Innobyte Services

Amazon Sales Report: Data-Driven Insights for Business Optimization

Prepared For: InnoByte Services
Prepared By: Koushik Bandyopadhyay
Internship Role: Data Analyst Intern
Date: June 20, 2025

Table of Content :

1.Sales Overview: Understand the overall sales performance, trends, and patterns over time. And visualizations

2. Product Analysis: Analyze the distribution of product categories, sizes, and quantities sold to identify popular products. And visualizations

3.Fulfillment Analysis: Investigate the fulfillment methods used and their effectiveness in delivering orders. And visualizations

4. Customer Segmentation: Segment customers based on their buying behaviour, location, and other relevant factors. And visualizations

5.Geographical Analysis: Explore the geographical distribution of sales, focusing on states and cities. And visualizations

6. Business Insights: Provide actionable insights and optimization ways

Executive Summary

This report presents a comprehensive data analysis of the Amazon Sales Report dataset, aiming to provide actionable insights for optimizing business operations, enhancing customer satisfaction, and driving revenue growth. The analysis, conducted using Python, meticulously examined over 128,000 sales records, focusing on sales performance, product popularity, fulfillment efficiency, and geographical distribution.

Key findings highlight:

- The dominant contribution of apparel, particularly T-shirts, to overall sales volume.
- A significant challenge posed by a 20% order cancellation rate.
- The superior efficiency and reliability of Amazon's fulfillment services compared to merchant fulfillment.
- Sales concentration in major Indian states and cities, alongside identification of high-Average Order Value (AOV) niche markets.

Based on these insights, a set of strategic recommendations has been developed. These include optimizing inventory for popular products, implementing targeted interventions to reduce cancellations, leveraging Amazon's robust fulfillment network, and tailoring marketing efforts to specific geographical segments.

Successful implementation of these recommendations is expected to foster sustained growth and improved operational excellence.

Introduction & Project Objectives :

Problem Statement

The core problem addressed by this project is to analyze and provide actionable insights from the Amazon Sales Report to support informed business decision-making.

Key Objectives

To achieve this, the analysis was guided by the following specific objectives:

- **Sales Overview:** To understand the overall sales performance, including trends and patterns over time.
- **Product Analysis:** To analyze the distribution of product categories, sizes, and quantities sold to identify popular products and optimize product strategy.
- **Fulfillment Analysis:** To investigate the effectiveness of different fulfillment methods in delivering orders and identify areas for logistical improvement.
- **Customer Segmentation (via Location & Behavior):** To gain insights into customer buying behavior and preferences based on geographical location and transaction type (B2B/B2C).
- **Geographical Analysis:** To explore the distribution of sales across different states and cities, identifying key markets and growth opportunities.
- **Business Insights & Recommendations:** To synthesize findings into actionable insights and provide concrete recommendations for optimizing sales strategies, improving customer satisfaction, and enhancing overall business performance.

Methodology

Data Source

The primary data source for this analysis was the "Amazon Sale Report.csv" file, a tabular dataset containing detailed information on various sales transactions.

Tools Used

The analysis was performed using Python, leveraging key data science libraries:

- **Pandas:** For robust data loading, manipulation, and cleaning.
- **Matplotlib:** For fundamental plotting and visualization.
- **Seaborn:** For creating aesthetically pleasing and informative statistical graphics.

Data Preprocessing

Rigorous data preprocessing was essential to ensure the accuracy and reliability of the analysis. The following steps were undertaken:

- **Column Removal:** The New and PendingS columns were identified as completely null and subsequently dropped to streamline the dataset.
- **Data Type Conversion:** The Date column, initially an object type, was converted to datetime objects to facilitate time-series analysis. The ship-postal-code column was converted to a string type, as it represents a code rather than a numerical value for calculation.
- **Handling Missing Values:** Missing Amount values were imputed with 0, primarily observed in Cancelled or Unshipped orders, aligning with their zero revenue contribution.

1. Missing currency values were filled with 'INR', given its exclusivity as the currency throughout the rest of the dataset.
2. Missing shipping details (ship-city, ship-state, ship-postal-code, ship-country) were imputed with 'Unknown'. This approach

retained complete order records while clearly marking unprovided shipping information, especially for non-shipped orders.

3. The fulfilled-by column, which had a significant number of missing values, was intelligently imputed: if Fulfilment was 'Amazon' and fulfilled-by was null, it was assigned 'Amazon FBA'. Remaining nulls were categorized as 'Unknown Merchant' to distinguish fulfillment types.

- **Duplicate Check:** A check for duplicate rows was performed, and no exact duplicate entries were found, ensuring unique transaction records. These preprocessing steps resulted in a clean, consistent, and well-structured dataset ready for in-depth analytical exploration.

Analytical Approach

The analytical approach involved an iterative process:

- **Initial Data Exploration:** Gaining a foundational understanding of data distribution, relationships, and potential anomalies.
- **Targeted Analysis:** Applying specific analytical techniques (e.g., grouping, aggregation, time-series plotting) to address each key objective.
- **Deeper Dive:** Performing more granular analysis on areas revealing significant patterns or anomalies to uncover underlying causes or opportunities.
- **Visualization:** Creating a variety of charts and graphs to visually represent findings, making complex data interpretable.
- **Insight Generation:** Interpreting the results of the analysis to formulate clear, concise, and actionable business insights.
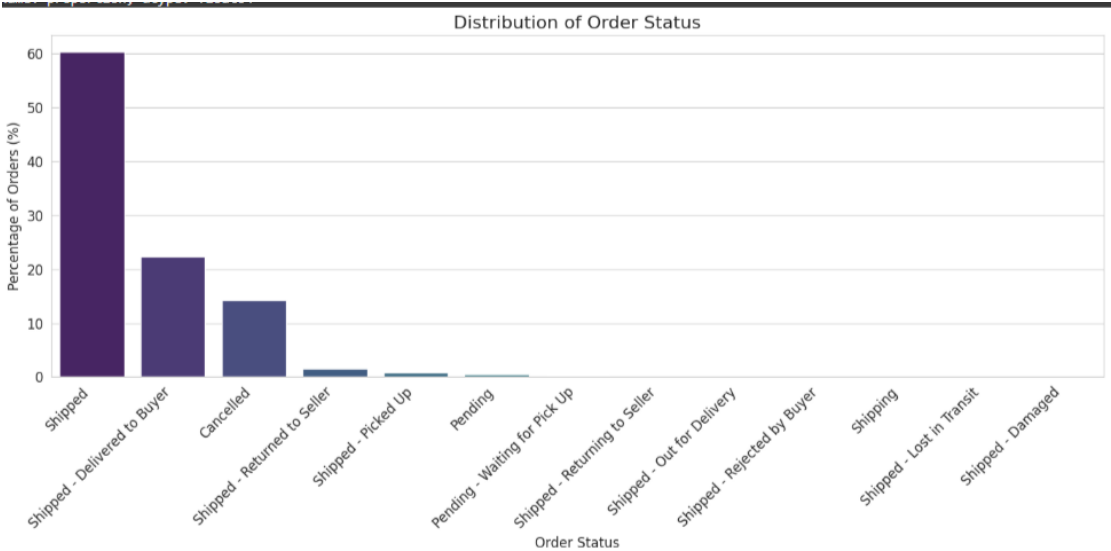
Detailed Analysis

Sales Performance Overview

The total sales amount derived from the dataset is INR 78,610,634.69, representing 116,634 units sold.

Order Status Distribution

The distribution of order statuses provides a crucial overview of operational efficiency and customer journey completion.

