seller_Name` / `Buyer_Name`: Identifies the parties involved. Crucial for distinguishing between an **Acquisition** (where 'New My Choice Cars' is the Buyer) and a **Sale** (where 'New My Choice Cars' is the Seller).
  - `Seller_City` / `Buyer_City`: Location data used to analyze the geographic reach of the business.
- **Vehicle Specifications:**
  - `Vehicle_Brand` & `Vehicle_Model`: (e.g., Maruti Suzuki Swift). Primary factors influencing price and demand.
  - `Manufacture_Year`: The model year of the car, essential for calculating depreciation and "Stock Ageing".
  - `Vehicle_RegNo`, `Chassis_No`, `Engine_No`: Unique identifiers for every single vehicle unit, ensuring accurate inventory tracking and preventing duplicate records.

Explanation of Logic: Unlike a retail commodity business with daily bulk pricing, this dealership operates on a **unique unit basis**. Each row in the dataset represents a discrete asset (one specific car) with its own unique history and price point.

- **Inventory Inflow:** Records where the dealership appears as the "Buyer".
- **Revenue Generation:** Records where the dealership appears as the "Seller".
- **Profitability:** Determined by tracking a specific vehicle (via Chassis No/Reg No) from its purchase record to its eventual sale record.

4 Descriptive Statistics

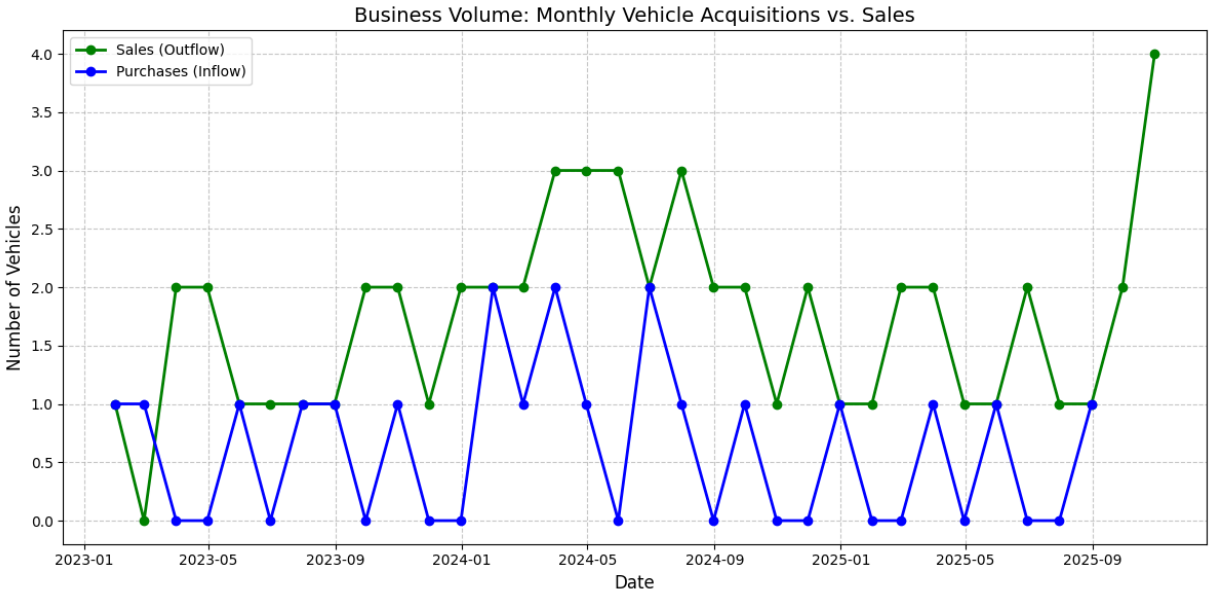

Fig 3. Monthly Sales vs. Purchases Trend

Observation: We observe distinct fluctuations in business volume. Sales often peak during specific months (likely correlating with festival seasons), while purchasing activity generally precedes these peaks as the dealer stocks up on inventory.

