Title: Enhancing Profitability for a Local Automotive Spare Parts Vendor and Garage Service

A Mid Term Submission report for the BDM capstone Project

Submitted by

Podey Om Prabhakar

21f2000968

IITM Online BS Degree Program,

Indian Institute of Technology,

Madras, Chennai Tamil Nadu, India, 600036

Contents

1.Executive Summary: 3

2. Proof of Originality 3

3. Meta Data 4

4. Descriptive Statistics 6

5 Methodology and Rationale for Method Selection 8

6. Results and Findings: 10

7. Expected Outcome: ..... **Error! Bookmark not defined.**

Declaration Statement

I am working on a Project titled **Enhancing Profitability for a Local Automotive Spare Parts Vendor and Garage Service**. I extend my appreciation to **Pawansut Automobiles** 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 from 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 principles 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 understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Signature of Candidate: (Digital Signature)

Name: Podey Om Prabhakar

Date: 10-10-2024

1.Executive Summary:

This capstone project aims to boost the growth of Pawansut Automobiles, a local provider of vehicle parts and garage services, while improving its operational efficiency. As an entity in the vehicle repair sector, the vendor currently encounters several significant challenges. The most urgent concerns involve low profit margins caused by excess inventory, inefficient sales tracking systems, and ineffective customer retention practices. These issues have led to financial instability, posing barriers to growth and profitability.

To tackle these challenges, a data-driven strategy was utilized to pinpoint operational inefficiencies and develop actionable insights. The analysis centered on sales and inventory data, looking into stock levels, sales trends of parts, and service costs. By employing advanced analytical techniques, such as inventory optimization, sales forecasting, and customer behavior analysis, this project aims to reduce surplus inventory, enhance the accuracy of sales tracking, and foster better customer engagement.

The analysis unveiled key operational inefficiencies, including inconsistencies between sales and inventory records, which may indicate potential problems such as manual entry mistakes or unmonitored product usage. Certain spare parts showed high demand but were understocked, highlighting difficulties in inventory management. Furthermore, the evaluation of service costs pointed out the necessity for a more adaptable pricing structure to align with actual service needs.

Based on these findings, several interventions have been recommended to improve operations. These comprise implementing better inventory management practices, refining sales reconciliation methods, and developing stronger customer engagement strategies. The primary goal is to diminish inventory-related expenses while boosting overall profitability, thereby ensuring a sustainable business model that promotes long-term success in the competitive automotive repair field.

2. Proof of Originality

Please Refer to the Link:- *Proof of Originality Link* (*https://drive.google.com/drive/folders/1vCvYHB7KvUEVmSVdwSNQTFoKZS-uT_iH?usp=sharing*)

Links Contain,

1. Primary Raw Data
2. Letter from Organization.
3. Shop Photos.
4. Interaction Video with owner.

3. Meta Data

1. Data Sources:

The information for this project came straight from Mr. Pranav, who owns Pawansut Automobiles This is a local company that focuses on automobile service & spare parts. The data was collected through direct interactions with the seller. This included looking at sales transactions and checking the inventory records. For gathering the information, we used human entry & export methods via Microsoft Excel. This program is now key for managing inventory and sales, which helps ensure that everything is both accurate & relevant to the business.

2. Data Structure:

2.1 Sales Data:

Description: This dataset holds detailed records of service and spare parts sales for two-wheelers. Each entry represents a unique transaction, which includes different parts put into a vehicle along with their prices and the final bill amount.

Table Name:

- Sales Transactions: This table captures the transactional details of each sale, broken down by part applied, quantity, price, and associated costs.

Key Columns:

- Sr. no: Each transaction gets a unique identification number (Sr. no).
- Date: The date when the service or sale was done.
- Vehicle Name: The model and make of the two-wheeler serviced or for which parts were sold.
- Parts Applied: Specific spare components that were either sold or used during this transaction.
- Quantity: How many units of each part were used or sold in that transaction.
- Price: This reflects the unit price of each part.
- Amount: Total cost for the parts used in the transaction (calculated as Quantity × Price).
- Servicing Cost and Labour Charges: Additional service or labour charges applied (typically calculated as 10% of the spare parts cost).
- Final Bill: The final bill includes all related costs and charges..
- Remark: Any additional notes or remarks related to the transaction or vehicle serviced.

