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

A Final 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

Contents

1 Executive Summary and Title

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 Analysis we found The sales revenue shows a steady increase from 2019 to 2022, with a peak in 2022. However, there's a significant drop in 2023.
Customers prefer shopping monthly and weekly, and they are attracted by affordable prices and a huge variety of quality products. The most desired products are Bangles, female clothing, and cosmetics.
Most customers are satisfied with the services and products provided by Saeed Mega Mart and are likely to recommend it.
The majority of the customers are female and belong to the age group of 25+ years.
Based on the analysis, **recommendations** include maintaining product quality and variety, offering special deals or discounts during peak shopping periods, implementing customer loyalty programs, and tailoring marketing strategies and product offerings to cater to the demographic of female customers aged 25-35 years. Regular data analysis and timely action based on the findings can help Saeed Mega Mart maintain its strong performance and customer satisfaction.

2 Detailed Explanation of Analysis Process/Method

Sales Analysis

The sales analysis process begins with the meticulous collection and recording of daily sales revenue data, captured over a significant period, in this case, from 16-11-2019 to 1-10-2023. This comprehensive dataset ensures a robust foundation for an in-depth and accurate analysis.

2.1 Data Collection:

- o Every day, at the close of business, the total sales revenue is recorded along with the date.
- o The data is manually entered into a ledger in an unstructured format.
- o [Raw data](#).

2.2 Data Entry:

4

- The collected data is manually entered into an Excel spreadsheet.
- Each entry consists of the date and the corresponding daily sales revenue.
- Also date entered is continuous from 16-11-2019 to 01-10-2023.
- Excel Data

Date	Daily Sales Revenue
16-11-2019	4600
17-11-2019	11100
18-11-2019	8600
19-11-2019	7100
20-11-2019	500
21-11-2019	4800
22-11-2019	7250
23-11-2019	5700
24-11-2019	11900

2.3 Data Cleaning:

- The data is cleaned using Python to ensure accuracy.

Column	Missing Values	Percentage Missing
Date	0	0
Daily Sales Revenue	426	30.084746

- This involves filling in missing values in the Sales column which 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 Sales variable.
- Cleaned Data.
- Google colab link.

2.4 Data Transformation:

o Additional columns for Day, Month and Year are derived from the Date column. This allows for more granular analysis of trends on a daily and monthly basis.

Date	Daily Sales Revenue	Day	Month	Year
16-11-2019	4600	Saturday	Nov	2019
17-11-2019	11100	Sunday	Nov	2019
18-11-2019	8600	Monday	Nov	2019
19-11-2019	7100	Tuesday	Nov	2019
20-11-2019	500	Wednesday	Nov	2019

2.5 Descriptive Analysis:

o Descriptive analysis is performed on the data using Excel. This involves creating pivot tables to summarize the data and identify patterns and trends.

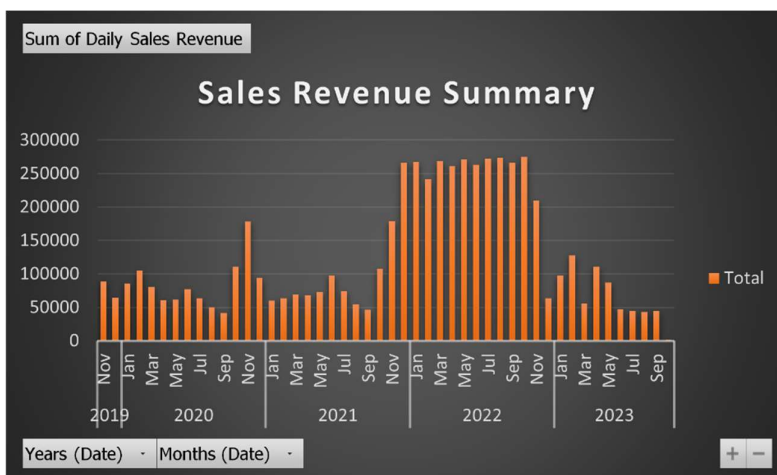|	Daily Sales Revenue
count	990
mean	2818.83939
std	2425.2331
min	100
25%	1300
50%	2025
75%	3500
max	23300

2.6 Data Visualization:

o I have created pivot tables to analyse the Monthly Sales Summary for all years combined and summaries by year as the combined summary would give an overall view of our sales trends, but it might obscure year-to-year variations.

