Delivery Delays, Customer Dissatisfaction & Revenue Risk in E-Commerce

An operational analytics study using "Olist Brazilian E-commerce data"

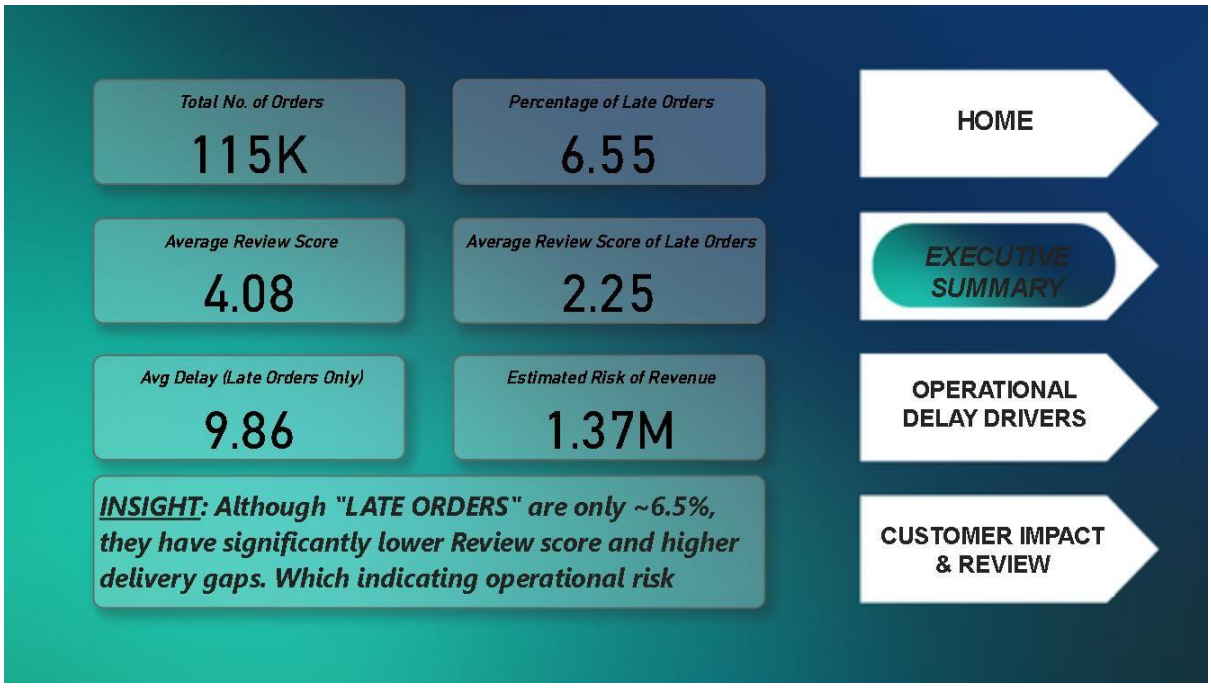
Asmit Swarnakar | Python, Power BI | 6thfeb, 2026

EXECUTIVE SUMMARY:

This analysis investigates the impact of delivery delays on customer satisfaction and revenue risk in an e-commerce platform.

Although late deliveries account for only ~6.5% of total orders, they a disproportionate negative impact on customer reviews and revenue exposure. The shipped -> delivery stage contributes the most to delivery delays, indicating potential logistics SLA (Service Level Agreement) issues.

Late orders receive an average review score of ~2.2 compared to ~4.1 overall, directly linking performance to customer dissatisfaction.



BUSINESS PROBLEM:

Why E-commerce orders fail even when system look fine?

OBJECTIVES:

1. Identify frequency and severity of delivery delays
2. Analyse operational stages causing delays
3. Measure customer sentiment impact
4. Estimate revenue at risk due to delays

DATA OVERVIEW & PREPARATION:

Multiple transactional datasets were cleaned, aggregated and merged into a single master dataset for sake of analysis. Delivery delays were calculated using actual delivery dates vs estimated delivery dates and classifying orders into ‘Early’, ‘On-Time’, ‘Late’, and ‘Very Late delivery’. Connect delays with review score and also analyse logistic gaps.

