

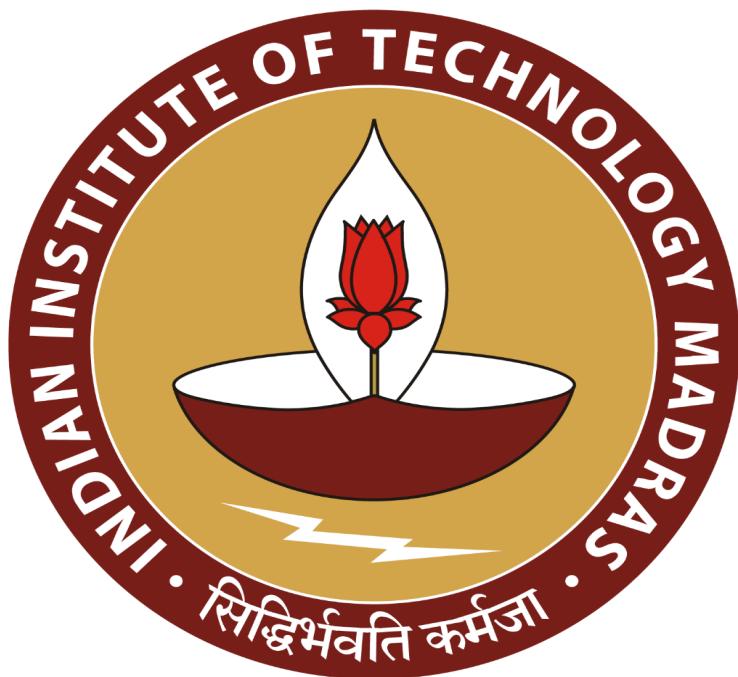
# **Data-Driven Optimization of Returns, Customer Trends and Sales for a Fashion Boutique**

A Mid Term Submission report for the BDM capstone Project

Submitted by

Name : BHABANI SHANKAR SAMAL

Roll number : 23F2005717



IITM Online BS Degree Program,  
Indian Institute of Technology, Madras, Chennai  
Tamil Nadu, India, 600036

## **Contents**

<b>Declaration Statement.....</b>	<b>3</b>
<b>1.Executive Summary.....</b>	<b>4</b>
<b>2.Proof of Originality.....</b>	<b>4-5</b>
<b>3.Metadata and Descriptive Statistics.....</b>	<b>5-7</b>
<b>3.1.Metadata.....</b>	<b>5-6</b>
<b>3.2.Descriptive Statistics.....</b>	<b>6-7</b>
<b>4.Analysis processes and methods.....</b>	<b>7-10</b>
<b>5.Results and Findings.....</b>	<b>10-13</b>

## **DECLARATION STATEMENT**

I am working on a Project titled “**Data-Driven Optimization of Returns, Customer Trends, and Sales for a Fashion Boutique**” . I extend my appreciation to **LIPU & LINKON DRESSES**, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through primary sources and carefully analyzed to assure its reliability.

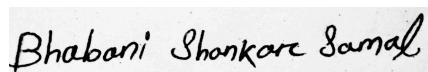
Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the information of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I agree that all the recommendations are business-specific and limited to this project exclusively, and cannot be utilized for any other purpose with an IIT Madras tag. I understand that IIT Madras does not endorse this.

Signature of Candidate:

A handwritten signature in black ink, appearing to read "Bhabani Shankar Samal".

Name: Bhabani Shankar Samal

## **1. EXECUTIVE SUMMARY**

The project titled “Data-Driven Optimization of Returns, Customer Trends, and Sales for a Fashion Boutique” is based on a men’s fashion boutique(**B2C model**) located near the Bus Terminal in Udaipur, Odisha, 757041 which is a semi-urban NAC area. The shop has been serving its local customers with a good variety of men’s clothing including formal and casual shirts, jeans, blazers, chinos, belts, t-shirts, kurtas, socks, and seasonal wear like sweaters and jackets. They are known in the area for offering trendy styles at reasonable prices, and many customers visit the shop regularly during weekends and festival seasons.