Data Types:

- Sr. no: Integer
- Spare Part: Text

- Type: Text
- Total Quantity Purchased: Numeric
- Currently Available: Numeric
- Price: Numeric

2.2 Inventory Data

Description: This dataset provides detailed records of inventory items related to two-wheeler & spare parts sales. Each entry details the parts, showing their purchase status as well as availability.

Table Name:

- Inventory Records: This table captures current stock levels & pricing information for each spare part, including quantities bought & available.

Key Columns: (can choose any one from this or above key columns (simple one)

- Sr. No: A unique identifier assigned to each inventory record. It's essential for indexing & referencing specific items efficiently.
- Spare Part: The name or ID of the spare part listed in the inventory. This field clarifies what parts are in stock and ready for sale.
- Type: This classifies the spare part, showing which type of vehicle it's linked to (for example: Bike, Scooter, Common). It’s helpful for organizing inventory based on vehicle types.
- Total Quantity Purchased: This indicates the total number of units purchased for each spare part. It reflects overall stock levels & aids in tracking purchasing trends.
- Currently Available: The number of units per spare part that are presently in stock. Crucial for managing inventory effectively and ensuring parts are available when needed.
- Price: The unit price for each spare part. Important for calculating total inventory value and future sales pricing.

Data Types:

- Sr. no: Integer
- Date: Date
- Vehicle Name: Text
- Parts Applied: Text
- Quantity: Numeric
- Price: Numeric
- Amount: Numeric
- Servicing Cost and Labour Charges: Numeric
- Final Bill: Numeric
- Remark: Text

4. Descriptive Statistics

4.1 Sales Data

To summarize the central tendencies, variability, overall distribution of these variables, we will calculate descriptive. This analysis will give into key metrics of sales performance.

Column wise insights:

Sr.no: The Sr.no column shows a unique number assigned to each transaction. It ranges from 1 to 78, which means we have 78 unique transactions in the sales data.

Date: The Date spans from 24-Jun-24 to 22-Jul-24. This means the data collected focuses on late June & extends into July for better overall sales capture.

Vehicle name: In the Vehicle Name column, there are 78 entries with 63 unique vehicle names. This highlights a nice variety of vehicles included in the data. The most common vehicle is "Honda CB Shine MY2020," appearing three times. While it's the most frequently listed, it's not excessively dominant. So, there's a well-rounded representation of various vehicle models.

Parts applied: The Parts Applied column shows 599 entries, likely representing the total count of parts recorded. Among these, there are 271 unique parts noted. This indicates a vast array of different parts being utilized. The part that appears most often is "SERVO 4T SAE 4 Stroke Engine Oil for Motorcycles." It shows up 18 times, marking it as commonly used but not overwhelmingly so.

Quantity: For Quantity, the average per transaction sits at 1.04 with a range between 1.0 to 4.0 and a standard deviation of 0.24. This suggests that most transactions include just one unit of spare part with little variation in numbers. The average being slightly over 1 hints at some transactions where more than one unit is bought.

Price: When looking at Price, it averages around 316.70 with a range from 4.3 to 4110.0. The median price is noted at 137.8 and there’s a high standard deviation of 485.99 showing significant price variability among parts. The average price being much higher than the median suggests that some very high-priced parts are skewing this average upwards—there are indeed outliers in play here.

Amount: In the Amount column, we find an average value of 573.44 ranging from 4.3 to 7203.0 with a substantial standard deviation of 978.88 indicating considerable variation in total amounts charged—this varies widely based on how many products are purchased & their market prices.

Servicing cost and labour charges: Servicing Cost & Labour Charges average around 57.34 with values ranging from 0.43 to 720.3; the high standard deviation here, at 97.89, reflects significant variations in servicing costs themselves. Because servicing cost is fixed at a percentage (10%) of

spare parts cost, any changes in spare parts prices will directly impact servicing charges—high costs on spare parts lead to higher servicing fees, while low prices result in lower charges.

Final Bill: For Final Bill amounts, we see an average value of about 2726.99 with a range spanning from 238.0 to around 7923.0; a high standard deviation of 1679.39 indicates notable variability in final bill totals—again pointing towards diverse transaction values overall! There's considerable fluctuation in bill amounts driven by high costs on certain services & spare parts; even if most bills fall below average, those few high-value bills can substantially influence both average & standard deviation figures.

Remark: As for the Remark column? It contains no data; its zero count indicates it wasn't used or filled out in this dataset at all.

