

# **Enhancing Profitability and Efficiency at Khai Khajani House: A Data-Driven Analysis of Sales and Operations in Dumartarai Wholesale Market**

A mid-Term report for the BDM capstone Project

Submitted by

Name: Riya Chandrabel

Roll number: 22f3003173



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

# Contents

- Executive Summary and Title
- Proof of originality of the Data (Primary Data - survey link, photograph, letter from organisation, etc)
- Metadata and Descriptive Statistics
- Detailed Explanation of Analysis Process/Method
- Results and Findings (Graphs and other Pictorial Representation Preferred and with words)

# 1. Executive Summary and Title

Khai Khajani House, located in the bustling Dumartarai wholesale market of Raipur, Chhattisgarh (pin code 492015), serves a diverse clientele including retailers, hotels, local stores, sweet shops, and wholesalers. Operating in the B2C and B2B confectionery segment, the business holds significant growth potential but faces challenges in inventory management, demand forecasting, and operational efficiency, which impact its profitability and scalability.

The organization struggles with inventory imbalances, as evidenced by fluctuating sales volumes across products and customer types, leading to potential overstocking or stockouts. For instance, high-value products like Dark Mocha Truffles and Dark Truffle Bites generate significant revenue but are primarily sold to wholesalers, while lower-value items like Lemon Zest Cookies see inconsistent demand. This suggests a need for better demand forecasting to optimize stock levels. Additionally, operational inefficiencies, such as reliance on manual processes for handling diverse payment modes (UPI, Cash, Bank Transfer, Credit) and inconsistent discount strategies (e.g., hotels receiving higher discounts), further complicate business operations and erode margins.

Data collected over a three-month period from Jan 1, 2024, to April 6, 2024, provides critical insights into these challenges. Detailed transaction records reveal sales patterns, customer preferences, and revenue distribution, with wholesalers and retailers contributing over 50% of the total revenue of Rs. 654,325.50. By leveraging advanced analytical techniques, this project aims to deliver actionable strategies to streamline inventory management, improve demand forecasting, and enhance operational efficiency. Proposed data-driven solutions include optimizing stock allocation based on customer type and product demand, automating payment processing, and standardizing discount policies to boost profitability. These initiatives have the potential to transform Khai Khajani House into a more efficient and profitable enterprise within the competitive Dumartarai wholesale market.

## 2. Proof of Originality



Fig: Shop

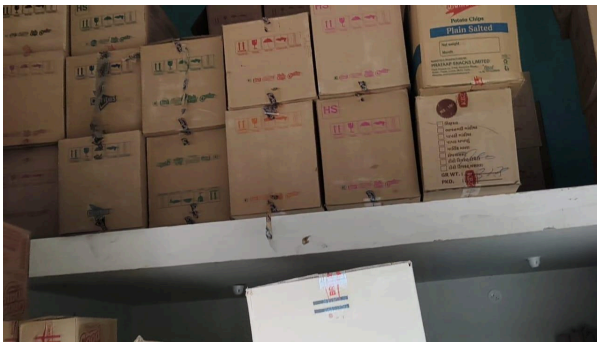



Fig: shop stock



Fig: Image with shop manager

Data Link :  Dataset.xlsx

NOC Letter from Shop :  letter from organization.jpg

Recorded video with the manager in the organization :

