

Amazon Sales Analysis

Shadha Mariyam

Intern, Tamimah

Abstract

Amazon, a global e-commerce giant, has established a significant presence in the Indian market, offering a vast array of products and services to millions of customers. Known for its customer-centric approach, Amazon provides a seamless online shopping experience with features like fast delivery, easy returns, and a wide selection of items ranging from electronics to household goods. The company continuously innovates to meet the diverse needs of Indian consumers, leveraging its advanced technology and logistics network. This project utilizes detailed Amazon sales data in India to analyze key metrics and uncover insights that can drive strategic business decisions. This project is an in-depth exploratory data analysis and visualization of a dataset with 128943 unique entries that contains detailed insights into Amazon sales data. The dataset covers a 90-day time frame, from 31 March 2022 to 29 June 2022, and includes information such as Status, Sales Channel, Category, Size, Courier Status, Amount, City, State and Country. The purpose of the analysis is to demonstrate the ability to derive actionable insights from sales data.

About the Dataset

The amazon dataset covers a 90-day time frame, from 31 March 2022 to 29 June 2022, and includes information such as

- Category: Type of product. (String)
- Size: Size of the product. (String)
- Date: Date of the sale. (Date)
- Status: Status of the sale. (String)
- Fulfilment: Method of fulfilment. (String)
- Style: Style of the product. (String)
- SKU: Stock Keeping Unit. (String)
- ASIN: Amazon Standard Identification Number. (String)
- Courier Status: Status of the courier. (String)
- Qty: Quantity of the product. (Integer)
- Amount: Amount of the sale. (Float)
- B2B: Business to business sale. (Boolean)
- Currency: The currency used for the sale. (String)

Data Cleaning

The dataset has 128976 entries. The data was highly imbalanced when column B2B was taken into consideration. We could see more than 99% of orders are B2C. The number of missing values is more in fulfilled-by variable which is about 70 % and 40% in promotion ids. Hence both are dropped out. There are missing values in courier status and amount. Courier status is filled with the mean while amount is filled with mean. 33 rows had the address details missing which includes city, state, pincode and country. Since the percentage of these rows are very small, removing these causes no harm.

Google colab notebook:

https://colab.research.google.com/drive/1eKQlZncyjvOCAcdWTTGphqmCxs9k_yrh?usp=sharing

