Optimising the Purchasing System of a Small Shop

Business Data Management Project – Final Report

Anirudh Gautham

21f2000384

IIT-M Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai,

Tamil Nadu, India, 600036

Table of Contents

Executive Summary

This Capstone project is title “Optimizing the Purchasing System of a Small Shop”. It focuses on Sun Agencies, a small shop in RA Puram area of Chennai that primarily sells stationary and snacks. The shop is run by Mr. Murugesh a spirited, longtime entrepreneur who has built the shop from scratch and has been consistently expanding it despite facing enormous difficulty due to his health problems and natural disasters, which tend to hit low-lying areas like RA Puram quite hard. Beyond the aforementioned items, Mr. Murugesh also runs a courier business, as an affiliate for Franch Express and tries to get the parts of his business to complement each other.

For this project, the purchases' data of the business, from November 2023 to Februrary 2024 was used as the basis for analysis conducted on the business. This data is used to analyse the purchasing activity of the business, the weakpoints of the purchasing system and provide recommendations that could be implemented to improve the purchasing mechanism.

To address these issues a comprehensive analysis of the purchases data was conducted. The data itself, contains a variety of columns which describe the various details related to a purchase, such as the product, quantity purchased, price and so on.

The analysis was conducted primarily using Excel/Google Sheets and Pandas, a Python library that enables us to work with tabular data and Matplotlib, another Python library, used for visualising data. Various features within Excel, such as charts, graphs, tables, pivot tables, sorting and filtering functions, numerical functions, etc., were used while a variety of Pandas functionality was used as well.

The majority of charts and graphs used this report were generated using the matplotlib library.

Please refer to this link for the raw data used for the analysis as well as short interview with the business owner along with some pictures of the shop.

Explaination of the Analysis Process and Method

The analysis of the purchases data that culminated with this report largely consisted of two parts:

Discussions with Business Owner

Once Mr. Murugesh agreed to collaborate with me on this project and was willing to share his purchases data with me, we had several short meetings where we spoke about the various parts of his business, its volumes, the challenges he faces and his future plans. While my learnings from these discussions weren't based on any data, they acted as useful priors for analysis that was conducted once the data was available. Large amounts of data can be intimidating to deal with and having a clear vision and thought process is crucial.

Collection and Cleaning of Data

Before performing any analysis, the data had to be collected into a neat, usable form that would be analysed. Mr. Murugesh does not maintain records of his purchases activity. This meant that the purchases' activity had to be inferred manually from the bills given by suppliers and entered into an Excel sheet. Once all the transactions were added, I then categorised them into the following categories – Stationary, Snacks, Entertainment, Cooking Ingredients and Household Items.

The final form looked like this:

	A	B	C	D	E	F	G
1	Date	Item	Quantity	Rate	Discount (if any)	Total	Category
2	03/11/23	Sketch Pen	20	11.46	0	229.2	Stationary
3	03/11/23	Sketch Pen	10	7.5	0	75	Stationary
4	03/11/23	Sketch Pen	50	2.33	0	116.5	Stationary
5	11/11/23	Zip Lock Cover	1	80	0	80	Household Item
6	11/11/23	Zip Lock Cover	1	90	0	90	Household Item
7	11/11/23	Milk Flavouring	1	56	0	56	Cooking Ingredients
8	11/11/23	Detergent	0.5	114	0	57	Household Item
9	11/11/23	Coffee Powder	2	72	0	144	Cooking Ingredients

Figure 1: Snapshot of Data

Exploratory Data Analysis (EDA)

Once the data was available in a useful form, I was able to explore and play around with it to get a feel of nature and volumes of the purchases' activity. For both purchases volume and purchases value variables, key metrics such as mean, median, mode, standard deviation, range, minimum, maximum, sum, and count were calculated. These metrics enable deeper insights and well-informed decision-making. This metadata is provided in detail and elaborated upon in the midterm report.

Some visual exploration was also conducted, where I used Excel features such as bar graphs, pie charts and histograms to explore total purchase volumes, trends across time and so on.

Comprehensive Analysis of Data

The previous step involved largely cursory looks at various aspects of the data without getting into the nitty-gritties. Once I got a sense of volumes and trends, both numerically and visually, I delved deeper into the data, seeking to perform analysis that would yield useful insights that could be used by the business to improve its purchases system.

There are multiple dimensions along which the analysis could be conducted:

1. Analysis of categories, individually and through comparisons.
2. Analysis of individual items, individually and through comparisons with other products.
3. Analysis of pricing, quantities and discounts, which were embedded into the above 2.
4. Time-series analysis: analysis trends in data over time.

All the analysis was conducted using Python and Excel.

This step allowed me to bridge the gap between merely learning about the exploring the data and using this analysis to make useful and actionable inferences. It involved making comparions across categories from various perspectives – distribution of total purchases, relationship between categories, analysis of individual categories and products and so on.

Recommendations and Changes to Implement

As mentioned before, exploring the data and learning it comprehensively is one thing, but being able to use these learning to provide actionable recommendations to the business is another. Early on in the analysis process, I found myself learning a lot but unable to see how this analysis could be useful for the business and how these learnings could be transformed into action statements and improve the purchasing system of the business. As time went by time, I refined my thought process and started looking at everything from the perspective how it could help the business.

