Optimising the Purchasing System of a Small Shop

Business Data Management Project – Midterm Report

Anirudh Gautham

21f2000384

IIT-M Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai,

Tamil Nadu, India, 600036

Table of Contents

1. Executive Summary and Title

Sun Agencies is a small shop in the RA Puram area of Chennai, which primarily sells stationary and snacks. The shop is owned and run by Mr. Murugesh. The shop also offers Xerox and courier services which are out of the scope of this project, which will focus on commoditties.

For this project, a comprehensive analysis of the shop's purchases data from November 2023 to February 2024 was conducted. As stated in the proposal report, the primary challenges faced by Mr. Murugesh were maintaining adequate levels of working capital, allocating money among different products and identifying optimal inventory levels among the various products. This report reveals findings from the four months of purchases data and seeks to address the aforementioned problems.

This analysis involves poring over the data and performing several functions, mostly in Excel, allowing me to gain a good understanding of the nature and scale of purchases over the four months concerned. Understanding the data and then being able to describe it properly is a crucial step in being able to address the problems highlighted above and achieve the objects of the project, as explained in detail in the Proposal report.

I computed several metrics relating to purchases, made several charts to visually desribe the relationship between the parts of the data and created tables to express information in a consise and readable manner.

2. Proof of Originality of Data

The data used for this project was collected with the explicit consent of Mr. Murugesh. He periodically gave me the bills he collected for his purchases. The data covers the purchases made between November 2023 through February 2024 and the first four days or March.

To prove the authenticity of the data used for this project, I have provided the following:

Interview with owner along with transcript	Link
Pictures of the shop	Link
Letter signed by owner allowing data to be used	Link
Pictures of a few purchases bills	Link
Excel Sheet Containing collected Data	Link

3. Metadata and Descriptive Statistics

3. 1. Metadata

Each row of the Excel sheet which records the purchases describes the purchase of an item. The variables/columns in the data are:

Column	Description
Date	The date on which the item was bought – ranging from 3/11/2023 – 4/3/2024.
Item	The item that was purchased – chips, ice cream, pencil, etc.
Quantity	The quantity of the item purchased.
Rate	The unit price at which the product was bought.
Discount	A discount, if any, offered by the vendor on the unit price.
Total	The total amount paid for that product, taking the quantity, rate and discount into account.
Category	The category into which the product falls – snacks, stationary, household items, cooking ingredients and entertainment.

Table 1: Metadata

- Item refers to a particular product like Chips or Pens and the brand is not recorded since the pricing and characteristics of day-to-day products tend to be similar across brands.

- Quantity isn't recorded as individual units or pieces that customers buy. Instead, it refers to boxes or packages containing several units that wholesalers sell to retailers.

- If **x** units of a product are bought at price **p** with discount rate of **d%** given, the total amount paid to the supplier is: **x * p – (d/100 * x * p)**.

- The categorisation of products was done by me based on the products that were purchased.

3.2 Descriptive Statistics

The following table describes some basic metrics, giving us a simple overview of the data collected.

Metric	Number
Total Purchases	₹57,716
Average Purchase Value	₹198
Median Purchase Value	₹148
Minimum Purchase Value	₹17
Maximum Purchase Value	₹771
Range	₹754
Number of Unique Products Purchased	61
Average Quantity Purchased	9 units
Average Per-Unit Price	₹100

Table 2: Descriptive Stats Summary

The following charts and graphs present a more nuanced view of the purchases data.

Stationary and snacks occupy the lion’s share of purchases. Together, they take up around 91% of the total purchase value.