Order Status Distribution (Bar Chart):

Insight: Approximately 60% of all orders are successfully Shipped, with Shipped - Delivered to Buyer forming a significant portion of successful deliveries. However, a notable 20% of orders are Cancelled. This high cancellation rate is a critical area that demands immediate attention for investigation and resolution. Other statuses like Pending, Shipped - Returned to Seller, and Shipped - RTO account for smaller proportions.
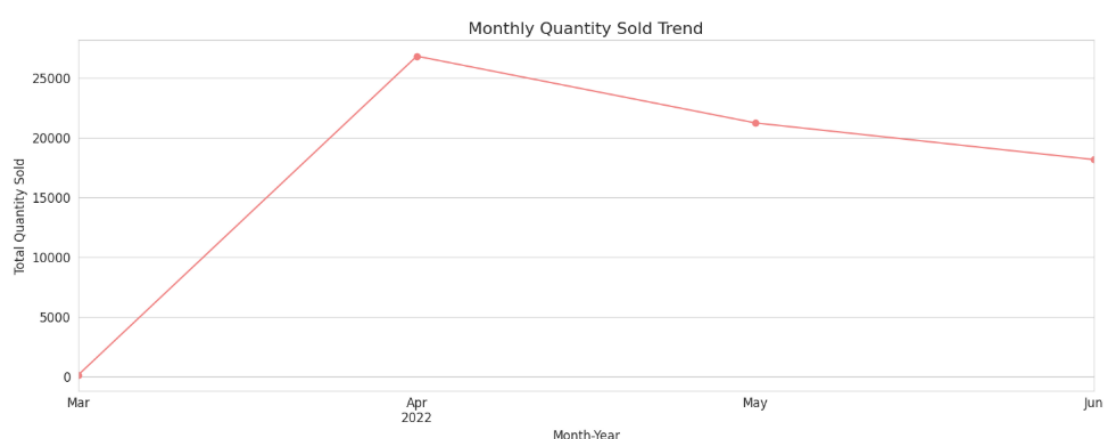
-

Sales Trends Over Time

Analyzing sales trends over time helps in identifying patterns and understanding seasonal or promotional impacts:

Monthly Sales Trend (Amount) (Line Chart):
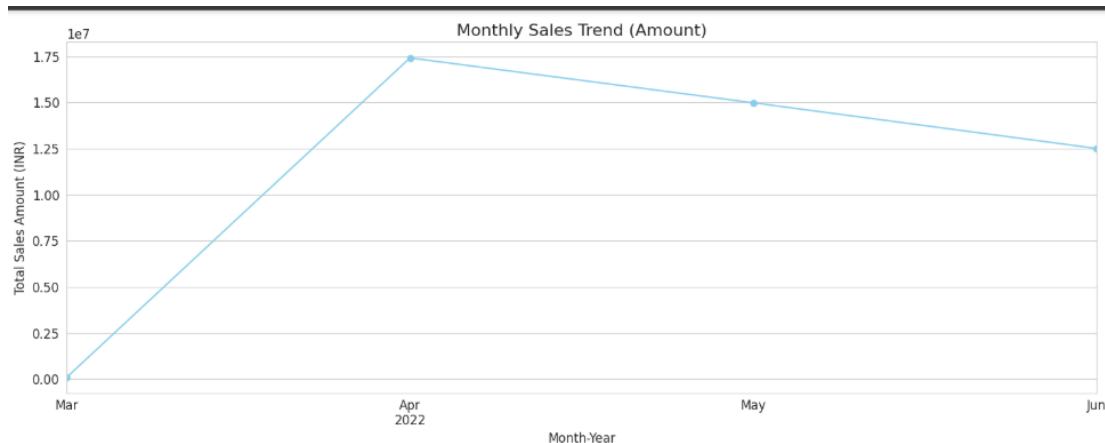
**Monthly Quantity Sold Trend (Line Chart):**

Insight: Both the total sales amount and quantity sold demonstrate noticeable fluctuations throughout the observed months. While a general downward trend in sales amount is observed towards the end of the period, the quantity sold exhibits intermittent peaks. These patterns suggest potential seasonal influences or responsiveness to specific marketing campaigns/promotions. Further granular analysis by specific dates could help correlate these peaks and troughs with actual events.

Daily & Weekly Sales Patterns

A deeper dive into daily and weekly patterns reveals consumer behavior within shorter timeframes.

Daily Sales Trend (Amount) (Line Chart):
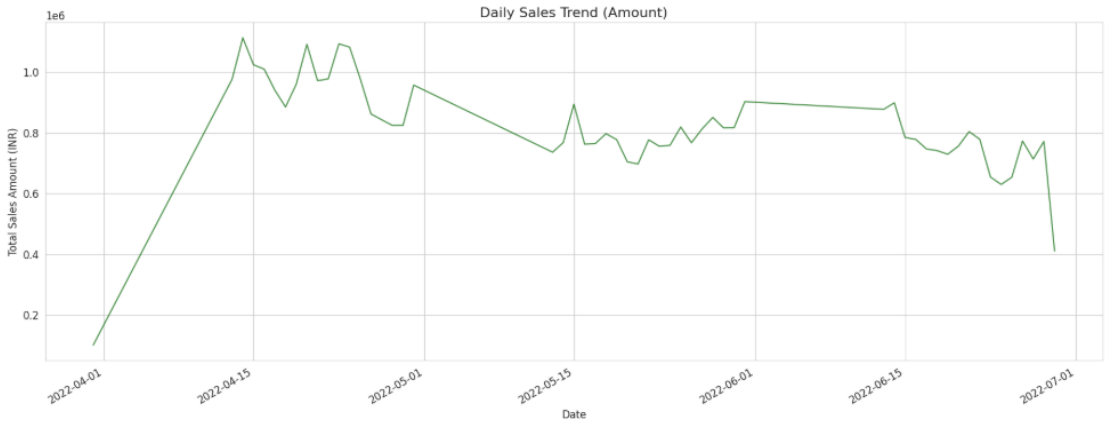
Insight: The daily sales figures show significant volatility, indicating that sales are not uniformly distributed throughout the month. This granular view can help in identifying specific high-traffic days.

## Total Sales Amount by Day of Week (Bar Chart):
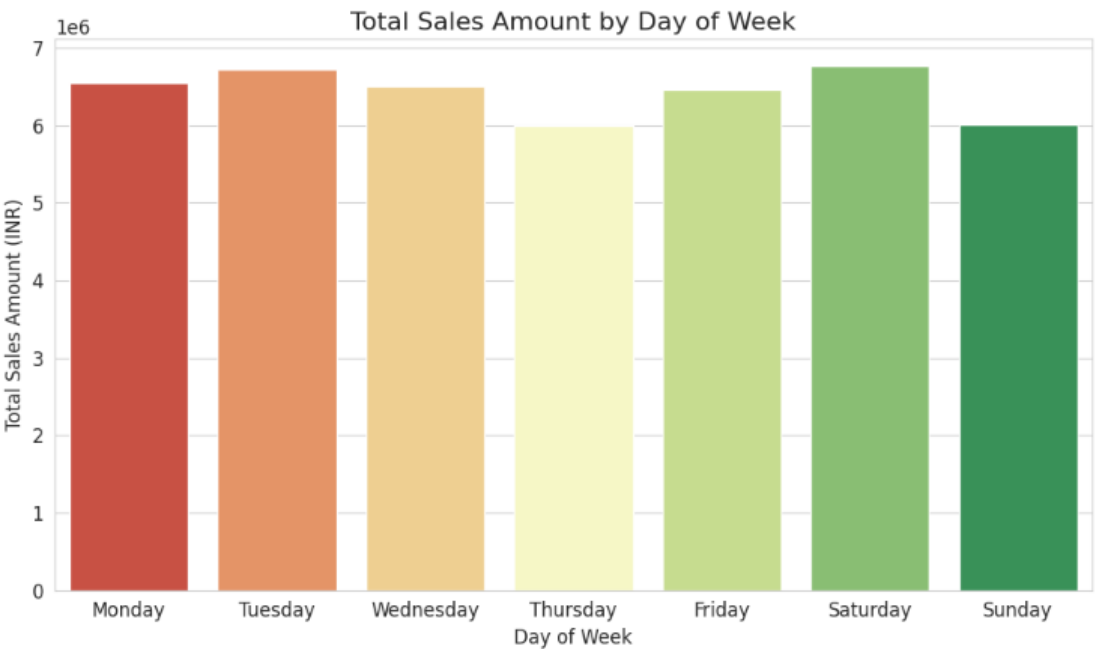
Insight: Sales amounts and quantities are relatively consistent across weekdays (Monday to Friday). Weekends, particularly Sunday, tend to exhibit slightly lower sales volumes. This is a common pattern in e-commerce, where consumer online activity might vary during leisure days. This information is valuable for scheduling promotional activities.

