Executive Summary

This project of Business Data Management aims to study the organization of a vegetable and fruit vendor, give a holistic view about their organization, identify and understand their business problem(s) and finally provide viable solutions obtained after thorough analysis. To commence with this report, we describe the background and nature of the business where we give an overview about the organization and its management.

Next, we define two problems which were bought up by the owner, problems which the business has been encountering and not been able to tackle by their regular course of action. The problems being reducing wastages in the business's supply chain of stock and understating the sales pattern of the business and how to optimize it.

After that, we explain our problem solving approaches, which we believe to be fit to solve the said problems. We aim to use the concept of Pareto Analysis and Average days of Inventory, so as to analyse and provide apt solutions to the business. We also explain the data we intend to collect and the tools we use to facilitate our analysis. We conclude by the expected timeline of this project, along with the expected outcome we aim to find.

Organization Background

The organization in this project is of a local vegetable vendor in our locality, namely *Santosh Vegetables and Fruit Supplier*. Its owner is Santosh, who has been in this business for the past 8 years, in the locality of Punjabi Bagh, New Delhi. Along with him, he also has his brother co-managing the store. Apart from both of them, they have employed two people as well, to help them in the delivery and transportation of stock.

Their business model is that of a regular vegetable vendor i.e. B2C. Every morning, they purchase new stock for the day from the wholesale market and aim to sell it to consumers by the ned of the day. They sell on credit to consumers, but do not purchase stock on credit. Being from the unorganized sector, they collect data primarily in book/account keeping application in their phones, with inwards and outwards of each day.

Problem Statement

The problem statement for this project is listed below, as follows:

1. I intend to identify and rectified major wastages in monetary terms, as well as raw stock in the business, over the course of the given timeline of the data to be analysed.
2. Along with this, I aim to see if there are ways in which more customers can be reached and taken in as regular customers in long term.

Background of the Problem

There were two meetings between me and the organization. In the first meeting, primarily the discussions were about how the business would only keep data of the stock they would purchase in the start of the day, and the total sales for the day. He would do this via a monitoring application and general book keeping.

Santosh expressed his concerns about how over the past few months, after looking into his accounts and expenses, he found unexpected outcomes. He was liable to pay a huge amount to his creditors, despite his earning and expenditure patterns not changing over that period. Apart from that, there was also a discussion about how some kind of strict monitoring about expenditures needs to be done, so as to identify these wastages.

In the next meeting, Santosh explained how after the pandemic, the reliance of consumers on the online delivery systems such as *Blink It, Swiggy Instamart etc*. has increased rapidly. In such a situation, it is important for the organization to increase their consumer base, if they intend to beat the inflation and keep their profit margins the same, if not grow. All in all, these two problems were brought up regularly by Santosh.

Problem Solving Approach

I will be using the concept of Pareto Analysis on the revenue and volume of the total sales done in the timeframe of 3 months. This analysis follows the Pareto Principal – that states 80% of a projects/organizations benefits or results come from 20% of the work/efforts/items. This implies the particulars which have the biggest payoff. If a business's revenue and volume of sales approximately follows this pattern, it can be assumed as the business heading in the right direction. This is not a guarantee that resources are being used optimally, however, it is a decent indication about the track of earnings of a business.

I also intend to find average days of inventory for each SKU, which will tell how many days does the stock of each item/SKU stay in storage, before being sold. Through this, I can understand the supply chain dynamics of the business, and how it can be optimized.

I intend to collect primarily three sources of data – revenue, volume of sales and opening stock details for each day. The total source of income each SKU (Stock Keeping Unit) of item generates is an important measure that must be taken in account, as well as the volume of each SKU sold each day. With both these data sets are taken into account and analysed, we can;

- Identify and prioritize problems
- Organize workloads efficiently
- Improve productivity and profitability

Apart from this, the details of opening stock and closing stock are of vital importance to us. This data will give us an idea about the average days of inventory. In the nature of perishable goods, it is very important that wastages are minimized, as storage options are limited.

I will be using Microsoft Excel as a tool for this project. We use Excel as a tool as it can perform a plethora of functions, which will be quite beneficial to the analysis. Pivot tables can be used to find the cumulative particulars for both revenue and volume, which are crucial to find the pareto analysis. We can also visualize our results by the help of scatter plots, where we can clearly see the results of our analysis, at a glance.

We can find average days of inventory, by which we can understand how many days can the average opening stock can last, for each SKU. Also, we can visualize those results as well, by the help of bubble chart.

Expected Timeline

For this project, I will be collecting data for 3 months from the organization. This projected is expected to conclude within the next two months i.e. end of November 2022.

Expected Outcome

With this project, I expect that the analysis will tell us whether the pareto analysis is being followed or not. By that, we can understand the SKUs which our supposed to be worked upon more, as they would be yielding more results. Also, we aim to understand the average days of inventory, so as to realize the supply of which items is supposed to be increased/decreased according to the sales pattern, so as to improve productivity and minimize unnecessary wastages.

This project also aims to help the business with some insights so as to how they can upscale their business by the help of marketing and optimum use of online resources. The outcome of this project lays heavy importance to help the organization compete with online vendors and increase their consumer base as well.

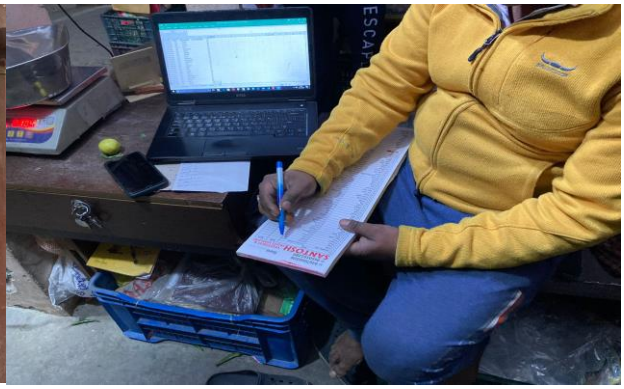
Executive Summary

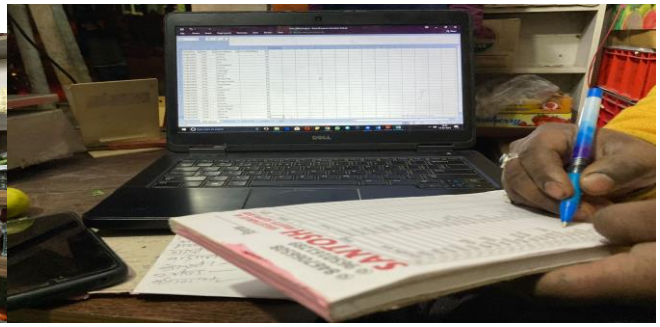
As we move ahead in the project, we define the primary raw data with the help of Metadata section, which gives path to the breakdown of the data in such a manner that it helps us see a story and work on the same. The latter is achieved in the section of Descriptive Statistics.

In our collected data, the SKU Master sheet contains information on the various unique items sold by our vendor, as well as their average price per kilogram. The Sales Data sheet shows the total sales of a given SKU on a day, for a total of two weeks (14 days). With that, we move into the Stock Data sheet, which describes the cumulative total of opening and closing stock of each SKU, along with the purchase and sales for that day for each given day. These sheets are elaborated further in the project, which support us to analyse the data and infer valuable insights from the same.

As far as our analysis approach goes, our primary tool would be PIVOT tables, which help us integrate and explain our data better to solve our problem. Along with that, we also use functions like VLookup and Charts to enhance our analysis process.

Proof of Originality



An interaction with the shop owner, Santosh. Here, we have collected the primary data from his books and monitoring application, Okay Credit in his presence. A video interaction with the owner can be seen on the link given below:

https://drive.google.com/file/d/1mi0LalKemARnLZ_2iq64GGgUtKRErZWm/view?usp=sharing

Metadata

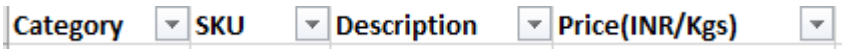
In our data, we have rounded of the decimal to one decimal point throughout our dataset for common and easier understanding. Our data can be found in the link below:

https://docs.google.com/spreadsheets/d/17KNJOzHCDOQ7XpL4df0817C2c7n8asKNPvF1GQ6ZpNE/edit?usp=sharing

In our main data workbook, there are three sheets which comprise of all the primary data collected from our vendor, which are as follows:

- **SKU_Master**: This sheet holds all the information of the various unique items sold by the vendor, along with their price per kilogram (INR/Kg)
- **Sales_Data:** In this sheet, there is the information on total sales in a day for each given SKU, with the units being in kilograms (Kgs)
- **Stock_Data:** With this sheet, we can see the opening and closing stock of all SKUs, as well as the purchases made on each day.

1. **SKU_Master**



In the work sheet SKU_Master, the metadata of the column header mentioned above is described as follows:

- Category: This describes the category of the given SKU, as our vendor deals in both Fruits and Vegetables, which are the two distinct categories for all our SKUs.
- SKU: Each item is given a unique alphanumeric code, so as to identify the said category and item. Kindly note that these are user defined values and not a part of the

primary data given by the shop owner. It is only to make our understanding and analysis more interpretable.

