

Data-Driven Decision Making for Saeed Mega Mart: A Case Study of Employing Sales Forecasting, Customer Segmentation, and Sentiment Analysis Strategies

A Midterm report for the BDM Capstone Project

Submitted by

Name: ANEEQA HAROON SIDDIQUI

Roll number: 21f2000377

Email Id: 21f2000377@ds.study.iitm.ac.in



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

1 Executive Summary

Saeed Mega Mart is a small family business that operates a small mart in **Sandila, Uttar Pradesh**, India. The business sells various products such as **bangles, cosmetics, handbags, jewellery**, and **gift items**. It also **rents out** some of its shops to other businesses.

The project aims to use data-driven methods to optimize the business performance of Saeed Mega Mart by addressing two problems: how to **improve the sales forecasting and analysis** using the daily sales data and external factors; and how to enhance the **customer segmentation and loyalty** using the customer data and feedback.

The in this project I have uses **MS Excel** and **Python** as the analysis tools. Python is used to perform the initial data cleaning and preparation. Excel is used for making pivot tables and charts for visualisation.

Google form is used to collect the customer data as it also provide basic visualisation of responses collected from customer.

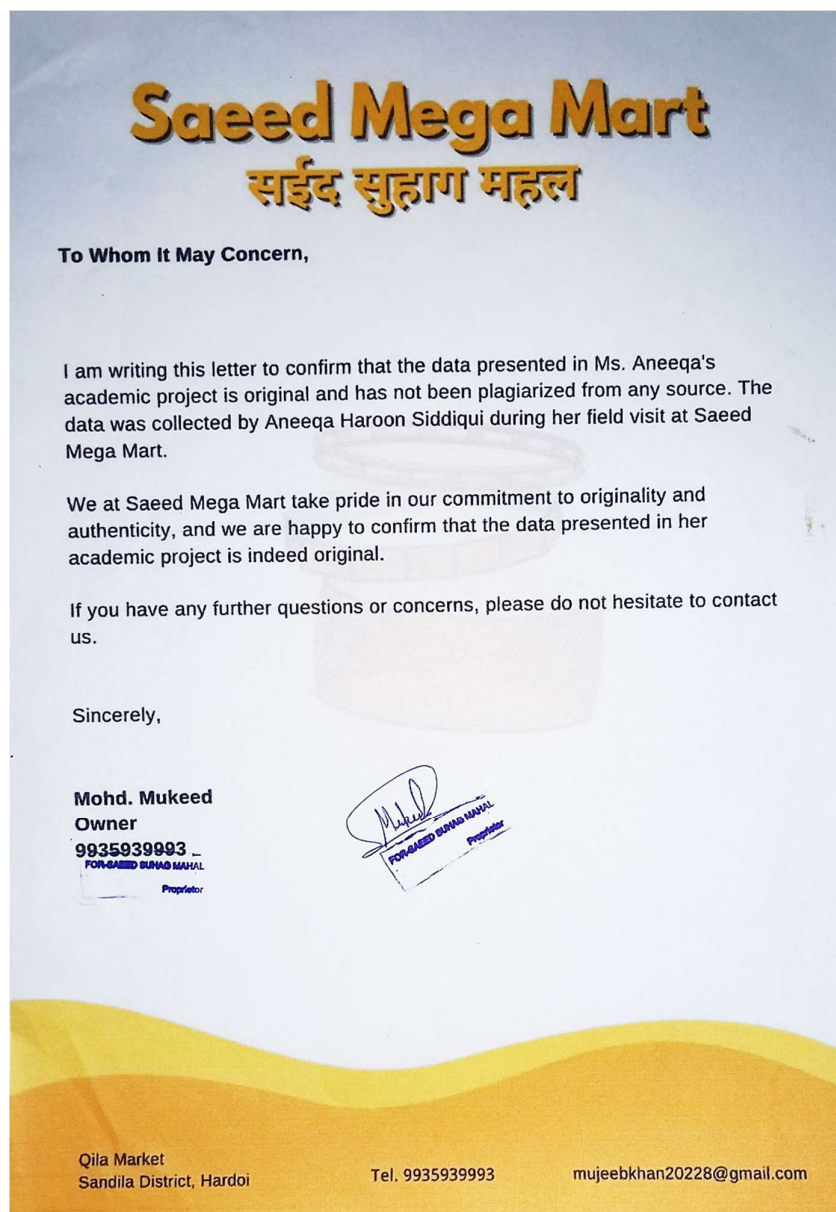
On performing **descriptive analysis** we conclude that the mean daily sales amount is **Rs 2818.8**. Missing data has been imputed using median values.

On Analysis we found that peak sales can be seen in **October-November** and sharp decrease in sales in July to September. From day-wise sales trend we can see peak sales in week days and Tuesdays and Thursdays whereas very few sales on **Wednesday**.

Most of the customers are satisfied with the services and they shop weekly. Larger part of the customer are female from **25-35 years** age group. **Customer retention** is good as 80% of the customers are old customers.

2 Proof of originality

Letter from organization in letter head with stamp and sign: All the data is provided by Mr. Mukeed owner of Saeed Mega Mart. Please find below letter to validate the authenticity of the data.



Images of the Mart:

Below are the images from the field visit during data collection.





A short video interacting with the owner:

<https://drive.google.com/drive/folders/105IEklTmmGLtkB2pzOVHtLa9hDeZ2rLF?usp=sharing>

3 MetaData

Sales Data:

Sales data provided by Mr. Mukeed was recorded in an old notebook. Data collected is from November 2019 to September 2023.

138150

12150

130300

AGE
DATE: 2/13/2020

131000

4600-16-11-19
11100-17-11-19
8600-18-11-19
7100-19-11-19
500-20-11-19
4800-21-11-19
7250-22-11-19
5700-23-11-19
11900-24-11-19
3550-25-11-19
3850-26-11-19
6500-28-11-19
2100-29-11-19
9000-30-11-19
7750-1-12-19
2100-2-12-19
6950-3-12-19
~~1250-4-12-19~~
4050-4-12-19
3700-5-12-19
1400-6-12-19
4200-7-12-19
2100-8-12-19
1500-9-12-19
1000-10-12-19
1300-11-12-19
900-12-12-19
650-13-12-19
1650-14-12-19
2500-15-12-19
800-16-12-19
1200-17-12-19
2100-19-12-19
600-20-12-19
1520-21-12-19
1750-22-12-19
800-23-12-19
400-24-12-19
400-26-12-19

2750-25-12-19
750-29-12-19
700-30-12-19
2350-31-12-19
1700-2-1-20
1750-3-1-20
1800-4-1-20
2600-5-1-20
2600-6-1-20
900-7-1-20
1800-9-1-20
800-10-1-20
900-11-1-20
3500-12-1-20
2950-14-1-20
400-16-1-20
2800-17-1-20
1450-18-1-20
5200-19-1-20
1950-20-1-20
1200-21-1-20
7300-23-1-20
1450-24-1-20
5750-25-1-20
1500-26-1-20
5300-27-1-20
5100-28-1-20
900-30-1-20
4400-31-1-20
1450-1-2-20
10200-2-2-20
2250-3-2-20
2850-4-2-20
1450-5-2-20
3100-8-2-20
2750-7-2-20
2650-8-2-20
6650-9-2-20
~~400-10-2-20~~
5100-10-2-20

5100-11-20
1200-12-20
6950-13-20
3600-14-20
3200-15-20
3300-16-20
1950-17-20
4100-18-20
4850-20-20
1300-21-20
3700-22-20
5950-23-20
1950-24-20
4150-25-20
5850-26-20
7600-28-20
2450-29-20
2500-31-20
2550-2-30
3200-3-30
2450-5-30
4900-6-30
4400-7-30
5500-8-30
1800-9-30
3600-12-30
1200-13-30
3400-14-30
5800-15-30
3300-16-30
700-17-30
1200-18-30
2300-19-30
2800-20-30
1100-21-30
1050-15-30
1200-16-30
500-17-30
1300-22-30
2750-23-30
2750-24-30
1000-25-30

Original data contains:

- Sales amount at the end of the day
- Date

Please use below link to see original data received.

https://drive.google.com/file/d/10e5v1mYHNHhIrpZSLbR6JFhoXU6Sa6Z8/view?usp=drive_link

Customer Feedback data:

Customer feedback is taken from a Google-form which was earlier circulated in WhatsApp groups.

Form link: <https://forms.gle/qhtXJYBe1GPxBndh9>

Data link: <https://drive.google.com/file/d/1eGI-yFJdOz0FTTrVuSX-oApeP0tlv4F3u/view?usp=sharing>

4 Descriptive Statistics

The descriptive statistics for the Sale variable are shown in below Table.

The mean sale amount was 2818.8. The minimum and maximum sale amounts were 100 and 23300, respectively. The median sale amount was 2050, and the interquartile range was 2200.

	Sale
count	990
mean	2818.83939
std	2425.2331
min	100
25%	1300
50%	2025
75%	3500
max	23300

The data set had 2 missing values in the Date column and 428 missing values in the Sale column. The missing values in the Date column were dropped, as they did not affect the analysis of the Sale variable.

The missing values in the Sale column were replaced with the median value of 2050, as the data was skewed and had outliers. This method reduced the impact of missing values on the mean and standard deviation of the Sale variable.

Cleaned Dataset:

<https://drive.google.com/file/d/1YJKxFqiQ7OVtJ4m7upTHrdTyHMXQdF09/view?usp=sharing>

5 Analysis Process

For Analysis I have used MS-Excel as my data just consists of Date and Sales.

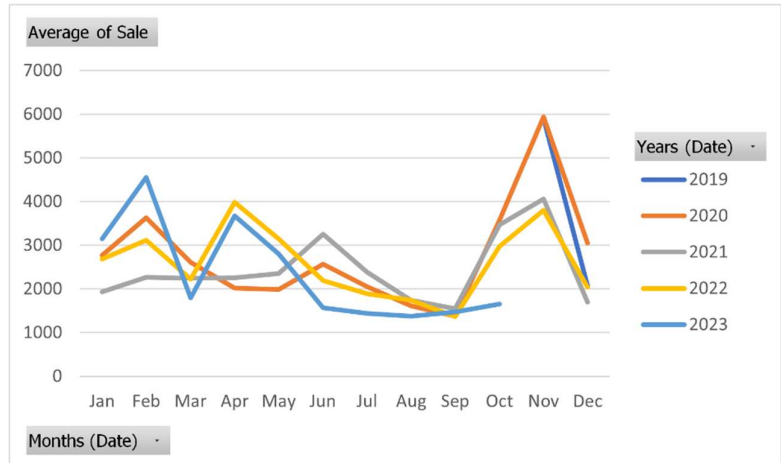
- I have done the basic descriptive analysis and cleaned the Data above using Python.
- I have derived Day, Month columns from Date column to analyse the trends appearing month-wise or day-wise.
- Firstly, I have created basic pivot tables, like Average monthly sales data from year late 2019 to late 2023.
- I have used Charts option in Excel to visualize my pivot table to analyse the trends.

6 Results and Findings

Average Monthly Sales Data:

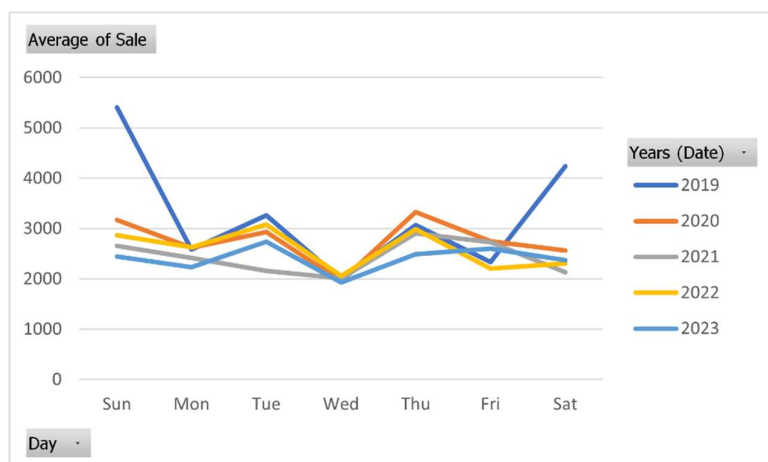
We can see the below pattern, for all the years 2019-23.

- shows a general decrease in sales from January to December for all years.
- We can see that there's a sharp decrease in March-May 2020 due to Covid-19. Similar trend can be seen for the next year but not much.
- Increase in sales can be seen in the last quarter probably due to increase in festivals and marriages.



Average Daily Sales for different years:

We can see that the sales vary by the day of the week, with Sunday, Tuesday and Thursday being the most profitable days and Wednesday being the least profitable day: possible reason for this could be on Wednesday's most of the shops are closed as they treat Wednesday as Week-off.

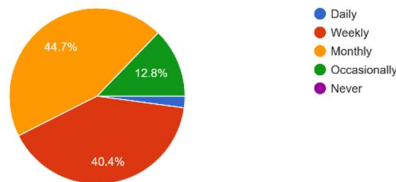


Customer Feedback Analysis:

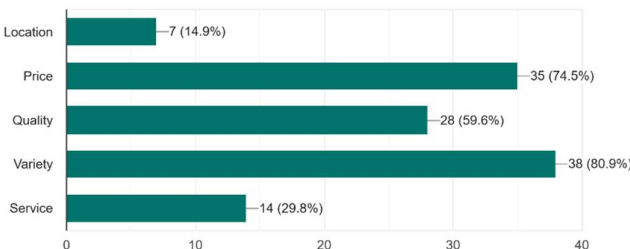
Below are the summary of the responses that I have obtained from customer feedback.

- We can see that most of the customer often shop weekly due to the variety of products and reasonable price at which they are sold.

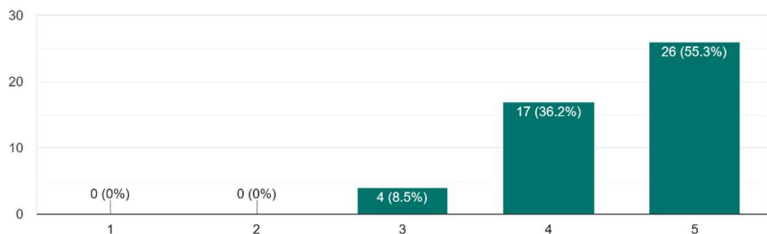
How often do you shop at Saeed Mega Mart ?
47 responses



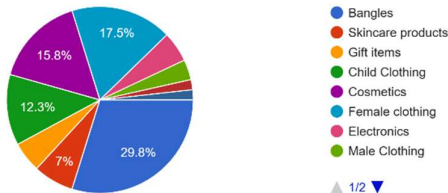
What are the main reasons you shop at Saeed Mega Mart? (Select all that apply)
47 responses



How satisfied are you with the following aspects of Saeed Mega Mart? (Rate from 1 to 5, where 1 is very dissatisfied and 5 is very satisfied)
47 responses



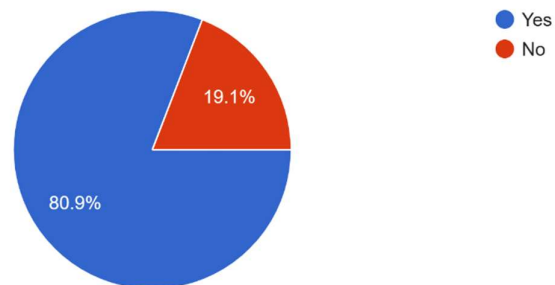
What are some of the products or services that you would like to see more of at Saeed Mega Mart?
47 responses



- More than 80% customers are satisfied with the services.
They are more interested in Bangles, clothing and cosmetics.

Are you an Old Customer?

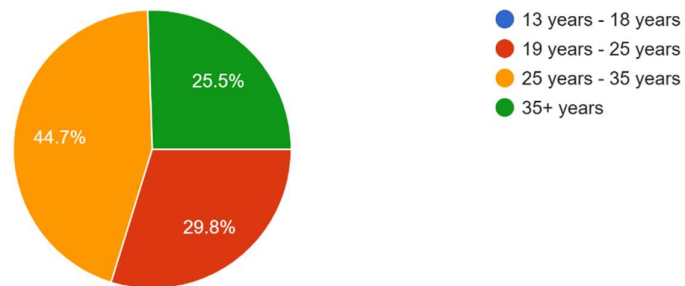
47 responses



- Also, we can see that most of the customers are female and they are old customers.

What is your age?

47 responses



Gender

47 responses