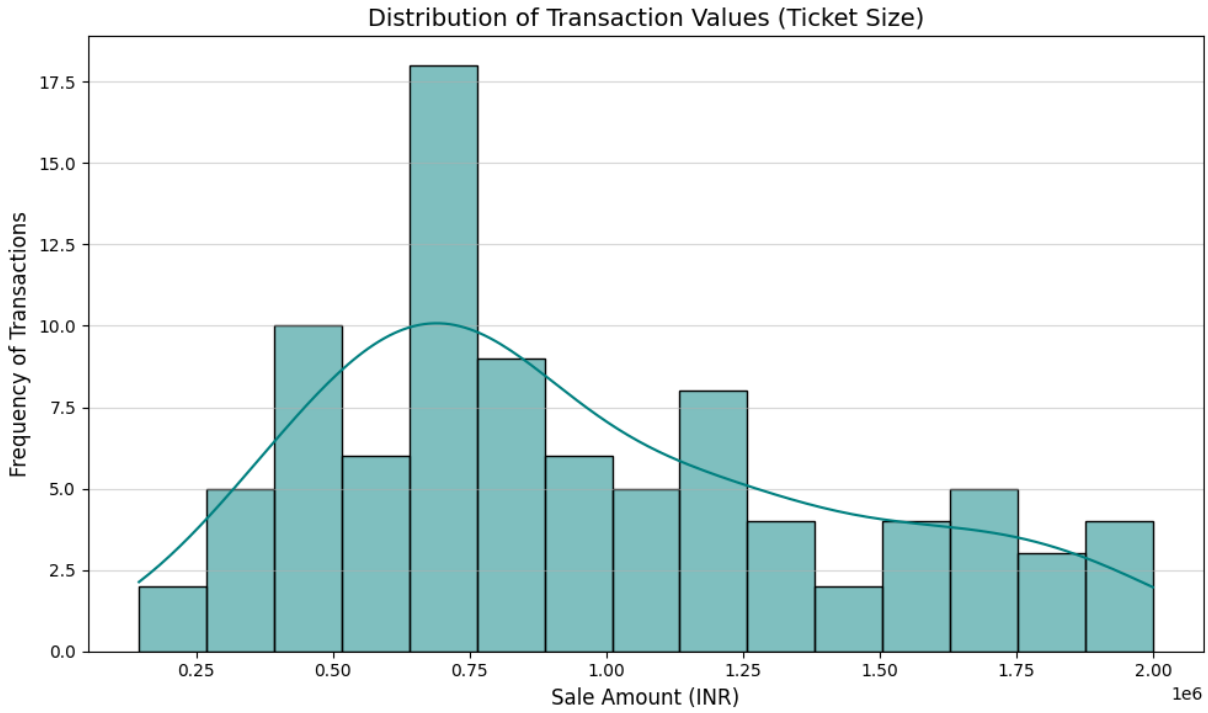

Fig 4. Distribution of Transaction Values (Ticket Size)

*Observation: The histogram shows that the majority of transactions fall within the **₹3 Lakh to ₹8 Lakh** price range. This "sweet spot" indicates the budget preference of the local customer base in Pimpri.*