Despite having a strong local presence, the shop has been facing some key problems. In my early conversations with the owner Mr. Dibakar, it became clear that frequent returns, lack of proper tracking on what gets returned and why, unclear customer preferences, and confusion around discounting and pricing strategies are affecting their day-to-day business. Also, the shop staff mentioned challenges in handling seasonal demand shifts and repeated issues with product sizing and quality complaints, especially during busy weeks.

To explore these issues properly, I met with the owner and visited the shop several times, building trust and understanding their working process. I explained how analyzing their own data even just using simple tools could reveal useful patterns. They eventually realize my point of view toward their pain points. With their support, I collected 4 months of sales and return data, from *November 2024 to February 2025* in spreadsheet format.

Using Google sheet-based analysis, I explored trends in customer preferences, reasons behind product returns, and patterns in sales performance. The initial results are helping the store identify opportunities for better inventory planning and improved customer service. As the project progresses, my aim is to continue applying data-driven insights through pivot tables and charts to help Mr. Dibakar’s business build a more sustainable and customer-friendly approach.

## **2. PROOF OF ORIGINALITY**

**Letter of Authentication:** [AUTHORISATION LETTER](#)

**Image of the Shop:** [GOOGLE DRIVE LINK OF LIPU & LINKAN DRESSES IMAGE](#)

**Video Chat with Owner:**  [VIDEO CHAT WITH OWNER.mp4](#)

Image with the owner: [GOOGLE DRIVE LINK OF IMAGE WITH OWNER](#)

Dataset Link: [PRIMARY DATASET](#)

### **3. METADATA AND DESCRIPTIVE STATISTICS**

#### **3.1 Metadata**

I collected data from November 1, 2024, to February 28, 2025, tracking sales and returns. The store provides data in spreadsheet format, making it easier for analysis. I obtained two distinct datasets, one for sales and another for returns. Sales dataset contains 990 rows and 9 columns, capturing various transactional details and the return dataset consists of 164 rows and 5 columns, focusing on returned items and their reasons.

#### **Metadata of Sales Data**

Variable	Description	Data Type	Relevance
SI_No	Unique identifier for each transaction.	Integer	Ensures proper tracking and referencing of sales data.
Age_Group	Age range of customer		Helps analyze purchasing behavior across different age demographics.
Purchase_Category	Category of the product purchased (e.g., Tie, Belt, Chinos).	String	Helps identify product demand and category-level performance.
Date_of_Purchase	The purchase date of the product.	Date	Useful for analyzing daily and seasonal sales trends
week	Week label of the purchase.(Derived column)	String	Supports weekly trend analysis.
Month	Month name of the transaction.(Derived column)	Date	Helps identify monthly patterns and seasonal demand.
No_of_Items_Purchased	Number of units purchased per item.	Integer	Tracks volume of individual product sales.
Price_per_Item	Selling price including GST per item.	Float	Determines per-item revenue.
Discount_Per_Unit	Discount offered per item.	Float	Helps evaluate promotion effectiveness.
Total_Sale	Total amount after discounts.	Float	Useful for calculating total sales.

### Metadata of Return Data

Variable	Description	Data Type	Relevance
SI_No	Unique identifier for each return transaction.	Integer	Helps track and reference returned products systematically.
Product_Name	Name or type of returned product.	String	Allows identification of frequently returned items.
Return_Reason	Reason stated for return (e.g., Fitting Problem, Stitching Issue).	String	Helps identify quality or sizing issues in products.
Days_Since_Purchase	Number of days between purchase and return.	Integer	Useful for tracking return cycles and eligibility windows.
Status	Whether the return led to an Exchange or Refund.	String	Helps in financial planning and inventory adjustment strategies.

