

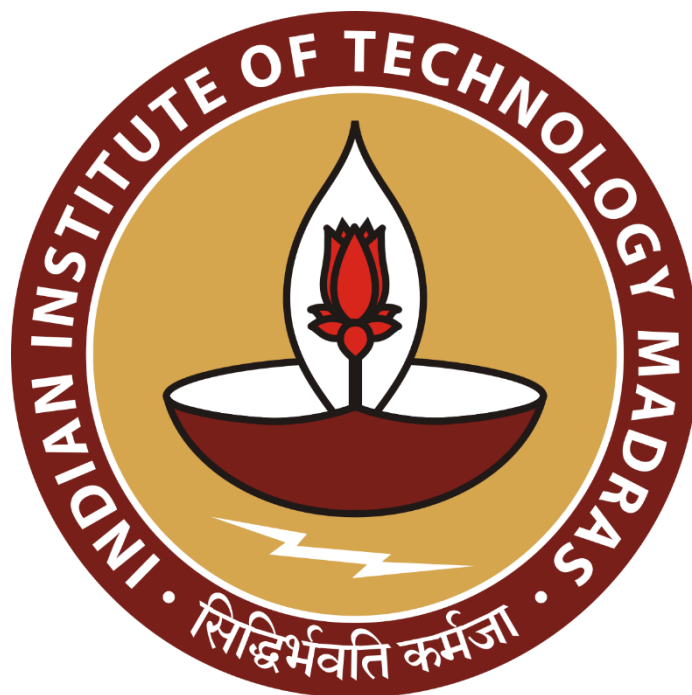
# Optimizing Sales Performance of an E-Commerce Company

Final report for the BDM capstone Project

Submitted by

Name: Lalit Kumar

Roll number: 21f3003123



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

## Contents

1	Executive Summary and Title.....	2
2	Proof of Originality.....	2
3	Meta data and descriptive statistics .....	2
4	Detailed explanation of analysis process/ method.....	5
	4.1 Data Collection, Cleaning and Preparation.....	4
	4.2 Sales performance analysis.....	5
	4.3 Target vs. actual sales comparison.....	10
	4.4 Customer behavior analysis.....	11
	4.5 Region wise sales analysis.....	14
5	Results and findings .....	15
6	Interpretation of results and recommendation.....	16

# 1 Executive Summary and Title

The organization which is a retail business operates mainly in India. It offers a broad ranges of products across multiple categories. Despite its online presence and market position company is facing challenges in achieving sales targets in certain categories which impacts overall revenue and profitability of the company.

We used a data driven approach to analyse and address the company's challenges. I started by comparing sales performance and found clothing as most profitable and highest performing category whereas company finds it difficult to generate profits from Furniture which is most underperforming category. Despite having multiple products in this category it contributes only 9.6% to company's profit. Next, we compared targets with actual sales and found it satisfactory.

Looking at region wise sales Madhya Pradesh and Maharashtra were identified as top performing states which implies company's strong market presence in the region. The regions with lower sales has the potential for company's future expansion. The results and recommendations given in the report aligns with the business issues faced by the company and it will help company to improve overall performance and profits.

## 2 Proof of Originality

The data that we are using for this project is derived from a publicly available dataset hosted on Kaggle. It contains one years of comprehensive data related to product sales organized in categories and sub categories, sales targets, and order details.

The dataset is sourced from the following link:

**Data Source:** <https://www.kaggle.com/datasets/benroshan/ecommerce-data/data>

## 3 Meta data and descriptive statistics

The dataset includes the following files in CSV format.

### 3.1 Meta Data

List of Orders.csv

Column	Data Type	Description
Order ID	String	A unique id assigned to each order
Order Date	Date	The date when the order was placed by the customer
CustomerName	String	The name of the customer who placed the order
State	String	The state from which order is placed
City	String	The city inside the state where the order is placed

*Table 1: Meta data list of orders*

#### Order Details.csv

Column	Data Type	Description
Order ID	String	A unique id for each order, used for cross-referencing with other table i.e. List of Orders.csv
Amount	Float	Price of the product
Profit	Float	The profit made from the order
Quantity	Integer	The number of units of products ordered
Category	String	The high-level classification of the product e.g. Electronics, Furniture
Sub-Category	String	A more detailed classification of the product within a category (e.g., Laptops under Electronics).

Table 2: Meta data order details

#### Sales target.csv

Column	Data Type	Description
Month of Order Date	String	Represents the month and year of the order date (e.g., "July-2024").
Category	String	The category of products for which sales target are set
Target	Float	The sales target amount for the given category and month.

Table 3: Meta data sales target

## 3.2 Descriptive Statistics

#### List of Orders.csv

	Order ID	Order Date	Customer Name	State	City
count	500	500	500	500	500
unique	500	307	332	19	24
top	B-25601	24-11-2018	Shreya	Madhya Pradesh	Indore
freq	1	7	6	101	76

Table 4: Descriptive statistics list of orders

Table 4 provides a summary of the order details which contains 500 entries. Order Date contains 307 unique values with most frequent order date was Nov 24, 2018. Shreya is the most frequent customer name that appears 6 times. There are 19 states with Madhya Pradesh being the top state for orders, similarly in 24 cities Indore is the city from which orders were placed most.

Order Details.csv (for numerical data)

	Count	Mean	Standard Deviation	Min	Max
Amount	1500	287.6680	461.0505	4.0	5729.0
Profit	1500	15.9700	169.1406	1981.0	1698.0
Quantity	1500	3.7433	2.1849	1.0	14.0

*Table 5: Descriptive statistics order details for numerical data*

In the above table Amount column contains the value at which each order was placed. The average order amount is 287.67 and maximum is 5729. The mean profit for orders is 15.97 with Standard deviation of 169.14. The average order quantity is 3.74 with minimum 1 which is going up to 14 items per order.

