Data Analysis of Product & Services of Agrotech Enterprise

A Midterm submission report of the Business Data Management Capstone Project

Submitted by

Name: Saurao Upare
Roll No: 21f1000355

IITM Online BS Degree Program

Indian Institute of Technology, Madras, Chennai Tamil Nadu, India, 600036

Contents

Executive Summary:

I had approached the responsible staff member and with the owner's permission collected the client invoices, including the sell of different tools, machinery,manure and the other agricultural related things.I convinced the owner about the idea of letting me go through some aspects of the data which are required to solve the problem the gym currently faces.

The Agricultural business fluctuates rapidly with the season as the rainy season have a high demand of all sorts of agricultural tools, fertilizers and seeds, whereas it decreases at the end of this season and somewhat rises in winter for a particular time period.And in the summers this business totally depends on the small scale business and plant nursery business people, whereas a homegrown plants and fruits and vegetables also sort of have a high demand nowadays but they seems to avoid the organic manure and the Gandul khat (manure made the help of earthworms and home cooked leftovers) as they have a high cost. As a requirement of time and environmental awareness and soil health etc..those topics are slowly being aware of the importance of organic manure, which will take time but this has a consequence on the business of the owner and a steady revenue has become an unpredictable scenario.

To solve the problem of unhealthy retention rate, which can be looked into via the invoices generated and collected between the said period of time, the necessary data has been cleaned and transformed into a clean spreadsheet which will aid me to draw some insights and gain a basic understand of the problem at hand which is the primary reason for dynamic and unstable revenue by analyzing the necessary data generated every month, I will figure out the necessary patterns and possible reasons behind those patterns and provide the business owner with valuable insights and suggestions which will in turn help the business to retain more and more of their clients and employ strategies which will help them do so.

Proof of originality of the Data:

1. Letter from organization in letter head with stamp and sign:

As part of the proof I have talked to the owner and asked to write a letter of authenticity which clearly clarifies that he has given me the data of his organization and will handle it carefully for academic purposes only.

Organization authenticity letter by owner:

https://drive.google.com/file/d/1XGgKiCKm4a8_e2C57BZJ3kLLvGBHgbKf/view?usp=share_link

2. Photograph of the shop with me and owner of the shop:

I requested him to take a photo with me and along with that I also included the main office of the shop and at the background we can see the whole shop. He being busy in his work I manage to take a single pic with me and also including a major portion of the shop.

The photo can be accessed from below link:

https://drive.google.com/file/d/1W53OZrXq_R2lVG9pCo3VkKafgjcidG_I/view?usp=share_link

3. Recorded video with founder:

I was so nice to me and quickly given permission to take a video of interaction, by the way the video of interaction was taken by me when I was talking for the data sharing with me, he was half convinced for the data at the first meet..but when I asked for the authenticity letter and shown the letter provided by the IIT M, he was totally convinced and I also gave him the assurance of use of data for academic purpose only.

Interaction video with owner:

https://drive.google.com/file/d/1XUY9aRAz_txmIT96OhAlM214L5OrgmPG/view?usp=share_link

4. Link for the primary data:

I have shared the raw / primary data which I had collected from the business with the exception of a few features like the owner has omitted some columns such as GST no, customer name, contact no for the privacy concerns and provided me with the data. The link of the data file is given below, which I imported in a google sheet and shared here:

https://docs.google.com/spreadsheets/d/1iXmSrWdUIJLkQSfgC3tRVYlQGxk0JLjDIAJy0I0kLrs/edit?usp=sharing

METADATA:

The List of Sales Vouchers containing all the voucher series, bill no and the quantity of products sales are included in the spreadsheet where the monthly sales are highlighted for the consecutive 1 and a half year. After performing data cleaning, transformation, and column elimination, I am left with some of the most significant data features that have been collected in a single sheet and which I think to be strongly relevant to my study.

Basic operations have been performed, including changing the supposedly numbers data type from text or a string of text to numbers, eliminating all useless columns that have no bearing on the analysis, getting rid of redundant columns, and adding a few meaningful columns to further simplify my own analysis.