Sales by Ship Service Level

Understanding the preference for shipping service levels can inform logistics and customer expectations.

Sales Amount by Ship Service Level (Bar Chart):
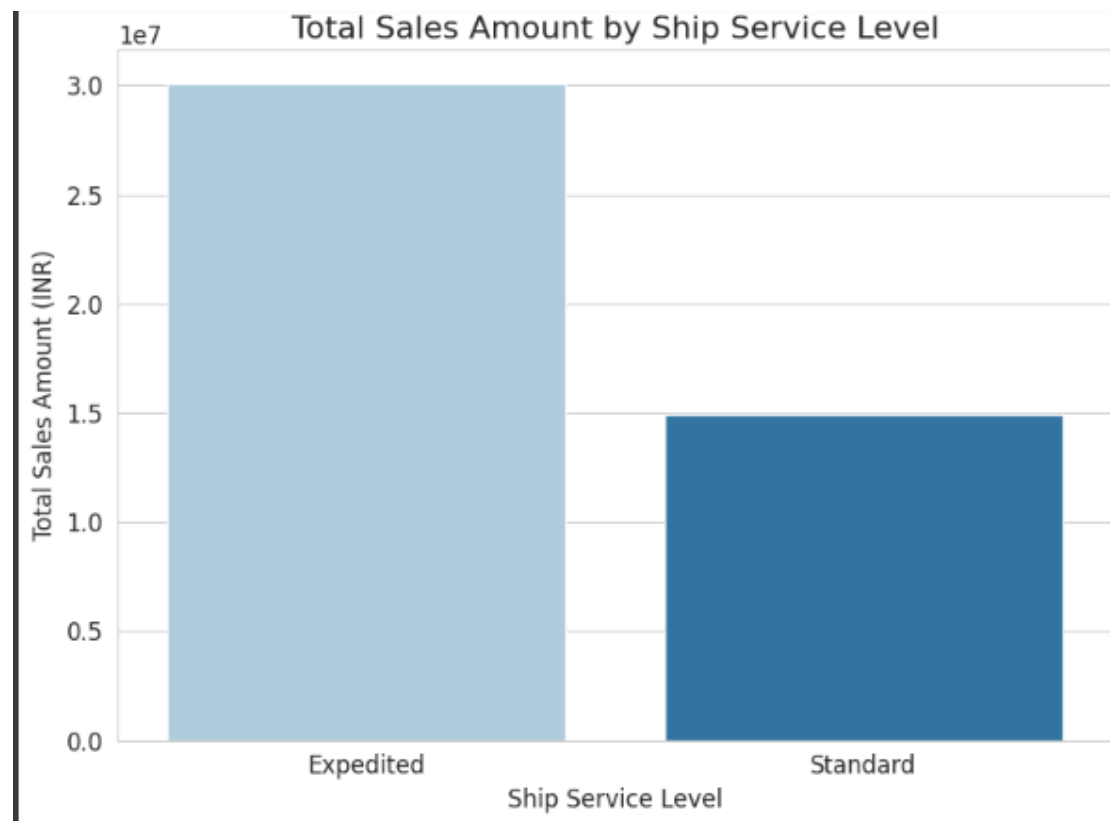
Insight: Expedited shipping accounts for the majority of both sales amount and quantity sold. This strongly suggests that customers either prefer faster delivery options or that higher-value and frequently purchased items are predominantly shipped via

expedited methods. This preference highlights the importance of efficient and quick delivery services.
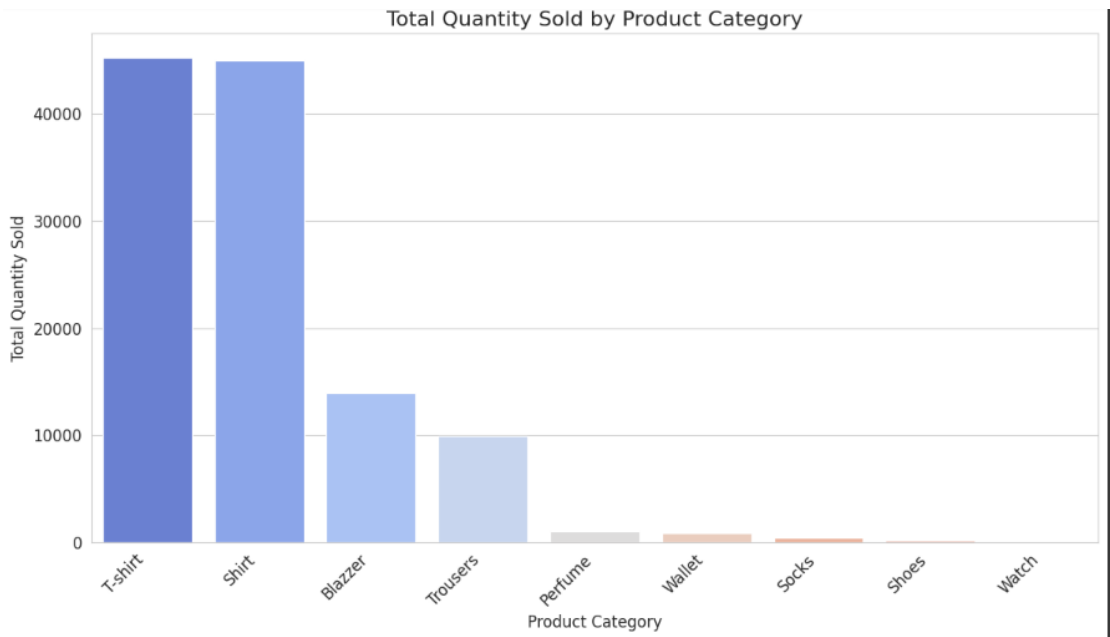
Product Analysis

Understanding the performance of different product categories and sizes is crucial for inventory management, product development, and marketing strategies.

Product Category Popularity

**Total Sales Amount by Product Category (Bar Chart):**

Total Quantity Sold by Product Category (Bar Chart):
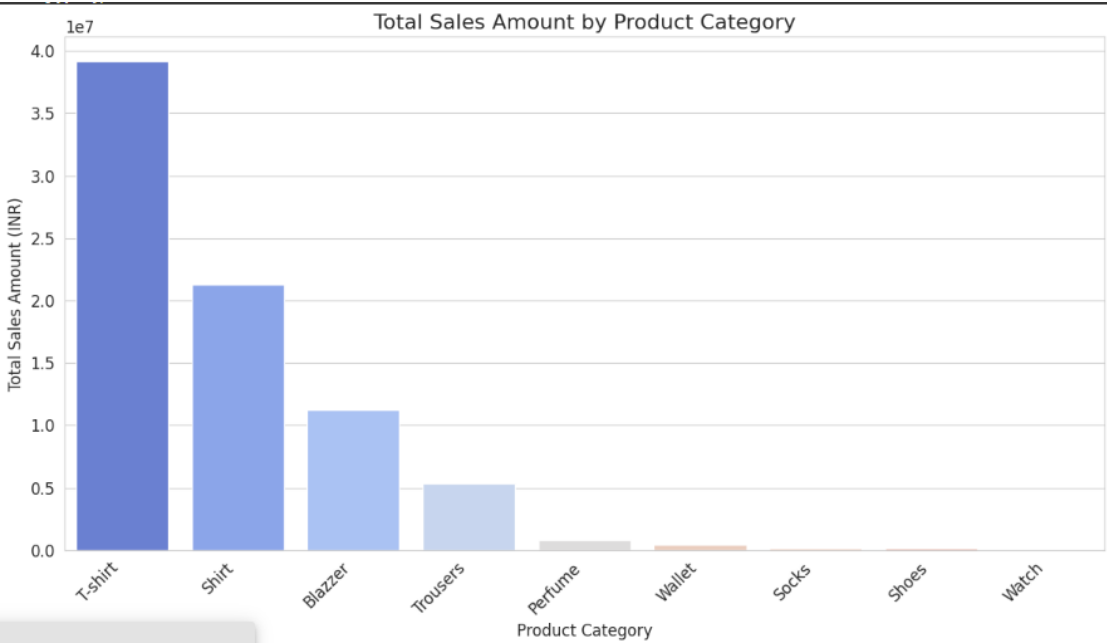
Total Quantity Sold by Product Category

Insight: T-shirt overwhelmingly dominates both in terms of total sales amount and quantity sold, followed by Shirt and Kurta. This confirms that apparel items are the core business and primary revenue drivers.

