

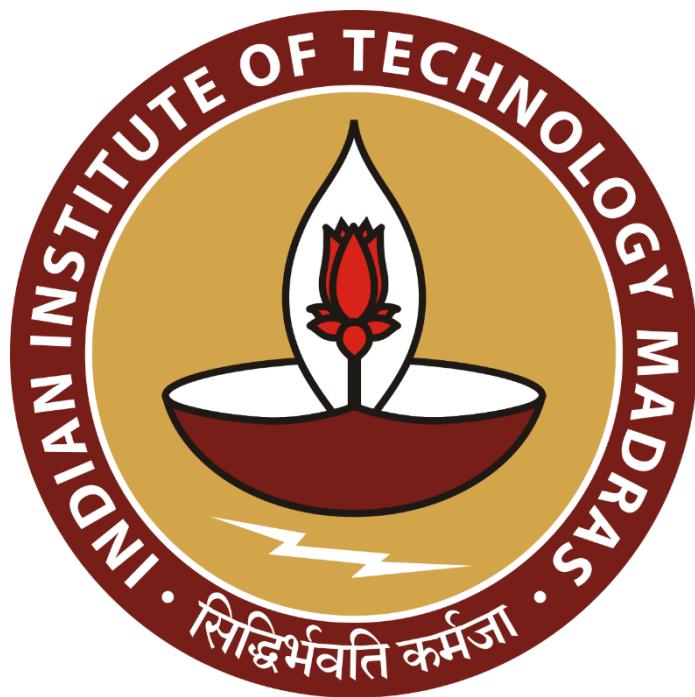
Title : Navigating Challenges: A Data-Driven Boost for 99 Market Mall

A Mid-term report for the BDM capstone Project

Submitted by

Name: RAJEEV KUMAR BHAGAT

Roll number: 21f3001527



IITM BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 6000361

Contents

1. Executive Summary	2
2. Proof of Originality of Data	2-3
● 2.1 Photo Proof	2
● 2.2 More Photo Proof	2
● 2.3 Video Proof	2
● 2.3 Authorization Letter	3
3. MetaData	4-6
● 3.1 Dataset Sheet Link	4
● 3.2 Test Run Results (Colab Link)	4
● 3.3 Stock and Sales Summary	4
● 3.4 Stock Organized by Categories	5
● 3.5 Month Wise Stock Data Amount	6
4. Descriptive Statistics	7-9
● 4.1 Descriptive Statistics for Monthly Stock	6
● 4.2 Descriptive Statistics for purchase and sales rate of each product	8
● 4.3 Descriptive Statistics for Sales and stock by Amount of each category	8
● 4.4 Descriptive Statistics of Sales by Amount of three months	9
5. Detailed Explanation of Analytical Process	9
● 5.1 Product Sale and Inventory Analysis	9
● 5.2 Customer Preference Analysis for Products	9
● 5.3 Monthly Sales Data Compilation for Products	9
● 5.4 Monthly Inventory Analysis for Products	9
6. Results and Findings	10-11
● 6.1 Pie Chart Showing Variety	10
● 6.2 Line Chart depicting Total Monthly Sales and Total Monthly Stock	10
● 6.3 Stock Amount and Sale Amount Comparison Chart	11
● 6.4 Pareto chart	11

1 Executive Summary

The challenge for 99 Market Mall, located in Maharajganj, Siwan, Bihar, is facing stiff competition from big retailers with limited resources. Shared here is a mid-term report that will tap into the data available at the POS system of this mall, which captures every transaction, not to mention the data of customers and inventory logs, to engage and enhance competitive advantage. In this analysis, different statistical methods are used to identify major trends and insights.