### Justification

The dataset provides transactional sales data and return data, which can be used to analyze:

- **Sales trends** over different time periods using **Week**, **Month** and **Total\_Sale** variables.
- **Customer Trends** and **Preferences** tracking according to different age groups using **Age\_Group** and **Purchase\_Category** variables.
- **Customer dissatisfaction** analysis using **Product\_Name** and **Return\_Reason** variables.

### **3.2 Descriptive Statistics**

#### Sales Data

- Total Number of Sales Record: 990 purchases.
- Average Quantity Sold per Day: 8 items.
- Mean Revenue per day: ₹11,925.
- Median Revenue per day: ₹12,758.
- Max Purchasing Age Group: 25-36(₹485,068).
- Min Revenue Generating Month: February.

- Average Order Value is ₹1,446.
- Total sales over four months from Nov-24 to Feb-25 were ₹14,31,045.

## **Return Data**

- Total Number of Return Record: 164 purchases.
- Max Return Status: Exchange(Total=129).
- Highest Returned Item: Formal Shirt(Total=13).
- Mode of Return Reason: Fitting Problem.
- Median Days to return Since Purchase: 2.

## **4. DETAILED EXPLANATION OF ANALYSIS PROCESS**

### **4.1 Data Cleaning and Preprocessing**

Data cleaning and preprocessing ensures that the analysis is based on accurate, consistent, and reliable information, preventing misleading results and poor decisions.

- I utilized Excel for Data Analysis and Visualization
- Null or empty cells were identified using filters and conditional formatting.
- Missing entries were handled by either filling with appropriate placeholder values or removing incomplete rows.
- Correct Typographical Errors: Fixed spelling and case inconsistencies
- Standardizing Formats: Ensured consistency in categorical fields such as Product\_Category, Return\_Reason, and Status (e.g., “Exchange” vs “exchange”) and standardized date formats (DD-MM-YYYY).
- Ensured consistent data types across columns, with Product Name and Return Reason as text, Date\_of\_Purchase as date, and No\_of\_Items\_Purchased, Price\_per\_Item, Discount, Total, and Days\_Since\_Purchase as integer or float.
- Derived the Week and Month column using the Date\_of\_Purchase column through the “=TEXT” formula.

## **4.2 Analysis Process**

### **ABC Analysis**

In this project, ABC Analysis was employed to categorize the product portfolio of LIPU & LINKAN DRESSES based on value and impact of different items on overall operational costs. I began by extracting data from sales data and classify items in three categories:

- *Category A* included high-impact identified through Pareto Analysis principles where Kurta , Winter Jacket and Blazer emerged as top performers contributing a huge share in total sales revenue. The analysis followed the traditional 80/20 rule where approximately 20% of products generated 80% of revenue.
- *Category B* consisted of moderately performing products like Jeans, Cargo Pants, Suit, Belt, T-shirt, Chinos showing stable performance across multiple age groups. These products still demanded attention and didn't have a significant impact like Category A.
- *Category C* included lower-impact products, such as Dhoti, Sports Wear, Socks, Baniyan, Underwear, Khadi Kurta. These products needed minimal oversight while maintaining availability for niche markets.

Through this three-tier classification, I was able to streamline procurement workflow of LIPU & LINKAN DRESSES by identifying key revenue driving Category A products should be consistently available during peak season and avoiding stock outs in critical periods.

### **Root-Cause Analysis**

For Root Cause Analysis, I used Excel and Python to study return data and find out why products are being sent back. My goal was to quickly identify the main problems and what's causing them, so the business owner can fix these issues and reduce future returns.

- In Excel, I made pivot tables and charts to see which products and reasons were most common for returns. “Fitting Problem” was the top reason, especially for Jeans, T-Shirts, and Formal Shirts. “Print Peeling” was a big issue for Designer T-Shirts, while “Zipper Broken (Low Quality)” showed up often for Hoodies, Jeans, and Winter Jackets.

