Profit Optimization through Data-Driven Analysis for Business Development and Understand Cycle shop trend Shop:Usman Cycle Works

A Mid Term Report For The BDM Capstone Project

Submitted by Name : Taniya Chouhan
Roll number : 23f2003584

IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036.

CONTENTS

EXECUTIVE SUMMARY

I am collaborating with Mr. Mehboob Ali, owner of Usman Cycle Works, a B2C cycle shop selling bicycles, tyres, tubes, and accessories from brands like Avon, Hero, Aqua, Holland, and Honda. The shop operates daily from 8 AM to 8 PM but faces declining profits, stock management issues, and a lack of customer insights due to rising expenses, shifting consumer preferences, and post-lockdown inventory challenges.

To address these issues, I collected and analyzed transactional data on sales and purchases. The sales data spans from April 1, 2024 to July 21, 2024 and includes 473 entries, while the purchase data covers April 4, 2024 to July 1, 2024 with 29 entries. The dataset captures key information such as transaction date, item details, quantity, unit price, and amount.

The analysis process involved data cleaning, descriptive statistics, and exploratory data analysis (EDA) using Python libraries (Pandas, NumPy, Matplotlib, Seaborn) and Excel. I standardized item names for consistency, calculated profit percentages, and created visualizations (bar charts and boxplots) to highlight sales performance, purchase patterns, and stock inefficiencies. Descriptive statistics (mean, median, standard deviation, range, skewness, kurtosis, and outliers) were applied to understand data distribution, variability, and detect unusual values.

Key findings showed that cycles were the main profit drivers, while accessories with 12% and 18% GST incurred losses. Accessories with 5% GST had small profits. Tubes and tyres sold well but had negative margins. Bells, locks, pumps, and horns showed low sales and minimal profit, indicating weak demand.

Next steps include time series analysis for demand forecasting and a detailed stock-to-sales ratio evaluation to guide inventory adjustments. These insights will support data-driven decisions to improve inventory efficiency, cut unnecessary costs, and increase profitability.

PROOF OF ORIGINALITY

Business Name : Usman Cycle Works
Business Owner name : Mr. Mahboob Ali
Business Address : Rani Satti Road, Gandhi Chowk 333001, Jhunjhunu, Rajasthan

To verify the authenticity of the data, I have submitted the following evidence:

The owner has provided me with a letter with his signature, stamp and contact detail.

1.IMAGES OF THE ORGANIZATION:

https://drive.google.com/drive/folders/1_3b64DTaaXeeLFFkYzsPC2Um7TbmrsTC?usp=sharing

I have uploaded the outer and inner view of the Usman Cycle Works shop, along with an image of the products and a picture of a sample bill book.

2.DATASET:

https://drive.google.com/drive/folders/1LflSN4YpeSaOVTjfamq1O4IWq5pyec9a?usp=sharing

3. Colab Notebook:

https://colab.research.google.com/drive/11Y-XnoMCmA7B8w6Xotgxjqw0LrxVDJGY?usp=sharing

Created the colab notebook using python libraries for Analysis of the data.

METADATA:

The owner provided sales and purchase transaction data for a bicycle-related business, organized into eight columns: Date, Vch/Bill No., Particulars, Item Details, Qty, Unit, Price, and Amount. The sales dataset contains a total of 473 rows, while the purchase dataset has 29 rows, with the first row in each serving as the title. The sales data spans from April 1, 2024, to July 21, 2024, and the purchase data ranges from April 4, 2024, to July 1, 2024. Dates are formatted in the Month/Date/Year (MM/DD/YYYY) style. Each record documents either a sale or a purchase of bicycle-related products such as cycles, accessories, tubes, pumps, bells, locks, tyres, and other related items. In the 'Particulars' column, sales entries mention the name of the buyer, while purchase entries list the company or vendor from which the item was procured. The 'Amount' column is computed as the product of the quantity (Qty) and the unit price (Price).