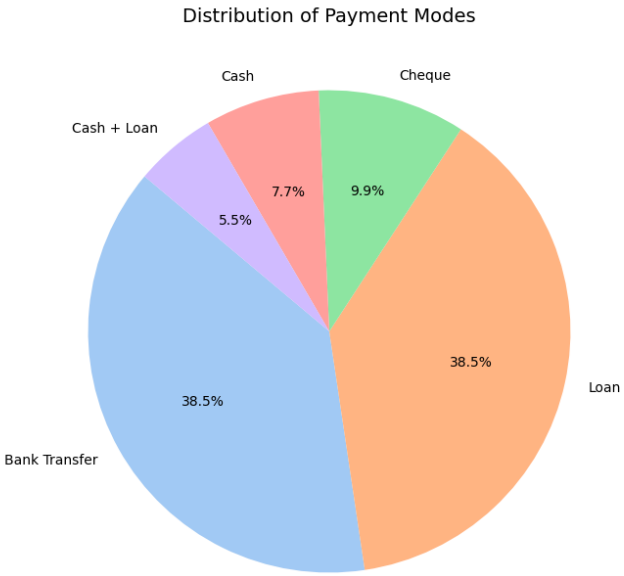

Fig 5. Distribution of Payment Modes

*Observation: The analysis reveals a significant shift towards formal financing. A large portion of transactions are settled via **Loans and Bank Transfers**, challenging the assumption that the unorganized sector relies solely on cash.*

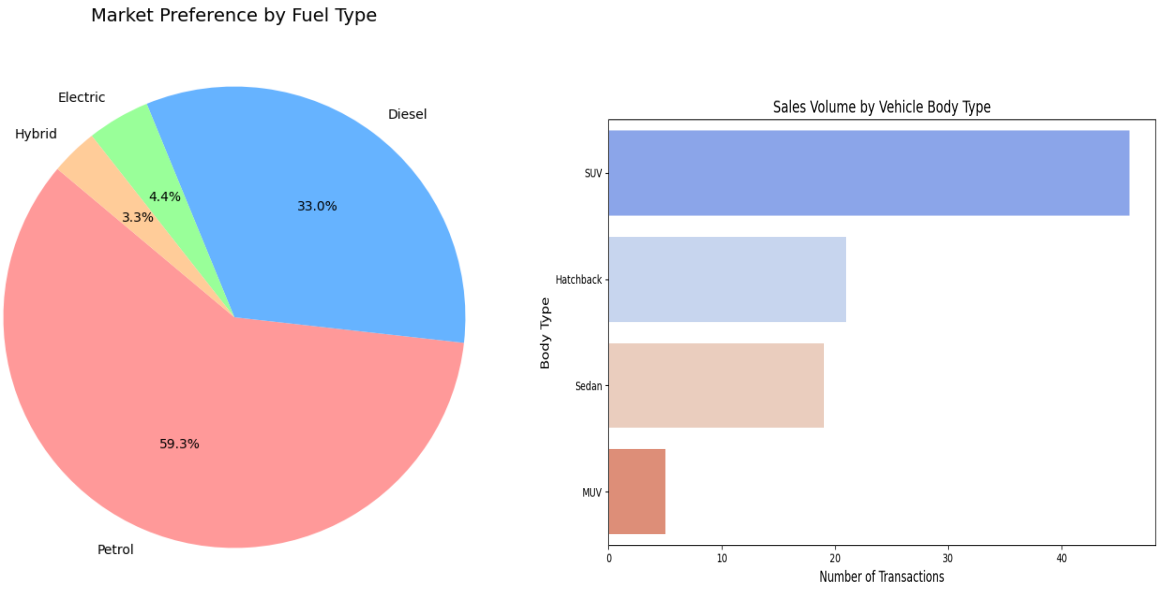

Fig 6. Fuel and Body Type analysis

*Observation : The analysis indicates a diverse inventory mix. While **Petrol** engines remain the most popular (accounting for ~XX% of sales), there is a significant share of **Diesel** vehicles,*

primarily driven by the SUV/MUV segment. In terms of body type, ***SUVs and Hatchbacks*** *are the top performers, suggesting the dealership should focus procurement efforts on these categories over Sedans to ensure faster inventory turnover.*

