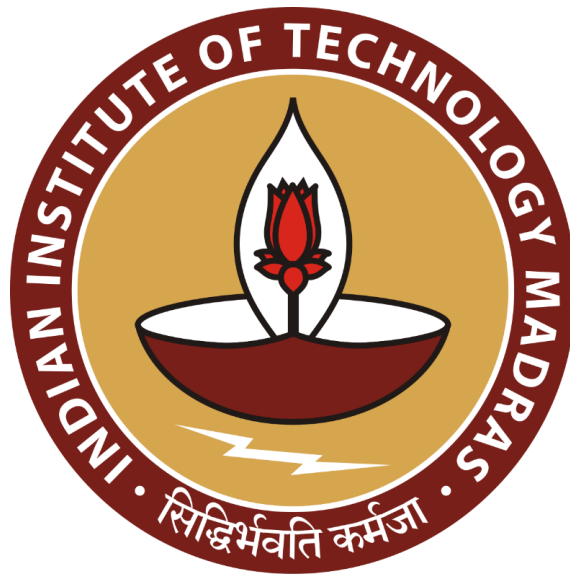# Sales Analytics for a Computer Parts Retailer

A Mid Term report for the BDM capstone Project

Submitted By :

Aditya Shanker Srivastava

Roll No. : 23f1001545

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

EXECUTIVE SUMMARY :

This report presents a comprehensive analysis of the sales data from WebTech System, showcasing the originality of the organization through relevant links and an attached letter of authorization. The metadata section contains the link to the dataset, along with detailed explanations of each column, ensuring clarity in understanding the data structure.

The report begins with a thorough explanation of the data cleaning and preprocessing steps undertaken to transform raw data into a usable format. This includes identifying and addressing missing values, correcting data types, and removing outliers. Initial assumptions were made, such as dropping certain values to focus on key products and customers. Following this, descriptive statistics were conducted to gain insights into the dataset's patterns.

Various analytical methods were employed, each accompanied by justifications for their use. Data cleaning and preprocessing facilitated a reliable foundation for analysis, while descriptive statistics allowed for a deeper understanding of product and customer behavior. Distribution analysis helped segment products and customers effectively, and trend analysis revealed significant sales patterns over time.

The report concludes by summarizing all relevant findings, highlighting key insights about sales trends, customer behavior, and product performance. Further analysis, along with tailored recommendations for WebTech System, will be provided in the final submission to enhance operational effectiveness and drive business growth.

PROOF OF ORIGINALITY :

- Organisation's Name : WebTech System
- Type of Organisation : B2C (Business-to-Consumer)
- Organisation's Location : LGF-25, Nirman Ambika Arcade, IT Chauraha, Lucknow
- Owner's Name : Mr. Ravi Srivastava
- Owner's Phone Number : +91 9415081584
- Owner's Mail ID : webtechsystemlko@gmail.com
- Number of Employees : 3
- Annual Turnover : Rs. 40,00,000 (approximately)
- Working Time : 11 AM to 7 PM

Photograph 1 : WebTech System Shop (front view)

Photograph 2 : Mr. Ravi Srivastava in WebTech System (inside shop)

- Interview Video : 🎬 VID-20240927-WA0001.mp4
- Some more photographs : 🖼 Photos
- Authorization Letter : 📄 Authorisation Letter.pdf

Photograph 3 : Authorization Letter from organization

META DATA :

Dataset Link : X BDM New.xlsx

Data collection duration : 2 financial years (FY 2022-23, FY 2023-24)

Data collection dates : 01-04-2022 to 24-03-2024

Dataset Dimensions : The dataset contains **13 columns** and **1,158 rows**. Each row corresponds to a single bill or transaction, providing detailed information about each purchase.

Columns :

1. **Date**: The date when the product was sold.
2. **Description of Goods**: A detailed description of the product, including its name and specifications.
3. **Name**: The name of the buyer/customer who purchased the product.
4. **Qty**: The quantity of the product sold in that particular transaction.
5. **Rate With Tax**: The price per unit of the product, including all applicable taxes.
6. **@**: The unit price of the product before taxes are applied.
7. **Total**: The total price for the product, excluding taxes, based on the quantity sold. **i.e. Qty * @**
8. **CGST**: The percentage of Central Goods and Services Tax (CGST) applied.
9. **CGST Amount**: The actual amount charged as CGST in the transaction.
10. **SGST**: The percentage of State Goods and Services Tax (SGST) applied.
11. **SGST Amount**: The actual amount charged as SGST in the transaction.
12. **Total Tax Amount**: The sum of the CGST amount and SGST amount, representing the total tax applied.
13. **With Tax Amount**: The final amount payable by the customer, which includes the total price plus all applicable taxes.