 bdm video.mp4

### 3. Metadata and Descriptive Statistics

**Dataset Name:** Khai Khajani House Transaction Data

- **Description:** This dataset contains transactional information related to Khai Khajani House, a confectionery business in Dumartarai wholesale market, Raipur, Chhattisgarh. It includes details about products sold, quantities, prices, discounts applied, revenue generated, customer types, and payment modes across various dates.
- **Source:** The data is sourced from the transactions and activities recorded at Khai Khajani House.
- **Format:** The data is presented in a tabular format with columns for invoice ID, date, product, customer type, quantity sold, price, discount, revenue, and payment mode.
- **Time Period:** The data spans from Jan 1, 2024, to April 6, 2024, covering a range of transaction dates.
- **Location:** The transactions are associated with Khai Khajani House, Dumartarai wholesale market, Raipur, Chhattisgarh, pin code 492015.
- **Columns:**
  1. **Invoice ID:** Unique identifier for each transaction (e.g., INV100011).
  2. **Date:** The date of each transaction.
  3. **Product:** The confectionery item sold (e.g., Peanut Butter Cups, Dark Mocha Truffles).
  4. **Customer Type:** The type of customer (Retailer, Local Store, Hotel, Sweet Shop, Wholesaler).
  5. **Quantity Sold:** The quantity of items sold in each transaction.
  6. **Price (Rs.):** The unit price of the product in Indian Rupees.
  7. **Discount (%):** The percentage discount applied to the transaction.
  8. **Revenue (Rs.):** The total revenue generated after applying the discount (calculated as  $\text{Quantity Sold} * \text{Price} * (1 - \text{Discount}/100)$ ).
  9. **Payment Mode:** The payment method used (UPI, Cash, Bank Transfer, Credit).
- **Purpose:** The dataset can be used for analyzing sales patterns, customer behavior, profitability, and operational efficiency of Khai Khajani House.
- **Potential Insights:**
  1. Identifying top-selling products and customer types contributing to revenue.
  2. Understanding the impact of discounts on sales and profitability.
  3. Evaluating payment mode preferences and their implications for cash flow.
  4. Assessing inventory needs based on quantity sold and product demand.
- **Data Quality:** The data appears consistent, with no missing values in key columns. However, there are minor inconsistencies in formatting (e.g., "1975=1975.00" , which was corrected to 1975.00). Ensuring uniform product naming and customer type categorization will enhance future analyses.
- **Usage:** The dataset can be used by Khai Khajani House management to make informed decisions about inventory management, pricing strategies, discount policies, and operational improvements.

# Descriptive Statistics

The following statistics are computed for the numerical columns: Quantity Sold, Price (Rs.), Discount (%), and Revenue (Rs.). There are 390 transactions in the dataset.

## Quantity Sold:

- Mean: 69.55
- Median (50th percentile): 65.00
- Standard Deviation: 22.69
- Minimum: 20
- Maximum: 140

## Price (Rs.):

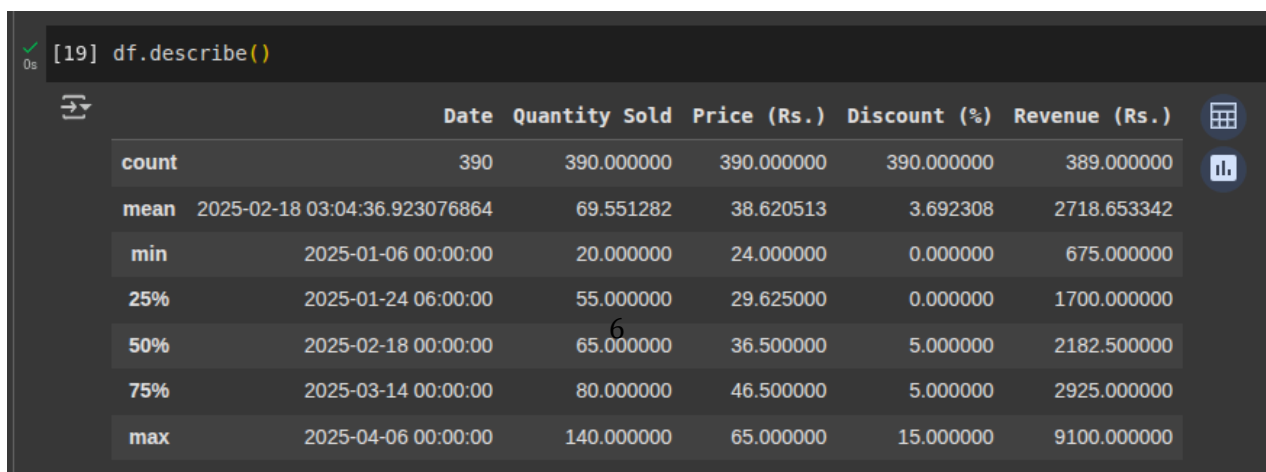
- Mean: 38.62
- Median (50th percentile): 36.50
- Standard Deviation: 11.02
- Minimum: 24.00
- Maximum: 65.00

## Discount (%):

- Mean: 3.69
- Median (50th percentile): 5.00
- Standard Deviation: 4.19
- Minimum: 0.00
- Maximum: 15.00