List of Sales Vouchers

In this sheet the featured columns such as the ‘Bill No’, ‘Particulars’, ‘TIN/GSTIN No.’, ‘MRP’ have been removed by me. The Particulars,bill no, GSTIN no, are of no use to me so I removed those columns but the MRP column was removed by me for the reason that some products have a zero MRP selling price because the product or the manure is produced locally and according to the market rate he is selling those products,so MRP was of no use for me.While I also add two new column for my analysis ie.Month and Type.Below is the picture of the systematized data I came up with, also there are two sheets of two consecutive financial year from 1/04/2021 to 31/03/2022 and other sheet from 1/04/2022 to 30/11/2022

Date	Month	Type	Item Details	Material Centre	Qty.	Unit	Price	Amount
05-04-2022	April	Tool	Agro Net 75%	Main Store	10.00	Metre	104.76	1047.62
05-04-2022	April	Tool	C76 Water Sprinkle Can 5 Ltr	Main Store	1.00	Pcs.	223.21	223.21
11-04-2022	April	Tool	Cap 30/75 400 X 400 SS	Main Store	5.00	Pcs.	7033.90	35169.49
15-04-2022	April	Tool	BRUSH CUTTER 2 WAY BLADE	Main Store	10.00	Pcs.	296.61	2966.10
15-04-2022	April	Tool	BRUSH CUTTER ROD	Main Store	1.00	Pcs.	2118.64	2118.64
23-04-2022	April	Machinery	Engine Operater Post Hole Dig	Main Store	1.00	Pcs.	18571.43	18571.43
05-05-2022	May	Machinery	Chaff Cutter	Main Store	1.00	Pcs.	16964.29	16964.29
05-05-2022	May	Machinery	Cham Power Weeder PTG500	Main Store	1.00	Pcs.	58035.72	58035.72
09-05-2022	May	Tool	BRUSH CUTTER ROD	Main Store	1.00	Pcs.	1694.92	1694.92
09-05-2022	May	Tool	CLUTCH DRUM	Main Store	1.00	Pcs.	635.60	635.60
09-05-2022	May	Tool	PIPE	Main Store	1.00	Pcs.	423.72	423.72
10-05-2022	May	Tool	Falcon Professional	Main Store	1.00	Pcs.	630.00	630.00
11-05-2022	May	Tool	Agro Net 90%	Main Store	257.00	Metre	119.05	30595.24
12-05-2022	May	Organic manure	COCO PIT	Main Store	79.62	Kgs.	50.00	3981.00
12-05-2022	May	Tool	Sq Tray- 13	Main Store	2.00	Pcs.	118.64	237.28
12-05-2022	May	Tool	Agro Net 75%	Main Store	6.00	Metre	95.24	571.42
24-05-2022	May	Tool	Sprinkler 3/4" 180 D/C Indust	Main Store	1.00	Pcs.	276.78	276.78
24-05-2022	May	Tool	P-2 Imp.sprinkler(3/4") M Wei	Main Store	1.00	Pcs.	178.58	178.58
24-05-2022	May	Tool	YELLOW STICKY TRAPS 25 CM	Main Store	2000.00	Pcs.	8.47	16949.16
24-05-2022	May	Tool	Carret 25 Kg	Main Store	150.00	Pcs.	183.04	27455.36
24-05-2022	May	Tool	Mulching 25 Micron 1 Mtr	Main Store	13.00	Roll	1737.29	22584.74
30-05-2022	May	Tool	Sprinkler 3/4" 180 D/C Indust	Main Store	10.00	Pcs.	267.86	2678.58
07-06-2022	June	Machinery	Aspee power Oprated Sprayer	Main Store	1.00	Pcs.	18571.42	18571.42
17-06-2022	June	Machinery	Cham Power Weeder PTG500	Main Store	1.00	Pcs.	58035.72	58035.72
18-06-2022	June	Machinery	Engine Operater Post Hole Dig	Main Store	1.00	Pcs.	18571.42	18571.42
23-06-2022	June	Tool	BRUSH CUTTER ROD	Main Store	1.00	Pcs.	1694.92	1694.92
23-06-2022	June	Tool	PIPE	Main Store	1.00	Pcs.	423.72	423.72
23-06-2022	June	Tool	Gear Assembly	Main Store	1.00	Pcs.	848.22	848.22
23-06-2022	June	Tool	Water Can 6 Litre	Main Store	1.00	Pcs.	223.22	223.22
23-06-2022	June	Tool	Trimerline wire	Main Store	2.00	Bundle	550.85	1101.70
24-06-2022	June	Tool	Water Sprinkler 3/4"	Main Store	30.00	Pcs.	291.83	8754.92
24-06-2022	June	Machinery	Cham Power Weeder PTG500	Main Store	1.00	Pcs.	58035.72	58035.72
30-06-2022	June	Tool	YELLOW STICKY TRAPS 25 CM	Main Store	200.00	Pcs.	27.97	5593.22
30-06-2022	June	Chem	Leucin Lure	Main Store	200.00	Pcs.	24.00	4800.00
16-07-2022	July	Machinery	Champ Power Weeder PTG 50	Main Store	1.00	Pcs.	126491.08	126491.08
16-07-2022	July	Tool	Aspee power Oprated Sprayer	Main Store	1.00	Pcs.	19196.42	19196.42
19-07-2022	July	Tool	Tap & Go (Heavy)	Main Store	1.00	Pcs.	593.22	593.22
19-07-2022	July	Tool	Tap & Go Light	Main Store	1.00	Pcs.	357.14	357.14
20-07-2022	July	Tool	Nursary Bag	Main Store	7.00	Kgs.	169.49	1186.44
21-07-2022	July	Tool	Aspee power Oprated Sprayer	Main Store	1.00	Pcs.	18571.42	18571.42
21-07-2022	July	Tool	Black Plastic Mat	Main Store	57.00	Kgs.	194.92	11110.16
22-07-2022	July	Tool	Safety Goggles	Main Store	2369.00	Pcs.	110.17	260991.52
22-07-2022	July	Tool	N95 Mask	Main Store	2369.00	Pcs.	76.19	180495.24

