BDM Capstone Project
Mid-term submission

Data-driven solutions for weak customer footfall at Annapurna Mega Mart

Rajya Lakshmi Jampani

22f3002986

30th June 2023

Table of Contents

Executive Summary

This project is aimed at identifying underlying factors behind crashing sales of Annapurna Mega Mart post relocation to a new outlet. Despite its presence in one of the prime commercial locations of Hyderabad, the store is not able to attract enough number of customers and boost sales.

It is identified that the store is not able to meet customer expectations in terms of a welcoming store-front and the quality of stock being sold. To study the above issues from data analysis grounds, monthly revenue data of the business (past 3 years data is available with the owner), customer footfall data for a period of one month, details of stock being kept on display and the capacity of containers used for the same, manufacturing and expiry dates for a sample of packaged products are collected.

Revenue data is plotted onto line graphs to study trends in sales. From total capacity and used capacity of each container on display, the %Utilization is calculated. It is then plotted onto bar of pie chart to obtain the percentage of empty containers and percentage of under-utilized containers (<25% filled) put on display. Expiry dates are used to calculated time left for stock which is then plotted to pie of pie chart to visualize how much % of stock is expired and how much % of stock is expiring soon (in coming 6 months).

Upon analysis, it was found that yearly sales trend remained same over the period of data collection (2020-2022). However, the average monthly sales recorded a drop of 61.8% in these two years. Weekday-wise customer footfall and revenue data revealed that Wednesdays are the most productive days for the business. Also, as expected, 32% of containers at display are empty creating an impression of No Stock. Stock quality data is under analysis (as of this document's creation date) and insights pertaining to it are yet to be drawn.

Data Collection

Mr. G. Narendar, the owner of the store has been very co-operative all through the course of data collection. He was open to discuss the challenges in his business, keen to understand what data analysis is about and how it could help him. Upon explaining the detailed intention of the project and the proposed approach to address the business problem, he agreed to share the required data and also provided a No Objection Certificate.

Four visits to the store, each one week apart, were conducted as a part of this project to gather information about the background of the business and problems being faced, to explain the objectives of project, to convince the owner to share the necessary data, to share templates for footfall data collection, to collect the raw monthly sales data, utilization data of jars put on display and the manufacturing and expiry dates of products. The customer footfall data being recorded was monitored on a weekly basis.

ANNAPURNA MEGA MART
Jaya Nagar, KPHB, Kukatpally, Hyderabad-500072

Date: 6/6/2023

TO WHOMSOEVER IT MAY CONCERN

This is to confirm that ~~Mr.~~/Ms. J. Rajya Lakshmi has collected data from our store for the purpose of ~~his~~/her BDM Capstone Project.

By this letter, we provide our consent for utilization of data for the above-mentioned project activities at Indian Institute of Technology, Madras. We hereby issue this **No Objection Certificate** on receiving confirmation from ~~Mr.~~/Ms. J. Rajya Lakshmi that the data is intended for IITMinternal-use only and will be disguised appropriately when disclosed for public purposes.

Sincerely,

Annapurna Mega Mart
7to11, Treasures,
Jayanagar Colony, KPHB,
Hyderabad

All the data gathered was loaded into MS Excel. The data cleaning involved removal of certain SKUs from the list, that were extremely old and do not contribute to active stock movement. The data points are then plotted onto appropriate visualization figures (pie charts, bar charts, line charts etc. where ever applicable). The intermediate findings are explained to the business and the

interaction is recorded for evidence. Below are the Google Drive links for Part1 and Part2 of the video. **Interaction Video Part1**

Interaction Video Part2

Metadata

The raw data consisting of five different sheets is placed in the location – **Raw Data**

Sheet 1: **Sales data**

Under Sales data, the monthly revenue details maintained at the store owner are collected. This is presented into two columns – Month and the corresponding amount. The data is available since Jan-2020 till Mar-2023. The revenue generated (in INR) is purely from sales and does not include any expenditures flowing out. This data did not require any cleaning since no anomalies were noted.