- Description: This defines the item as it is sold to the consumers i.e. the name of the fruit/vegetable.
- Price(INR/Kgs): This column contains the price in Indian National Rupee per kilogram (INR/Kg) for the said SKU.

2. Sales_Data

In the work sheet Sales_Data, the metadata of the column header mentioned above is described as follows:

- Date: This column contains the date for which the sales entry is being recorded
- Day: This column contains the day for which the sales entry is being recorded
- SKU: Each item is given a unique alphanumeric code, so as to identify the said category and item. Kindly note that these are user defined values and not a part of the primary data given by the shop owner. It is only to make our understanding and analysis more interpretable.
- Sales(Kgs): This column contains the sales made in one day for the given SKU, the unit for this column is Kilograms (Kgs).

3. Stock_Data

In the work sheet Stock_Data, the metadata of the column header mentioned above is described as follows:

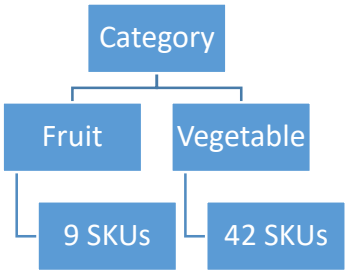
- Date: This column contains the date for which the stock entries are being recorded
- Day: This column contains the day for which the sales entry is being recorded
- SKU: Each item is given a unique alphanumeric code, so as to identify the said category and item. Kindly note that these are user defined values and not a part of the primary data given by the shop owner. It is only to make our understanding and analysis more interpretable.

- Opening_Stock(Kgs): This column defines the opening stock in Kilograms for each given SKU which was present with the shop owner at the beginning of each day.
- Inward_Stock(Kgs): This column contains information on the purchases made in Kilograms for each given SKU at the beginning of each day.
- Outward_Stock(Kgs): This column contains information on the sales made in Kilograms for each given SKU by the end of each day.
- Closing_Stock(Kgs): This column defines the closing stock in Kilograms for each given SKU which was present with the shop owner at the end of each day.

Descriptive Statistics

With further interactions with Santosh, the basic description of the data we have, was a particularly interesting one at that. He explained how people living in apartments/flats preferred their vegetables as fresh as possible. However, that did not mean that the vegetables/fruits went to waste or unused at the end of that very day. Located in a slum area, Santosh also had a recurring demand from people in his locality for vegetables even if they were a few days old, given that it was winters. Apart from that, restaurants also would buy a day or two older vegetables with relative ease.

So, on an average, mostly all vegetable SKUs had a shelf life of 2-3 days and on the other hand, the SKUs of mostly all fruit category had a shelf life of 2-3 days. Kindly note that these are estimates from the shop owner and not part of an analysis. This can be used as a prior belief to be incorporated with our analysis at a later stage. As far as our SKU distribution goes, it can be understood better from the flowchart below:



If we see the data, the average volume of kilograms of vegetables and fruits traded is **6.27 Kgs** each day, which gives us an estimate that of all the items in the business inventory, only a small portion is traded. As they would increase their growing opportunities, they would clearly need to invest in cold storage (such as a fridge) as this number would be going up.

Also the variance in sales of all products over the given time period is roughly **2.40**, which shows that the sales are fairly consistent and homogenous in nature. This gives us an indication that there is a presence of regular customer with recurring demands in our business.

With this information in hand, we can now proceed further to understand in detail the steps we will take to do our analysis and draw useful insights from the same.

Detailed Explanation of Analysis Process/Method

For our analysis, the critical tool of use has been Pivot Tables, VLookup, Filters and sorting feature, which has helped immensely in understanding the data in a categorised manner. The above inferred data has led way to representing our data in a pictorial manner, which gives us a clear and direct understanding/description of our data.

With the help of Pivot table, we are able to extract a cumulative for sales made throughout the given time frame, for each SKU. Post that step, we sort the values in descending order. We also find the cumulative percentage contributor of our SKUs in terms of volume of sales. This enables us to get a sense of what the required demand consists of across all given SKUs. The result is then depicted on a clustered column chart to check for Pareto principle.

Following the above, we also intend to do the same for our revenue to analyse that aspect as well.

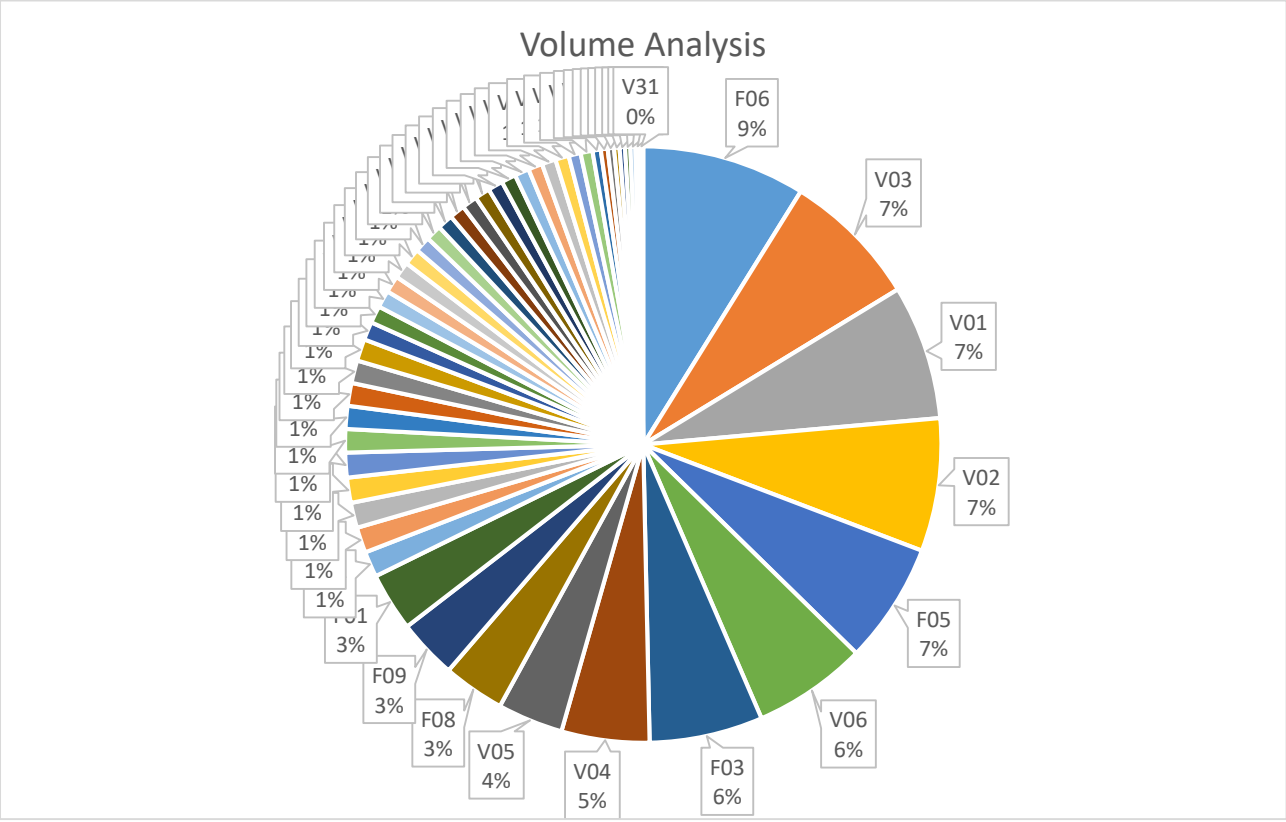
For that, we pick up the price column from our SKU_Master sheet and use it in Sales_Data, a step which is done with the help of VLookup function. After that, we create a new column in our Sales_Data sheet, labelled ‘Revenue’ which is computed as the product of price and sales for each SKU on a given day.

The above steps were used with the previous strategy to create a revenue analysis displayed on a column chart. With that, we could understand which SKUs contribute to our revenue more so than others and whether they tend to follow Pareto Principle or not.