Calculation of some descriptive statistics, such as mean, minimum, maximum, and standard deviation, helped in understanding the stock data for three months. These numbers cannot just be presented as they are. Each statistic was checked against the business to derive meaningful insights that may turn out to be very helpful in decision-making. For instance, if there is an understanding of how much the stock levels vary, along with the sales patterns, then the optimum management of the inventory can be ensured with the assurance of the availability of the items in demand. The Key-findings suggest that categories like "Women's Wear" and "Backpack" drive a substantial portion of sales, suggesting that targeted marketing could enhance revenue. The sales were highest in the festival season in the month of November.

The proposed data-driven strategies will target improving the customers' shopping experiences and building customer loyalty to better the position of 99 Market Mall in the local market. This will ensure further growth in business and help the mall face competition effectively within Maharajganj, Siwan, and surrounding areas. In the near future, expansion in variables and time frame of the dataset will be done to increase the depth of analysis and further validate these findings.

2 Proof of Originality of Data:



- [Click here for more photo proof of the shop](#)
- [Interaction video with the shop owner](#)

Signed Authorization Letter by the Shop Owner

99 Market Mall

Opp.canara bank maharajganj
Siwan,bihar 841238
Contact:+91 7250072591

01-05-2024

To whom it may concern,

Subject: Consent for the Utilisation of data in the context of a Business Data Management Capstone Project.

This correspondence is to officially confirm that Mr. Rajeev Kumar Bhagat (Student ID: 21F3001527), an esteemed student currently pursuing the BS degree program specialising in Data Science and Applications at IIT Madras, has been duly authorised to utilise our data for his Business Data Management Capstone Project. In our capacity as representatives of 99 Market Mall, we hereby grant Mr. Rajeev full access to and approval for the usage of this data for his college project. We are pleased to provide this official confirmation, granting Mr. Rajeev the requisite permissions to access the data required for the duration of his project. The data supplied is genuine and authenticated to the best of my knowledge. We are enthusiastic about offering our support to Satish as he progresses through his educational journey and works on his project.

Yours truly,
Brayesh kumar shah
Proprietor
99 Market Mall

Brayesh shah
05-05-24

3 MetaData:

The Data Can be Validated from here : [DATASET SHEET LINK](#)

Test Run Results for Google Colab File can be Validated from here : [Google Colab](#)

NOTE : Only a few items bills were generated by the shop over the months(due to the unavailability of electricity at the shop most of the times). For other sales, they give physical handwritten receipts to the customers for their purchase. For their sales record, they note this thing in their diary. Therefore, the data present in the sheets as “Raw Data” contains only the electronically generated bills data, for other sales, I have done the calculations manually from their diary(whose photo was not allowed to capture by the shop owner) and have added the amounts in the “Cleaned Data” section of the Google Sheets.

Duration of collected data : November 2023 - January 2024 (3 months).

The data collected consists of Stock Data and Stock-Sales Summary data for 99 Market Mall. Metadata pertains to descriptive information regarding this dataset. The product data consists of three columns. (Fig 3.1)

Items: Individual products listed.

Items Category: Classification of the items into respective categories.

Rate: Cost or price assigned to each item.

Quantity: Amount or number of items available.

Price: Total cost calculated based on the rate and quantity.

Selling Price per Piece: Price set for each individual item.

Selling Price: Total price for all items sold.

CLEANED DATASET							
99 MARKET MALL							
ITEMS, STOCK, PURCHASING PRICE AND THEIR SALES							
SNo	ITEMS	ITEM_CATEGORY	RATE	QUANTITY	PRICE	SELLING PRICE/PIECE	SELLING PRICE
1	SOAP	PERSONAL CARE	30	10	300	36	360
2	T-SHIRTS	APPAREL	320	30	9600	384	11520
3	JEANS	APPAREL	900	20	18000	1080	21600
4	SKIRTS	WOMENS WEAR	480	20	9600	576	11520
5	JARS	MASON JARS	300	30	9000	360	10800
6	SOCKS	UNDERGARMENTS	80	40	3200	96	3840