- I used Python (Pandas) to count how many returns each product-reason pair had. The worst combinations were “Designer T-Shirt – Print Peeling” and “Jeans – Discoloration or Stitching Issue.” These products had the highest refund requests, not just exchanges.
- The Exchange Refund chart showed that most returns (79%) are exchanged, but 21% lead to refunds. Refunds happened more when the problem was about quality, like “Print Peeling” or “Broken Zipper,” which means customers lost trust in those products.
- Using the 5 Whys method:
  - For “*Fitting Problem*”: The real issue is likely inconsistent sizing or unclear size guides for customers.
  - For “*Print Peeling*”: The root cause is probably poor print quality or bad colouring materials.
  - For “*Zipper Broken (Low Quality)*”: Likely due to cheap zippers or lack of quality checks.
  - For “*Stitching Issue*”: Most likely rushed or careless stitching during manufacturing.
- These above findings show that the main reasons for returns are sizing problems and quality issues with materials or stitching. Fixing size charts, improving quality checks should help reduce returns.

The insights from this root cause analysis allowed me to pinpoint not only which return reasons and product combinations were most problematic but also the likely operational or quality control gaps responsible.

## **Trend Analysis**

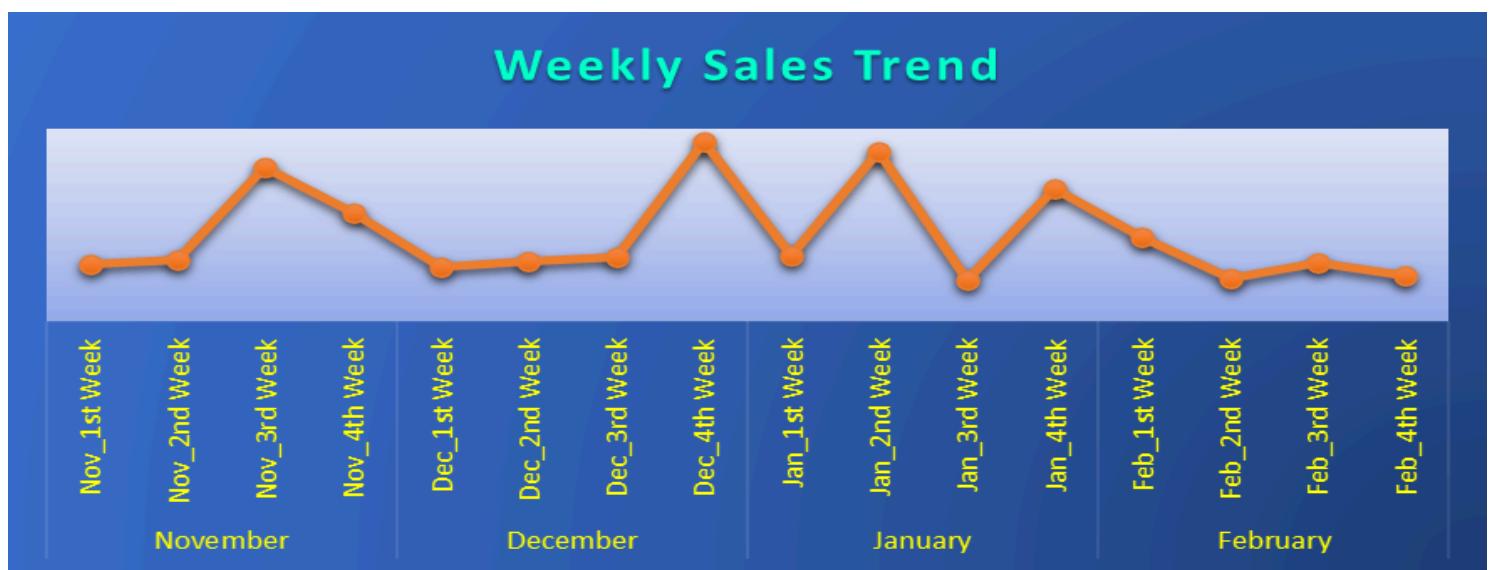
For Trend Analysis, I used sales data from November to February to find patterns in product demand. I plotted weekly and monthly sales using Excel.

