Amazon Sales Analysis Report

Aaryan Jain
IIT Kanpur
Project ID: 78G0OL

June 28, 2025

1 Executive Summary

This report presents a comprehensive analysis of Amazon sales data. The objective is to identify product performance trends, fulfillment efficiency, geographic sales distribution, and basic order-based segmentation. Insights and strategic recommendations are provided to support business decisions.

2 1. Data Preparation

The dataset was cleaned and preprocessed by:

- Converting `Order Date` and `Amount` columns to appropriate formats.

- Removing rows with missing or invalid data.

- Creating derived columns like `Order Month` for time-based analysis.

3 2. Sales Overview

The monthly sales trend showed a peak in April 2022, followed by slight decline in subsequent months.
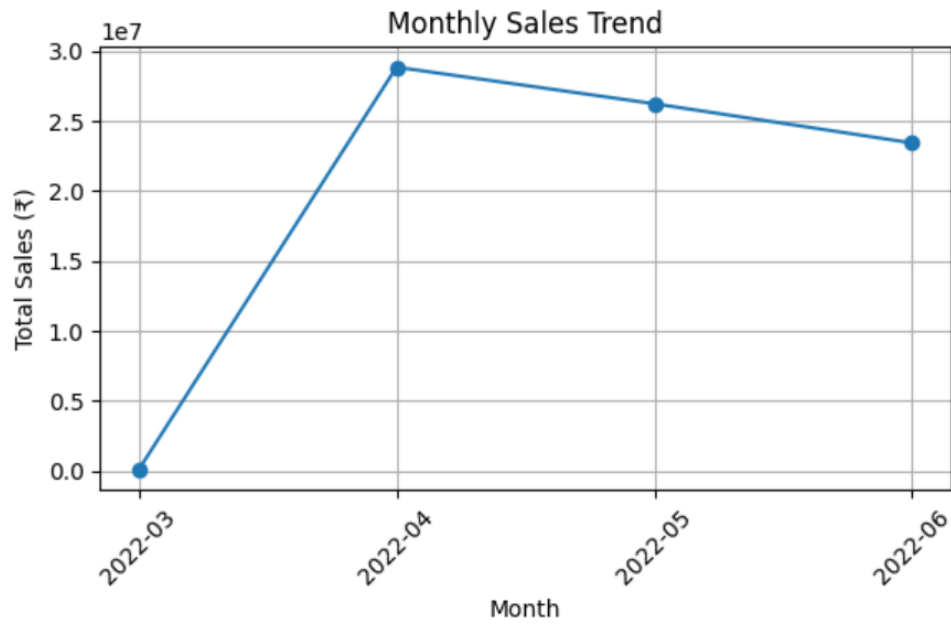
Figure 1: Monthly Sales Trend

4   3. Product Analysis

T-shirts generated the highest revenue, followed by shirts and blazers.
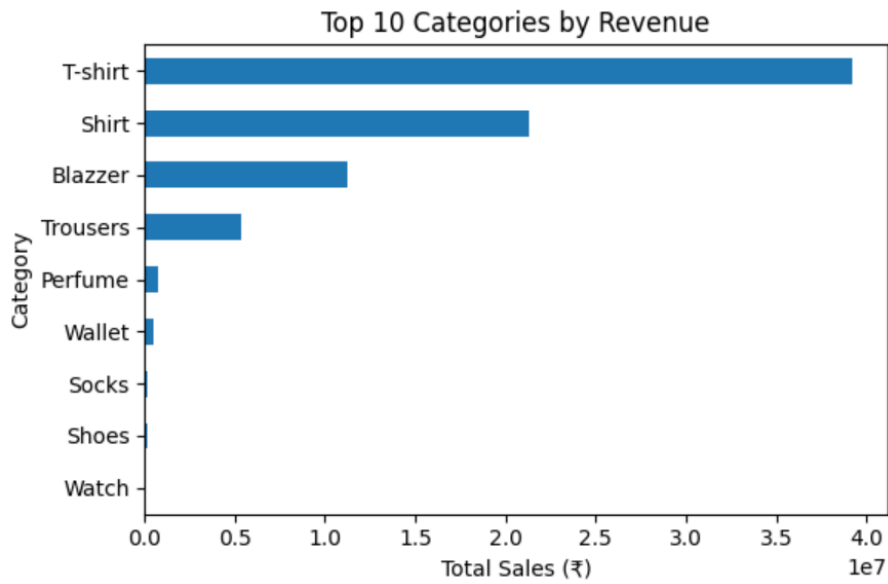

