

Amazon Sales Analysis Report

Dataset Name : Amazon Sales Report

Datasets: https://drive.google.com/file/d/1YrjYKtS1WHmINL6eafRsrDzrZaw2_WvX/view?usp=sharing

Problem Description:

The provided dataset contains information about sales transactions on Amazon, including details such as order ID, date, status, fulfilment method, sales channel, product category, size, quantity, amount, shipping details, and more. The objective is to conduct a comprehensive analysis of the data and extract actionable insights to support business decision-making.

Prepared by: **Atharva Mathe**

Tools: Python, SQL, Tableau

Key Objectives :

1. **Sales Overview:** Understand the overall sales performance, trends, and patterns over time.
2. **Product Analysis:** Analyze the distribution of product categories, sizes, and quantities sold to identify popular products.
3. **Fulfilment Analysis:** Investigate the fulfilment methods used and their effectiveness in delivering orders.
4. **Customer Segmentation:** Segment customers based on their buying behaviour, location, and other relevant factors.
5. **Geographical Analysis:** Explore the geographical distribution of sales, focusing on states and cities.
6. **Business Insights:** Provide actionable insights and recommendations based on the analysis to optimize sales strategies, improve customer satisfaction, and enhance overall business performance.

Dataset Overview & Methodology

Dataset Overview & Methodology

Dataset Overview

- The analysis is based on **Amazon sales transaction data** covering order-level details, time-based trends, product attributes, customer segmentation, fulfilment methods, and order outcomes.
- Key variables include sales metrics (**Amount, Quantity**), product details (**Category, Size**), operational fields (**Fulfilment, Ship-Service-Level, Courier Status, Fulfilled-By**), and order status indicators (**Delivered, Cancelled**).
- **Geographical dimensions** such as **Ship-City, Ship-State, Ship-Postal-Code, and Ship-Country** enable location-based performance and cancellation analysis.

Methodology

- **Python** was used for data cleaning, preprocessing, and exploratory data analysis to identify trends and patterns.
- **SQL** supported metric validation and structured aggregation for key KPIs.
- An **Interactive Tableau Dashboard** was developed to visualise insights and support data-driven business decision-making.

Strategic Objectives

This project aims to transform raw transactional data into actionable business intelligence across five key analytical dimensions:



Sales Performance

Comprehensive revenue analysis and temporal trend identification



Product Analysis

Category performance and inventory optimisation insights



Fulfilment Efficiency

Logistics evaluation and cancellation pattern analysis



Customer Segmentation

B2B versus B2C behavioural comparison



Geographic Distribution

Regional revenue patterns and market opportunities

Sales Overview: Performance & Trends

₹78.46 Cr

Total Revenue

1.28 Lakh

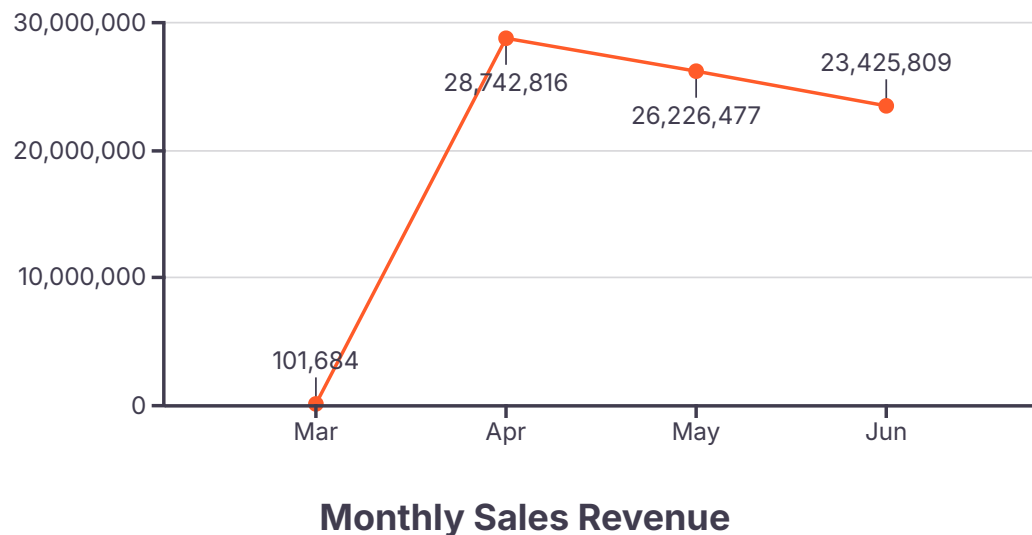
Total Orders

₹609.41

Average Order Value

14.22%

Cancellation Rate



◆ Insights

- Revenue and order volume **peaked in April**, indicating a strong seasonal or campaign-driven performance.
- Slight decline in May and June suggests **post-peak normalization** or reduced promotions.
- A **14.22% cancellation rate** is relatively high and represents a clear revenue leakage opportunity.