Date	Month	Type	Item Details	Material Centre	Qty.	Unit	Price	Amount
Apr 1, 2021	April	Tool	Carborater	Main Store	1.00	Pcs.	848.22	848.22
Apr 1, 2021	April	Tool	C35 FOUR ARMS SPRINKL	Main Store	10.00	Pcs.	118.64	1186.44
Apr 1, 2021	April	Tool	PACKING KIT	Main Store	1.00	Pcs.	127.12	127.12
Apr 1, 2021	April	Tool	Clutch Assembly	Main Store	2.00	Pcs.	223.21	446.42
Apr 2, 2021	April	Tool	Grass Weed Slasher	Main Store	1.00	Pcs.	2678.58	2678.58
Apr 2, 2021	April	Tool	Garden Hand Tools	Main Store	1.00	Pcs.	2542.38	2542.38
Apr 2, 2021	April	Tool	Grass Weed Slasher	Main Store	1.00	Pcs.	2321.42	2321.42
Apr 2, 2021	April	Tool	Harvesting Tools	Main Store	1.00	Pcs.	3750.00	3750.00
Apr 7, 2021	April	Tool	Spray Pump	Main Store	1.00	Pcs.	2142.86	2142.86
Apr 15, 2021	April	Tool	Battery Spary Pump 18 Lt	Main Store	1.00	Pcs.	2142.86	2142.86
Apr 20, 2021	April	Tool	Falcon Garden Tools Gras	Main Store	1.00	Pcs.	580.00	580.00
Apr 20, 2021	April	Tool	Spray Pump 5 Ltr	Main Store	1.00	Pcs.	900.00	900.00
Apr 20, 2021	April	Tool	Digger	Main Store	1.00	Pcs.	250.00	250.00
Apr 20, 2021	April	Tool	Golf King Hedge Shears	Main Store	1.00	Pcs.	600.00	600.00
Apr 20, 2021	April	Tool	Garden Gloves	Main Store	1.00	Pcs.	152.54	152.54
Apr 22, 2021	April	Tool	Mulching 25 Micron 1 Mt	Main Store	15.00	Roll	1677.97	25169.50
May 1, 2021	May	Tool	Green Net 50% (Barish)	Main Store	2.00	Bundle	2380.95	4761.90
May 1, 2021	May	Tool	Green Net 50% (Double I	Main Store	4.00	Bundle	2857.14	11428.58
May 6, 2021	May	Tool	Spray Pump	Main Store	1.00	Pcs.	2142.86	2142.86
May 16, 2021	May	Tool	Nursary Bag	Main Store	50.00	Kgs.	169.49	8474.58
May 16, 2021	May	Tool	Nursary Bag	Main Store	5.00	Kgs.	169.49	847.46
May 16, 2021	May	Tool	Agro Net 75%	Main Store	50.00	Metre	95.24	4761.90
May 16, 2021	May	Tool	Net Rasi	Main Store	2.00	Bundle	150.00	300.00
May 16, 2021	May	Tool	Gadasa	Main Store	1.00	Pcs.	400.00	400.00
May 20, 2021	May	Tool	Battery Spray Pump	Main Store	1.00	Pcs.	1964.28	1964.28
Jun 1, 2021	June	Tool	Flora Prunning Shears	Main Store	2.00	Pcs.	500.00	1000.00
Jun 1, 2021	June	Tool	Flora-2 Pruning Shears	Main Store	3.00	Pcs.	700.00	2100.00
Jun 1, 2021	June	Tool	Trimerline wire	Main Store	1.00	Bundle	508.48	508.48
Jun 1, 2021	June	Tool	2 Teeth BLADE	Main Store	6.00	Pcs.	296.61	1779.66
Jun 1, 2021	June	Tool	Nursary Bag 6x9 Inch	Main Store	50.00	Kgs.	167.80	8389.84
Jun 1, 2021	June	Tool	Black Plastic Mat	Main Store	100.00	Kgs.	190.68	19067.80