Product Size Distribution

Total Quantity Sold by Product Size (Bar Chart):

Total Quantity Sold by Product Size
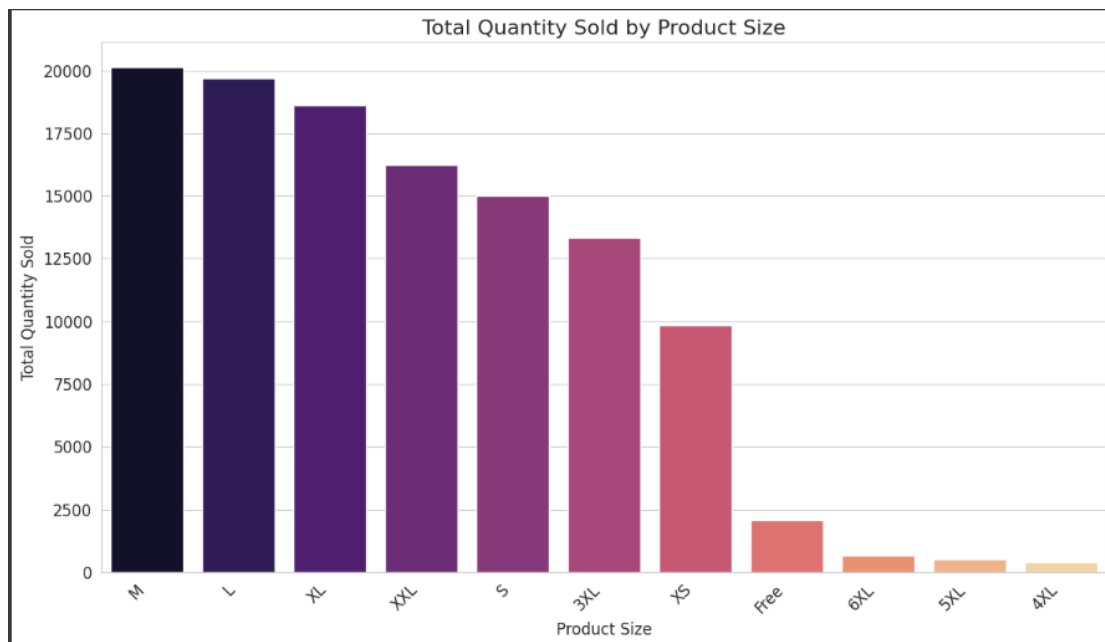
Insight: M (Medium) and L (Large) are the most popular sizes, aligning with general clothing market demands. XL and S also show considerable demand. This information is vital for maintaining optimal stock levels for each size.
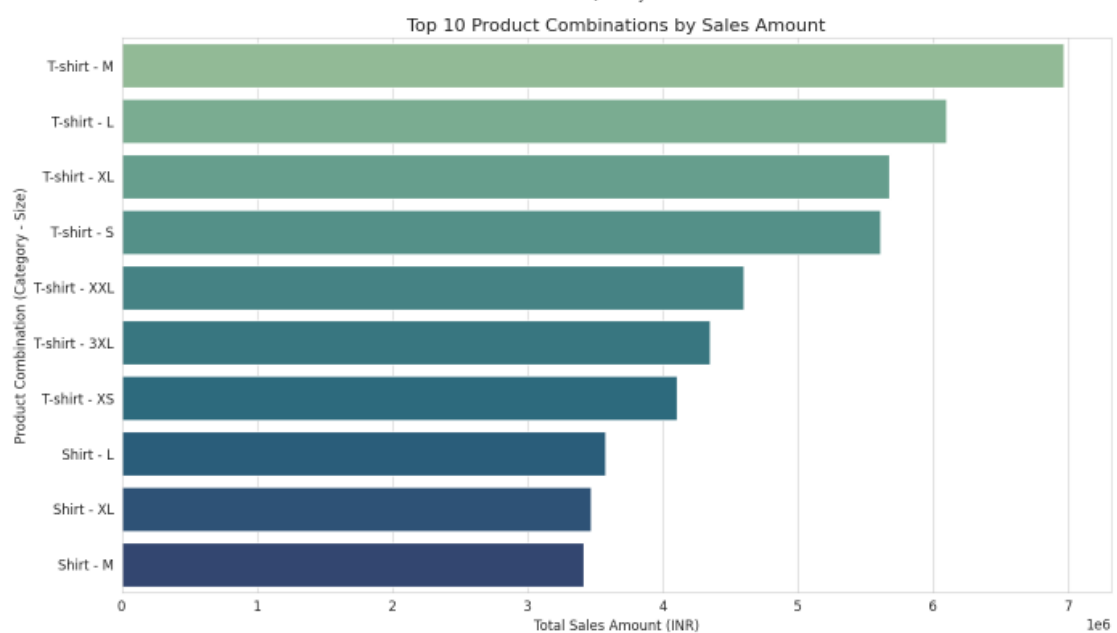
Top Product Combinations

Analyzing specific category-size combinations provides a granular view of consumer preferences.

Top 10 Product Combinations by Quantity Sold (Bar Chart):

Top 10 Product Combinations by Quantity Sold

Top 10 Product Combinations by Sales Amount (Bar Chart):
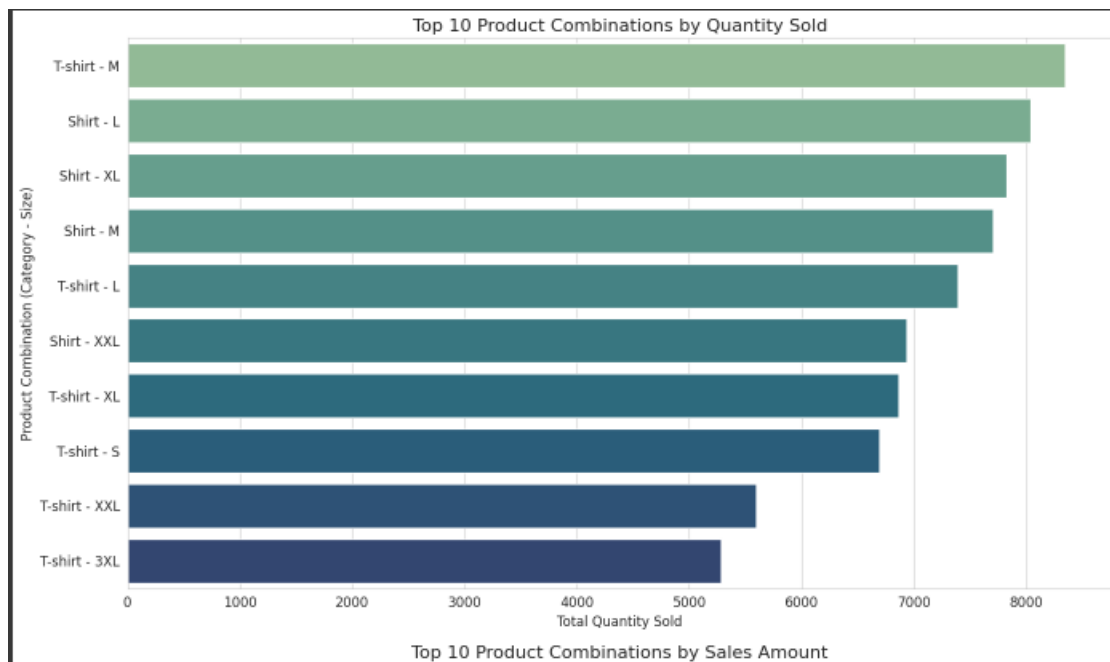
Top 10 Product Combinations by Sales Amount

Insight: T-shirt - M, T-shirt - L, and T-shirt - XL consistently appear as the top-performing product combinations. This identifies highly demanded stock-keeping units (SKUs) that require meticulous inventory planning.