Date	Description of Goods	Name	Qty	Rate With Tax	@	Total	C G S T	CGST Amount	S G S T	SGST Amount	Total Tax Amount	With Tax Amount
2022-04-01	Microtek Inverter	Anuj Srivastava	1	4650.00	3940.68	3940.68	9.00%	354.66	9.00%	354.66	709.32	4650.00
2022-04-01	Exide Battery Toll Tubler	Nishant Varshney	1	14500.00	11328.13	11328.13	14.00%	1585.94	14.00%	1585.94	3171.88	14500.00
2022-04-02	HP Printer Repair	Ajay Kumar Singh	1	800.00	677.97	677.97	9.00%	61.02	9.00%	61.02	122.03	800.00
2022-04-02	Blutooth Spkr	PS Patauna	1	2000.00	1694.92	1694.92	9.00%	152.54	9.00%	152.54	305.08	2000.00
2022-04-03	Inverter Exide	Anil Singh	2	11500.00	9745.76	19491.53	9.00%	1754.24	9.00%	1754.24	3508.47	23000.00
2022-04-03	Blutooth Spkr	Vinay Kumar	1	2000.00	1694.92	1694.92	9.00%	152.54	9.00%	152.54	305.08	2000.00
2022-04-04	Dlink CAT 6 Cable 1 bun	Anjali Saxena	1	2850.00	2415.25	2415.25	9.00%	217.37	9.00%	217.37	434.75	2850.00
2022-04-04	Intex Blutooth Mouse	Anjali Saxena	1	270.00	228.81	228.81	9.00%	20.59	9.00%	20.59	41.19	270.00
2022-04-04	CP+ Cam Indoor	Shree Jee	2	1400.00	1186.44	2372.88	9.00%	213.56	9.00%	213.56	427.12	2800.00
2022-04-04	CP + Cam Outdoor	Shree Jee	1	1600.00	1355.93	1355.93	9.00%	122.03	9.00%	122.03	244.07	1600.00
2022-04-04	DVR -4CH CP+	Shree Jee	1	2500.00	2118.64	2118.64	9.00%	190.68	9.00%	190.68	381.36	2500.00
2022-04-04	Supply 4CH Dlink	Shree Jee	1	600.00	508.47	508.47	9.00%	45.76	9.00%	45.76	91.53	600.00
2022-04-04	HDD 1 TB	Shree Jee	1	3500.00	2966.10	2966.10	9.00%	266.95	9.00%	266.95	533.90	3500.00
2022-04-04	Cable	Shree Jee	1	1600.00	1355.93	1355.93	9.00%	122.03	9.00%	122.03	244.07	1600.00
2022-04-04	Foxin Key Board	Ajay Mishra	1	325.00	275.42	275.42	9.00%	24.79	9.00%	24.79	49.58	325.00
2022-04-05	Blutooth Spkr	Manoj	1	2000.00	1694.92	1694.92	9.00%	152.54	9.00%	152.54	305.08	2000.00
2022-04-06	Exde imtt1500	Ganesh	1	14200.00	11093.75	11093.75	14.00%	1553.13	14.00%	1553.13	3106.25	14200.00
2022-04-06	V-Gard Stablizer 4KVA	Neeraj Verma	1	3850.00	3262.71	3262.71	9.00%	293.64	9.00%	293.64	587.29	3850.00
2022-04-06	Hicvison DVR-4CH	Naveen Srivastava	1	3350.00	2838.98	2838.98	9.00%	255.51	9.00%	255.51	511.02	3350.00
2022-04-07	Micritek Inverter 1050/12v	Nishant Varshney	1	5200.00	4406.78	4406.78	9.00%	396.61	9.00%	396.61	793.22	5200.00

Table 1 : A snapshot of how the data looks like (top 20 entries)

DESCRIPTIVE STATISTICS :

During the data cleaning process, it was discovered that out of 49 unique products, only 29 were sold more frequently. The remaining 20 products had just 1 or 2 sales over the two-year period. Upon further discussion with Mr. Srivastava, it was revealed that these 20 products are not regularly sold by WebTech System but are made available on special demand for certain customers. According to him, these are the least-selling products and take up space that could be better utilized for other SKUs. As a result, we have separated the data for these 20 products and proceeded with the analysis focusing on the descriptive statistics for the remaining 29 products

From the descriptive analysis of the numerical values, it was observed that most products sold are in single units, as indicated by a mean of 1.15. However, the maximum value of 4 suggests that there are instances where multiple products are sold together, which warrants further investigation.

