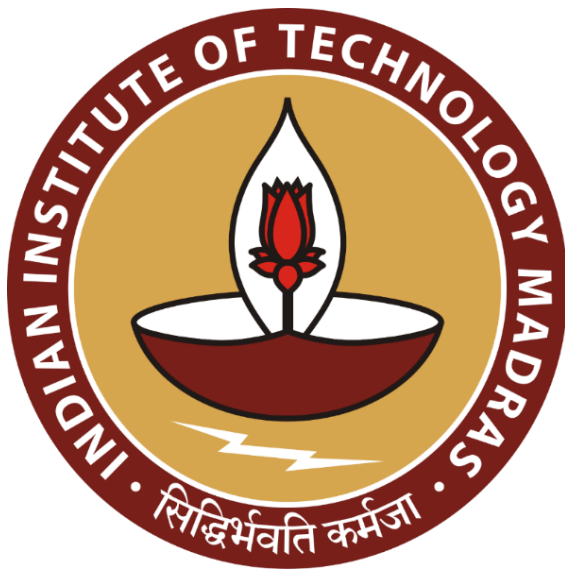


Analyzing Stock Sales of Retail Clothes Seller

Final Report



Name: - Anmol Sarraf
Roll: - 22f3002374

Table of Contents

<i>Title</i>	<i>page no</i>
1. Executive summary	2
2. Detail explanation of analysis/method	3
2.1. Data collection	3
2.2. Data Preprocessing/Cleaning	3
2.3. Methods used and why	4
3. Results and findings	6
3.1. Time series analysis	6
3.2. Pareto analysis	9
3.3. Bar chart	10
3.4. Scatterplot	12
3.5. Item-wise Comparison	14
4. Interpretation of result & Recommendation	16
4.1. Result Interpretation	16
4.2. Recommendation	17

1. Executive summary

The Focus of this project was a small retail shop “New Dolly Dress Lok” located at bara bazaar, Madhubani which sells traditional Indian readymade clothes like *salwaar suit*, *kurta*, *frock*, *shirt etc.* in B2C mode. The owner believes that the business is losing its revenue and wants to know why. He also believes that we need to have better stock management to address stock issues and he also wanted to move toward online business.

All data required was collected daily. Although no major preprocessing was required to clean the data. Some small step like categorization of multiple same items in same name and manually allocating week number. Descriptive stats, bar charts, pivot tables were used to draw conclusions.

Data revealed that starting week 1 to 4 they were highest in revenue and lowest in missed sales and latter weeks were lowest in revenue but highest in missed sales. This happened because of stock unavailability as it ran out of stocks over time. Salwaar suit and kurta were highest sold in 3-month span, but they were also highest in missed sales indicating stock issues.

All this happened due to the stocking issue, as the owner did not have any record of how much stock they have. This creates a bad perception of shop in the market and due to low sales overstocking also started kicking in. I recommend that the owner should collect data daily and study on weekly basis to get trends in the market and he should offer some discounts and other incentives to boost the sales of low demand items.

If the owner successfully implements my recommendation, then I am sure that business will bloom.

2. Detailed Explanation of Analysis Process/Method

2.1. DATA COLLECTION: -

All the data required to address the business problem was manually collected with the help of the shop owner. Sales records were noted down either by me or the owner in a notebook daily. To ensure accuracy and completeness, we cross-verified the total earnings recorded each day with the actual cash collected, thereby confirming the reliability of the data.

Name of item sold, quantity of item sold, price of each item and date it was sold on was collected which together forms sales data set which was latter logged into excel worksheet manually. This also goes same for customer survey. Feedback from customers who didn't buy anything and item which they were looking for was taken.

With all these I have three sets of datasets collected over the Span of 3 months from **1st April 2025 to 30th June 2025**. Sales dataset for all entries of sales, Customer feedback dataset which contain feedback of customers and Stock dataset which contain details of all stocks available taken on 31st March 2025.

Data :- [Click here](#)

2.2. DATA PREPROCESSING/CLEANING

All values in the dataset were recorded in a well-structured format from the beginning, which eliminated the need for extensive data cleaning or handling of missing values. However, certain **preprocessing steps** were performed to simplify analysis.

To ensure consistency and simplify analysis, similar product variants were grouped into broader item categories. For instance, all types of **Kurtas**—including cotton, silk, and party-wear—were consolidated under the unified category “**Kurta.**” Similarly, different styles of **Salwaar suits** such as cotton, printed, and embroidered varieties were grouped as “**Salwaar suit.**” Various types of **Jeans** including cotton, denim, and polyester were merged into the single category “**Jeans,**” and all shirt variants like cotton, polyester, and casual were classified under “**Shirt.**”

This categorization approach was applied not only to the sales dataset but also extended to the customer feedback and stock data. By standardizing item names across datasets, it became easier to analyse trends, compare performance, and identify patterns without confusion caused by product name variations.

A new column indicating the **week** was also added to the dataset. All transaction dates from **1st April to 30th June 2025** were categorized based on the week they fell into, starting from Week 1. This enabled **weekly trend analysis** for both sales and revenue patterns. This was done manually categorizing week number according to dates. This was done on both sales and customer feedback dataset.

To calculate Total Price of stock sold in a day. I have taken product quality sold and per item to get total Price.