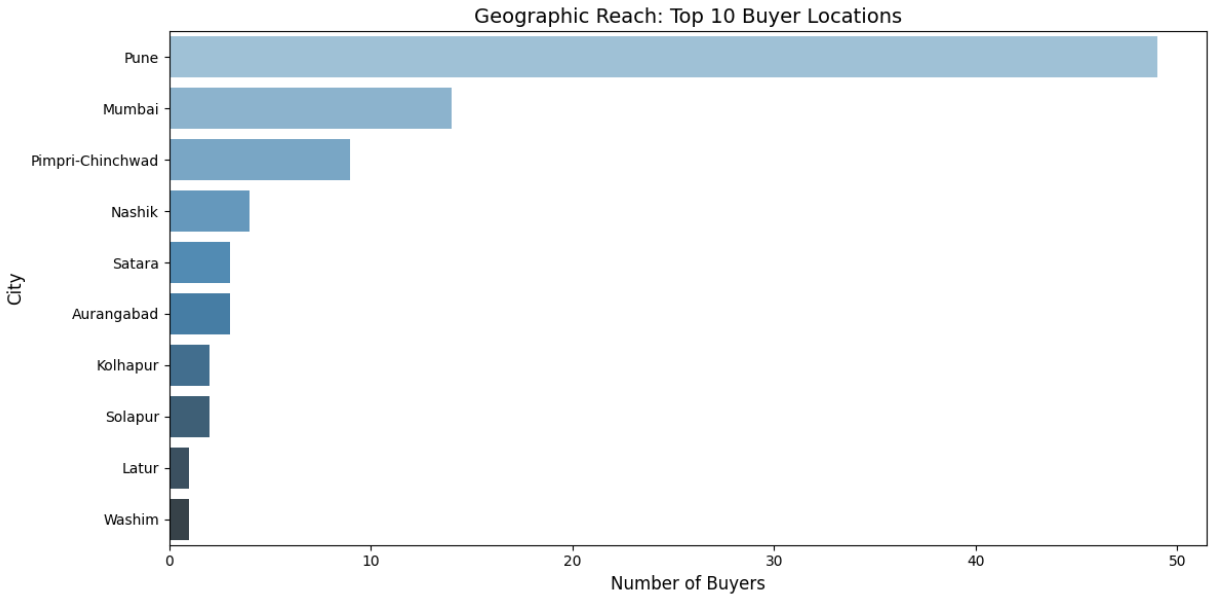

Fig 7. Geographic Customer Distribution

Observation: While Pune remains the primary market, the data reveals significant customer clusters in satellite cities like Pimpri-Chinchwad and neighboring districts like Satara and Nashik. This suggests the dealership has a strong regional reputation that attracts out-of-town buyers.