Sheet 2: **Month-wise sales**

The sales data collected in the previous sheet is categorized month-wise for identifying sales trends in different months. For each calendar month from January to December listed in rows, 3 columns are created for the years 2020, 2021 and 2022. The corresponding revenue details are obtained from the Sales Data sheet. The data of Jan-2023 to Mar-2023 is ignored for convenience. The annual revenue (sum of revenues of each month of particular year) and the average monthly revenue (Annual revenue/12) across the years are calculated.

Sheet 3: **Customer footfall**

The footfall data at the store was manually captured for a period of one month starting from 01-May-2023 to 31-May-2023. The details recorded include Date, Weekday, No. of customers who purchased something from the store on the day, Total sale amount of the day, No. of customers who left empty handed on the day and the items they asked for (i.e., the items the store failed to make available to the customer). On days when the store remained closed, a zero is recorded in no. of customers and sale amount columns.

Sheet 4: **Display Container Utilization**

This is the data related to the first variable of interest. The containers put on display are manually observed one-by-one and their contents are noted. Also, the total capacity of the

container (in Kg) and the quantity present in it (in Kg) are noted. For empty containers, the name of the content is noted as NA. The %Utilization of each container is calculated using (Quantity present/Total Capacity)*100. At the end, Total number of jars, number of empty jars, number of utilized jars are calculated for analyzing.

Sheet 5: **Stock Quality Data**

To capture this data, few of the packaged SKUs from the store are selected. Their corresponding categories (like Household items, Snacks, Cookies, Chocolates etc.) are manually identified and noted. The manufacturing and expiry dates (mmm-yyyy), quantity available of the SKUs are collected by inspection and listed down. Using the above data, the freshness of the stock being sold will be analyzed with parameters such as % of stock expired, category of items corresponding to the expired stock, % of stock expiring soon and the category of items in it etc.

Descriptive Statistics

With 39 data points from Jan-2020 to Mar-2023, the revenue of Annapurna Mega Mart stood at an average of ₹10.64 lakhs per month. Out of the total 39 months, only 17 months generated above average revenue. The other 22 months, which correspond to 56.4% of total duration, recorded below average sales. The highest revenue was ₹21,49,219 which was observed in April 2020. A grouped frequency distribution table shows that 14 months recorded a revenue of 5-10 lakhs (Mode of the sales data). With a range (max sales – min sales) of ₹18.71 lakhs the sales data shows a high standard deviation of ₹4.93 lakhs and an even higher Inter Quartile Range (IQR) of ₹8.55 lakhs.

Upon categorizing the full list of sales data into month-wise sales across years, a monthly average of ₹16.94 lakhs in 2020, ₹10.4 lakhs in 2021 and ₹6.46 lakhs in 2022 are observed. This shows a drop of 61.8% in sales during the two-year duration. Year-wise standard deviations of revenue were almost consistent at ₹1.77 lakhs in 2020, ₹1.55 lakhs in 2021 and ₹1.6 lakhs in 2022. Range (max – min) of sales was ₹7.07 lakhs in 2020, ₹5.53 lakhs in 2021 and ₹5.71 lakhs in 2022.

A total of 256 customers visited the store in a period of one month (May 2023). The maximum number recorded per day was 13. Total sales for the month were ₹2.23 lakh which approximates to ₹873 per head. Only 8 customers left the store without any purchase and the items they asked for did not repeat.

From the Display Container data, out of total 38 jars, 12 jars are kept empty which contributes to around 32%. Total available capacity is 135.25kg out of which 25.3% i.e., 34.3kg is used. Best utilized containers have Mint, Green Elachi (100% used) and Edible Gum (90% utilized). Frequency Distribution Table of %Utilization data shows that Mode of the data is 23 containers that are 0-25% filled. With just 38 data points, a standard deviation of 31.22 is observed in the Utilization patterns showing a wide variation in the usage of containers.

Under Quality Data, a collection of 90 SKUs from 21 different categories is selected as the sample of study. Sum of their individual quantities gives 597 out of which Rala Spices (Under Spices Category) have maximum quantity (30) available. It is observed that 37 SKUs have crossed their expiry dates already and 31 SKUs are about to expire in another 6 months. Quantity wise expired SKUs contribute to 210 out of 597 (35%) and about to expire SKUs add up to 277 out of 597 (46.3%). When studied age-wise, 42.88% of the stock is less than 6 months old.