02-04-2025	Frock	2	₹ 500.00	=C5*D5	W1
------------	-------	---	----------	--------	----

All the values were written in excel sheets onto daily basis. Later data was imported into Jupyter notebook using python panda's library to perform more statistical analysis and tables for creating graph and visualization.

2.3. METHODS USED AND REASON

Descriptive statistics were used to calculate key metrics such as mean, median, total revenue, quantity sold, and standard deviation for each item category. This method allowed me to quickly identify **top-selling items**, **average pricing trends**, and the maximum and minimum quantities sold. These insights help in identifying high-demand items, which may lead to understocking if not replenished, and low-demand items, which can result in **overstocking**. Additionally, standard deviation was useful in assessing the variability and price consistency of each item, as well as identifying potential skewness in the sales distribution.

Time series analysis was performed using **week-wise grouping of data**. By categorizing all entries by week, I analysed **revenue trends**, **sales variation**, and **customer feedback patterns** over the 13-week span. This enabled me to spot critical trends such as **increasing missed sales** in later weeks and their **direct impact on revenue**, as well as **seasonal demand surges** (e.g., for Salwaar suit and Kurta). This is also done on months basis to study revenue generation by each item. I collected all weekly and monthly data using pivot table and then use the graphs.

A **pareto analysis** was done on basis on 80/20 rule to identify which item contributed most of revenue. This was crucial in recognizing that a few items like Salwaar suit and Kurta generated most of revenue. Hence it should be in priority while inventory planning. Using pivot table all the sales of individual clothes were taken and after that cumulative sum was calculated and later cumulative percentage was calculated by dividing the cumulative sum at individual item with total of cumulative sum. And then using combo in excel pareto chart was drawn.

Items	Sales	Cumulative sum	Cumulative percentage
Salwaar suit	₹ 1,05,350.00	₹ 1,05,350.00	0.27
kurta	₹ 1,02,750.00	=M134+L135	0.53

Fig-2.1 calculation of cumulative sum

Items	Sales	Cumulative sum	Cumulative percentage
Salwaar suit	₹ 1,05,350.00	₹ 1,05,350.00	0.27
kurta	₹ 1,02,750.00	=M134+L135	0.53
undergarments	₹ 3,100.00	₹ 3,93,650.00	1.00

Fig-2.2 calculation of cumulative percentage

A **bar chart** comparison between successful and missed sales was used to identify revenue loss due to stock issues, especially for high demand items. After that calculating ratio between sales and missed sales revealed which item had more stock issues.

A **scatterplot** used to study the relationship between item price and customer demand which resulted in a negative correlation, indicating low sales for higher prices items and vice versa. Average prices of all items and count of total sales were collected using pivot table and scatter plot was created.

A **reason-wise breakdown of feedback** (both weekly and item-wise) was also conducted to understand the main cause of missed sales. This includes issues like no stock, price negotiation failures, staff unavailability and budget constraints. This was also made by using pivot table.

Pivot tables were used extensively to create tables that was used to create charts and graphs. All the charts that were described above were made using pivot table

3. Results and findings

After applying all the methods, here are some of the interesting findings and facts.

3.1. TIME SERIES ANALYSIS

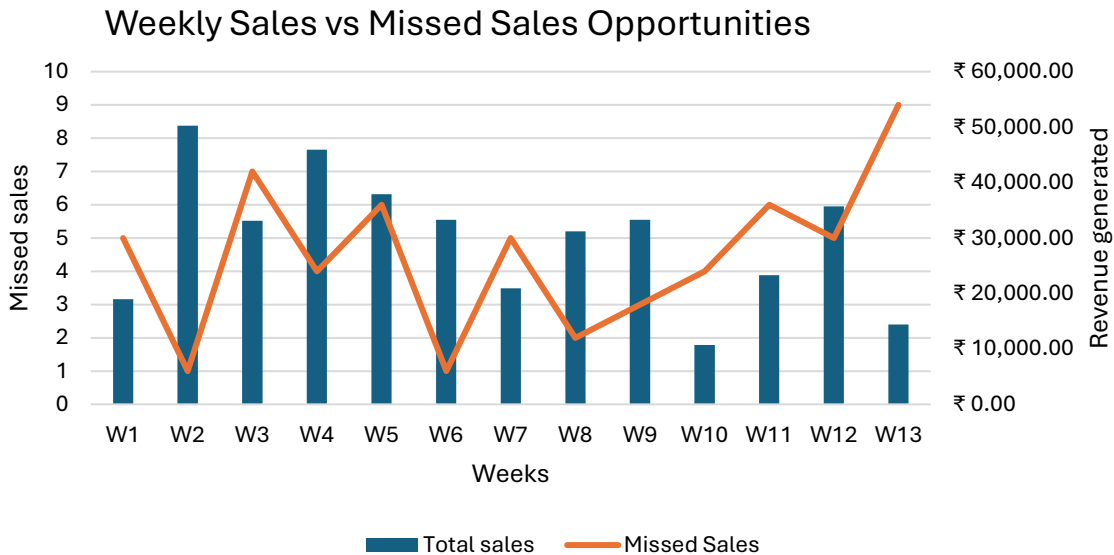


Fig-3.1 "This chart shows weekly total revenue (bars) and number of customers who didn't buy (line). A visible inverse trend indicates sales loss due to stock or service issues."