Order Details.csv (for categorical data)

	Order ID	Category	Sub-Category
count	1500	1500	1500
unique	500	3	17
top	B-25656	Clothing	Saree
freq	12	949	210

*Table 6: Descriptive statistics order details (categorical data)*

The order details table contain 1500 entries with clothing is the most frequent category, order placed for this category was 949 times. Similarly, Saree is the most frequent ordered product.

Sales target.csv (for numerical data)

	Count	Mean	Standard Deviation	Min	Max
Target	36	12108.3333	2667.8375	9000.0	16000.0

*Table 7: Descriptive statistics for sale target (numerical data)*

The table 7 contains the details for sales target's Target column. The mean target is 12180.33 the target value range from 9000 to 16000.

Sales target.csv (for categorical data)

	Month of Order Date	Category
count	36	36
unique	12	3
top	Apr-18	Furniture
freq	3	12

*Table 8: Descriptive statistics for sales target (categorical data)*

Sales targets were assigned monthly for each category, resulting in a total of 36 entries. April 2018 was the month with the highest frequency of sales targets.

## 4 Detailed explanation of analysis process/ method

To perform data analysis and address the organization's problem we used different tool and techniques primarily I used Python and Jupyter Notebook. We start with exploratory data analysis in notebook where I plotted graphs and used tables to understand the patterns in data. We did sales performance analysis to see trends, seasonality and find any outliers. We analysed different categories and sub categories and their profit contribution to company. Plotted chart to compare the set targets and actual sales over the year. Also I studied how customers are buying products, identified profitable and most frequent purchase group of customers for targeted marketing. We did region wise sales analysis to understand the places where company is performing well and where it can expand in future.

### 4.1 Data Collection, Cleaning and Preparation:

Our first steps was to clean the collected data and prepare it for data analysis. We began by loading all the files in Pandas Data Frames and checked data for null values and other inconsistencies, found List of Orders.csv contains some null values. We took the necessary steps and removed rows that contains only null values.

The 'Order Date' column of List of Orders.csv was originally in String format, we have converted it to Python date format to perform accurate date based analysis. Whenever required we combined the data based on Order ID or Order Date to facilitate further analysis. We have added a few columns such as Total Sales as it required for sales performance analysis.

We have calculated total sales or revenue by using the given formula.

$$TotalSales = Amount \times Quantity$$

### 2. Sales Performance Analysis:

The sales performance analysis of the company shows key insights. There is high fluctuation in total sales over the months in year 2018 and 2019, Jan 2019 is the high performing month whereas July 2018 is lowest selling month.

Increase in sales in December and Jan could suggest high demand possibly due to holidays and seasonal promotions. The month of July could suggest a seasonal decline. We can further investigate

to find the reason for dip during Jun July other possible reason could be inventory shortage or issues in business operations.

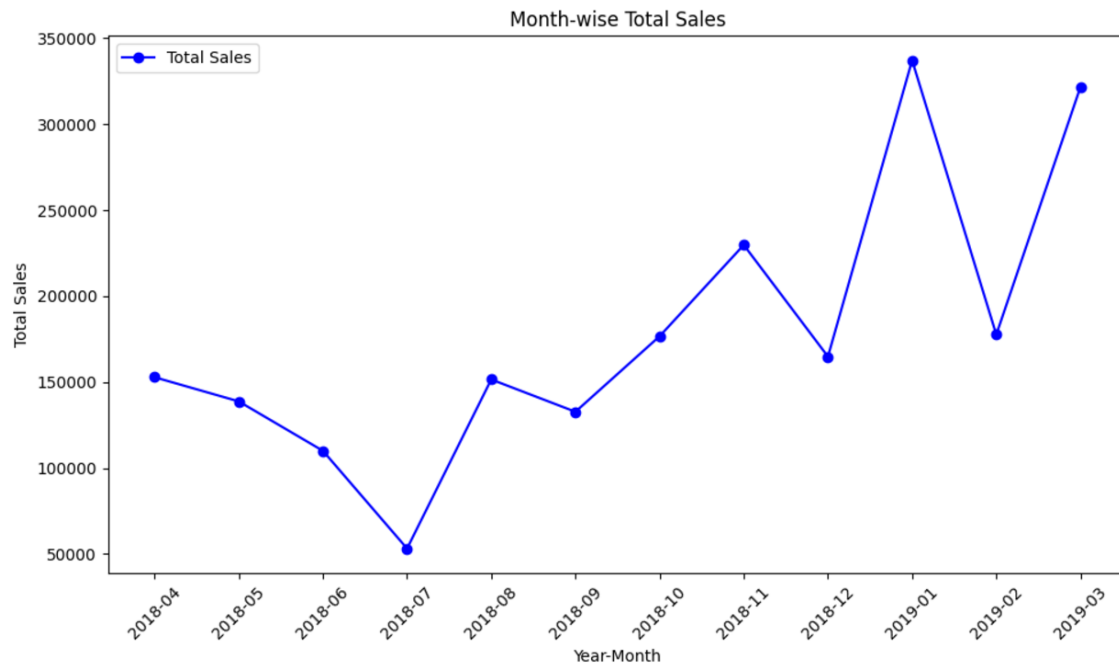


Figure 1: Month-wise Sales Trend

When looking at monthly sales by category electronics is the highest selling category among all in the month of Jan 2019 and Furniture is highest selling category in the month of March 2019. Again July is the lowest selling month for all the categories. Clothing is contributing to overall sales during the peak sales months.

The figure 2 shows sales trend of different categories over the months. Clothing shows relatively stable sales throughout the year. Electronics and furniture exhibits a sharp rise during the months of Dec and Jan. Sales for Furniture is consistently increasing and makes highest sale in March.