Average Price per Item by Category

Examining the average price per item across categories offers insights into revenue potential beyond just volume.

Average Price Per Item by Product Category (Bar Chart):

Average Price Per Item by Product Category

Insight: While T-shirt and Shirt drive high volume due to their affordability, categories like Blazzer, Western Dress, and Ethnic Dress command significantly higher average prices per item. This suggests that these higher-value products could contribute disproportionately to profit margins and should be considered for strategic promotion to increase overall revenue per transaction.

Fulfillment Analysis

Investigating fulfillment methods is critical for assessing logistical efficiency, identifying bottlenecks, and improving delivery reliability.

Fulfillment Method Distribution

Distribution of Fulfillment Methods (Bar Chart):

Distribution of Fulfillment Methods
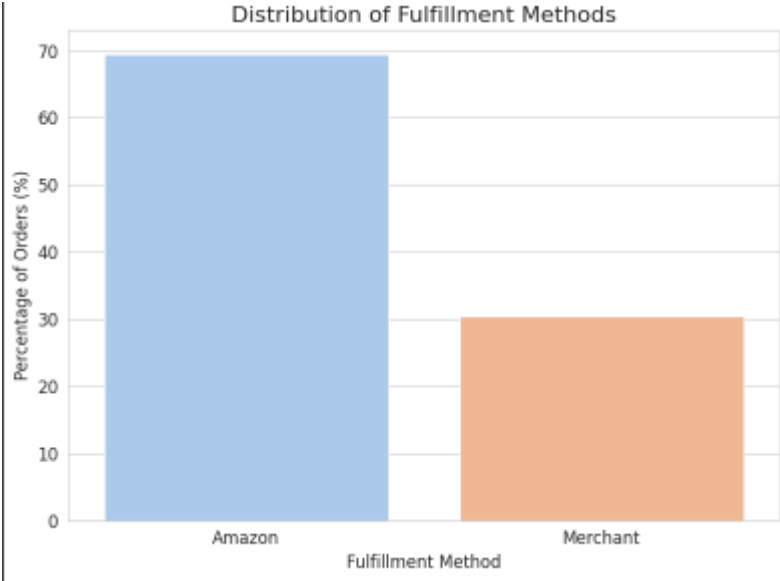
> **Insight:** Amazon is the predominant fulfillment method, handling approximately 70% of all orders. Merchant accounts for the remaining volume. This indicates a significant reliance on Amazon's established logistics infrastructure.

Fulfillment Effectiveness by Order Status

Comparing order statuses across fulfillment methods provides direct insights into their performance.

Order Status Breakdown by Fulfillment Method (Stacked Bar Chart):

Order Status Breakdown by Fulfillment Method
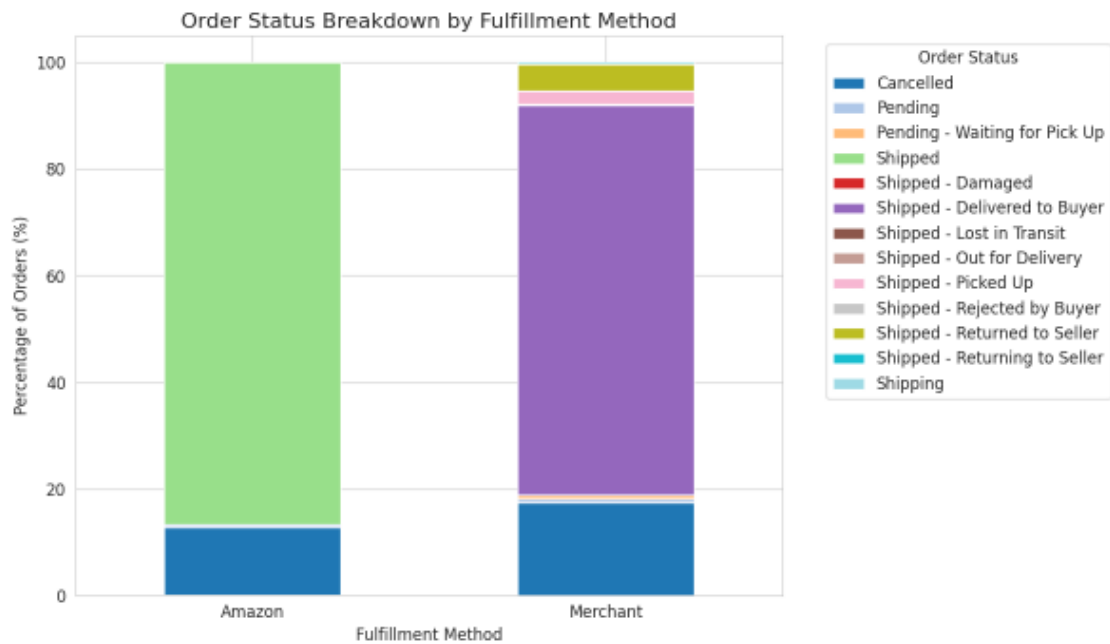
Insight: Amazon-fulfilled orders exhibit a substantially higher proportion of Shipped - Delivered to Buyer orders and a lower percentage of Cancelled orders compared to merchant-fulfilled orders. This clearly indicates that Amazon's fulfillment services are more efficient in processing and delivering orders, leading to better successful delivery rates and fewer cancellations.

Cancellation and Return Patterns by Product

Understanding which products contribute most to cancellations and returns can guide product improvement efforts.

Top 10 Products by Quantity of Cancelled/Returned Items (Bar Chart):

Top 10 Products by Quantity of Cancelled/Returned Items
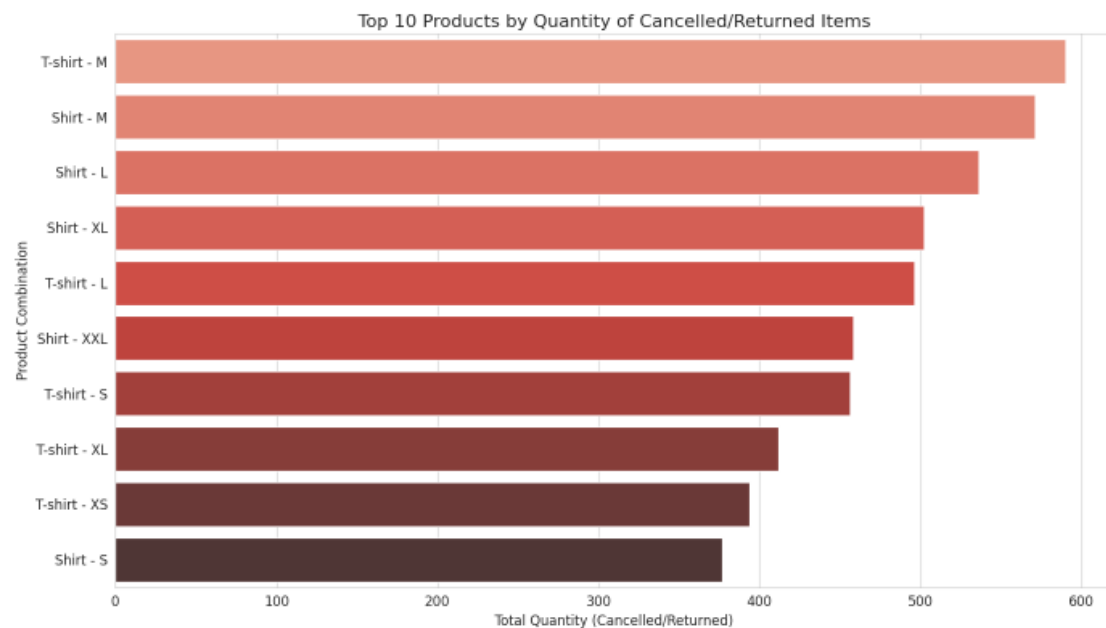
Insight: While T-shirt and Shirt in popular sizes (M, L, XL, S) appear most frequently in cancelled or returned orders, this largely correlates with their high sales volume. However, this trend warrants further investigation into specific reasons beyond just high quantity, such as common fit complaints, quality issues for particular designs, or discrepancies in product descriptions.