For example, if we look at the share of various products in the total purchases of the business, what can this tell us about the product mix in the future? Should the product line be narrowed, expanded or kept the same? If we look at how snacks purchases trended over the data collection period, how can these trends be used to make better decisions during these months in the next year?

Finally, I was able to come up with a number of recommendations for the business, which can hopefully improve the purchasing system of sun agencies.

The results of all the above steps are elaborated in detail from here on. I found this project to be rather enjoyable, even if frustrating at times, and it was wonderful to get an opportunity to work with a business in real life and spend some in the real world and interact with entrepreneurs I would like to emulate someday.

Results and Findings

The analysis on the data was done along different lines as mentioned in the previous section, each of which is elaborated upon in this section. This analysis can reveal important trends regarding the purchases of various products, categories, quantities and so on. This can help in addressing and coming up with solutions to the business problems mentioned in an earlier section. The results of this analysis served as the basis for the recommendations made to the business.

Category-Wise Findings

In this section, we explore the inferences made when the data is analysed category-wise and each category is assumed to be a single entity. As mentioned earlier, there are six categories in total – Stationary, Snacks, Entertainment, Cooking Ingredients and Household Items.

First, we start off with some analysis of the category-wise distribution of purchases. The chart below shows us the share of various categories in the total purchases over the period for which data was collected:

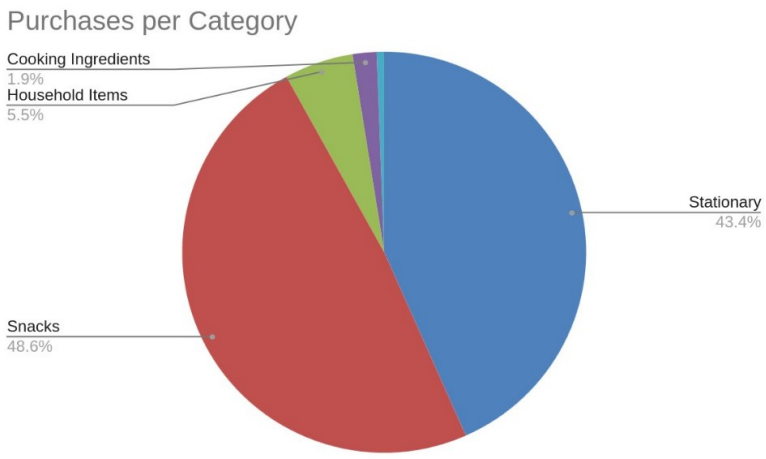

Figure 2: Purchases by Category

As we can see from the chart, stationary and snacks occupy the majority of purchases – around 93% combined.

If we look at the distribution of purchases for each category, we get a more nuanced picture than the above pie chart. This allows us to understand each category and its distribution more clearly:

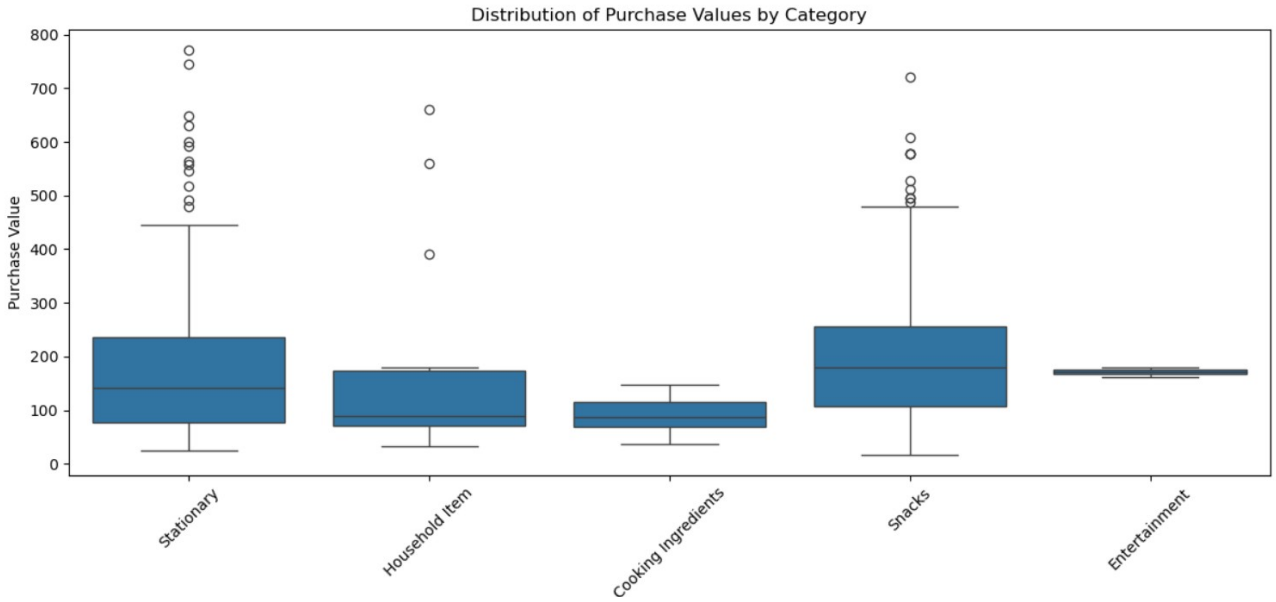

Figure 3: Distribution of Purchase Values

As we can see from the chart, stationary and snacks occupy the majority of purchases – around 93% combine. Snacks has the highest median purchase value and a higher Interquartile Range (IQR) of all the categories. Stationary products have a wider range of purchase values due to the diversity of stationarproducts, which includes ₹5 pens and laminated notebooks. Household items follow an asymmetric distribution with the median purchase value being in the lower end of the range. This is due to the fact the purchases of such items are highly inconsistent and quite random. This is addressed in the next section.

