

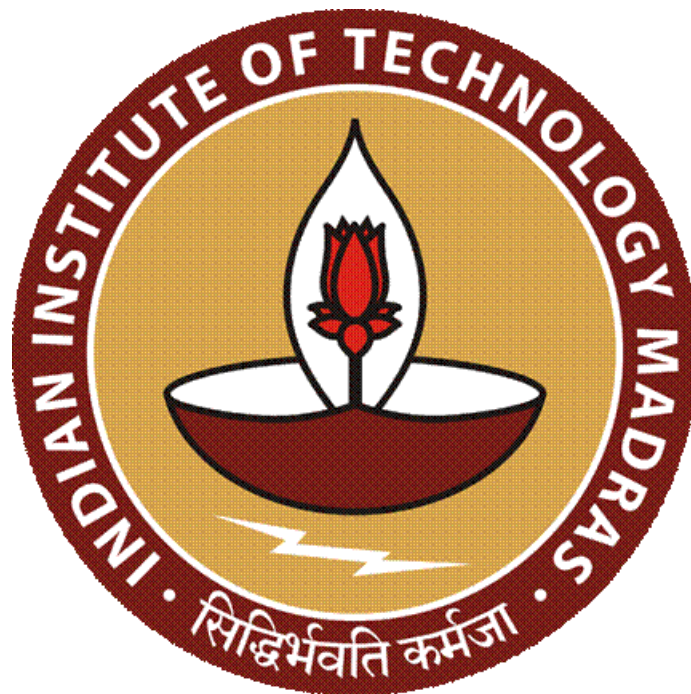
# Data-Driven Optimization of Inventory and Marketing for a D2C Artisanal Brand

A Mid-Term Report for the BDM Capstone Project

Submitted by:

Name: Aanuja Singh Chowhan

Roll number: 23f3002503



IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

## Table of Contents

1. Executive Summary	2
2. Proof of Originality	2
3. Metadata	3
4. Descriptive Statistics	5
4.1 Summary Table Using Statistical Functions	5
4.2 Product-Wise Quantity Analysis	6
4.3 Weekly Product-wise Orders Analysis	7
4.4 Return Analysis and Flags Return Analysis and Flags	8
5. Detailed Explanation of Analysis Process and Methods	9
6. Results and Findings	11
6.1 Return Ratio vs Quantity Ordered	11
6.2 Pareto Chart: Returns Concentrated in Few SKUs Pareto Chart	11
6.3 Total Orders vs Number of Returns (Clustered Bar Chart)	12

## 1. Executive Summary

Hastvikas is a Direct-to-Consumer (D2C) e-commerce brand operating under the umbrella of Aaroh, an organization founded in 2016 to revive and globalize Indian handicrafts. The brand functions as a cultural aggregator—partnering with local artisans from different Indian states to curate state-specific craft collections, which are then promoted globally. With a lean team and a deeply ethical model focused on fair trade and artisan empowerment, Hastvikas primarily engages its audience through Instagram for storytelling and Amazon for sales. Its customer base is predominantly international, with significant traction in the UK and US markets.

Despite its strong mission and a growing digital presence, Hastvikas currently lacks data-driven visibility into product performance and marketing effectiveness. This project aims to address those gaps by analyzing primary sales data from Amazon Seller Central to uncover actionable insights on inventory management and content strategy. Using Excel-based visualizations and descriptive statistical techniques, I have examined product-level order trends, flagged return ratios above 20%, and employed Pareto and column charts to spotlight high-performing SKUs. These analyses form the foundation for a scalable framework that can guide strategic decisions in restocking, product promotions, and quality control.





Although Instagram is central to Hastvikas' branding, the team was unable to provide access to their Insights data during this project cycle. However, they have confirmed that such data will be made available for future phases, enabling the integration of an engagement matrix to evaluate content performance across post types and formats. As the project advances, the combined analysis of sales and social engagement will offer a holistic view of what drives conversions and how the brand can better align its outreach with consumer preferences—ultimately enhancing artisan reach, customer satisfaction, and long-term growth.