Fulfillment Method and Ship Service Level

The relationship between fulfillment methods and chosen shipping service levels reveals operational and customer experience dynamics.

Ship Service Level Breakdown by Fulfillment Method (Stacked Bar Chart):

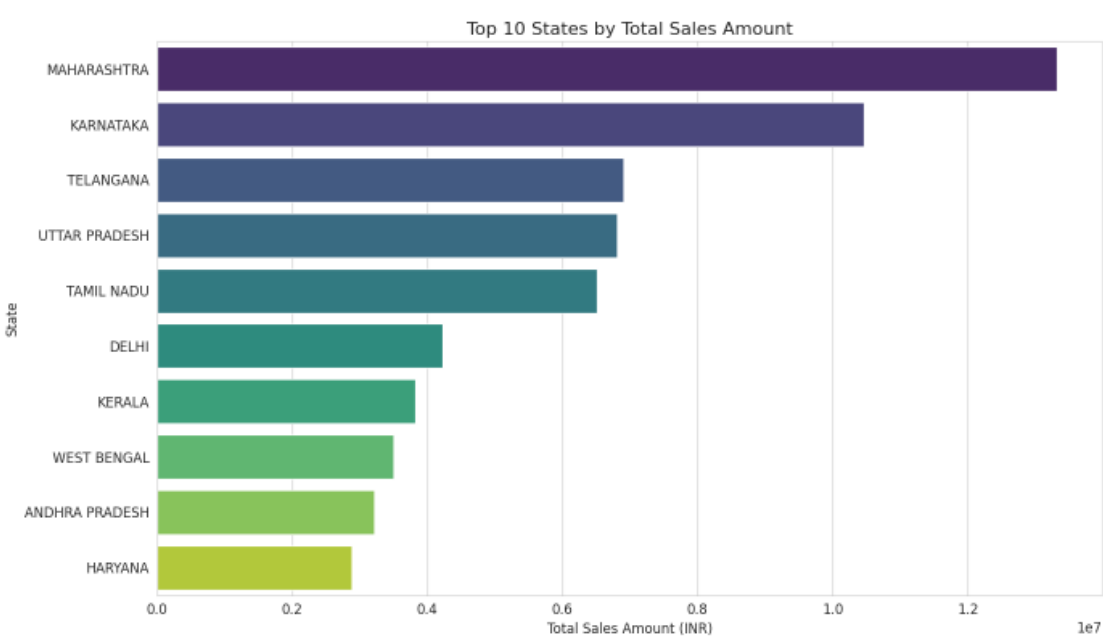Ship Service Level Breakdown by Fulfillment Method
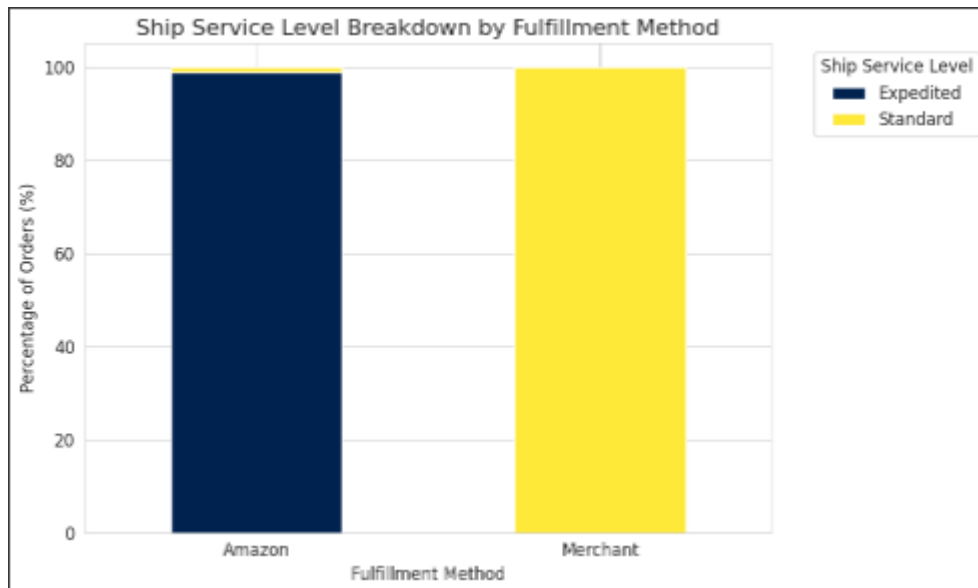
Insight: Amazon-fulfilled orders show a significantly higher utilization of Expedited shipping. This aligns with Amazon's commitment to faster deliveries and suggests that customers associate Amazon fulfillment with speed. Conversely, merchant-fulfilled orders rely more predominantly on Standard shipping. This disparity in service levels could contribute to differences in customer expectations and satisfaction regarding delivery times, potentially impacting cancellation rates for merchant-fulfilled orders.
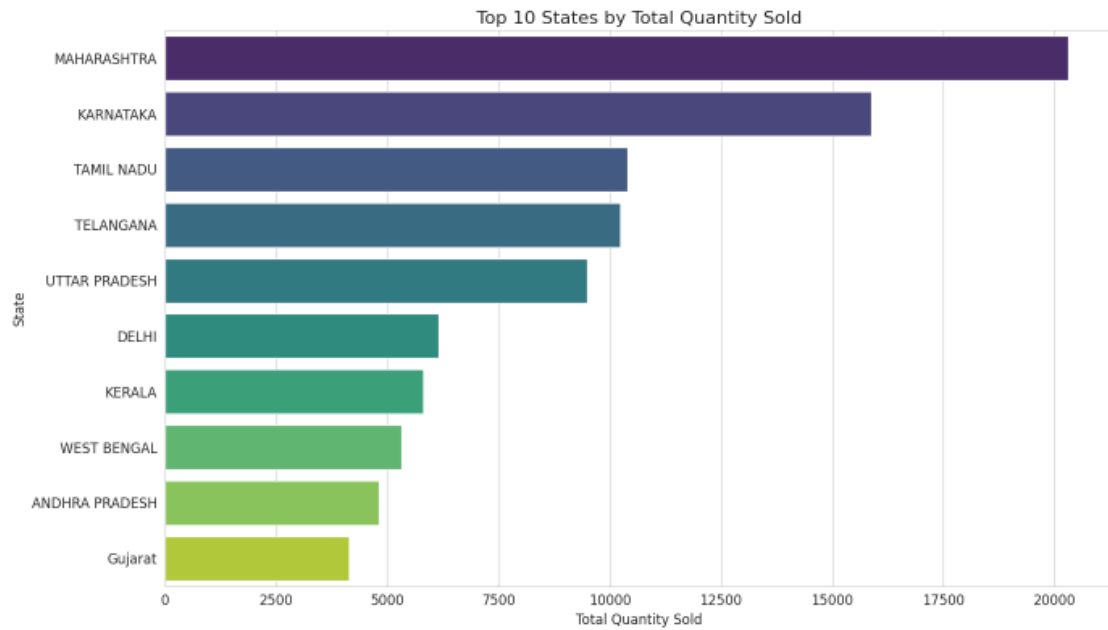
Geographical Analysis & Customer Segmentation

Analyzing sales geographically helps in understanding market concentration, identifying growth opportunities, and segmenting customers based on location.