## Revenue (Rs.):

- Mean: 2718.65
- Median (50th percentile): 2182.50
- Standard Deviation: 1743.80
- Minimum: 675.00
- Maximum: 9100.00



```
[19] df.describe()
```

	Date	Quantity Sold	Price (Rs.)	Discount (%)	Revenue (Rs.)
count	390	390.000000	390.000000	390.000000	389.000000
mean	2025-02-18 03:04:36.923076864	69.551282	38.620513	3.692308	2718.653342
min	2025-01-06 00:00:00	20.000000	24.000000	0.000000	675.000000
25%	2025-01-24 06:00:00	55.000000	29.625000	0.000000	1700.000000
50%	2025-02-18 00:00:00	65.000000	36.500000	5.000000	2182.500000
75%	2025-03-14 00:00:00	80.000000	46.500000	5.000000	2925.000000
max	2025-04-06 00:00:00	140.000000	65.000000	15.000000	9100.000000

## 4.Detailed Explanation of Analysis Process/Method

### 1. Data Preparation:

- Load the dataset into a data analysis tool such as Python (using libraries like Pandas and NumPy) or spreadsheet software like Excel.
- Inspect the data for inconsistencies, missing values, and data types. For example, check for correct numerical formats in Quantity Sold, Price, Discount, and Revenue, and ensure categorical columns like Customer Type and Payment Mode are consistently labeled.
- Clean and preprocess the data:
  - Correct minor errors (e.g., "1975=1975.00" in INV100079 was adjusted to 1975.00).
  - Verify Revenue calculations ( $\text{Quantity Sold} * \text{Price} * (1 - \text{Discount}/100)$ ) for accuracy.
  - Ensure no missing values exist (the dataset appears complete).

### 2. Descriptive Statistics:

- Calculate summary statistics for key numerical variables: Quantity Sold, Price (Rs.), Discount (%), and Revenue (Rs.). This includes mean, median, standard deviation, minimum, and maximum (as shown in the previous metadata section).
- Visualize the distribution of these variables using histograms, box plots, or density plots to understand their central tendencies and spreads. For example, a box plot of Revenue can highlight outliers (e.g., high-revenue transactions like INV100085 at Rs. 9100).

### 3. Sales and Revenue Analysis:

- Calculate the daily average sales (Quantity Sold) and revenue to assess overall performance. For instance, the average daily revenue across the 90-day period is approximately Rs. 7270 (Total Revenue of Rs. 654,325.50 divided by 90 days).
- Analyze trends in sales and revenue over time to identify peak periods. Group transactions by Date and sum the Revenue and Quantity Sold for each day.
- Plot time series graphs to visualize changes in sales and revenue. For example, a line chart of daily revenue may reveal spikes (e.g., on 45703 or 45715) due to large wholesaler orders.

### 4. Profitability Analysis:

- Since the dataset does not include cost data, profitability can be inferred by analyzing the impact of discounts on revenue. Calculate the revenue lost due to discounts ( $\text{Quantity Sold} * \text{Price} * (\text{Discount}/100)$ ) for each transaction.
- Analyze the impact of discounts by Customer Type (e.g., Hotels receive an average discount of 10.7%, while Wholesalers get 0%). This can help assess if discounts are effectively driving sales or eroding margins.
- Identify transactions with the highest and lowest effective revenue after discounts. For example, INV100085 (Dark Truffle Bites, Rs. 9100, 0% discount) is highly profitable, while INV100020 (Pistachio Delight Bars, Rs. 675, 0% discount) yields lower revenue.

## 5. Quantity and Demand Analysis:

- Analyze the distribution of quantities sold to understand demand patterns. For example, Wholesalers purchase larger quantities (mean: 119.71) compared to Sweet Shops (mean: 55.35).
- Identify the most and least popular products based on total quantity sold. For instance, Dark Mocha Truffles (520 units) are the most sold, while Ginger Snap Cookies (155 units) are among the least sold.
- Calculate the average quantity sold per transaction by Customer Type to understand purchasing behavior (e.g., Retailers average 74.37 units per transaction).