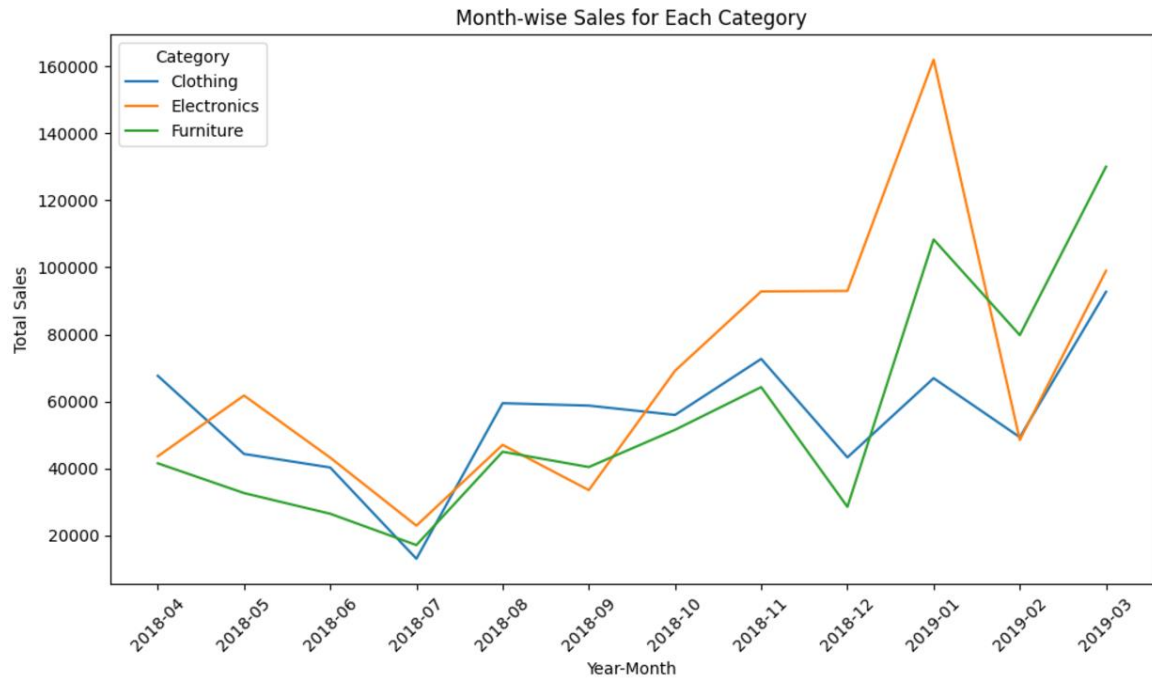


Figure 2: Month-wise sales trend for each category

The monthly sales trend for top 3 subcategories with highest and lowest total sales are given in figure 3 and figure 4 respectively. Printers in Electronics, Bookcases in Furniture and Saree in Clothing are highest performing subcategories. The sale of Printer and Bookcases was highest in the month Jan.

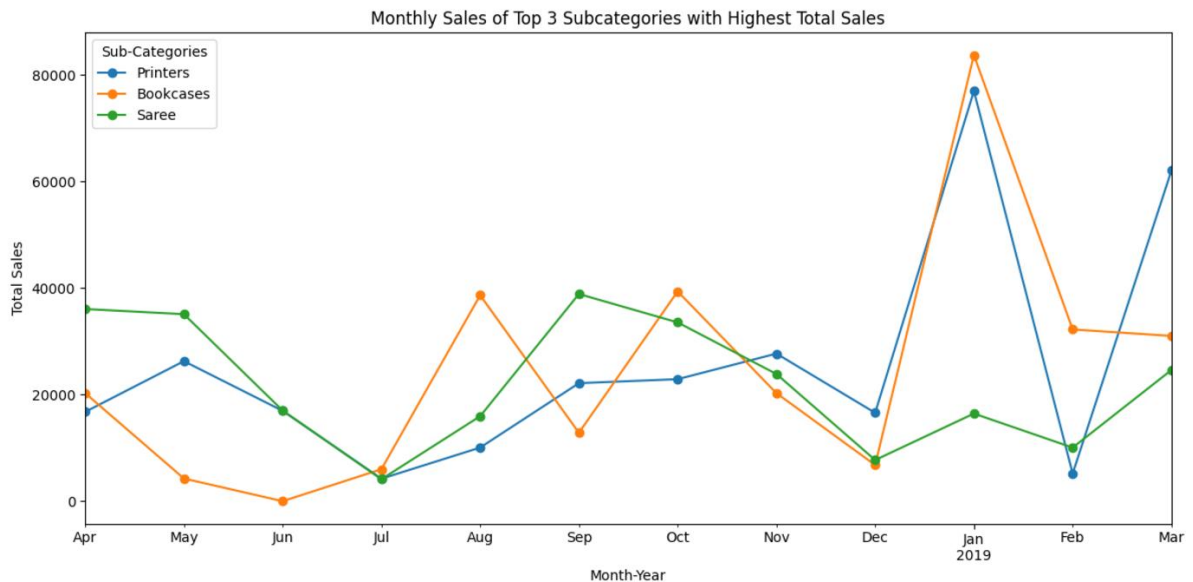


Figure 3: Top 3 subcategories with highest total sales

Similarly the lowest performing sub categories are Leggings, skirt and kurti all of them belongs to the same category i.e. Clothing. Furnishing in furniture and accessories in Electronics are lowest performing subcategories.