Top States by Sales

### Top 10 States by Total Sales Amount (Bar Chart):

Top 10 States by Total Quantity Sold (Bar Chart):

Top 10 States by Total Quantity Sold

Insight: MAHARASHTRA, KARNATAKA, UTTAR PRADESH, and TAMIL NADU consistently emerge as the top-performing states in terms of both total sales amount and quantity sold. These states represent the core and most lucrative geographical markets for the business.

Top Cities by Sales

Top 15 Cities by Total Sales Amount (Bar Chart):

Top 15 Cities by Total Sales Amount
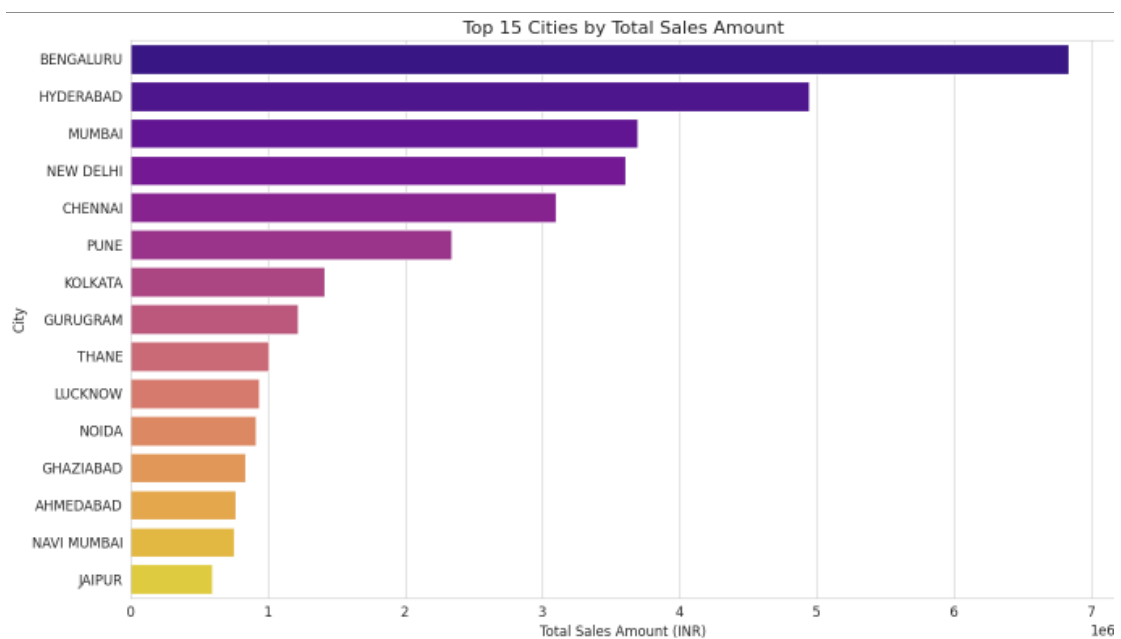
Top 15 Cities by Total Quantity Sold (Bar Chart):
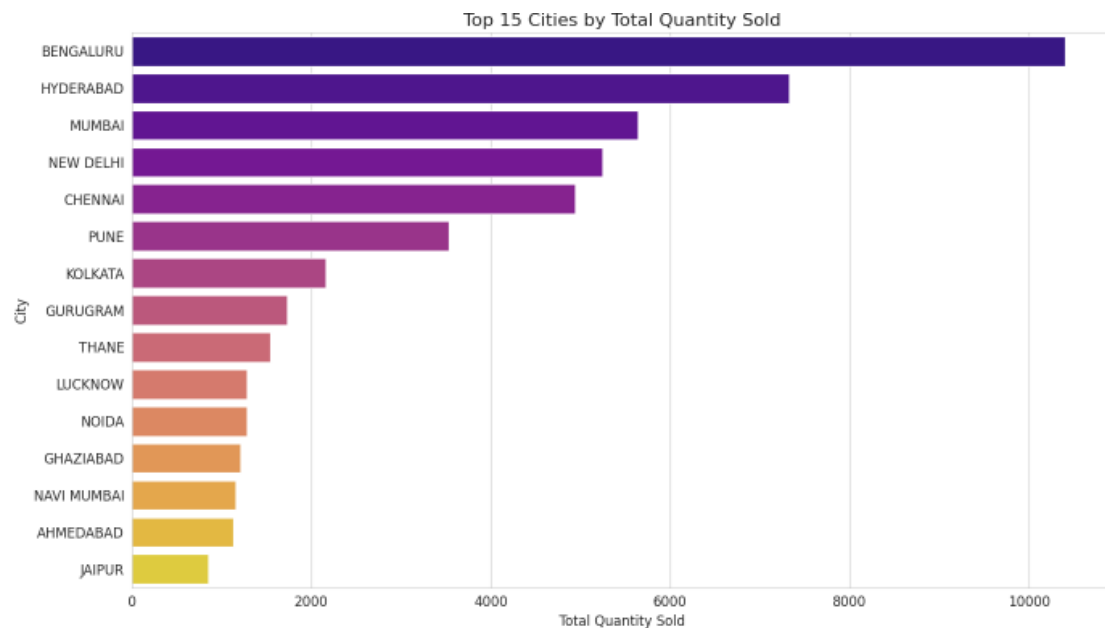
Insight: Major metropolitan areas such as BENGALURU, MUMBAI, HYDERABAD, and NEW DELHI are the leading cities contributing to sales. This is a common and expected pattern, reflecting these cities' status as significant commercial and population centers.

Average Order Value by Location

Examining Average Order Value (AOV) by state reveals segments with higher spending potential.

Top 10 States by Average Order Value (AOV) (Bar Chart):

Top 10 States by Average Order Value
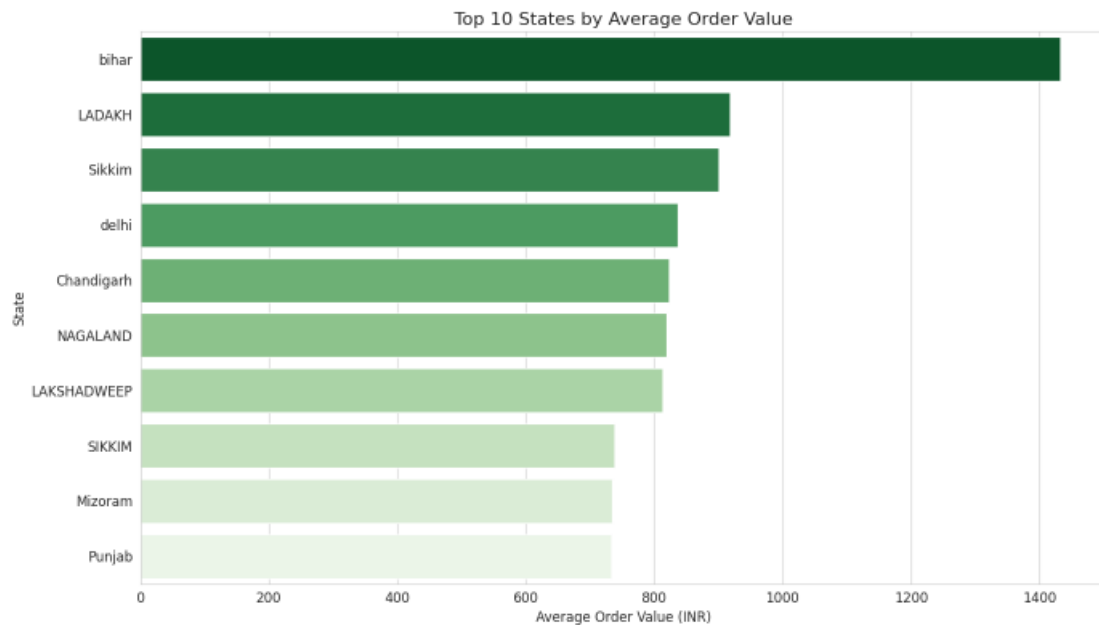
Insight: While the major states drive overall sales volume, some smaller states like LADAKH, MEGHALAYA, and SIKKIM demonstrate a notably higher Average Order Value. This indicates that customers in these regions, though potentially fewer in number, tend to place higher-value orders. These areas represent potential niche markets for targeted high-value product promotions.