The dual-axis line chart illustrates a comparison between weekly revenue generation and the number of missed sales over the same period. Notably, Week 2 recorded the highest revenue along with the lowest number of missed sales. In contrast, Week 13 showed the lowest revenue and the highest number of missed sales. Similar trends can be seen in week 7, week 10 and week 11.

Although there some weeks like week 4 and week 5 which has high revenue along with high missed sales this indicates that this week had potential to earn more. If we look at graph after week 8 to week 13 there is increasing trend in missed sales and hence it has made a huge impact on revenue generation.

Finally, we can say that from this chart that Loss of revenue which the owner was talking about was indeed correct. Missed sales was major problem which has cause loss of revenue to this business.

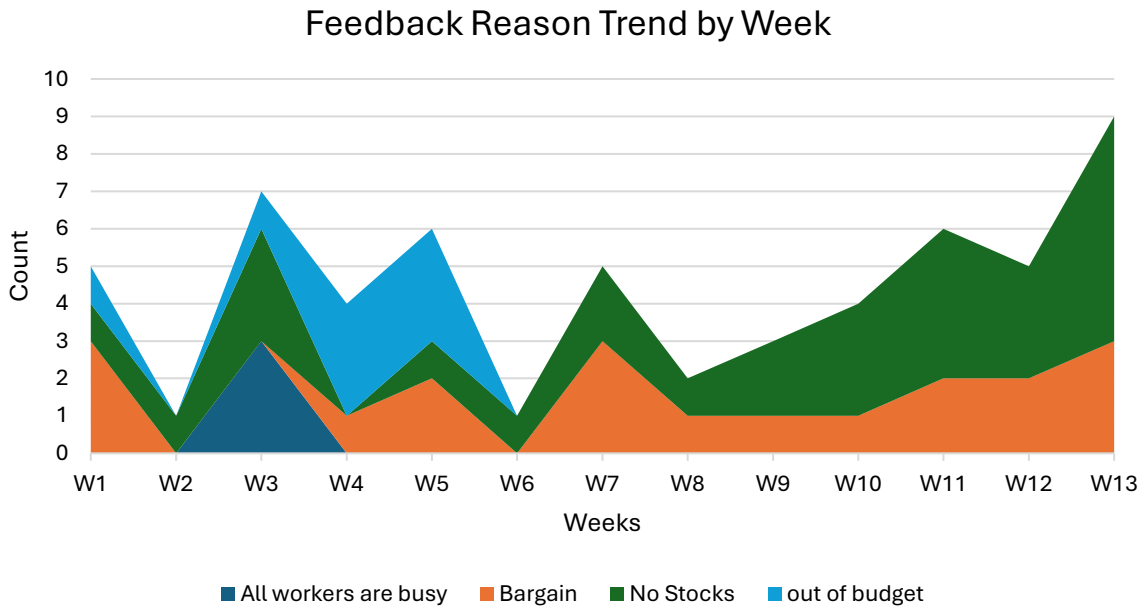


Fig-3.2 Weekly Trends in Customer Feedback Reasons

The Above graph shows week-wise distribution of customer complaints based on different reason so far. As we can see that from starting weeks customer have been complaining about “Stock issues” which became worse after week 8, indicating that there is a critical gap in stock management from the start. This suggests that business was unable to meet the customer’s expectation possibly due to the unavailability of preferred item, colour or size).

In addition to stock issues, Bargain failure also contributed to missed sales starting and more in latter weeks. Even if they do find their desired stock cost may be another issue for customers. Furthermore, some customers also complained about the **lack of staff or all workers being busy** appeared prominently in the middle weeks (e.g., W3- W5), pointing towards operational bottlenecks. Some people also cited that the store did not have items in their budget range, indicating the need for **more variability** in some items.

Overall, the rising trend in missed sales due to multiple overlapping reasons creates a **domino effect**, where poor stock availability, staff shortages, and price inflexibility collectively impact customer satisfaction and revenue.