As the pie chart showed, stationary and snacks are the most dominant categories. What about the other three categories which account for 7% or ₹4641 of purchases, a meagre amount for a four-month period? Further inspection into these three categories shows us that a substantial portion of products bought under them are those one wouldn't expect to see in a shop that mostly sells stationary and packaged snacks, like playing cards, detergent powder and knives. This raises questions about the shop's unique-selling point and the kind of customers it seeks to attract. The following chart showcases some of the least-purchased items.

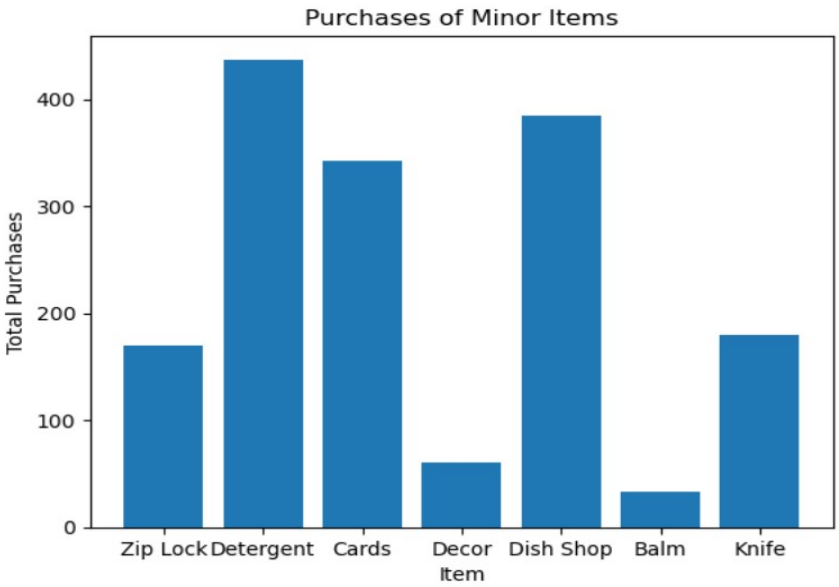

Figure 4: Purchases of Minor Category Products

Analysis of the relationship between the purchases of various categories reveals that there is negative or insignificant positive correlation between them. This is exhibited in the following heatmap:

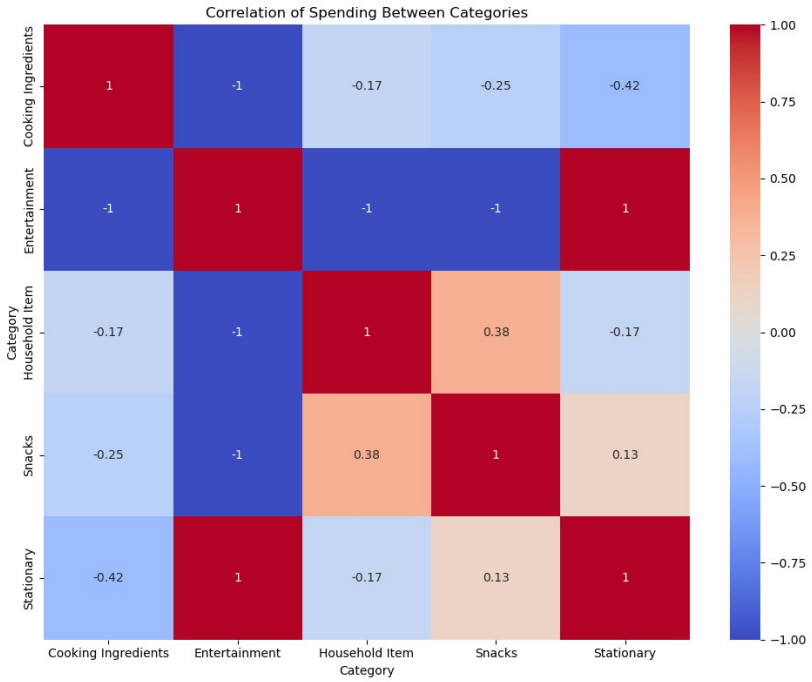

Figure 5: Relationship between Categories

This shows that purchasing decision for particular categories are made separately from the others. Negative correlations suggest that expenditure on one category comes at the cost of others. This further adds to the assertion that Mr. Murugesh should narrow down on his shop's product line and not spread himself as thin as he currently has.

