

Amazon Sales Report

Project Id: 78G00L

1.Introduction: *In today's competitive e-commerce landscape, understanding sales data is crucial for businesses seeking to optimize their operations and improve customer satisfaction. This report delves into a comprehensive dataset of Amazon sales transactions, which includes valuable details on order ID, order date, fulfillment status, sales channel, product category, size, quantity, order amount, and shipping details, among other variables.*

The objective of this analysis is to extract actionable insights from these sales records to support data-driven business decisions. By examining key aspects such as sales trends, product performance, fulfillment efficiency, and customer purchasing patterns, this report aims to identify areas for growth, highlight potential operational inefficiencies, and suggest strategic adjustments to enhance revenue and customer experience.

With a detailed look into metrics across various product categories, channels, and fulfillment methods, this report serves as a vital tool for Amazon and its stakeholders to make informed decisions that align with market demands and business objectives. The following sections provide in-depth findings, visualizations, and recommendations based on data analysis, offering a roadmap to optimize sales strategies and improve overall business performance.

Key Objectives:

- i. Sales Overview:** *Understand the overall sales performance, trends, and patterns over time.*
- ii. Product Analysis:** *Analyze the distribution of product categories, sizes, and quantities sold to identify popular products.*
- iii. Fulfillment Analysis:** *Investigate the fulfillment methods used and their effectiveness in delivering orders.*
- iv. Customer Segmentation:** *Segment customers based on their buying behaviour, location, and other relevant factors.*

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

vi. Business Insights: Provide actionable insights and recommendations based on the analysis to optimize sales strategies, improve customer satisfaction and enhance overall business performance.

2. Data Cleaning and Preparation:

a) Understand the Data Structure

- **Review columns:** Check column names, data types, and understand what each column represents.
- **Inspect data types:** Use `df.info()` to understand the data types and detect any issues like numeric data stored as text.

3. Handle Missing Values

- **Identify missing values:** Use `df.isnull().sum()` to find columns with missing values.
- **Decide on an approach: Impute:** Replace missing values with suitable substitutes (e.g., mean or median for numerical data, mode for categorical data, or forward-fill for time series)
- **Drop rows or columns:** If a column or row has excessive missing values and isn't essential, it might be best to remove it.

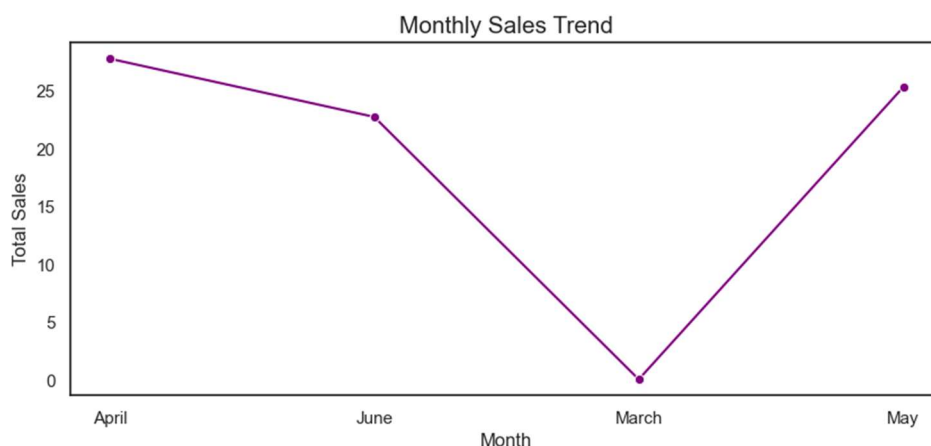
4. Analysis and Findings:

a) Sales Trends Analysis:

Objectives:

Identify Monthly Sales Patterns:

- Analyze monthly sales data to observe trends and identify months with high and low sales.



Results:

1. **Significant Drop in Sales from April to March:**

- Sales experienced a steep decline from April to March, reaching the lowest point in March. This may indicate a seasonal dip or other factors affecting demand in March.

2. **Recovery in Sales by May:**

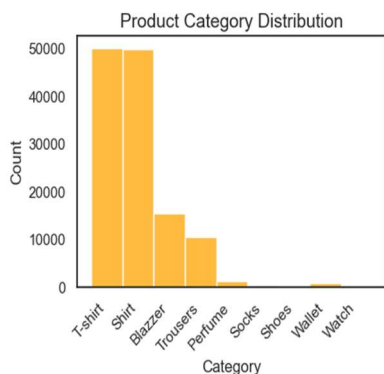
- After the March low, sales rebounded sharply by May, suggesting renewed customer demand or effective marketing efforts.

