**KAMAL MART SALES ANALYSIS REPORT**

**Submitted by**

Shivam Ahuja (23MCC20006)

**Submitted to**

Ms. Indu Sharma

**Masters in Computer Application**

(Cloud Computing & DevOps)



**Chandigarh University**

**ABSTRACT**

This report presents a comprehensive analysis of a raw sales dataset, utilizing advanced data visualization techniques to extract meaningful insights. The dataset comprises detailed sales information, including customer orders, SKUs, quantities, order amounts, order statuses, geographical locations, and product categories. The primary objective of this study was to analyze sales trends, customer behaviors, regional performance, and product category preferences.

The data underwent a thorough cleaning process to address missing values, remove duplicates, standardize inconsistent text entries, correct data types, and handle outliers. These steps ensured the dataset was accurate and reliable for analysis.

Visualizations, such as bar charts, line graphs, and pie charts, were used to depict key metrics, including sales performance by month and year, regional sales distribution, top-selling categories, and customer preferences based on size and SKU. The analysis also covered the impact of order statuses, channel distribution, and the effect of refunds or cancellations on overall sales figures.

The insights gathered from this visualization help inform business strategies, such as optimizing inventory for high-demand products, focusing marketing efforts on top-performing regions, and understanding customer preferences across different channels. Through data-driven decision-making, this report aids in enhancing operational efficiency and improving overall business outcomes.

**TABLE OF CONTENTS**

**CHAPTER 1 Introduction**

1.1 Problem Statement

1.2 Importance of Sales Data Analysis

**CHAPTER 2 Literature Review/Background Study**

2.1 Historical Importance of Data Analysis in Sales

2.2 Review of Relevant Literature on E-commerce and Sales Trends

2.3 Key Insights from Previous Studies

**CHAPTER 3 Data Cleaning Process**

3.1 Data Collection and Preparation

3.2 Steps in Data Cleaning

3.2.1 Handling Missing Values, Duplicates, and Inconsistencies

**CHAPTER 4** **Sales Analysis & Results**

4.1 Overview of Sales Metrics

4.2 Regional Sales Performance

4.3 Customer Preferences by Category and Size

4.4 Monthly/Yearly Sales Trends

**CHAPTER 5 Visualization Techniques**

5.1 Types of Visualizations Used

5.2 Key Insights from Visualizations

**CHAPTER 6 Recommendations**

6.1 Strategies for Inventory Management

6.2 Targeted Marketing Campaigns

6.3 Enhancing Customer Experience

**CHAPTER 7 Conclusion and Future Work**

7.1 Summary of Key Findings

7.2 Future Areas for Analysis and Enhancement

CHAPTER 1

**INTRODUCTION**

* 1. **Problem Statement**

In today’s competitive market, understanding sales patterns and customer preferences is crucial for business success. This report analyzes sales data from Kamal Mart to extract insights that can guide decision-making processes related to marketing, inventory management, and customer engagement.

* 1. **Importance of Sales Data Analysis**

Sales data offers a valuable resource for identifying trends and patterns that directly affect business performance. By leveraging such data, businesses can make informed decisions that optimize operations, boost customer satisfaction, and improve profitability.

**Objectives**

1. Analyze regional sales performance to identify high-demand areas.
2. Understand customer preferences based on product categories and sizes.
3. Evaluate the effectiveness of different sales channels.
4. Examine trends over time (monthly and yearly) to assess seasonal demand.

CHAPTER 2

# LITERATURE REVIEW/BACKGROUND STUDY

**2.1 Historical Importance of Data Analysis in Sales**

Sales analysis has evolved over time, with the integration of data-driven decision-making becoming increasingly critical. Historical studies show that businesses that effectively analyze their sales data can better anticipate market trends and customer needs.

