
Order Fulfillment Funnel & Delivery Performance Analysis

Brazilian Olist Dataset | SQL (BigQuery) + Tableau | Portfolio Project — Part 2

Giovanni Valdes

2026

Executive Summary

This report is the second in a two-part analysis of the Brazilian Olist e-commerce dataset. The first report established that 97% of customers are one-time buyers and that the top 10% of customers drive 38% of revenue. This report investigates *why* customers may not return — by examining the order fulfillment pipeline, measuring delivery performance, and connecting delivery experience to customer satisfaction and repeat purchase behavior.

The central finding is a paradox: the fulfillment funnel is operationally excellent, with conversion rates above 98% at every stage — but the customer experience inside that efficient pipeline reveals a meaningful problem. Late deliveries, affecting 8.1% of orders, are associated with a dramatic drop in review scores (4.29 to 2.57). Crucially, delivery performance alone does not explain the low retention rate, pointing to deeper structural factors in marketplace purchasing behavior.

99,441 Total Orders	12.1 Avg Delivery Days	8.1% Late Delivery Rate
-------------------------------	----------------------------------	-----------------------------------

Project Context

How This Connects to Report 1

The first analysis revealed a 97% one-time buyer rate — the defining challenge for the Olist business. A natural follow-up question is whether operational failures in fulfillment are driving customers away. This report was designed to answer that question directly: does the order pipeline lose customers, and does delivery performance explain the retention problem?

Dataset & Tools

The same Brazilian E-Commerce Public Dataset (Olist) from Kaggle was used, with two additional tables joined: **order_reviews** (customer satisfaction scores) and the timestamp fields within **orders** (purchase, approval,

shipment, delivery, and estimated delivery dates). All analysis was performed in Google BigQuery using SQL. Findings were visualized in Tableau.

- SQL techniques: CASE WHEN, DATE_DIFF, CTEs, multi-table JOINS, AVG, COUNT DISTINCT
- Tables used: orders, order_reviews, customers, order_items
- Scope: all orders with a recorded delivery date (96,476 orders)

Dashboard Overview

The Tableau dashboard below presents the four core findings in a single view: total order volume KPIs, the fulfillment funnel, on-time vs. late delivery breakdown, and the customer satisfaction impact of delivery delays.

Order Fulfillment Funnel & Delivery Performance Analysis

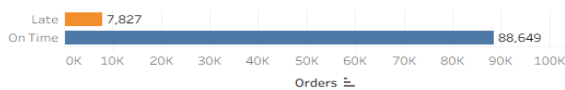
Analyzing delivery performance and its impact on customer experience

99,441
Total Orders

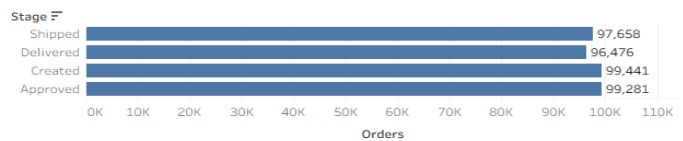
12.1
Avg Delivery Days

8.1%
Late Delivery Rate

8% of Orders Experience Delivery Delays Despite High On-Time Rate



Order Funnel Shows Minimal Drop-Off Across Fulfillment Stages



Late Deliveries Significantly Reduce Customer Satisfaction



Figure 1 — Order Fulfillment Funnel & Delivery Performance Dashboard (Tableau)

Findings & Analysis

Finding 1 — The Fulfillment Funnel Is Operationally Excellent

Mapping orders through each fulfillment stage reveals a highly efficient pipeline. Of 99,441 total orders, the vast majority successfully progress from creation through approval, shipment, and final delivery with minimal drop-off at any individual stage.

Stage	Orders	Conversion Rate
Created (Total)	99,441	—
Approved	99,281	99.84%
Shipped	97,658	98.37%
Delivered	96,476	98.79%

All three stage-to-stage conversion rates exceed 98%. The largest drop occurs between approval and shipment (1.63%), suggesting that seller fulfillment delays are the primary source of pipeline loss — not payment processing or final-mile delivery.

Key SQL Query — Funnel Conversion Rates

```
WITH funnel AS (  
  SELECT  
    COUNT(*) AS total_orders,  
    COUNT(order_approved_at) AS approved_orders,  
    COUNT(order_delivered_carrier_date) AS shipped_orders,  
    COUNT(order_delivered_customer_date) AS delivered_orders  
  FROM `ecommerce_data.orders`  
)  
SELECT  
  total_orders, approved_orders, shipped_orders, delivered_orders,  
  approved_orders / total_orders AS pct_approved,  
  shipped_orders / approved_orders AS pct_shipped,  
  delivered_orders / shipped_orders AS pct_delivered  
FROM funnel;
```

Finding 2 — Average Delivery Time Is 12.1 Days

While the funnel shows high completion rates, the time required to move through that funnel tells a different story. The average time between order placement and customer delivery is **12.09 days** — a meaningful lag for an e-commerce experience in which customers increasingly expect delivery within a few days.

This extended delivery window creates a structural disadvantage for customer satisfaction even when orders arrive on or before their estimated date. It also provides important context for the late delivery findings below: when the baseline is already 12 days, any additional delay is felt acutely by the customer.