Given that snacks is one of the two dominant categories, let us look at it individually. The following chart shows the change in snacks purchases over the period for which data was collected:

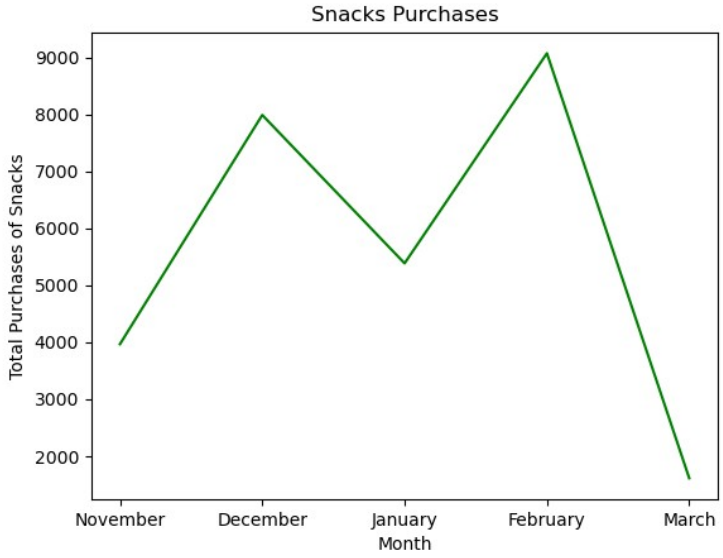

Figure 6: Monthly snacks purchases

We see that the purchases of snacks peaked in February. This is because of the onset of the summer months in Chennai, during which time people tend to consume more of Ice Cream, soft drinks and other cold items that people normally have in cold weather. Please note that data for March was collected only for three days. However, snacks purchases for December doesn't lag too far behind either and if we look at just Ice Cream and Soft drinks, which are the two most commonly purchased snack items and their purchases for December and February, we see an interesting trend:

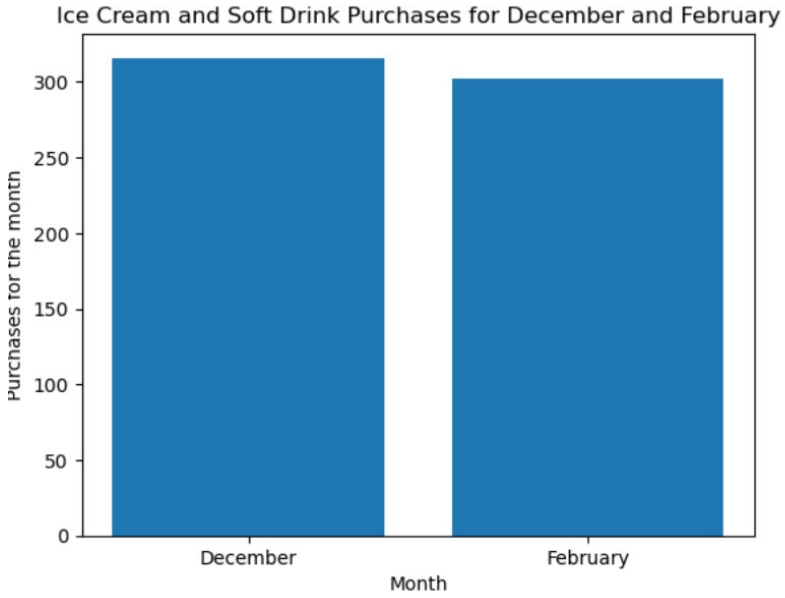

Figure 7: Ice Cream and Soft drinks

The purchases for the two items are almost identical for the two months, with December being slightly higher by a negligible margin of ₹13! This reflects two possibilties:

1. Mr. Murugesh has evaluated that the demand for the two items doesn't fluctuate based on the season and remains high all year. This could make sense given that Chennai is rather hot and humid for most of the year and doesn't experience Winters in the same manner as cities in the North.
2. The demand for the two products does indeed fluctuate based on the season the business is making a mistake purchasing similar quantities of cold items through the year.

Upon discussing the same with Mr. Murugesh, he confirmed that the sales of cold items during the winter months is less than the summer months and that the purchase volumes are sometimes decided arbitrarily leading to a situation like this one. However, it is also possible that December 2023 was an outlier, compared to typical winter times.

Stationary is the other dominant category, and it exhibits the following trend:

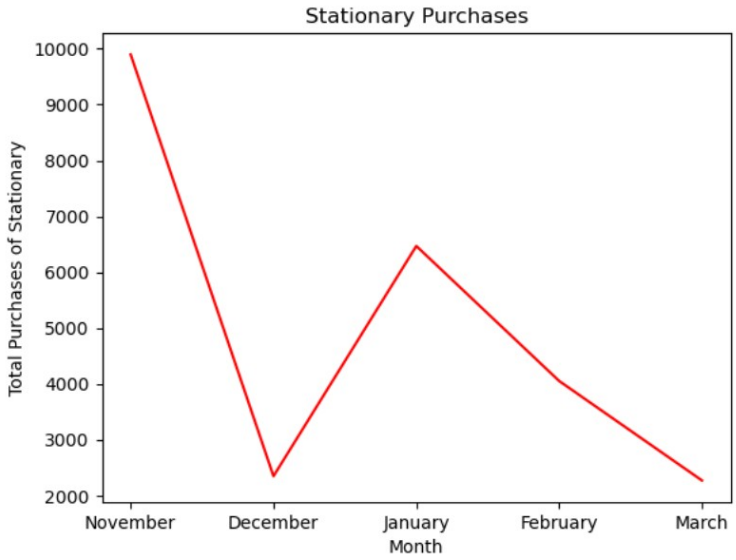

Figure 8: Monthly Stationary Purchases

The purchases of stationary is far more variable than snacks across the period. This is because of two reasons – floods that ravaged Chennai in early December, giving schools several days off and the Christmas holidays, which also closed schools for an extended period of time. The purchases shot up again in February, to meet demand for the January term.

