Harnessing Data-Driven Inventory Management to Boost Profits in Pipes and Sanitary Ware Store

A Final report for the BDM capstone Project

Submitted by

Name: PIYUSH KANT

Roll Number: 22F3002119

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Executive Summary

This project aimed to optimize inventory management for Chaudhary Pipes and Sanitary Ware by analyzing six months of sales and purchase data. Initially, 631 items were categorized into 14 groups with the owner's input, facilitating detailed analysis. Data preprocessing included cleaning and adding a "Category" column to datasets.

Monthly sales trend analysis revealed seasonal patterns, with peak sales in September and a decline in early months, highlighting the need for strategic inventory planning. Category-wise sales analysis identified "Pipe Fittings" as the highest revenue generator, while categories like "Motors and Pump System" generated the least.

ABC analysis categorized items based on revenue, with Category A items generating 80% of revenue, requiring stable stock levels. XYZ analysis classified items based on demand variability using the Coefficient of Variation (CV), with Category X items showing stable demand, Y items moderate variability, and Z items high unpredictability.

Nine combined categories (AX, AY, AZ, BX, BY, BZ, CX, CY, CZ) were created to tailor inventory management strategies. For example, AX items, high-value with stable demand, require safety stock and automated replenishment, while CZ items, low-value with unpredictable demand, should be phased out or stocked on-demand.

Key findings include peak sales in September and high demand variability in 80% of items (category Z). Recommendations include maintaining safety stock for AX items, performing demand forecasting for AY items, and adopting make-to-order approaches for AZ items. Implementing these strategies will streamline operations, enhance profitability, and improve customer service, ensuring sustainable growth for Chaudhary Pipes and Sanitary Ware.

Detailed Explanation of Analysis Process

Meeting with the shop owner

The project started by a meeting with Mr. Sanjay Kumar the owner of “CHAUDHARY PIPES AND SANITARY WARE”. During the first meet, I introduced myself and explained him the purpose of that meeting which was this capstone project. I asked him for a brief information about his shop and the items he sells in that shop. So, the shop is situated on the

NH22 in a small town Jehanabad (Bihar). The following point he told about his shop during the meeting :

- His shop contains the items essential for water supply in kitchens and bathrooms, including pipes, taps, water tanks, sanitary products, and hardware for their installation.
- He deals with both retail and wholesale. Means the business is both B2B and B2C. Also, he is running that firm for last seven years.
- The shop offers products in a range from local to international brands like Pearl, Parryware and many more.
- Most of the items he sells are of Ashirvad brand. For this he has got many awards in the recent years. This year he got the award for Highest sales in Plumbing in Ashirvad and also got free trip to Malaysia.

- The vision of his firm is to give quality products to his customers and be the top choice among the sanitary ware shops in Jehanabad.

Then, for the main purpose of this meeting I asked him about the business problems he might be facing in his firm. An hour long discussion went on this topic and he concluded that his shop is going well without any problem and if I want then, I can improve his shop performance and help him boost his overall profitability and increase customer satisfaction.

Problem Identification

According to Mr. Sanjay, his shop is doing well. But, as his firm comes in plumbing and water supply sector, the main challenge faced is inefficient inventory management. Because, the shops under this sector sells large variety of different products used in the plumbing

works. And, inventory overstocking and stockouts are very common which leads to loss, break in cash flow, and also affects customer satisfaction badly. So, considering all these I identified the following problems :

- The first problem is inefficient inventory management. This is mainly due to a huge variety of products in this field.
- The second problem is obviously the profitability. This problem may be due to many reasons. But, the inventory management is also a cause.
- The third one is a depression in customer satisfaction. This is also due to constant stockouts.

Data Collection

After problem identification, the time was to collect the data which was relevant to my analysis and appropriate for solving the problems identified. But, the owner told me that he only had sales data and purchase data of his products. Since the problem is on inventory management, the stock-in and stock-out data was necessary for the analysis.

I got the last six months sales data and purchase data of the shop. So, this data can give me insights on the sales trend, so that overall profitability can be increased. But, for the inventory optimization it doesn't suffice. So, I got these two data in the csv format and decided to go with it and extract the information for the analysis which would help in the inventory optimization, profitability boosting, and customer satisfaction.

Data Pre-processing and Cleaning

I used both MS Excel and Python (Pandas) for all the analysis of this project. The data collected form the shop was in csv format. The following steps gave the cleaned data for analysis :

- I opened both the sales and purchase data on excel sheet and removed the contents of first three rows. These rows was containing the information about data like Shop name as heading, and the dates from when to when the data is collected.
- 

Figure 1

These are the columns present in sales data. I removed the columns “Vch/Bill No”, “TIN/GSTIN No.”, and “Material Centre”. These columns are of no use for the analysis.