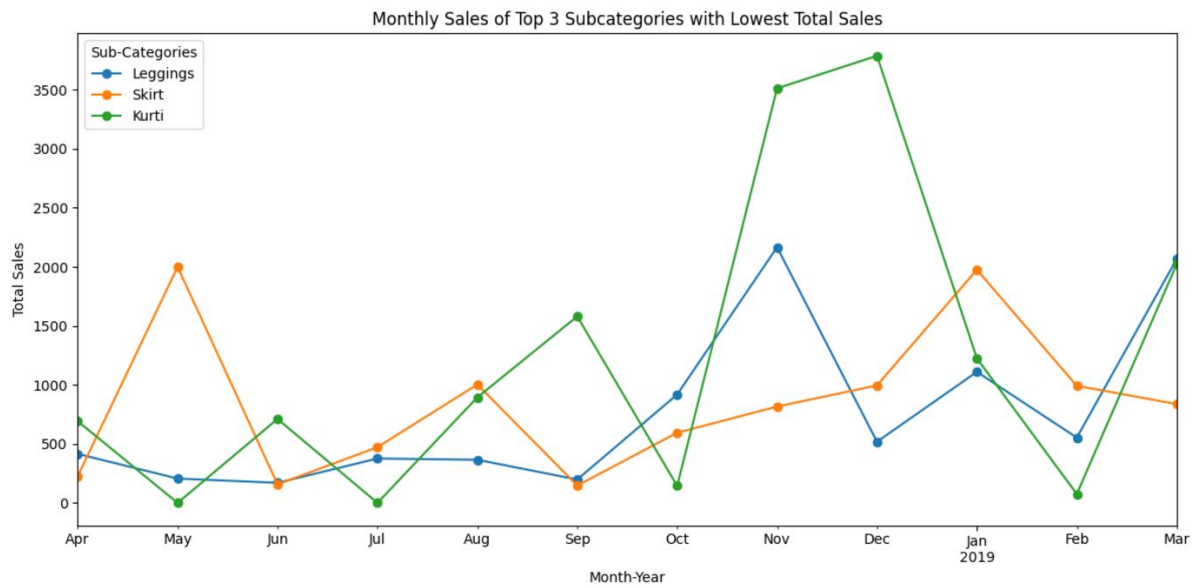


Figure 4: Top 3 subcategories with lowest total sales

As we can see in figure 5 the company was actually making loss before Oct 2018 but did well in later months. In March Clothing was most profitable category similarly in the month of January Electronics was highest profit making category. Before Nov Furniture was loss making category and in later months its profit declining once again.

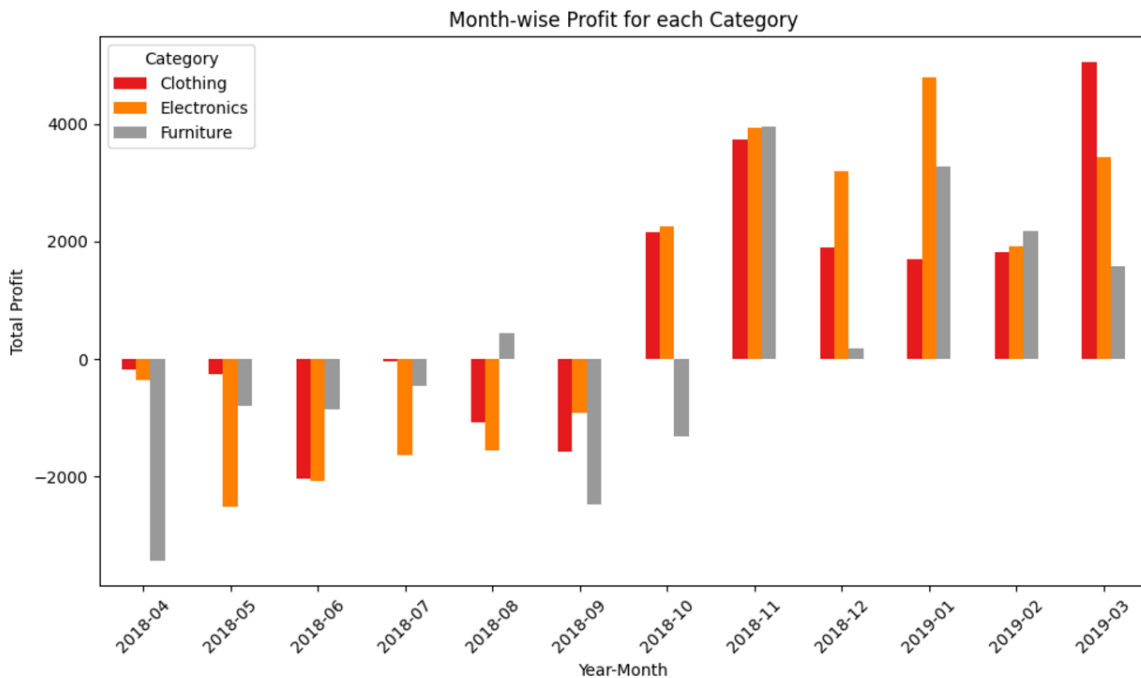


Figure 5: Month-wise profit for each Category

The overall profit for each category is given below. Clothing is most profitable category whereas Furniture is lowest.

Category	Profit
Clothing	11163.0
Electronics	10494.0
Furniture	2298.0

Table 9: Categorywise profit

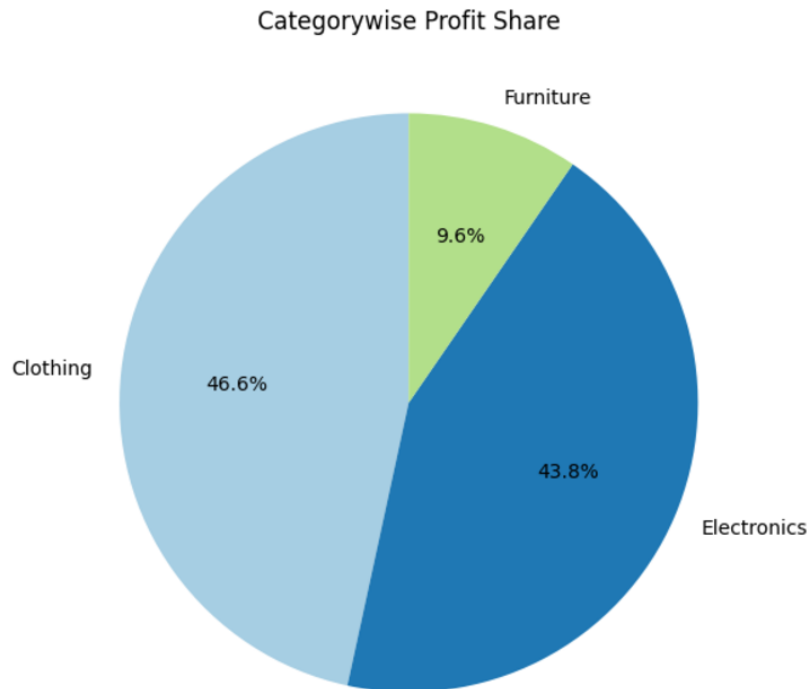


Figure 6: Pie chart - profit share by category

Overall Profit & loss for each subcategory is given in the below table. The subcategories are organized in the table by descending order of their profit. As we can see here Printers is highest profit making subcategory whereas Tables is lowest.