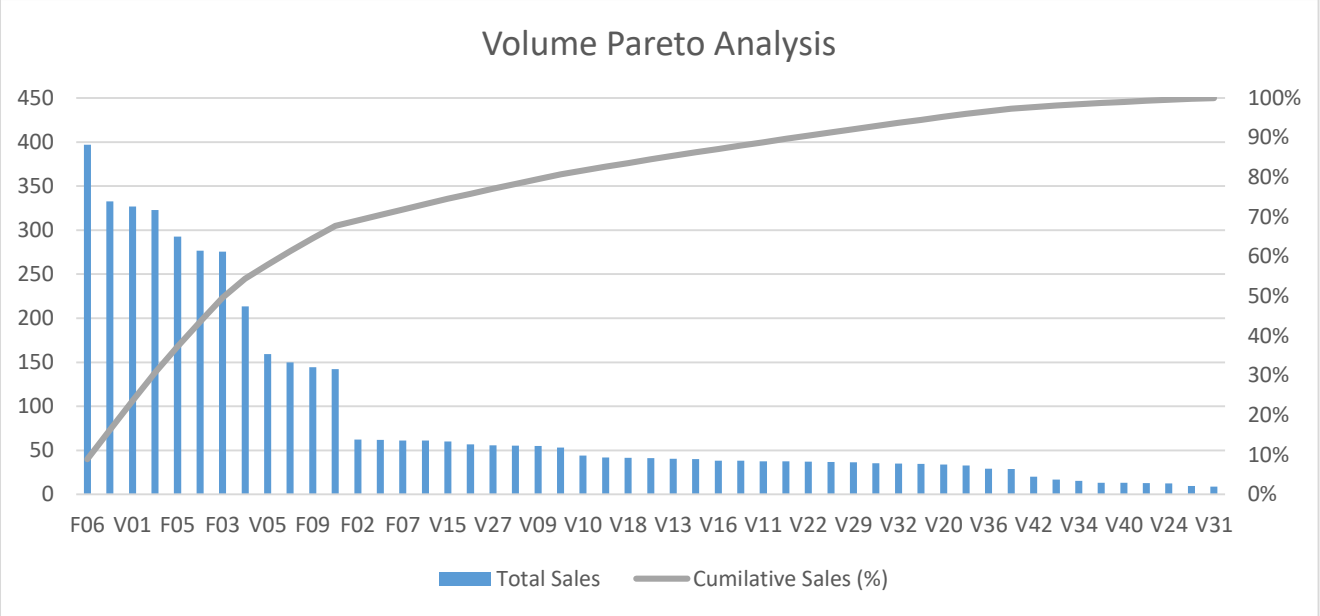
It is important to understand that the Pareto principle is not a business rule carved in stone. It is an observation which indicates how well a business is performing with a subset of items. If it is not followed or overshoots the 80/20 principle by a lot, it is just an indicator that things are possibly not going in the correct direction i.e. some deliberate actions need to be taken so as to how we organise our business SKUs in a more efficient and effective manner.

Also, we like to identify the SKUs which we could pay a key focus on as we move ahead in our project, so as to understand how to work on SKUs which satisfy our aim to improve on both revenue and volume optimisation.

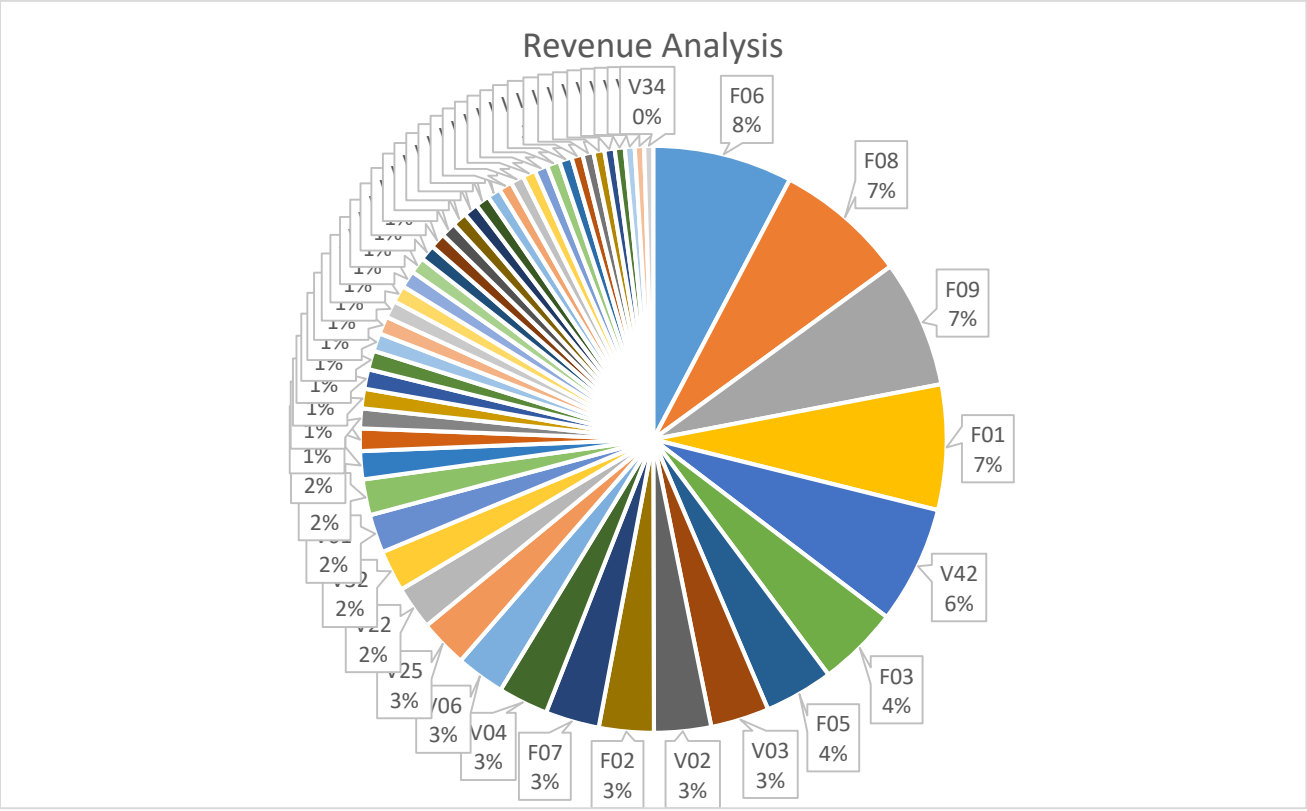
Results and Findings



In the above pie chart, we can see the total sales generated by each SKU as a slice of the pie. From this and the data we have observed that Banana, Tomato, Potato, Onion, Watermelon, Carrot, Papaya, Pea, Cauliflower, Orange, Guava, Apple, Grapes, Cabbage, Sweet Melon, Muskmelon, Bottle Gourd, Spinach, Pumpkin, Cucumber and Methi contributed to 80% of the volume. This clearly shows that the Pareto Principle is not being followed here, as it takes over 21 SKUs out of 51 (roughly 41%) to generate 80% of the sales. We can see this clearly in the chart given below.

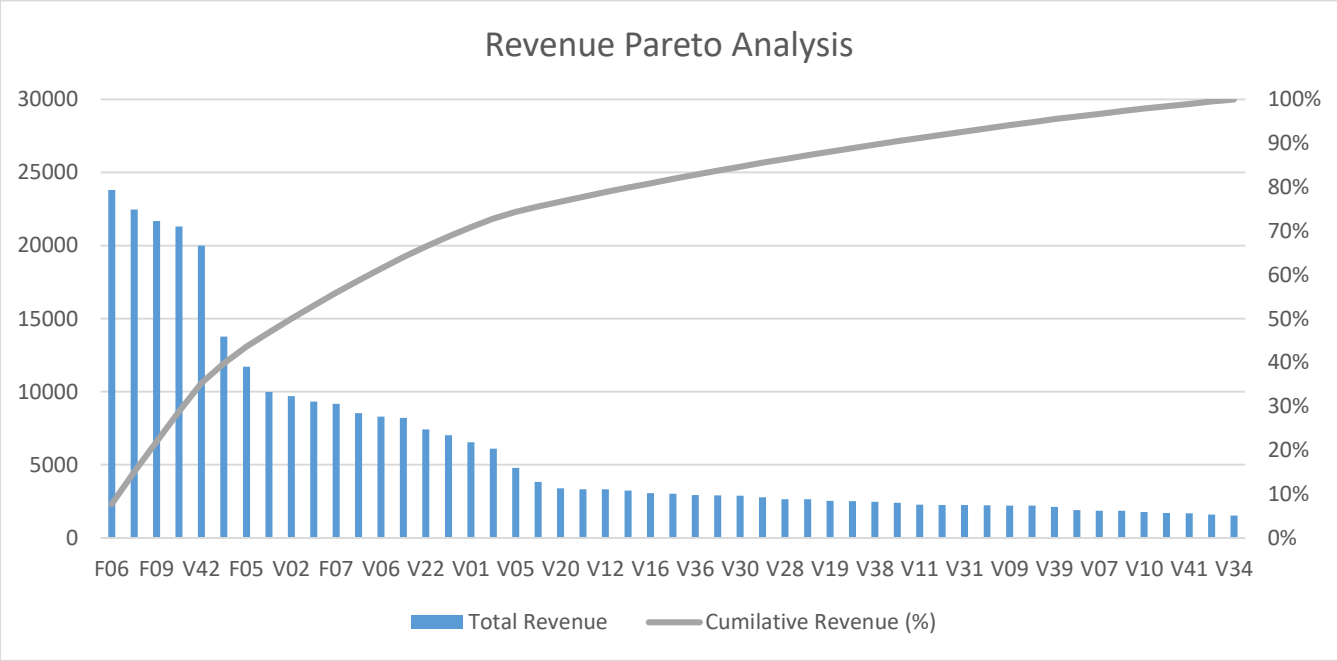


In the above column chart, we can see that on the X-axis are the SKUs while on the Y-axis is the volume of items sold in units defined (Kgs). We can clearly see how the business is not following the Pareto Principle, at least as far as sales is concerned. We should also see the revenue analysis to confirm the same for revenue generated.