Table 1.1: For Sales and Purchase Datasets

KEY	**DESCRIPTION**
Date	The date on which an item was sold or purchased.
Vch/Bill No.	A unique identifier (voucher or bill number) for each transaction.
Particulars	For sales: Name of the buyer. For purchases: Name of the supplier/vendor.

Item Details	Describes the item type (e.g., Cycle, Accessories, Tube, Bell, Lock, etc.).
Qty	Quantity of items involved in the transaction.
Unit	Measurement unit for quantity (always in "Pcs").
Price	Per-unit price of the item.
Amount	Total transaction amount, calculated as Qty × Price.

Vendors/Companies Mentioned:- Hindustan Tyre Co., GURU RAM DASS CYCLE STORE, Hero Cycle L.M.T, Jaipur Cycle, Ridhi Shidhi Enterprises, M.K Cycle Industries, Avon Cycle L.M.T, Single Traders.

Item Categories:- Includes a variety of bicycle-related products such as Cycle, Accessories, Tube, Bicycle Horn, Kid Toy Motorcycle, Bell, Lock, Pump, Tyre.

Further analysis of this metadata will focus on key business insights. Sales trends will be identified by analyzing data over time to detect peak sales periods and seasonal patterns. Top-selling products will be determined to highlight the most popular items. Customer preferences will be examined using buyer-specific sales data to understand purchasing behaviors. Revenue analysis will assess total sales amounts to identify high-value transactions and overall profitability. Lastly, inventory management will be optimized by tracking stock movement through quantity data, ensuring efficient stock control.

DESCRIPTIVE STATISTICS:

After organizing and cleaning the data, I decided to perform descriptive statistics to summarize and understand various characteristics of data. I have performed descriptive statistics on Sales data and Purchase Data like mean, median, mode,standard deviation, variance, range, IQR, Skewness, Kurtosis, 25th percentile, 75th percentile, outliers and also created outlier plot and histogram plot. It will be helpful to know about the variability, spreadness, average sales, average quantity sold.

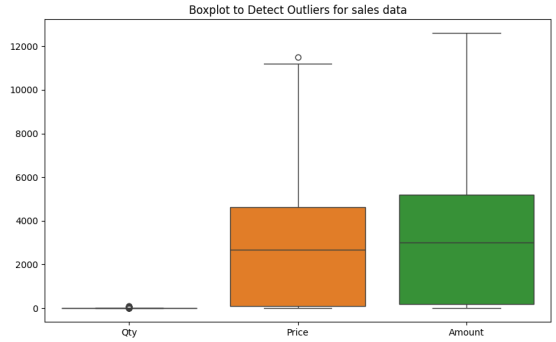
Table 2.1: Sales Data:-

Descriptive Statistics	Quantity of sale Item	Per sold Item Price	Total Amount Earn
Sum	1425.0	₹1278001.855	₹1535215.69

Mean	3.019068	₹2707.631049	₹3252.575614
Median	1.0	₹2676.0	₹3000.0
Mode	1.0	₹20.0	₹20.0
Standard Deviation	8.963814	₹2504.491936	₹2928.915322
Variance	8.034995e+01	₹6.272480e+06	₹8.578545e+06
Range	99.0	₹11498.01	₹12605.06
IQR(Interquartile range)	1.0	₹4520.8375	₹5030.06
Skewness	7.473712	0.625700	0.589553
Kurtosis	66.059392	-0.149952	-0.416485
25th percentile	1.0	₹90.0	₹170.0
75th percentile	2.0	₹4610.8375	₹5200.06
Outliers(z-score >3)	11	4	2

From performing the descriptive statistics we can identify that for months April 2024 to July 2024 total revenue generated is ₹ 1535215.69 by selling 1425 products.The range shows a significant variation in demand across different products. By analyzing the sales data, we can determine which products sell more frequently than others. The variance and standard deviation indicate moderate variability in sales, with some transactions involving much higher quantities and prices. The mean reveals that the average quantity sold per transaction is 3, with an average selling price of ₹2707.63, leading to an average transaction value of ₹3252.58. In Quantity highly right-skewed 7.47, indicating rare bulk purchases. In Price, slight right skew 0.625, indicating some higher-priced items. In Amount, Slight right skew 0.589, meaning occasional high-value transactions.