B2B vs. B2C Sales Distribution

Understanding the business model focus is crucial for strategic alignment.

Total Sales Amount: B2B vs B2C (Bar Chart):

Total Sales Amount: B2B vs B2C
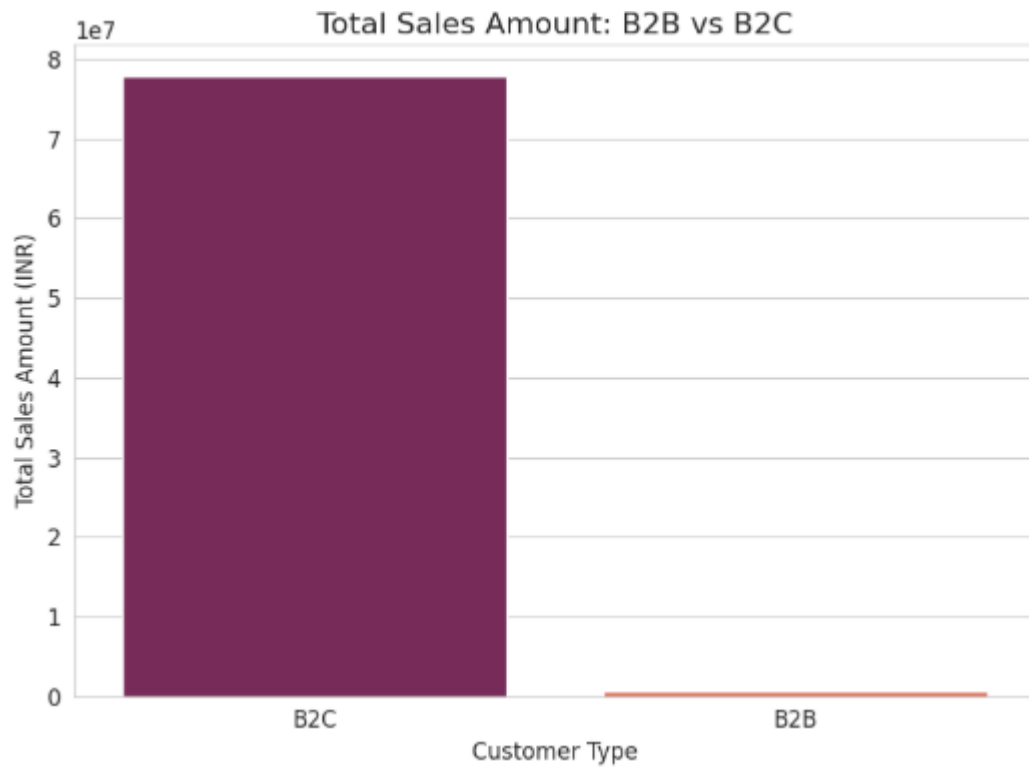
Insight: The analysis clearly shows that the vast majority of sales, both in terms of amount and quantity, are derived from B2C (Business-to-Consumer) transactions. B2B (Business-to-Business) sales constitute a very small and negligible fraction of the total. This confirms that the business is primarily consumer-focused and its strategies should be aligned accordingly.

Key Findings

This comprehensive analysis has yielded several critical insights into the Amazon sales operations:

- **Apparel Dominance & Price Segmentation:** T-shirts and Shirts are the primary volume drivers, with popular sizes (M, L) leading. However, categories like Blazzers and Dresses command significantly higher average order values, indicating higher-margin opportunities.

- **Persistent Cancellation Challenge:** A substantial 20% of all orders are cancelled. While some cancellations are inherent to e-commerce, this rate is significant and warrants a deep investigation into underlying causes across products and fulfillment methods.

- **Superior Amazon Fulfillment:** Amazon's fulfillment services demonstrate higher efficiency, resulting in a greater proportion of successfully delivered orders and fewer cancellations compared to merchant-fulfilled orders.

- **Customer Preference for Speed:** There is a clear customer preference for expedited shipping, particularly evident in Amazon-fulfilled orders.

- **Concentrated Geographical Demand:** Sales are highly concentrated in major Indian states (Maharashtra, Karnataka, Uttar Pradesh, Tamil Nadu) and their metropolitan cities (Bengaluru, Mumbai, Hyderabad, New Delhi).

- **Emerging High-Value Niches:** Some geographically smaller states exhibit a higher Average Order Value, pointing towards potential high-value niche markets.

- **B2C Dominance:** The business is overwhelmingly focused on Business-to-Consumer sales, with minimal B2B activity.

Recommendations

Based on the key findings, the following actionable recommendations are proposed to optimize sales strategies, improve customer satisfaction, and enhance overall business performance:

Product & Inventory Optimization

- **Strategic Stocking for Core Products:** Ensure robust and agile inventory management for high-volume apparel categories like T-shirts, Shirts, and Kurtas, specifically focusing on Medium, Large, and XL sizes, to prevent stockouts and meet consistent demand.
- **Targeted Promotion of High-AOV Categories:** Develop and execute targeted marketing campaigns for higher Average Order Value categories such as Blazzer, Western Dress, and Ethnic Dress. This could include premium branding, exclusive collections, or bundles to enhance revenue per transaction.

Fulfillment & Operational Efficiency

- **In-depth Cancellation Root Cause Analysis:** Conduct a detailed investigation into the reasons behind the 20% order cancellation rate. This should involve analyzing specific products, customer feedback (if available), payment issues, address inaccuracies, and lead times to pinpoint and address systemic issues.
- **Enhance Merchant Fulfillment Performance:** Collaborate closely with merchant partners to improve their fulfillment accuracy and speed. Provide clear guidelines on inventory management, order processing, and adherence to service level agreements (SLAs). Consider performance incentives or penalties to drive improvements.

- **Maximize Amazon FBA Utilization:** Given its proven efficiency and higher delivery success rates, explore strategies to further expand the use of Amazon's Fulfillment by Amazon (FBA) services, especially for high-volume and high-demand products, to leverage Amazon's robust logistics network.

Targeted Geographical Strategies

- **Reinforce Core Market Penetration:** Continue to invest in marketing, promotions, and potentially localized logistics infrastructure in high-volume states (Maharashtra, Karnataka, Uttar Pradesh, Tamil Nadu) and major cities (Bengaluru, Mumbai, Hyderabad, New Delhi) to solidify market share.
- **Develop Niche High-Value Markets:** Implement targeted strategies for states/cities exhibiting a higher Average Order Value despite lower overall volume (e.g., Ladakh, Meghalaya). These could involve premium product offerings, personalized outreach, or specialized delivery options to cater to this segment.

Sales & Customer Experience Enhancement

- **Strategic Promotional Scheduling:** Leverage the consistent weekday sales patterns by planning daily deals or flash sales during peak weekday hours. Additionally, explore specific weekend-focused promotions or campaigns to stimulate sales on traditionally slower days (e.g., Sunday).
- **Highlight & Expand Expedited Shipping:** Clearly communicate and emphasize the availability and benefits of expedited shipping options to customers. Explore opportunities to expand expedited delivery capabilities, particularly for merchant-fulfilled orders, to align with customer preference for faster service and potentially reduce cancellations related to delivery times.

Conclusion

This analysis of the Amazon Sales Report has provided a robust, data-driven foundation for strategic decision-making. By understanding core product performance, identifying operational efficiencies in fulfillment, and recognizing key geographical market dynamics, the business is well-positioned to optimize its e-commerce operations. The recommendations provided offer actionable pathways to address current challenges, capitalize on strengths, and unlock new avenues for revenue growth and enhanced customer satisfaction. Continuous monitoring of these metrics and adaptation to evolving market trends will be crucial for sustained success.

Github link : https://github.com/KOUSHIKB315/Amazon-Sales-Report-Data-Analysis_innobyte