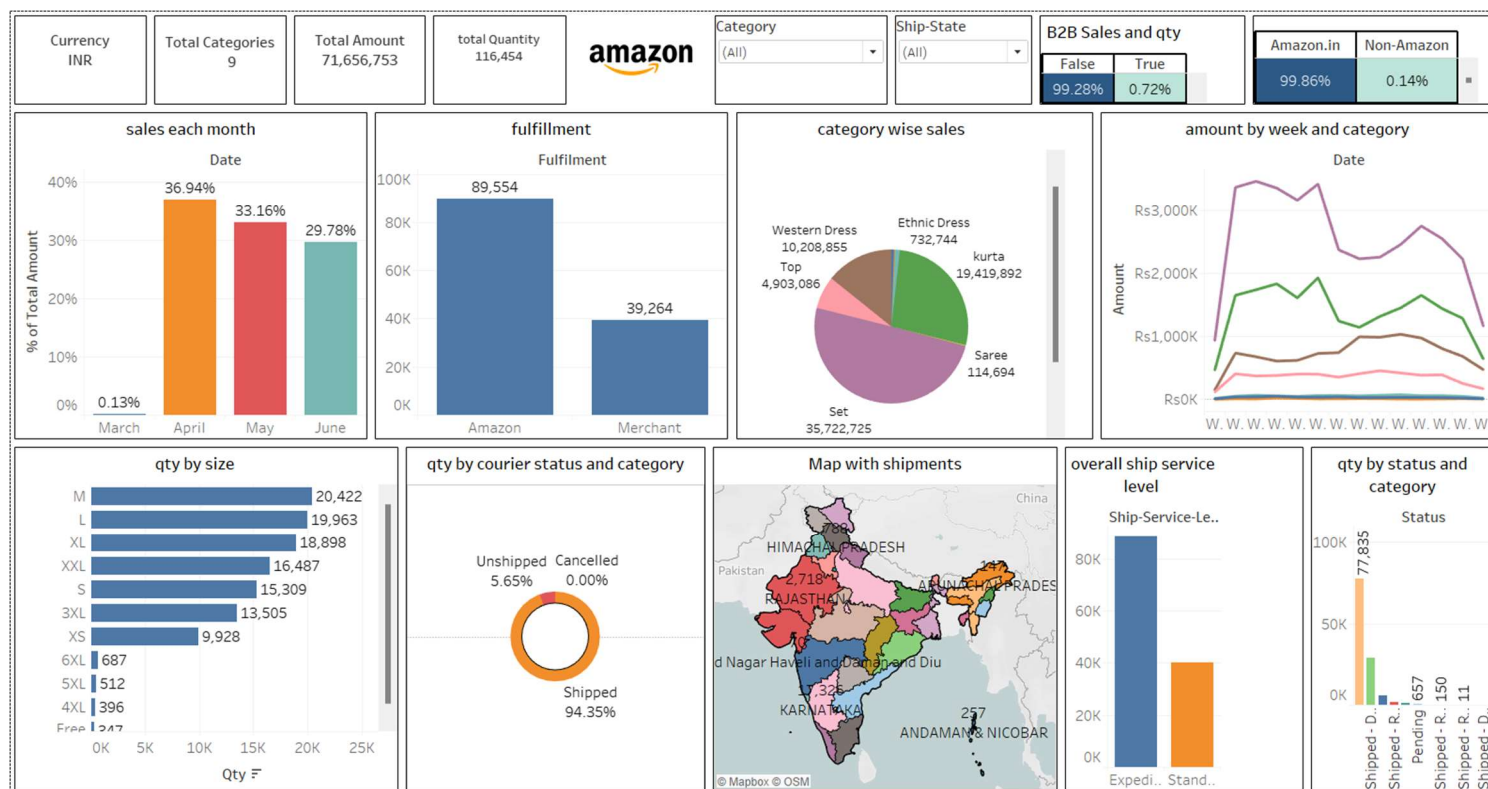
Analysis

Total Revenue generated over a specific month:



At Rs 26,413,120.00, April had the highest revenue followed by May and June in India. April accounted for 36.92% of Revenue. There is a fall in revenue in India.

Let's dive into the dashboard:

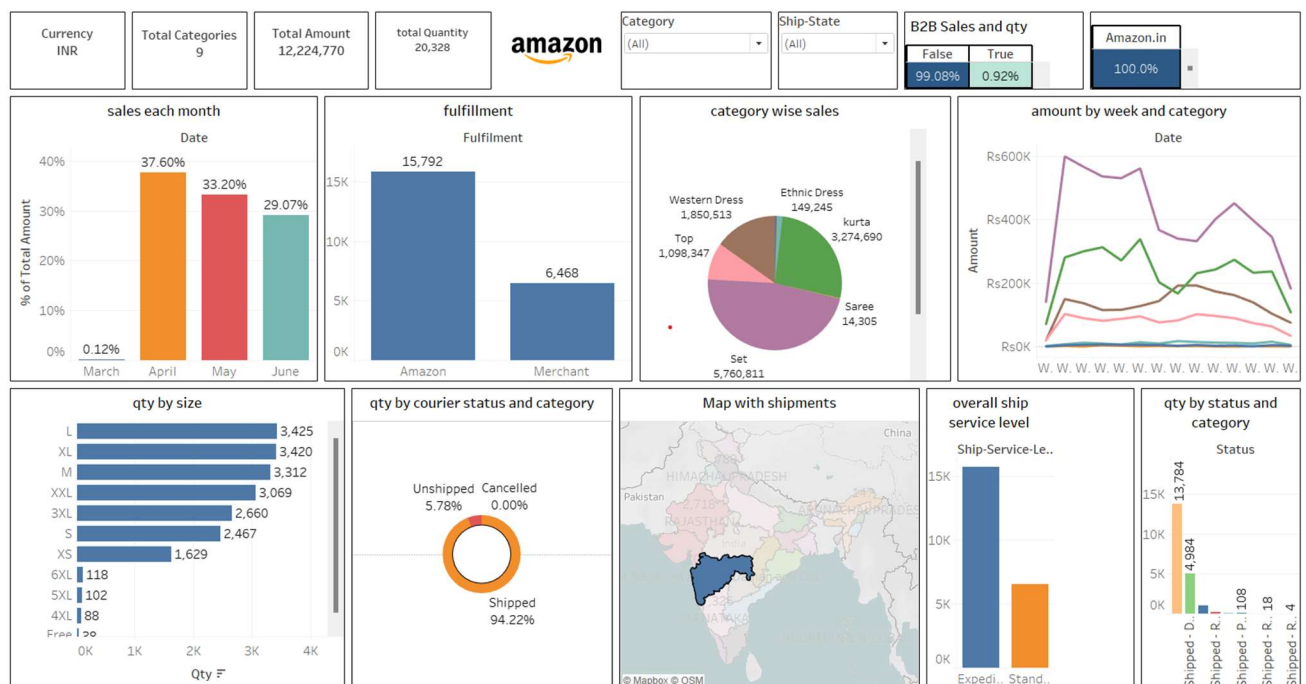


The dashboard visualizes various aspects of the sales data, including monthly sales performance, fulfillment methods, category-wise sales, quantity sold by size, courier status, shipping service levels, and geographic distribution of shipments. During this period, the total sales amount reached INR 71,656,753, with a total quantity of 116,454 items sold across nine categories.

April and May emerged as peak sales months, driven primarily by the high demand for sets and kurtas. Amazon's fulfillment services dominated with a significant portion of orders being handled directly, and expedited shipping emerged as the preferred delivery option, reflecting customers preference for quick service. Geographic distribution highlighted Maharashtra as the top destination for shipments, followed by Karnataka and Tamil Nadu, indicating strong market presence in these regions. In contrast, regions like Lakshadweep, Orissa and Pondicherry saw minimal activity, suggesting potential for market expansion. Pondicherry and Orissa had only 1 and 3 shipments respectively.

Despite the overall success, areas for improvement include addressing the 5.65% of unshipped orders and optimizing inventory for less popular sizes to minimize stockouts and overstocking. Would suggest exploring B2B opportunities to broaden revenue streams. These findings can guide strategic decisions aimed at improving customer satisfaction, operational efficiency, and market expansion for Amazon India.

Now let's summarize based on Maharashtra which has the strong market presence.