Product Analysis: Category Performance

♦ Best Sellers

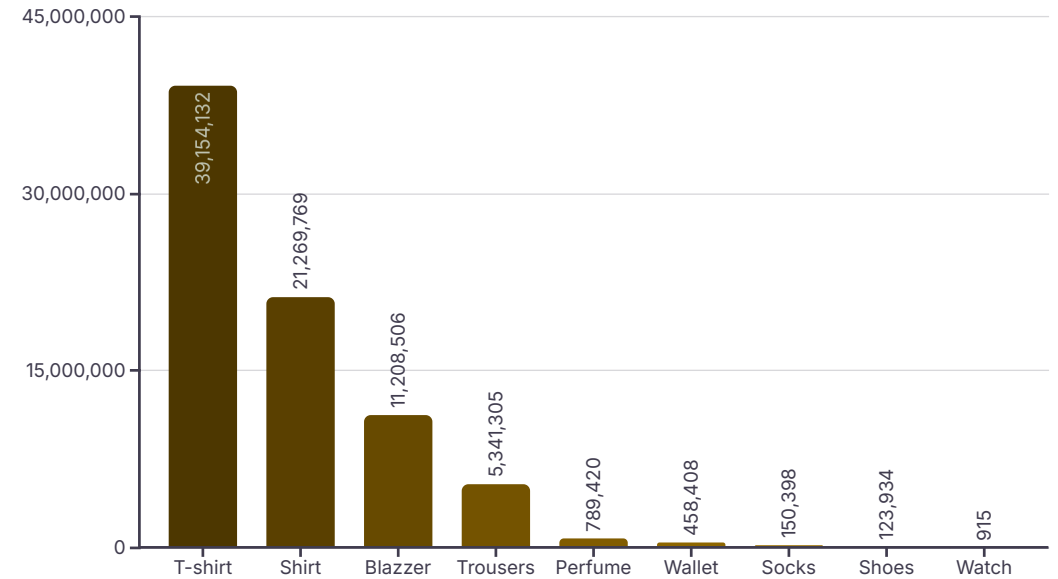
- **Best-Selling Product:** T-shirt (45,228 units)
- **Highest Revenue Product:** T-shirt – ₹39.15M

♦ Category Contribution

- Top 3 categories (T-shirt, Shirt, Blazer) generate **~90% of revenue**
- Long-tail categories (Wallet, Socks, Shoes, Watch) have minimal impact