Figure 2: Top 10 Product Categories by Revenue
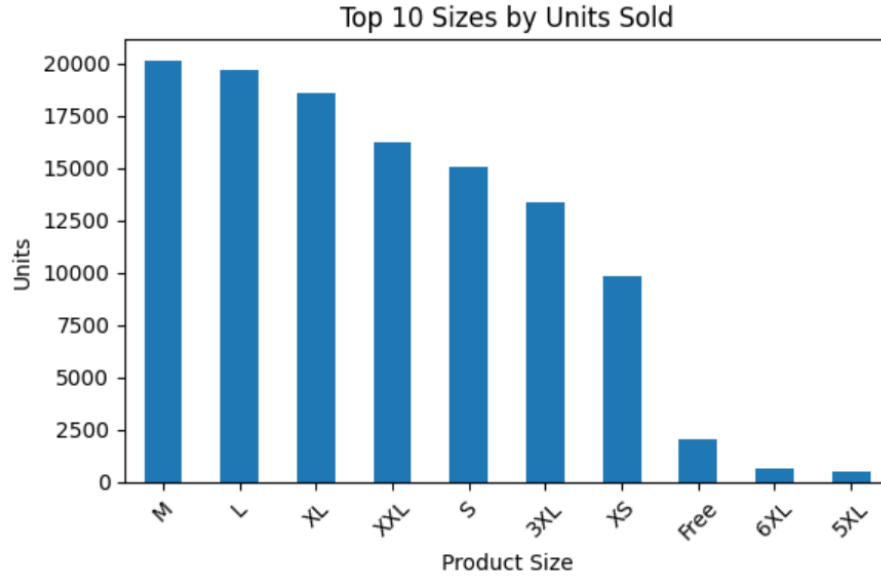
Sizes M, L, and XL dominated in unit sales.

Figure 3: Top 10 Product Sizes by Units Sold

5 4. Fulfillment Analysis

Amazon (FBA) and Merchant Fulfillment were compared:

Fulfillment	Orders	Revenue ()	AOV ()	Cancel Rate
Amazon	78,211	54,327,540.00	694.63	6.7%
Merchant	34,676	24,262,630.25	699.70	13.7%

Table 1: Fulfillment Method Metrics

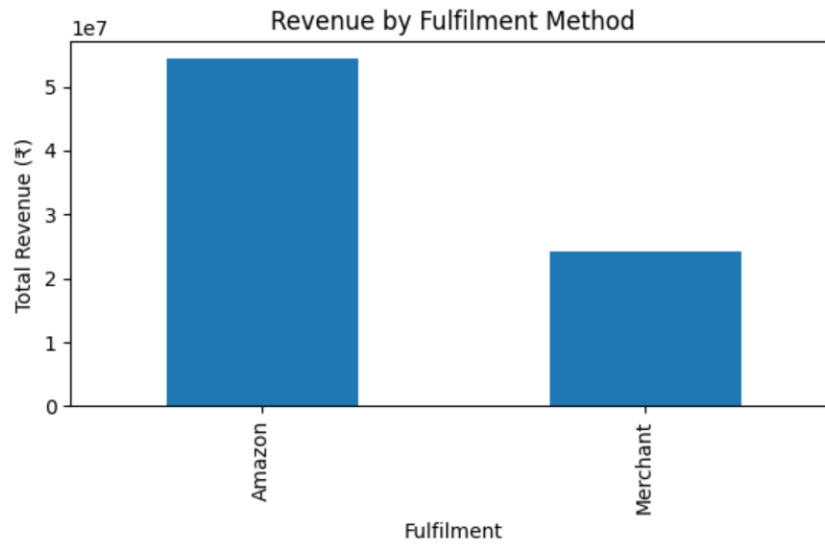Amazon Fulfillment yielded higher reliability and revenue.
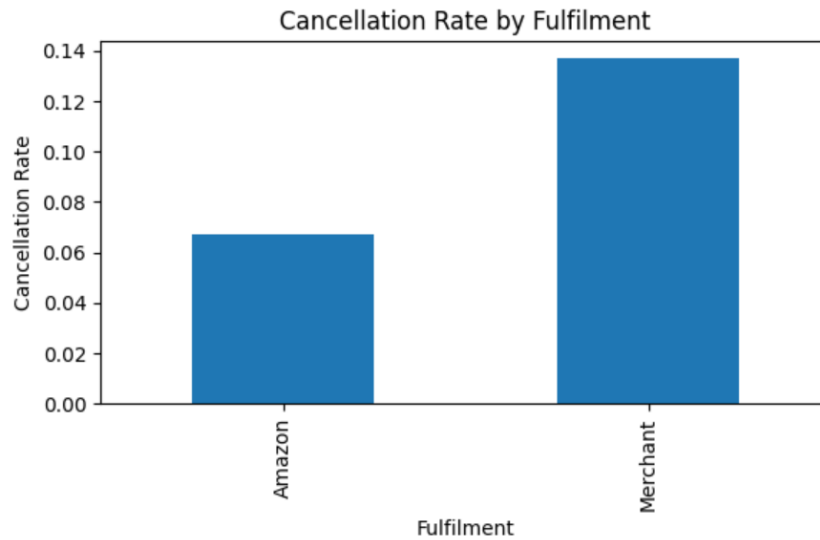