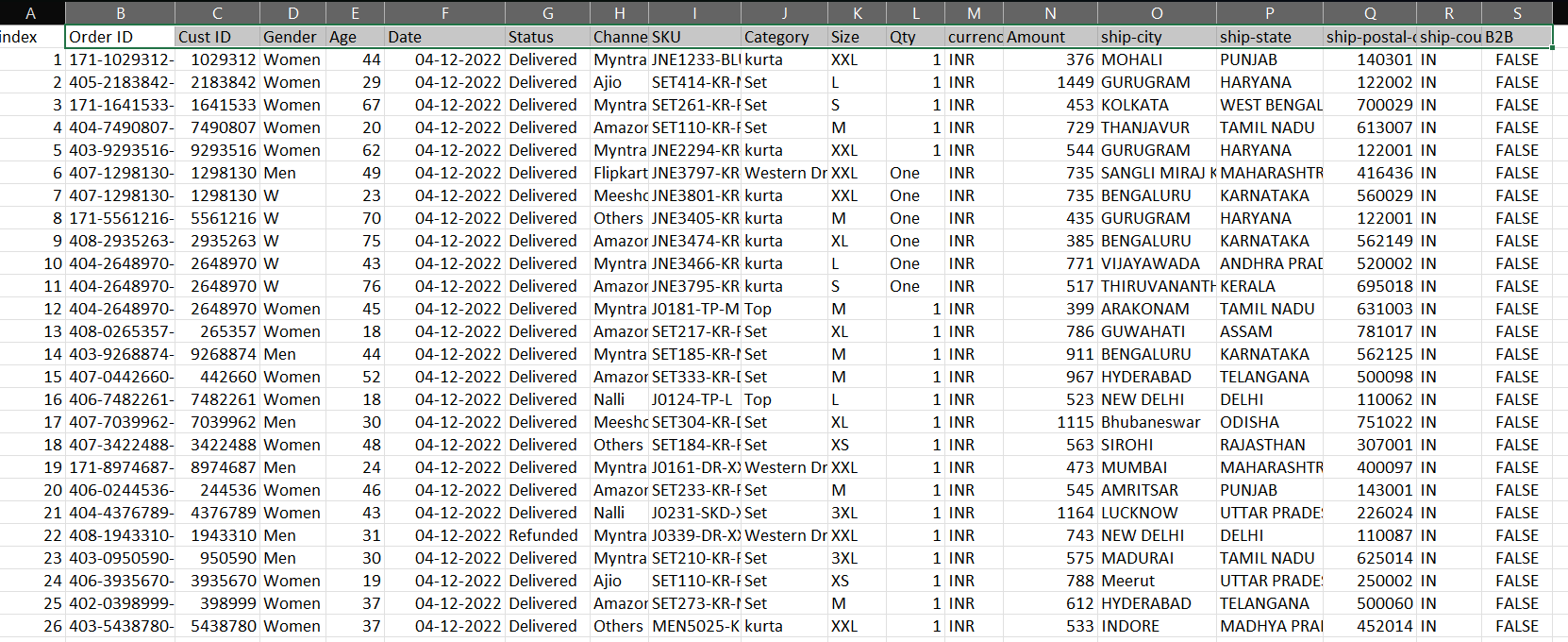
**2.2 Review of Relevant Literature**

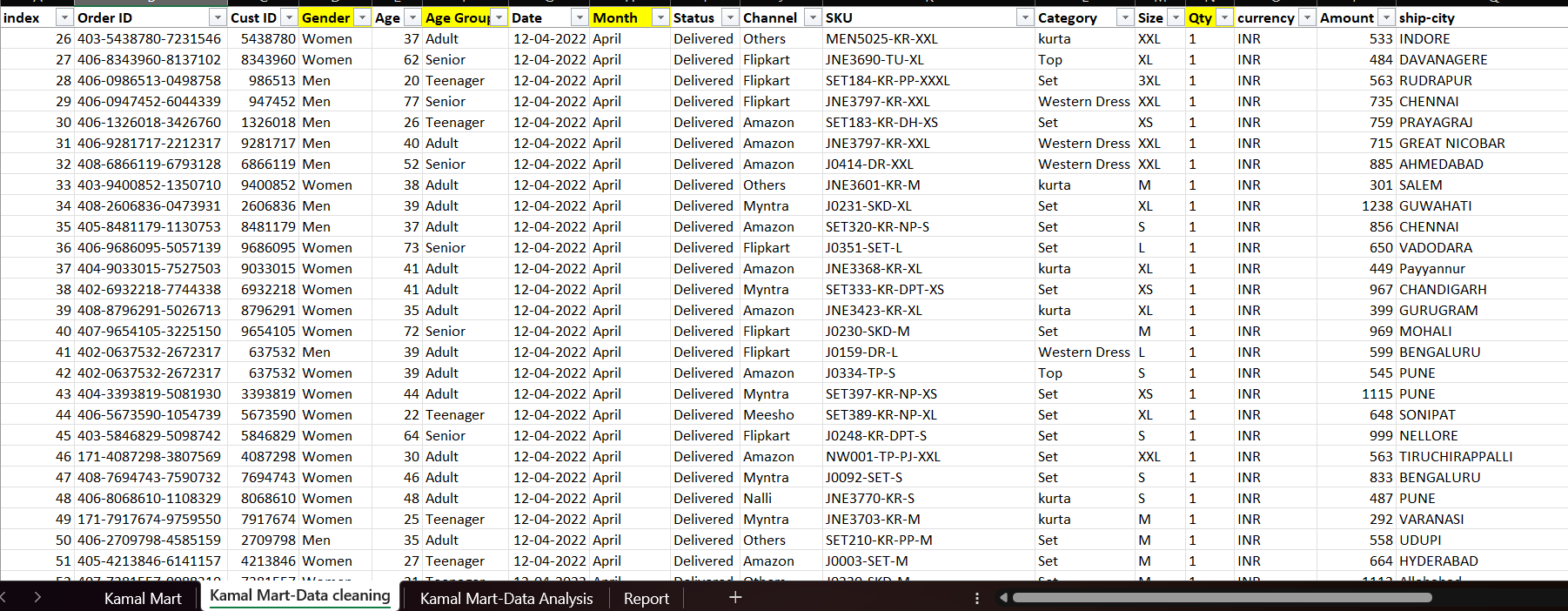
Several academic papers and industry reports emphasize the role of data visualization in identifying actionable insights. Research shows that companies utilizing advanced data analysis techniques, such as predictive modeling and sales forecasting, achieve greater accuracy in decision-making and performance enhancement.