In the above pie chart, we can see the total revenue generated by each SKU as a slice of the pie. From this and the data we have observed that Banana, Orange, Guava, Apple, Avocado,

Papaya, Watermelon, Tomato, Onion, Grapes, Sweet Melon, Pea, Carrot, Lady Finger, Garlic, Red & Yellow Capsicum, Potato, Muskmelon, Cauliflower, Lemon, Ginger, Bitter Gourd, Cucumber and Arbi contributed to 80% of the volume. This clearly shows that the Pareto Principle is not being followed here, as it takes over 24 SKUs out of 51 (roughly 47%) to generate 80% of the sales. We can see this clearly in the chart given below.



In the above column chart, we can see that on the X-axis are the SKUs while on the Y-axis is the revenue generated by the shop in units defined (INR). It is evident how the business is not following the Pareto Principle, in the terms on revenue generated as well.

So, the above two charts clearly show that the shop owner should try and optimize their fruits volume and increase it as much more as they could, since they contribute significantly to both revenue and volume. It is not to say that vegetables should be neglected, however, fruits such as ***Banana, Orange, Guava, Apple, Papaya, Watermelon, Grapes, Sweet Melon and Muskmelon*** are common among both the analysis.

With that, vegetables such as ***Tomato, Onion, Potato, Carrot, Pea, Cauliflower and Cucumber*** are common amongst both the analysis.

That makes up to 16 SKUs out of the total 51 SKUs (roughly 31%) which show up interesting results from both the revenue and volume pareto analysis. We can try to focus more on these SKUs and move closer to satisfying the pareto principle, which could help the business optimise their time and resources. As it would be ideal that the business could

generate maximum results from roughly 20% of the SKUs. It does not mean that we neglect the other SKUs, we would also be analysing them through the same process.

As we move ahead into the final submission, there we would like to see how the average days of inventory is for these SKUs, apart from other SKUs as well. We would keenly like to see whether the business is equipped to handle the demand of these SKUs and others in general.

Apart from those, further on in our project we would also like to see the trends and correlations among our SKUs with other factors such as with days, other SKUs and understand why certain SKUs perform better than others in our locality. This would surely help us achieve our target of revenue growth as well as increasing our customer base given the variables at play.

Executive Summary

In this portion of the project, I have tried to analyse various trends in terms of both volume and revenue, to understand how potential trends can be observed in terms of certain days of the week. I have achieved this using PIVOT analysis in Excel and clustered column charts to show the relationship between revenue and sales regarding each day.

Then the seaborn and matplotlib libraries of Python are used to create a heatmap of the correlation amongst the days, in terms of both revenue and sales.

Moving ahead with those observations, I have used the concept of average days of inventory to estimate how long the opening stock survives as per the average sales trend. For this, I have calculated the average opening stock and average sales and plotted the average days of inventory in a bubble chart.

As a concluding remark from my end, I have suggested that if the owner wants to compete with the big players in the market, he must focus on increasing his primary customer base and gain more traction from that end, before he could move to a stage of investing in a cold storage system or warehousing services.

Detailed Explanation of Analysis Process/Method

My primary tool of focus in this analysis has been using the pivot tables feature in Excel to get aggregates of both, revenue and sales data for each day of the week, across all SKUs. Doing this exercise has allowed me to view how the sales and revenue are spread across the week, which can further help us draw useful insights from our data.

We had the raw sales data for each day over the period of 14 days, from which I used PIVOT to find out the cumulative sales for each day. Post this, I have found out the percentage contribution of that day to the total sales generated by the business in the given period. Kindly note that all these can be seen in the sheet '***Sales_Data_Analysis***'. With this data in hand, I have created a clustered column chart, which represents the sales (in Kgs) on the y-axis and the day on the x-axis. Another interesting facet of this chart is that it represents the percentage contribution as a line for each day of the week, with the same being represented on another vertical axis as well.

A similar strategy has been followed for the revenue generated by the organization, just to validate any findings from the above and create a common comparison level for both, sales, and revenue.

I have used this approach as I want to find out on which days can the owner expect high sales and high revenue. Of course, Santosh had an intuition of the same, however, validating it mathematically gives a strong base to work ahead on this. With this, we can find special/key days of the week when Santosh must equitably mobilize his resources.

Moving ahead, I have also tried to find the average days of inventory and express that with the ideal graph. To find out the average days of inventory, the initial step has been to collate the Stock Data in such a way that we can view all variables (Opening Stock, Purchase, Sales, and Closing Stock) in a day-wise manner, which has been achieved in the '***Day_Wise_Stats***'.

With that process being done, I have also tried to get a sense of how the days are related to each other, from perspectives of both sales and revenue. This reason for doing the same is that I completed the previous stage, and some interesting observations were found which led to a few hypotheses, which needed to be evaluated.

For this analysis, I have used the ***Seaborn*** library in Python to create a heatmap of the correlation of days to themselves, given the sales and revenue. For this, we used a PIVOT to first get a day-wise estimate of sales for each SKU and converted that sheet to a CSV file. All

the above results are stored in the sheet '***Heatmap_Analysis_Sales***' and '***Heatmap_Analysis_Revenue***' in our datasheet.

Moving into Python, I used the pandas library to convert the CSV into a data frame that represented our CSV file format. Moving from there, I used the matplotlib library along with seaborn to get a heatmap. All this was then done for revenue as well. The above results were achieved using Google Collaboratory, the link for which has been appended below.

Google Collab Link – BDM_Heatmap.ipynb

https://colab.research.google.com/drive/115wCWEazeHipNNOhLl8Cx7ivkljKF0o7?usp=sharing

Once this sheet was ready, I tried to calculate the average of opening stock and the average of sales throughout the given period. This has been made possible due to the **AVERAGEIF()** function in MS Excel. This allows us to take an average of one column, where rows are iteratively processed, given that our column satisfies are of the same category we seek to find the average for, in our case Opening Stock and Sales. All these changes have been made to the sheet '**AverageDOI_Analysis'.**

Before moving into the average days of inventory, I made a line chart representing the average opening stock and average sales, as well as a line chart representing average sales and average purchases. This has been done as a previous exercise to the average days of inventory, so as to get a first glance at our inventory patterns.

Post finding the average opening stock and average sales for each SKU, we have moved ahead to fit a bubble chart to this data, where the x-axis represents the value of the average opening stock and the y-axis represents the average sales in a day. A bubble is created from the ratio of these both, whose formula is as given below:

$$\boldsymbol{Average\ Days\ of\ Inventory = \frac{Average\ Opening\ Stock}{Average\ Sales}}$$

The average days of inventory have been expressed as a bubble in the chart we have made, with the size of the bubble representing the number of days the opening stock can handle the sales requirement of Santosh's market.

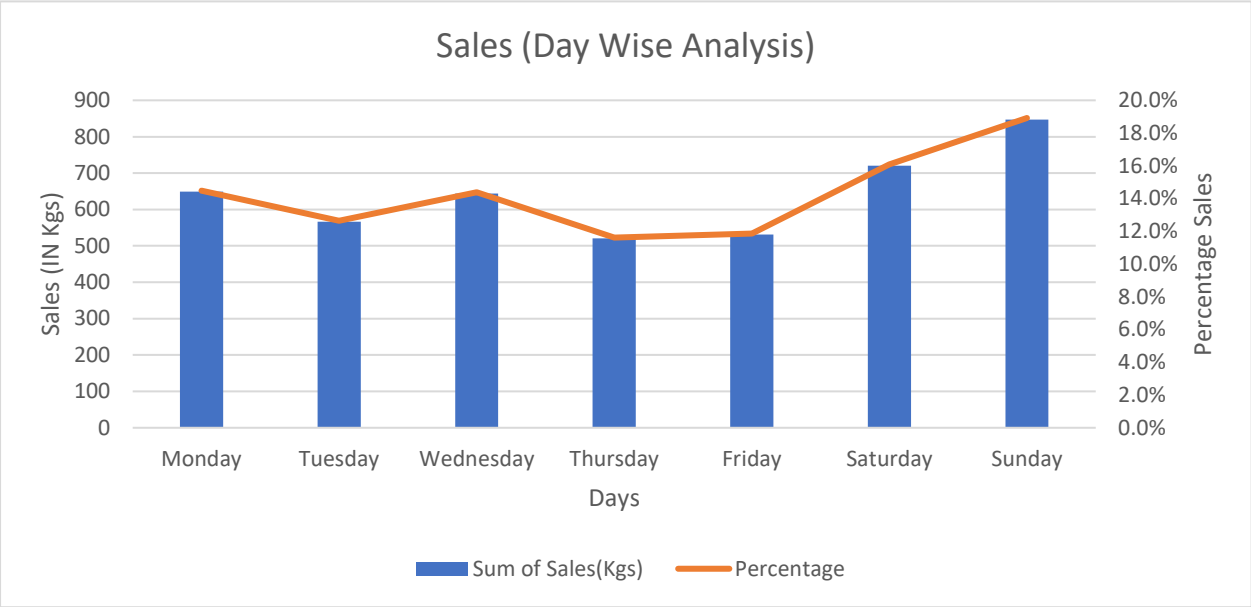
I have taken this approach to analyse his stocking needs as I am of the firm belief that it is critical to know how well Santosh is dealing with his stock, as he is in the business of perishable goods, any kind of mismanagement can have an adverse effect on the business.