2019-23	Sum of Daily Sales Revenue
2019	152800
2020	1009600
2021	1157596
2022	2931249
2023	657915
Grand Total	5909160

o The pivot tables are then visualized using Excel's charting functionality. This provides a clear and concise view of the sales trends over time.

Average Daily Sales for different years
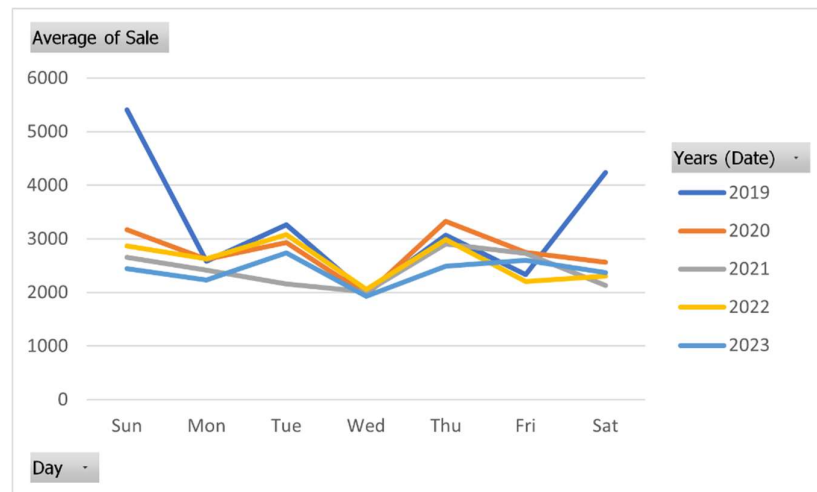
Sales Revenue Trend over past few years.

Average Monthly Sales

Average of Sale
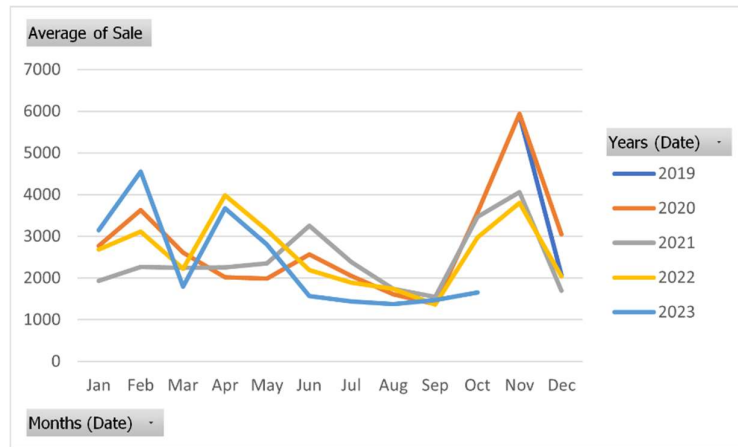
3 Detailed Explanation of Analysis Process/Method

Customer Feedback Analysis

3.1 Data Collection:

- o Data collection is a crucial step in our customer feedback analysis.
- o We created a Google Form to gather valuable information from our customers.
- o This form was circulated in various WhatsApp groups, reaching a wide range of our customer base.
- o The form included questions about shopping frequency, reasons for choosing Saeed Mega Mart, satisfaction levels, loyalty, and more.
- o This method allowed us to collect a total of 47 responses till now, providing us with a dataset to analyse and gain insights from.
- o [Form Link](#).
- o [Excel Data.](#)