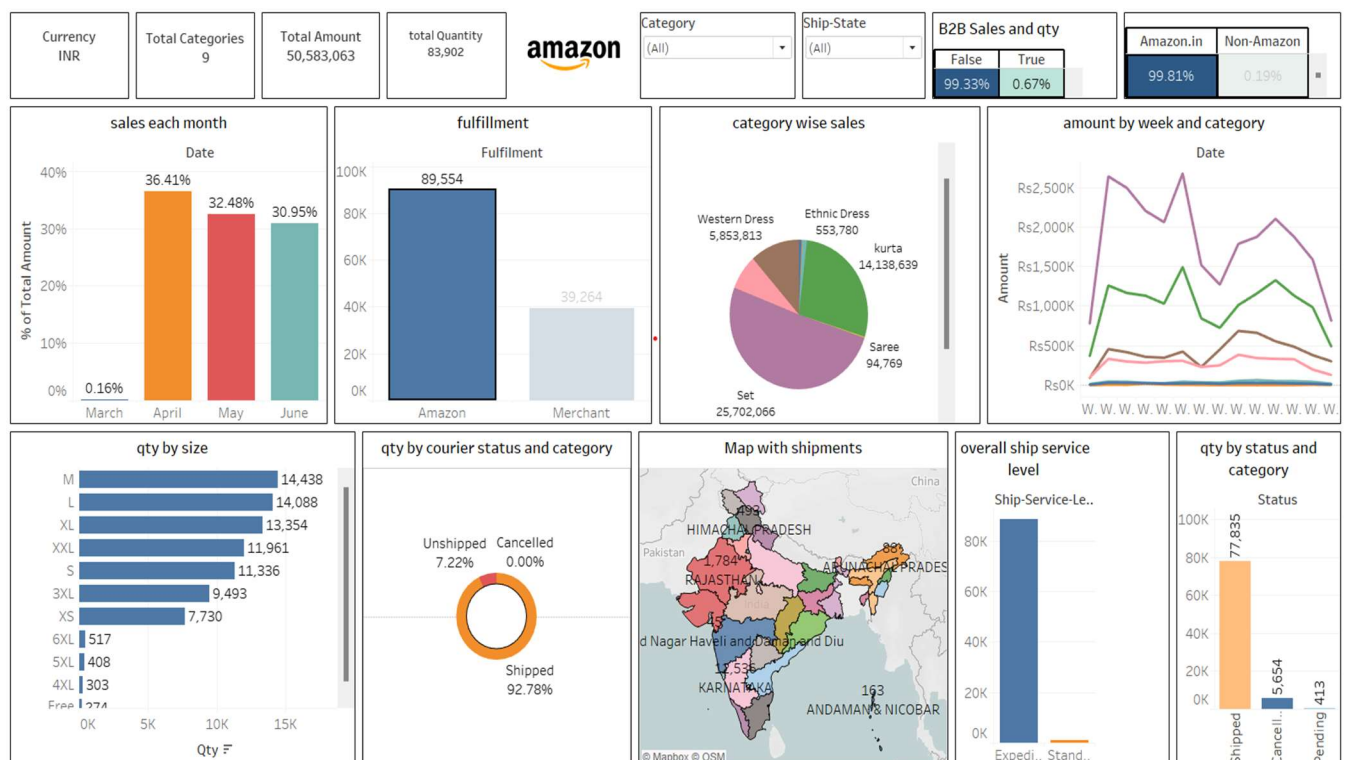
As similar to overall sales, April emerges as the peak sales month in Maharashtra, contributing 37.60% to the total sales in Maharashtra amounting to ₹12,224,770, followed by May and June with 33.20% and 29.07%, respectively. Fulfilment data highlights Amazon's significant role, handling 15,792 orders compared to 6,468 fulfilled by merchants within Maharashtra. "Set" and "Kurta" categories lead in sales, collectively representing over 70% of total sales within the state.

The dashboard also shows high operational efficiency, with 94.22% of orders successfully shipped and minimal cancellations in Maharashtra. This efficiency is crucial for maintaining customer satisfaction. Geographically, sales within

Maharashtra show concentrated activity, suggesting opportunities for targeted market penetration strategies. Size-wise L and XL are most popular, providing essential insights for inventory management decisions specific to the state. Expedited shipping emerges as the preferred service level, reflecting Amazon's commitment to prompt delivery and customer satisfaction in Maharashtra.

Now let's see how amazon fulfilment works across the country:

Amazon has two options for sellers to handle their order fulfillment: Fulfillment by Amazon (FBA) and Fulfillment by Merchant (FBM). **FBA** allows sellers to outsource storage, packaging, shipping, and customer service to Amazon, making products eligible for Amazon Prime and benefiting from Amazon's logistics network. On the other hand, **FBM** requires sellers to manage these tasks themselves, offering greater control and potentially lower costs, but without the automatic Prime eligibility and logistical support from Amazon.



The sales analysis indicates a strong performance in April, May, and June, with April

achieving the highest percentage of total sales of amazon fulfilment at 36.41%. Fulfilment data shows that Amazon handles most orders, with 89,554 orders fulfilled compared to 39,264 by merchants. This reflects Amazon's robust infrastructure in managing logistics efficiently. The category-wise sales breakdown reveals that the "Set" category dominates sales with Rs 25,702,066, followed by "Kurta" and "Western Dress.". It is also highlighting the shipment efficiency, showing that 92.78% of orders are shipped, with only 7.22% remaining unshipped and 0% cancellations. This indicates a high level of operational effectiveness. The geographic map illustrates shipment distribution across different Indian states, with significant volumes in Maharashtra, Karnataka and Tamil Nadu, suggesting these are key markets for Amazon. Additionally, the analysis of overall ship service levels shows a preference for expedited shipping, reinforcing the importance of quick delivery in customer satisfaction.