3.2. PARETO ANALYSIS

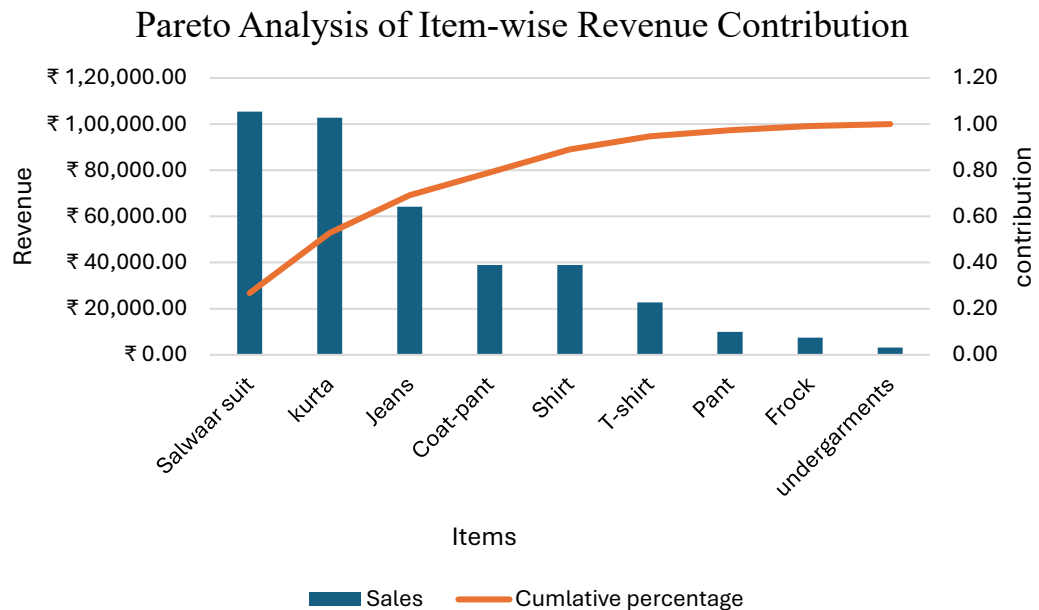


Fig-3.3 pareto analysis of items and revenue

Above Pareto chart provides clear, visual explanation of which item contribute most to total sales - based on 80/20 rule (Pareto Principle). Bars represent total revenue per item. **Item on the left (Salwaar suits, Kurta, Jeans) generate the highest sales. Item on the right (Pant, Frock, Undergarments) contribute very little to overall revenue.**

Line (Orange – Cumulative Percentage) Shows the cumulative contribution of each item to total revenue. It shows a steep early rise that means first few items contributes heavily. Later item shows flat line that means remaining items contributes little additional revenue.

Items	Customer Looking for
Coat-pant	2
Frock	1
Jeans	2
Kurta	13
Pant	7
Salwaar suit	21
shirt	7
T-Shirt	5
Total	58

Table-3.1

The above table shows item – wise distribution of clothes that customers were looking for. As we can see that **total number of 58 customer didn't find what they are looking for**. 21 customers were looking for Salwaar suit which is highest, 13 customers were looking for kurta. A very little number of customers were able to find what they were looking for. This also shows us that there are stocking issues.

From figure 3.5 we can see that Kurta and Salwaar suit were highest revenue generator but from this table we also see that the greatest number of missed sales were of Kurta and Salwaar suit which shows revenue loss for product which is in demand.

Reason why customer didn't buy

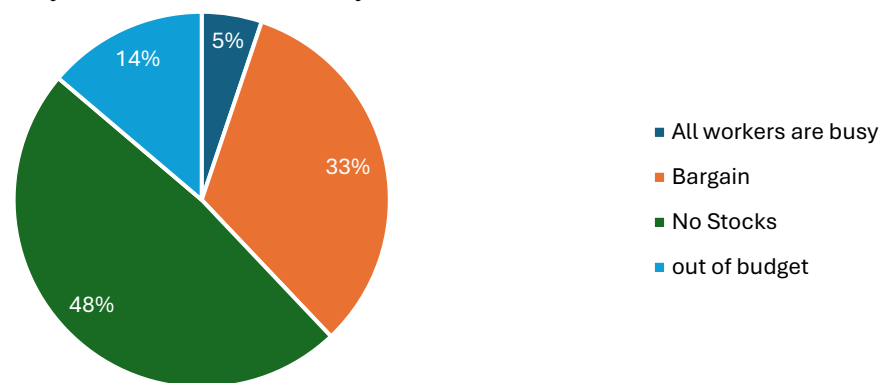


Fig- 3.4 Pie chart representing contribution of reason for which customers didn't buy

As seen in **Figure 3.1**, it is evident that revenue loss occurred, with missed sales being a major contributing factor. The accompanying **pie chart** highlights the reasons why customers chose not to make a purchase. Notably, **48%** of customer responses cited "**No Stock**" as the reason, followed by **bargaining failure (33%)**. This strongly indicates stock-related issues as a key driver of revenue loss.

Overall, after analysing both charts, we can conclude that the business experienced revenue loss primarily due to an increasing number of missed sales, which were largely caused by **stock unavailability**. Customer remarks mentioning the absence of the desired **size or colour** point toward **understocking**, indicating a critical gap in inventory management.