Product-Wise Findings

As was explained in the midterm report, ice cream was the single-most purchased product for the window in which data was collected, by a healthy margin. This chart, taken from the midterm report, shows us the top 10 most purchased items:

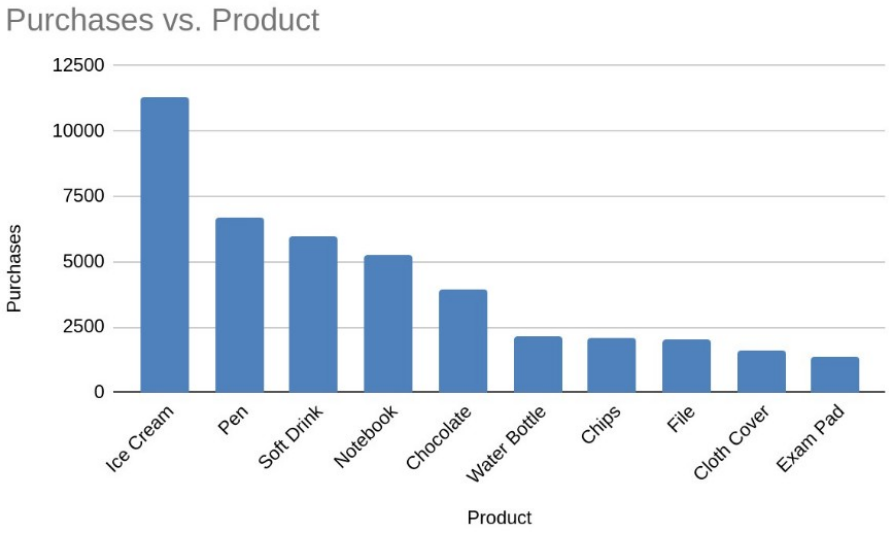

Figure 9: Most purchased products

Ice cream, as expected, accounts for close to 20% of the total purchases while pens account for more than 10%. The 10 products in this chart account for 74% of the total purchases. We will now delve deeper into this product-based analysis, beyond the cursory look we had in the midterm.

The scatter plot below display the total purchases of each of the top 10 most purchased items against how frequently they are purchased. This can help the business optimize how frequently certain products are purchased based on how much is spent on them.

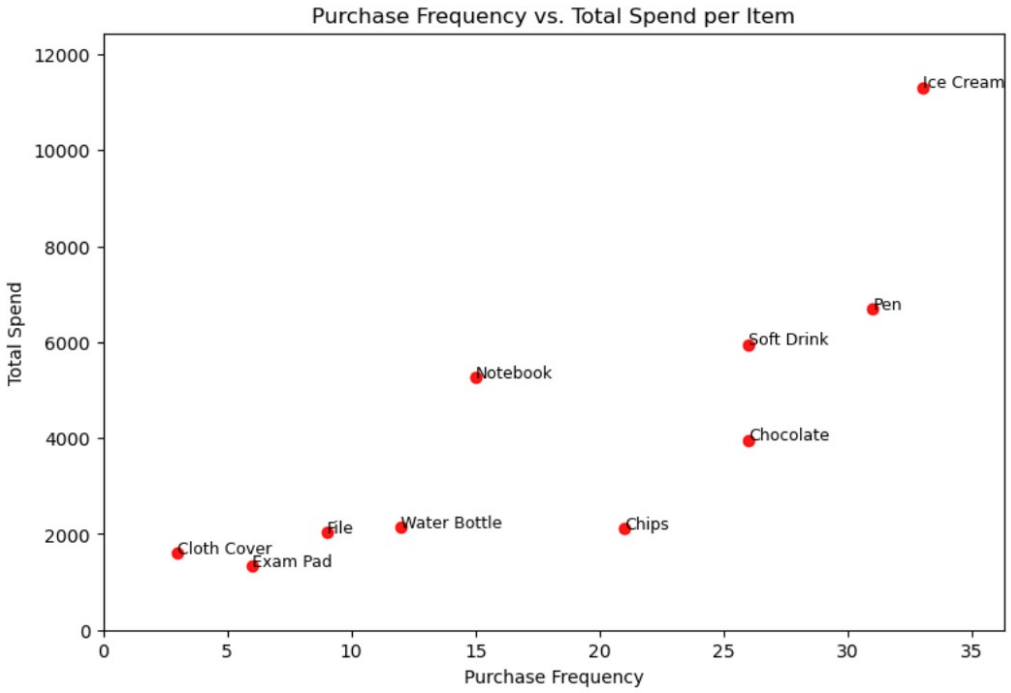

Figure 10: Purchase Frequency vs Total Spend

The scatter plot throws up some interesting findings. Chocolate and soft drinks are purchased as frequently as each other, but total expenditure on chocolates is higher. Mr. Murugesh could afford to buy chocolates at a lower frequency without reducing its availability. Given that he has been nursing a severe leg injury for several years, this could be a welcome relief. This will mean buying them in higher quantities which will also allow him to avail discounts from willing suppliers, which is something that will be touched upon later. Notebooks represent a good benchmark in this regard, in that they are bought relatively rarely, despite having high expenditure. Files, water bottles and chips also have similar total expenditure. Yet, they differ in how frequently they are purchased, despite none of them being perishable items.