- 

Figure 2

These are the columns present in the purchase data. I removed the “Material Centre” column from this dataset. The material centre has only one type of entry in all rows “Main Store”. So, this column is of no use.

- Then removed the last row of both the datasets. The last row contained the totals of quantity and amount. This may cause problem during analysis.
- Saved both the datasets as sales.csv and purchase.csv. Then, used these csv files for further data cleaning and analysis using python.
- Now, I used pandas to fill the null values into each columns of both datasets.
- First imported pandas and used it to read the csv files.
- Using the info() function, I checked for the details of the datasets. Found that, there are 12738 rows and 8 columns in sales.csv, and 1751 rows and 7 columns in purchase.csv.
- Also, this function gave me the exact number of null values in each columns. There were null values in “Date” and “Particulars” columns of sales_data, and in “Date” and “Name” columns of purchase_data.
- The null values in the Date column is due to the Item details entries for a particular customer on a particular date. The different items purchased by a customer on a particular date are added in different rows where but the date is mentioned in the row only for the first item in each transaction. Hence, the null values in the date column must be filled with the previous entry in the date column.
- To fill the null entries in date column I used the fillna() function with “ffill” method. This fills the null entries with the previous entry.
- Due to the same condition as of “Date” column, the “Particulars” column of sales_data and “Name” column of purchase data have null values. So, these null values are again filled using the “ffill” method in “fillna” function.

Product Categorization

After preprocessing and cleaning of data I come to know that there are 631 different types of items present in the dataset. So, analyzing this wide range of products is a very complex task. Hence, I decided to categorize these products into different categories so that the analysis can be done on category basis.

I again met the owner of the firm Mr. Sanjay. After a long discussion, categorized each items into 14 different categories. The categories of products are following :

1. Pipes
2. Pipe Fittings
3. Traps and Bends
4. Plumbing fixtures and Accessories
5. Valves and Taps
6. Lubricants and Adhesives
7. Fasteners and Connectors
8. Toilet and Sanitary Accessories
9. Electrical Components and Wiring
10. Drainage and Waste management
11. Motors and Pump System
12. Structural and Support Rods
13. Storage and Water supply
14. Miscellaneous Tools

The items in each category : Product Category

Then added a column “Category” to both the datasets. For this I first created the Product_Category.csv file and the created a dictionary which maps product names to categories. Used this dictionary to map each entries of the datasets to the appropriate product category.

Monthly Sales Trend Analysis

The monthly sales trend analysis is important for identifying the sales pattern of a business. This helps in understanding the overall sales performance for business decisions related to inventory management.

For this analysis I used matplotlib for the visualization of monthly sales trend. The following process helped me in finding the insights from this analysis:

- I first converted the "Date" if sales_data into pandas datatime format using to_datetime.
- Then created a dataframe named monthly_sales in which there are three columns "Month" , "Amount", and "Moving_avg". The 'Month' contains the month names, 'Amount' contains the total revenue collected in that particular month and "Moving_avg" contains the average of the consecutive three months.
- Then calculated the monthly total amount from the sales data. Then, converted the months into month names and calculated the "Moving_avg" using the code below.
- This created the monthly_sales dataframe which contain all the details for the visualization of monthly sales trend. Then, I used the line plot with 3 month moving average to visualize these data .

Month	Amount	Moving_Avg
April	247850.38	NaN
May	400162.89	355883.43
June	419637.02	379041.9666666666
July	317325.99	350943.19333333336
August	315866.57	357689.11333333334
September	439874.78	NaN

Figure 3

Category-wise Sales Analysis

There are 631 different types of products. The analysis of each items one by one is not possible. So, I am doing a category-wise sales analysis. The categories for each items are already added into the sales data. Hence, the following process is used for the category wise sales analysis:

- I grouped the total amount by the “Category” column to analyze the revenue generated under each category.
- The total sales amount for each category is calculated which gave the summary of sales performance for each product category.
- For the visualization I used created a bar chart displaying the total sales amount for each category.

ABC Analysis

For optimization of inventory, the focus must be on the most critical items in the shop. In order to do this, I have to do ABC analysis in which the items are divided into three categories A, B, and C based on their contribution to the revenue generated. To do this, I performed the following methods:

- I grouped the dataset by “Item Details” to calculate the total revenue generated by each item. This gives a record of the total revenue generated by each item in last six months.
- Then sorted the items in descending order based on their total revenue. Such that the highest revenue generating items are at the top.
- For a better understanding of the contribution of each item to the total revenue, I calculated cumulative revenue and cumulative percent.
- Now, the classification of items into A, B, and C is done on the basis of following criteria: The items whose cumulative percentage is less than or equal to 80% are categorized in A, the items with cumulative percentage greater than 80% but less than or equal to 95% are categorized under B, and the remaining items are categorized under C.
- Then, summarized the number of items in each category, their total revenue, and their percentage contribution to overall revenue.
- For visualization I used a pareto chart to plot the cumulative revenue curve and a bar chart for each category to see the revenue contribution of A,B, and C visually.