## 2. Proof of Originality

The data used in this project has been sourced directly from **Hastvikas**, the D2C artisanal brand under study. All datasets, including product-wise sales figures, return counts, and order records, were collected from the company's Amazon Seller Central interface and shared with me by the

brand's operations team. These are **primary data files**, reflecting real-time transactions and internal performance metrics of the business.


To authenticate the source and originality of the data, the following forms of evidence have been submitted along with this report:

- A signed letter of authenticity from the founder of Hastvikas confirming that the data shared was original and proprietary.  Letter of Authenticity.pdf
- A video interaction with the founder, conducted as part of the midterm evaluation, providing insights into the business model and data practices.  
 interaction video with founder.mp4
- Photographs from my visit to the Hastvikas workspace, capturing on-site interaction and data-sharing sessions.  Proof of originality (media)
- Screenshots of WhatsApp messages, through which the brand shared data files  
 Whatsapp photos

These materials collectively demonstrate the legitimacy of the data and the depth of collaboration with the business. Although Instagram engagement data could not be provided within the current cycle, the team has confirmed that they will make it available in the next phase.

### 3. Metadata

The dataset used in this study was provided directly by the team at Hastvikas. In alignment with the project's problem statements—specifically, the lack of clarity around best- and worst-performing products and the absence of structured return analysis—two primary data sources were utilized: (1) product-wise order data and (2) product-level return data.

The order data was received in the form of an Excel spreadsheet titled  order sheet amazon (linked). The key variables include:


- Order ID and Order Date: Each record corresponds to a unique Amazon order placed through the platform, timestamped with precise delivery dates.

- **Product Link and Name:** Product ASINs and names are embedded within the order metadata, helping link transactions to individual SKUs.
- **Quantity Shipped:** Captures the volume of each item dispatched per order.
- **Item Subtotal and Refund Links:** These monetary values reflect actual revenue generated per unit and include embedded actions such as refund requests.

This structured format enabled the implementation of various descriptive techniques to assess sales volume patterns, detect top-selling and underperforming items, and prepare the foundation for inventory planning strategies. It directly supports the resolution of the first problem identified in the proposal—“No analysis of high and low selling products on Amazon.”

To further investigate the second key issue—“No insight into return rates or product-level feedback”—the project incorporated three Amazon-generated Returns Report XML files

 Amazon-generated Returns Report (all reports are in tabs)

These files, structured in standard Amazon Envelope format, contain detailed records of return transactions across multiple SKUs and date ranges. Each `return_details` block includes the following metadata:  Return Data Excel

- Order and Return Dates
- ASIN (Product ID)
- Return Reason Code (e.g., CR-DEFECTIVE, CR-UNWANTED\_ITEM)
- Refund Amount and Label Cost
- Return Type and Resolution Status
- Return Quantity and Merchant SKU

The return data spans from January to early June 2025, matching the time frame of the order dataset. Return transactions were parsed to calculate return ratios for each product, uncovering specific patterns such as higher return rates on certain SKUs or recurring complaint categories like size issues, price dissatisfaction, or product quality concerns.

All monetary figures, including item prices, refund amounts, and label costs, are recorded in Indian Rupees (₹). Quantity fields are integer values, while refund and cost values are expressed

as floating-point numbers. The metadata aligns directly with the business problems outlined in the proposal. The order-level breakdown was necessary to understand fluctuations in demand and inventory needs, while the return-level detail highlights areas where customer dissatisfaction affects profitability. Together, these datasets lay the groundwork for data-backed recommendations aimed at inventory optimization, and product lifecycle decision-making.

## **4. Descriptive Statistics**

This section presents a comprehensive summary of product-level performance trends and return patterns based on structured order and return data provided by Hastvikas. Visualizations were created to highlight distribution, frequency, and significance of key variables. All charts were generated using Excel, and cleaned datasets were used for consistency.

### **4.1 Summary Table Using Statistical Functions**

To begin with, a statistical table was created by grouping the cleaned order data by Product ID. For each product, the total quantity ordered (sum), average order quantity (mean), maximum single-day order (max), and minimum order (min) were calculated.