Going beyond the frequency of purchases, the following two charts compare the quantities and prices of the top five most purchased items. The top five account for more than 50% of total purchases. Looking at the chart below, which shows the average quantity purchase for each of the top five items, we see an interesting fact – while the total purchases of Ice Cream is the highest by a distance, the average quantity purchased is rather small compared to the other items. A unit of a particular product doesn't refer to a singular item, but instead boxes or packages that wholesalers sell to retailer.

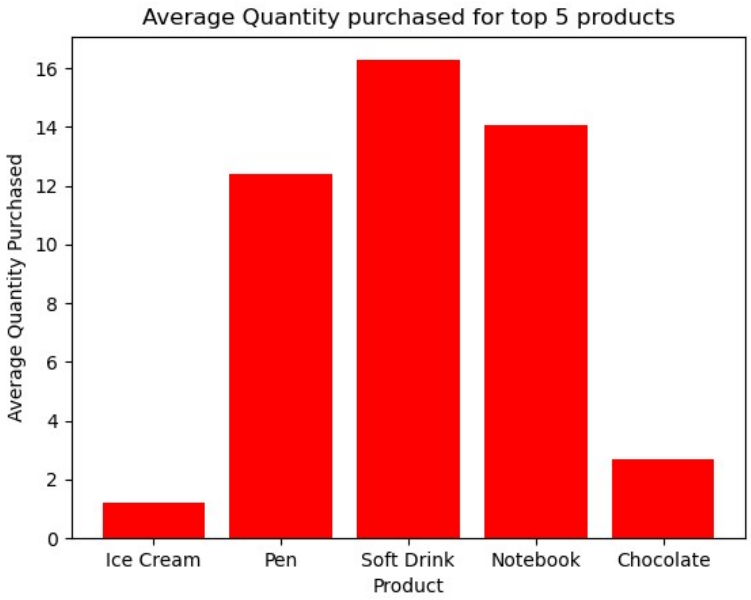

Figure 11: Average purchase quanitity per product

The average price per purchase clears up any possible confusion from the above chart – Ice Cream's average is significantly higher while pens and notebook, which stand tallest in the above chart, are rather low.

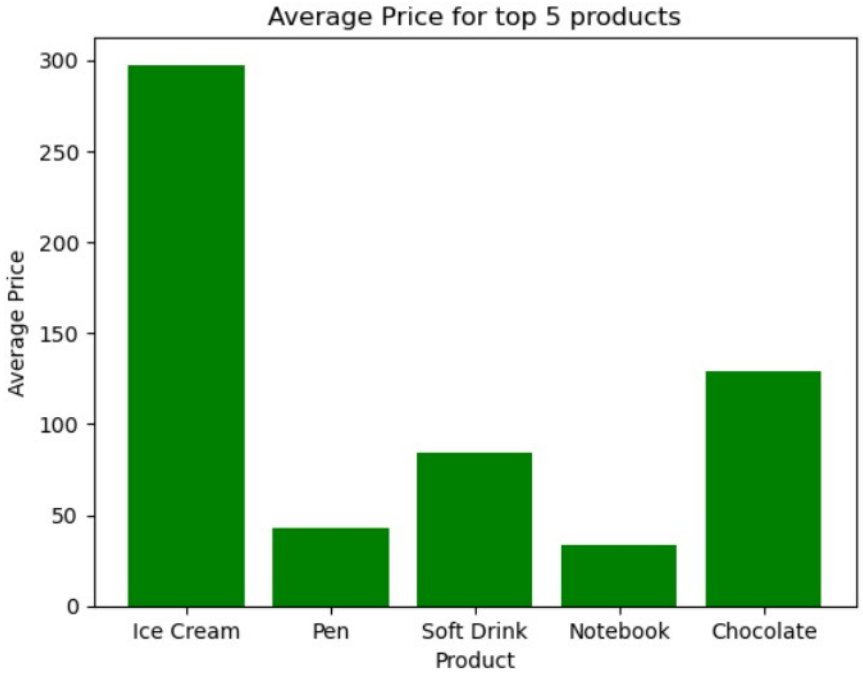

Figure 12: Average price per product

The figures for Chocolates call for some more exploration. Similar to Ice Cream, chocolates too, are purchased in relatively low quantities but are purchased for a high price. This suggests that the chocolates being stocked are the more expensive ones like Dairy Milk Silk, Kitkat and so on. A quick visit to the shop would confirm the same visually. Soft Drinks are purchased in higher quantities and for higher prices than pens. Yet, the total expenditure on pens is more than soft drinks. This is due to the fact that pens are purchased more frequently, as the scatter plot in the previous page shows us.

Let us now analyse the prices at which these items are purchased. The following table displays some basic data on the prices of the top five most purchased items (all amounts are in INR):

Product	Minimum Unit Price	Maximum Unit Price	Average Unit Price	Standard Deviation
Ice Cream	90	720	297	156
Pen	7.5	160	42	32
Soft Drink	4.13	220	84	82
Notebook	4	75	33	19
Chocolate	8	308	129	66

Table 1: Pricing metadata

It is not advisable to use the absolute values of the pricing metrics to make comparisons across products as one unit is defined differently across the board. For example, one unit in soft drinks could mean 4 bottles, while in Ice Cream it could mean 5 cones. In the instance where soft drinks were purchased for ₹4 apiece, 100 units were purchased, meaning that each individual bottle was counted as one unit. However, what stands out is the volatility. The high standard deviation for Ice Cream suggest that prices deviate significantly from the mean, which suggests that suppliers tend to change

their prices frequently. This further backed up by the fact the Ice Cream doesn't have too many varieties, especially when sold at a small shop. This is probably due to seasonal demand for it. Pens and notebooks, on the other hand, deviate relatively less from their means, displaying less volatility.