3.3. BAR CHARTS

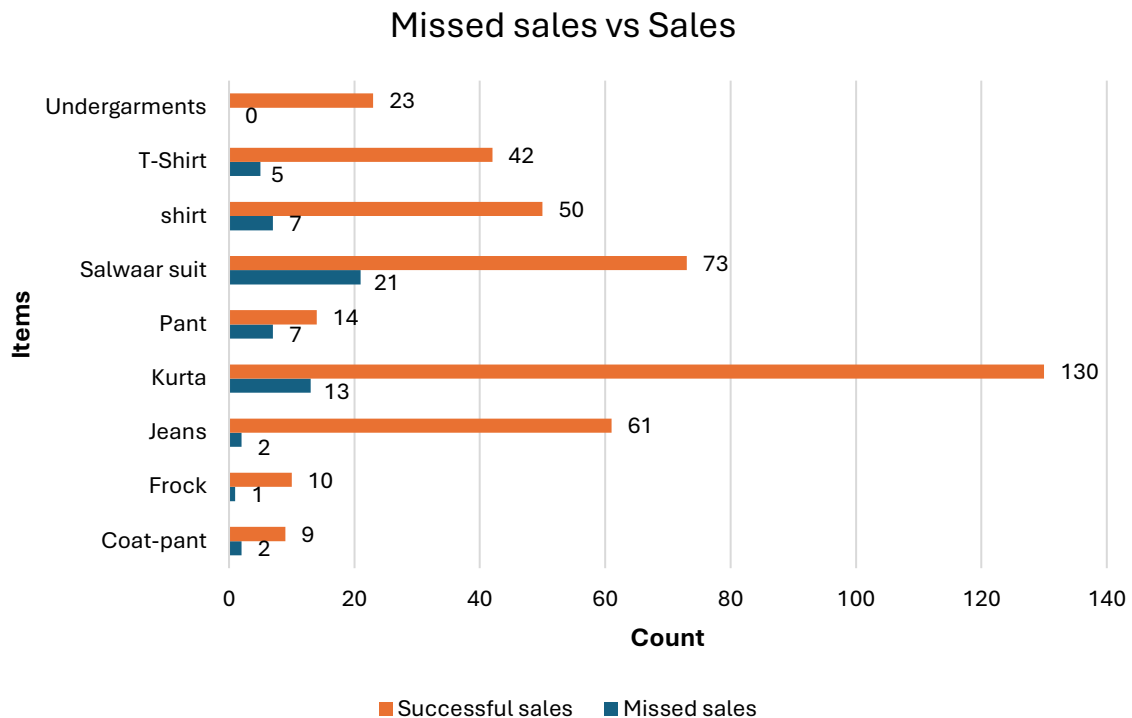


Fig-3.5 grouped bar chart showing comparison between missed sales and successful sales

The above graph is comparison between missed sales and successful sales made. Salwaar suit, Kurta and Jeans were sold in highest numbers. Moderate sales for T-shirt, shirt and undergarments. Frock and Coat-pant were in least number. Similarly, we can see that Salwaar suit and kurta had highest number of missed sales. Rest of them have a very number of missed sales.

If we Calculate ratio (sales / missed sales) for kurta and Salwaar suit.

Kurta: - 130 sales vs 13 sales = 10:1

Salwaar suit: - 73 sales vs 21 sales = 3.47:1

This shows kurta has higher demand reliability, while Salwaar suit more severe stock issues relative to its demand. The shop should prioritize **restocking Salwaar suit and Kurta** more frequently and ensure buffer stock exists to reduce lost opportunities.

Revenue generated vs Revenue Loss

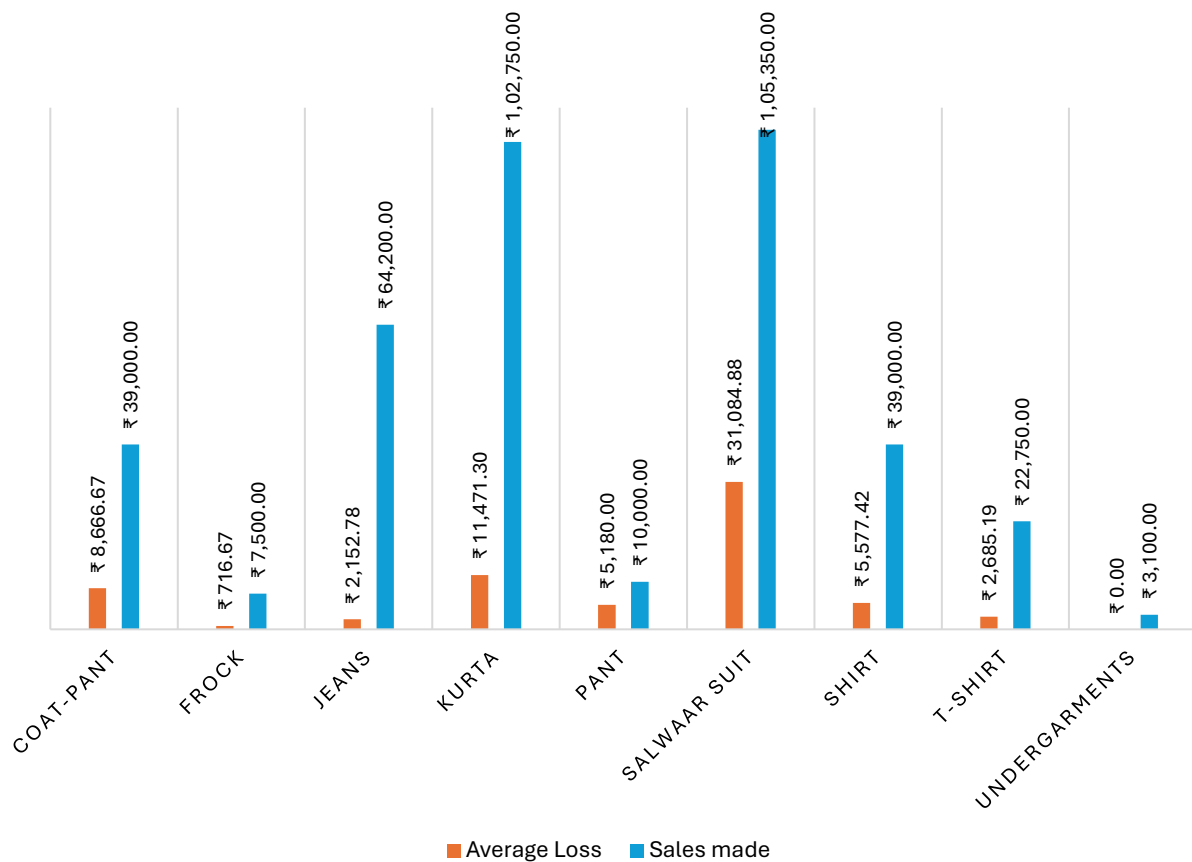


Fig-3.6 Grouped bar chart showing comparison revenue total sales and missed sales (average)

The above-mentioned bar chart illustrates a comparison between the total revenue generated and the estimated average revenue loss for each item. As shown, **Salwaar suit** and **Kurta** were the top revenue-generating items, contributing ₹1,05,350.00 and ₹1,02,750.00 respectively. These two items played a crucial role in the overall business performance. In contrast, items like **Jeans**, **Shirt**, **T-shirt**, and **Coat-pant** contributed moderately to revenue, while **Frock** and **Undergarments** remained at the lower end.