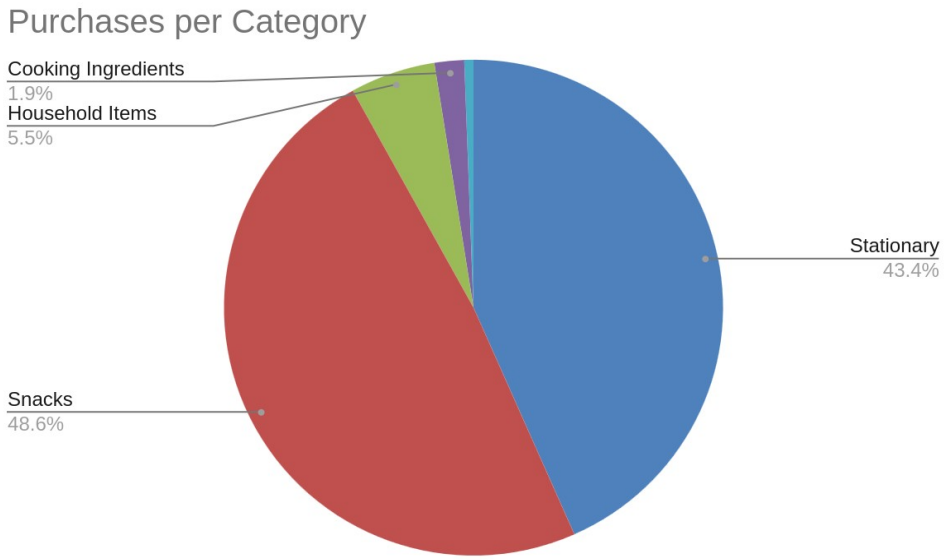

Figure 1: Purchases by Category

However, while stationary and snacks account for a similar proportion of the total purchases, there is a wide disparity in the quantity purchased between the two of them, with stationary alone accounting for more than half the total quantity purchased.

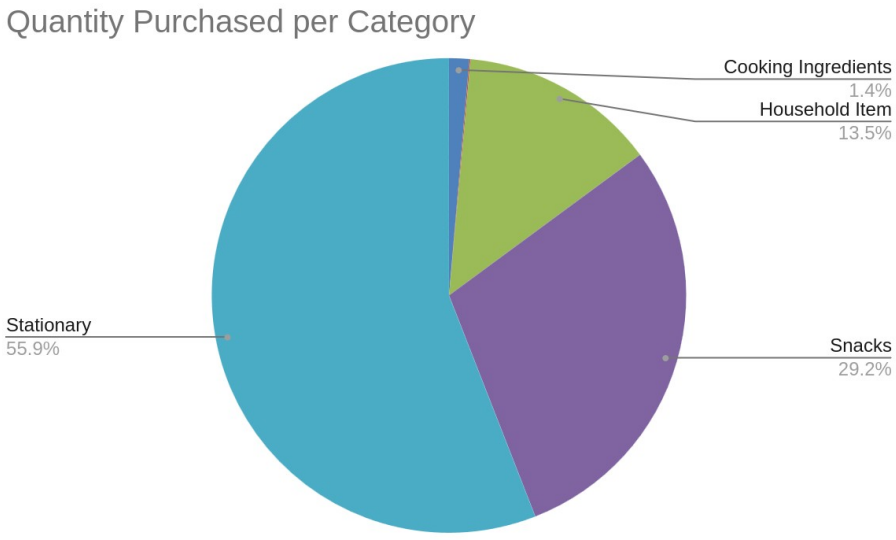

Figure 2: Quantity Purchased Per Category

Further inspection into the snacks and stationary categories shows us differing trends. Purchases of stationary took a sharp dip in December, presumably because of school and college closures due to the Chennai floods and Christmas holidays. Snacks sales peaked in February, marking the onset of

Summer and increased demand for Ice Cream and Soft Drinks, which are the two most dominant products in that category. The following chart displays the total purchases for both categories for each month.

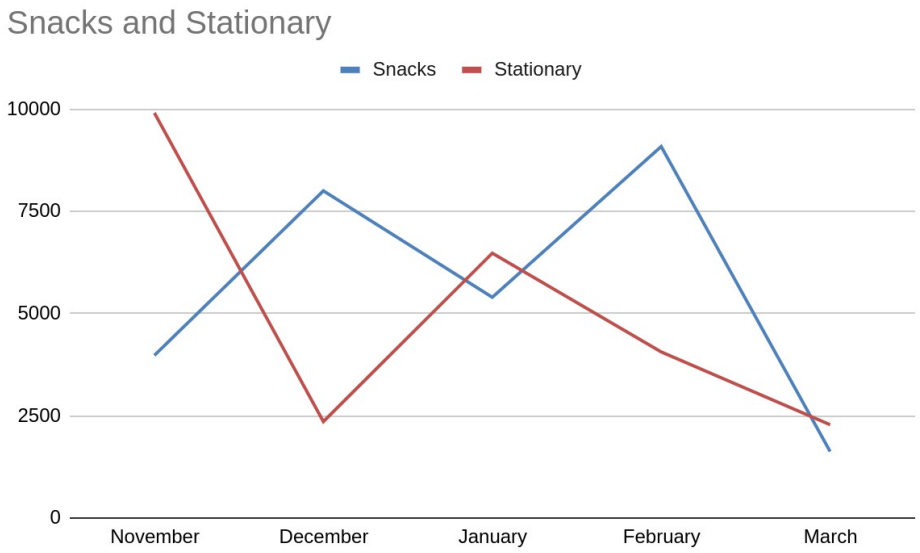

Figure 3: Snacks vs Stationary

Looking at the overall purchases, across products, timewise, November recorded the highest purchase value among all products, followed by a dip in December, due to the floods in Chennai. Post the floods, activity picked up again. Please note that data was collected for only 4 days in March, due to which it's tally appears deceptively low.