KEY ANALYSIS:

A. DELIVERY DELAY DISTRIBUTION:

1. Early deliveries dominate
2. Late deliveries are 6.5 of total orders

```
import pandas as pd
df = pd.read_csv('Master_data_Ecommerce.csv')
Late_order_total = df[df['delivery_delay_days']>0].shape[0]
total_no_orders = df['delivery_delay_days'].shape[0]
print("Late_order_total:",Late_order_total)
print("total_no_orders:",total_no_orders)
percetage_of_late_orders = ((Late_order_total/total_no_orders)*100)
print("percetage_of_late_order:",percetage_of_late_orders)
```

```
Late_order_total: 7264
total_no_orders: 110145
percetage_of_late_order: 6.594943029642744
```

B. OPERATIONAL DELAY DRIVERS:

1. Shipped → Delivered stage has the highest average delay
2. Indicates logistics partner or last-mile issues
3. SLA performance degradation visible for late orders

```
print("Avarage purchase_approved_gap:",df.loc[df['delivery_delay_days'] > 0, 'purchase_approved_gap'].mean())
print("Avarage approved_shipped_gap:",df.loc[df['delivery_delay_days'] > 0, 'approved_shipped_gap'].mean())
print("Avarage shipped_delivered_gap:",df.loc[df['delivery_delay_days'] > 0, 'shipped_delivered_gap'].mean())
```

```
Avarage purchase_approved_gap: 0.3519614590502409
Avarage approved_shipped_gap: 5.235132158590308
Avarage shipped_delivered_gap: 27.022577092511014
```

C. **CUSTOMER IMPACT AND REVIEWS:**

1. Average review score drops from ~4.1 to ~2.2 for late orders
2. Late deliveries have higher share of ≤2-star reviews
3. Clear correlation between delivery delay and dissatisfaction

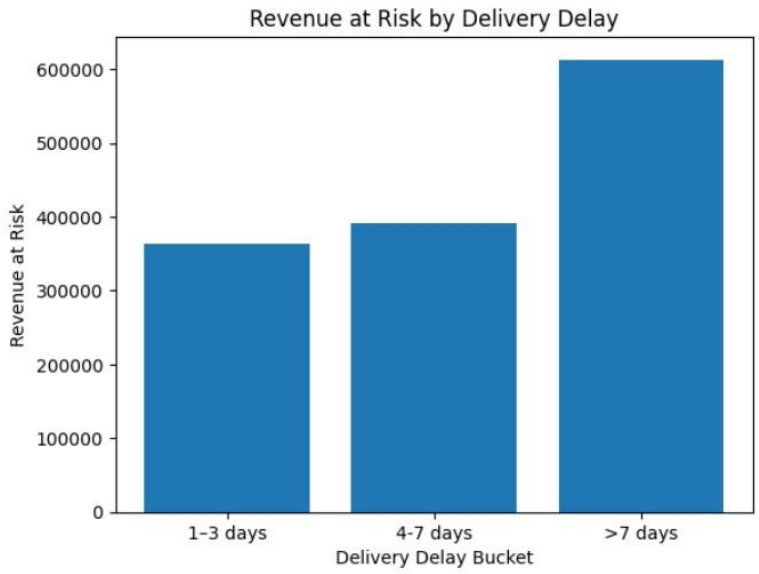
```
df[df['delay_days'].isin(['LATE'])].groupby('delay_days')['review_score'].value_counts(normalize=True)* 100
```

```
delay_days  review_score
LATE        1.0            46.464646
            5.0            20.913682
            4.0            12.603306
            3.0            11.730946
            2.0             8.287420
Name: proportion, dtype: float64
```

ACKNOWLEDGEMENT: Late deliveries have a significantly lower average review score (≈2) compared to on-time deliveries. Additionally, around 54% of late orders receive a rating of 2 stars or below, indicating that delivery delays strongly contribute to customer dissatisfaction and negative reviews

REVENUE RISK ASSESSMENT:

Late and very late orders contribute disproportionately to revenue risk, with in estimated exposure of 1.37M. Although operationally small in volume but delayed orders pose a significant reputational and financial risk.



DASHBOARD & DECISION SUPPORT:

1. Multi-page dashboard
2. Executive summary view
3. Drill-down into operational and customer impact
4. Navigation enabled for stakeholder consumption

BUSINESS RECOMMENDATIONS:

1. Review logistics SLA (Service Level Agreement) for shipped -> delivered stage. Renegotiate with them if needed.
2. Flag high risk courier routes. Take penalty if threshold of delivery date over.
3. Predict late deliveries at shipping stage.

LIMITATION:

1. Analysis limited to historical delivery data**.**
2. External factors (e.g, weather, regional logistics) not modelled**.**

FUTURE WORK:

As a future enhancement, a delay risk prediction model can be built to proactively identify SLA-breaching orders.

FINAL CONCLUSION:

Although the overall proportion of late deliveries is relatively small, their impact is significant. An order is considered failed if it is delivered more than 10 days late—around **31% of late orders exceed this threshold**—or if it is cancelled after payment approval or receives a review score of **≤ 2**.

The analysis shows that the **shipped → delivered stage is the primary contributor to delivery delays**, indicating logistics and last-mile fulfilment as the key operational risk area.