The mean value of the products sold is Rs 6,793.05, which is relatively high. Another notable observation is the wide range in the "Rate with Tax" column, spanning from Rs 250 to Rs 17,500. This range is well distributed, as evidenced by the standard deviation of Rs 5,484.77 and the 50th percentile sitting at Rs 4,700. This suggests that the majority of products sold by WebTech System fall on the higher price end.

```
[12]    1 filtered_df.describe()
```

	Qty	Rate With Tax	@	Total	CGST	CGST Amount	SGST	SGST Amount	Total Tax Amount	With Tax Amount
count	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000	1137.000000
mean	1.153914	6793.051891	5526.456614	6483.410709	0.102841	750.884795	0.102841	750.884795	1501.769590	7985.180299
std	0.551937	5484.778702	4355.968183	6218.062698	0.021853	831.802803	0.021853	831.802803	1663.605607	7840.954831
min	1.000000	250.000000	211.864407	211.864407	0.090000	19.067797	0.090000	19.067797	38.135593	250.000000
25%	1.000000	2000.000000	1694.915254	1694.915254	0.090000	152.542373	0.090000	152.542373	305.084746	2000.000000
50%	1.000000	4700.000000	3983.050847	3983.050847	0.090000	358.474576	0.090000	358.474576	716.949153	4700.000000
75%	1.000000	11500.000000	9745.762712	10898.437500	0.140000	1421.875000	0.140000	1421.875000	2843.750000	13950.000000
max	4.000000	17500.000000	14830.508475	28906.250000	0.140000	4046.875000	0.140000	4046.875000	8093.750000	37000.000000

Table 2 : A snapshot of descriptive statistics of numerical columns

The average total amount paid by customers (mean bill value) is Rs 7,985.18, with a range from Rs 250 to Rs 37,000. Tax rates seem to vary between 18% and 28%, split between CGST and SGST, with each tax component ranging from 9% to 14%.

In the descriptive analysis of the "Description of Goods" and "Name" columns, it was found that Bluetooth Speakers are the best-selling product at WebTech System, with 120 units sold over a two-year period. On the customer side, Abhishek Chauhan emerged as a key buyer, having made 54 purchases. This raises the question of whether Abhishek is a regular customer or perhaps a contractor or business purchasing products in bulk to resell. Further analysis is needed to understand his buying patterns and the types of products he is purchasing.

```
1 filtered_df[['Description of Goods', 'Name']].describe()
```

	Description of Goods	Name
count	1137	1137
unique	29	67
top	Blutooth Spkr	Abhishek Chauhan
freq	120	54

Table 3 : A snapshot of descriptive statistics of Products and Customers

ANALYSIS METHODS :

DATA CLEANING & PREPROCESSING :

Process :

The first step in the analysis was cleaning and preprocessing the raw data to ensure accuracy and consistency. This involved handling missing values, correcting data types, and removing or separating outliers. For instance, the data received from Mr. Srivastava came in the form of individual Excel sheets per bill. To make the data suitable for analysis, it was preprocessed and consolidated into a single sheet, removing irrelevant details such as the shop name, address, and contact information, which were part of the basic bill format. Also, after discussing with Mr. Srivastava, products with very few sales (1 or 2) over the span of 2 years were separated to focus the analysis on key products that were sold regularly.

Justification : Data cleaning ensures that the analysis is conducted on reliable and meaningful data. By removing inconsistencies and outliers, this step makes it easier to identify patterns and trends accurately. In this case, focusing on the core products that are sold regularly prevents skewing the analysis toward products that do not reflect the business's typical operations. Additionally, consolidating the data into a single sheet streamlines the analysis process, allowing for more effective exploration of sales trends and customer behavior.

DESCRIPTIVE STATISTICS :

Process : Descriptive statistics were used to summarize and understand key characteristics of the data. We looked at the mean, standard deviation, and distribution of important variables such as product prices, quantities sold, and total sales amounts.

Justification : This method provides a high-level overview of the dataset, helping to identify central tendencies (mean) and variations (standard deviation). It highlights key trends, such as the fact that most products are sold individually but some are sold in bulk, and that WebTech deals in high-priced products, as shown by the mean sales price of Rs 6793.05.

DISTRIBUTION ANALYSIS :

Process: A frequency distribution analysis was conducted on categorical columns like 'Description of Goods' and 'Name.' This helped identify the most popular products and most frequent customers.