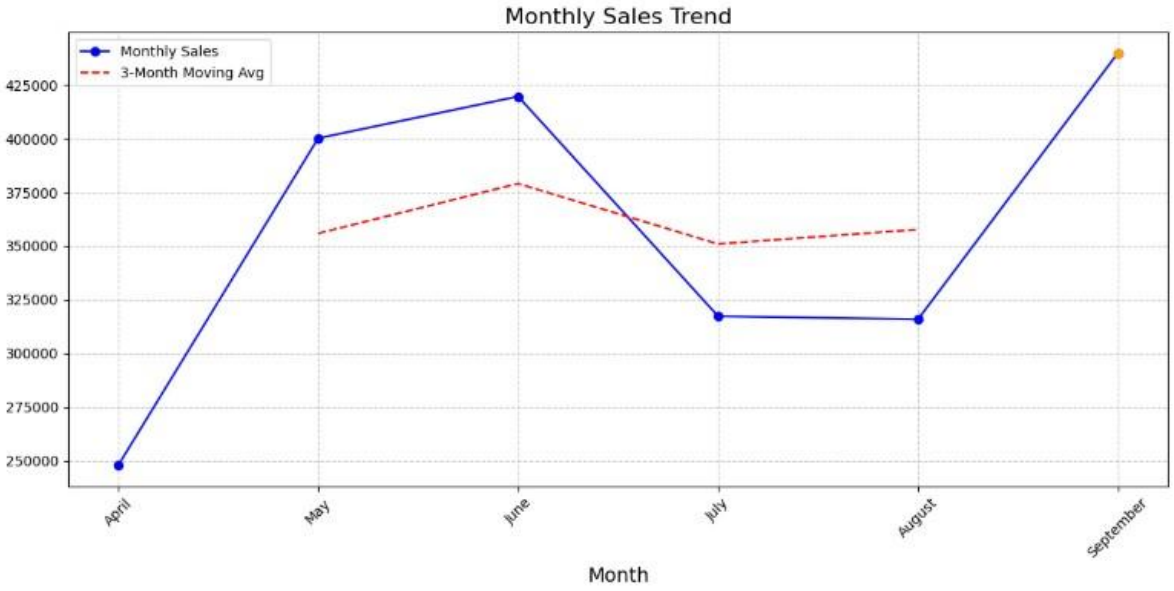
XYZ Analysis

Unlike ABC analysis, this analysis categorize items based on their demand variability. This analysis is done by following the process below:

- Created a dataframe by grouping the data by each item. Then calculated the mean demand and standard deviation.
- Then, calculated the Coefficient of variation (CV). The formula to calculate the CV is, CV = (Standard Deviation) / (Mean Demand).
- After calculating the CV, I classified items into X, Y, and Z categories in such a way that items with CV less than or equal to 0.10 comes under X, items with CV greater than 0.10 but less than or equal to 0.25 comes under Y, and the highly variable items means those with CV greater than 0.25 comes under Z.
- For the visualization of distribution of items in categories X, Y, and Z, I used the pie chart. Also, I used a violin plot that shows the distribution of CV for each category.