Analysis Process

The sales data (Month, corresponding revenue) which was collected from the store owner was loaded into MS Excel. An overall sales trend was plotted as a line-chart. The trendline of the plot was also included to see how fast the revenue is dropping. Average monthly revenue, standard deviation of the data, IQR of the data are found using the built in functions AVERAGE, STDEV and QUARTILE available in Excel. No. of months with above average sales and number of months with below average sales are also calculated using the COUNTIF function of Excel with appropriate criteria.

This data is then segregated year wise. Average monthly revenue of each year and total revenue of each year are calculated using AVERAGE and SUM functions. The data is then plotted onto line chart with 3 lines, one representing each year. Also, the revenue is calculated quarter wise by summing up corresponding monthly values. The percentage drop in revenue for each quarter is also calculated. This data is plotted onto a combo chart (bar + line) with one bar for each quarter and a line representing % change of revenue.

From the footfall data, a pivot table is created by grouping the customer count and revenue values weekday wise. This data is used to draw a pivot combo chart (bar + line) with bars representing day wise footfall and line representing day-wise revenue. It was observed that

Wednesday contributes to maximum footfall and maximum revenue among all days of the week. The number of customers who left empty handed was also plotted onto a bar chart.

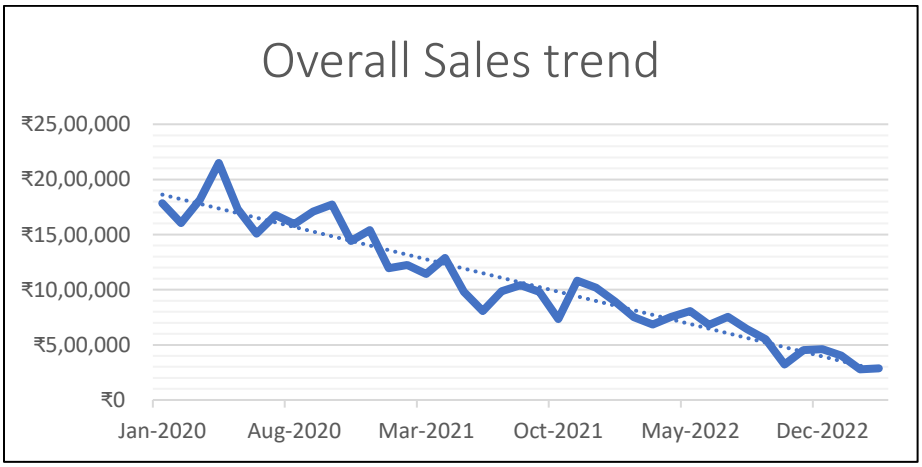
The Display Container Utilization contains Contents of the jar, Total Capacity and Utilized capacity of each jar put on display. From these data, %Utilization is calculated and a frequency distribution table with classes of Utilization% as 0-25%, 25-50%, 50-75%, 75-100% is formed using the FREQUENCY function. The frequencies thus obtained are also measured as fraction of whole (in decimals). With this information, a bar of pie chart is constructed with pie chart representing % of Empty and Utilized jars and bar of Utilized pie representing % of Utilization.

Under the Quality Data head, SKUs along with their categories, manufacturing date, expiry date and quantity available are listed. Age of the SKU (Today – Mfg Date) and Months to expire (Exp Date – Today) are calculated for the items. The Stock Quality analysis is under process and the analysis method is as follows:

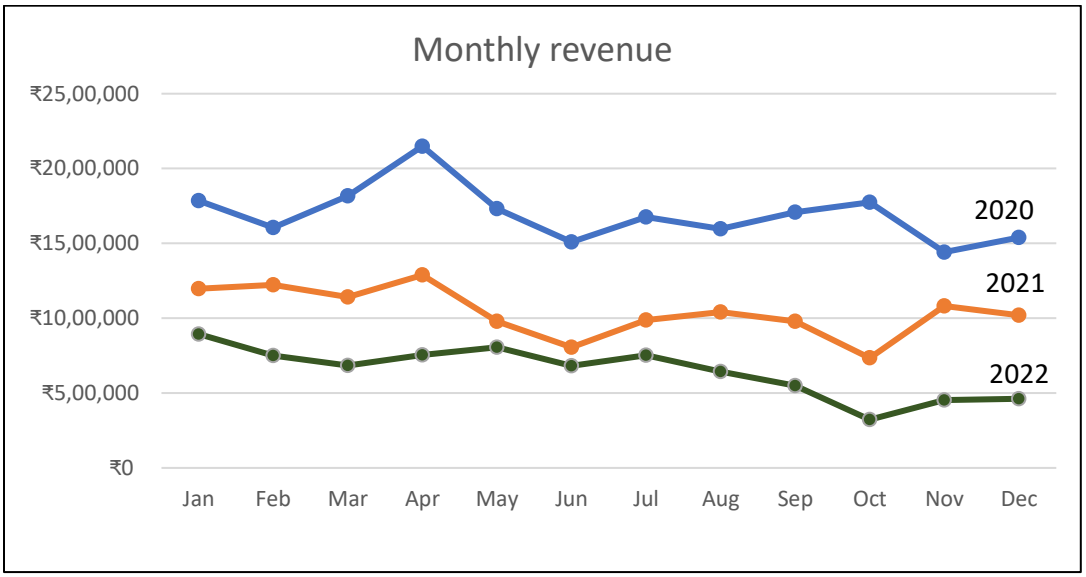
A pivot table with months to expire and the no. of SKUs will be constructed a pie chart will be drawn out of it to know how much % of SKUs are Expired, 0-6 months left, 6-12 months left and >1 year left. Out of the expired stock, a pivot table showing category of items and their quantity will be formed. This data helps break down which categories occupy more % among the expired stock and the stock expiring soon. This analysis will be useful in identifying what items to clear soon and what items to replace so that fresh quality stock will be available to customers.

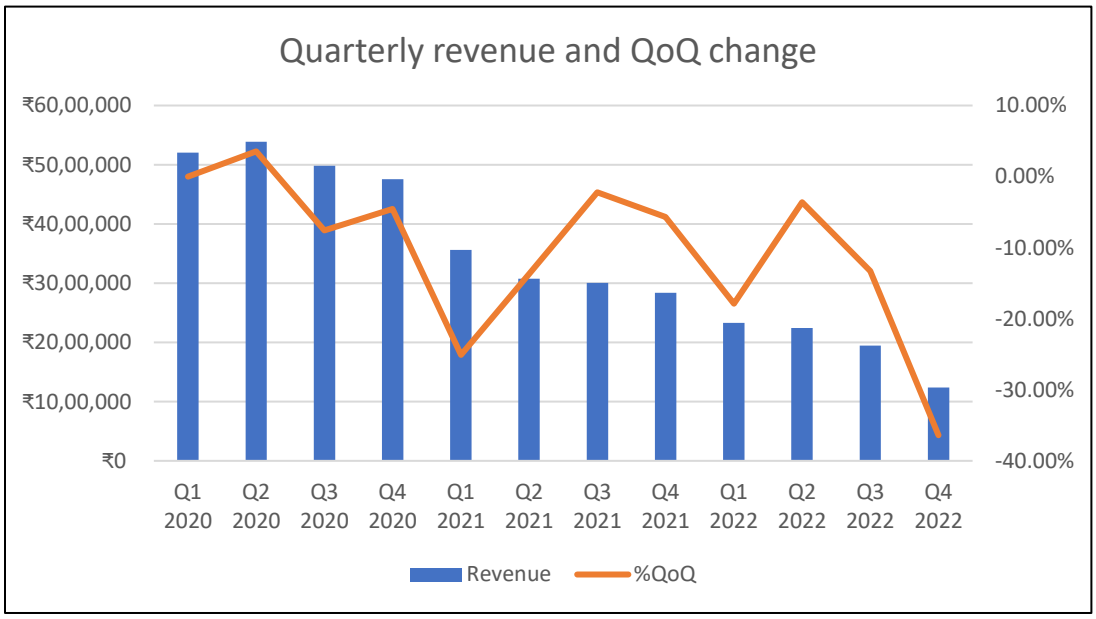
Results and findings

The overall sales graph shows a drastic fall in the revenue generated by Annapurna Mega Mart within a short span of two years.