Sub Category	Profit
Printers	5964.0
Bookcases	4888.0
Accessories	3559.0
Trousers	2847.0
Stole	2559.0
Phones	2207.0
Hankerchief	2098.0
T-shirt	1500.0
Shirt	1131.0
Furnishings	844.0
Chairs	577.0
Saree	352.0
Leggings	260.0
Skirt	235.0
Kurti	181.0
Electronic Games	-1236.0

Table 10: Sub category wise profit

### 3. Target vs. Actual Sales Performance Comparison:

The target vs. actual sales comparison highlights areas where sales targets are not being met. In figure 8 which shows the actual sales vs. target for clothing we can see that target is met in every month except July 2018.

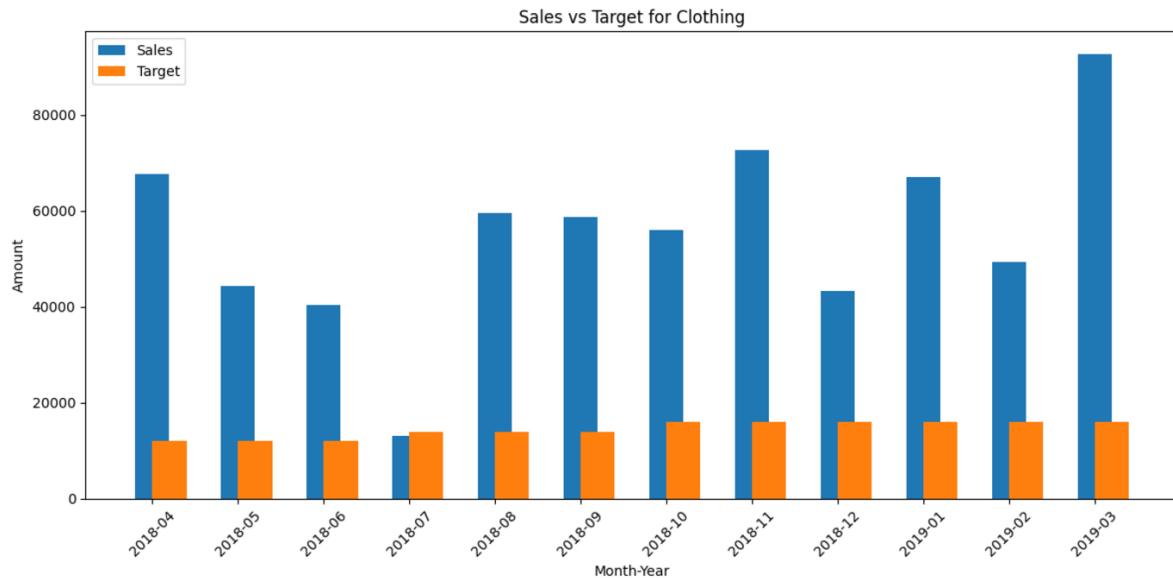


Figure 7: Sales Vs. Target for Clothing

Similarly, the actual sale vs. target for Electronics is given in figure 8. Here target is met in every month. The number of products sold in this category were above the expectation and set target.

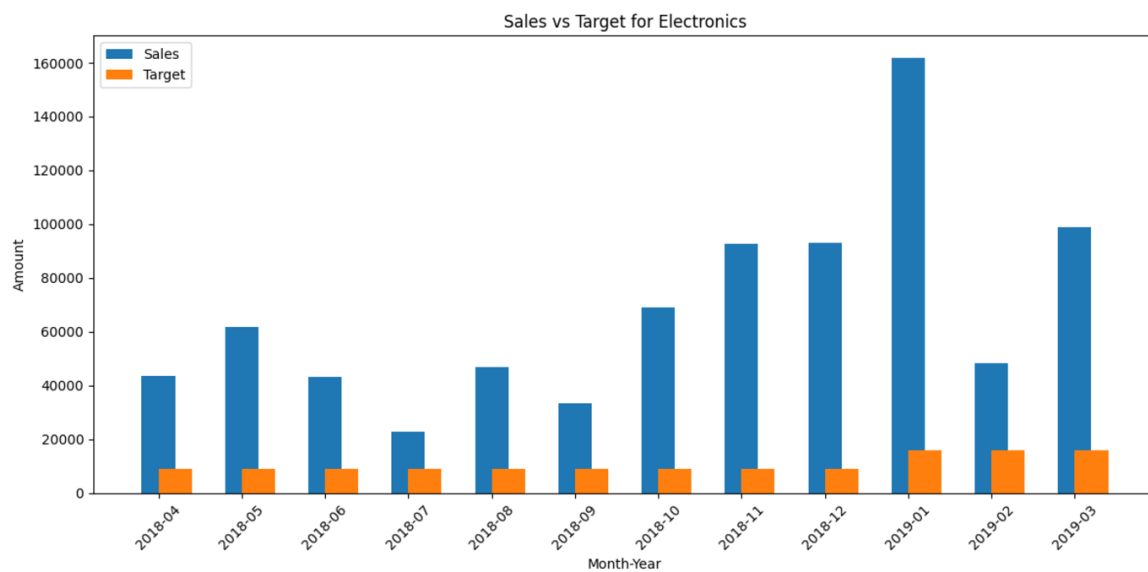


Figure 8: Actual Sales Vs. Target for Electronics

Figure 9 shows the sales vs. target for furniture here also target is met in every month.