- I saw a big sales jump in the third week of November during **Prathamastami**, a festival in Odisha where families buy new clothes for the eldest child. This led to high demand for traditional and children’s wear. Category A products like Kurta and Winter Jacket sold very well during this time.

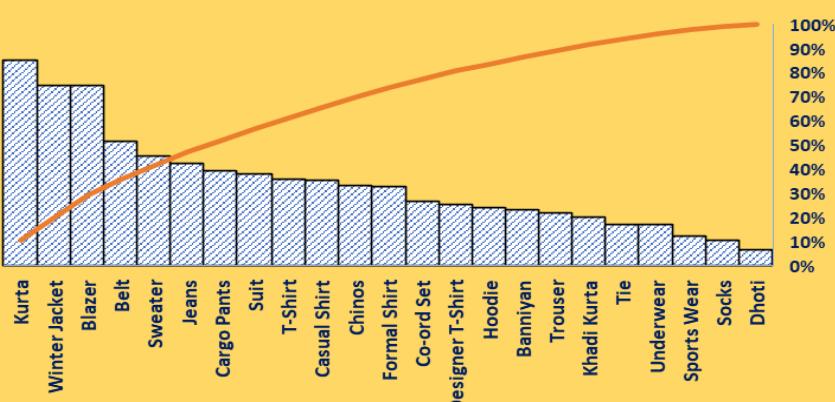
- In the fourth week of December, sales peaked again because of **New Year celebrations**. People bought new outfits to welcome the new year. Category A products, especially Winter Jacket and Cargo Pants, were top sellers.
- Sales went up again in the second week of January for **Makar Sankranti**, another festival where buying new clothes is common. Kurta and Cargo Pants continued to be popular Category A products.
- There was a smaller sales rise in the fourth week of January during **Saraswati Puja**, when families often buy new clothes for students.
- Monthly sales trends matched these festivals: higher sales in November and December, a drop in January after the festivals, and a recovery in February.
- Throughout these months, Category A products like Kurta, Winter Jacket, and Cargo Pants stayed as best sellers, even outside festival weeks.
- Most buyers were in the **18-24** and **25-36** age groups, showing strong demand from young customers during both festival and regular times.

This trend analysis helped me see that festival weeks always bring sales peaks, especially for *Category A* products. Knowing this, I can plan stock and marketing better for the next festival season and avoid overstocking in regular months.