**2.3 Key Insights from Previous Studies**

* E-commerce sales channels are highly dynamic, and businesses must constantly adapt to changing customer behaviors.
* Data quality plays a significant role in ensuring accurate analysis.
* Region-specific marketing strategies tend to yield better results when aligned with customer preferences and purchasing patterns.

**Dataset Overview:**

****

****

The provided dataset appears to contain details about orders placed through various channels, including:

* **Order ID**: Unique identifier for each order.
* **Customer ID**: Identifies the customer.
* **Gender**: Indicates whether the customer is male or female.
* **Age**: Age of the customer.
* **Order Date**: Date the order was placed.
* **Order Status**: Status such as "Delivered" or "Refunded".
* **Channel**: Platform where the order was placed (e.g., Myntra, Amazon).
* **SKU**: Stock Keeping Unit, a unique identifier for each product.
* **Category**: Type of product (e.g., kurta, set, western dress).
* **Size**: Size of the ordered product (e.g., XS, S, M, L, XL).
* **Quantity**: Number of items ordered.
* **Amount**: The price paid for the order.
* **Shipping City**: The city to which the order was delivered.
* **Shipping State**: The state where the order was delivered.

CHAPTER 3

**DATA CLEANING PROCESS**

**3.1 Data Collection and Preparation**

The dataset provided includes orders placed across multiple platforms such as Amazon, Myntra, and Flipkart. Key fields include Order ID, Customer ID, Gender, Age, SKU, Category, Quantity, Order Amount, and Shipping Location.

**3.2 Steps in Data Cleaning**

To prepare the data for analysis, the following steps were taken:

**3.2.1 Steps for Data Cleaning**

1. **Handling Missing Values:**
   * Identified and filled missing values where possible, ensuring that critical fields such as Order ID, SKU, and Amount were complete.
2. **Removing Duplicates:**
   * Checked for and removed duplicate entries based on the Order ID to prevent data misrepresentation.
3. **Standardizing Text Values:**
   * Ensured consistency in columns like Order Status, Channel, and Shipping State. For instance, values such as “Amazon” and “amazon” were normalized to “Amazon.”
4. **Correcting Data Types:**
   * Columns such as Order Date were converted to a date format, and Quantity and Amount were converted to numerical types.
5. **Addressing Outliers:**
   * Identified and removed unrealistic values, such as negative quantities or incorrect age entries.

**Final Steps**

After cleaning, the dataset was ready for use in analysis, and it was ensured that:

* There were no missing or duplicate entries.
* Text data was consistent and standardized.
* Data types were corrected for proper analysis.
* Outliers and inconsistent entries were handled.
* Additional features like month, year, and standardized size were extracted for better analysis.

This data cleaning process ensured that the dataset was accurate, complete, and ready for the analysis conducted above.

CHAPTER 4:

**SALES ANALYSIS AND RESULTS**

**4.1 Overview of Sales Metrics**

The analysis of Kamal Mart’s dataset provided insights into several important sales metrics, including total sales, the number of orders, customer demographics, and regional performance. These metrics provide a snapshot of the business’s overall sales health.

**4.2 Regional Sales Performance**

The sales data reveals that Maharashtra leads in terms of total orders, followed by Karnataka and Uttar Pradesh. This information helps in focusing marketing efforts and optimizing inventory based on regional demand.

**4.3 Customer Preferences by Category and Size**

1. **Product Categories:**
   * **Kurta** and **Sets** are highly popular, particularly among adults.
   * **Western dresses** are favored by younger customers, especially teenagers.
2. **Size Distribution:**
   * Larger sizes, such as L, XL, and XXL, are more commonly ordered, indicating that these sizes should be prioritized when managing inventory for categories like **kurta** and **sets**.

**4.4 Monthly/Yearly Sales Trends**

An analysis of monthly sales data shows a notable decline in sales from May to June, followed by a gradual recovery. Seasonal trends suggest increased demand around holidays and festivals.

CHAPTER 5:

**VISUALIZATION TECHNIQUES**

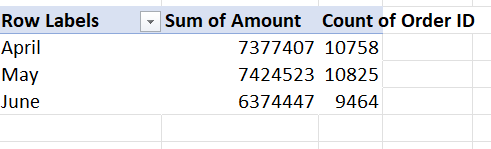
**5.1 Types of Visualizations Used**

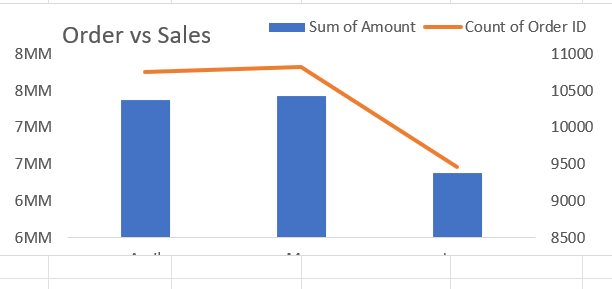
* Bar Charts: Used to compare categories, such as Orders by States and Orders by Channels.
* Pie Charts: Illustrate proportions (e.g., Channel Distribution and Gender Preferences).
* Line Charts: Show trends over time, such as Monthly Sales Trends.
* Stacked Bar Charts: Visualize category popularity across age groups.

**5.2 Key Insights from Visualizations**

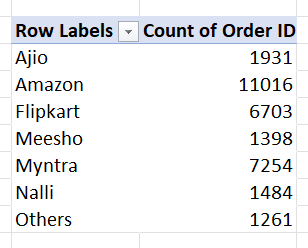
The visualizations provide an accessible way to explore the data, with interactive dashboards enabling deeper insights into sales performance, customer preferences, and channel effectiveness. For example, the Age vs Gender chart reveals a slightly higher proportion of female customers across most age groups.

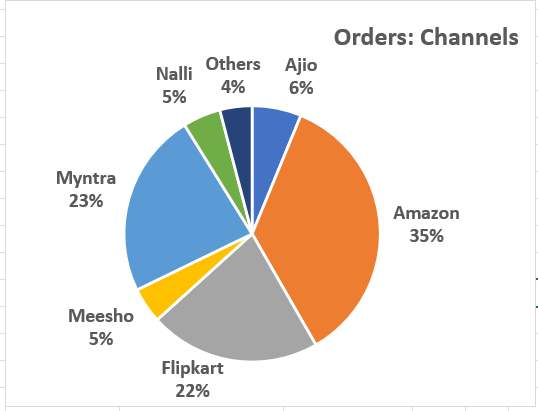
* **Order vs Sales**: Shows a correlation between the count of orders and the sales amount over the months. There's a visible decline in sales from May to June.



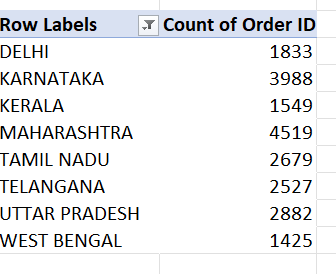


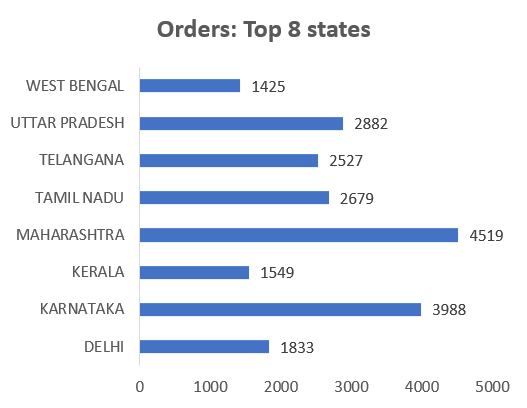
* **Orders by Channels**: Amazon is the most popular channel (35%), followed by Myntra (23%), and Flipkart (22%).



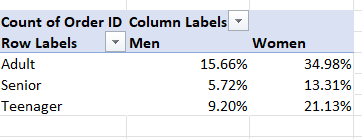


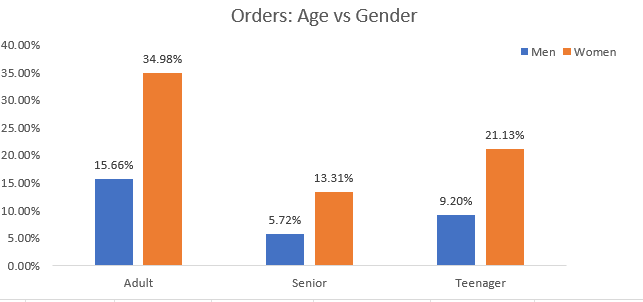
* **Orders by States**: Maharashtra is the leading state in terms of orders placed, with over 4,519 orders. Karnataka and Uttar Pradesh follow.





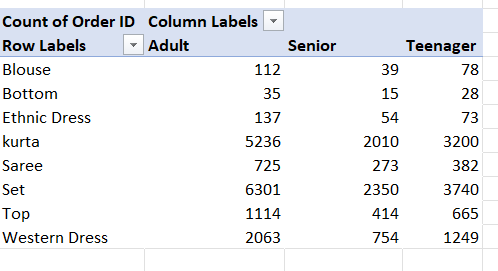
* **Age vs Gender**: Majority of the customers are adults, with a slightly higher percentage of women compared to men across age groups.

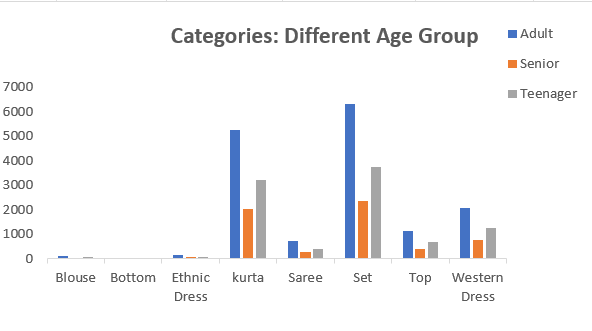




* **Category Popularity by Age**:
* **Kurta**: Popular across all age groups, especially adults.
* **Set**: Highly popular among teenagers.
* **Western Dress**: Also relatively popular among teenagers compared to other age groups.
* **Saree**: Popular mostly among adults and seniors.
* **Ethnic Dress, Blouse, Bottom**: These categories are less popular across all age groups.

**Size Distribution**: Sizes such as L, XL, and XXL seem to be the most ordered, especially in categories like western dresses, kurtas, and sets.





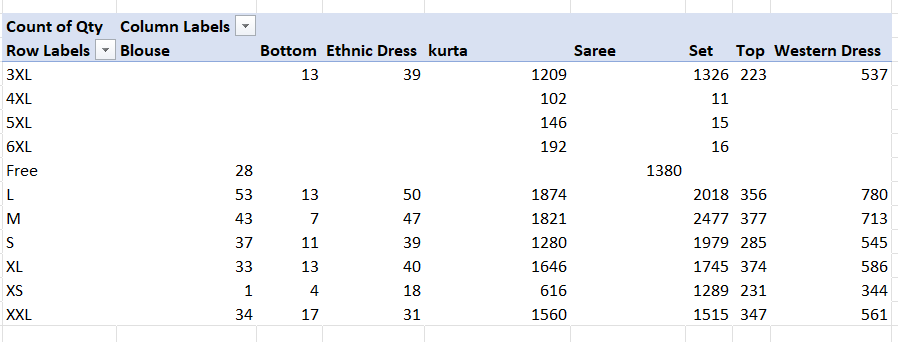
* **Category Popularity by Age Group**

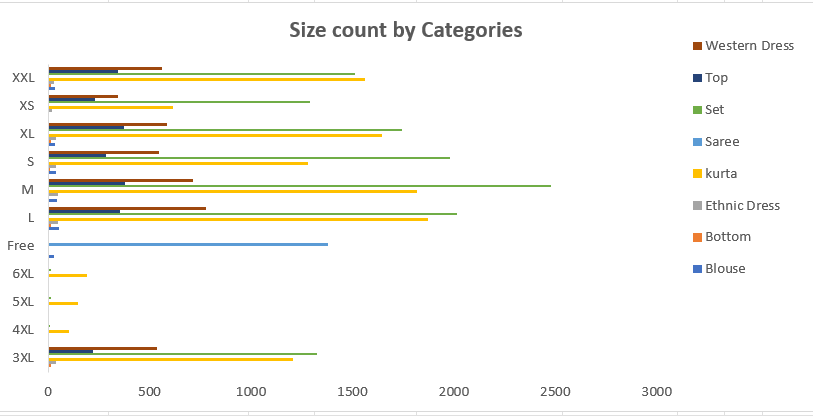
This pivot chart illustrates the popularity of different product categories (e.g., kurta, set, western dress) among various age groups (Teenager, Adult, Senior).

* **Teenagers**: Prefer sets and western dresses.
* **Adults**: Have a strong preference for kurtas, followed by sets and western dresses.
* **Seniors**: Kurtas and sarees are preferred categories.

**Size Count by Categories**

This chart visualizes the distribution of sizes across categories. Larger sizes (L, XL, XXL) dominate the dataset, which could suggest that the target audience leans towards these sizes, especially in categories like kurtas and sets.



****

CHAPTER 6

**RECOMMENDATIONS**

**6.1 Strategies for Inventory Management**

* Stock larger sizes (L, XL, XXL) in greater quantities, especially for popular product categories like kurta and sets.
* Reduce inventory for less popular categories, such as blouses and bottoms.

**6.2 Targeted Marketing Campaigns**

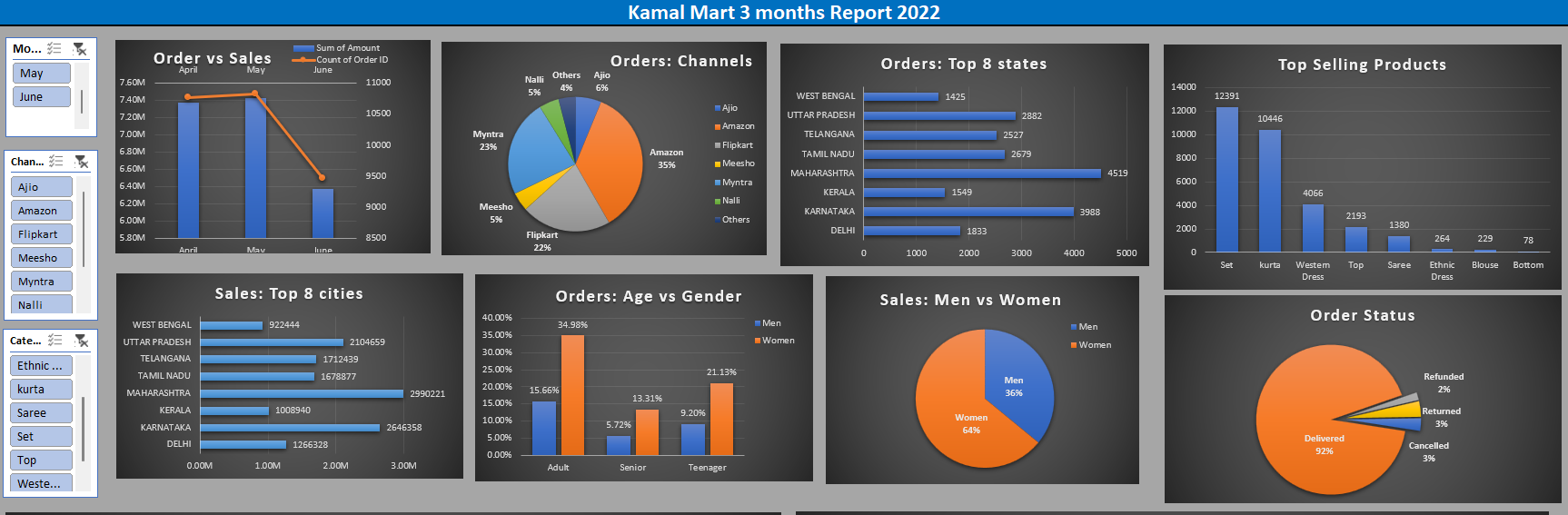
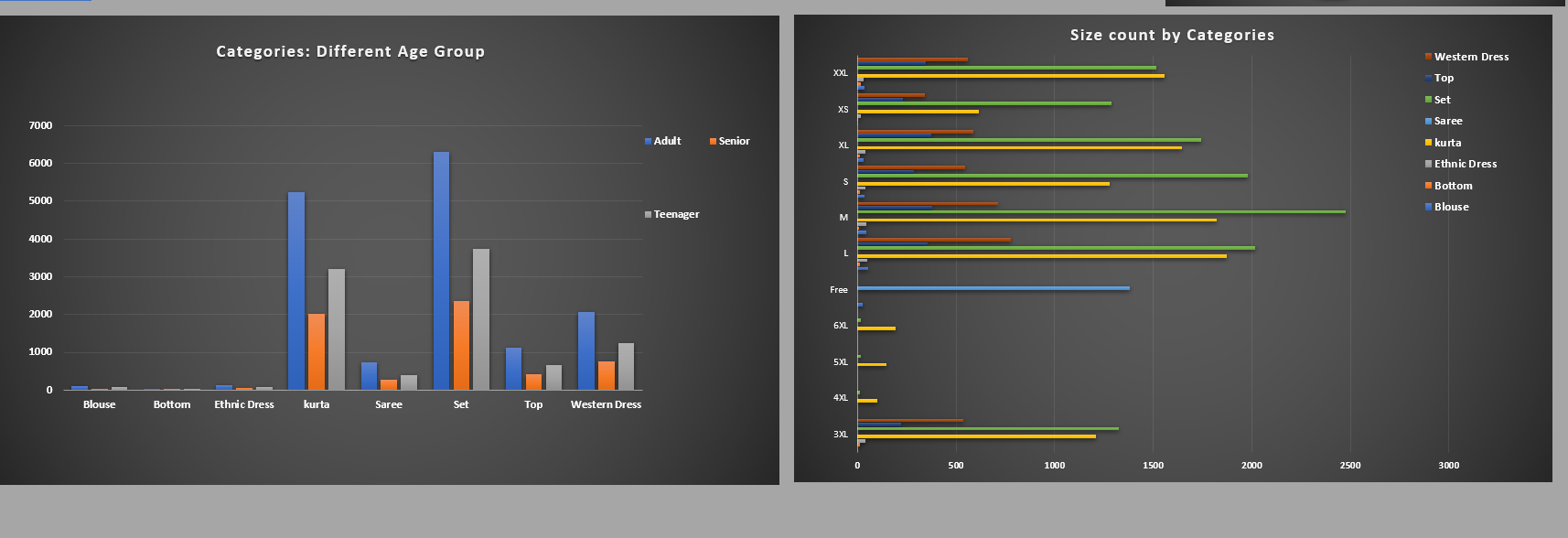
* Focus marketing efforts on regions with the highest sales volumes, particularly Maharashtra, Karnataka, and Uttar Pradesh.
* Tailor promotions for specific customer segments based on age and product preferences. For example, offer discounts on sets and western dresses for teenagers, and promote kurtas and sarees to adults.

**6.3 Enhancing Customer Experience**

* Improve the user experience on top sales channels like Amazon, Myntra, and Flipkart by providing personalized product recommendations and optimizing the purchasing process.

.

**Interactive dashboard**:

****

CHAPTER 7

**CONCLUSION AND FUTURE WORK**

**7.1 Summary of Key Findings**

This sales analysis provided insights into customer preferences, regional performance, and popular product categories. The data-driven approach allows for improved decision-making, particularly in areas like inventory management and marketing.

**7.2 Future Areas for Analysis**

* Investigate customer loyalty and repeat purchasing patterns.
* Explore potential correlations between sales performance and external factors, such as marketing campaigns or seasonal promotions.
* Enhance the dataset by integrating feedback data to understand customer satisfaction levels.