If we look at the **average loss**, Salwaar suit also recorded the **highest average revenue loss of ₹31,084.88**, indicating that despite high demand, a significant portion of potential sales could not be fulfilled—likely due to stock shortages. This suggests a **critical gap between demand and inventory availability** for high-performing items. The loss graph highlights that addressing stock issues, especially for top sellers like Salwaar suit and Kurta, could lead to a substantial

increase in overall revenue. Optimizing stock levels for these items should be a top priority in future planning to reduce missed opportunities and improve profitability

3.4. SCATTERPLOT

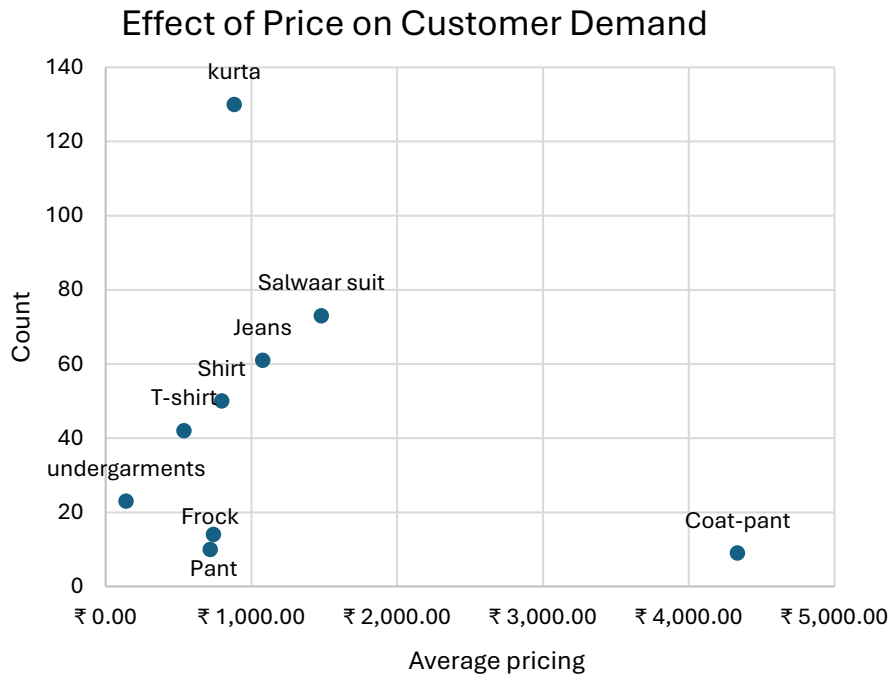


Fig-3.7 Scatterplot showing relation between sales count and Average pricing

From this scatterplot we can demonstrate that there is **negative correlation between Average price and sales made (demand)**. Here lower-priced items such as kurta, Jeans, T-shirt and shirt shows higher number of sales. Salwaar suit maintains moderate number of sales for moderate price. Coat-pant shows a low number of sales for high prices. Frock, pant and undergarments were the lowest number of sales for the lowest price which indicates some more must be done to increase sales.

Furthermore, this negative correlation suggests that the business must be cautious while **introducing or pricing high-ticket items**. Any product with a higher price must offer clear value in terms of **design, quality, or occasion-specific need**; otherwise, it risks underperforming. The chart also emphasizes that **items within a ₹700–₹1,500 range perform best**, and this price segment may represent the **ideal pricing sweet spot** for this customer base.

This plot also serves as a useful diagnostic tool for future pricing strategies. If the shop wishes to introduce premium items, it must pair them with **added-value offerings** like festive discounts, loyalty points, or bundling with popular items. Conversely, for low-performing, low-priced items, **aggressive promotions** may be necessary to drive awareness and test their true market potential