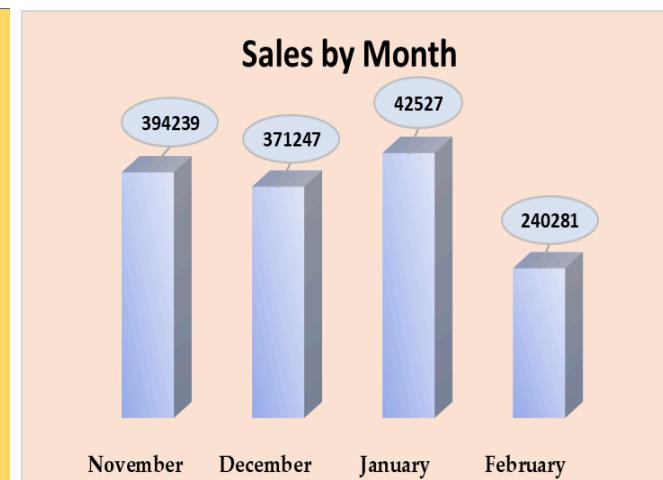
## **5. RESULTS AND FINDINGS**



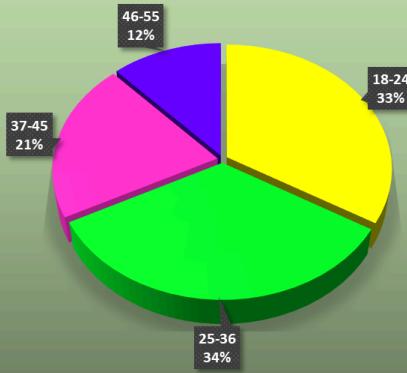
## PARETO ANALYSIS OF PRODUCT SALES



## Sales by Month



## Sales Overview by Age Group



## Monthly Order Volume and Average Order Value Trends



## Top Age Group Purchaser by Product



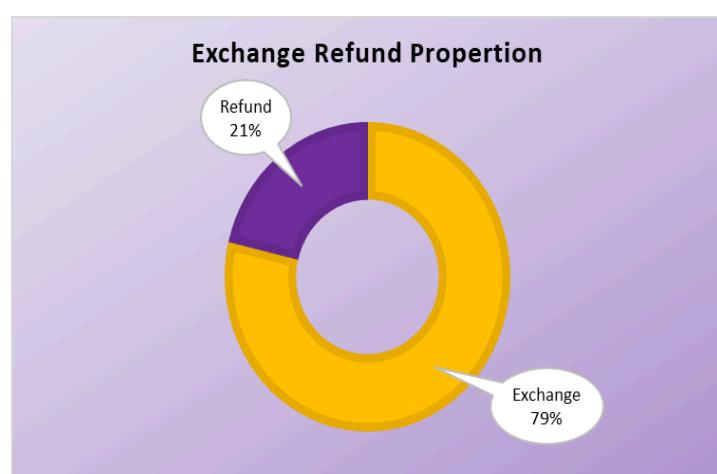
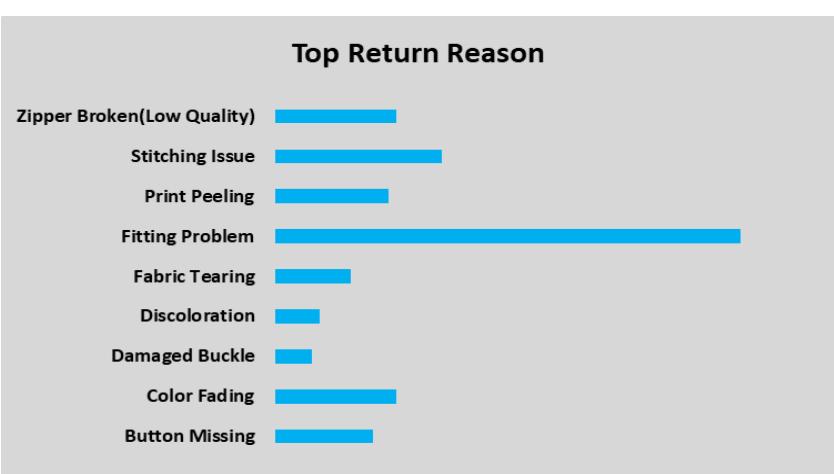
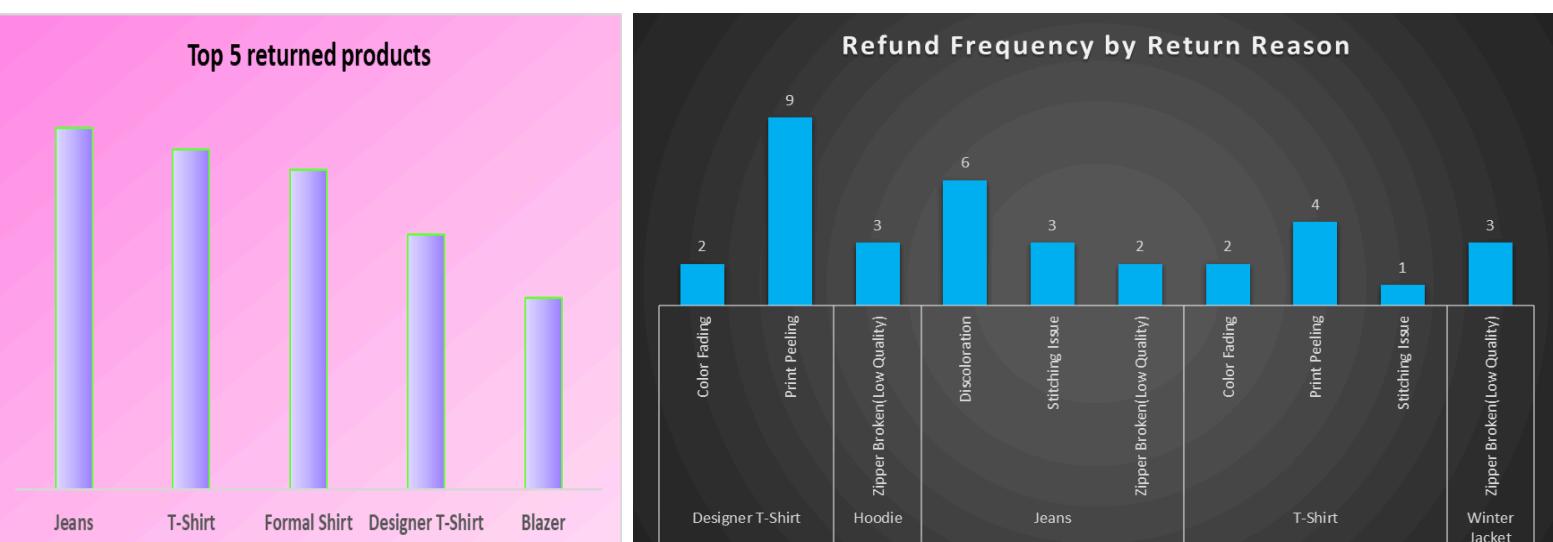
## Findings of Sales Data