Analysis of Discounts

One thing that stands out from the purchases data is that discounts were given very sproadically. A short analysis of the same shows us that purchases without discounts are overwhelmingly higher than purchases with discounts. The pie chart below displays this discrepancy by comparing the percentage of purchases made with discounts as opposed to without it.

Figure 13: Purchases with and without discount

This is unfortunate and represents a clear missed opportunity save costs. If we quickly skim through the data, we see that the average purchase quantity is only 9 units. In most instances where discount is given, the purchase quantity exceeds this minumum. The scatter plot in an earlier page also shows that several goods are purchased rather frequently. This frequency could be reduced to increase the per-visit purchase quantity and get higher discounts.

Findings from Cumulative Data

Finally, we move onto to looking at the data as a whole rather than analysing it item-wise or category-wise. This helps us get a sense of the overall purchases activity of the business. The total purchases during the period between November 2023 and February 2024 (and 3 days of March) was ₹57,716.

The following chart shows us each month's share in the total:

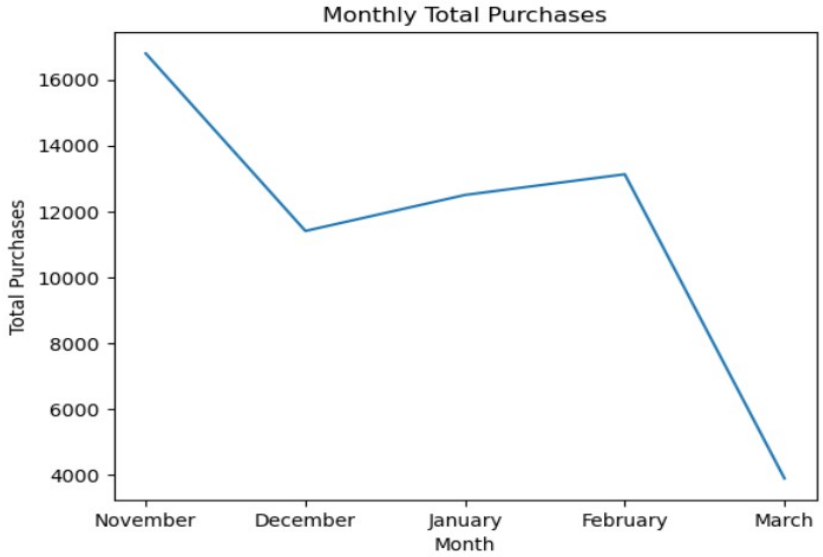

Figure 14: Total Purchases Per Month

November recorded the highest purchases, while January and February didn't lag too far behind. December saw a dip due to the floods in Chennai and school closures.

Moving on, we can observe that purchases are made quite frequently and in small quantities. Each unique date in the date column usually represents a trip to one or many suppliers' places or deliveries from them. Given this, Mr. Murugesh has made a purchase order on 31 different occasions over the data collection period of four months, meaning that he makes purchases approximately every four days. The average purchase amount for each trip is ₹1861 and the average quantity purchased is 87 units, comprising various products. This is arguably too frequent and Mr. Murugesh must make visits to suppliers in a more systematic and organized manner. Such frequent visits to suppliers also impacts sales as the shop is often closed when Mr. Murugesh isn't there and can't find someone to fill in for him. As mentioned earlier, purchasing more rarely in larger quantities could also get larger discounts.

The chart below displays this inconsistency by visualising the total amount spent for each visit to suppliers:

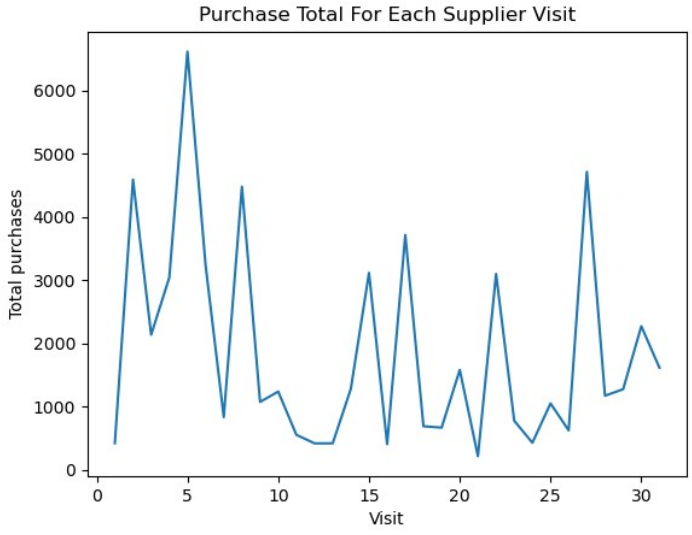

Figure 15: Purchases per visit

Given that demand for products like stationary and snacks are largely predictable and not too volatile, purchases must be made in more consistent levels. This will enable products to be available in a more predicable manner, as opposed to variability in whether a particular product is sufficiently available on a particular day or not.