Also, expecting a small vendor like him to make investments in even a refrigerator would not be suitable or practical. Not only does it increase the cost as a fixed cost, a recurring expense of electricity is also an issue which only adds. Since he is a small vendor, my idea is to first ensure that he is able to use the best of his resources, minimise any wastages and try to get volume first, so that he can have some money to incur these costs.

In the next section, we will be seeing how our insights have shaped up and will try to draw logical inferences to the same. For that, we have used the help of multiple charts such as Clustered Column Chart, Pie Chart, Line Chart, Bubble Chart, Heatmaps as described above.

Results and Findings

Sales Perspective



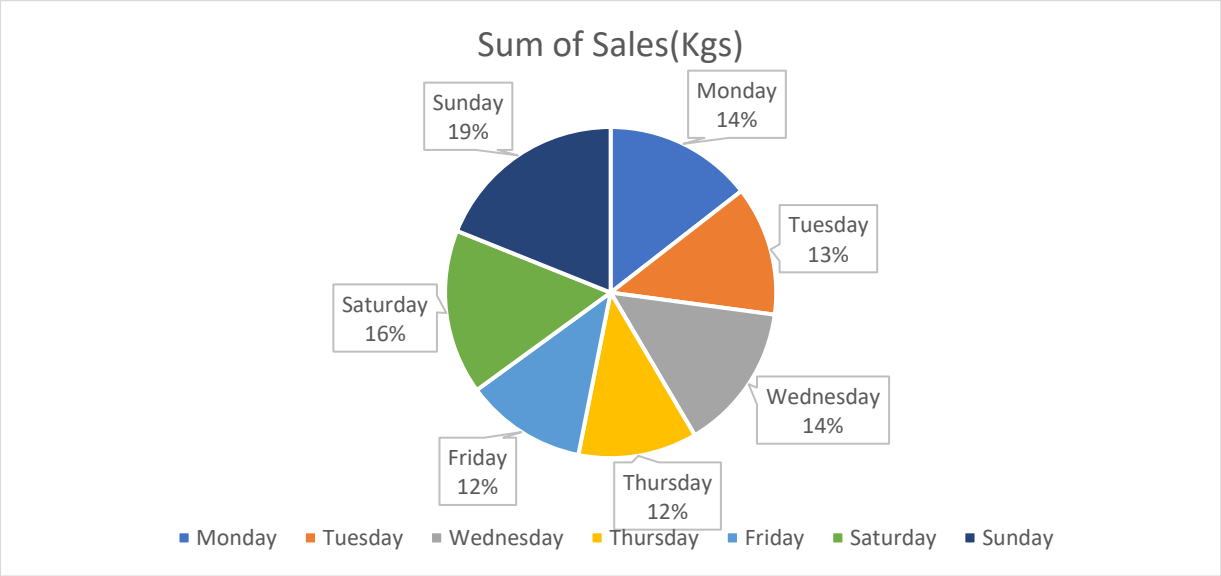
In the above graph, we can see the aggregate of sales on each day and how they individually contribute to the overall sales. The above picture gives us a very interesting observation, for if we see both the percentage line as well as the individual bars.

On weekends i.e. Saturdays and Sundays, the shop has been experiencing increased demands across 51 SKUs and is able to grow his volume traded, as there has been jump in customers buying his vegetables and fruits. So, it is reasonable to say that these two days hold high importance for Santosh in a workweek as they yield maximum outputs, both in absolute and percentage terms.

Apart from that, even Mondays and Wednesdays, his sales have been on the higher side, as compared to the other days of the week. Of course, not as big as the weekend sales, but quite higher than the other 3 days in the working week.

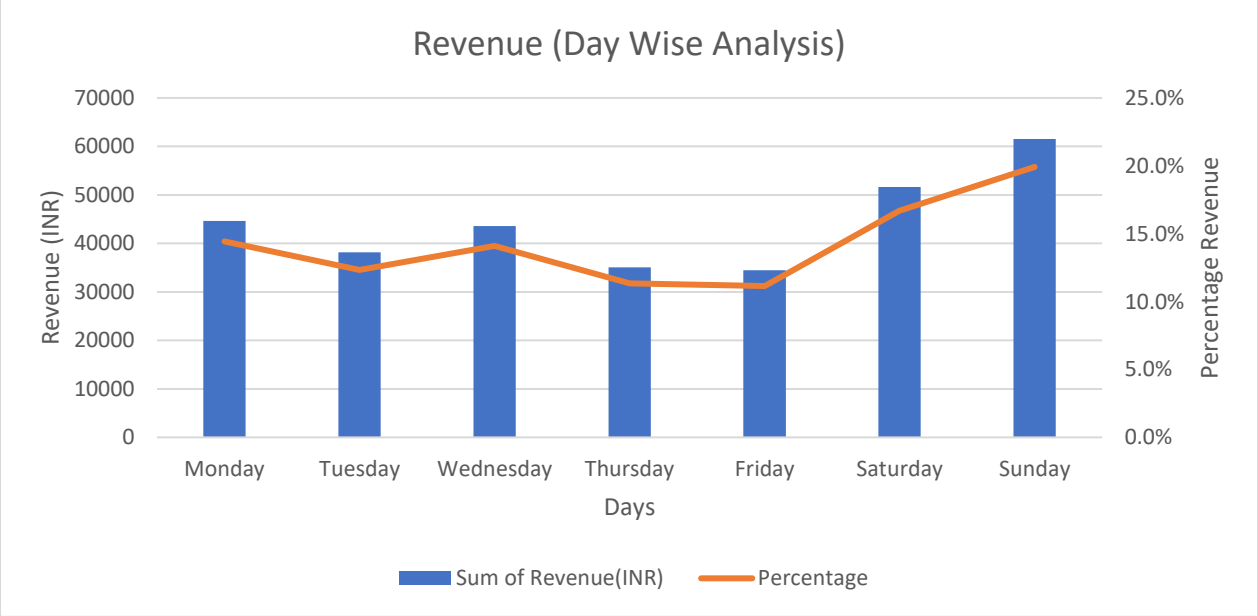
However, on Tuesday, Thursday and Friday the sales have taken a dip, as clearly visible in the figure above. With all these observations, we have decent reason to believe that the shop is gaining more customers on the weekend and shows good traction on even Monday and Wednesday, however it does take a dip on Tuesday, Thursday and Friday.

We could also understand the contribution of each day via pie charts, as shown below.



Here, we can see how all the week days (Mon-Fri) contribute equally to the total sales, with Saturday and Sunday being the highest contributors.

Revenue Perspective

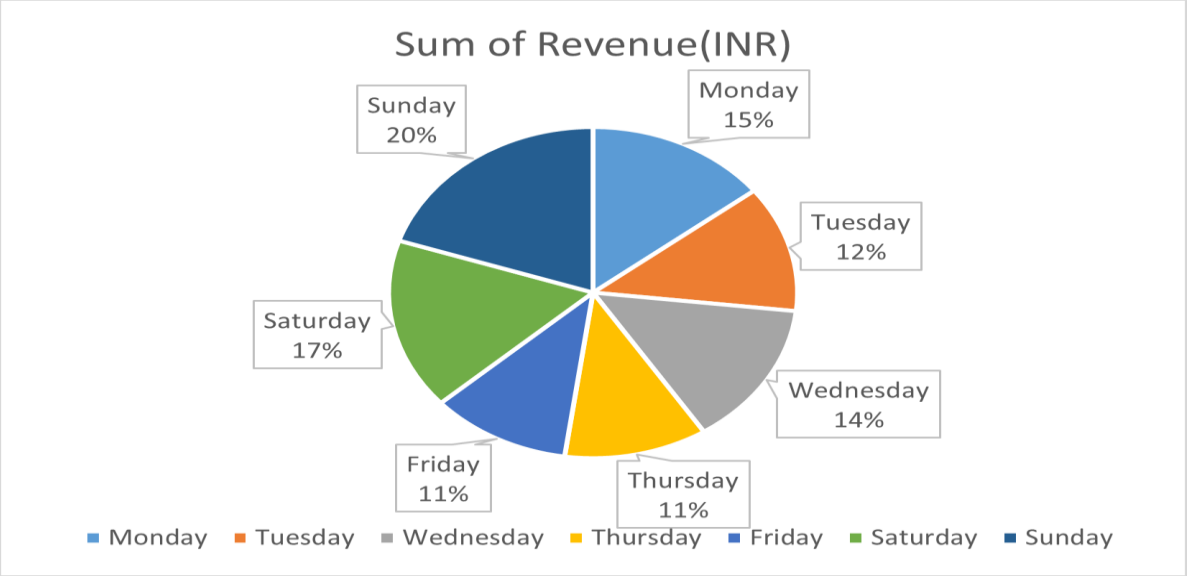


As we can see in the above graph, the above inferences also hold true in this analysis. The days generating the highest revenues are Saturday and Sunday, with the revenue crossing INR 50,000 on both days. Once we think about it, it is only obvious that both days will be performing highly on both sales and revenue front.

Since these days are holidays, most people generally evaluate their weekly consumption of fruits and vegetables, making purchases on the weekends. Apart from that, many people clear their credit due for the week with Santosh. Both these factors contributed to an increment in sales and revenue.