- Top Products: Kurtas, Winter Jackets, and Blazer bring in most sales (Pareto rule).
- Monthly Sales: Highest in November and December—likely due to holiday shopping.

- Age Groups: 18–24 and 25–36 age groups buy the most—together they make up over 65% of total sales.
- Order Trends: January-25 has the most orders; December-24 has the highest average order value.
- Product by Age: Age 25–36 is the top buyer group for most products.

### Preliminary Insights on Sales Data

- Focus marketing and stock on top-selling products.
- Plan promotions around November–December to boost seasonal sales.
- Target ads and offers to age groups 18–36, they are the main buyers.
- Keep pricing steady but boost engagement in low-sales months with offers.
- Age 25–36 is the key group for product-specific promotions.



## **Findings of Return Data**

- Most Returned Products: Jeans, T-shirts, and Formal Shirts are returned the most.
- Top Return Reasons: Common issues include Zipper Broken (Low Quality), Print Peeling, and Stitching Problems.
- Product-Specific Issues:
  - *Designer T-shirts*: High returns due to Print Peeling.
  - *Jeans*: Discoloration and Stitching Issues are frequent.
  - *Winter Jackets & Hoodies*: Zipper problems are common.
- Refund vs Exchange: 79% of customers prefer exchanges over refunds—only 21% ask for refunds.

## **Preliminary Insights on Return Data**

- Improve quality checks for zippers, stitching, and prints—these are top return triggers.
- Focus on quality control for Jeans and T-shirts, as they lead in returns.
- Since most customers prefer exchanges, streamline the exchange process to retain buyers.

Leveraging these insights and data-driven forecasting strategies will help the store enhance performance, achieve sales targets, optimize pricing and taxation, and improve customer engagement for sustained growth.