Metadata

1	Purchase date	This column consist of the date on which the product has been sold
2	Month	This column says about the month in which the purchase was made
3	Type	This column tells us about the product type which is purchased on that particular date
4	Item Details	This column tell us about the details of the item which is to be purchased by the customer
5	Material centre	This column tells us about the location from where the purchase is being made as the enterprise has a second plant nursery farm near to the village where the selling also takes place.
6	Quantity	This column tells us about the number of units sold to the buyer of a particular item
7	Unit	This column tells us about the specification of the standard measurement unit that is used to count inventoried items and express them in specific quantities.
8	Price	This column tells us about the price at which the retailer sell the product, it is for a single quantity amount
9	Amount	This column tells us about the total price of the items that is will be purchasing (comprises of the price of the items x quantity of the particular item)

Descriptive statistics

The total revenue generated by the shop for the financial year 2021-22 is ₹ **18721907.07.** The product sale was major in the month of the December 2021 with revenue of ₹ **15311146.02** as this time Covid third wave was gone and farmers were being ready for the 2nd crop of year

(kharip crop), while the revenue in the month of monsoon season (which had the peak for the covid third wave) was ₹ **567404.74.** Also there was a significant revenue generation for march 22 ie. ₹ **801374.28.**

Sum of Amount	Month												
Type	January	February	March	April	May	June	July	August	September	October	November	December	Grand Total
Chem							395	812	8000		220	5320.2	14747.2
Machinery	175892.84	255357.16	519670.1			62500	196250.02	132589.3	269596.56		419642.85	13526795.72	15558294.55
Organic Manure		7500					10588.4	311.92	330.36	11847.14	428.58	1674.1	32680.5
Plant							1838	393		1200		1125	4556
Tool	53585.34	40297.28	281704.18	45838.34	35081.56	504904.74	31461.32	67838.32	140258.22	96258.19	38170.33	1776231	3111628.82
Grand Total	**229478.18**	**303154.44**	**801374.28**	**45838.34**	**35081.56**	**567404.74**	**240532.74**	**201944.54**	**418185.14**	**109305.33**	**458461.76**	**15311146.02**	**18721907.07**

The total revenue for the next year was ₹ **2114667.09** and highest for the month of July 22 ie ₹ **1066815.69** and lowest being for the month of November ie ₹ **37848.56**

Sum of Amount	Month								
Type	April	May	June	July	August	September	October	November	Grand Total
Chem			4800	33337.02					38137.02
Machinery	18571.43	75000.01	153214.28	193008.04	119553.58	19642.86	22321.42	22914.64	624226.26
Organic manure		3981							3981
Plant					21000		8000		29000
Tool	41525.06	104911.38	18639.92	840470.73	316146.64	40712.08	41983.18	14933.92	1419322.91
Grand Total	**60096.49**	**183892.39**	**176654.2**	**1066815.79**	**456700.22**	**60354.94**	**72304.6**	**37848.56**	**2114667.19**

The organic manure has a significant sale in the year 21-22 ie for the month of July, August, September, October, November, December and February and highest being for the month of October & July ie ₹ **11847.14** & ₹ **10588.4** respectively

The Machinery has a significant sale all over the year (21-22) and the revenue generated by sale is highest being for the month of December ie ₹ **13526795.72** whereas lowest for the month of June ₹ **62500**. For machinery there is no proper trend the sale of every product is effected by covid wave, but we see a peak at the month of December.