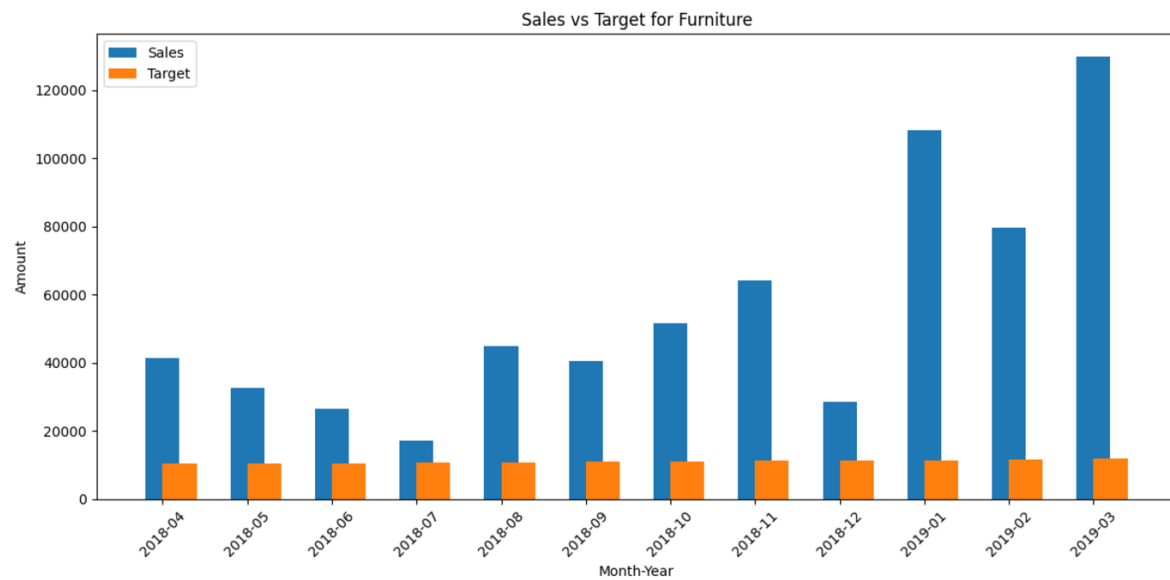


Figure 9: Actual Sales Vs. Target for Furniture

#### 4. Customer Behaviour Analysis:

We can optimize the company's product offering by understanding the customer's purchase behaviour. Also we can personalize the offers and discounts on products or categories for each customer. We calculated purchase frequency of unique products and the top 10 customer with their purchase frequency is given in the figure 10. Shreya has bought products 6 times from the company and the Sheetal 4 times from the company.

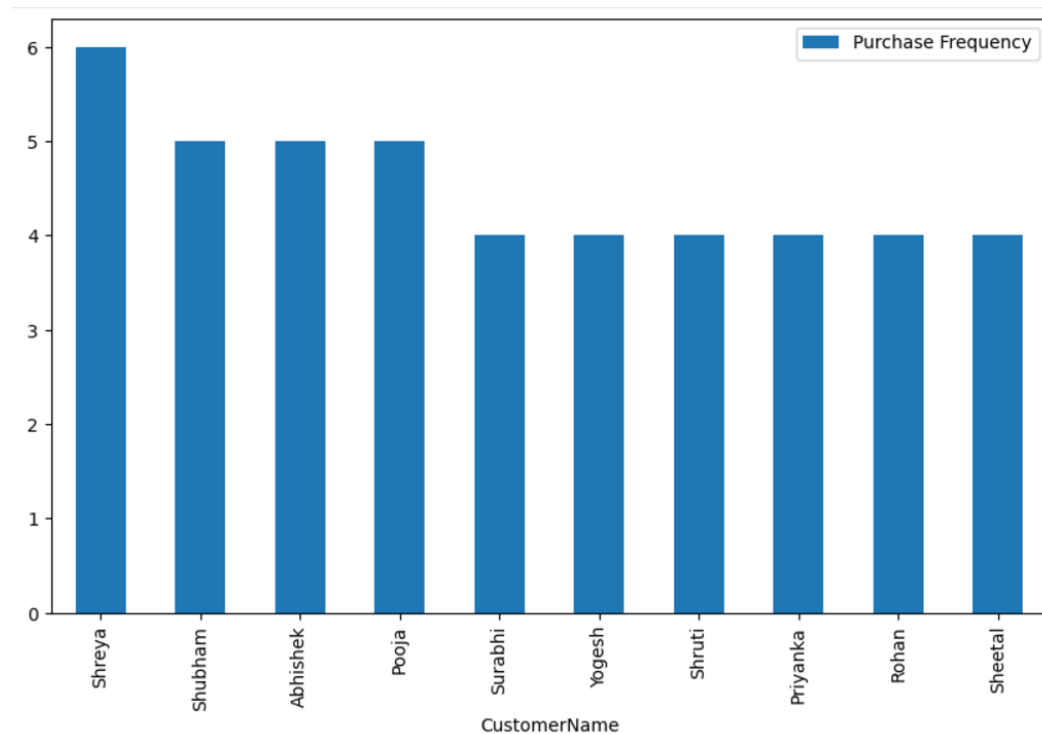


Figure 10: Top 10 customer by purchase frequency

Figure 11 shows the customers with highest average order value, Seema is the top customer who is buying products with average order value of 12844.