7	PURSE	HANDBAG	240	20	4800	288	5760
8	SHAMPOO	PERSONAL CARE	210	30	6300	252	7560
9	WALL WATCH	CLOCK	150	20	3000	180	3600
10	SCHOOL BAG	BACKPACK	450	20	9000	540	10800
11	FLOWER POT	PLANTERS	600	20	12000	720	14400
12	TOOTH BRUSH	DENTAL CARE	25	40	1000	30	1200
13	CHAIR	FURNITURE	320	30	9600	384	11520
14	TROLLEY	BACKPACK	1200	20	24000	1440	28800
15	LADIES KURTI	WOMENS WEAR	660	20	13200	792	15840
16	TOOTH PASTE	DENTAL CARE	20	20	400	24	480
17	SKIRTS	WOMENS WEAR	480	10	4800	576	5760
18	HAND WATCH	CLOCK	300	20	6000	360	7200
19	DUPATTA	WOMENS WEAR	280	30	8400	336	10080
20	BUCKET	HOME ACCESSORIES	150	40	6000	180	7200
21	LUNCH BOX	PLASTIC	180	50	9000	216	10800
22	NOTEBOOK	STATIONARY	40	100	4000	48	4800
23	WATER BOTTLE	ACCESSORIES	18	50	900	20	1000
24	HAIR OIL	ACCESSORIES	35	40	1400	42	1680
25	DETERGENTS	HOME ACCESSORIES	90	50	4500	108	5400
26	TONGUE CLEANER	DENTAL CARE	5	50	250		300
					TOTAL		213820

Fig 3.1

The monthly Stock Data for 99 Market Mall is organized by category and comprises five columns. (Fig 3.2)

Category Name: Names of product categories.

Stock Amount: Amount for Quantity of items available in stock.

Sale Amount: Total value of items sold.

Profit: The monetary gain achieved from sales.

Number of Variety: Count of different types of products available across categories.

CATEGORY WISE					
SNo	CATEGORY NAME	STOCK AMOUNT	SALE AMOUNT	PROFIT	NUMBER OF VARIETY
1	MASON JARS	9000	10800	1800	1
2	PLASTIC	9000	10800	1800	1
3	PERSONAL CARE	6600	7920	1320	2
4	APPAREL	27600	33120	5520	2
5	CLOCK	9000	10800	1800	2
6	BACKPACK	33000	39680	6680	2
7	HOME ACCESSORIES	10500	12600	2100	2
8	DENTAL CARE	1650	1980	330	3
9	ACCESSORIES	6800	8080	1280	3
10	WOMENS WEAR	36000	43200	7200	4

Fig 3.2

The monthly Stock Data Amount for 99 Market Mall is given below and has 3 columns. (Fig 3.3)

Months: Particular month of whose stock amount has been given.

Stock Amount: Amount for Quantity of items available in stock for a particular month.

Duration: Dates covered in a particular month.

OVERALL MONTHLY STOCK		
MONTHS	DURATION	MONTHLY_STOCK_AMOUN T
November 2023	1-11-2023 to 30-11-2023	178000
December 2023	1-12-2023 to 31-12-2023	158600
January 2024	1-01-2024 to 31-01-2024	156016. 8

Fig 3.3

The data now got cleaned using Google Colab. Upon observing all the rows and columns, we removed all the null or empty values. Also removed some of the mistyped data from our dataset that could have caused problems in doing our analysis. The Results can be seen in the below screenshots.

```

import numpy as np
import pandas as pd
df=pd.read_excel('/content/BDM DATA.xlsx',sheet_name='ITEMS')
df.head()

```

SNo	ITEMS	ITEM_CATEGORY	RATE	QUANTITY	PRICE	SELLING PRICE/PIECE	SELLING PRICE	
0	1	SOAP	PERSONAL CARE	30	10	300	36	360
1	2	T-SHIRTS	APPAREL	320	30	9600	384	11520
2	3	JEANS	APPAREL	900	20	18000	1080	21600
3	4	SKIRTS	WOMENS WEAR	480	20	9600	576	11520
4	5	JARS	MASON JARS	300	30	9000	360	10800