Figure 1.1: Boxplot of sales dataset



The boxplot (Figure 1.1) shows that Quantity is highly skewed with most sales involving very few items, but a few extreme outliers represent large quantity sales. Price and Amount have wider distributions, indicating greater variability. Some high-value outliers are present, especially in Price, where a few items were sold at significantly higher prices. Amount also shows a broad range due to combined effects of quantity and price, with fewer outliers. Overall, the plot highlights typical sales patterns and pinpoints unusual transactions.

Table 2.2: Purchase Data:-

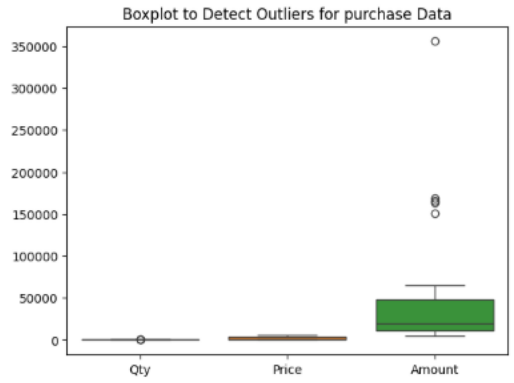
Descriptive Statistics	Quantity of purchased Item	Per purchased Item Price	Total Amount Paid
Sum	4779.0	₹47269.17	₹1502630.45
Mean	170.678571	₹1688.184643	₹170.678571
Median	76.50	₹152.95	₹19573.77
Mode	4.0	₹19.94	₹4475.0
Standard Deviation	239.521754	₹1979.482305	₹78768.482305
Variance	5.737067e+04	₹3.918350e+06	₹6.204551e+06
Range	946.0	₹5364.06	₹351174.67
IQR(Interquartile range)	182.75	₹3684.805	₹36146.365
Skewness	2.213889	0.573183	2.570355
Kurtosis	4.565907	-1.503756	7.368877
25th percentile	26.5	₹70.015	₹11523.45
75th percentile	209.25	₹3754.82	₹47669.815

Descriptive Statistics	Quantity of purchased Item	Per purchased Item Price	Total Amount Paid
Sum	4779.0	₹47269.17	₹1502630.45
Outliers(z-score >3)	1	0	1

The total purchase cost for the analyzed period is ₹15,02,630.45, with a total of 4779 items purchased. Using the mean calculation, the average purchase price per item is ₹1688.18, and the average quantity purchased per transaction is 170.68. The highest purchased quantity in a single transaction is 946, while the lowest is 1, indicating significant variations in purchase volumes across different products. The range of ₹5364.06 in item prices shows a wide disparity in spending on different products. The standard deviation of ₹1979.48 for per-item price suggests moderate price variability, while the skewness of 0.57 indicates that some high-priced items are pulling up the average. The total amount paid is highly variable (Std Dev: ₹78,768.48, Skewness: 2.57), indicating that while most purchases fall within a typical range, a few large transactions significantly impact overall expenditure.

The boxplot (Figure 1.2) shows that Quantity and Price in purchase data are mostly consistent with few or no major outliers. However, Amount displays several extreme outliers, with some purchases exceeding ₹150,000 and even ₹350,000, indicating a few unusually large transactions likely due to bulk buying or high-value items.