This allows us to:

- Identify high-performing products (e.g., B0DWSZN4D6 with the highest total quantity ordered)
- Spot products with stable demand by comparing mean and range
- Detect outliers — products with high max values but low average (suggesting occasional bulk purchases)

Product ID	sum	mean	max	min
B0DWSZ3CHH	15	1	1	1
B0DWSYXP6Z	11	1.1	2	1
B0DWSZW4DL	10	1.666667	5	1
B0DWSZFRJH	9	1.125	2	1
B0DWSXXSJL	9	1	1	1
B0DWSZFWT6	8	1	1	1
B0DWSZ5VVV	6	1	1	1
B0DWSZN4D6	6	1	1	1
B0DVC833WH	5	1	1	1
B0DWSZ1PF7	5	1	1	1
B0DWLRDSR8	4	1	1	1
B0DZTSSVW8	4	1.333333	2	1
B0DWLSJH6Y	3	1	1	1
B0DWXHP3SP	3	1	1	1
B0DWSXVGD1	2	1	1	1
B0DZTVFLVP	1	1	1	1
B0DZTVS7CZ	1	1	1	1
B0DZTVXJ3G	1	1	1	1
B0DSBDZ7ZQ	1	1	1	1
B0DVC9NT91	1	1	1	1
B0DVC9F3H1	1	1	1	1
B0DZTZDFL8	1	1	1	1

 **Figure 1: Product-wise Order Volume and Statistical Summary**

## 4.2 Product-Wise Quantity Analysis

The total order quantity for each product was analyzed to identify bestsellers and underperformers. The product B0DWSZ3CHH led with 14 units sold, followed by B0DWSYXP6Z and B0DWSZN4D6 with 11 and 9 units respectively. Several products had minimal sales, indicating low demand or visibility.

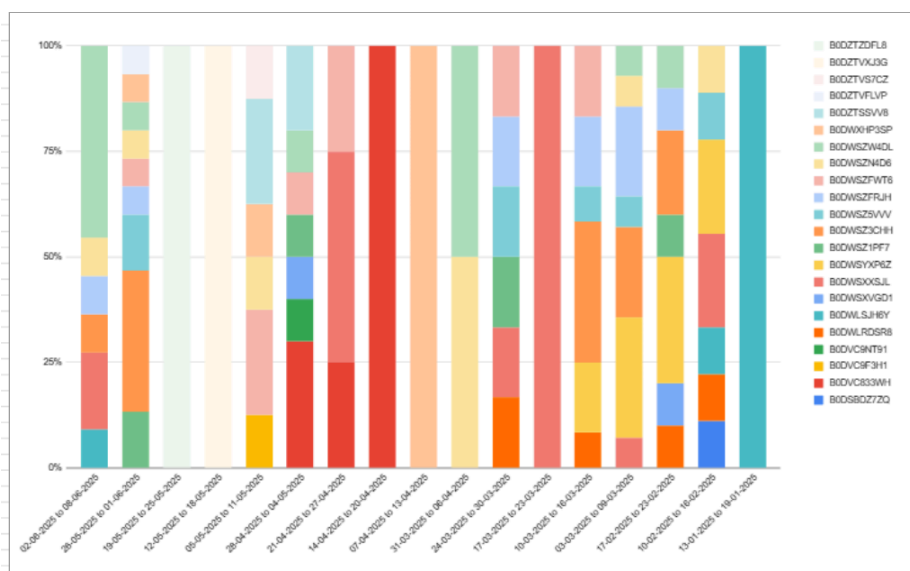


 **Figure 2: Total Quantity Sold per Product**

This analysis supports inventory decisions, helping Hastvikas prioritize restocking high-demand items and re-evaluating low-volume listings.