Results and Findings

Monthly Sales Trend Analysis

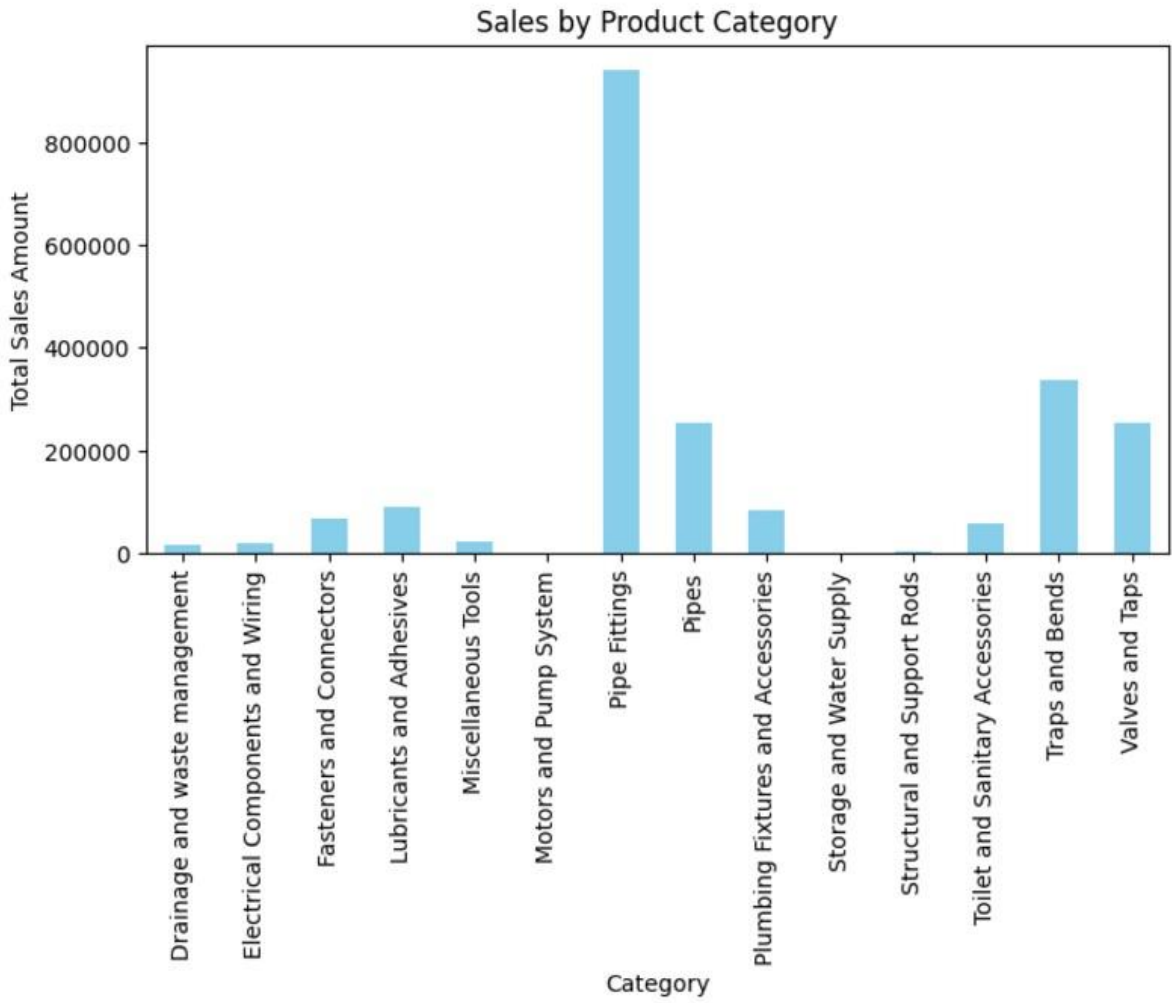


This graph provides a detailed view of monthly sales trend. I added a monthly moving average line (red dashed), that is rolling over a three months period means it takes the

average of that month and the previous two months. The following insights are drawn from this analysis :

- The orange point shows the peak of the monthly revenue. In the month of September, the shop collects highest revenue.
- The lowest revenue collected in the first month of this financial year.
- Since the red-dashed line is overall decreasing, it shows a decline in overall sale.
- The red-dashed line indicates that there is need to conserve resources or optimize stock levels to boost demand and hence profitability.

Category-Wise Sales Analysis



The bar chart above gives the brief visualization of the category wise revenue collection.

- The "Pipe Fittings" items contributes to the highest overall revenue collection in last six months.
- Motors and Pump system, Sorage and swater supply, and Structural and support rods are providing the least revenue collection in last six months.
- Most of the revenue collection in last six months are due to sales of "Pipe Fittings", "Traps and Bends", "Valves and taps". These product categories are the top 3 contributors to the six months sale.
- Toilet and sanitary accessories are very vulnerable products. Also, this product category is one of those which are giving the least revenue. So, these products should be very less in stock and use the resources occupied by these products to focus on other higher revenue generating products.
- "Fasteners and connectors" and "Lubricants and adhesives". These products contributes low revenue. However, these are very essential as almost every customers who buy any other products alse need this for the pupose of installation. Even if these products are providing low revenue, the stocks for these products must be full. If a customer doesn't get these items, he'll go to a different shop giving a chance to the competing shop to attract the customer.

ABC Analysis

The main purpose of performing this ABC analysis in my sales data is to categorize the items in the shop from most to least critical. This enables to focus on mainly on the items that come under category A.

Summary of Item Categorization:

ABC Category	Total Items	Total Revenue (₹)	Percentage Contribution (%)
A	95	9,264,436.21	79.87
B	155	1,752,736.65	15.11
C	381	582,209.17	5.02

Table 1

The following points are drawn from the summary above:

- Almost 80% of the total revenue is generated by the items in category A. The 95 items present in the category A are high-priority items and their availability must be stable in the shop.
- The items under category B accounts for almost 15% of the total revenue collected in the last six months. A periodic check is must for items in this category to avoid unbalance inventory.
- The C category items contributes to the lowest percent of the total revenue. However, it contains the largest number of items among other two. These items are least critical items. These items should be stocked in minimum amount to use the resources on the other two categories.