4.2 Inventory Data:

We will perform descriptive statistical analysis on the inventory dataset, focusing on metrics such as Total Quantity Purchased, Currently Available, and Price and other variables if necessary. This analysis aims to provide a comprehensive understanding of the inventory levels, purchase quantities, and pricing structures.

Column wise insights:

The inventory data encompasses 430 records, with the Total Quantity Purchased ranging from 1 to 40 units, indicating a broad spectrum of purchase volumes. The average quantity purchased is approximately 13.89 units, with a standard deviation of 5.97, reflecting significant variability in purchase volumes.

Sr.no: The serial number represents the number of transactions or log entries in the inventory. There are 430 entries, which contain information regarding product purchases and stock levels.

Total Quantity Purchased: The total quantity purchased ranges from 1 to 430 units. The mean remained at approximately 13.89 units. This means that most purchases cluster around this average; also, a standard deviation of 5.97 indicates considerable variability overall.

Spare Parts: The "Spare Part" column contains 430 entries with 424 unique values, indicating a high level of diversity in the spare parts recorded. The most frequently listed spare part is "AIR Filter Assy. This high uniqueness suggests that the dataset covers a wide range of spare parts, with a small number of parts being more common than others.

Type: The "Type" column consists of 430 entries with 3 unique values namely Bike, Scooter and Common. The most prevalent type is "Scooters," which appears 248 times. This indicates that the dataset primarily focuses a bit more on scooter parts, with relatively smaller representation of other vehicle types.

Currently available: The Currently Available amounts show a minimum of 1 and a maximum of 40 units. The mean is around 12.37 units, with a standard deviation close to 5. This illustrates some variation in stock levels, but it also shows that a sufficient amount of stock is generally maintained.

Price: Prices vary significantly, ranging from 4.30 to 4110.00. The mean price is 449.07; however, the large standard deviation of 660.28 indicates significant variance within pricing and multiple outliers in this variable category.

5 Methodology and Rationale for Method Selection

5.1 Python for Descriptive Statistics

In the initial phase of the analysis, we used Python to calculate descriptive statistics for dataset. Python is a one of the powerful programming languages. It has many great libraries like pandas & numpy. These libraries are known for being efficient when it comes to handling & analysing data.

Rationale for Using Python in Descriptive Analysis

Python was a good choice here because it efficiently manages large datasets. It also helps automate processes, making it easier to repeat analyses while ensuring accuracy & consistency. Plus, Python's large collection of libraries means we can easily connect with advanced data processing & visualization techniques. It adapts well to changing analytical needs. Unlike traditional spreadsheet tools that can be limited, Python allows us to plot many different kinds of graphs quickly. This makes it better for handling big and complex datasets too—which reduces the chance of mistakes.

Using Python doesn't just improve accuracy and reliability for this initial set of descriptive statistics; it also creates a strong base for more advanced modelling work later on and for making decisions based on data. That's why we picked Python for this stage of the project; it gives us a thorough and solid method for analysing our data.

5.2 Excel for Sales and Inventory Data Reconciliation:

During this phase of the analysis, Excel was used to line up sales data with inventory data. The goal is to find any mismatches in total parts versus current inventory levels. The first step involved figuring out how many each part was sold using the sales data. This gave a broad picture of sales activities. At the same time, inventory data was looked at to establish how much had been purchased and what's available for each part now.

Rationale for Using Excel in Sales and Inventory Data Reconciliation

Excel helped reconcile the sales vs. inventory numbers accurately. First, it gathered the total amount sold for each part through the sales data using Pivot Tables. This gave a neat, organized snapshot of sales figures. Also, we analysed how much had been bought & what the current inventory levels are from the inventory data. By comparing total purchases and current stock, we estimated how much had been sold. To match these estimates with the sales numbers, we brought in Excel's VLOOKUP function. This made it easy to directly compare the records from sales & inventory. We picked Excel for its strong ability to analysed data and its easy-to-use tools that help us aggregate information and match data precisely. Using tools like Pivot Tables & VLOOKUP meant we could spot discrepancies accurately—things like unexpected shortages in stock that might point to data entry mistakes or even theft. Furthermore, Excel's built-in features made it simpler to show and explain these discrepancies clearly. This strengthened our overall analysis approach. In short, this method allowed for a complete reconciliation of sales & inventory data, offering valuable insights while keeping our inventory records accurate and trustworthy.