Figure 1.2: Boxplot of purchase dataset



ANALYSIS PROCESS AND METHODS

Analyzing data has several steps. Defining the objective, collection of data , data cleaning and organizing , Exploratory Data analysis process using descriptive statistics,charts, Finding

trends, patterns and getting insights and Interpreting results. I started with defining the objective of my analysis process. When I approached the Usman Cycle Works shop owner and explained about this project, He was ready to provide full support to my project enthusiastically. The owner and I have talked over several problems. I found out that the two major issues are profit and Stock. I decided to collect data and work on those problems.To understand the problem much more I talked with some of the employees and customers . To proceed with analysis, I asked the owner to share the possible data. He has shared data about sales and purchase. The analysis of the shop's sales and purchases was conducted using a structured approach that involves data collection, visualization, and interpretation. Below is an outline of the process and methods used in this analysis:

Data Collection:- The data was collected through a primary data collection method. Sales and purchase records were directly obtained from Usman Cycle Works in the form of Excel spreadsheets for April 2024, May 2024, June 2024, July 2024, the dataset included details such as item names, purchase quantities, sales quantities, purchase costs, and selling prices. Profit percentages were derived from the difference between sales revenue and purchase costs.

Data Preprocessing:- In the preprocessing step, inconsistent item names in the "Item Details" column were standardized for clarity. Items like "ACCESSORIES@12%", "18%itam", and "Accessories" were cleaned and renamed to clear formats such as "Accessories (12% GST)" ,"Accessories (18% GST)" and "Accessories (5% GST)" respectively, since the accessories data included prices categorized by GST tax rates. Variations like "BICYCLE HORN" and "bicycle horn" were unified as "Bicycle Horn", and all other entries were converted to title case. This improved consistency and made the data ready for accurate analysis.

Data Visualization:- To interpret trends effectively, various visualization techniques were used: Bar Charts, used to represent total quantities of sales and purchases per item. This helped identify which products had high demand and which ones were overstocked.Total Purchase & Sales Amount Analysis, bar plots were used to compare the monetary value of items purchased versus sold. This provided insights into whether investments in stock were translating into revenue. Profit Percentage Analysis, a bar chart showing profit margins per item helped identify which products were profitable and which were incurring losses.

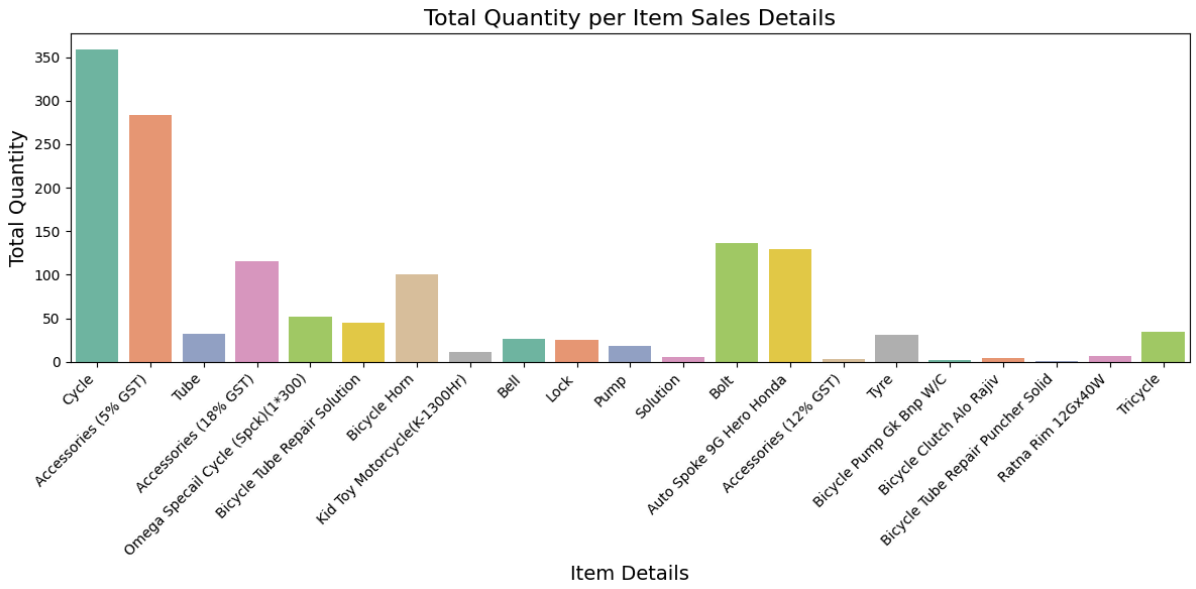
Key Analysis Methods:- Comparative Analysis, sales and purchase data were compared to determine if there was overstocking or understocking of items.
Profitability Analysis, profit percentages were calculated using the formula:

$$Profit\,\% = ((Sales\ Amount - Purchase\ Amount)/Purchase\ Amount)\times 100$$

This helped identify loss-making and profit-generating items. Trend Identification, patterns in sales and purchase data were examined to determine seasonal demand and slow-moving stock.Stock Optimization Suggestions, based on sales performance, recommendations were made to adjust purchase quantities, apply discounts, or increase stock for high-demand items.

RESULTS AND FINDINGS:

Figure 2.1:Bar Chart for Quantity of item sold in between April 2024 to July 2024

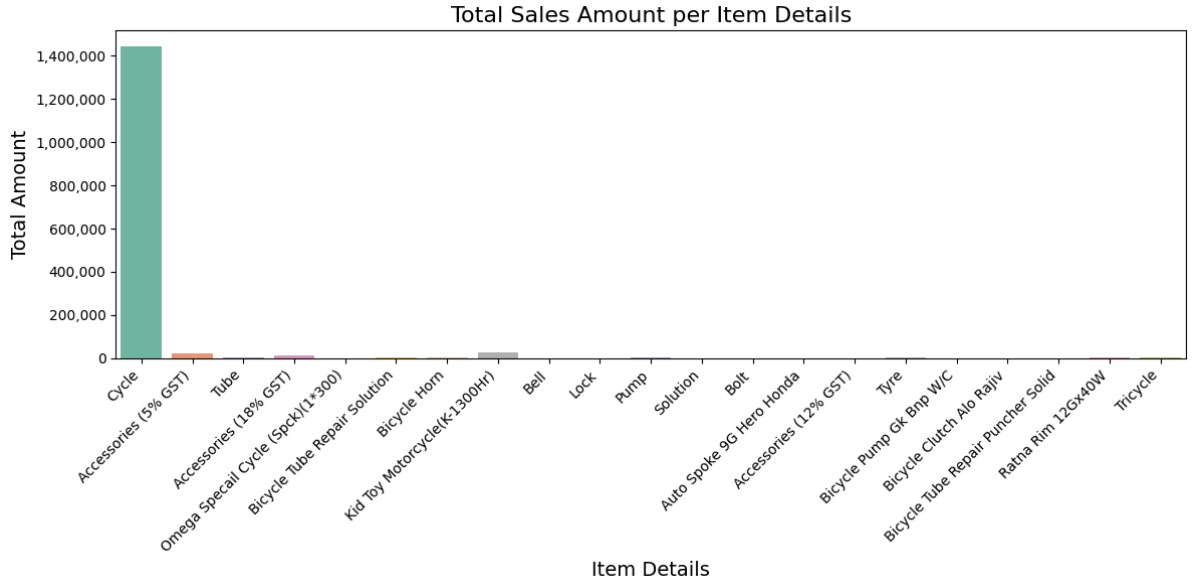


The sales and purchase data of Usman Cycle Works highlight key inventory and demand insights. Cycles, Accessories (5% GST), and Tubes are the top-selling items, while Bicycle Tube Repair Solution and Bicycle Horn have moderate sales (Figure 2.1). Pump, Lock, Belt, and Tyre show significantly lower demand. In purchases (Figure 2.2) , Accessories (12% GST) lead in quantity, surpassing all other items, followed by Tubes, Accessories (18% GST) and Cycles, indicating possible stock accumulation. The shop should optimize inventory to prevent overstocking slow-moving items while ensuring high-demand products remain available. To boost sales, promoting slow-moving products through discounts or bundles

could help. Effective stock management and demand forecasting are essential for maintaining balanced inventory and maximizing profitability.