Timestamp	How often do you shop at Saeed Mega Mart ?	What are the main reasons you shop at Saeed Mega Mart? (Select all that apply)	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)	How likely are you to recommend Saeed Mega Mart to a friend or family member? (Rate from 1 to 5, where 1 is very unlikely and 5 is very likely)	What are some of the products or services that you would like to see more of at Saeed Mega Mart?	How do you compare Saeed Mega Mart with other similar businesses in your area? (Rate from 1 to 5, where 1 is much worse and 5 is much better)
21-11-2023 08:23	Daily	Price	5	4	Cosmetics	5
15-11-2023 05:48	Occasionally	Variety	4	4	Gift items	5
15-11-2023 20:49	Monthly	Location	5	5	Skincare products	5
16-11-2023 01:55	Monthly	Price, Variety, Serv	3	4	Electronics	4
16-11-2023 10:37	Weekly	Location, Price, Qu	4	3	Makeup products	4
16-11-2023 17:05	Occasionally	Location, Quality, \	5	5	Female clothing	5

| How loyal do you feel to Saeed Mega Mart? (Rate from 1 to 5, where 1 is not loyal at all and 5 is very loyal)	Are you an Old Customer?	What is your age?	Gender	Do you have any suggestions or comments on how Saeed Mega Mart can improve its services?
5	Yes	25 years - 35 y	Female	
5	Yes	19 years - 25 y	Female	
5	No	25 years - 35 y	Male	
4	Yes	25 years - 35 y	Female	Online shopping
3	No	35+ years	Male	
5	Yes	25 years - 35 y	Female	

3.2 Data Cleaning:

- In the data cleaning phase, we made two significant changes to our dataset using python. First, we renamed the columns to make them more manageable and easier to understand. This step is crucial as it simplifies the data handling process and makes the data analysis more straightforward.

- Second, we decided to remove the 'customer_suggestions' column from our dataset. This decision was based on the fact that this column had more than 70% missing values.

- Handling missing data is a critical step in the data cleaning process. Since a large portion of the data is missing, it might be more beneficial to remove the column altogether. This is because the missing data might introduce bias or inaccuracies into our analysis.

- [Cleaned data.](#)

3.3 Correlation Analysis:

- Correlation analysis is done using python.
- A high correlation means that two or more variables have a strong

Correlation Heatmap
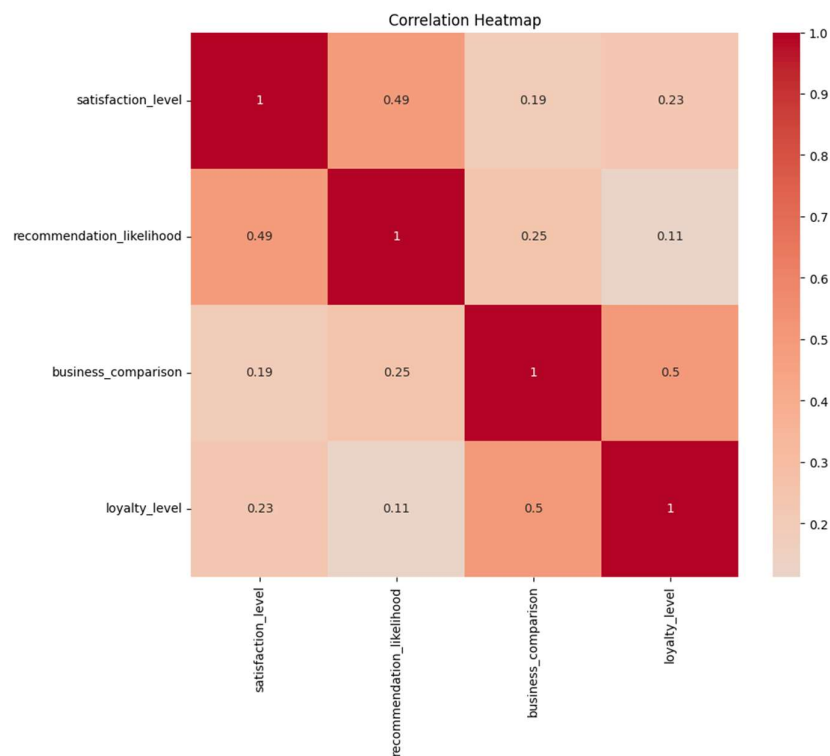
relationship with each other, while a weak correlation means that the variables are hardly related.