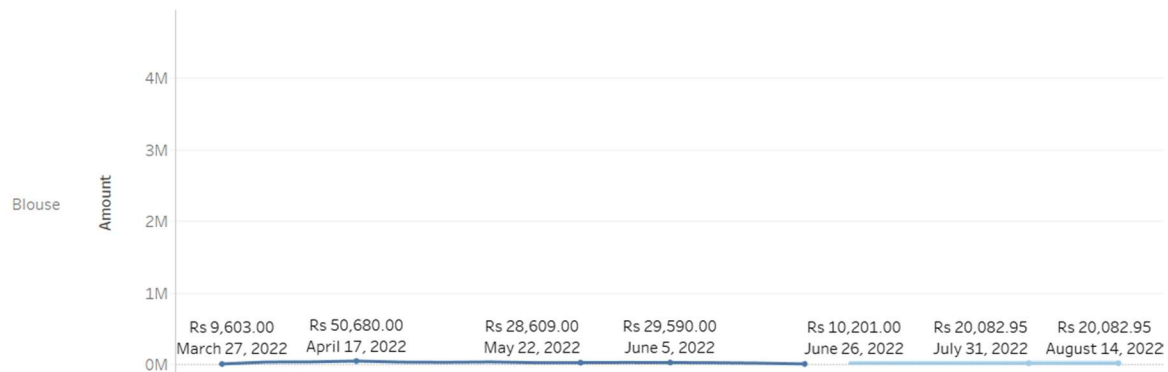
Prediction

The prediction for this data is done using exponential smoothing since this time series data has no clear trend or seasonality. The time period is very small. Exponential smoothing is a time series forecasting technique that uses weighted averages of past observations to forecast future values. The weights decrease exponentially over time, giving more importance to recent observations while not entirely discarding older observations. This method is particularly useful for data with no clear trend or seasonal pattern.

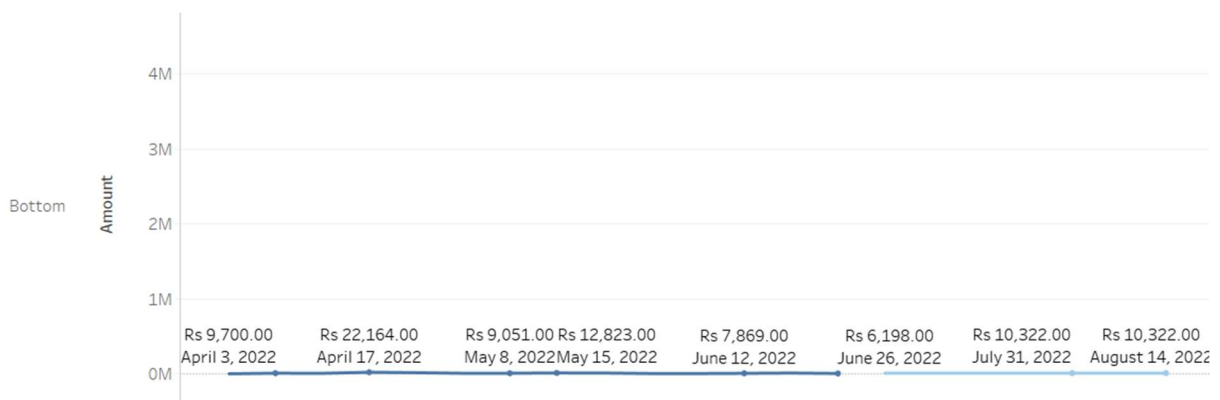
We have predicted the sales for the next 7 weeks (that is from July 3, 2022, to August 14, 2022). Since the time series is ridiculously small, the interval of prediction is also shown (in light blue color)