Interpretation of Results and Recommendations

Diversity of Product Line

Sun agencies is a store that has been steadily expanding for the last decade. Initially, Mr. Murugesh ran the business purely as a courier provider and has since expanded into the stationary and snacks segments. This expansion might've been carried out for arbitary reasons and the business might now be operating without a clear vision on what it sells and the kind of customers it seeks to attract. We remarked earlier that products purchased under the Cooking Ingredients, Household Items and Entertainment categories account for around 7% of total expenditure for the given period. These categories consist of items like playing cards, which one would hardly expect to see in a stationary and snacks shop. Customers usually go to a particular shop having a particular item(s) in mind and map shops to certain products and certain categories. Those who wish to buy playing cards, for example, would not visit a shop like Sun Agencies and those who visit Sun Agencies usually wouldn't be looking by cards.

From our product analysis, we saw that 10 products accounted for 74% of all purchases. This closely mirrors the pareto distribution. The remaining 26% is spread out across several products that are purchased in small quantities. The business must narrow down its product line and focus on a few items like Ice Cream, Soft Drinks and some stationary items and double down on the purchases of these rather than spread itself thin by selling several different products.

Building Relationships with Suppliers

Another integral part of optmizing the business' purchasing system is to build a network of high-quality, trusted suppliers to provide the items that are frequently purchased. This enables the business to avail higher and more frequent discounts. It also standardizes the purchasing activity of the firm since all activity is conducted with a select group of known suppliers. Such standardization can also curb some of the pricing volatility we observed earlier with products like Ice Cream and chocolates as these products are often sold at varying prices by different suppliers and in differing quanitites.

As an aside, since Mr. Murugesh also runs a courier business, he could use that to build supplier relations. One way to do this would be to offer discounts on courier to suppliers who supply goods beyond Chennai. Not only will this help build good relations with suppliers, it will also scale up the courier business and allow it to cater to a new type of customer.

Bulk Purchasing

A recurring theme throughout this report has been the low volumes at which business is conducted with suppliers. Products are purchased rather frequently and in low quantities. Beyond the physical

strain associated with frequent trips, this compromises the business' ability to avail discounts due to the low quantities and imposes a high overhead cost on the business due to the frequency in which purchases are made.

By consolidating these small purchases into larger, less frequent orders, the business can:

a) Potentially negotiate better prices due to higher volume.

b) Reduce the administrative costs associated with frequent ordering.

c) Minimize delivery costs.

It is important to note the limitations of this. Several goods purchased by the business like Ice Cream are perishable and the space available in the shop isn't a lot, which is amplified by the sheer diversity of products being sold. This places constraints on the quantities at which purchases can be made since products need to last for a long time and space will need to be available at the given instance.

Negotiate Consistent Discounts

We saw from an earlier section that discounts were obtained sproadically, abeit on a variety of products and categories.

Obtaining higher discounts, to a large, extent, depends on the previous two points – building sustained relationships with suppliers and purchasing goods in large quantities. This presents a standard opportunity to reduce costs.

Seasonal Planning

From our analysis of categories and products, we saw significant differences in the purchases of various products based on the season. Stationary, for example, experienced a significant blip in December. Snacks, similarly saw a dip in January. By reading and analyzing these fluctuations in the future, the business could:

a) Place orders in advance of peak seasons and later in lull periods.

b) In the process of doing the first point, the business could avail better prices by committing to purchases earlier. Suppliers often provide concessions to guaranteed business.

For example, as we near the end of 2024, the business could purchase large amounts of stationary in October, given that purchases will spike in November. Suppliers will raise prices when the demand increases, meaning that the business will be able to obtain goods at lower prices before they shoot up during peak times. Such planning will also give Mr. Murugesh a competitive advantage over other shops in the area as they often do not analyse their activity with the same depth.

Consolidated and Standardized Purchasing System

Tying all the above points together, Mr. Murugesh must develop a comprehensive system which governs purchasing activity. The data exhibits a reactive approach to making purchases such that goods are purchased in arbitrary intervals and non-standard quantities. A systematic approach to purchasing will ensure that orders are placed periodically, in the right quantities and with a trusted

group of suppliers. Reducing chaos from the purchasing process will also enable the business to high-quality data on its purchases, which can be used for future analysis to fine-tune the system.

Recording and Collection of Data

One of the biggest weakpoints in the business the lack of data being collected and recorded. Mr. Murugesh maintains no record of his purchases and sales activity. All the data used for this project was taken from disjointed purchases bills, which would've been disacarded if not for my request to keep them for this project. Ideally, a tally of all the purchases and sales that are made should be kept in an Excel file or a dedicated notebook that will enable the business to get a clear picture of its financial health, volumes, profitability and efficiency. For a business that is seeking to expand beyond the domains covered in the project, this is paramount.

Maintaing records of historical data can enable analyses like this report to be performed frequently and with a larger trove of data.

When I raised this point with Mr. Murugesh, he remarked that at the very least, maintaining records would help future students with their college projects!

Periodic Review and Planning

Along with creating and maintaining records of business activity, Mr. Murugesh must periodically spend sometime every week or so to review the business' performance and introspect on how the business performed and rectify the problems it faced over the period.