For the next year the revenue generated by the sale of the organic manure is ₹ **3981** which is very low than what was in the last financial year. And the sale of manure is only for the month of May in the year 2022.

The revenue generated through Machinery is ₹ **624226.26** for the year 2022, where highest for the month of July and lowest being for April ₹ **18571.43**. Unlike organic manure the sale of machinery is significant in evey month, despite of high cost. For Machinery there is a proper trend which has a peak for the month of July.

The sale of other product is normal whereas tools contribute significantly for the given revenue generation, while the revenue generation by the tools for the year 22 is astounding. The above statistics is very helpful in understanding the business revenue fluctuations and current retention rates.

Detailed explanation for analysis method:

There are two different voucher sale list of year 21-22 to 22-nov22, and seen two major issues in the sale and also when I discussed with owner the following two major issues he pointed out to me. So I worked out on the following data and looked for the solution for the issues.

Decrement in the sale of Organic manure

I have taken the data and imported it to the excel, both the data sheet where cleaned up by me, removed the redunctant values, null values, data missing rows or columns. Then I added up two more columns of month and type. The column of month was added by me for the purpose that when I will compare data over the year it will more easy for me to look around the month as compared to looking particularly for the respective days. The column of type is very important for the analysis as I am looking for the sale of organic manure and machinery the data was not labelled or the identification of data was not there. When I first saw data it took like 2 hours for pointing out the sale of organic manure and machinery. So I decided to add the column of Types in the data so whenever I make a pivot table then the comparative study will be more easy.

After processing of the data I will make a pivot table and will see the trends of the sale of different products over the month for both years. From that I will make a trend plot or a scatter plot which will give me a spread of sale over the months and in which month farmers/customers and buying the organic manure for there work.Will make some conclusions and then from the trend will try to arrive to a solution.

Subsidized Agricultural Machineries leading to decrement of sale:

To approach this problem , I again take the data both the voucher list and clean the data remove null values and missing values for the rows and columns. Add two columns for the data analysis ie Month and Type, so I can effectively do the analysis.

The cleaned up data will be arranged in a table and then converted to pivot table and and will look for the sales over different months and look for the month of higher purchase whereas will also see why particular machinery is being purchased by the customers over a long time. Also I saw a machinery named Cham Power Weeder which is being priced to 50k to 65k but apart from such high costing, customer is buying this machine with subsidized amount. So for the higher sale of machinery which is less purchased I will make a list and arrange a process by which they can buy this machinery for a lesser amount.

Pareto Analysis

Pareto analysis is premised on the idea that 80% of a project's benefit can be achieved by doing 20% of the work—or, conversely, 80% of problems can be traced to 20% of the causes. Pareto analysis is a powerful quality and decision-making tool.

Volume Pareto:

I have taken up the cleaned data and converted to the pivot table. Now I took the item details and given it to the rows and also took the quantity and made a column of sum of all quantity which is sold for the given particular item. Then I took the cumulative sum of the quantity of the sale of each item and then took the cumulative percentage in the next column and I saw that there total 252 items in total in the dataset 1 and the 80% of the revenue was generated by the 45 topmost item which is approximately the 20% of the item. So the Volume pareto is satisfying for the given dataset 1 of year 21-22

Similarly I took the volume pareto of the dataset 2 of the year 22 and saw that there are total 71 items and the total revenue generated by the topmost item which summed up to 80% was the top 10% of the items which roughly matches the volume pareto principle.

