

Final Summary:

Brazilian E-Commerce

1. Key drivers of delivery delays

The primary drivers of delays in deliveries are the distance between the seller's state and the customer's state. Delays also depend on where the product is being shipped from (seller state), where it is being delivered to (customer state) and the overall freight value of the product

2. Geographic weaknesses in the logistics pipeline

Orders delivered to the northeastern region of Brazil experience the most delays (17.72% of orders are delayed). Orders where the seller's state is based in the north-to-northeast experience the most delays (24.38% of orders are delayed), although these delays are still lower than those observed for deliveries to the same region.

3. Customer segments most at risk for delays

Customers receiving orders in the northeastern region are most likely to experience delays (17.72% delay in orders). Customers purchasing items with higher freight values are also at a greater risk of late deliveries (24.38% delay). Customers shipping across regions, where the seller state and customer state are different. High freight orders may also experience more delays.

4. How confident are you in the model?

Confidence in the model is poor. The performance metrics are poor with the precision at 0.1, F1 at 0.17 and ROC AUC at 0.586. The model would benefit from additional, more constrained and meaningful features that better capture the complexity of the delivery process. For example, incorporating actual distance between the sellers and customers, stock location (which products are being shipped from which state), investigating features within the delayed-only data (is it dimension-based? Is it distance based?) Additionally, including looking at the relationships between features (features are currently individualised in the model), for example, the relationship between seller state and logistics partner.

5. What data should the business collect next?

Recommended next data points include:

- *Stock or seller location data – identifying where the inventory is being shipped to and from.*
- *Route complexity data – understanding how many transfers occur along the route, the accessibility of infrastructure in each region, the frequency of transport to the area, and the typical load on transportation networks*
- *Seasonality – weather-based (e.g., delays from high flood regions during rainy season) and the influx of orders (e.g., there is an influx of orders during the festive season as well as an increase in delays)*

6. Recommendations for reducing delays operationally

Stock placement: pinpoint which products are in high demand in which regions, and stock or adjust sellers accordingly. This applies both to cross-state deliveries and to relocating inventory or using seller companies closer to states with high order volume, where viable.

Track carrier and seller performance: Monitor the performance of respective logistics partners across different regions, prioritising carriers with a stronger performance. Underperforming carriers can be reviewed or replaced. This is also applicable to sellers.

Prepare for seasonality: Implement season-specific logistical planning. This could be increasing the carrier frequency in peak season or adjusting the state sellers for certain products during the rainy season.