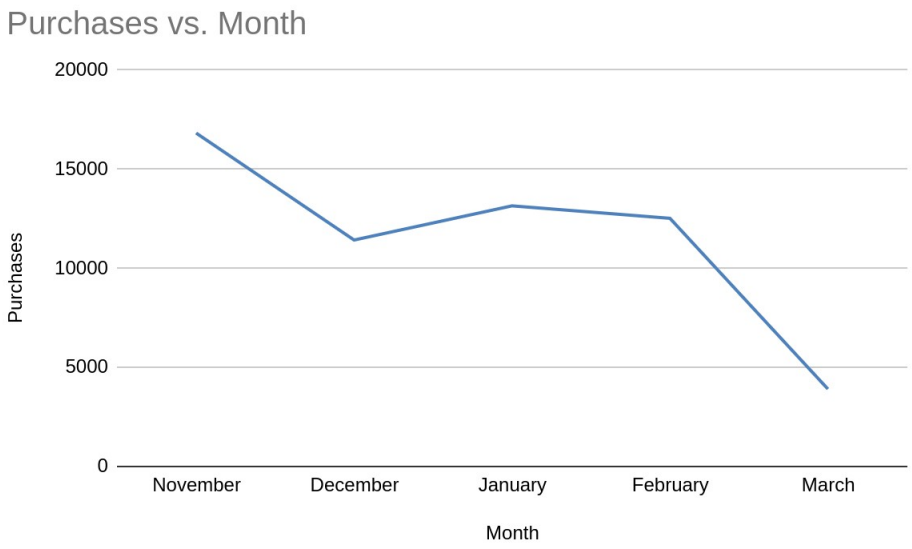

Figure 4: Purchases by Month

Looking at the top 10 most purchased products, Ice cream takes the proverbial cake here, with Pens and Soft Drinks being purchased in ample quantity too.

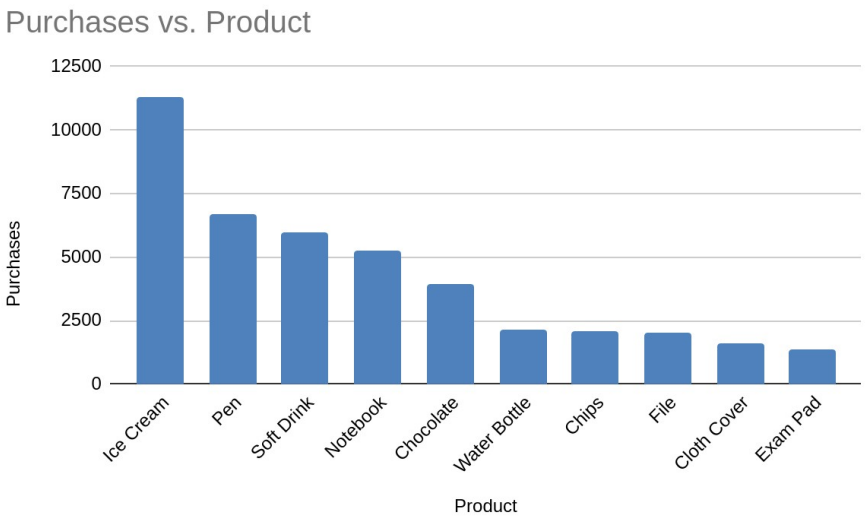

Figure 5: Top 10 most purchased products

Further analysis of the product-wise purchases shows us that these 10 products account for around 74% of total purchases, meaning that of 61 different products purchased, about 16% of them account for three-quarters of the total purchases.