5.3 Linear Regression for Servicing Cost Analysis:

To tackle issues with final billing servicing costs, we conducted an analysis using linear regression in Python First, we took a deep dive in the sales data This helped us see the connection between product costs and servicing charges or any additional charges associated with it. We discovered that servicing costs were always at 10% of the product cost. This raised questions about whether this fixed rate captured the changing nature of servicing costs.

Rationale for Using Linear Regression in Servicing Cost Analysis:

We used linear regression to examine discrepancies in servicing costs. The goal was to find out if the servicing cost should remain a fixed percentage or if a more dynamic model might reflect real service needs better. Linear regression gave us a solid method for understanding how servicing costs change with product costs. Thanks to Python's powerful libraries, like scikit-learn, we could plot and analyze these relationships with precision.

Python was explicitly used in this process because it makes automated calculations easy. Also, it provides detailed visualizations, which are crucial for accurate analysis. Using linear regression in Python brought many benefits: it improved accuracy & offered insights into whether we should stick to a fixed percentage model or consider changes. Moreover, it allowed for efficient data management & analysis and let us build dynamic models that can adjust to varied information.

5.4 Graphical Analysis of Sales Trends and Inventory Patterns:

Analyzing & sales data thoroughly involved using both excel and Python. The main agenda is to identify what's available against the quantities ordered up until now. By plotting these trends we could see the between ordered items and what's on hand, which helped spotlight any discrepancies

or big orders that might not make sense. We also took a close look at sales trends for bikes & scooters. This provided clear insights into which vehicles were selling better including the most selling products through some simple excel analytics. Overall, this detailed approach was key for improving our understanding of inventory management practices along with sales dynamics.

Rationale for Using Graphical Analysis in Sales Trends and Inventory Patterns:

To study sales trends & inventory patterns available in the collected raw data effectively, we will use both Excel and Python for visualization. Initially, Excel created simple charts & graphs showing ordered versus available inventory. This made it easier to spot basic patterns & discrepancies right away. For deeper analysis, we turned to Python's plotting libraries like Matplotlib & Seaborn for detailed trend assessments, especially regarding bike & scooter sales data. By blending Excel's user-friendly visual tools with Python's advanced capabilities, we carried out a well-rounded analysis of both inventory and sales data. This combination offered better visualization along with valuable insights into sales performance & inventory management practices.

6. Results and Findings:

6.1 Sales and Inventory Analysis:

In our analysis of sales & inventory data, we used Excel along with pivot Tables. This helped us find gaps between how much was sold and what inventory is available with accuracy.

Analysis Overview:

We started by using Pivot Tables to sum up the total quantity sold for each product, pulling this from the sales data. This step ensured we had accurate total sales figures. At the same time, we looked at the inventory data to match the total quantity purchased against what was left in stock.

Key Findings:

1. **Discrepancies Identified:** Our review uncovered notable discrepancies between sold quantities and available inventory levels. For instance, with Product Front Shield SET, it showed that 2 units were sold while 8 units were expected in stock; however, only 6 units were actually listed as available in our records. This difference suggests there might be an error in reporting. A detailed report can be accessed here: *click_link*
2. **Financial Implications:** The discrepancies noted led to a financial variance of around Rs. 16,917 in the value of products, that raises concerns. Such a variance often indicates potential problems like wrong data entry, technical mistakes, or even theft that goes unreported.
3. **Potential Causes:** Several factors could have caused these errors. These include issues from manual data entry or glitches in our inventory tracking systems, plus cases of

unregistered theft. Investigating these areas is vital to solve and correct the discrepancies we found.

6.2 Results and Findings: Linear Regression Analysis of Service Costs

We used regression to analyze service expenses. This strategy allowed us to see how customers responded to spare parts costs, labor, servicing fees, and other potential expenses included in the final bill. Our goal was to determine whether the set 10% service charge is appropriate or if a more flexible pricing model is required.

Analysis Overview:

We used linear regression to investigate the relationship between total bill amount and service prices. We wanted to know if the 10% service charge on spare parts should vary based on the actual expenses and complexity of the service. Refer to link: *click_here*

Key Findings:

1. **Linear Relationship Observed**: The findings revealed a straight relationship between what clients pay and the service expenses. Specifically, the service price regularly accounted for 10% of the spare parts expenditure. For example, if the entire cost of spare components is 20,000 currency units, the associated service charge is 2,000 Rs. (or 10%).
2. **Inadequacy of Fixed Percentage**: Our data suggest that this set rate of 10% may not accurately reflect how complex or time-consuming a service is. For example, if a car requires a full day of extensive servicing, charging only 10% of the replacement parts (such as 2,000 Rs.) does not accurately reflect the amount of labour performed. This discrepancy suggests that using a dynamic pricing strategy could be better.
3. **Need for Dynamic Pricing**: The analysis clearly demonstrates the need to alter the current service price structure. A dynamic pricing model that varies based on actual service time and complexity would provide a more accurate picture of expenditures expended and services offered.