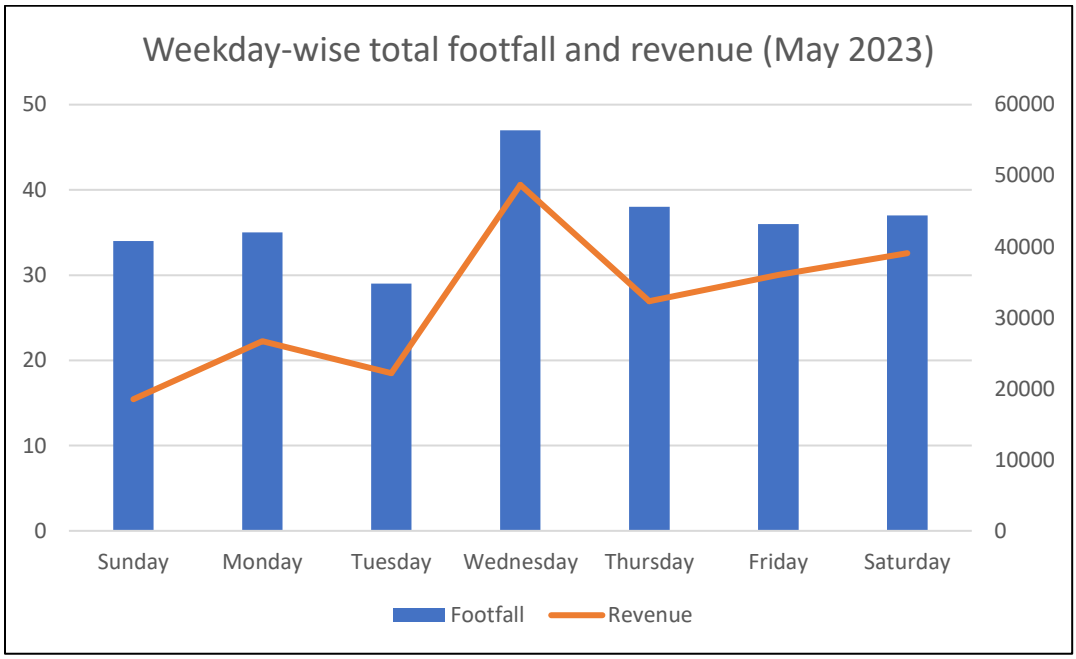
The linear trendline moves downwards with a negative slope indicating a fall in the revenue. The monthly revenue plot inserted shows year-wise revenue trends for the years 2020, 2021 and 2022. From the graph, it is evident that the trend of sales remains consistent across the years. In any year, highest sales are recorded around the months of April and May and least sales around October and November.



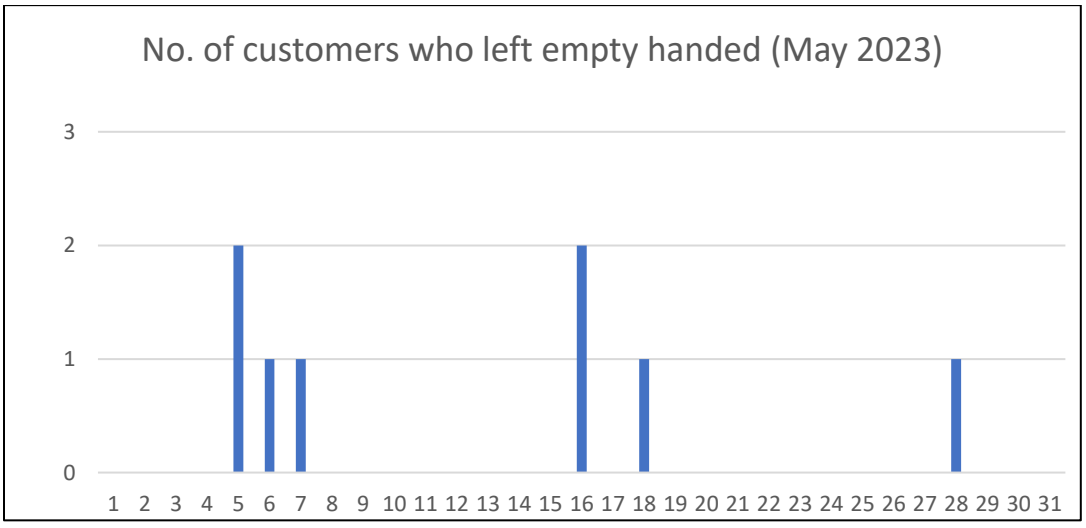


From the quarterly revenue plot, we see that during Q2 2020 there was a slight increase (3.54%) in sales. All other quarters recorded falling revenues. Taking the initial duration of available data, Q1 2020, as the basis, the percentage change of revenue is also plotted on secondary