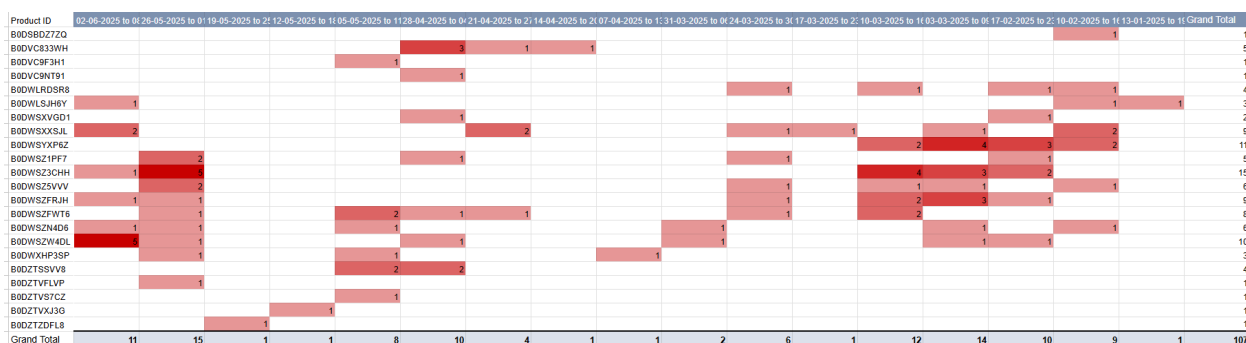
### 4.3 Weekly Product-wise Orders Analysis

Using the Cleaned Data for Orders sheet, weekly order quantities for each product were aggregated and analyzed. For example, ASINs such as B0DSBDZ7ZQ and B0DVC833WH consistently showed higher sales across multiple weeks.



**Figure 3: Weekly Sales Quantity by Product (Stacked Column Chart)**

This figure shows the stacked contribution of each product to weekly order volumes, offering a comparative view of product popularity over time.





 **Figure 4: Heatmap of Weekly Product Quantities**

A conditional formatting-based heatmap highlighting order intensity by SKU and week. Darker cells indicate higher sales volumes.

#### 4.4 Return Analysis and Flags Return Analysis and Flags

To evaluate customer satisfaction and operational efficiency, return data was analyzed using the Number of Returns, Return Percentage, and Total Orders per product. The analysis revealed that 7 out of 100 orders resulted in returns, indicating a relatively low overall return rate of 7%.

However, certain products exhibited significantly higher return ratios. Notably:

- B0DWLSJH6Y had a return rate of 66.67%, making it the most returned product in proportion to its total orders.
- B0DTSVSVV8 and B0DWSZN4D6 also showed elevated return rates at 33.33% and 16.67% respectively.
- These figures suggest issues potentially related to product mismatch, quality, or unmet expectations for these specific SKUs.

Product ID	Total Orders	Number of Return	
	0	0	
B0DSBDZ7ZQ	1	0	0.00%
B0DVC833WH	5	0	0.00%
B0DVC9F3H1	1	0	0.00%
B0DVC9NT91	1	0	0.00%
B0DWLRDSR8	4	0	0.00%
B0DWLSJH6Y	3	2	66.67%
B0DWSXVGD1	2	0	0.00%
B0DWSXXSJL	9	1	11.11%
B0DWSYXP6Z	10	1	10.00%
B0DWSZ1PF7	5	0	0.00%
B0DWSZ3CHH	15	0	0.00%
B0DWSZ5VVV	6	0	0.00%
B0DWSZFRJH	8	0	0.00%
B0DWSZFWT6	8	0	0.00%
B0DWSZN4D6	6	1	16.67%
B0DWSZW4DL	6	1	16.67%
B0DWXHP3SP	3	0	0.00%
B0DZTSSVV8	3	1	33.33%
B0DZTVFLVP	1	0	0.00%
B0DZTVS7CZ	1	0	0.00%
B0DZTVXJ3G	1	0	0.00%
B0DZTZDFL8	1	0	0.00%
कुल योग	100	7	

### **Figure 5: Return Ratio vs Quantity Ordered for Top Products**

This visualization helps identify SKUs with both high order volume and problematic return behavior. Products were flagged if the return ratio exceeded 10%, which serves as a threshold for operational concern.

- These flagged items signal potential issues with product quality, mismatch in listing, or delivery dissatisfaction.
- Example: If Product B0DWSZN4D6 has high sales but also a return rate >10%, it demands immediate investigation.