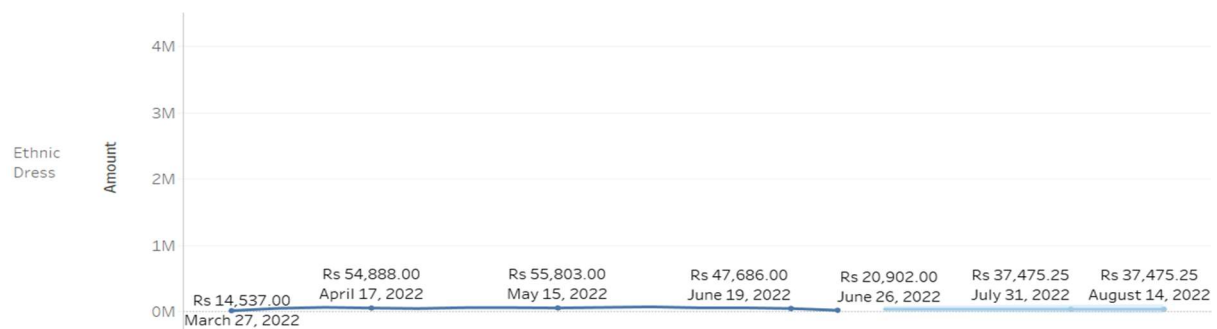
Now let's see the prediction for each category



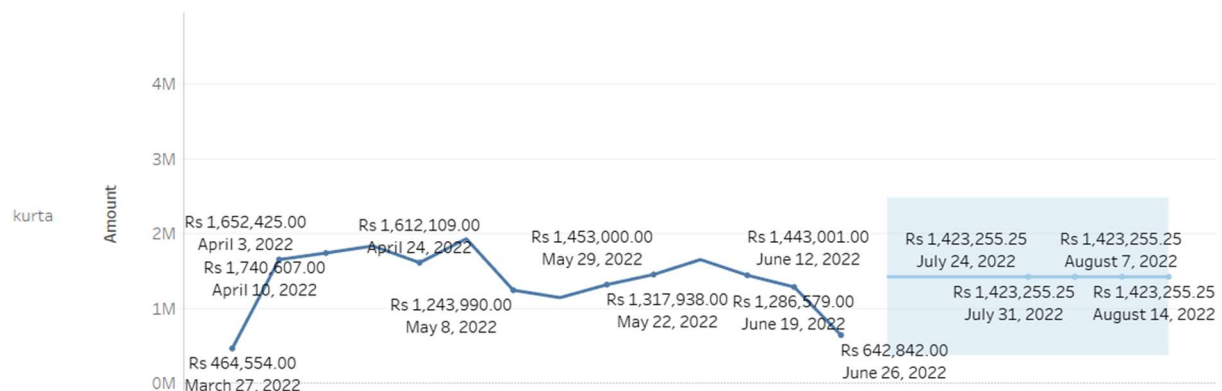
This is the prediction for Blouse, it is almost linear. Predicted sales show a steady forecast of Rs 20,082.95 on the interval from July 31 to August 14, 2022, indicating an expectation of consistent sales around that amount following the actual data trends.



This is the case for Bottom Category. Predicted sales, shown in light blue, indicate steady forecasts of Rs 10,322.00 on the interval from July 31, 2022, to August 14, 2022. The actual sales show fluctuations, with lower values such as Rs 7,869.00 on June 12, 2022, and Rs 6,198.00 on June 26, 2022.



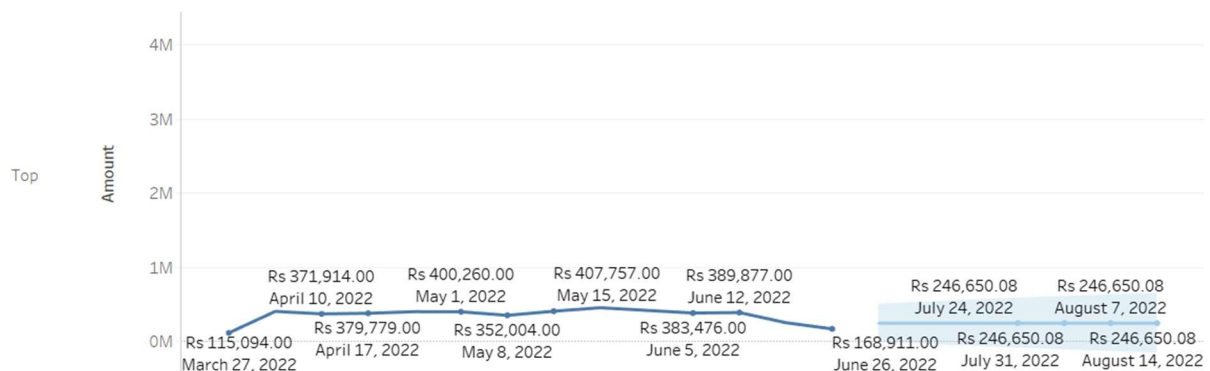
This is the case for ethnic dresses. There was a dip to Rs 20,902.00 on June 26, 2022, indicating some variability in actual sales.