Justification: This analysis allows for a straightforward understanding of which products are in high demand and which customers are contributing the most to sales. For example, the identification of Bluetooth Speakers as the most sold item, and Abhishek Chauhan as a frequent buyer, provides direction for targeted marketing or inventory strategies.

SALES ANALYSIS :

Process: Using the preprocessed data, a line graph will be plotted to analyze sales trends and identify seasonal patterns throughout the year. From distribution analysis it was found that Bluetooth speakers were the most sold product, their unit price is low compared to the average bill amount, suggesting a focus on volume rather than revenue. The analysis will therefore prioritize high-value products that contribute more significantly to revenue, examining their sales trends to detect patterns. This will help segment products based on their impact on revenue and provide insights for further analysis.

Justification: Analyzing sales trends and revenue contributions helps the business understand demand fluctuations and identify peak sales periods. By focusing on high-revenue products, WebTech System can align inventory and marketing strategies more effectively, ensuring stock availability during high-demand periods and reducing overstocking. This approach helps prioritize products that drive profitability, allowing the business to optimize resources and develop targeted strategies to promote key products that have the most significant financial impact.

RESULTS AND FINDINGS :

RESULT 1 : Top 15 Purchased Products

This graph highlights the most popular products sold by WebTech. **Bluetooth Speakers** stand out significantly as the highest-selling product with 120 units, followed by **Hicvision DVR-4CH** at 70 units. This large gap suggests that Bluetooth Speakers are a major product for WebTech. However, all top 3 selling products, including **CP+ Cam Indoor**, have a low unit price compared to the average bill value. This suggests that while these products are frequently sold, they do not contribute significantly to the overall revenue. Further analysis is required for high-value, low-volume products.
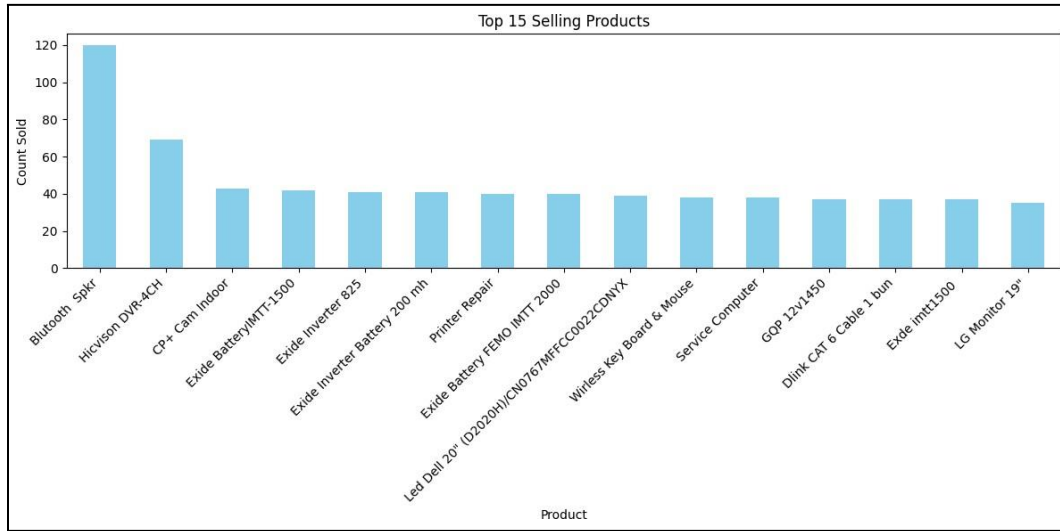
Chart 1 : Top 15 selling products at WebTech System