Apart from that, as Saturdays and Sundays are weekly offs, the consumption of vegetables in restaurants also increases, which leads them to make high purchases on that day as well. As Saturdays and Sundays are weekly offs, the consumption of vegetables in restaurants also increases, which leads them to make high purchases on that day as well. This gives us somewhat sense that a mix of both, households and restaurants type of customers is present in weekend sales, which drives up both the facets of the business on these days.

Here also, we can see that revenue increases on Monday and Wednesday. Now, this one is a very interesting observation, as despite people making purchases on Saturdays and Sundays, there is also an increase in both sales and revenue on Monday. We can also see how each day contributes to the total revenue.



Here also, we can see that more or less the days contribute similarly to the total revenue generated by the shop in the weekdays, with Saturdays and Sundays being the only exceptions as they have almost increased by 50%.

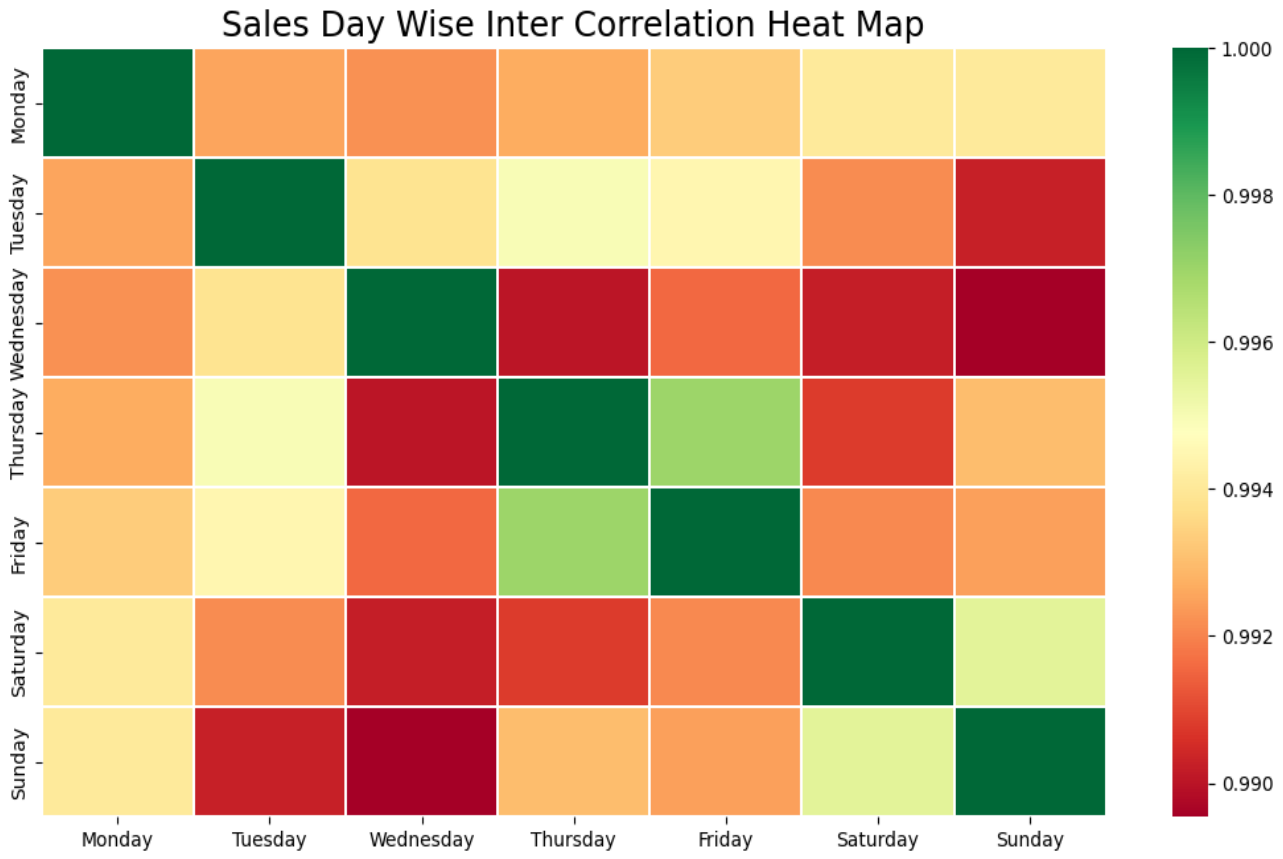
This leads me to believe that many customers also take their weekly consumption evaluation on Mondays and purchase on that day as well. A decent chunk of people does the same on Monday, instead of just Saturday and Sunday. Even in the middle of the week i.e. Wednesday, people are purchasing more, as compared to Tuesday, Thursday, and Friday. We

can now see that there is a clear pattern in the days and how both the factors i.e. revenue and sales are also affected by the day of the week, and that too at a significant level.

Kindly note that all the above observations about how people and restaurants are making their consumption are not truths yet. Yes, they can be considered hypotheses now, however, we would need some other metric to validate/nullify the above.

With these observations, it would also be interesting to see how each of these days is correlated to each other. That can be achieved via a heatmap, which shows how the days are correlated to each other from both revenue and sales perspective.

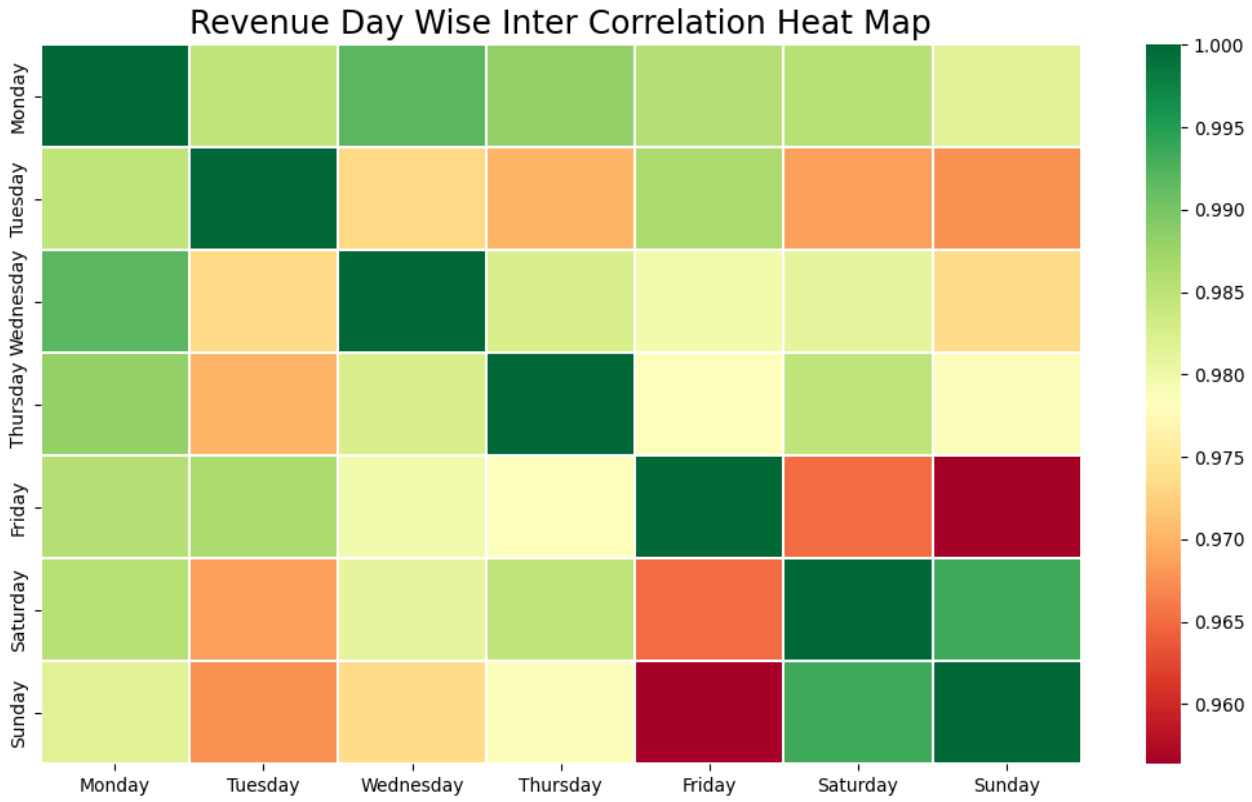
This would allow us to reach a logical conclusion to the hypothesis that have been stated above. Of course, the above hypothesis seems true given the clustered column charts that we have seen, however I am of the view that another level of scrutiny to them would only help us to validate our beliefs/hypothesis and even find faults if any.



In the above heatmap, we can get a sense of how the days are correlated to each other, in terms of sales. As we pay attention to our graph, the Sunday row is showing a neutral-esque correlation with Monday, Thursday, and Friday. As far as Tuesday goes, it shows a lower correlation and when it comes to Wednesday, it shows that it is of dark red colour, implying further degradation of the correlation between the two. It is only with Saturday that it follows a similar pattern.

Only the Saturday row is slightly green to indicate a decent/strong correlation between both days. This goes on to show that how if the sales are increasing on Saturday, they are likely to increase on Sunday as well. This proves that our hypothesis of restaurants making their purchase on weekends has some justification for it.