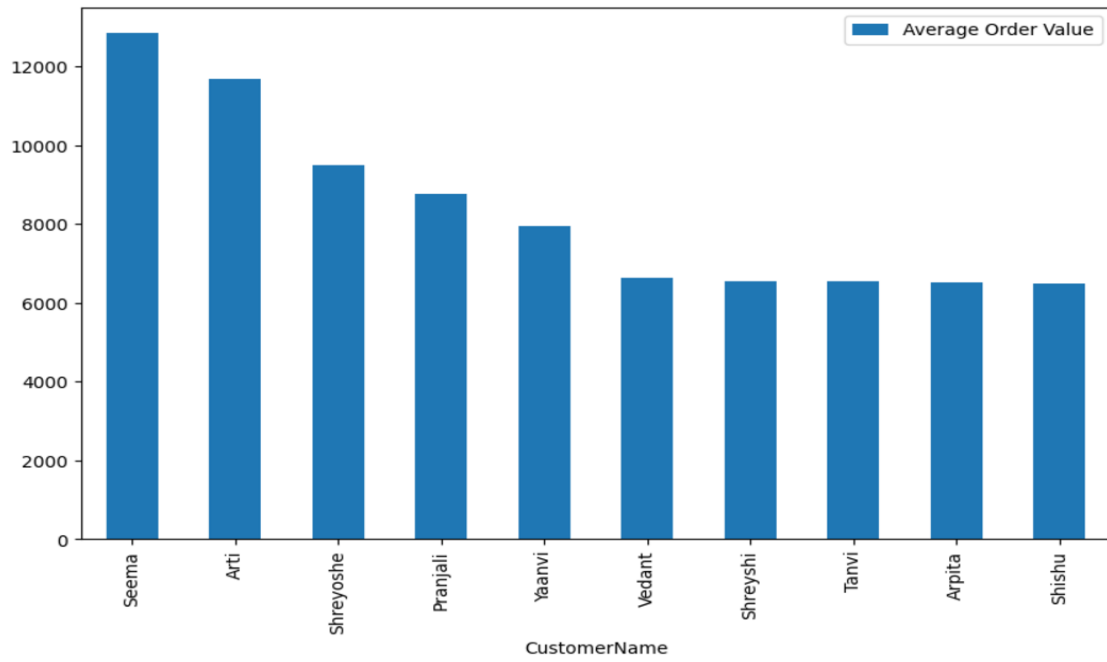


Figure 11: Top customers by average order value

Company got highest profits from customers that are given in the figure 12.

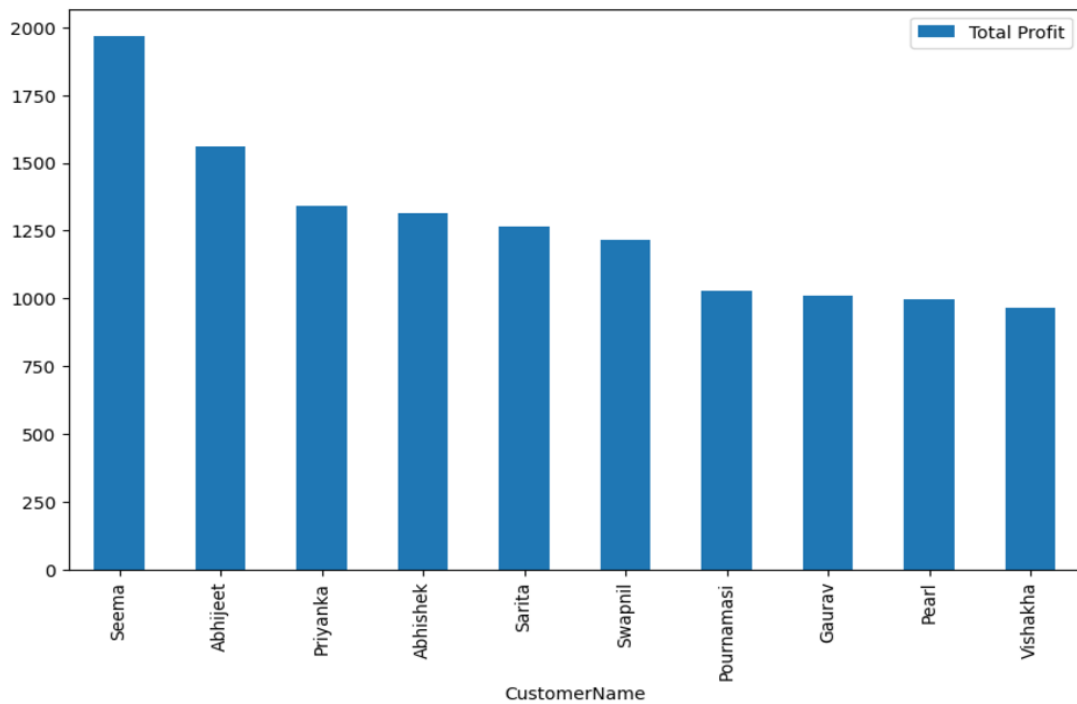


Figure 12: Top 10 customers contributing to profit

#### 4. Region wise Sales Analysis:

The largest revenue is coming from Madhya Pradesh followed by Maharashtra whereas lowest revenue is coming from Sikkim and Goa. Figure 13 show top 10 states by total sales.