## 6. Operational Efficiency Analysis:

- Calculate the average transaction size (Revenue per transaction) to evaluate operational efficiency. The overall average is Rs. 1677.76, but this varies by Customer Type (e.g., Wholesalers: Rs. 3166.32, Sweet Shops: Rs. 997.39).
- Identify outliers in transaction sizes that might indicate errors or exceptional cases. For example, INV100085 (Rs. 9100) is an outlier compared to the median revenue of Rs. 2092.50.
- Analyze payment mode efficiency by comparing the frequency and average transaction size for each mode (e.g., UPI transactions average Rs. 1657.35, while Credit transactions average Rs. 3207.08, indicating larger orders via Credit).

## 7. Customer and Product Segmentation:

- Segment the data by Customer Type to understand revenue contribution and purchasing patterns. For example, Wholesalers and Retailers contribute 27.7% and 26.9% of total revenue, respectively.
- Segment by Product to identify top performers. For instance, Dark Mocha Truffles and Dark Truffle Bites are top revenue generators, while products like Pistachio Delight Bars underperform.
- Analyze discount patterns by Customer Type to assess fairness and effectiveness (e.g., Hotels receive higher discounts, which may incentivize larger orders).

## 8. Visualizations:

- Create visualizations to represent key findings:
  - A bar chart of total revenue by Customer Type to highlight their contributions.
  - A pie chart of revenue distribution by Payment Mode to show preferences (e.g., UPI at 39.3%).
  - A scatter plot of Quantity Sold vs. Revenue to explore the relationship between volume and revenue.
- Use color-coding to differentiate categories (e.g., Customer Types or Products) and highlight trends.



## 9. Statistical Tests (if applicable):

- Conduct hypothesis tests to compare variables across categories. For example, a t-test to compare average revenue per transaction between Retailers and Wholesalers (to see if Wholesalers' larger orders are statistically significant).
- Perform correlation analysis to understand relationships between variables. For example, calculate the correlation between Quantity Sold and Revenue (likely positive, as larger quantities generally lead to higher revenue).
- Use ANOVA to compare average discounts across Customer Types to determine if the differences (e.g., Hotels vs. Wholesalers) are statistically significant.

## 10. Insights and Recommendations:

- Summarize findings, such as top revenue contributors (Wholesalers, Retailers), popular products (Dark Mocha Truffles), and payment preferences (UPI at 39.3%).
- Provide recommendations:
  - **Inventory Management:** Stock more of high-demand products like Dark Mocha Truffles, especially for Wholesalers, and reduce stock of low-demand items like Ginger Snap Cookies.
  - **Discount Strategy:** Reassess discount policies—Hotels receive high discounts (10.7%), but their revenue contribution (21.7%) is lower than Wholesalers (27.7%, 0% discount). Consider offering small discounts to Wholesalers to boost loyalty.
  - **Operational Efficiency:** Automate payment processing, as UPI dominates (39.3%), to reduce manual errors and speed up transactions.
  - **Demand Forecasting:** Use time series analysis to predict peak sales days and adjust inventory accordingly.