- Moderate positive correlation can be seen between "**business_comparison** and **loyalty_level**" and "**recommendation_likelihood** and **satisfaction_level**".

3.4 Data Visualization:

- I have created several pivot tables in Excel to analyse the customer data which are then visualized using Excel charts:

  - Shopping Frequency by Age and Gender

Count of shopping_frequency	Shopping habits of different age groups and genders		
	Female	Male	Grand Total
19 years - 25 years	14		14
25 years - 35 years	19	2	21
35+ years	11	1	12
Grand Total	**44**	**3**	**47**

  - Average Satisfaction Level by Shopping Frequency

Shopping Frequency	Average of satisfaction_level
Daily	5
Monthly	4.333333333
Occasionally	4.5
Weekly	4.578947368

  - Loyalty Level by Old/New Customer Status

Old Customer	Average of loyalty_level
No	3.444444444
Yes	4.263157895

4 Results and Findings

Sales Analysis

Missing value Analysis:

	From	To	Missing record continuously for more than 5 days
1.	21-03-2020	14-05-2020	55
2.	18-07-2020	26-07-2020	9
3.	30-12-2020	20-01-2021	22
4.	01-05-2021	09-05-2021	9
5.	14-05-2021	26-05-2021	13

1. **21-03-2020 to 14-05-2020 (55 days)**: This period falls under the first nationwide lockdown in India due to the COVID-19 pandemic, which started on 25 March 2020.
2. **18-07-2020 to 26-07-2020 (9 days)**: This period does not seem to coincide with any major national holidays or lockdowns in India maybe the data was not recorded here.
3. **30-12-2020 to 20-01-2021 (22 days)**: This period falls under lockdown.
4. **01-05-2021 to 09-05-2021 (9 days)**: This period falls under the lockdown in Uttar Pradesh, which was extended till 10 May 2021.
5. **14-05-2021 to 26-05-2021 (13 days)**: This period also falls under a lockdown.

Average Monthly Sales Data:

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

- We can see that there's a sharp decrease in March-May 2020 due to Covid-19. Similar trend can be seen for the
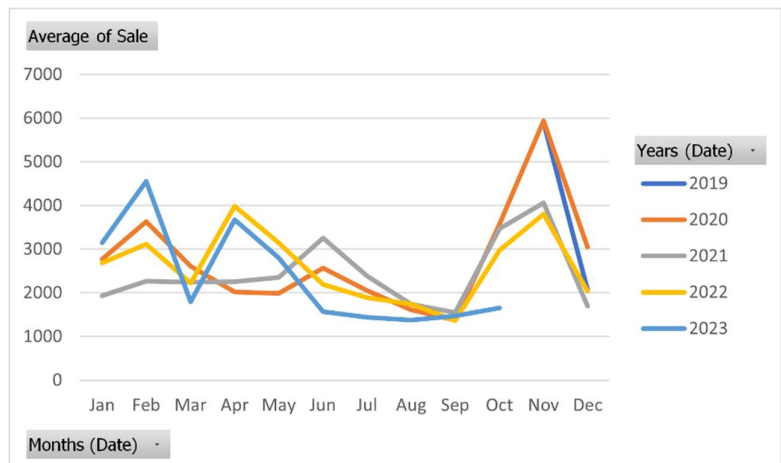
next year but not much.