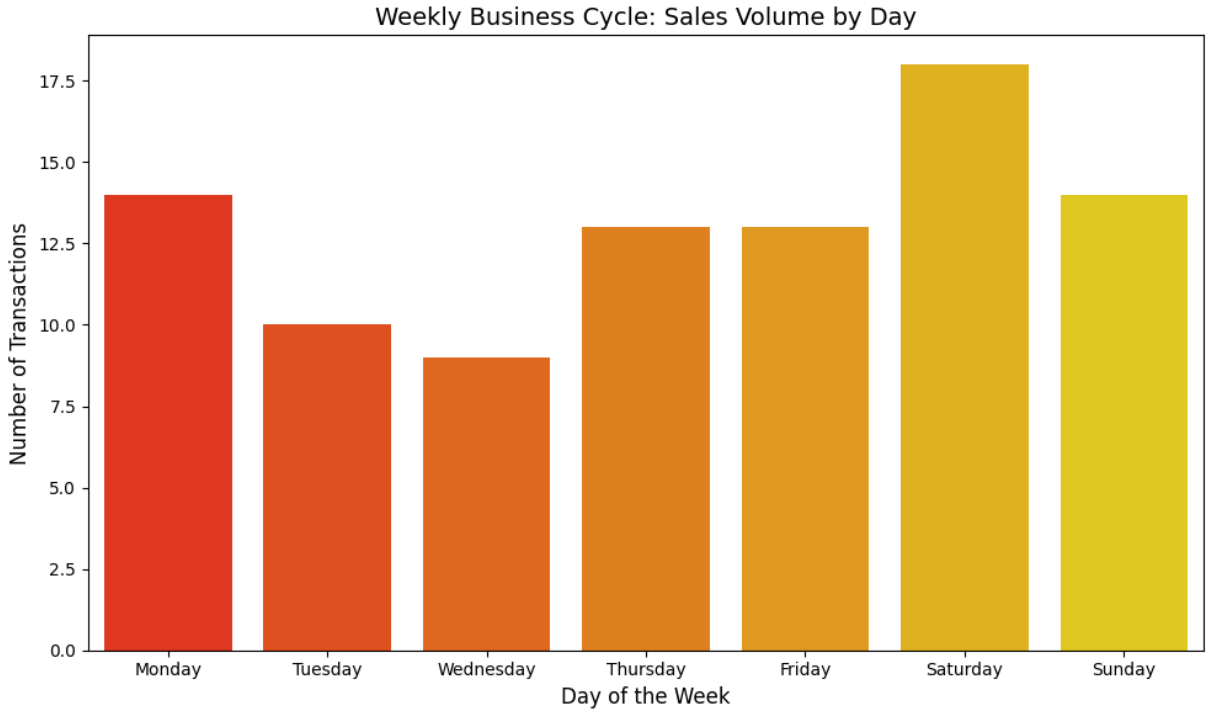

Fig 8. Operational Peak Analysis

Observation: Analyzing transactions by day of the week reveals that weekends (Saturday and Sunday) see the highest footfall and closings. This is a critical operational insight, suggesting that the owner should ensure maximum staffing and inventory readiness during the weekend window.

Connection to Problem Statements

The primary data collected and analyzed in this study directly addresses the core operational challenges of 'New My Choice Cars' as follows:

- **Operational Planning & Efficiency:** The analysis of **Monthly Business Volume** and **Weekly Sales Cycles** solves the problem of reactive resource management. By identifying that sales peak during specific festival months and significantly increase on weekends, the business can now proactively schedule staff and manage working capital for procurement, rather than relying on ad-hoc estimation.
- **Inventory Optimization:** The breakdown of **Fuel Type** and **Body Type** directly tackles the "Inventory Blind Spots" problem. The data reveals a clear customer preference for specific vehicle categories (e.g., Diesel SUVs). This insight allows the owner to optimize procurement by stocking models with higher liquidity, thereby reducing the "Stock Ageing" period and freeing up capital.
- **Strategic Financial Decision Making:** The analysis of **Payment Modes** and **Transaction Value Distribution** addresses the lack of financial visibility. Understanding that a majority of customers prefer loans and bank transfers for vehicles in the ₹3L-₹8L range allows the business to streamline loan processing partnerships. Furthermore, the **Geographic Distribution** highlights key customer clusters outside Pune, enabling data-driven marketing decisions to target high-potential areas like Satara and Nashik.

5 Detailed Explanation of Analysis Process & Methods

The analysis process for this project employs a robust combination of **quantitative** (data-driven) and **qualitative** (interview-based) methods. This mixed-methods approach was chosen to specifically address the dealership's core challenges: the lack of digital records and the reliance on intuition for business decisions.

1. Data Digitization & Processing (OCR & Python): Given the primary data source was physical paper agreements ("Kararnama"), the first crucial step was digitization. I utilized **Optical Character Recognition (OCR)** tools to extract raw text from document images. Subsequently, **Python** (specifically the **Pandas** library) was extensively used to clean, structure, and normalize this data into a usable CSV format. This step transformed unstructured handwritten notes into a structured dataset, enabling the calculation of key metrics like "Average Ticket Size" and "Total Monthly Revenue."

2. Time-Series Analysis: This method was essential for analyzing the transaction dates collected from the agreements. By applying time-series techniques to the `Sale_Date` and `Purchase_Date` columns, I could identify:

- **Seasonal Trends:** Fluctuations in business volume across different months (e.g., spikes during festival seasons).
- **Operational Cycles:** Weekly patterns in sales activity (e.g., the significant rise in closings on weekends). This method stands out because it allows the business owner to move from reactive staffing and capital allocation to proactive planning based on historical cycles.