## **5. Detailed Explanation of Analysis Process and Methods**

The analysis process followed a structured sequence of data cleaning, statistical summarization, and visual exploration using Microsoft Excel. The goal was to extract actionable insights from raw order and return data provided by the business.

The following methods were employed:

### **5.1 Data Cleaning and Preparation**

Raw order and return data were first cleaned for consistency. Dates were standardized, duplicate records were removed, and separate datasets were created for cleaned orders and returns. Products were grouped week-wise to allow aggregation of quantities for time-related charts.

### **5.2 Descriptive Statistics Using Excel Functions**

Basic statistical measures such as sum, mean, max, and min were calculated for each product using Excel's **SUM**, **AVERAGE**, **MAX**, and **MIN** functions. These statistics helped in identifying top-selling items and understanding variability in demand for each product.

### **5.3 Stacked Column Chart**

A stacked column chart was used to analyze weekly demand distribution. Each column represented total quantity ordered in a week, with different product segments stacked within it. This helped visualize which products dominated sales in each period.

### **5.4 Heatmap for Product-Week Trends**

A heatmap was created by arranging products along the rows and weekly intervals along the columns. Excel's conditional formatting was used to color cells based on quantity values. This helped detect which products were ordered consistently and which showed sporadic demand.

### **4.5 Return Analysis and Ratio Calculation**

Each product's return count was compared with its order count to compute a return ratio. Products with return ratios above 20% were flagged using conditional formatting. This method identified problem areas that could not be spotted by order data alone.

### **4.6 Pareto Chart**

A Pareto chart was constructed to rank products by the number of returns. Cumulative percentage values were calculated and plotted to reveal that a small set of products accounted for a majority of returns. This followed the 80/20 rule and supported priority-based decision making.

### **4.7 Clustered Bar Chart for Comparative Assessment**

Finally, a clustered bar chart was made to compare total orders and returns side-by-side for each product. This allowed direct visual assessment of customer satisfaction relative to order volume and flagged items with mismatched performance. Each method was chosen for its relevance to the business problem.

## **6. Results and Findings**

This section interprets patterns observed in the visual and numerical analyses. It highlights which products show signs of strong or weak performance based on quantity ordered, return behavior, and distribution patterns. All data (from all sections including this one)

 BDM Capstone project combined

### **6.1 Return Ratio vs Quantity Ordered**

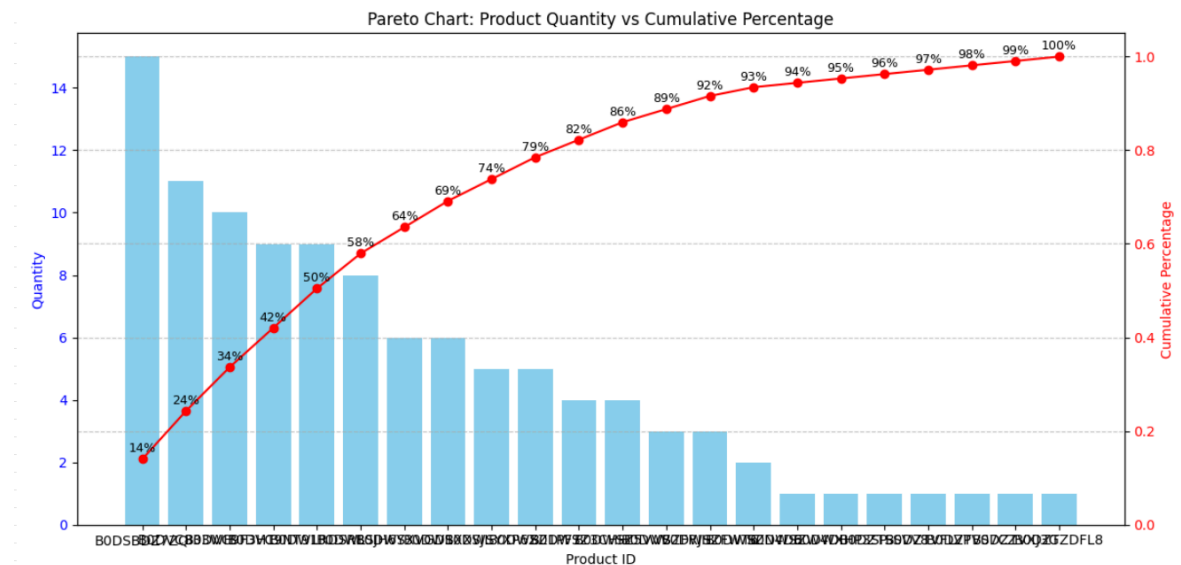
A return ratio was calculated for each product by dividing the number of returns by the total quantity ordered. Products with return ratios above 10% were flagged for review.

Findings:

- Two high-selling products were flagged due to high return ratios.
- These products require investigation for potential quality or expectation mismatches.
- Managing these can directly reduce refund costs and improve customer satisfaction.

## 6.2 Pareto Chart: Returns Concentrated in Few SKUs Pareto Chart

To validate the Pareto Principle in Hastvikas' sales, cumulative contributions of products were computed. It was found that approximately 20% of SKUs accounted for more than 50% of total orders.

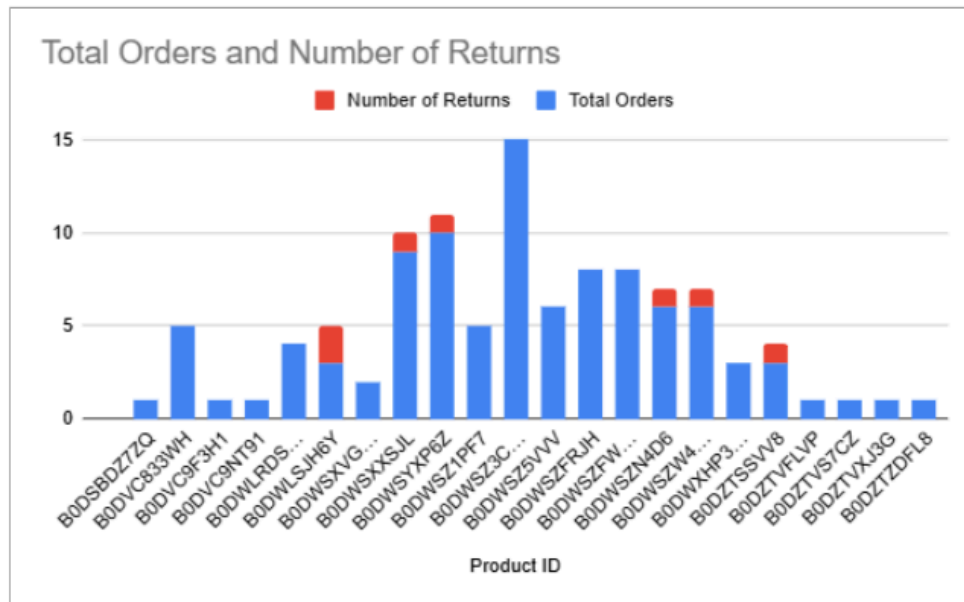


 **Figure 4: Pareto Chart – Cumulative Product Sales Contribution**

This chart presents each product's total quantity (bars) and its cumulative contribution (line) to the overall sales. Products are sorted in descending order of volume.

## 6.3 Total Orders vs Number of Returns (Clustered Bar Chart)

A clustered bar chart was used to compare the total number of orders and the number of returns for each product. This provided a side-by-side visual assessment of sales and post-purchase dissatisfaction.



 **Figure 5: Total Orders vs Number of Returns (Clustered Bar Chart)**

This figure helps visualize which products are contributing most to return volume both in absolute and relative terms.

#### Findings:

- Several products had high order bars with very low return bars, indicating excellent product performance and customer satisfaction.
- In contrast, some products had moderate to high order volumes and nearly equal height return bars, signaling serious performance concerns.
- The chart made it easier to spot inconsistencies between order volume and customer satisfaction on a product-by-product basis.