- 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.
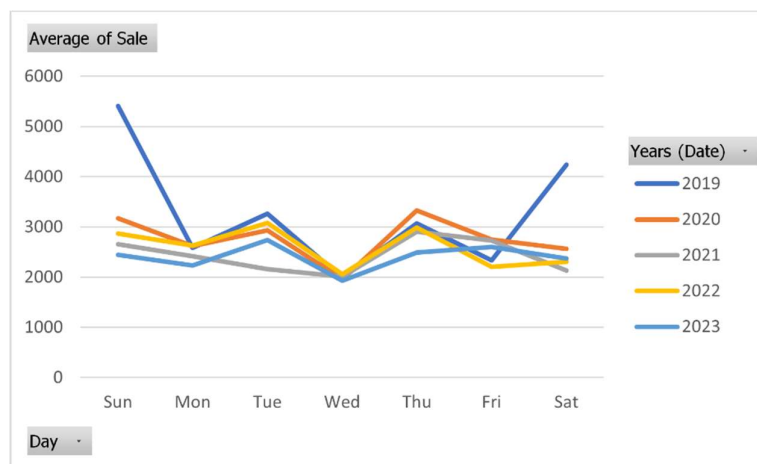
Sales Revenue Trend over past few years.

- ○ **Yearly Analysis**: The yearly sales revenue shows a steady increase from 2019 to 2022. The highest revenue was recorded in 2022 with a total of 2,931,249. However, there's a significant drop in 2023 with a total of 657,915. This could be due to various factors such as changes in market conditions, product demand, or business strategies.
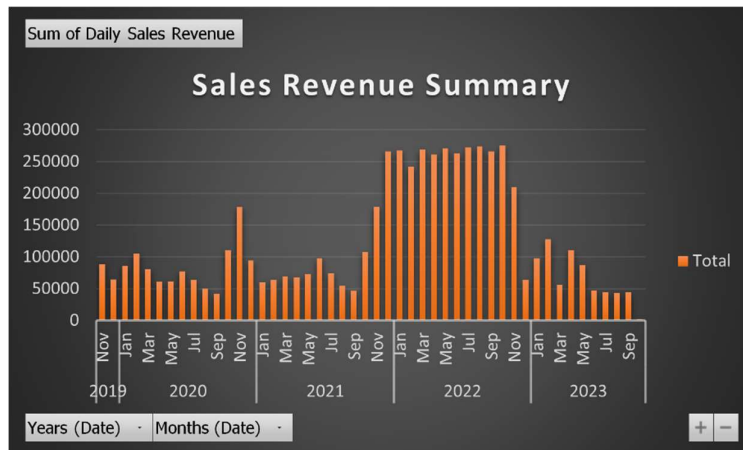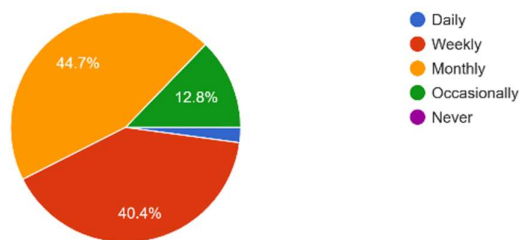
- **Monthly Analysis**: The monthly sales revenue in 2022 shows a consistent performance with the highest revenue recorded in October (275,342) and the lowest in December (63,275). However, in 2023, there's a noticeable decline starting from June.
- **Overall Trend**: The overall trend shows a peak in sales revenue in 2022 followed by a decline in 2023. This suggests that while the sales performance was strong in 2022, there might have been challenges or changes in 2023 that affected the sales negatively.
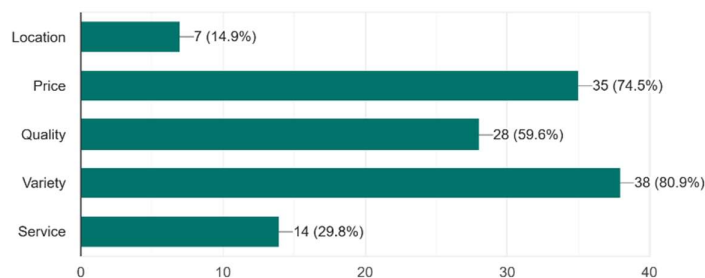
Customer Feedback Analysis:

Results of Google form responses:

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

|	
Location	7 (14.9%)
Price	35 (74.5%)
Quality	28 (59.6%)
Variety	38 (80.9%)
Service	14 (29.8%)