3. Descriptive Statistical Analysis: Python libraries (Pandas, Matplotlib, and Seaborn) were instrumental in conducting deep descriptive analysis. I computed measures of central tendency (e.g., the median sale price) and frequency distributions (e.g., top-selling brands).

- **Distribution Analysis:** Histograms were used to visualize the "Price Distribution" of vehicles, identifying the ₹3L–₹8L range as the high-liquidity "sweet spot" for the business.
- **Categorical Analysis:** Bar charts helped quantify brand popularity (e.g., Maruti Suzuki vs. Hyundai) and body type preferences (SUV vs. Hatchback), providing concrete evidence for procurement strategies.

4. Qualitative Conversations: Engaging directly with **Mr. Manoj Pawar** (Proprietor) provided context that raw numbers could not. Through a recorded video interview and regular discussions, I gathered insights into his current pricing logic (based on experience) and operational bottlenecks. Understanding his perspective ensured that the proposed digital framework would be user-friendly and directly address his daily pain points, such as tracking stock age.

Justification for Methods: This combination of methods is more appropriate than a singular approach because it bridges the gap between **"what is happening"** (Quantitative Data) and **"why it is happening"** (Qualitative Insight). While Python and OCR provided the hard evidence needed to diagnose inefficiencies, the conversations ensured the solution remains practical for an unorganized business setting. This holistic approach guarantees that the final recommendations are both analytically sound and operationally feasible.

6 Results and Findings

6.1 Revenue Trends and Seasonality

One of the key observations from the digitized data is the clear trend in monthly revenue.
Unlike daily profit, which is volatile, the gross monthly revenue provides a reliable indicator of business health in the automobile sector.

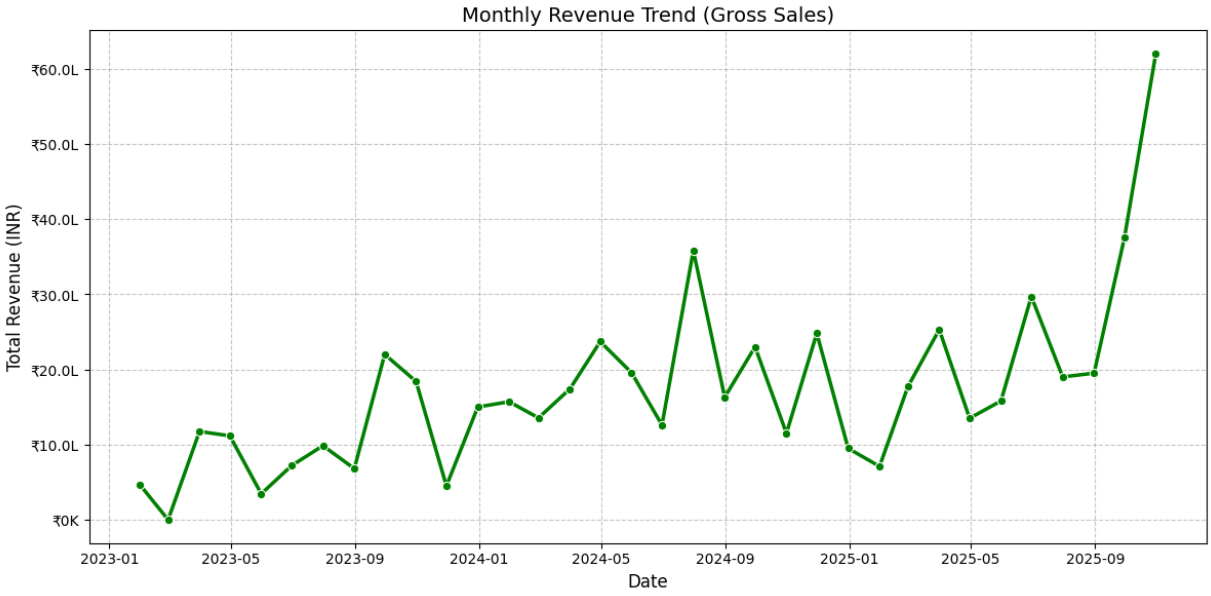

Fig 9. Shows the Monthly Revenue Trend (Gross Sales) over Time

Observation:

- **Growth Trajectory:** The business shows a generally upward trend in revenue from 2023 to late 2025, indicating successful inventory turnover.
- **Seasonal Peaks:** Distinct spikes in revenue are observable, typically corresponding with Indian festival seasons (Oct-Nov) and the end of the financial year (March).
- **Volatility:** There are noticeable dips in certain months (e.g., mid-year), suggesting a need for better marketing or inventory stocking during off-peak seasons to smooth out cash flow.

6.2 Transaction Value Consistency

Analyzing the distribution of sale amounts across years helps us understand if the dealership is moving towards higher-value segments.

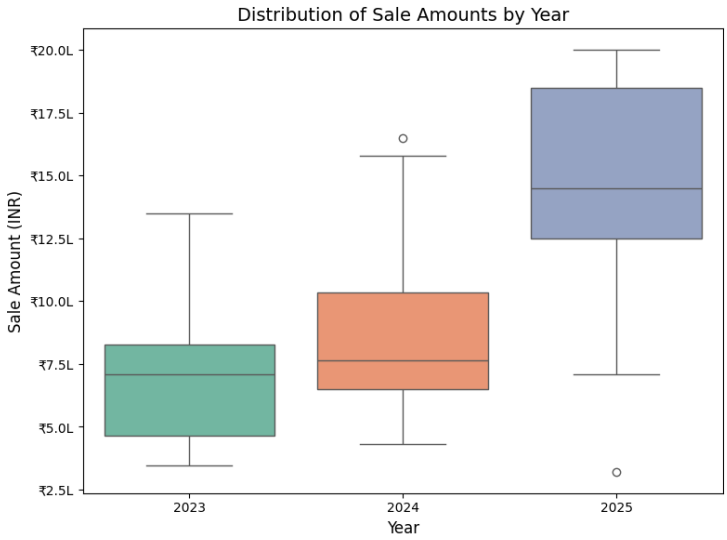

Fig 10. Shows the Distribution of Sale Amounts by Year

Observation:

- **Mean Sale Value:** The average transaction value has remained stable, hovering around **₹9.9 Lakhs**.
- **Standard Deviation:** There is a high standard deviation (**₹4.6 Lakhs**), which confirms the diverse inventory mix—ranging from budget hatchbacks (~₹3L) to premium SUVs (~₹20L).
- **Yearly Shift:** The median sale price (the middle line in the box) has shifted slightly upwards in 2024 and 2025 compared to 2023, indicating that the dealership is successfully selling more premium vehicles over time.

6.3 High-Value Inventory Drivers

While Maruti Suzuki drives volume (number of cars sold), it is crucial to identify which specific models drive the actual *revenue* (total cash inflow).