The following visualizations gives a better view of the analysis.

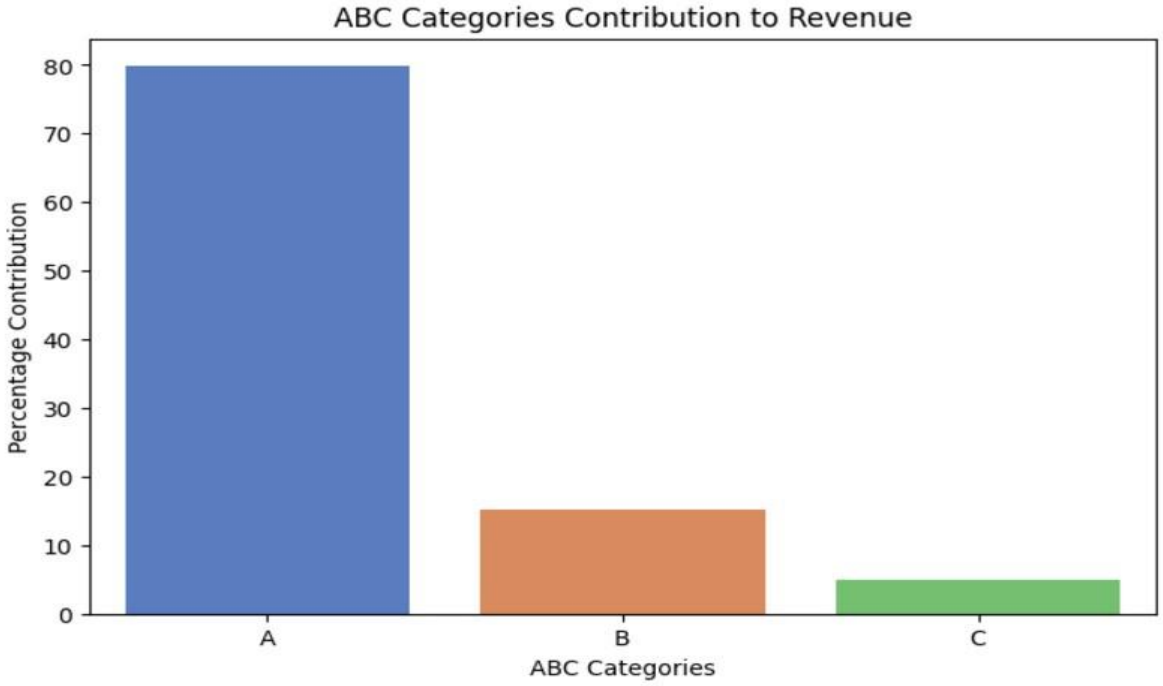


Figure 4

To see the items in each categories: Click Here

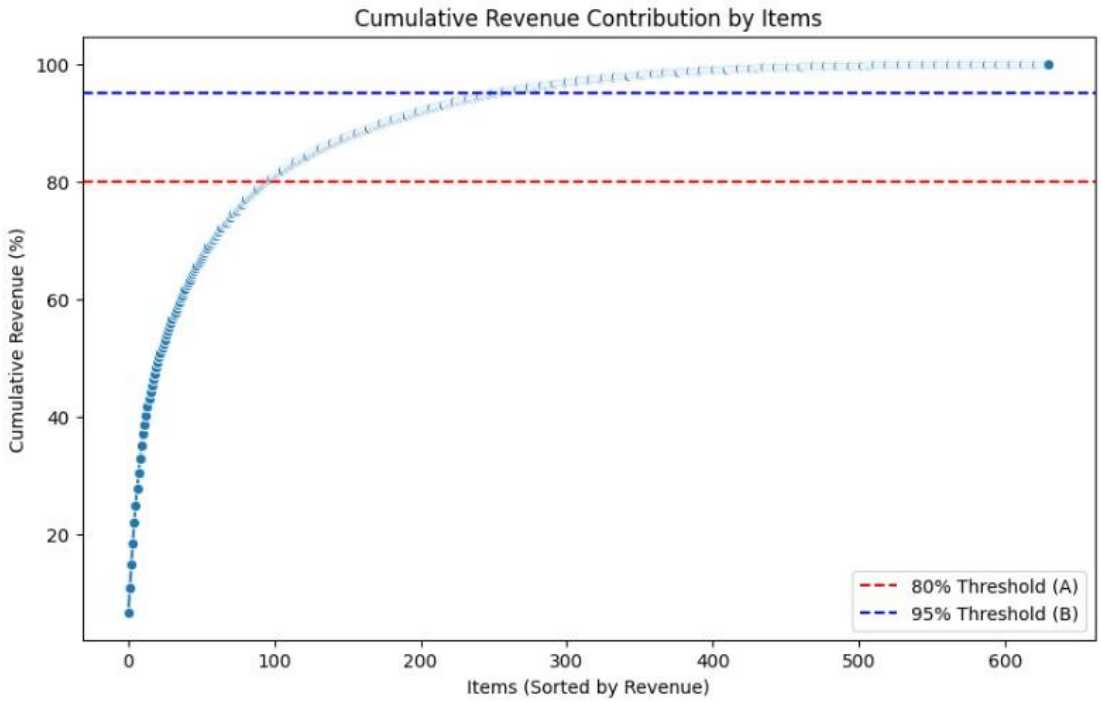


Figure 5

Figure 5 gives the following information:

- The leftmost item on the X-axis is the item that generates the most revenue, the next item generates the second most revenue and so on. X-axis is representing the index of the items.
- Y-axis is representing the cumulative percentage of revenue contributed by the items. Starting from 0% and gradually increasing to 100%.
- The cumulative revenue percentage is shown on the line plot connecting the data points.
- The red line indicates the point where the cumulative revenue reaches 80%. Items to the left of this line are typically categorized as “A” items.
- The blue line indicated the point where the cumulative revenue reaches 95%. The items between the red and blue lines are categorized as “B” items. Items to the right of the blue line are categorized as “C” items.

The ABC method is also known sometimes by “Always Better Control” method. This analysis enables a business to spend their time and energy primarily on the items in “A” category. It doesn’t mean to not pay attention on the items in B or C. But, the main focus needs to be on the items in A.