Also, Sunday and Monday are neutral/moderately correlated, showing that the numbers from Sunday to Monday do not take a significant drip and tend to somewhat follow the same pattern. Another interesting pattern we can observe is that how Thursday and Friday have a strong correlation amongst themselves. It just goes on to show that both these days follow similar trends in terms of sales, may it be in the positive or negative direction. People generally do not prefer to purchase on Thursday and do the same on Friday as well.



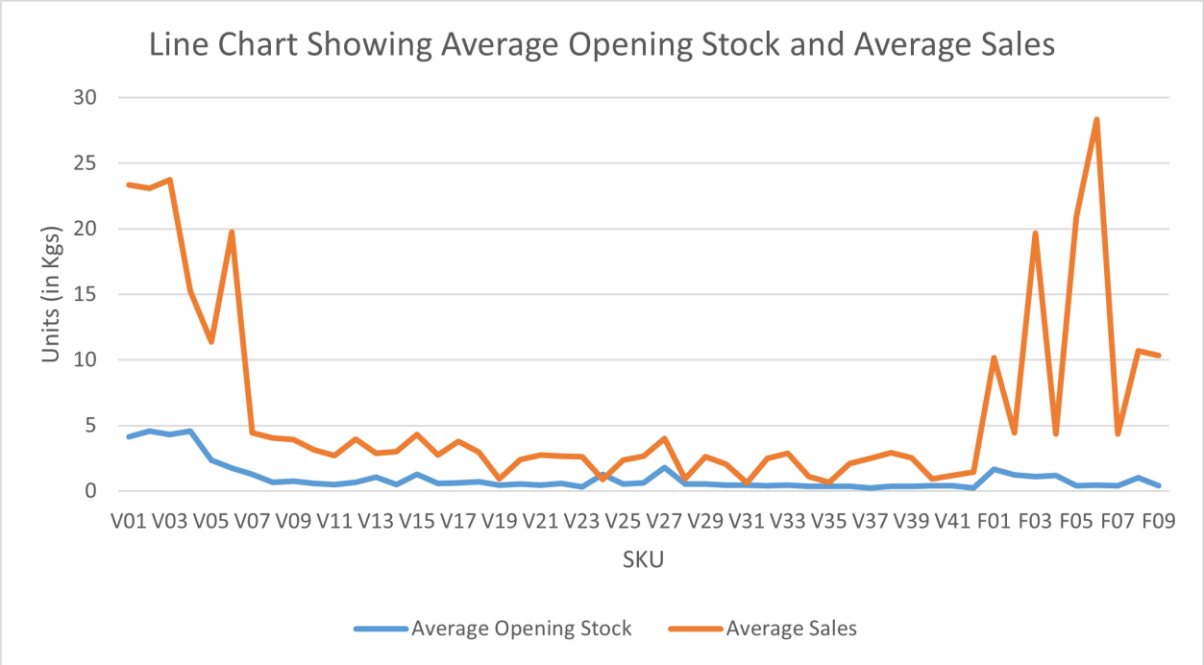
As we can see in the above graph, there are quite a few many interesting patterns that can be observed in the heatmap.

If we look at Saturday and Sunday, we can see how both of them are highly correlated to each other and move in the same direction. This further strengthens the hypotheses of how both types of customers, restaurants, or regular customers, do purchase in a similar trend on both days, in our case which is that they make high purchases on Saturdays and Sundays, leading to stronger revenues on the weekend. Also, in comparison, Monday is more correlated to Wednesday than Saturday or Sunday for that matter when it comes to revenue.

Also, revenue is also driven by previous credit, which the shop owner receives on the weekend. This is because most households for that matter settle their debts with Santosh on weekend's, days when they themselves are at home.

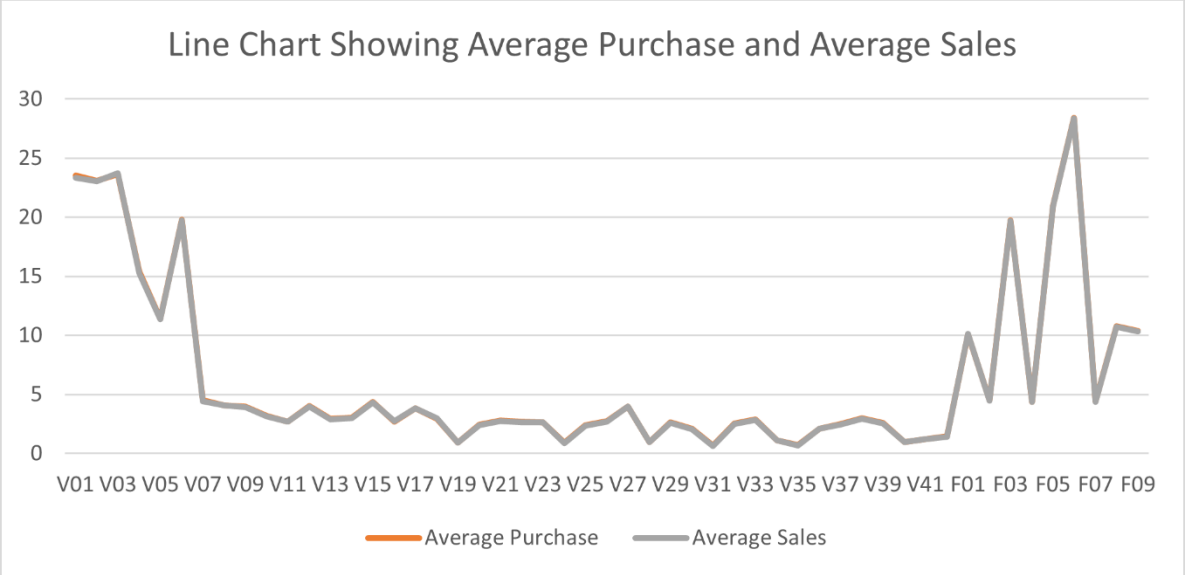
If we look at the correlation between Sunday and Friday, as well as Saturday and Friday, we can see that how both are inversely correlated to each other, in comparison to other days. With this, we can rightly say that towards the end of the working week, people do not follow/purchase in the same pattern as that of weekend days. Similar to this, we can say the same for Tuesday and Thursday as well, when talking about their correlation with the weekend days.

Now, before we move into the average days of inventory concept, I thought it would be better to get a sense of how our average opening stock and average sales look like. For that, I have used Line chart to represent both these things on a single graph.



As we can see in the graph above, there is a very big gap between average opening stock and average sales. If we see carefully, none of the average opening stock from all SKUs crosses a value greater than 5 kgs, whereas quite a few many SKUs leap above 10 Kgs, with some even going above 20.

This gives us good reason to believe that our shop owner is making purchases on a daily basis to meet the demand of the business. Again, it would be better if we can justify the above statement using a graph, for which we can plot a line chart of average purchase and average sales.



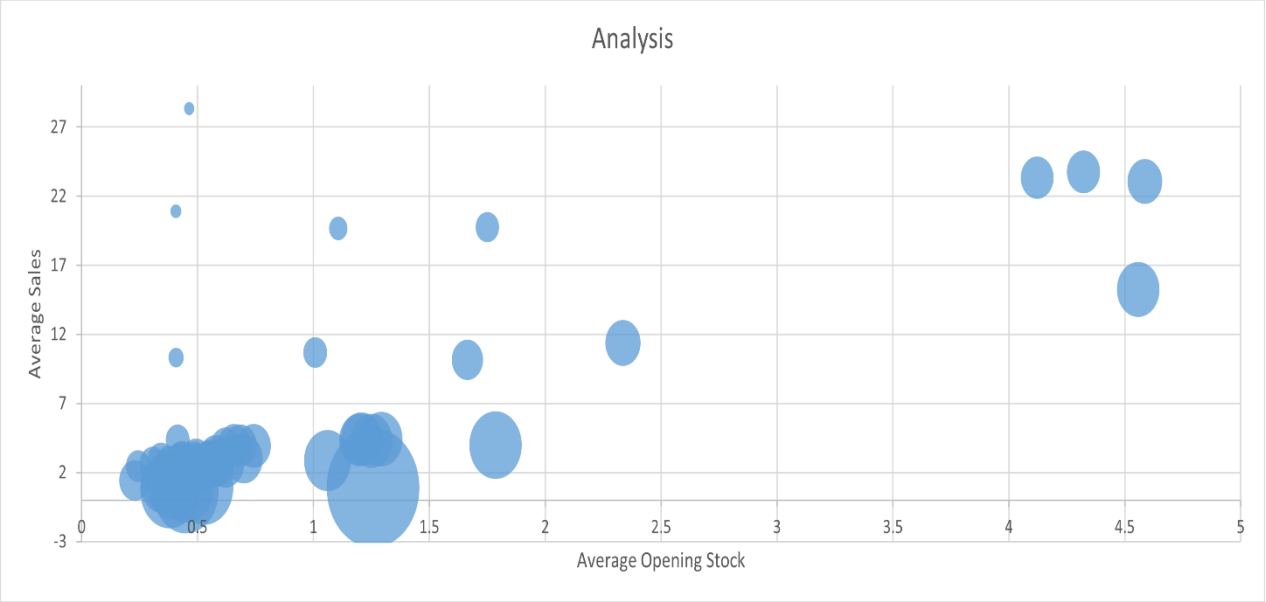
As we can see in the above graph, it is quite evident that the shop owner is making the purchase and sales hand to hand, the option of keeping an active inventory is not being exercised in this business.