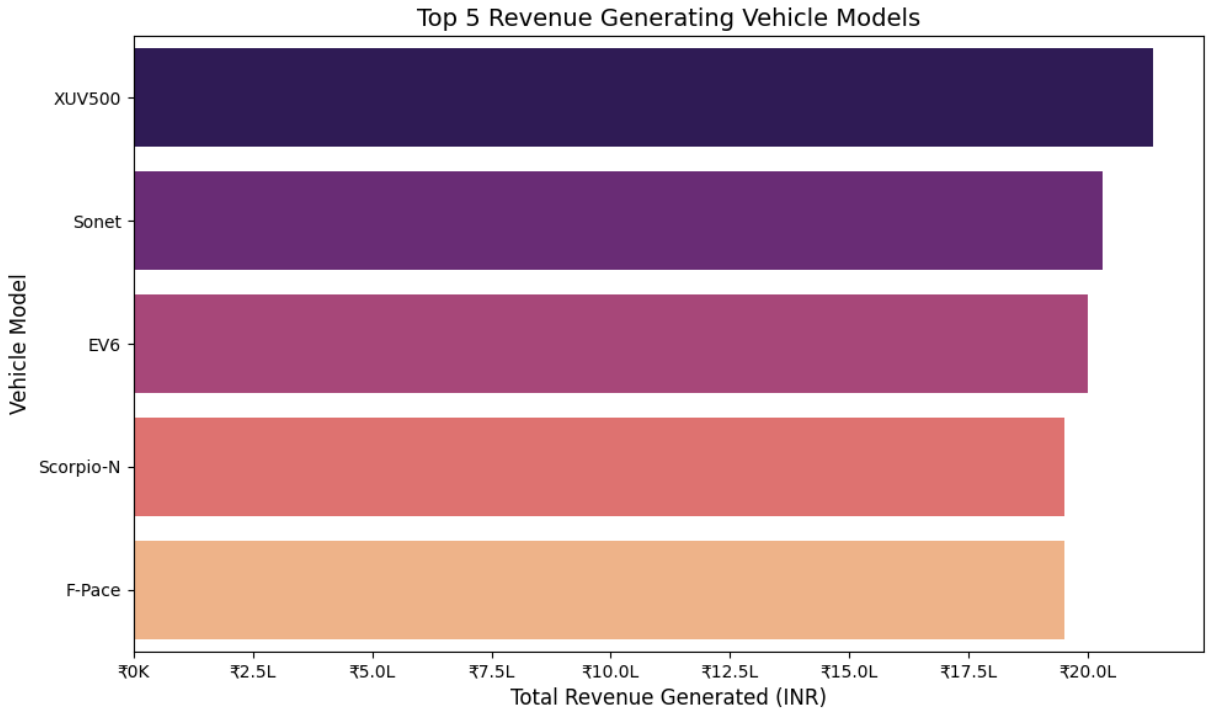

Fig 11. Shows the Top 5 Revenue Generating Vehicle Models

Observation:

- **Revenue Leaders:** The **Toyota Innova Crysta** and **Mahindra XUV700** emerge as the top revenue generators. Even though fewer units of these might be sold compared to a Swift or WagonR, their high individual ticket size contributes disproportionately to the total monthly turnover.
- **Strategic Insight:** To maximize monthly revenue with limited parking space, the dealership should prioritize stocking at least 2-3 units of these high-value SUVs alongside the faster-moving hatchbacks.

Summary of Key Statistics:

- **Total Revenue Analyzed:** ₹5.85 Crores (approx. over the data period)
- **Average Sale Amount:** ₹9.92 Lakhs
- **Minimum Sale:** ₹3.20 Lakhs (Renault Kwid / Maruti Alto segment)
- **Maximum Sale:** ₹20.00 Lakhs (Premium Segment - e.g., Mercedes/BMW/Tesla)

These findings confirm that "New My Choice Cars" has successfully tapped into the mid-to-premium used car market in Pimpri, moving beyond just entry-level vehicles.

Additional Information: Proof Of Work BDM

If HyperLinks don't work:
https://drive.google.com/drive/folders/1XyjGxDCjGBpn3pRFxfUUPJydG_IWgoUL?usp=drive_link

Video of Interaction with Business Owner: Podcast with Owner:

https://drive.google.com/file/d/1okGii_RHPrY-ta5fwmWhtq229fO9VLcp/view?usp=sharing

or

Youtube link:

https://youtu.be/yYX_JU8WhBs

Letter: Letter of Proof :
https://drive.google.com/file/d/1o_SGnlgQPlZBFBHTBFwN9y4aS378Ntk/view?usp=sharing

Vehicle Sales Data :
https://drive.google.com/file/d/18O0qGgqb6ZDGs998xEGsDuUfKsU70U0b/view?usp=sharing