XYZ Analysis

In order to identify which items have the most and least stable demand, the XYZ analysis is used. During this analysis the inventory items are categorized on the basis of demand variability. The items are categorized into X, Y, and Z in such a way that items in X have stable demand, items in Y have moderate demand variability, and items in Z have the most unpredictable demand.

Coefficient of Variation is key here. All classification is done on the basis of CV. It measures the demand variability. Low CV means demand is stable and High CV means unpredictable demand. The pie chart below shows the distribution of items in each category based on their CV.

The pie chart in figure 6 gives the following information:

- About 80% of the items are categorized into Z. Hence, 80% of the items have high CV, means the items in this category are unpredictable.
- The items in X covers almost 18% of the total items in the shop. The most stable items in the shop are 18% of the total items. These items are predictable.
- The items in Y are very less almost 1.6%. It shows that the moderately stable items are very few.

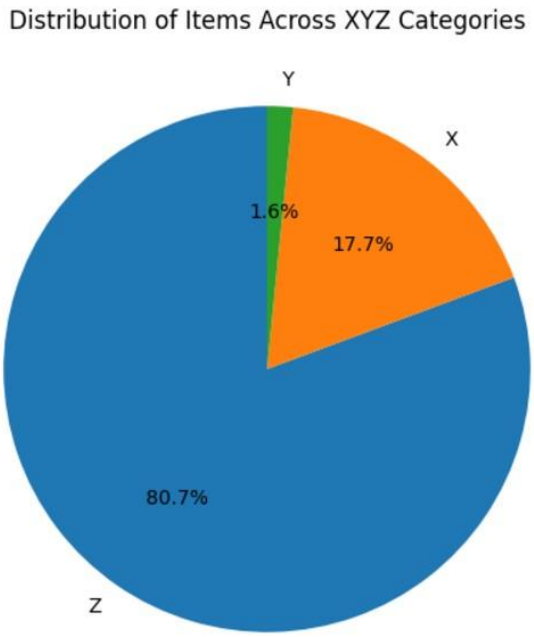


Figure 6

For a deeper analysis of the variability of demand for each category, the following violin plot and heatmap is created.