Fig 3.4

We are reading the data. The columns present inside the database can be seen in the above screenshot. The excerpt of all the columns can be seen above. (Fig 3.4)

```

[ ] import numpy as np
import pandas as pd
df=pd.read_excel('/content/BDM DATA.xlsx',sheet_name='SALES AMOUNT')
df.head()

```

	MONTHS	DURATION	SALES AMOUNT	NET PROFIT
0	2023-11-01	1-11-2023 to 30-11-2023	213600	35600
1	2023-12-01	1-12-2023 to 31-12-2023	198250	31251
2	2024-01-01	1-01-2024 to 31-01-2024	195021	300456

Fig 3.5

Similar thing has been done on the Sales Data of our Dataset and can be seen above(Fig 3.5).

3 Descriptive Statistics :

Once our data is cleaned (null values and empty columns removed), we are trying to describe our data. The 'df.describe()' function is used to obtain the relevant statistics for the dataset.

1. Descriptive Statistics of the product (including variety) monthly stock(Fig 4.1) :

```

[5]: import pandas as pd
df = pd.read_excel('/content/21F3001527 BDM DATA.xlsx', sheet_name="T_STOCK")
df.describe()

```

	SOAP	T-SHIRTS	JEANS	SKIRTS	JARS	SOCKS	PURSE	SHAMPOO	WALL WATCH	SCHOOL BAG	...	SKIRTS .1	HAND WATCH	DUPATTA	BUCKET	LUNCH BOX	NOTE BOOK	WATER BOTTLE
count	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	...	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	
mean	113.333333	3316.666667	6306.666667	3366.666667	3110.000000	1106.666667	1686.666667	2180.000000	1056.666667	3156.666667	...	1763.333333	2106.666667	2903.333333	2063.333333	3076.666667	1380.000000	322.666667
std	161.957075	5443.457847	10136.278981	5403.122554	5102.675769	1812.990164	2698.486702	3569.195956	1684.230784	5065.040309	...	2640.309325	3374.630844	4761.839381	3409.696956	5130.169398	2269.184875	500.241275
min	10.000000	30.000000	20.000000	30.000000	40.000000	20.000000	30.000000	20.000000	20.000000	30.000000	...	10.000000	20.000000	30.000000	40.000000	50.000000	40.000000	18.000000
25%	20.000000	175.000000	460.000000	250.000000	165.000000	60.000000	130.000000	120.000000	85.000000	235.000000	...	245.000000	160.000000	155.000000	95.000000	115.000000	70.000000	34.000000
50%	30.000000	320.000000	900.000000	480.000000	300.000000	80.000000	240.000000	210.000000	150.000000	450.000000	...	480.000000	300.000000	280.000000	150.000000	180.000000	100.000000	50.000000
75%	65.000000	960.000000	9450.000000	5040.000000	4650.000000	1640.000000	2520.000000	3255.000000	1575.000000	4725.000000	...	2640.000000	3150.000000	4340.000000	3075.000000	4590.000000	2050.000000	475.000000
max	300.000000	9600.000000	18000.000000	9600.000000	9000.000000	3200.000000	4800.000000	6300.000000	3000.000000	9000.000000	...	4800.000000	6000.000000	8400.000000	6000.000000	9000.000000	4000.000000	900.000000

8 rows x 26 columns

Fig 4.1

- Upon analyzing the statistics, it was observed that the item 'JEANS' has the highest mean, while 'TONGUE CLEANER' exhibits the lowest mean.
- An excerpt of the result of the above mentioned steps in Colab has been attached below. \