3. **Volatile Sales Trend:**

- The V-shaped trend in the chart highlights monthly volatility. This fluctuation suggests that sales may be influenced by external factors such as seasonality, promotional events, or consumer behavior shifts.

b). **Product Performance:**

Objective: Evaluate the popularity of different product categories and sizes to determine top-performing items.



i)Results:

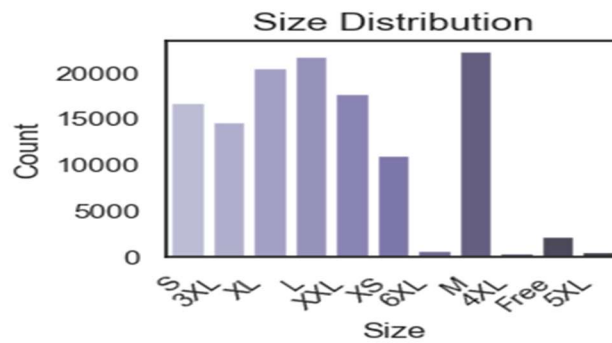
Top Sellers: T-shirts and Shirts have the highest sales, each around 50,000 units.

Moderate Sales: Blazers and Trousers show moderate demand.

- **Low Sales:** Perfume, Socks, Shoes, Wallet, and Watch have low sales numbers. Focus on high-demand items and consider promotions for lower-selling categories.

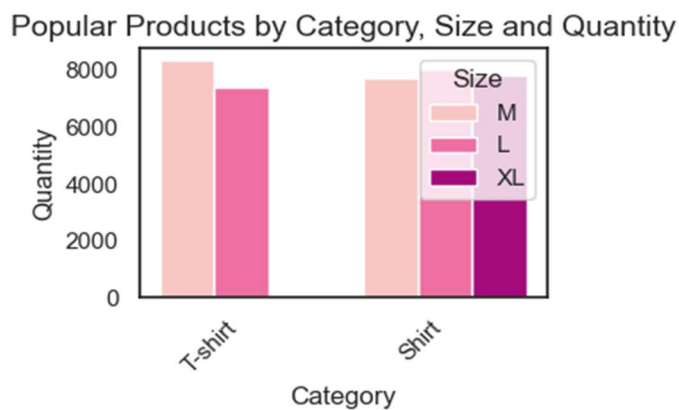
ii)Results:

- **Popular Sizes:** M, L, and XL have the highest counts, with M leading.
- **Least Popular Sizes:** S, Free, and 5XL have very low counts.
- Focus inventory on M, L, and XL sizes, and limit stock for lower-selling categories.



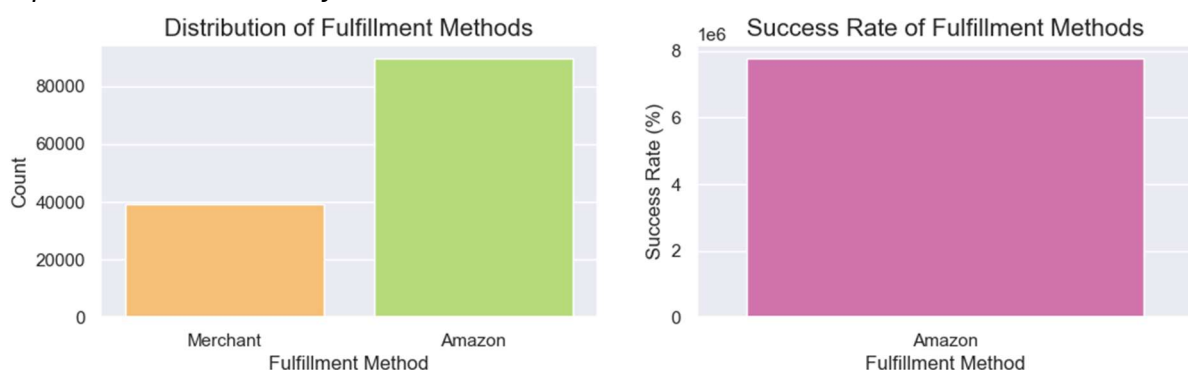
-
- **iii)Results:**

Top Categories: T-shirts and Shirts are popular in sizes M, L, and XL. Prioritize inventory for M, L, and XL in T-shirts and Shirts due to high demand.



c) Fulfilment Analysis:

Objective: Assess the effectiveness of various fulfilment methods and their impact on customer satisfaction.