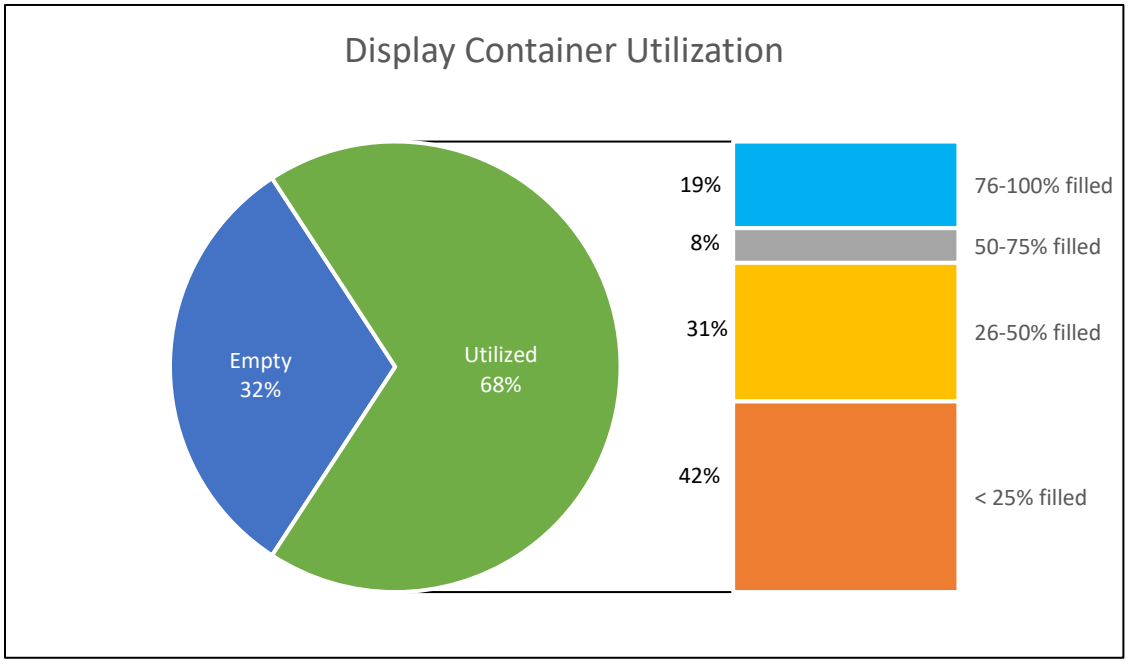
axis. It shows that the drop was considerably high from Q4 2020 to Q1 2021 (-25.08%) and even higher from Q3 to Q4 of 2022 (-36.39%).



The weekday-wise total footfall and revenue for May 2023 are plotted together on two axes as above. It shows that Wednesday is the most productive day for the business both in terms of footfall and revenue. Tuesday marks least numbers in terms of both the parameters. Footfall as well as revenue looks consistent towards the end of the week. Also, the second half of any week is producing more revenue than the first half. The graph showing number of customers who left empty handed do not seem to add much of the insights as such instances are very few in number. This can be taken as a good indicator of least possible stock outs in the business.



The Display Container Utilization analysis is the crucial aspect in identifying how the store is failing to attract customers. It is evident from the data that the store has neglected arranging the display to a good extent post relocation. Out of all the containers put up on display counter, 32% are left empty. This will create an assumption of No Stock and will demotivate customers from entering the store. The corresponding combo chart is inserted below.



Even among the utilized containers, a major portion of 42% are under-utilized with only <25% filled. 31% are moderately filled with 26-50% of volume utilized. 8% of the containers have 50-75% of their capacity utilized. Only 19% containers have good amount of stock in them with 76-100% utilization.

With these analyses, it can be strongly recommended that the store definitely needs to rework on arranging the display better so as to create a good a store-front there by increasing customer footfall leading to improvised revenue.