Figure 2.2: Bar Chart for Quantity of item purchase in between April 2024 to July 2024

Figure 3.1: Bar Chart for Total Amount Earn for each item in between April 2024 to July 2024



The purchase and sales amount data reveal key financial insights into the shop's performance. Cycles account for the highest purchase spending, followed by Accessories (12% GST), Tyres, Accessories (18% GST) and Tubes, making cycles the primary inventory investment. In sales (Figure 3.1), Cycles generate the highest revenue, while Accessories (5% GST),

Tubes, and repair solutions contribute minimally. A notable mismatch exists between purchase (Figure 3.2) and sales amounts-Accessories (12% GST) is heavily stocked but has low sales, indicating slow turnover or over-purchasing. Similarly, small repair items and bicycle components generate low revenue, suggesting weak demand. To improve profitability, the shop should optimize inventory management, ensuring high-cost items like Accessories (5% GST) drive sales. Low-revenue, high-cost items should be reassessed through promotions or procurement adjustments. With Cycles as the core revenue driver, maintaining a balanced stock while controlling excess inventory in other categories is crucial.

Figure 3.2: Bar Chart for Total Amount Spend for each item in between April 2024 to July 2024

The Figure 4.1 breaks down profit percentages by item and GST category, revealing important differences. Among the accessories, those taxed at 5% show a small profit, while accessories with 12% and 18% GST result in significant losses, with the 12% GST category showing the most negative profit. Cycles show the highest positive profit percentage, indicating good profitability. Tubes and Tyres both reflect high losses, similar to accessories under higher GST rates. In Figure 4.2, all accessories are grouped together under "Accessories (All GST)," resulting in an average loss of around 80%, which conceals the variation among the individual GST categories.

Figure 4.1: Bar Chart for Profit Percentage per Item when Accessories analysed according to GST %

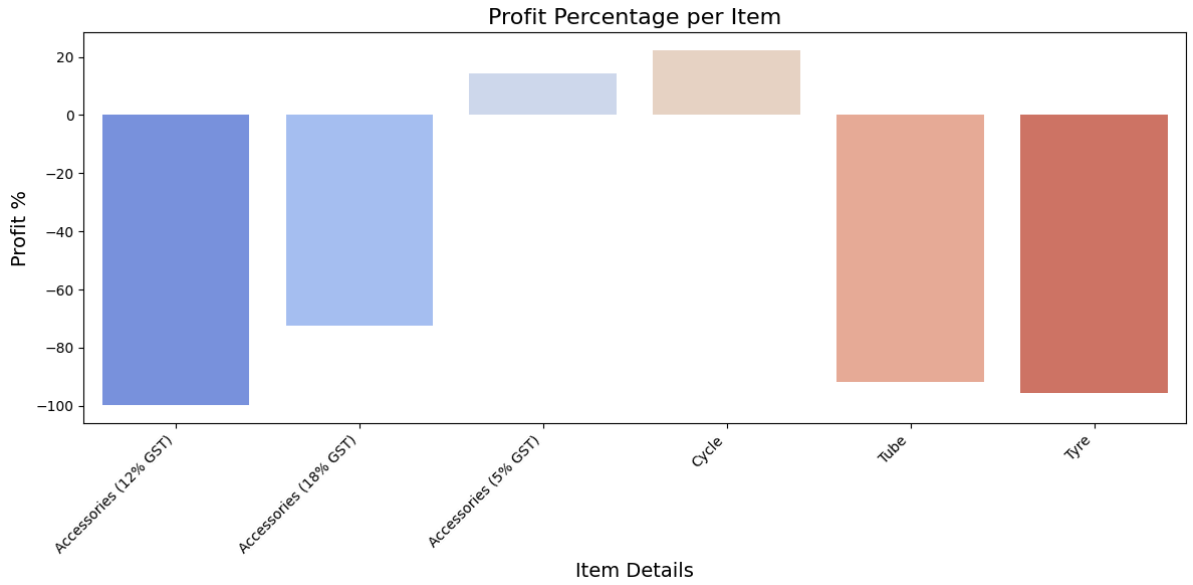


Figure 4.2: Bar Chart for Profit Percentage per Item when Accessories were combine