2. Descriptive Statistics of purchase and sales rate of each product(Fig 4.2) :

Upon analysis of the statistics, it was revealed that the product "**Tongue Cleaner**" has the lowest sale rate valued at Rs 06, while the product "**Trolley**" holds the highest sale rate of Rs 1440. Similarly, regarding the purchase rates, "**Tongue Cleaner**" has the minimum rate of Rs 5, whereas "**Trolley**" boasts the maximum rate of Rs 1200.

```
[10] import pandas as pd
    df = pd.read_excel('/content/21f3001527 BDM DATA.xlsx', "PURCHASEvSALES")
    df.describe()
```

	PURCHASE RATE	SELLING PRICE/PIECE
count	26.000000	26.000000
mean	290.884615	349.000000
std	294.675323	353.669789
min	5.000000	6.000000
25%	50.000000	60.000000
50%	225.000000	270.000000
75%	417.500000	501.000000
max	1200.000000	1440.000000

Fig 4.2

3. Descriptive Statistics of Sales and stock by Amount of each category(Fig 4.3):

After analyzing the statistics, it was determined that the category "**Women's Wear**" has the maximum variety with a value of 8, while "**Mason Jars**" has the minimum variety with a value of 2.

```
[8] import pandas as pd
    df = pd.read_excel('/content/21f3001527 BDM DATA.xlsx', "COMPANY_COLAB")
    df.describe()
```

	SNo	STOCK AMOUNT	SALE AMOUNT	PROFIT	NUMBER OF VARIETY
count	10.000000	10.000000	10.000000	10.000000	10.000000
mean	5.500000	14915.000000	17898.000000	2983.000000	2.200000
std	3.02765	12328.648885	14813.30468	2484.834222	0.918937
min	1.000000	1650.000000	1980.000000	330.000000	1.000000
25%	3.250000	7350.000000	8760.000000	1440.000000	2.000000
50%	5.500000	9000.000000	10800.000000	1800.000000	2.000000
75%	7.750000	23325.000000	27990.000000	4665.000000	2.750000
max	10.000000	36000.000000	43200.000000	7200.000000	4.000000

Fig 4.3

4. Descriptive Statistics of **Sales by Amount of three months (November, December, January)**(Fig 4.4):

Upon analysis, it was discovered that the average sales value across months amounted to 202290.33. Notably, in three months, the maximum sales amount, reaching 2,13,600, occurred specifically in November(festival seasons in Bihar) while the lowest was in January, reaching 1,95,021 only(due to extreme cold).

```
[7] import pandas as pd
df = pd.read_excel('/content/21f3001527 BDM DATA.xlsx', sheet_name="SALES COLAB")
df.describe()
```

	MONTHS	SALES AMOUNT	NET PROFIT
count	3	3.000000	3.000000
mean	2023-12-01 08:00:00	202290.333333	122435.666667
min	2023-11-01 00:00:00	195021.000000	31251.000000
25%	2023-11-16 00:00:00	196635.500000	33425.500000
50%	2023-12-01 00:00:00	198250.000000	35600.000000
75%	2023-12-16 12:00:00	205925.000000	168028.000000
max	2024-01-01 00:00:00	213600.000000	300456.000000
std	Nan	9926.632376	154185.465464

Fig 4.4

4 Detailed Explanation of Analytical Process:

- Product Sales and Inventory Analysis:** Utilize Excel's SUM function to aggregate monthly stock quantities for each product category and related items available at 99 Market Mall. Extract monthly sales quantities from the stock-sales summary. Employ a Column chart to visually compare monthly stock and sales quantities. Introduce Analysis charts(pie chart/bar graph(attached below)) to identify the most impactful product categories in terms of sales volume and revenue generation.
- Customer Preferences Analysis for Products:** Compile monthly sales quantities for each product category and compute the total sales amount per month. Present Sales Quantity and Sales Amount using Line and Bar charts, respectively. These visualizations will highlight top-selling products and customer preferences at 99 Market Mall, aiding in inventory management and marketing strategies.
- Monthly Sales Data Compilation for Products:** Sum up daily product sales to generate monthly sales data for analysis. Utilize a Line Chart in Excel to illustrate overall performance and trends in product sales, providing a clear overview of monthly sales patterns at 99 Market Mall.
- Monthly Inventory Analysis for Products:** Sum up the total inventory for each product category on a monthly basis. Use a column plot in Excel to depict these monthly stock levels and variations. Additionally, employ another column plot to showcase the total variety of products available per specific department within 99 Market Mall, aiding in inventory optimization and shelf management strategies.