3.5. ITEM WISE COMPARISON

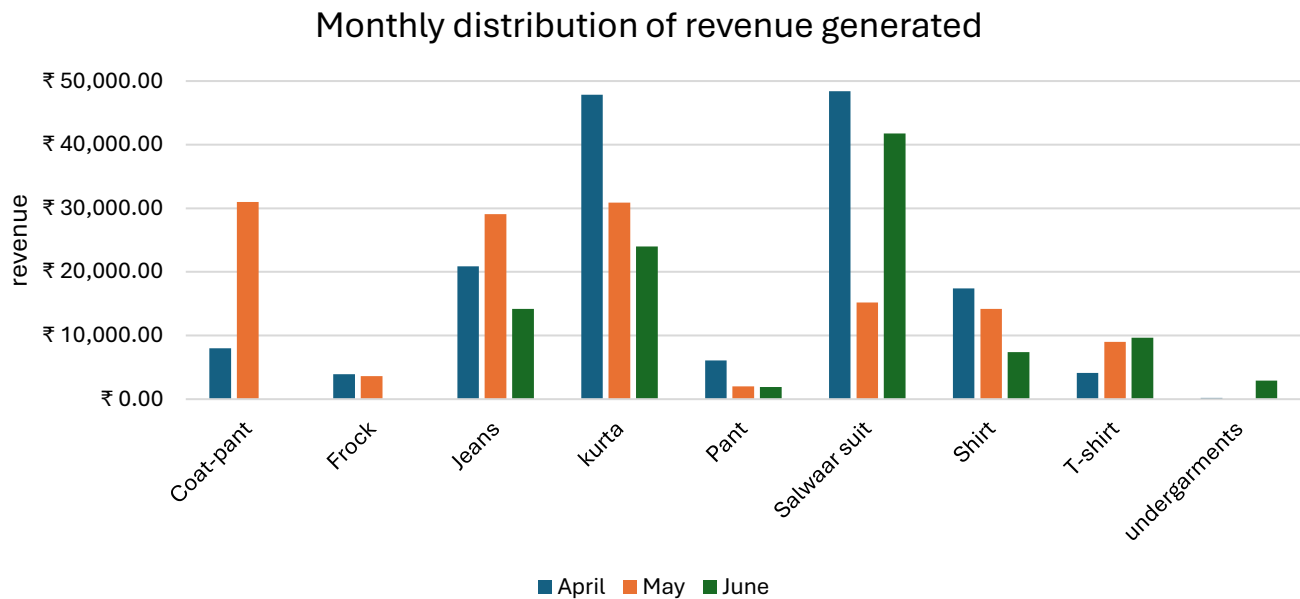


Fig-3.8 Monthly distribution of revenue generated

From fig- 3.8, we can see item-wise monthly distribution of Revenue generated.

In April

Kurta and Salwaar suit were the highest revenue generator in April. Jeans and shirts generated moderate revenue, and the rest were in very low demand. The undergarments and frock revenue generation were nil.

In May

Salwaar suit sees a significant drop in demand. Kurta sees a moderate drop in demand. Coat and jeans see a surge in demand which create more revenue. The undergarments revenue generation was still nil.

In June

Salwaar suit sees a huge surge in demand and has the highest revenue in month. Kurta and jeans see a small drop again. A huge drop is seen in demand for coat-pants. Shirts and T-shirts are very low demand. Frock demand was nil.

High demand of Kurta and Salwaar suit in April, May and June is mostly likely a **seasonal surge** which had happened because of start of wedding season according to Hindu customs (*called as lagan in local language*). Same can be said for the surge in sales of Coat-pant in May.

Overall, Coat-pant and jeans generated moderate revenue. The shirts and T-shirts were more than little in all three months. Frock, pants and undergarments were the lowest in demand in all 3 months, which could possibly create an overstocking issue and for Kurta and Salwaar suit there could be understocking issue.

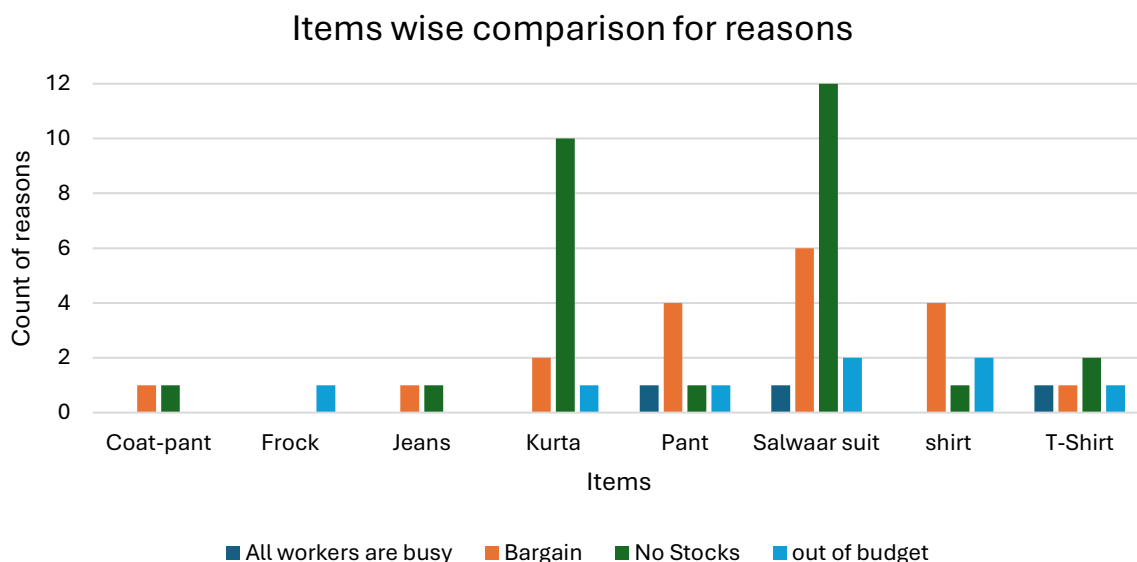


Fig-3.9 Item -wise comparison for reasons for not

The above stacked bar chart provides a detailed **item-wise breakdown of customer feedback** regarding reasons for not purchasing. It clearly shows that **"No Stock"** was the most frequently cited reason, especially for **Kurta** and **Salwaar suit**, with both items having **the highest count of stock-related complaints** among all categories. This reinforces the earlier findings (Fig. 3.5), where both items also showed **the highest revenue and highest average revenue loss**, confirming that stock unavailability of popular items directly impacts overall revenue performance.