Figure 4: Revenue by Fulfillment Method

Figure 5: Cancellation Rate by Fulfillment Method

6   5. Geographic Analysis

Top cities and states by sales were identified. A bar-chart visualizes the state-wise sales distribution.
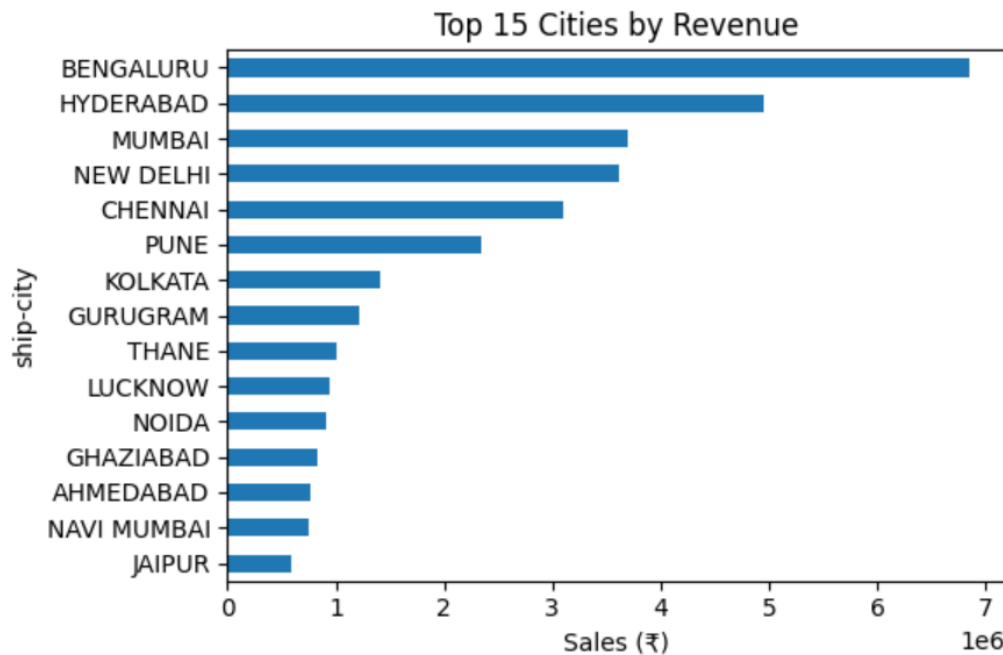

Figure 6: State-wise Sales Bar-chart

7   6. RFM-based Order Analysis

Due to lack of Customer ID, Order ID was used as a proxy to analyze order-level recency and value. High-value recent orders were flagged for marketing potential.
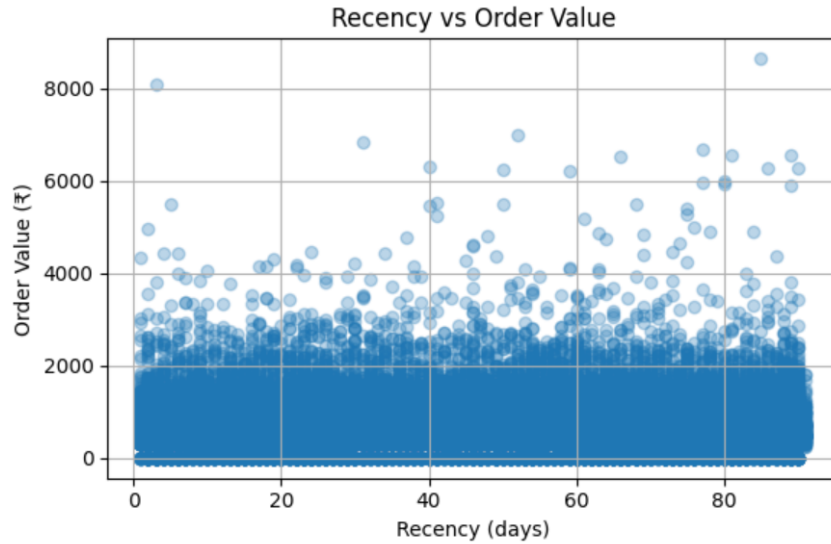
Figure 7: Recency vs Order Value Scatter Plot

8   7. Business Recommendations

- Shift high-selling products to Amazon FBA to minimize cancellations.

- Address Merchant fulfillment issues contributing to higher cancellation rates.

- Maintain higher stock of M, L, and XL sizes.

- Engage recent high-order-value customers with promotional offers.

- Bundle related products to increase AOV (e.g., shirts + blazers).

- Investigate causes behind April sales spike to replicate in future campaigns.

9   8. Limitations and Future Scope

- Lack of Customer ID limited individual customer analysis.

- Address-based proxy segmentation may be explored.

- Future integration with CRM data is recommended for deeper insights.