RESULT 2 : Top 15 Customers

There was no significant difference between the top customers. **Abhishek Chauhan** appears frequently, but the distribution of customer purchases does not suggest that he is a contractor buying in bulk for resale purposes. Rather, he is likely a loyal customer.
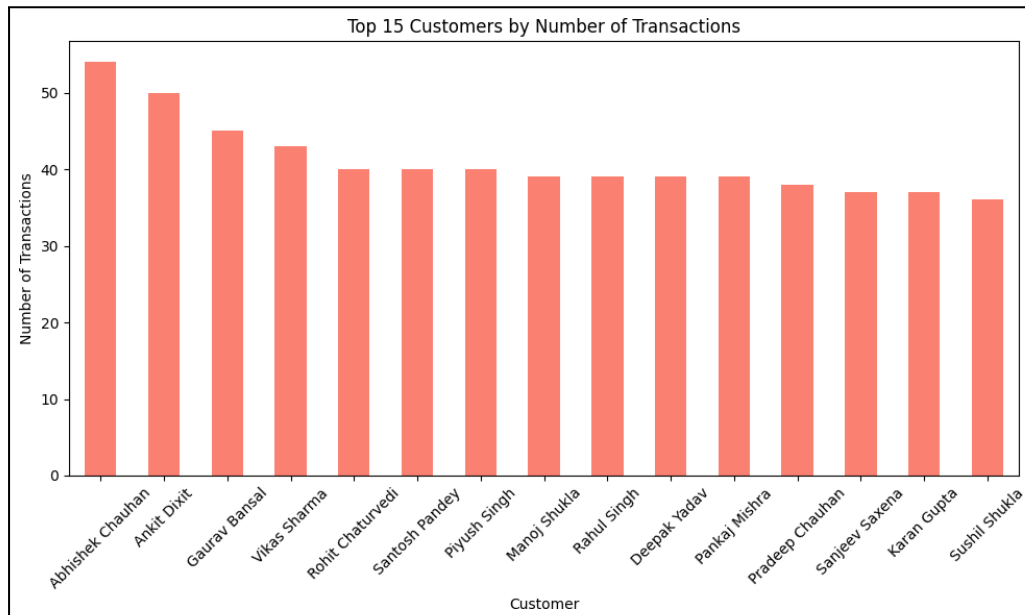

Chart 2 : Top 15 customers at WebTech System

RESULT 3 : Trend Analysis

The highest sales occurred in **February 2023**, surpassing ₹550,000, likely driven by a seasonal surge or promotional event. Another significant rise in sales is observed in **December 2023**, which could be attributed to the holiday season. Conversely, there are notable drops in **October 2022**, **July 2023**, and **February 2024**, with sales dipping to around ₹250,000–₹300,000, suggesting periods of lower demand. Despite these fluctuations, the overall average sales remain steady at around ₹400,000, indicating consistent performance during regular months. These insights suggest the need for strategic planning, including inventory optimization and targeted promotions, to capitalize on peak periods and stabilize sales during low-demand months.
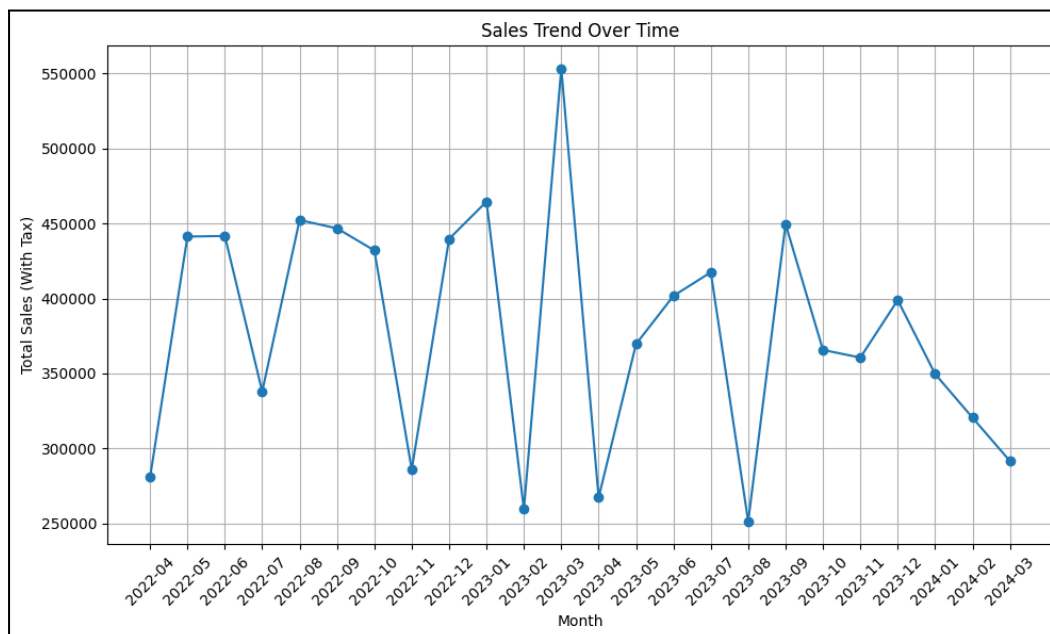
Chart 3 : Sales trend from Apr-22 to Mar-24

All the results and findings included in the report are derived from the initial analysis. A more in-depth examination will be conducted on specific issues to gain deeper insights. The final results, along with appropriate recommendations, will be provided in the concluding submission.