Pant and Shirt stand out with relatively high number of complaints related to bargaining failure, indicating Price may high for these items. This also proves why these items were in low – demand. That means that price flexibility must revisit.

Interestingly, items such as Frock, Coat-pant, and Undergarments received very few complaints overall, which could indicate lower customer interest or limited visibility in-store.

Additionally, a few items like T-shirt and Shirt received notable mentions under "Out of Budget" and "All Workers Are Busy", suggesting that operational issues like staffing delays and pricing mismatches also play a role in lost sales

4. Interpretation of Results and Recommendations

4.1. RESULT INTERPRETATION

The analytical results of this capstone project present a clear picture of the challenges and opportunities that the retail clothing shop faces. Several key patterns emerged from the sales data, customer feedback, and stock records that can directly inform inventory, staffing, and pricing decisions.

Firstly, it was observed that that despite being highest revenue generators, Salwaar suit and kurtas were also accounted for most missed sales. This indicates that these items were in consistently high demand, they frequently go out of stock, leading to missed revenue opportunities. This pattern shows an underlying issue in stock management. Supply and demand are not being tracked on real-time basis which causes repeated shortage in most profitable categories which had a potential for more profitable.

In addition to stock issues, pricing-related complaints were frequently observed in feedback data. Items such as Shirts, Pants, and T-shirts often received complained regarding bargain failures, indicating a level price sensitivity among customers. Although these items fall into the mid-range price bracket, customers were still reluctance when they could negotiate prices. This means that a small variation in prices could lead missed sales. **perceived value versus cost** is a significant factor influencing purchase decisions.

Moreover, some complaints related to “Staff unavailability” during peak hours is also a operational bottleneck. Complaints like “All workers are busy” was mentioned, especially during week 2 to 4 which also coincided with increase customer footfall and missed sales. This highlights the importance of service experience as customer will leave without purchasing if they are not attended properly, especially in small retail settings where customer engagement often determines conversion.

The price – demand scatterplot (fig-3.6) gives us additional insights about purchasing behaviour. It shows that higher-priced items, such as Coat-pants and frocks, consistently had lower sales volumes, despite being available. This means that items do not align with customer preferences or are seen as higher priced, leading to poor sales performance. In contrast, items within a

moderate price range showed higher transaction volumes, conforming that customers in the market are more inclined to value-for-money purchases.

When combining the revenue trends, missed sales data, and feedback analysis, a pattern emerges: “stock unavailability” created dynamo effect. Stock unavailability (understocking) causes customers to not buy which created revenue loss for some items like Kurta and salwaar suit. Poor sales of some items like shirts, frock causes revenue loss which in result in overstocking. These findings suggest that addressing these issues will have a tangible, positive effect on overall performance of the shop.

4.2. RECOMMENDATIONS: -

Based on the result and interpretations above, several data-driven recommendations are given to help the shop and owner to resolve its current issues and improve operational efficiency.

1. Collect data and study.

First, owner should maintain a register of all sales made in day. This will help him to identify trends, sales patterns and demand spike due to festive season. This will create data that can be used to forecast analysis to predict demand for each product which will help him in countering overstock and understocking of items. He can also use basic excel sheet to collect data.

Customer feedback should be considered regularly so it can help him identify reason by customers are not buying. It will also improve customer experience which will help shop in future.

2. Enhance stock management for highly demanded items.

Items like salwaar suit and kurta suit should be prioritized in stocking schedules. The shop should maintain a weekly stock review process and order all the product before they reach to their minimum stock level. By forecasting demand based on past sales data, shop can prevent frequent stock-outs and retain more customers.

3. Implement demand-based ordering

Monthly demand-based ordering should be used to create a dynamic ordering system. This will help them in find trend in sales during festival season or local event when demand is high. Sales and feedback data should be analysed to determine which items should be ordered frequently, reducing missed sales chances.

4. Make price more flexible for sensitive items

For items commonly linked with bargaining failures (such as Shirts and Pants), the shop should explore offering **limited-time discounts**. These strategies can help close the gap between expected and offered prices, improving conversion without significantly reducing profit margins.

5. Reevaluate and reduce low-demand inventory.

Items like coat-pant and frocks, which consistently performed low in both sales and customer feedback, should be reevaluated. If these items are not demand, then these items should not be purchased or purchased in very low amount or forecasted by past data.

6. Improve Staff allocation and service efficiency

The shop should allocate staff in way that no customer leaves without buying anything during peak hours. As we saw in customer feedback data that some customers are leaving because they were not attended by any staff because all the workers were busy. This also create a bad perception of shop. **If it is okay than shop can hire part-time workers to help during peak hours of festive season.**

7. Establish a regular feedback Monitoring system.

The customer feedback analysis should be done every month. Feedback should be reviewed to detect trends early, such as a rise in complaints about stock or price. This will improve the service. Data collected from this can reveal more information about customers purchasing patterns and operational challenges.

8. Moving Toward online business.

I have suggested him to open account on amazon business and other e-commerce platform. These platforms are beginners friendly, and he should try selling products online. Especially, product which have low sales in local markets and are getting overstocked.