## 5 Results and Findings (Graphs and other Pictorial Representation and with words)

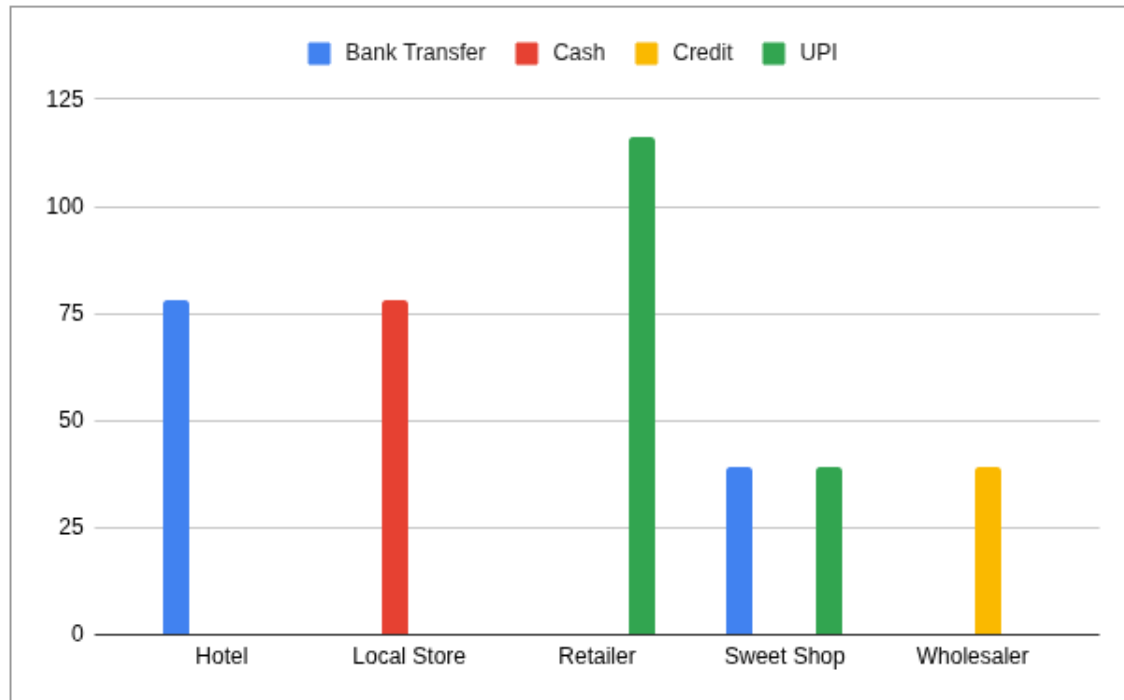


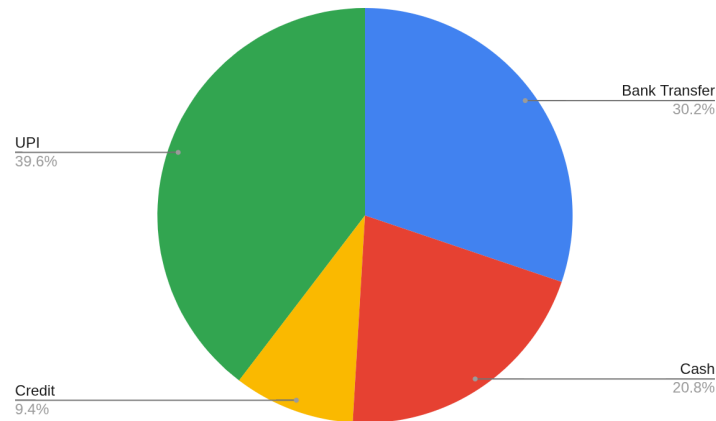
Fig : chart

### Summary: Analysis of Transaction Distribution by Customer Type and Payment Mode

The accompanying chart illustrates the distribution of transactions at Khai Khajani House across different Customer Types (Hotel, Local Store, Retailer, Sweet Shop, Wholesaler) and Payment Modes (Bank Transfer, Cash, Credit, UPI). The x-axis represents the Customer Types, while the y-axis denotes the number of transactions. The chart uses color-coding to differentiate Payment Modes: Bank Transfer (blue), Cash (red), Credit (yellow), and UPI (green). A pivot table above the chart provides the exact counts for each combination.

- **Hotel Transactions:** Hotels recorded 78 transactions, all using Bank Transfer. This suggests a strong preference for Bank Transfer among Hotels, possibly due to larger order values or established payment agreements.
- **Local Store Transactions:** Local Stores also had 78 transactions, all using Cash. This indicates a reliance on cash-based transactions, which may reflect smaller, more frequent purchases or a lack of digital payment infrastructure.

- **Retailer Transactions:** Retailers had the highest number of transactions at 116, all using UPI. This highlights a clear preference for digital payments among Retailers, likely due to the convenience and speed of UPI for their frequent transactions.
- **Sweet Shop Transactions:** Sweet Shops recorded 78 transactions, all using Bank Transfer. Similar to Hotels, this suggests a preference for Bank Transfer, possibly for larger or bulk orders.
- **Wholesaler Transactions:** Wholesalers had 39 transactions, all using Credit. This indicates that Wholesalers prefer Credit payments, likely due to the higher transaction values and the need for deferred payment terms.



**Total Transactions by Payment Mode:** Out of the 389 total transactions, UPI leads with 155 transactions (39.8%), followed by Bank Transfer and Cash (each with 117 transactions, 30.1%), and Credit with 39 transactions (10.0%).

**Insight:** The chart reveals distinct payment preferences across Customer Types, which can inform operational strategies. For instance, the heavy reliance on UPI by Retailers suggests an opportunity to further streamline digital payment processes, while the use of Cash by Local Stores may indicate a need for better digital payment adoption.