6.3 Results and Findings: Graphical Analysis of Sales Trends and Inventory Patterns

Overview

We performed a detailed examination of the sales and inventory. This analysis gave us valuable insights into managing inventory & assessing sales performance. It allowed us to spot key trends, along with some discrepancies between what was ordered and what was available in stock. The findings pointed out challenges like overstocking issues and differences in stock levels when comparing bikes to scooters. This led us to propose ways to improve inventory practices.

Key Findings:

1. **Discrepancies and Stock Management Issues:** Our review uncovered notable discrepancies between the amounts ordered versus the amounts actually consumed. In particular, some spare parts had considerable order quantities, yet the consumed quantities were quite low. This discrepancy indicates potential challenges in how inventory is managed, such as overstock situations & not enough consumption. Refer Link: *click_here*

2. **Low Order Quantities and High Consumption Rates:** The analysis showed that several items had very low stock levels when compared to their actual consumption. For example, spare parts like AIR Filter Rcf/12, BAJAJ 4T Premium Engine Oil (20W50) for BS4 and BS6, High Performance Four Stroke Bikes Engine Oil 1L, S E M Two Wheeler Brake Shoe Front and Rear Brake Shoe, along with SERVO 4T SAE 4 Stroke Engine Oil intended for Motorcycles (BS-4 and BS-6), displayed large consumption quantities between 9 to 18 (as identified through Excel lookups), while their ordering rate from vendor were still low overall. This imbalance suggests a problem with how stock is ordered, leading to inventory management issues. Refer Link: *Click_here*

3. **Bikes vs. Scooters:** Additionally, we noticed that the number of scooters in inventory was substantially higher compared to bikes. This pattern hints at a possible overstock situation for scooters, which could negatively impact overall efficiency in inventory management. Refer fig.6.3.a

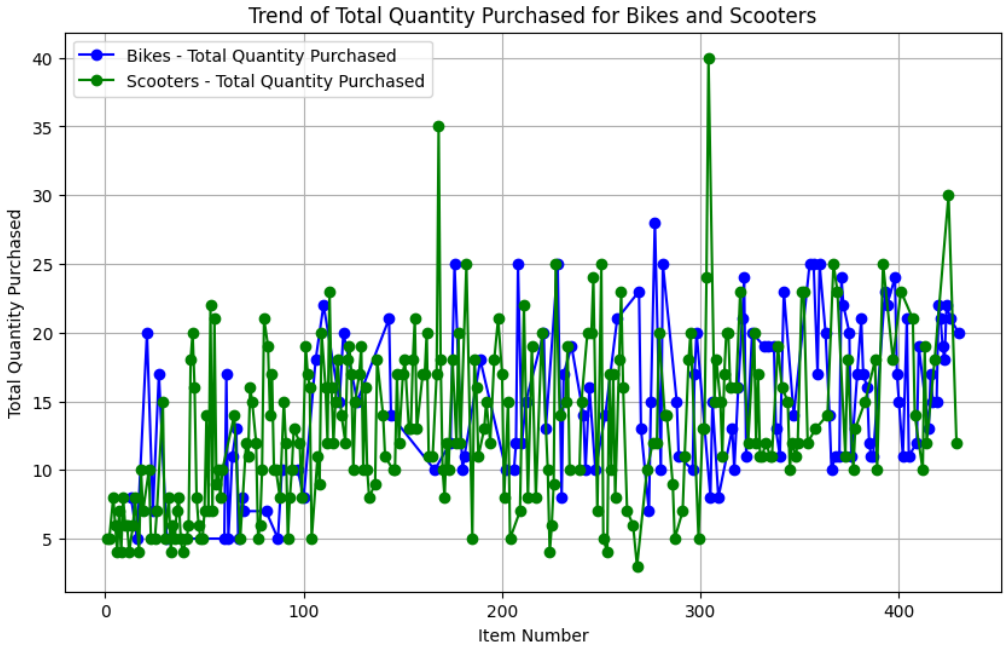

fig.6.3.a : Trend of Total Quantity Purchased for Bikes and Scooters.