♦ Size Insights

- **Most Popular Size:** M
- Sizes M, L, XL dominate across products



Sales By Product

45,228

T-shirt

Best-Selling Product Quantity

20,116

M - Size

Most Popular Size

₹3.91 Cr

T-shirt

Best-Selling Product Revenue

Fulfilment Analysis: Delivery & Logistics

Amazon Fulfilment

Market Share: 57.7% of total orders

Performance: Fastest delivery times with 17.47% cancellation rate

Customer Satisfaction: Highest ratings for reliability and tracking accuracy

Merchant Fulfilment

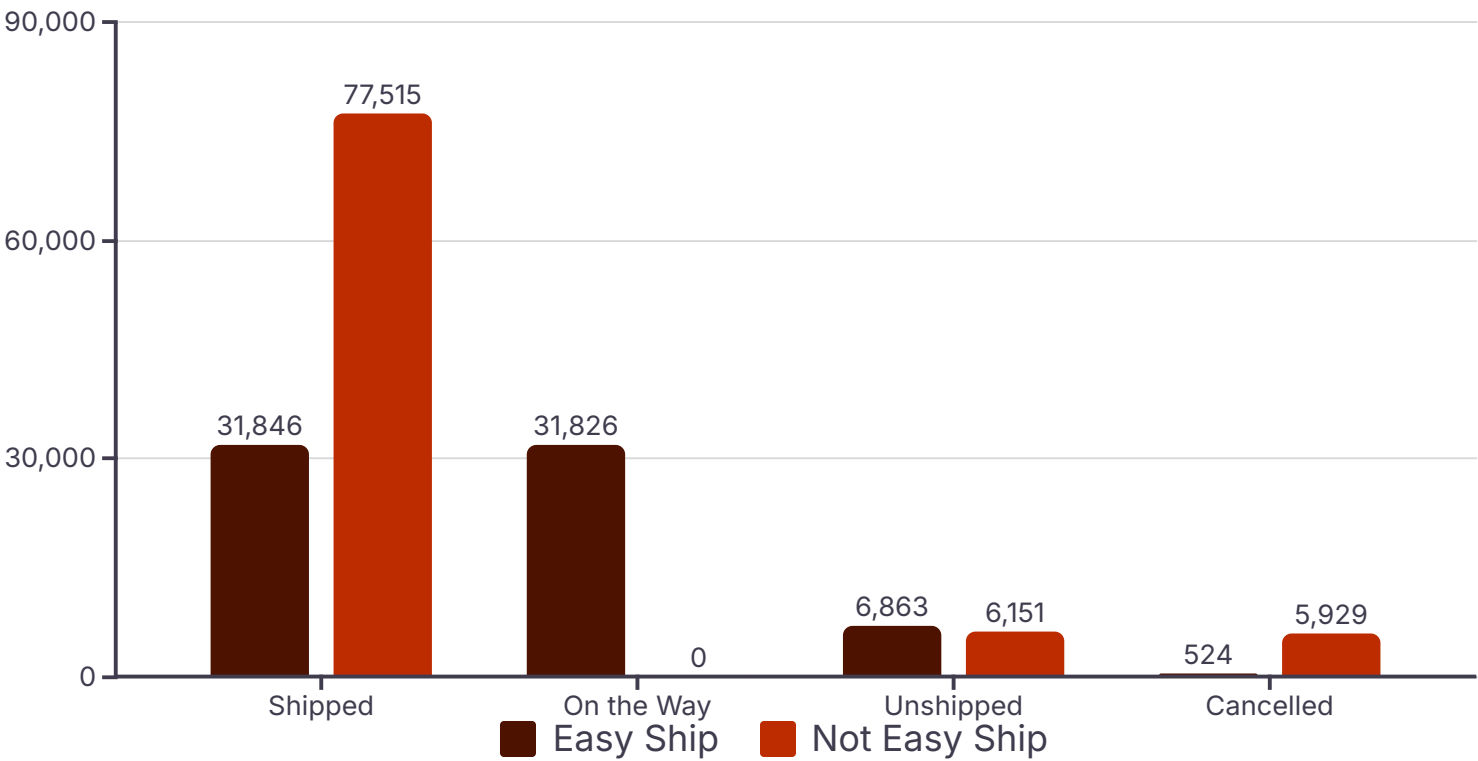