Results: The fulfilment analysis in the provided image highlights two main points:

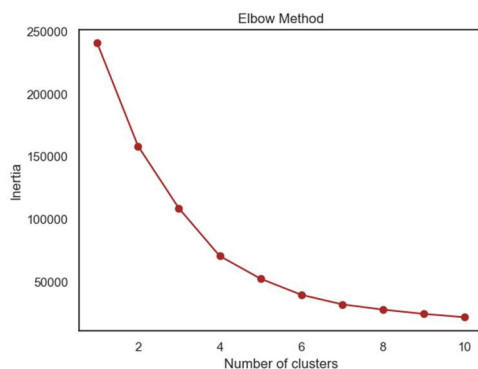
Distribution of Fulfillment Methods: The bar chart shows a comparison between the number of orders fulfilled by Amazon versus those fulfilled by merchants. Amazon has a significantly higher count of orders compared to merchants.

Success Rate of Fulfillment Methods: The second chart, although incorrectly labeled, appears to suggest that the success rate (or another metric related to the number of orders) is being compared between fulfillment methods. However, it only shows data for Amazon with a large value, implying a strong performance or high volume for Amazon.

In summary, Amazon handles more orders than merchants and seems to dominate in terms of fulfillment success or volume.

d) Customer Segmentation:

Objective: Segment customers based on their buying behaviour, location, and other relevant factors.



The elbow in the plot is around 3 clusters, this suggests that using 3 clusters is a reasonable choice for segmenting the customers.

Results:.



The scatter plot shows the relationship between the Total Amount Spent and the Total Quantity Purchased by different customer segments.

The data points are color-coded to represent three distinct customer segments (0, 1, and 2).

Segment 1 (green) has the highest representation, indicating a wide range of spending and purchasing behavior.

Segment 0 (purple) appears to have lower total amounts spent and quantities purchased.

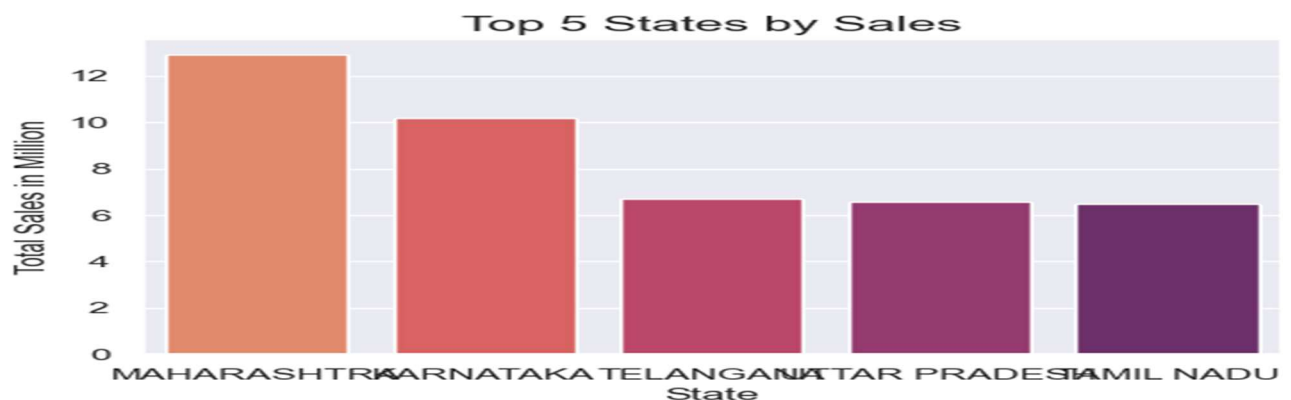
Segment 2 (yellow) also exhibits lower spending and purchasing, clustered at the lower end of the chart.

In summary, segment 1 customers tend to spend and purchase more, while segments 0 and 2 have relatively lower engagement.

d) Geographical Analysis:

Objective: Explore the geographical distribution of sales, focusing on states and cities.

i)Results:



The bar chart titled "Top 5 States by Sales" shows the total sales (in millions) for the top five states. The states listed are:

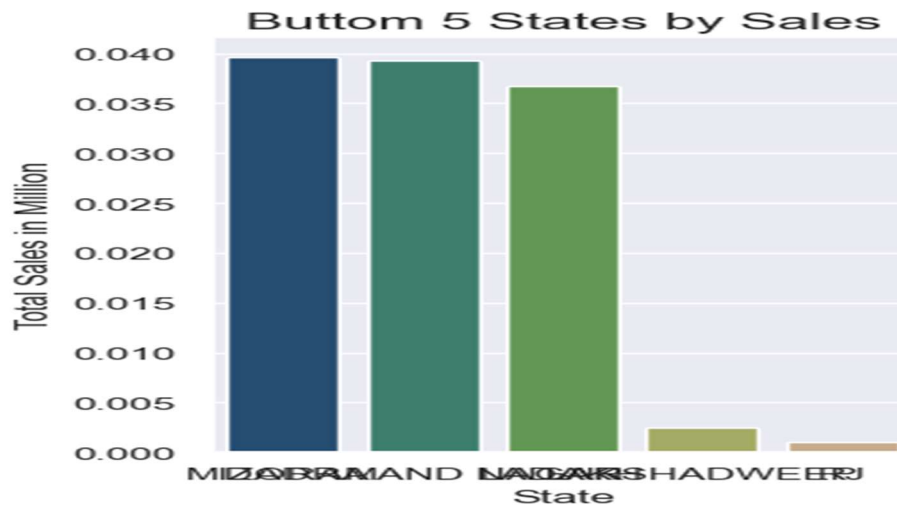
Maharashtra - Leading with the highest sales, approximately 12 million.