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
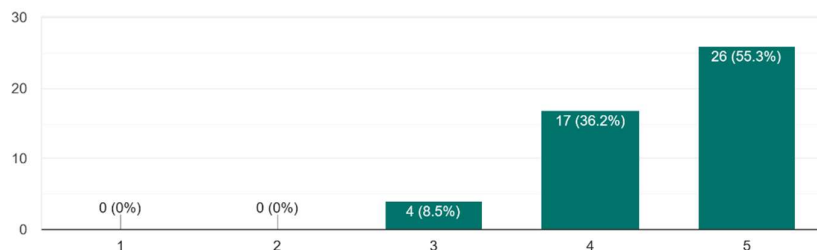
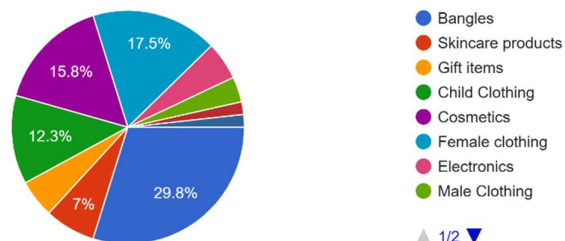| Rating	Count
1	0 (0%)
2	0 (0%)
3	4 (8.5%)
4	17 (36.2%)
5	26 (55.3%)

How likely are you to recommend Saeed Mega Mart to a friend or family member? (Rate from 1 to 5, where 1 is very unlikely and 5 is very likely)
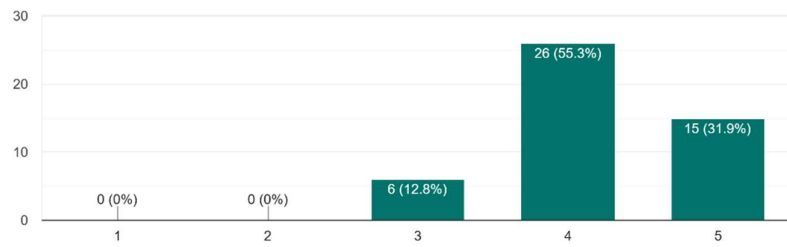47 responses
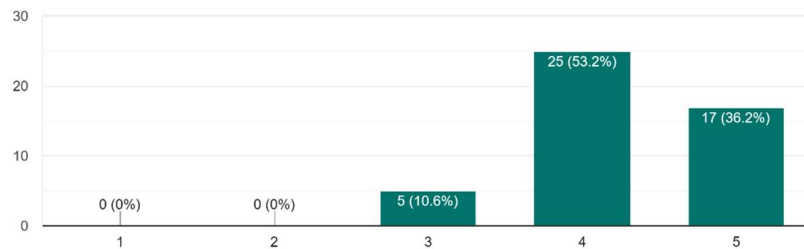
Rating	Count
1	0 (0%)
2	0 (0%)
3	5 (10.6%)
4	25 (53.2%)
5	17 (36.2%)

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

- Bangles — 29.8%
- Skincare products — 7%
- Gift items
- Child Clothing — 12.3%
- Cosmetics — 15.8%
- Female clothing — 17.5%
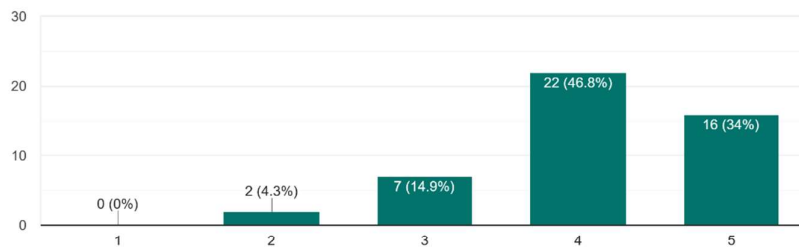- Electronics
- Male Clothing