Market Share: 42.3% of total orders

Performance: Extended delivery windows with 12.79% cancellation rate

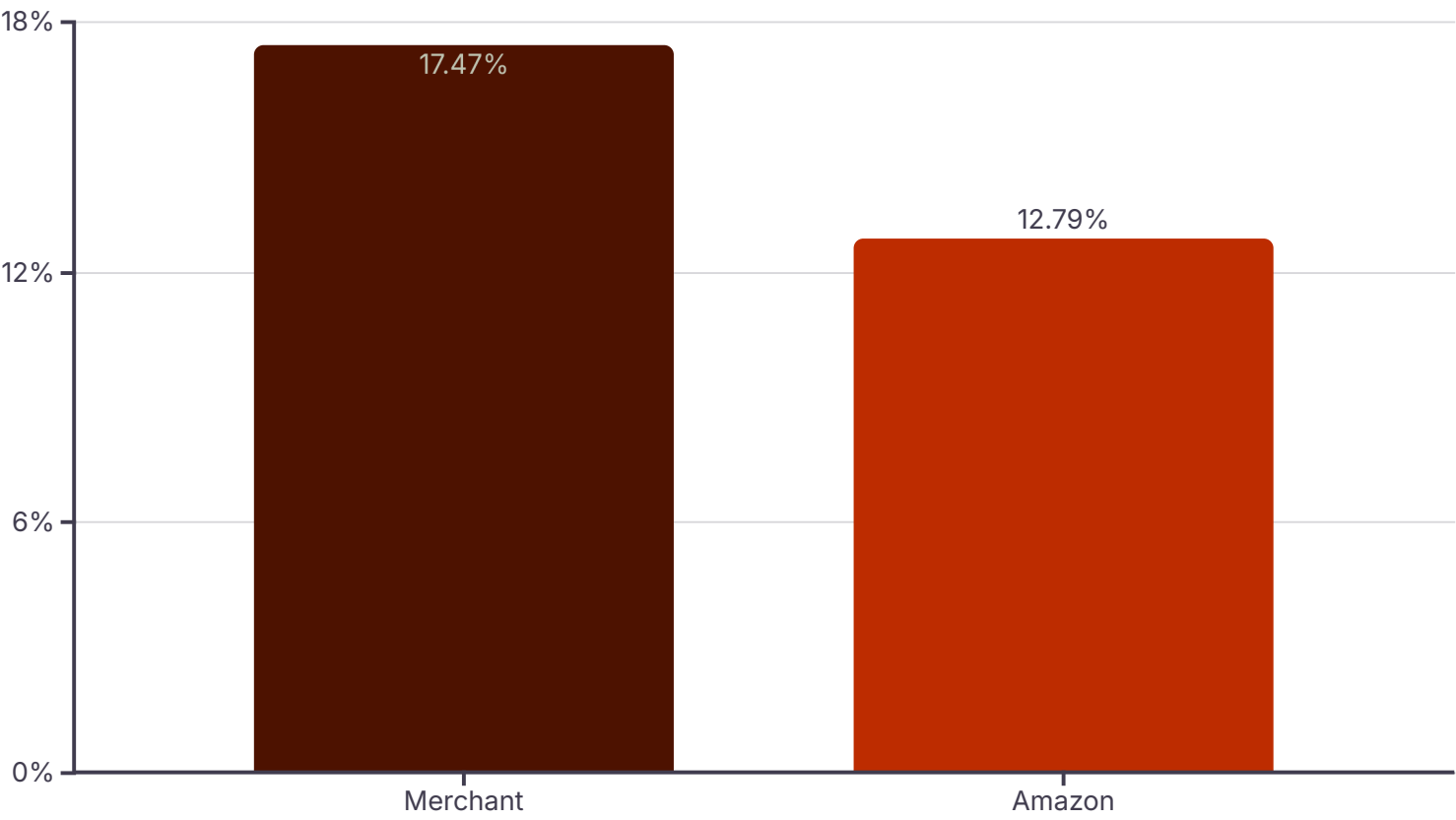
Challenge Areas: Inconsistent courier performance and delayed shipments

Business Insight:

- Despite **Amazon’s faster delivery** and **higher reliability perception**, **cancellation rates** remain significant across both fulfilment types, indicating scope for operational improvement.
- Shifting more orders to Amazon fulfilment can **reduce cancellations and improve CX**.

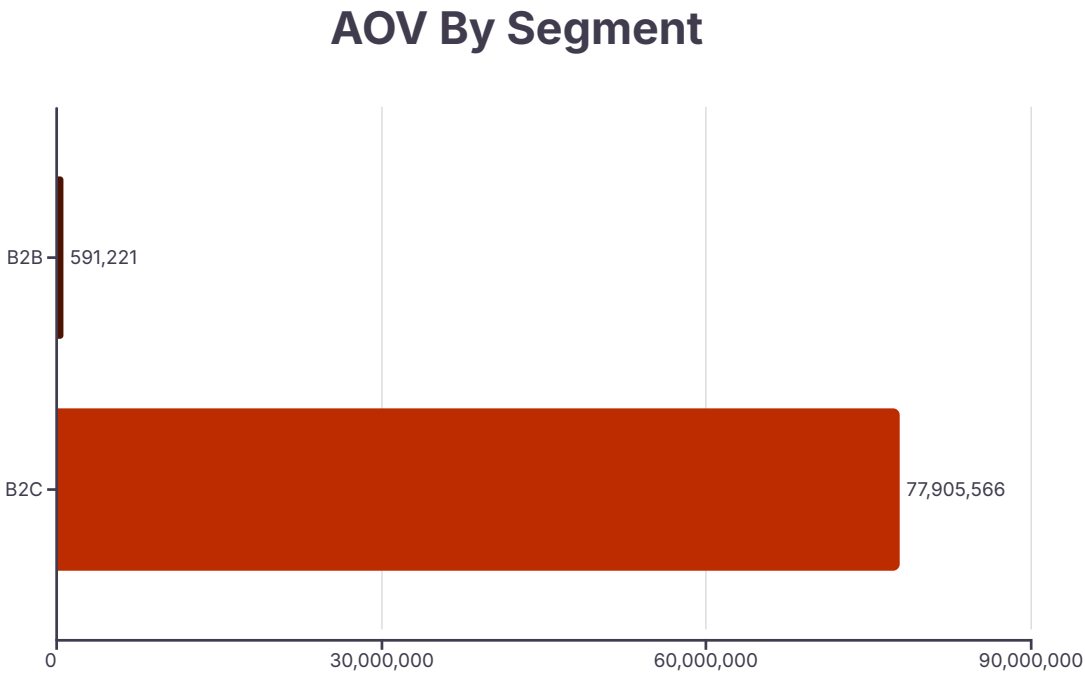
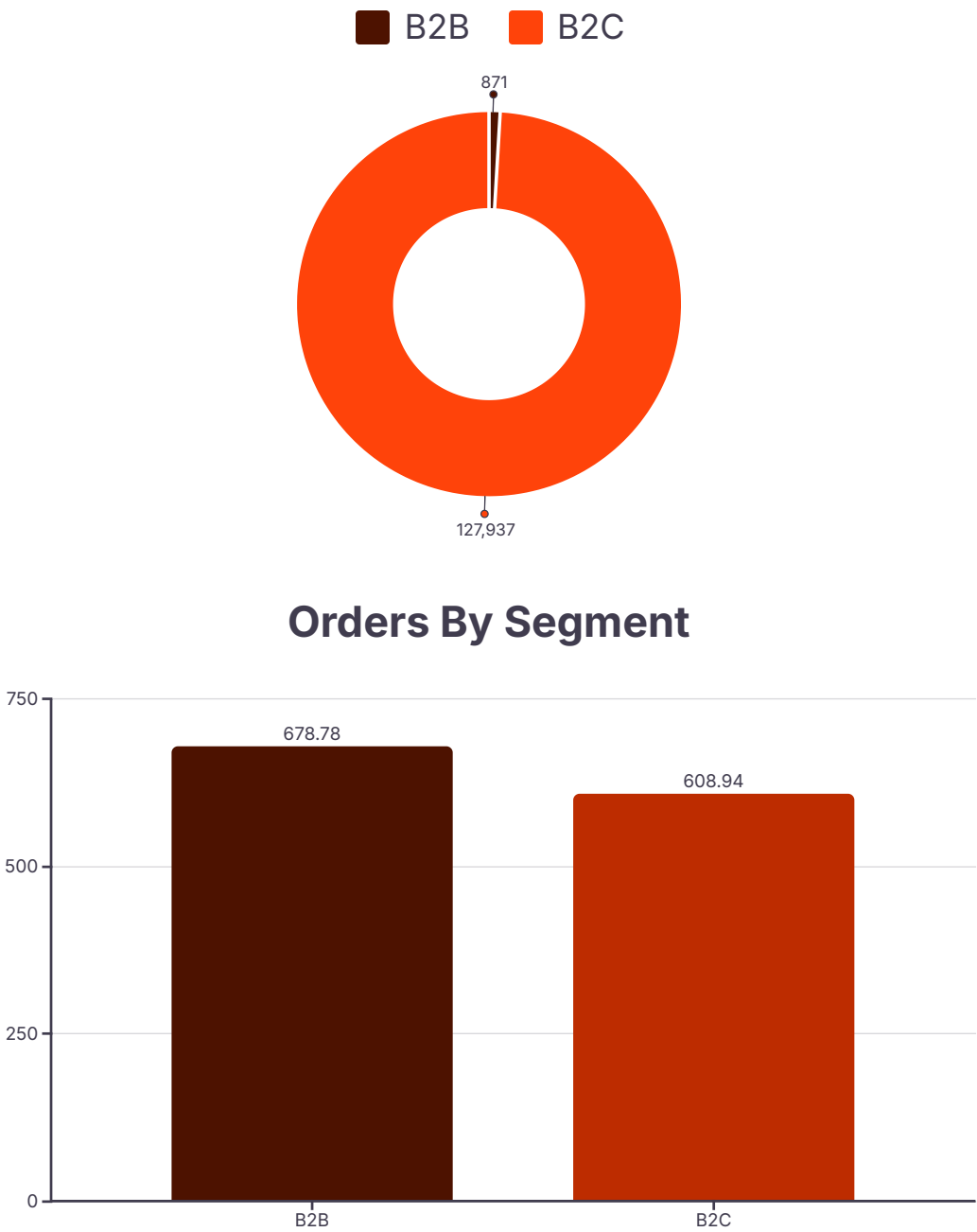


Courier Status Breakdown



Cancellation Rate By Fulfiment

Customer Segmentation: B2B vs B2C Insights



Revenue By Segment

B2B Customer Behaviour

1. Very **low order volume** but **high order value**
2. **Higher AOV (₹678.78)** than B2C
3. **Low cancellation rate (8.54%)**
4. **Stable and predictable** monthly demand
5. Strong **growth opportunity** with better profitability



B2C Customer Behaviour

1. **~99% of total orders** and main revenue driver
2. **Lower AOV (₹608.94)** compared to B2B
3. **Higher cancellations**, causing revenue leakage
4. **Seasonal spikes** (peak in April)
5. High **operational and logistics load**

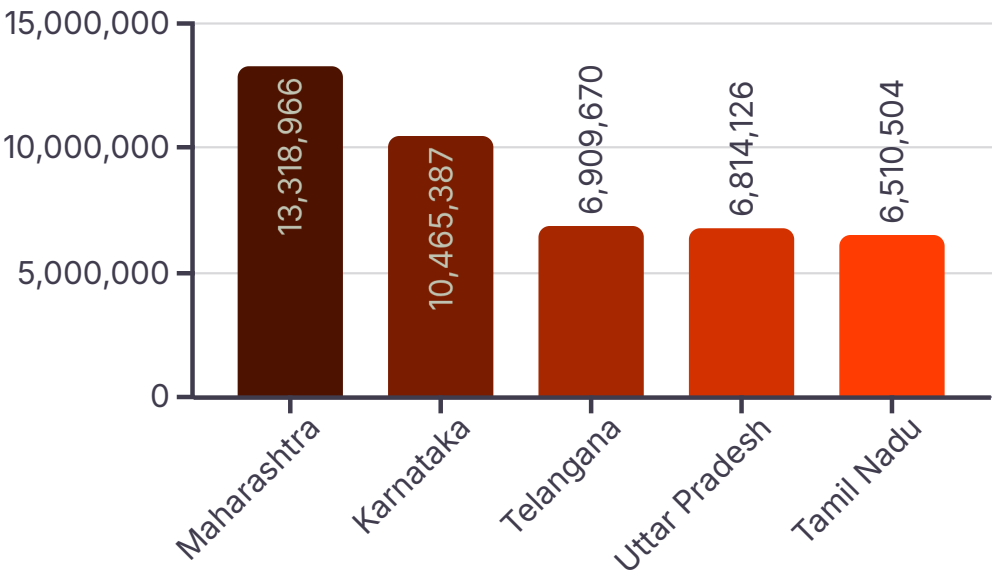
Business Insight:

- B2B customers are **higher quality, higher value, and more reliable**.
- There is a strong opportunity to **scale B2B sales** for stability and profitability.

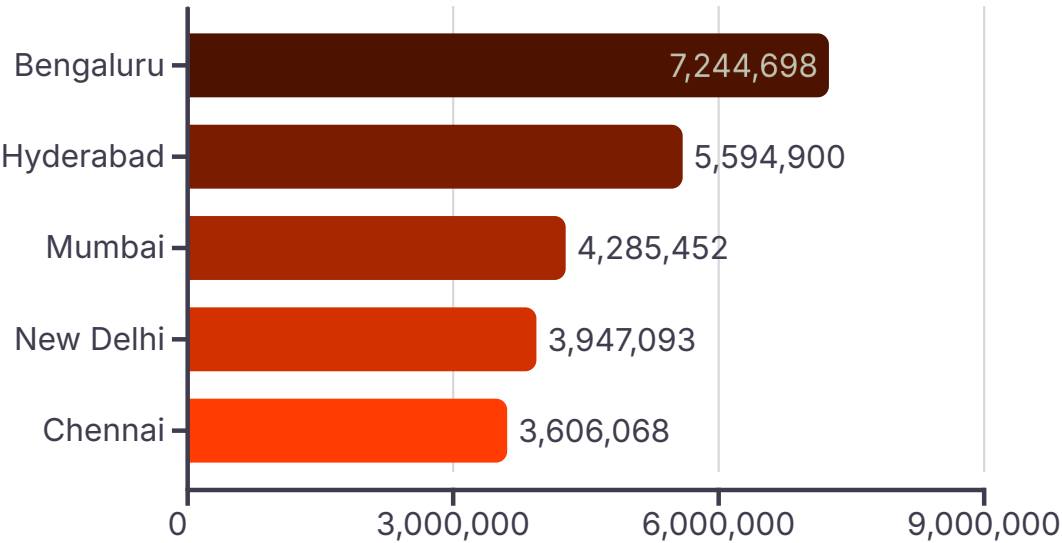
Geographical Analysis: Regional Performance

State-Level Revenue Distribution

From the **Geographical Analysis**, it is clear that sales and revenue are **highly concentrated in a few key states and metro regions**. **Maharashtra** emerges as the **top-performing state** in terms of both order volume and revenue, with Bengaluru standing out as the **highest revenue-generating city**, highlighting the strong contribution of urban and economically developed regions. The **top five states contribute 56.08% of total revenue**, indicating a significant **geographic dependency risk**, as business performance is heavily reliant on limited regions. On the risk side, certain low-volume regions such as **Lakshadweep show exceptionally high cancellation rates (25%)**, pointing to potential logistics, delivery reach, or customer experience challenges in remote locations. Overall, the dashboard suggests that while core states drive growth and stability, there is a clear need to **diversify geographically and address high-cancellation regions** to reduce risk and improve overall operational efficiency.



Top 5 States By Revenue



Top 5 Cities By Revenue

Strategic Business Recommendations

Based on comprehensive data analysis, we recommend the following strategic initiatives to enhance performance across key operational dimensions:



Reduce Cancellations by Optimizing Fulfilment & High-Risk Regions

- Focus on **high-cancellation states** (e.g., Lakshadweep and select North-East regions) and **merchant-fulfilled orders**, which show higher cancellation rates.
- Increase **Amazon-fulfilled coverage** and improve last-mile logistics in remote regions.



Expand and Nurture the B2B Segment for Stable Growth

- Although B2B contributes low volume, it has **higher AOV and significantly lower cancellations**.
- Introduce **bulk pricing, dedicated B2B SKUs, and account-based selling**.



Focus Inventory & Marketing on High-Performing Products and Sizes

- Apparel—especially **T-shirts and Shirts**—drives the majority of revenue, with **M, L, and XL** being the most popular sizes.
- Reduce investment in **low-performing categories** (Wallets, Socks, Watches).



Reduce Geographic Dependency Through Market Diversification

- Over **56% of revenue comes from the top 5 states**, increasing business risk.
- Use **targeted promotions and faster fulfilment options** to grow Tier-2 and Tier-3 markets.

Implementation Priority:

- Direct revenue recovery without increasing marketing spend.
- Improves profitability, demand predictability, and reduces volatility.
- Better inventory turnover, lower holding costs, and higher ROI.
- More resilient revenue base and long-term scalability.

Conclusion & Key Takeaways

Overall, the analysis shows that the business is **strongly driven by B2C demand in metro and top-tier states**, with apparel—especially T-shirts—being the primary revenue contributor. While sales performance is healthy, **high cancellation rates, geographic concentration, and underutilization of the B2B segment** limit profitability and scalability. Addressing operational inefficiencies and rebalancing growth levers can significantly improve long-term performance.

B2C Drives Scale, B2B Drives Quality

B2C contributes most orders and revenue, while B2B offers higher AOV and lower cancellations, making it a strong growth opportunity.

Revenue Is Geographically Concentrated

Over half of the revenue comes from the top five states, increasing dependency risk and highlighting the need for regional diversification.

Cancellations Are a Major Revenue Leakage

Higher cancellations in certain states and merchant fulfilment reduce net revenue and must be addressed operationally.

Core Products and Sizes Power the Business

A small set of apparel products and popular sizes (M, L, XL) generate most revenue, enabling focused inventory and marketing strategies.

Implementation of these recommendations, supported by continued analytical monitoring, positions the business for sustained competitive advantage and accelerated revenue growth in an increasingly dynamic e-commerce landscape.