Karnataka - Following closely, with sales slightly below 12 million.

Telangana, Uttar Pradesh, and Tamil Nadu - These states have similar sales figures, all hovering around the 6 to 7 million mark.

This graphical analysis indicates Maharashtra and Karnataka are significantly ahead in sales compared to the other three states.

ii)Results:



The bar chart titled "Bottom 5 States by Sales" displays the total sales (in millions) for the five states with the lowest sales.

The states listed are:

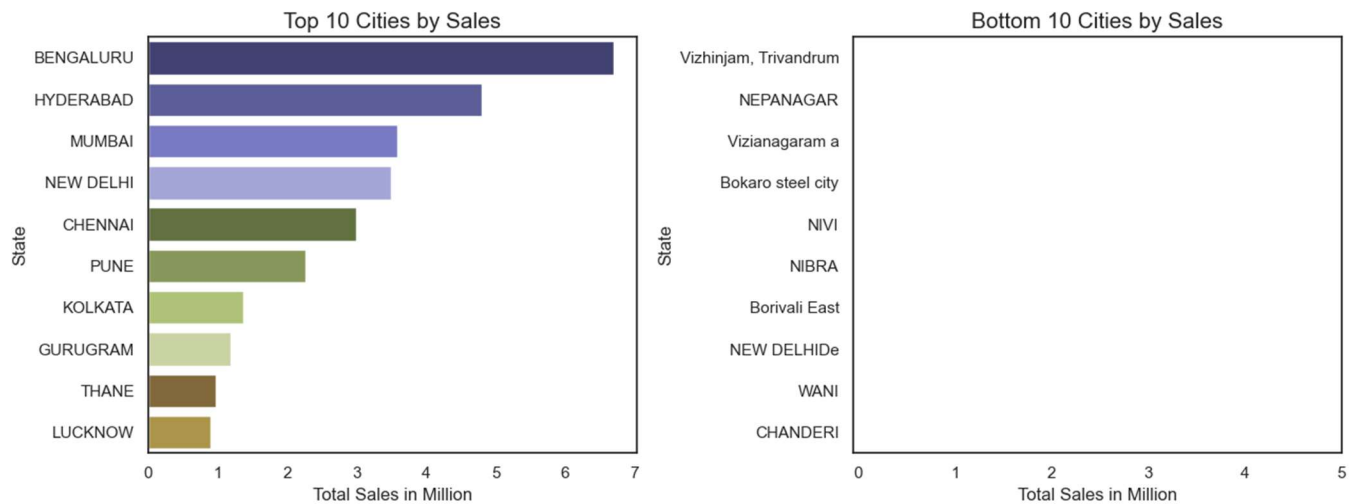
Mizoram, Nagaland, and Manipur - Each with sales around 0.04 million.

Dadra and Nagar Haveli and Daman and Diu - Sales slightly above 0.005 million.

Arunachal Pradesh - The lowest, with sales close to 0.002 million.

This analysis shows a significant disparity in sales among the bottom states, with Mizoram, Nagaland, and Manipur having higher sales compared to the minimal sales in Dadra and Nagar Haveli and Arunachal Pradesh.

iii)Results:



The combined bar charts show "Top 10 Cities by Sales" and "Bottom 10 Cities by Sales":

Top 10 Cities by Sales:

Bengaluru leads with the highest sales, around 6 million.

Hyderabad and Mumbai follow with sales between 4 to 5 million.

New Delhi and Chennai have similar sales, slightly above 3 million.

Other cities like Pune, Kolkata, Gurugram, Thane, and Lucknow range from about 1.5 to 2.5 million.

Bottom 10 Cities by Sales:

Cities like Vizhinjam (Trivandrum), Nepanagar, Vizianagaram, and Bokaro Steel City show significantly lower sales, with each under 0.1 million.

Other cities listed (e.g., Nivi, NIBRA, Borivali East, New DelhiDe, Wani, Chanderi) also have minimal sales.

Summary:

The top cities show a strong concentration of sales, particularly in major metropolitan areas, while the bottom cities reflect minimal sales, likely due to smaller populations or lower economic activity.