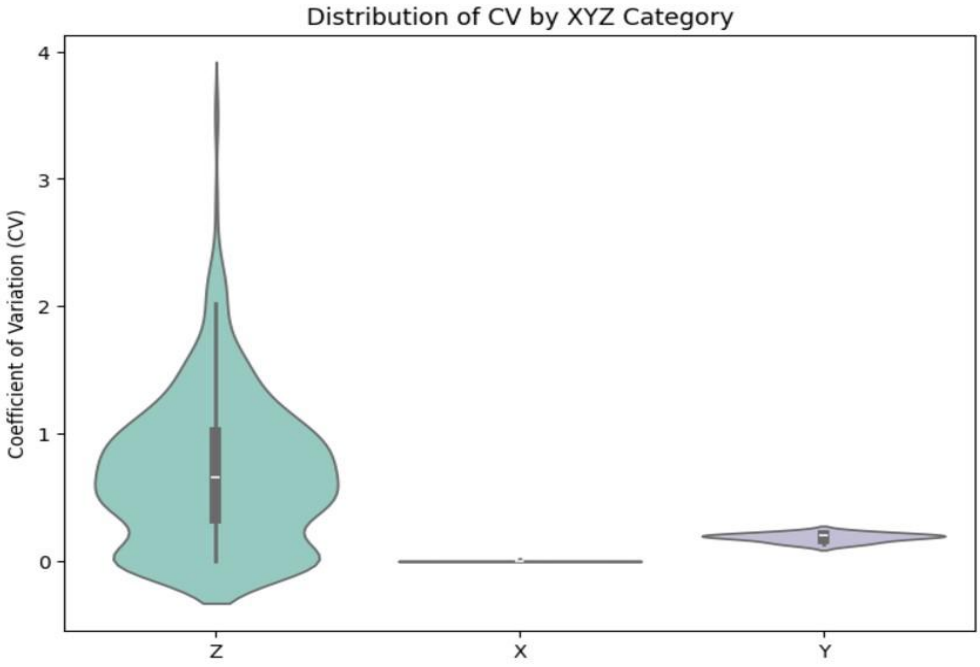


Figure 7

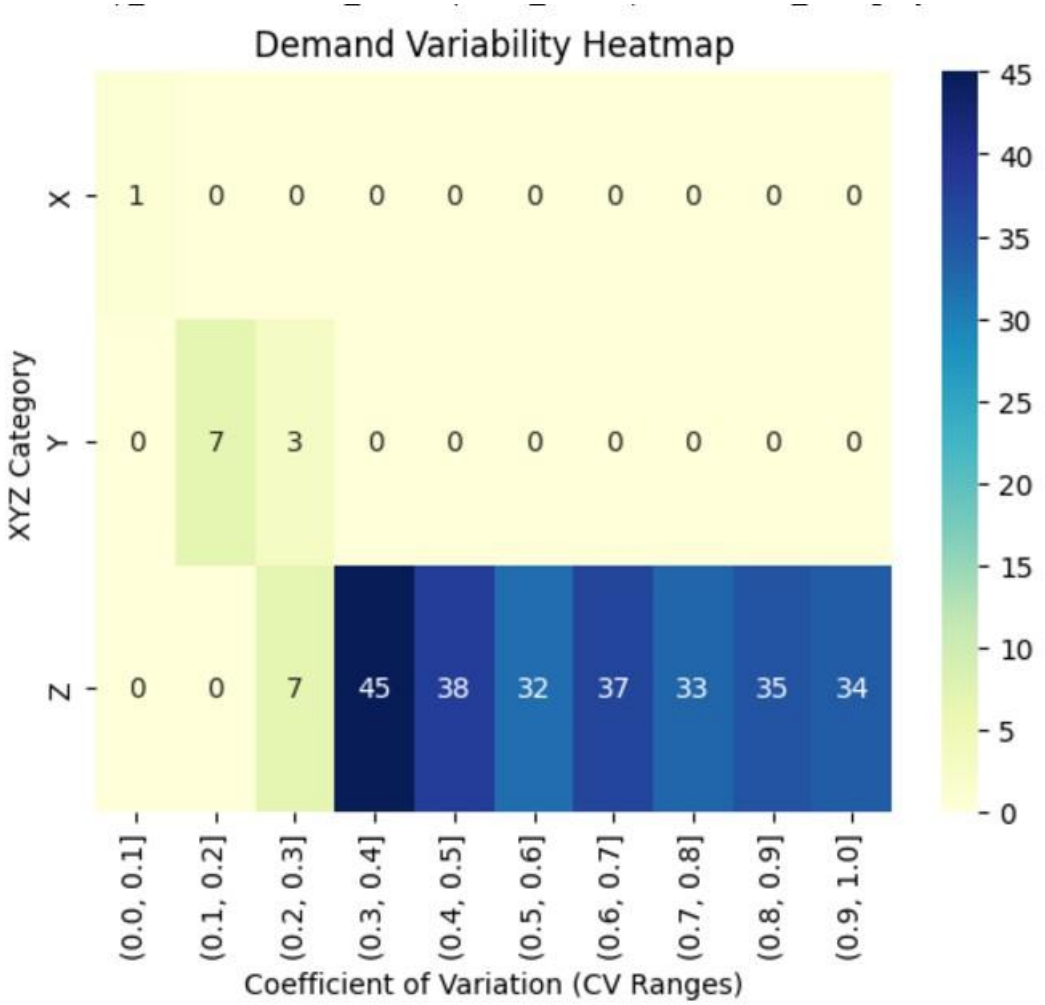


Figure 8

The following observations are drawn from the violin plot in figure 7:

- For X, the CV is almost zero. This suggests that the products in this category have variability as good as zero. So, these items can be maintained without the need to advanced inventory management techniques.
- The white dot at the centre is the median basically. In Y, there is a narrow spread around the median CV. This suggests that the items in this category have occasional spikes in demand.
- The wide spread in the Z category tells that there is high demand variability. Due to high variability the items in this category demands frequent restocking. This leads to impractical expenditure.
- The long tail in the Z category symbolizes the demand uncertainty. Hence, to manage the loss there should be variable pricing like demand variability.

- The outliers in Z category represents some rarely sold items.

Combined Categories

From ABC Analysis, I have three categories of items based on contribution to the total revenue. The categories are A, B, and C.

From XYZ Analysis, I again have three categories of items based on their CV. The categories are X, Y, and Z.

Now,

I am again categorizing the items based on the combined categorization of items in both ABC and XYZ category. After this combination the following categories are generated: **AX, AY, AZ, BX, BY, BZ, CX, CY, and CZ.**

The items in these categories can be seen here: Combined Category

	X	Y	Z
A	High Value, Low Variability	High Value, Moderate Variability	High Value, High Variability
B	Medium Value, Low Variability	Medium Value, Moderate Variability	Medium Value, High Variability
C	Low Value, Low Variability	Low Value, Moderate Variability	Low Value, High Variability

Table 2

Interpretation of Results

After performing both the ABC and XYZ analysis, I combined the categories obtained in both the analysis to get 9 combined categories. Then, listed the items in each combined categories to interpret the results category wise.

AX Category:

Products in this category generates the most revenue and also have low variability in sales. The following insights are drawn:

- The items in this category are high-priority items because these are both profitable and predictable, ensuring steady revenue collection.

AY Category:

These items are valuable but experience fluctuating demand, possibly due to season changes or varying customer preferences.

AZ Category:

These items are high-value but highly unpredictable in demand, which can lead to challenges in maintaining optimal stock levels.

BX Category:

These items have a moderate impact on revenue and stable demand, requiring regular but straightforward management.

BY Category:

These items have moderate value and demand variability, indicating some level of risk and the need for careful monitoring.

BZ Category:

These items are moderately valuable but highly unpredictable, posing a risk of obsolescence or stockouts.

CX Category:

These low-value items have stable demand and are low priority, though they require minimal effort to manage.

CY Category:

These low-value items experience fluctuating demand but pose minimal risk due to their limited contribution to revenue.

CZ Category:

These items are low-value and unpredictable, contributing minimally to revenue and often leading to unnecessary holding costs.

Recommendations

AX Category:

Maintain safety stock to ensure the availability of these high-priority, high-demand items. Implement automated replenishment systems to manage inventory efficiently and prevent stockouts.

AY Category:

- Use historical sales data and market trends to perform demand forecasting, accommodating seasonal and preference-based fluctuations.
- Maintain a moderate level of safety stock to avoid overstocking while ensuring availability during peak demand periods.

AZ Category:

- Avoid overstocking by adopting a make-to-order or just-in-time approach where feasible.
- Conduct risk analysis for demand forecasting and establish partnerships with suppliers who can offer flexible lead times to respond quickly to changes in demand.

BX Category:

- Maintain steady stock levels based on reliable demand forecasts and implement regular reordering processes.
- Regularly monitor these items to ensure supply meets demand without excessive inventory buildup.

BY Category:

- Implement a buffer stock strategy to handle fluctuations in demand and ensure product availability.
- Monitor these items seasonally or based on observed trends to adjust inventory levels accordingly.

BZ Category:

- Avoid overstocking by using a demand-driven restocking approach.
- Assess if these items can be replaced with more predictable alternatives to reduce risk and streamline inventory management.

CX Category:

Utilize bulk ordering to reduce procurement costs and maintain stable stock levels with minimal monitoring effort.

CY Category:

- Keep low stock levels and monitor sales trends occasionally to adjust inventory as needed.
- Minimize focus on these items as their impact on overall profitability is limited.

CZ Category:

- Consider phasing out or replacing these items if they contribute minimally to revenue and lead to high holding costs.
- Stock only on-demand or in very small quantities, and conduct a cost-benefit analysis to determine their necessity in the inventory.

By implementing these recommendations, Chaudhary Pipes and Sanitary Ware can streamline its operations, enhance profitability, and provide a better customer experience. The adoption of data-driven strategies will ensure sustainable growth and success in the competitive market.