1/2

15

How do you compare Saeed Mega Mart with other similar businesses in your area? (Rate from 1 to 5, where 1 is much worse and 5 is much better)
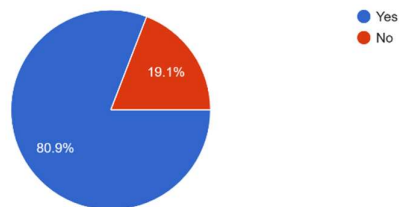
47 responses

How loyal do you feel to Saeed Mega Mart? (Rate from 1 to 5, where 1 is not loyal at all and 5 is very loyal)
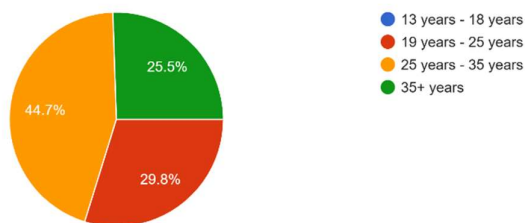
47 responses

Are you an Old Customer?

47 responses

What is your age?

47 responses

Gender
47 responses

● Male
● Female

93.6%
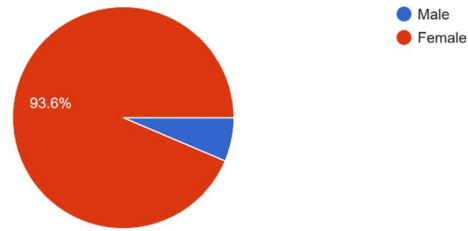
Loyalty Level by Old/New Customer Status

Average of loyalty_level

Loyalty Level by Old/New Customer Status

3.444444444

4.263157895

■ Total

NO   YES

old_customer

Average Satisfaction Level by Shopping Frequency

Average of satisfaction_level

Average Satisfaction Level by Shopping Frequency

5

4.333333333

4.5

4.578947368

■ Total

DAILY   MONTHLY   OCCASIONALLY   WEEKLY

shopping_frequency

Customer preferences:
- o Customer prefers shopping monthly and weekly.
- o Shopping reasons are affordable prices and huge variety of quality products.
- o Majority of the customer is Satisfied with the services and products provided by Saeed Mega Mart.
- o Recommendation likelihood is very high.
- o Most desired products are Bangles, female clothing, cosmetics etc.

Findings

- o Mega mart on comparison with similar business is much better.
- o Most of the customers are loyal to the shop and they are old customer.
- o Majority of the customers are female and belongs to the age of 25+ years.
- o Satisfied customers of Saeed Mega Mart are likely to recommend it, but their satisfaction doesn't strongly influence their loyalty or comparison with other businesses.
- o Customers who rate the Mart better than others tend to be more loyal.

5   Interpretation of Results and Recommendations

The data analysis provides valuable insights into the sales performance and customer preferences of Saeed Mega Mart. Here are the interpretations and recommendations based on the findings:

1. **Sales Performance**: The sales revenue shows a steady increase from 2019 to 2022, with a peak in 2022. However, there's a significant drop in 2023. This suggests that while the sales performance was strong in 2022, there might have been challenges or changes in 2023 that affected the sales negatively.

It's recommended to investigate the factors contributing to this decline and address them promptly.

2. **Customer Preferences**: Customers prefer shopping monthly and weekly, and they are attracted by affordable prices and a huge variety of quality products. The most desired products are Bangles, female clothing, and cosmetics. It's recommended to maintain the product quality and variety, and consider offering special deals or discounts during peak shopping periods to attract more customers.

3. **Customer Satisfaction and Loyalty**: Most customers are satisfied with the services and products provided by Saeed Mega Mart and are likely to recommend it. However, their satisfaction doesn't strongly influence their loyalty or comparison with other businesses. It's recommended to implement customer loyalty programs or offer exclusive benefits to loyal customers to enhance customer retention.

4. **Demographics**: The majority of the customers are female and belong to the age group of 25+ years. It's recommended to tailor marketing strategies and product offerings to cater to this demographic.

5. **Comparison with Similar Businesses**: Saeed Mega Mart performs better compared to similar businesses. It's recommended to continue leveraging its strengths and look for areas of improvement based on customer feedback and market trends.