Revenue Pareto:

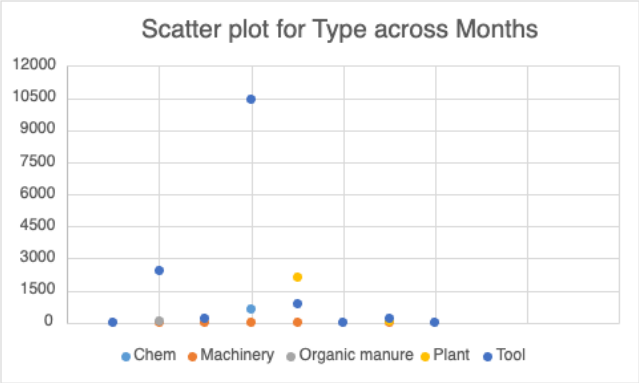
I have taken up the cleaned dataset 1 and converted to the pivot table. Now I took the item details and given it to the rows whereas the amount column was added to the values and then sorted the items list descending order and then made a cumulative sum column, then made the percentage column which shows that the first 24 items have summed up the total revenue upto the 80% , hardly satisfying the pareto principle

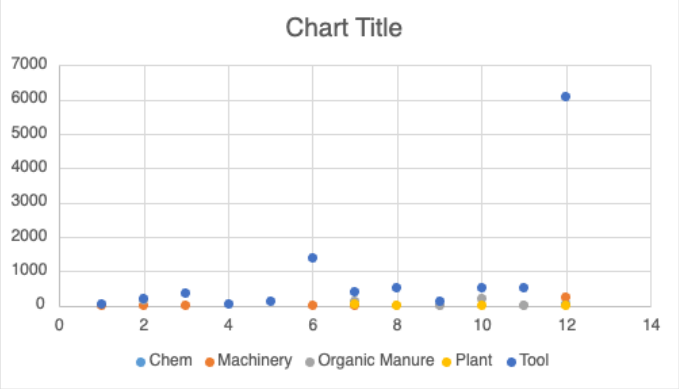
Similarly took the dataset 2 for the year 22, and I done the same steps which was for the dataset 1 and saw that the total items in this list is 71 and the top items 14 items sum up the total revenue to the 80% of the total revenue generated in that year. So the 20% of the total items in the list sum up the 80% of total revenue of the sale. Hence pareto principle exactly matches the values.

I also saw that the revenue generated was significantly due to the machinery items which make up the top 20% of the item list in the both volume and revenue pareto analysis, which tells us that the machinery is a important aspect of the customers over the year.

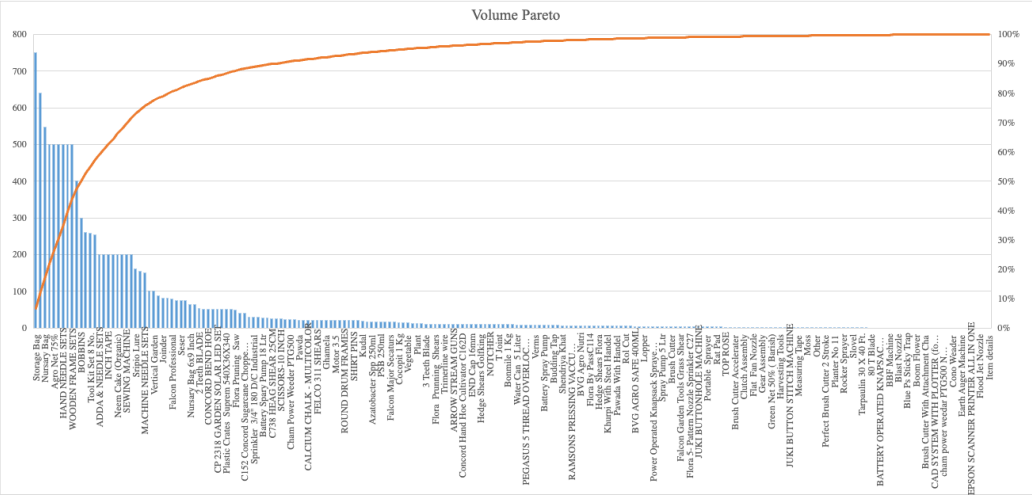
Results and Findings:

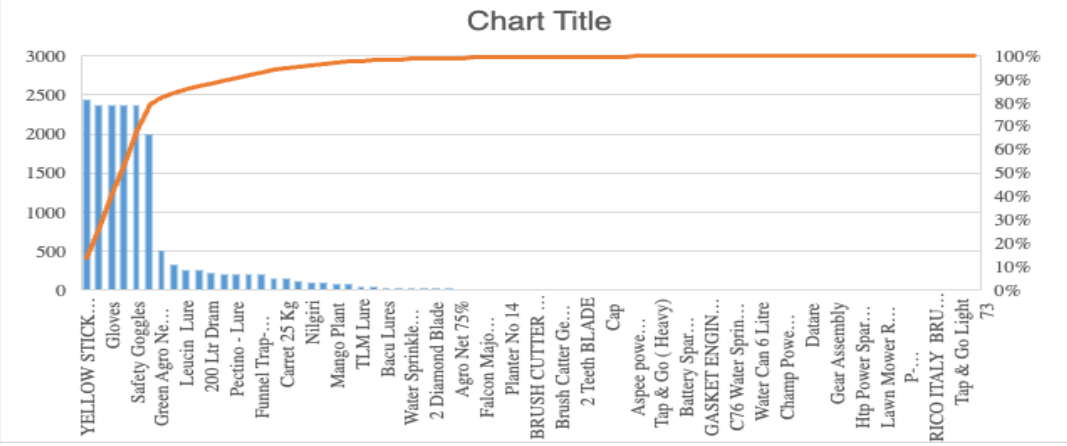
Scatter plot for both years showing the amount of items purchased in particular year 21-22(left) and year -22 (right)





Pareto Analysis : The Volume pareto of both the dataset of both years:

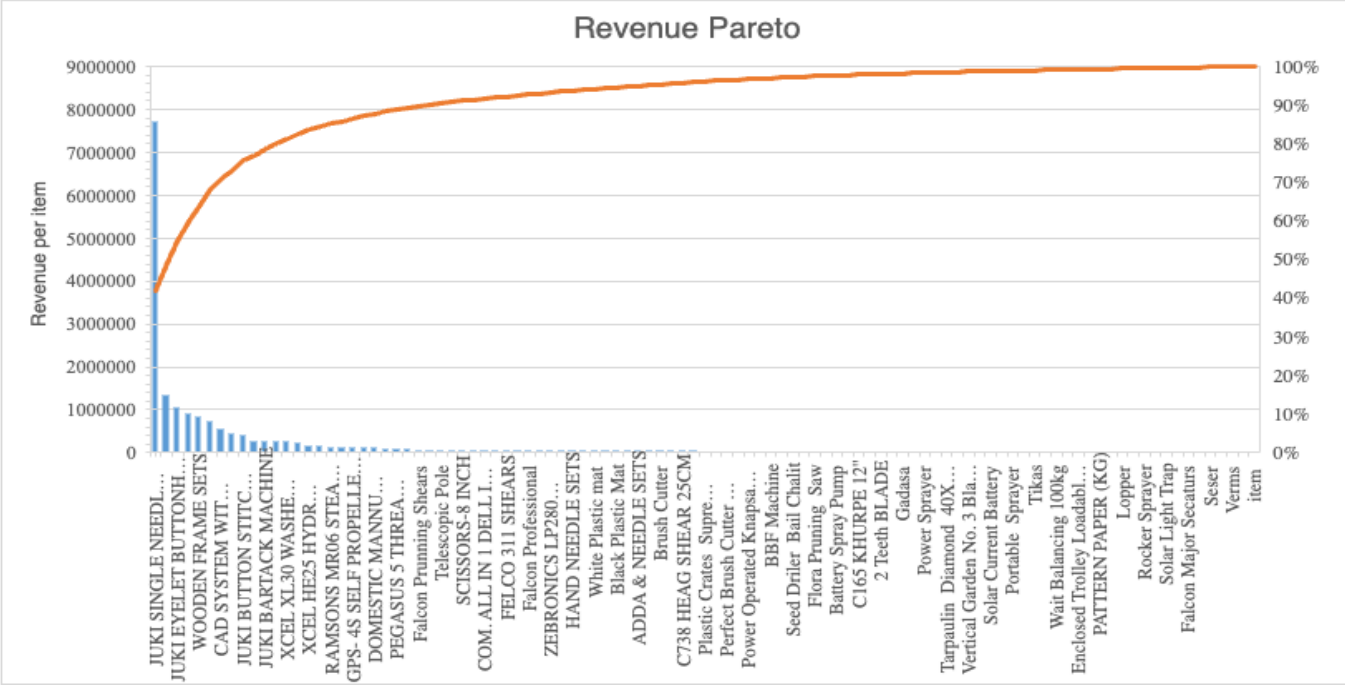




Revenue Pareto:

Dataset 1 doesnt matches the principle but the dataset2 matches the principle smoothly.

Dataset1:



Dataset 2: