

# **Optimizing Revenue Streams and Operational Efficiency: A Data-Driven Approach for a Clothing Manufacturing and Export Company**

A Final Report for the BDM Capstone Project

Submitted by  
Name: Navya Tyagi  
Roll number:22f1001270



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

## Declaration Statement

I am working on a Project titled "Optimizing Revenue Streams and Operational Efficiency: A Data-Driven Approach for a Clothing Manufacturing and Export Company".

I extend my appreciation to JK Clothing, 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 ensure its reliability.

Additionally, I affirm that all procedures employed for 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. If 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 'N. Tyagi', is displayed on a light blue rectangular background.

Name: Navya Tyagi

Date: 05/12/2023

## **1. Executive Summary and Title**

JK Clothing is a distinguished garment manufacturer and exporter under JK Handicrafts, situated in the vibrant city of Jaipur, Rajasthan. Renowned for delivering high-quality products globally, JK Clothing boasts a legacy of 15 years in the industry.

However, JK Clothing encounters critical challenges impacting its operational efficiency. The issues primarily revolve around stagnant sales, inefficient inventory management, and a restricted understanding of evolving customer preferences. These challenges arise from rapidly changing consumer trends and fluctuations in tourist footfall in Jaipur throughout the year. Moreover, external factors such as global competition and economic fluctuations further alleviate these challenges.

This report presents an analysis of manufacturing and sales data obtained from JK Clothing spanning four months i.e. May-August 2023. While the data lacks specific details on categorized sales and quantities, it offers valuable information for addressing some identified problems. The proposed solution involves a data-driven approach. By harnessing purchase data, sales data, customer behavior insights, and inventory trends, this report offers conclusions and recommendations to refine marketing strategies, optimize inventory levels, and align product offerings with evolving customer needs. This data-centric strategy targets enhanced sales performance, reduced return rates, and improved operational efficiency by implementing the recommended solutions.

This strategic endeavor aligns with JK Clothing's commitment to maintaining its standing as a distinguished global supplier. The project endeavors to propel JK Clothing towards an adaptive and data-informed business model, fostering growth and sustainability in an increasingly competitive market landscape.

## **2. Detailed explanation of analysis process/method**

The analysis in this project aimed to derive meaningful insights from the data and address the challenges faced by JK Clothing. The primary objectives of the analysis were to gain insights into sales and purchase patterns and analyze sales trends over time. By conducting a comprehensive and holistic analysis of the available data, we aimed to provide actionable recommendations to overcome the challenges and drive business growth.

The analysis process plays a crucial role in understanding the underlying factors impacting the business's performance. By examining sales and purchase data, we can identify trends, patterns, and potential areas of improvement. The insights gained from

this process enable us to make informed data-driven decisions, optimize inventory management, and develop strategies to increase sales for the company.

**Data cleaning and preprocessing:** The initial phase involved data cleaning. The dataset initially available in spreadsheets was imported to Google colaboratory Notebook using the pandas library in Python for further analysis. We meticulously reviewed the data for any missing values, ensuring consistent data types and upholding data integrity throughout the dataset.

**Descriptive Statistics:** The subsequent phase involved conducting descriptive statistics on the data corpus. The statistics of the dataset were summarized using measures like mean, standard deviation, min, and max for a better understanding of the distribution of the variables in the dataset. This was facilitated by the *describe()* and *info()* methods of the Pandas dataframe.

The analysis involved in addressing each problem statement is as follows:

### **Problem Statement 1: Enhancing Sales Performance**

**Method:** To tackle this problem, we did a thorough purchase and sales analysis to gain a better understanding of the purchased and sold products and their relationship. The monthly purchase and sales data was used to determine the most purchased and sold products using bar plot visualization to gain a better understanding of cost share among the products. Pareto charts were plotted to represent and study the monthly purchase and sales distribution of the products. Pareto charts showcased which products hold the most influence on total purchases or sales, allowing quick identification of the most impactful items among a range of products. This visualization aids in prioritizing attention toward the most critical contributors, facilitating strategic decision-making for product management and sales strategies.

Additionally, we did Sales forecasting using the ARIMA model which is explained in detail further below. This ARIMA model can be used to predict sales for the next 12 months based on past data. This is an effective way to observe sales behavior and implement and enhance optimal sales performance.

### **Problem Statement 2: Inventory Optimization for Cost Efficiency**

**Method:** To address this particular problem, we did a price distribution analysis to explore the distribution of prices ('RATE') for products, analyze price ranges, and identify any outliers or common price points. This understanding aids in setting competitive yet profitable pricing strategies and price point optimization to maximize revenue. A histogram was created to provide a clear visual depiction of the

concentration of prices within different ranges, making it easier to identify common price points and outliers within the dataset.

Additionally, a Time Series Analysis was conducted to illustrate the trend of total purchase amounts over time, providing insights into overall purchase patterns and trends across months or time intervals.

These trends equip us with actionable insights to streamline inventory management practices to enhance cost efficiency.

### **Problem Statement 3: Customer Behavior Analysis for Product Enhancement**

**Method:** Customer behavior analysis was done by utilizing the gender and size variables to discern customer preferences based on these criteria. Bar plots were used to visualize this distribution of purchases and sales across different gender and size categories. Additionally, a pivot table was also created to analyze quantities bought/sold for different sizes based on gender and a heatmap was utilized to vividly display and interpret the distribution patterns across different gender and size categories within the dataset. Using this particular analysis, we can effectively understand consumer interest and make data-driven decisions to tailor our products effectively to meet consumer demands and enhance overall customer satisfaction.

These analyses provide an overview of the monthly trends, descriptive statistics offer quantitative measures of the data, and the rest analysis methods highlight any specific patterns or trends associated with the data.

## **3. Results and Findings**

The detailed analysis of the purchase and sales data of JK Clothing gave critical insights into the data and shed light on various patterns and trends.

**Data Preprocessing:** The first step involved cleaning and preprocessing the data. As the data was imported from google spreadsheets, there were a few minor issues with the data format which were resolved as a part of the cleaning of the data such as there was a whitespace error in one of the column names of the dataframe and the date format from the spreadsheet was not supported by the pandas dataframe. These issues were conveniently resolved using Python's pandas library.

```
[ ] df.columns

Index(['DATE ', 'PRODUCT', 'GENDER', 'SIZE', 'PIECES', 'RATE', 'TOTAL'], dtype='object')

[ ] df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 123 entries, 0 to 122
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0    DATE        123 non-null    object
1    PRODUCT     123 non-null    object
2    GENDER      123 non-null    object
3    SIZE        123 non-null    object
4    PIECES      123 non-null    int64
5    RATE        123 non-null    float64
6    TOTAL       123 non-null    float64
dtypes: float64(2), int64(1), object(4)
memory usage: 6.9+ KB
```

Fig 1: Data preprocessing

#### ✓ Cleaning and pre-processing the data

```
# Noticed that in date column name an extra whitespace character was present and it was not in correct datetime format
# Remove leading/trailing spaces from column names
df.columns = df.columns.str.strip()

# Display the column names to verify the exact names
print(df.columns)
```

```
Index(['DATE', 'PRODUCT', 'GENDER', 'SIZE', 'PIECES', 'RATE', 'TOTAL'], dtype='object')
```

```
[ ] # Convert 'DATE' column to datetime format
df['DATE'] = pd.to_datetime(df['DATE'], format='%m/%d/%Y', errors='coerce')

# Check the data types after conversion
print(df.dtypes)
```

```
DATE      datetime64[ns]
PRODUCT    object
GENDER     object
SIZE       object
PIECES     int64
RATE       float64
TOTAL      float64
dtype: object
```

Fig 2: Data cleaning

**Descriptive statistics:** The table showcases descriptive statistics summarizing the purchase/manufacturing and sales data of JK Clothing between May 2023 to August 2023. It provides key insights into the distribution of 'PIECES', 'RATE', and 'TOTAL' variables.

**Table 1: Purchase & Manufacturing data**

	PIECES	RATE	TOTAL
count	123.000000	123.000000	123.000000
mean	48.089431	1000.820325	30376.747967
std	178.207744	438.892689	21192.129289
min	10.000000	100.000000	10000.000000
25%	17.500000	830.000000	15450.000000
50%	30.000000	1000.000000	26100.000000
75%	50.000000	1100.000000	40037.500000
max	2000.000000	2500.000000	200000.000000

**Table 2: Sales Data**

	PIECES	RATE	TOTAL
count	123.000000	123.000000	123.000000
mean	48.089431	1601.312520	48602.796748
std	178.207744	702.228302	33907.406863
min	10.000000	160.000000	16000.000000
25%	17.500000	1328.000000	24720.000000
50%	30.000000	1600.000000	41760.000000
75%	50.000000	1760.000000	64060.000000
max	2000.000000	4000.000000	320000.000000

**Problem Statement 1: Enhancing Sales Performance**

**Most Purchased or Sold Product Analysis:** As a part of our comprehensive data analysis, we created a plot to visualize the distribution of the purchase or sold volumes across various items. Through this visualization, we identified top performers who help in decision-making related to inventory management, marketing strategies, and product development to enhance the overall sales performance. Notably, the product ‘SCRUNCHIE’ (See Fig 3 & 4) comes up as the top performer attributed to its compact nature and everyday utility. This type of product does not follow seasonal trends and

turns out to be a reliable and good investment. However, for long-term strategies, we can try to diversify our focus towards other products to capitalize on customer preferences and seasonal trends, thereby further boosting sales performance.

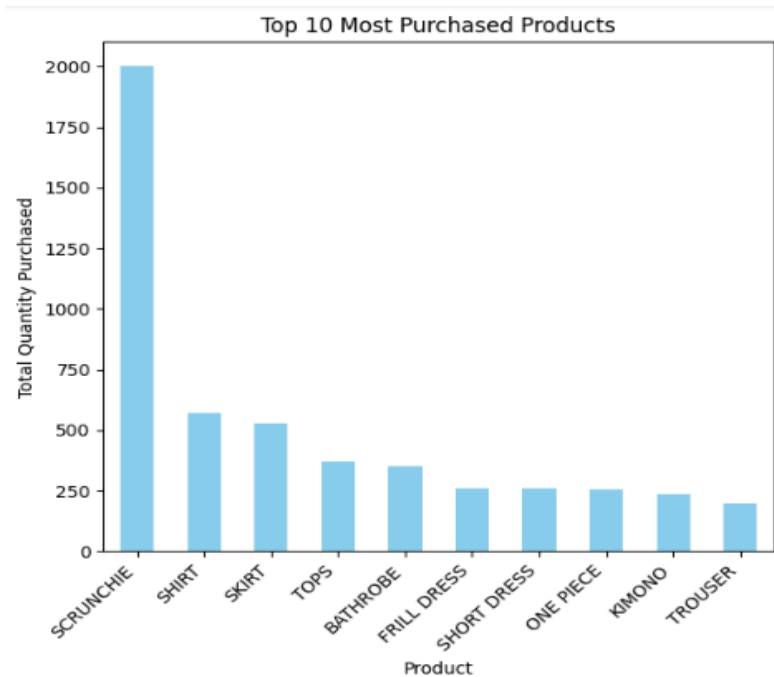


Fig 3: Top 10 Most Purchased Products

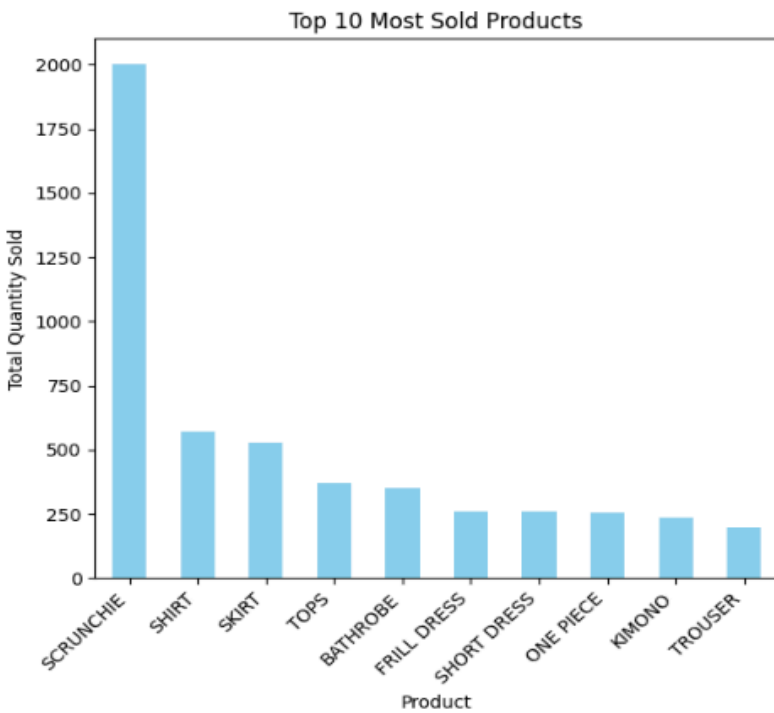


Fig 4: Top 10 Most Sold Products



**Pareto Analysis:** By using Pareto charts, we identified the most influential products and months contributing significantly to overall purchases and sales, aiding in strategic inventory management. Knowing the top contributing products helps in decision-making regarding pricing strategies or promotions to maximize revenue.

Unlike the bar charts in Fig 3 & 4 which identify top-performer products in a more primitive way, this is a more relative method to identify the top products. The Pareto chart in Fig 5 & 6 showcases the varying contributions of different months to total sales/purchases. It indicates that certain months, such as May 2023, significantly drive the overall purchase/sales volume. Afterward, there is a steep decline in purchase/manufacturing and sales volumes which is associated with the decline in tourism due to the peak summer months approaching signifying the impact of tourism on the sales of JK Clothing We can focus our effort on tourism-peaking months to maximize sales/profits. Additionally, there is a need for more data to study tourism-peak trends more in-depth.

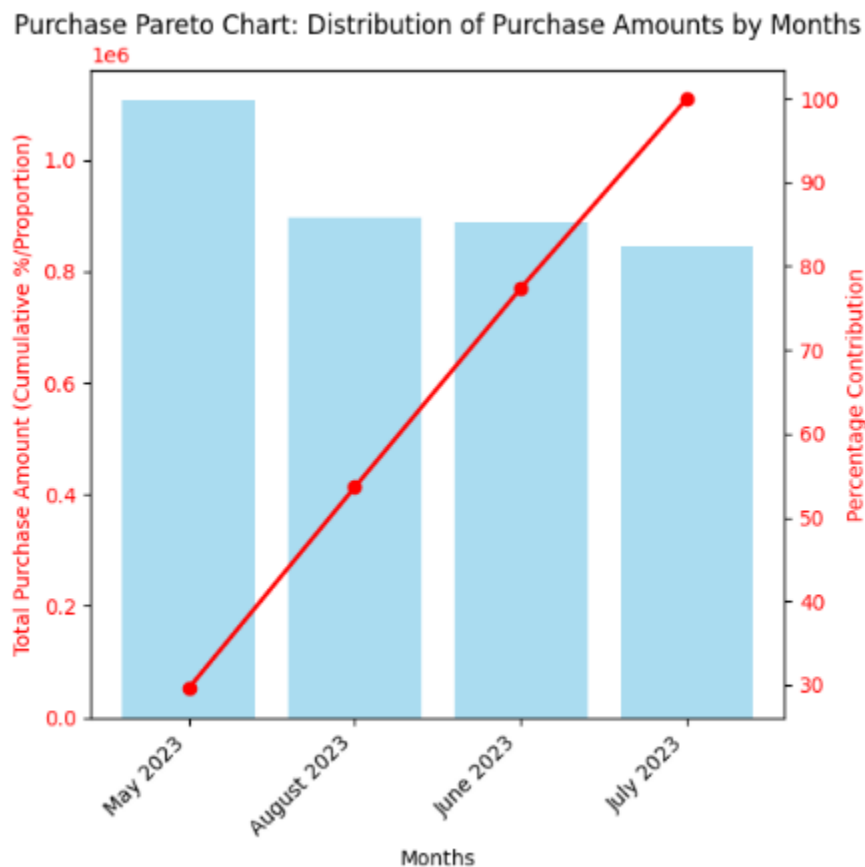


Fig 5: Purchase Pareto chart: Distribution of Purchase amounts by months

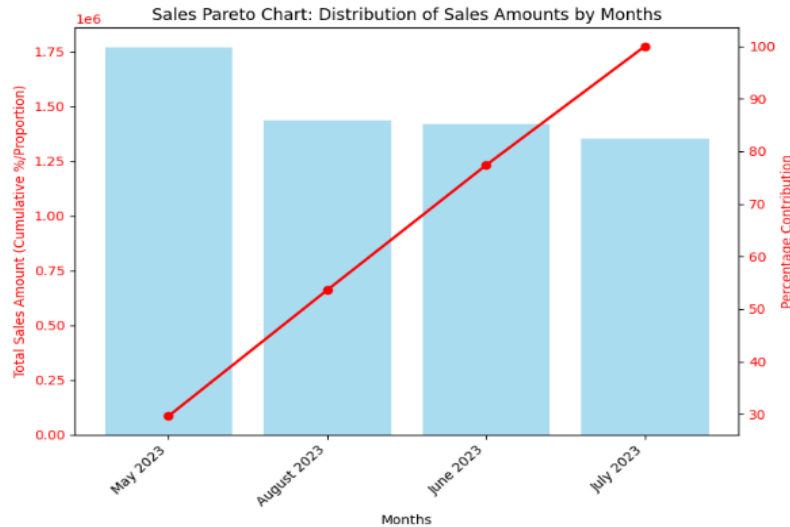


Fig 6: Sales Pareto chart: Distribution of Purchase amounts by months

**Sales forecasting using the ARIMA model:** ARIMA (AutoRegressive Integrated Moving Average) is a popular time series forecasting method that is used to predict future values based on historical data. We employed this model to do a simple sales forecasting for our dataset.

1. AutoRegressive(AR) Component: The ARIMA model considers previous observations of the dataset to predict a future observation using the linear combination of the observed values.
2. Integrated(I) Component: This step makes the time series stationary by removing seasonal fluctuations to ensure a more predictable pattern.
3. Moving Average(MA) Component: This considers the residual error from observations and accounts for short-term fluctuations in the series not attributed to the current and past values.

This analysis suggests that forecasted sales values are relatively consistent and stable across time, with minimal fluctuation. This may indicate a uniform projection, assuming a constant change in sales without significant peaks or drops (see Fig. 7).

This indicates a need for further investigation for suitable parameters for the ARIMA Model. This nearly straight line raises concerns about the fitting of the model and the suitability of the model parameters and implores the necessity for refining the forecasting model to get more accurate predictions for implementing effective business strategies.

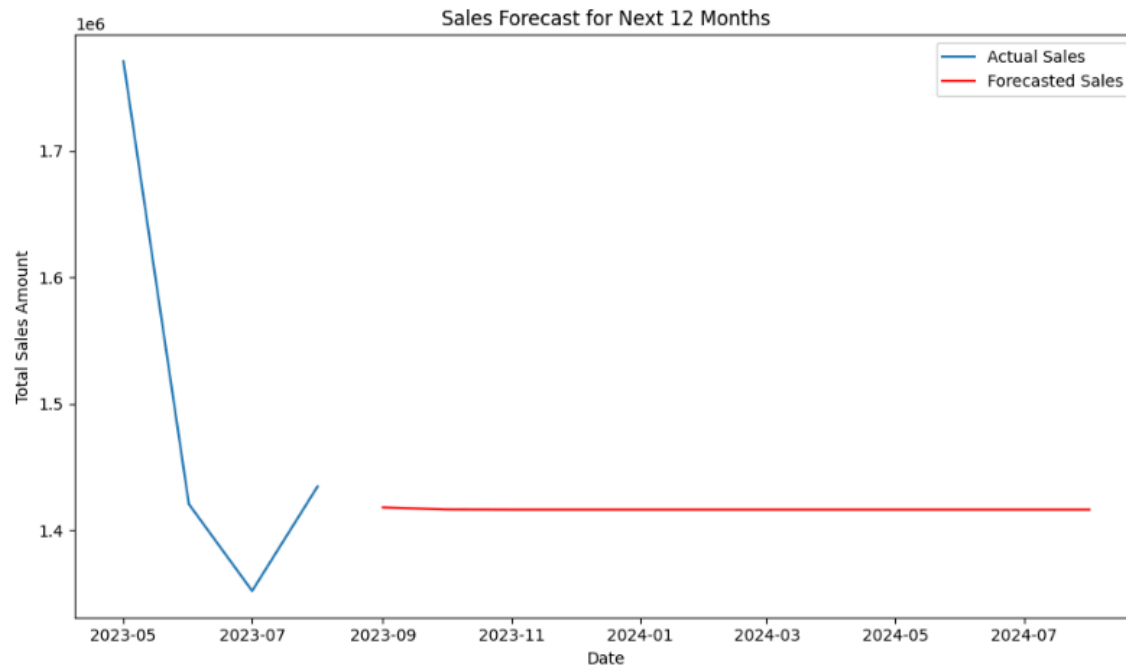


Fig 7: Sales forecast using ARIMA model

For optimum parameter selection for the ARIMA model, we defined values as:

AR component( $p$ ) = 1

I component( $q$ ) = 0

MA component( $d$ ) = 1

(Earlier, we were using default values for  $p$ ,  $q$ ,  $r$  as 1, 1, 1 respectively.)

This is the refined plot (see Fig 8) we created after optimal parameter selection using iterative refinement. This showcased enhanced accuracy compared to the previous model.

The updated model better captures the underlying patterns and observations in the sales data. By fine-tuning the parameters, the model's probability of being underfitted or overfit has been reduced, thus, this model strikes a better balance i.e. avoids oversimplification in the data. This forecast can be used to make data-driven strategies in the business.

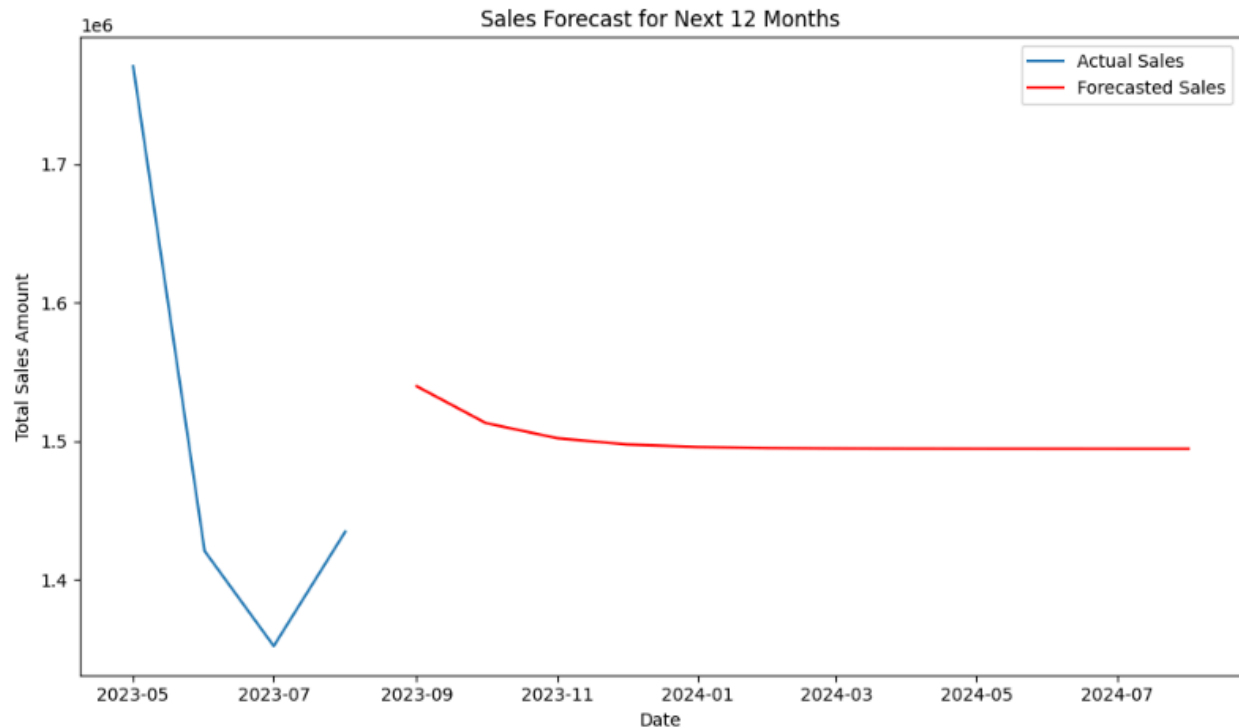


Fig 8: Modified sales forecast using ARIMA model

## Problem Statement 2: Inventory Optimization for Cost Efficiency

**Price distribution of products:** Price distribution helps to identify the range of costs incurred or procured. This helps in understanding the variability of price and negotiation for better pricing. This is also utilized to reveal profit margins and identify price bands. Insights gained from this analysis are crucial for various aspects of business strategy. Our analysis of price distribution highlights the top price points within the dataset. For purchased or manufactured products (Fig. 9), 1000 came as a frequently occurring price point. Similarly, for sold products (Fig. 10), 1500 came as a frequently occurring price point. These points highlight the prices preferred by the customers.

The price range for purchased or manufactured products and sold products of JK Clothing, a B2B business is 200-2500 and 200-4000 respectively. This range difference sheds light on the profit margin of the business and indicates the diverse pricing spectrum within the products.

These findings can be used to refine pricing strategies to align with customer preferences and market dynamics.

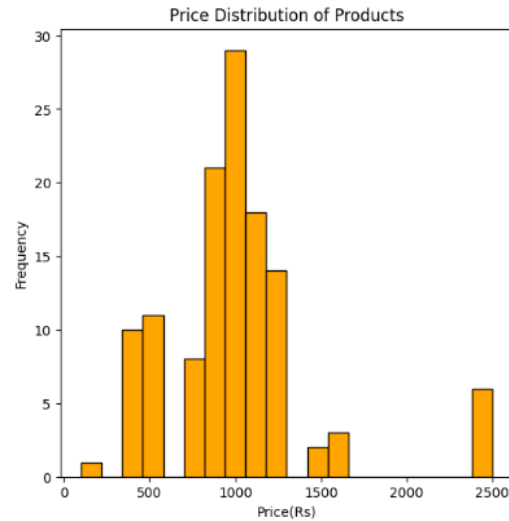


Fig 9: Price distribution of Purchased products

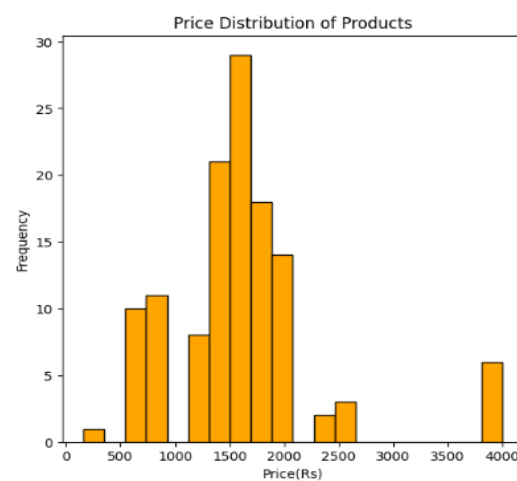


Fig 10: Price distribution of Sold products

**Time series analysis:** Time series analysis helps in providing insights into overall purchase and sale patterns and trends across months or time intervals. The time series analysis of monthly purchase/sale trends revealed a distinct pattern in the data. There is a steep decline in the monthly trends suggesting a seasonal behavior around the peak summer months when there is a sharp decline in tourism in Jaipur, Rajasthan where JK Clothing company is located. There is a need for additional data to confirm this claim and marketing strategies to solve this particular challenge. These insights provide the foundation for predictive analysis.

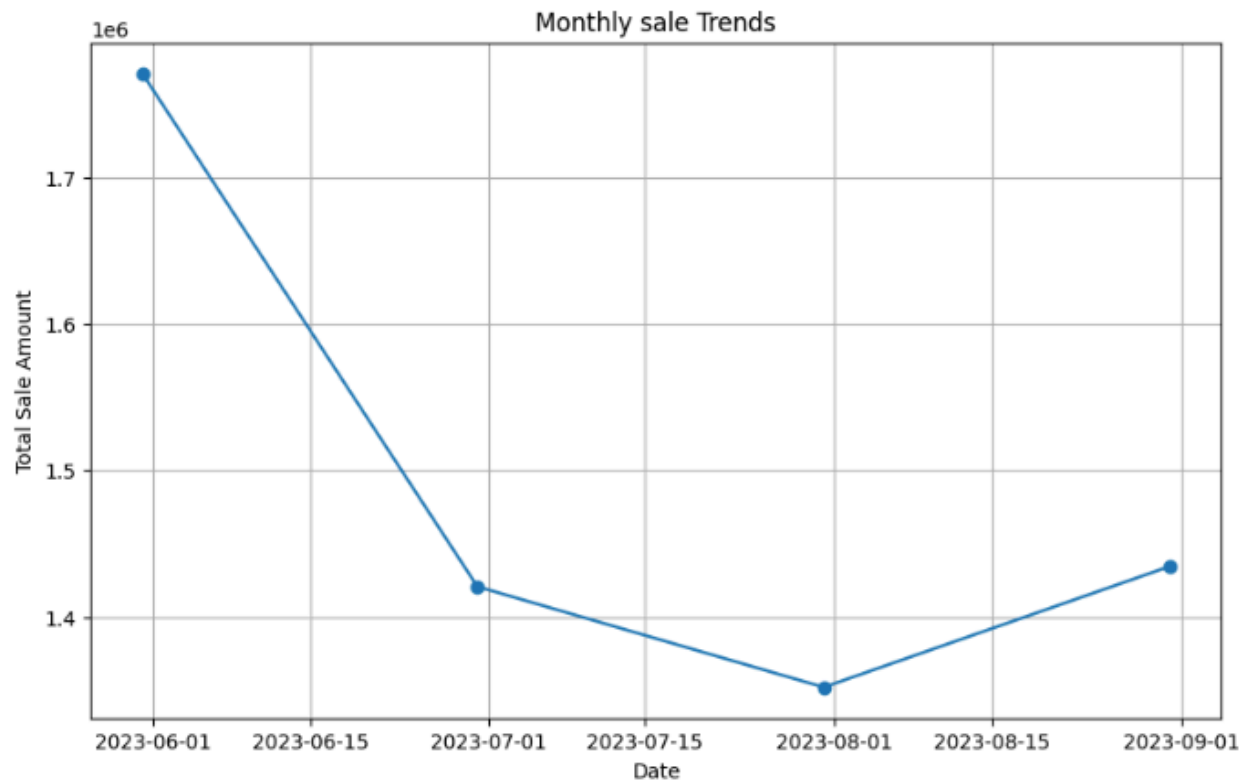


Fig 11: Monthly sale trends

### Problem Statement 3: Customer Behavior Analysis for Product Enhancement

**Gender & size analysis:** This analysis gave more insight into the customer preferences for optimum inventory management, product development, and targeted marketing strategies.

Upon analysis, the data revealed varying purchasing patterns between genders. Specifically, it was observed that females tend to show a higher preference for products than males. The analysis highlighted distinct size preferences where free size is the top performer. The gender vs size analysis heatmap (Fig 14) highlighted distinct size preferences among different gender groups. For example, Females demonstrated a stronger affinity towards free size compared to Males, indicating notable variations in size preferences.

Insights derived from gender and size analysis shed light on customer behavior, revealing nuanced preferences and purchase tendencies. For instance, Females showcased a major share in the product volume compared to other groups i.e. Males. These findings open avenues for targeted marketing initiatives. They provide the groundwork to devise personalized campaigns, discounts, or product variations that

resonate more effectively with distinct gender and size preferences.

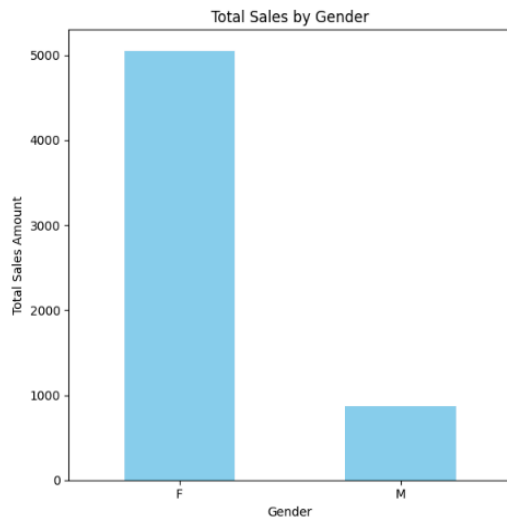


Fig 12: Total sales by Gender

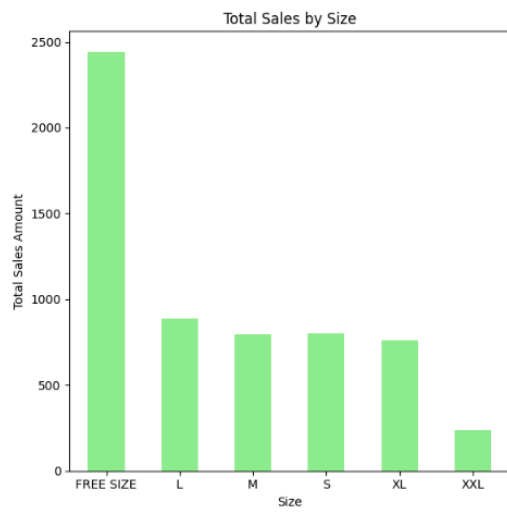


Fig 13: Total sales by Size

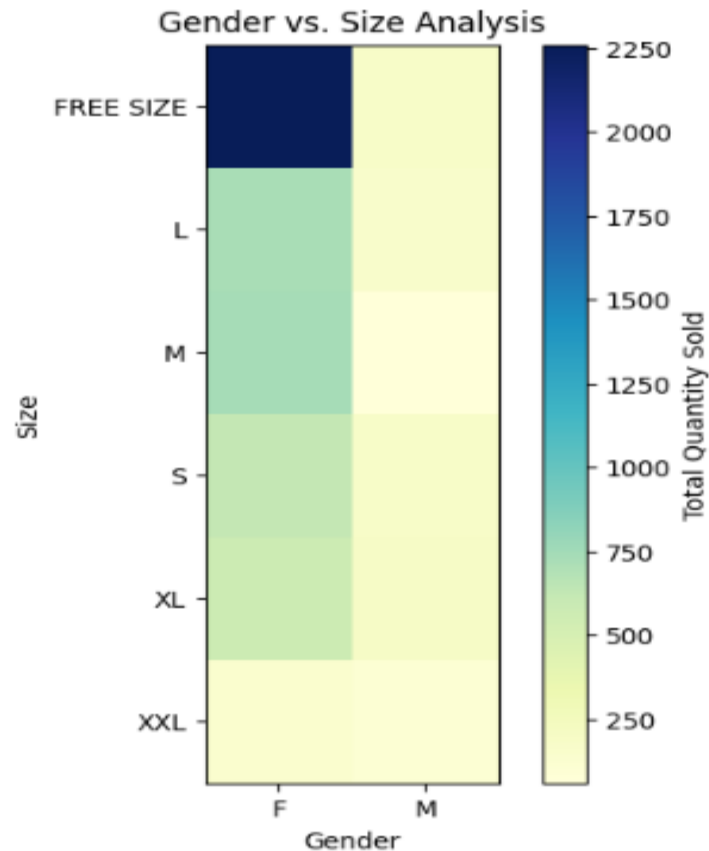


Fig 14: Gender vs. size analysis heatmap

#### 4. Interpretation of Results and Recommendations

The detailed analysis of the sales and purchase data from the JK Clothing company has provided valuable insights into the sales trends, customer behavior, and price variability associated with the dataset.

Based on the findings discussed above, the following recommendations can be made:

##### 1. Impact of Tourism:

The geographical location of JK Clothing in Jaipur, Rajasthan, a prominent tourist destination, significantly impacts the business. During peak summer months, the scorching weather in Jaipur significantly impacts tourism, leading to a sharp decline in footfall. Consequently, the company experiences a slump in overall production and sales during this period as observed in Fig 5 and 6.



## **Expanding Reach through E-commerce:**

To address this dependency on local tourism, it's essential to explore alternative paths. Given the rising trend of online shopping, ensuring various mediums of sales is essential. Leveraging established e-commerce platforms like Amazon, Flipkart, Meesho and other prominent marketplaces can serve as a gateway to reach a wider audience beyond the local region. The shift towards online retail not only widens the customer base but also opens doors to new markets, potentially increasing sales performance.

## **Addressing Initial Challenges:**

While venturing into e-commerce may pose initial logistical and manpower challenges, such as establishing a robust delivery network and managing online inventories, the long-term benefits overshadow these obstacles. Investing in a well-structured online platform can significantly reduce the adverse impact of tourism or seasonal fluctuation on sales, offering stability to the business.

## **2. Need for Additional Data:**

The current data from May to August 2023 provides valuable insights into inventory management practices during slower tourism. However, for a comprehensive understanding, data from peak tourism months in Jaipur is essential. Extending the dataset to cover broader time intervals, specifically during peak tourist seasons, would facilitate a holistic review of consumer behavior and purchasing trends.

## **Enhancing Data-Driven Strategies:**

Access to additional data covering peak tourism months gives a more unhindered analysis, helping to formulate data-driven strategies. By observing consumer preferences and purchasing behaviors during peak tourism, the company can more effectively tailor marketing and inventory management strategies. Moreover, this extended dataset enhances the accuracy of the predictive models, allowing for more precise sales forecasts.

## **3. Customer Behavior Insights:**

The comprehensive analysis of gender and size preferences, price distribution, and time series data offers invaluable insights into customer behavior.

## **Short-Term and Long-Term Strategy Alignment:**

Identifying the top-performing categories, such as the FEMALE category and FREE SIZE products, presents opportunities for increased short-term production to meet immediate demand. However, for sustainable growth, diversification and promotion of other product categories are essential. Utilizing these insights will aid in tailoring marketing strategies to align with customer preferences effectively.

#### **4. Supply Chain Optimization:**

The analysis of purchase and sales data presents opportunities for strategic supply chain optimization. The data highlights trends showcasing varying consumer preferences and seasonal demand fluctuations. By leveraging these insights, JK Clothing can strategically recalibrate its supply chain operations to align better with these seasonal dynamics.

##### **Understanding Seasonal Demand Fluctuations:**

During peak seasons or periods of increased tourism, there's a noticeable surge in demand for specific garment categories or sizes. Conversely, off-peak months witness a slowdown in consumer activity, resulting in reduced sales for certain products. This pattern implies a direct correlation between demand and external factors like tourist arrivals or seasonal preferences.

##### **Flexible Supply Chain Strategy:**

Implementing a more adaptive and responsive supply chain strategy can be advantageous. A flexible approach allows the company to efficiently manage inventory levels by adjusting production volumes based on anticipated demand. For instance, during peak tourist seasons or periods of high demand, production can be scaled up to meet the expected surge in sales. Conversely, during quieter periods, reducing production can prevent excess inventory buildup.