5 Results and Findings:

The below pie chart shows the number of varieties each category possesses. We can see, Women's wear has the highest number of variety of products available while Mason Jars and Plastic has the lowest number of variety for products in our shop(Fig 6.1).

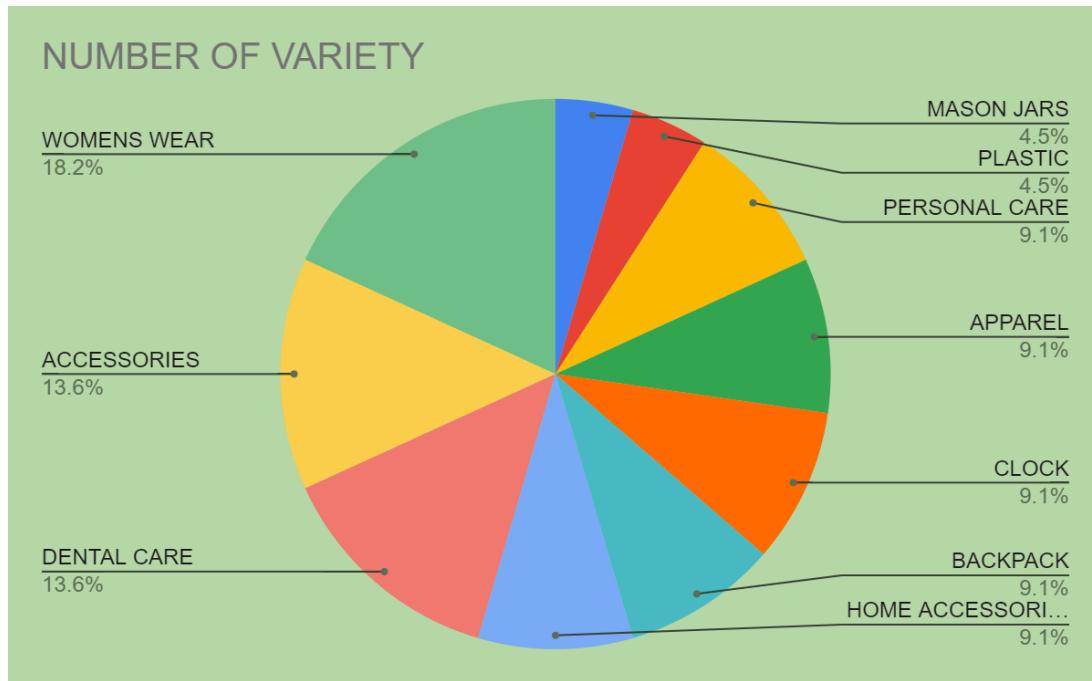


Fig 6.1

Upon analyzing the Line Chart depicting Total Monthly Sales and Total Monthly Stock, it's noticeable that compared to the sales in November, the total sales during the other two months were lower. (Due to less customers in the market in the extreme cold season). (Fig 6. 2).



Fig 6.2

The next finding is about the stock amount we have spent vs the sale amount(to ensure inventory management). We can see from the chart, the sale amount is always higher than the stock amount which ensures we are ordering the items in the correct amount. We can say, all our ordered items are being sold(Fig 6.3).