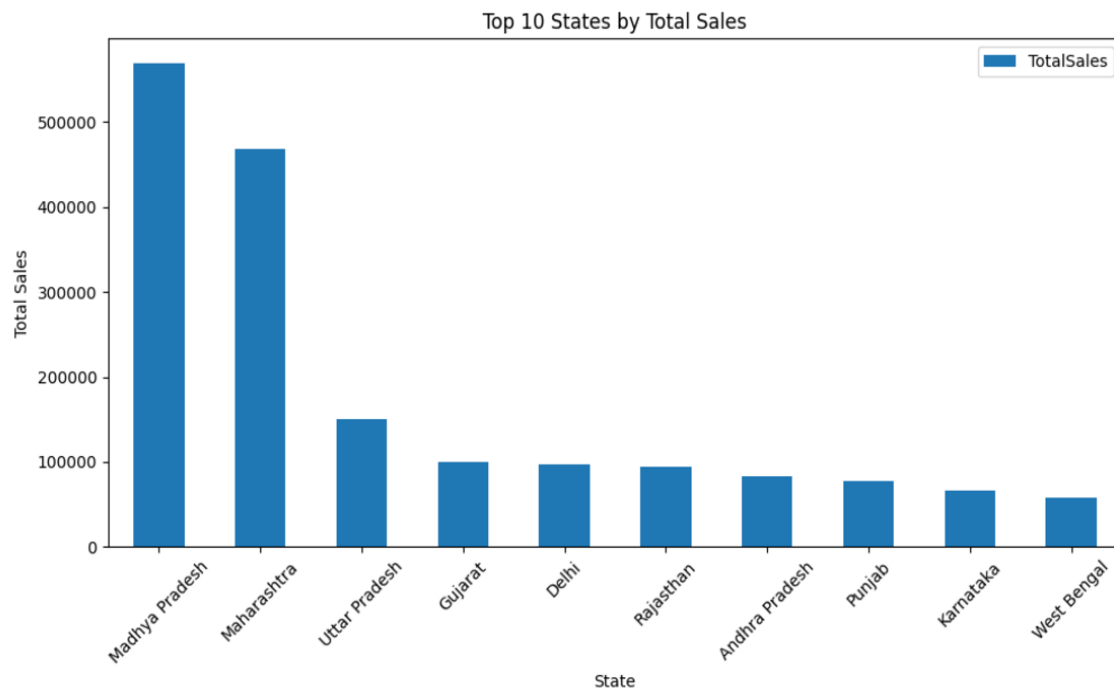


Figure 13: Top 10 States by Total Sales

Indore is the largest revenue making city whereas Amritsar was lowest revenue making city. Figure 14 show top 10 cities by total sales.

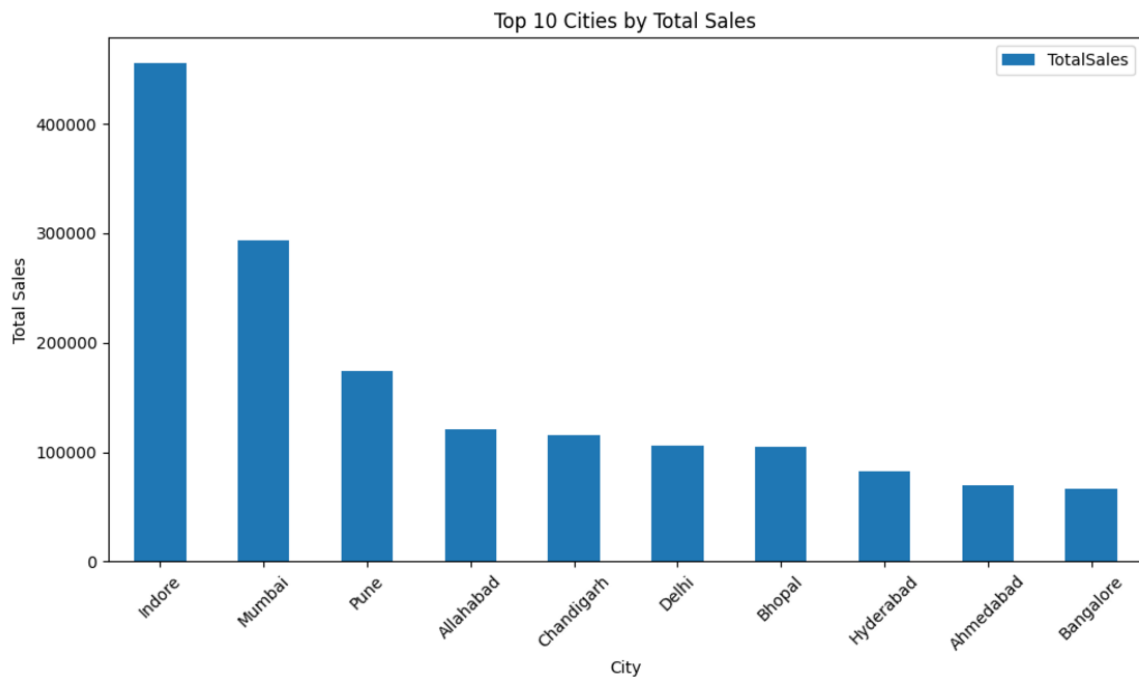


Figure 14: Top 10 cities by total sales

## 5 Results and findings

From the above analysis we found the given results –

- By total sales Furniture is most underperforming category whereas Clothing is highest performing category.
- Similarly Leggings, Skirt and Kurti are most underperforming subcategories whereas Printers, Bookcases are top performing sub-categories.
- Company is making 46.6% profit from Clothing, 43.8% from Electronics and 9.6% from Furniture.
- The target vs. actual sales comparison analysis shows company is meeting targets in every month except July 2018 for clothing
- The company is making highest revenue from Madhya Pradesh whereas lowest from Sikkim and Goa.
- Similarly in cities company is making highest revenue from Indore whereas lowest from Amritsar
- Seema is customer with highest average order value and contributing companies profit whereas Shreya is most frequent buying customer



## 6 Interpretation of results and recommendation

Based on the above analysis and results we recommends –

- Clothing is highest performing category in both sales and profit indicates high demand and efficient sales strategies were used in this segment. The possible cause for Furniture category's underperformance is lack of demand, pricing or quality company need to revisit these things and improve its sales strategies for Furniture and subcategories like Legging, Skirts and Kurtis.
- Again the profit share of 9.6% of furniture shows inefficiency or low sales volume. We recommend company to review its sales strategy for Furniture.
- The exception of July for not meeting target in clothing category is possibly due to seasonality, low demands or high market competition
- In Madhya Pradesh and Maharashtra company is making highest revenue shows the strong market presence. And low sales in Sikkim and Goa indicate less competitive and untapped market company can expand here.
- The high average order value from customer like Seema suggest premium or bulk purchases company can offer exclusive and premium deals to such customers to maintain their engagement. Similarly company can provide personalized offers to customers like Shreya to encourage more spending.