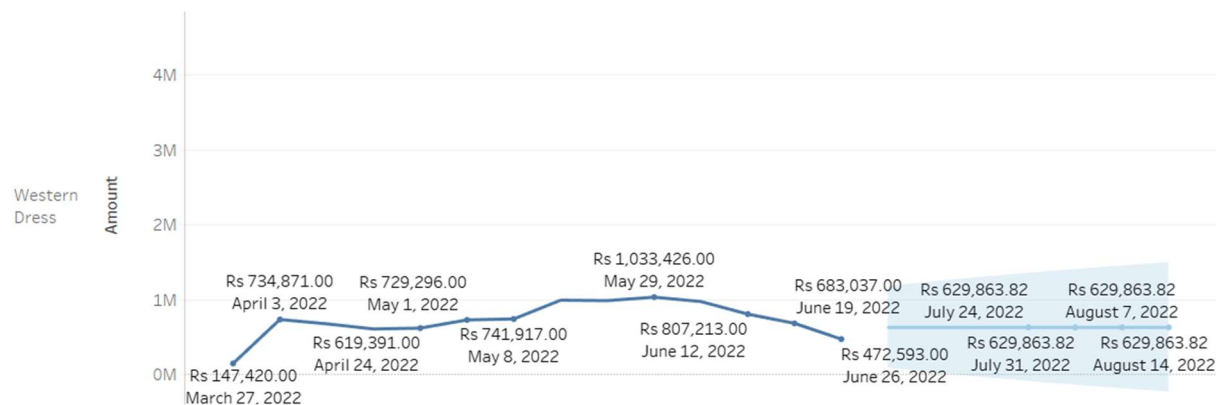
This is the case for kurta category we can clearly see that the demand for kurtas is likely to increase in the coming years.



In the case of Set, there is a very long prediction interval indicating that there is a lot of fluctuations in the market. But the interval is constant throughout.



When it comes to the case of Top, it moreover has a trend or the sales is almost uniform throughout.



Actual sales fluctuated throughout the period, with a peak of over Rs 1,033,426 on May 29, 2022, and a low of Rs 472,593 on June 26, 2022. Predicted sales generally followed the same trend, but there were some deviations, particularly in June and July. It appears that actual sales were higher than predicted sales in June and lower in July.

Conclusion

The analysis of Amazon sales data in India from March to June 2022 reveals valuable insights. April stood out as the highest revenue-generating month, with Maharashtra leading in sales of sets and kurtas. The predictive analysis using exponential smoothing provides a glimpse into future sales trends. Let's see how improvements can be made: Enhance demand

forecasting for popular sizes (M, L) to avoid stockouts. For less popular sizes, optimize inventory to reduce overstocking costs. Consider expanding product lines in high-demand categories while refining offerings in less popular ones based on customer feedback and sales trends. Develop region-specific promotions and product offerings to tap into local preferences and cultural trends. Create dedicated marketing strategies for B2B customers, highlighting the benefits of purchasing through Amazon. Additionally, focus on targeted market expansion by understanding regional sales trends and preferences to strengthen its market presence in specific regions.

References

1. Dataset:

<https://www.kaggle.com/datasets/thedevastator/unlock-profits-with-e-commerce-sales-data/data>

2. Tableau :

<https://www.udemy.com/course/tableau10/>