Of course, we have to understand that this market is of perishable goods so storage cannot be done on a large scale, however, that can also lead to a lot of problems in the future.

Santosh basically makes his purchases through a local wholesale market, as he gets the best margin for his vegetables/fruits there only. Since the product is homogenous and the market is perfectly competitive, he may not be in a position of increasing costs to ensure profits. Sellers in such a market may not be able to add value to their product because adding value may not increase prices which would be generally accepted by other customers. With added competition from *BlinkIt* and other online platforms, a price war may not be a war Santosh could afford in the long run.

As this would eventually increase the cost to the seller while having minimal impact on the revenue. Ultimately, the profit is the thing which suffers as that margin decreases. And, if he increases his prices owing to the better quality of items, consumers who are not his regular customers may just shift to other options.

Daily procurement of vegetables is ideal now as the operation is small, however if he wants to scale up this may not be the best strategy. This procurement is being done daily, which brings an additional cost to our business. Eventually, we reach a stage where Marginal Cost to Procure an item becomes greater than the Marginal Revenue to sell an item and the graph of Marginal Cost goes upwards. This leads to a loss-making business.

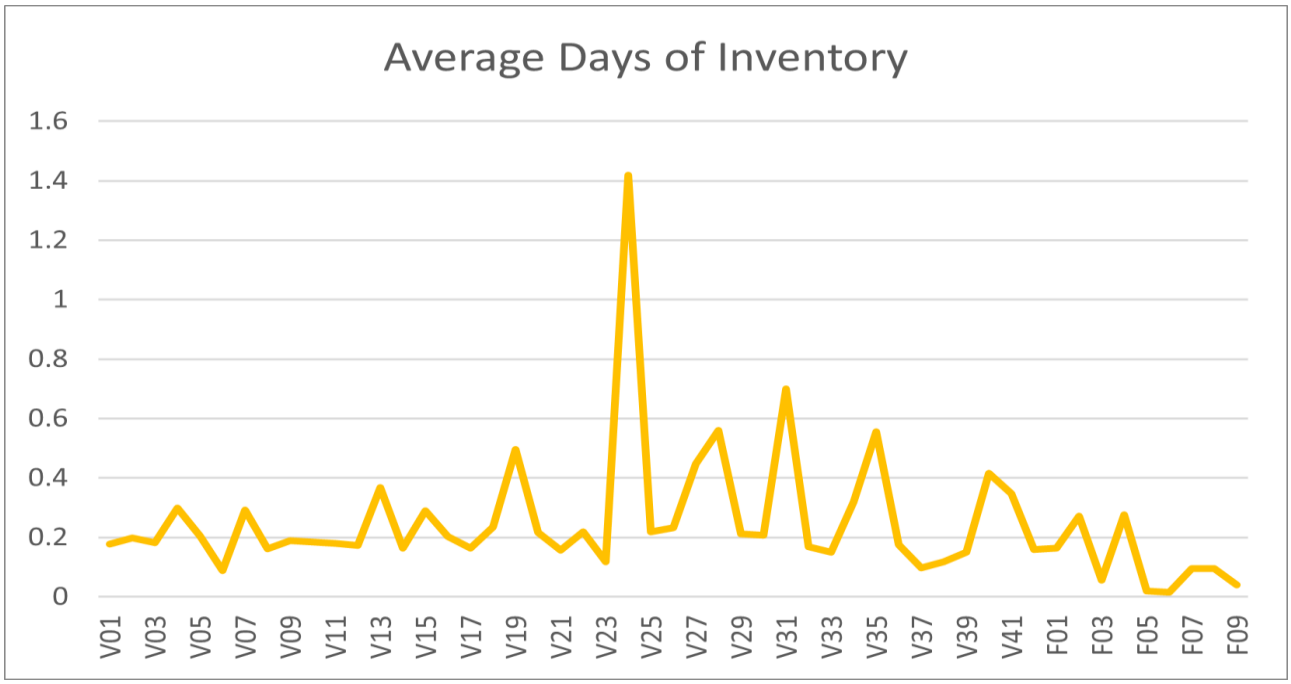


In the chart above, all the inferences we made above can be seen here and further strengthen our analysis. Here, the x-axis represents the average opening stock and y-axis represents the average sales for all SKUs.

The bubbles here represent the average days of inventory for each SKU. Since we are in the business of perishable goods, the values will not be very large. However, what we were expecting, this is even lesser than that.

As per Santosh's belief, he is of the opinion that mostly all items have a shelf life of 2-3 days at least, as the households of the slum area do not mind if the vegetables/fruits are not as fresh as they should be. We should look at the graph given below to understand this argument and prove that this prior belief is not correct.

Another way would be to view the average days of inventory in some other chart, individually, as shown below.



As we can see in the above chart, the y axis represents the number of days, and the x-axis represents the SKUs. Contrary to our prior belief, our business does not have a shelf life for products for 2-3 days, but rather less than 1 day.

This is leading Santosh to make purchases on every day, which is leading to a high cost as time goes by.

Interpretation of Results and Recommendations

With all that we have gathered and analysed, I was able to understand the data and its nuances, as well as draw some key insights to how to increase the revenue of the organization.

One key thing which we need to understand is our business is of perishable goods, that too in a perfectly competitive market. There are infinitely many buyers and many sellers in our business, with the entry point not being a major barrier. Santosh can only reduce his price by a limited amount if he must maintain a decent profit margin, let alone grow and upscale his business. Of course, with that being said, he is already facing tough competition from large players in the market such as Blink It, Swiggy Instamart and other online vendors who deliver vegetables and fruits than cheaper rates than his.

Let us also factor in the type of customers he has. As per his observations, he primarily has three types of customers which are as follows:

1. **Recurring Middle/High Class**: Basically, these are the people which live in apartments/flats or bungalows and have a recurring demand of vegetables and fruits. With them being privileged, the 20-30% increment in the price of the product does not affect them as long as they are satisfied with the quality of the product. The only USP that Santosh has to offer is that his vegetables/fruits are fresh and of high quality. Any person who gets used to his products, stays on as a regular and recurring customer and does not mind paying an extra fee as compared to online vendors.
2. **Low Class/Slum Area**: These people are also those who would prefer to have fresher vegetables and fruits, however cost is a big factor to them. Since most of them are earning their income daily, they do hesitate before paying a higher sum, so much so that they ready sacrifice on the quality of the goods up to a certain extent, if it does not hamper their daily budget.
3. **Restaurant**: This class basically consists of the restaurants/hotels who need vegetables primarily to make their broth and curries. Since in that use case, it is not necessary for the vegetables to be fresh. They buy it at the same rates the slum people buy it for.

In the last two classes of customers, Santosh cannot expect to increase his revenue or volume after a certain point, as they both are not the ideal market for the USP his product as they do not value it over cost.

The only way Santosh can increase his earnings is by an increment of volume. Since he cannot play on the price component, his efforts should be focused on the procurement of more customers and building his customer base in such a way that it helps him to upscale his business and gain more profits.

For that, I recommend the strategies to be adopted by him, which are as follows:

1. Door to Door Service and Asking for Recommendations
    a. Here, a face-to-face connection with the customer helps to build an authentic relationship with the customer. An online vendor may be able to provide better prices and faster delivery, however, they will not be able to provide a personal connection. Going personally to deliver vegetables whenever possible and maintaining a cordial connection with them is a thing that pays greatly as time passes by.
    b. There are certain times when vegetables come rotten or with some defect to the customer, at that time, having someone known to reach out to and get them replaced only adds to customer satisfaction. It is a very long process to even reach a customer care representative, let alone get the product replaced.
    c. Santosh can make sure his post-selling services make up for the additional cost borne by the customer and leaves a great impression. In the end, it is finer details that determine whether a customer will stay or leave.
    d. Every once in a while, Santosh could ask the customer to give their other known people a small recommendation of his services. The online vendors have a great marketing and advertisement budget, through which they are able to reach a lot of people. However, the power of word of mouth is unparalleled even in today's digital age.
    e. Especially with these customers, they give great importance to someone whom they know, and with that, they themselves reach out to experience a good product. Sure, a huge hoarding about '*fresh vegetables in 15 minutes*' can be effective, but a word from a close friend or relative carries much more weight and is acted upon whenever the opportunity arises. This could surely help Santosh to reach out to other localities and grow his customer base.

With this, Santosh can definitely hope to increase his customer base, which helps to increase his sales, and that in turn, would boost his revenue. As his revenue increases, he can then

make fixed purchases such as buying a small scooter/e-rickshaw for deliveries, a small refrigerator to store high price products which can help to increase their shelf life, a small room or storage facility to keep his products for later use and to handle future demands etc.

One more thing which Santosh must ensure is that the wastages are minimized. Since the business is of vegetables and fruits, there is always a high chance that wastages are present if the sales and not met with purchases at an ideal level. Incorrectly estimating sales and purchasing more goods can lead to massive wastages, which bleed the business out.

It is not a certain fact, however if Santosh tries his level best to increase his user base organically, he can surely increase his revenue. Basically, in this business, greater the turnover and money circulation, higher the chances of profits.