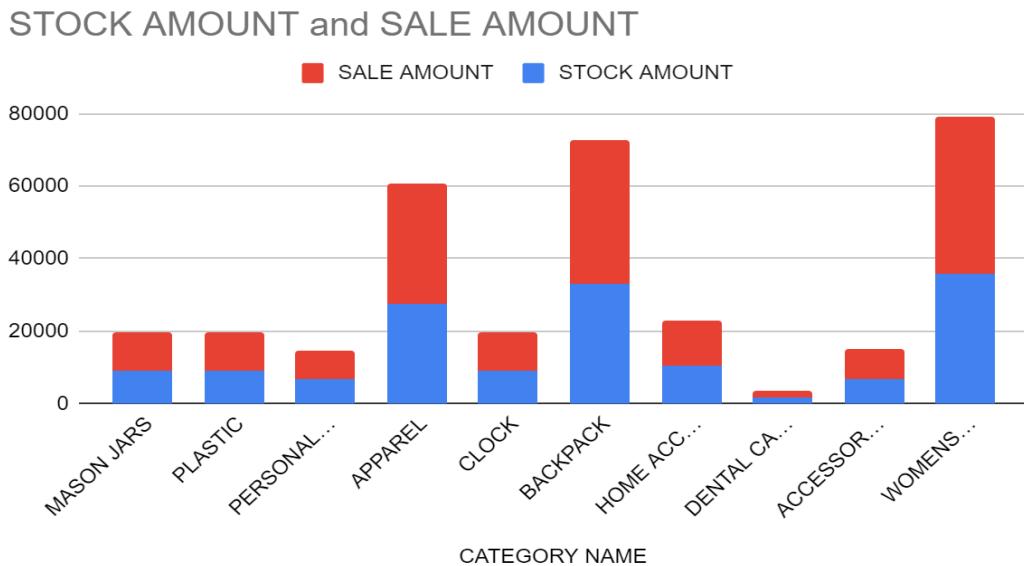


Fig 6.3

The Pareto chart illustrating category stock amounts does not completely align with the Pareto Principle. Roughly one-third of the items sold originate from only three specific Categories: Women's Wear, Backpack, and Apparel(Fig 6.4).

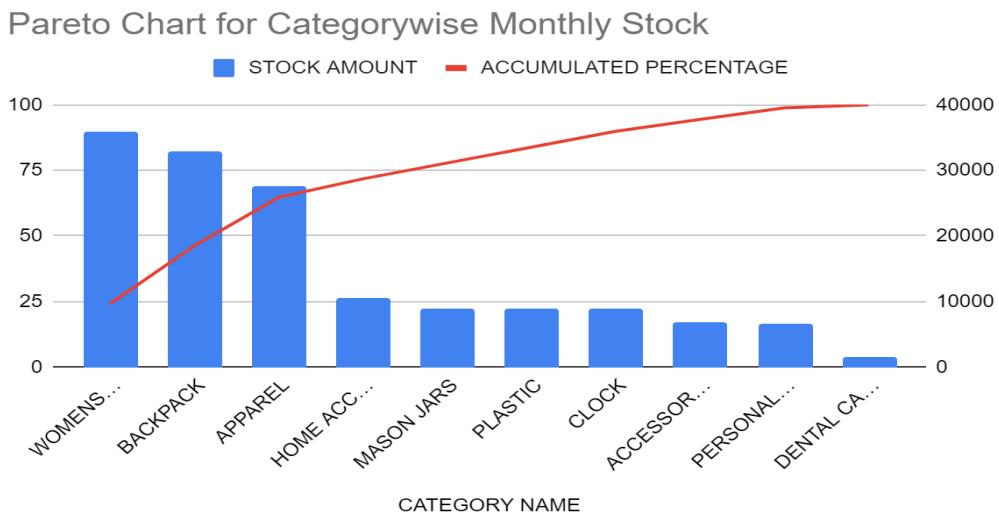


Fig 6.4

Customer Preference- Analyzing the Sales Quantity and Sales Amount chart it is clear that customers prefer to buy "Women's Wear" category products. Customers like to buy "Backpack" and "Apparel" category products but "Women's Wear" contributes the most to the 99 market mall with its large variety and sales.