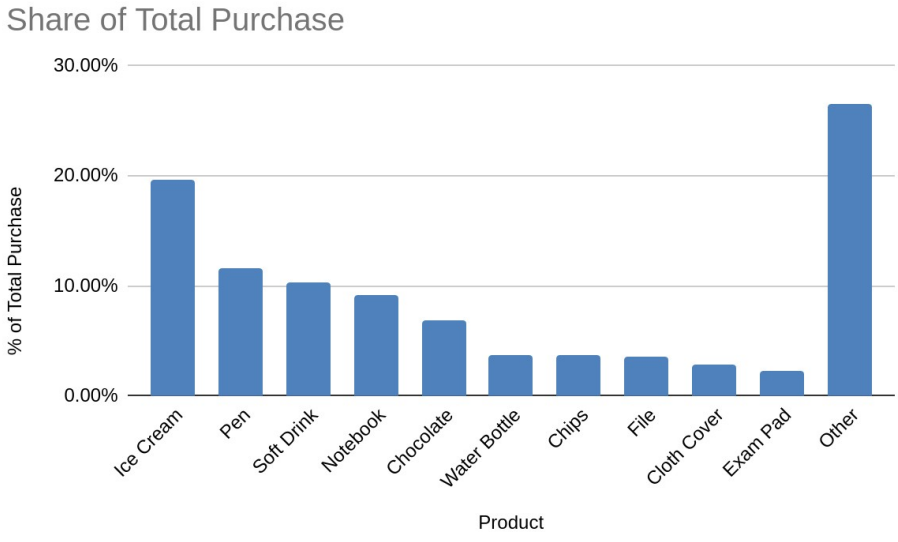

Figure 6: Share of total purchases by product

4. Explaination of Analysis Process and Method

The precursor to the analysis process was periodically collecting purchases bills and entering them into an Excel sheet to make them organised and useful for analysis. I added each transaction of each bill one-by-one into the sheet. The entry of data into a sheet was a rather time-consuming process. Once the data was cleaned and transformed into a useful form, I was able to work on the analysis.

The first step in analysing the data was deciding the various categories of products. What is the ideal level of precision in the categorisation? For example, should ‘snacks’ be split into individual categories like Ice Cream, Chips and Chocolate? Similarly, should stationary be split into Pen, Pencil, Paper and so on? Ultimately, I decided to go with a relative broad scheme of categorisation since I felt it would be easier to perform comprehensive analysis. To a large extent, people’s consumption of major stationary and snack items, for example, tends to be quite similar across products.

Once the data was cleaned and categorised, I conducted Exploratory Data Analysis (EDA) on it. This involved applying filtering and aggregation methods on the data. I calculated some basic metrics such as the total purchases over the period, average purchase amount and so on, giving me a sense of the scale of purchase activity and the kind of amounts with which I am dealing. I also looked at several entries of the raw data to check if the averages and summaries provided a good representation of the data.

The next step in the analysis process was to review the relationship between the different variables. This was done using visual analysis methods. Visual methods of analysis allow to understand relationships and descriptions of the data in a manner that is intuitive and understandable to us. Without much effort, we can gauge several important details regarding the data.

For this, I primarily used Line charts, Bar Graphs. Line charts were used since they allow us to trace how a variable trends over time. In this context, understanding the monthly and seasonal changes in purchase levels is a crucial part of this project. Bar Graphs enable us to make comparisons between different subgroups and categories of data. I used it to compare the purchase levels of products, to understand which products were being purchased the most. Allocating investment optimally among the various products is one of the objectives of this project. Pie charts are usually used to describe how a whole is divided among its parts. Since the purchases data is divided into five categories, I used pie charts to understand how the total purchases was distributed among the various categories.

After gaining a comprehensive understanding of the data using numbers, tables and graphs, I set out to interpret the data and make meaningful inferences that would be useful in decision-making. This is described in detail in the next section.

5. Results and Findings

The findings from the data show us that the bulk of expenditure is on stationary and snacks. The other three cateogries comprise around 9% of expenditure and in fact, under the ‘entertainment’ category, there is only item, which is a pack of playing cards. Buying items in such low quantities is arguably not a good strategy. People usually associate shops with specific product lines. Those who are looking to buy cooking ingredients like oil and milk and household items like detergent would not go to a stationary shop looking for them, especially when their stock is very low and hardly visible in the shelves.

When we look at the split between the categories in terms of quantity, we see that stationary’s tally is almost twice that of snacks, which means that stationary items of nominal prices are bought in large quantities. This could be used to foster relationships with suppliers and gain larger discounts. A quick look at the raw data shows that discounts are currently given very rarely.

Looking at how overall purchase levels changed over time, we see the dip in December due to the floods in Chennai, which damaged the shop badly due to knee-height water and caused substantial loss. Such calamities are arguably unavoidable in a low-lying area with poor infrastructure. With rains being a prominent part of Chennai weather, such events are inevitable and steps should be taken to minimise their damage.

The product-wise distribution of purchases is possibly the most interesting part of the analysis. As seen in *figure 5* and *figure 6*, an overwhelming majority of purchases are driven by a minority of products, closely mirroring the famous Pareto Distribution. When I raised this with Mr. Murugesh, he remarked that sales too, follows a similar pattern. However, the demand for some of the dominant products like Ice Cream and Soft Drinks can fluctuate based on the season, not to mention that the cost of maintaining them due to their perishability and the need for refridgeration, is rather high.

Overall, one can notice that Mr. Murugesh is running a diverse business, as opposed to most other shops that focus on one ‘category’ or product line. This is on top of the courier and xerox business which are major components in their own right. On the one hand, this could be a good strategy since being a one-stop-shop for all customer needs can create loyalty and combined purchasing. On the

other hand, since the shop isn’t operating at large scale, it’s possible that the sheer diversity of categories and products undercut each other and focusing on area would be a better tactic.