Key SQL Query — Average Delivery Time

```
SELECT
  AVG(DATE_DIFF(order_delivered_customer_date,
               order_purchase_timestamp, DAY)) AS avg_delivery_days
FROM `ecommerce_data.orders`
WHERE order_delivered_customer_date IS NOT NULL;

-- Result: 12.09 days
```

Finding 3 — 8.1% of Orders Are Delivered Late

Comparing actual delivery dates to estimated delivery dates shows that 7,827 orders (8.11%) were delivered after their promised date. While 91.89% of orders arrive on time, the late-delivery segment is large enough — nearly 1 in 12 orders — to represent a meaningful and recurring customer experience problem.

Delivery Status	Orders	% of Delivered
On Time	88,649	91.89%
Late	7,827	8.11%

Key SQL Query — Late vs. On-Time Classification

```
SELECT
  CASE
    WHEN order_delivered_customer_date > order_estimated_delivery_date THEN 'Late'
    WHEN order_delivered_customer_date <= order_estimated_delivery_date THEN 'On Time'
    ELSE 'Unknown'
  END AS delivery_status,
  COUNT(*) AS order_count,
  COUNT(*) / (SELECT COUNT(*) FROM `ecommerce_data.orders`
              WHERE order_delivered_customer_date IS NOT NULL) AS pct_of_total
FROM `ecommerce_data.orders`
WHERE order_delivered_customer_date IS NOT NULL
GROUP BY delivery_status
ORDER BY order_count DESC;
```

Finding 4 — Late Deliveries Dramatically Lower Customer Satisfaction

This is the most actionable finding in the analysis. Joining the orders table with order_reviews and segmenting by delivery status reveals a stark gap in customer satisfaction: on-time orders average a review score of **4.29 out of 5**, while late orders average just **2.57** — a drop of 1.72 points, or 40%.

Delivery Status	Avg Review Score	Review Count
On Time	4.29 / 5.00	88,658
Late	2.57 / 5.00	7,701

A 40% drop in review score for late orders is not a marginal finding — it is a strong signal that delivery timing is a primary driver of customer dissatisfaction. Even though only 8.1% of orders are late, those orders generate disproportionately negative customer experiences that likely suppress ratings, discourage referrals, and damage platform reputation.

Key SQL Query — Review Score by Delivery Status

```
WITH delivery_status AS (  
  SELECT o.order_id,  
         CASE  
           WHEN o.order_delivered_customer_date > o.order_estimated_delivery_date THEN 'Late'  
           WHEN o.order_delivered_customer_date <= o.order_estimated_delivery_date THEN 'On Time'  
           ELSE 'Unknown'  
         END AS delivery_status  
  FROM `ecommerce_data.orders` o  
  WHERE o.order_delivered_customer_date IS NOT NULL  
)  
SELECT d.delivery_status,  
       AVG(r.review_score) AS avg_review_score,  
       COUNT(*) AS review_count  
FROM delivery_status d  
JOIN `ecommerce_data.order_reviews` r ON d.order_id = r.order_id  
GROUP BY d.delivery_status  
ORDER BY avg_review_score DESC;
```

Finding 5 — Delivery Delays Do Not Directly Reduce Repeat Purchases

The most analytically interesting result of this project: customers who received late deliveries actually show a *slightly higher* repeat purchase rate (4.81%) compared to on-time customers (3.40%). This appears to contradict the strong negative impact on review scores seen in Finding 4.

Delivery Status	Customers	Repeat Customers	Repeat Rate
On Time	85,896	2,917	3.40%

Late	7,772	374	4.81%
------	-------	-----	-------

IMPORTANT — Selection Bias: Customers who place multiple orders are more likely to experience at least one late delivery simply because they have more orders in the system. This inflates the late-delivery repeat rate artificially. The result should NOT be interpreted as evidence that late deliveries are neutral or beneficial to retention.

The correct interpretation is that delivery performance alone cannot account for the 97% one-time buyer rate identified in Report 1. The low retention rate is likely driven by structural marketplace factors — customers using Olist for one-off purchases, limited loyalty mechanisms, and the nature of the product categories (gifts, specialty items) rather than fulfillment failures alone.

Key SQL Query — Repeat Rate by Delivery Experience

```
WITH delivery_status AS (
  SELECT o.order_id, c.customer_unique_id,
         CASE WHEN o.order_delivered_customer_date > o.order_estimated_delivery_date
              THEN 'Late' ELSE 'On Time' END AS delivery_status
  FROM `ecommerce_data.orders` o
  JOIN `ecommerce_data.customers` c ON o.customer_id = c.customer_id
  WHERE o.order_delivered_customer_date IS NOT NULL
),
customer_orders AS (
  SELECT c.customer_unique_id, COUNT(DISTINCT o.order_id) AS total_orders
  FROM `ecommerce_data.orders` o
  JOIN `ecommerce_data.customers` c ON o.customer_id = c.customer_id
  GROUP BY c.customer_unique_id
)
SELECT d.delivery_status,
       COUNT(DISTINCT d.customer_unique_id) AS customers,
       COUNT(DISTINCT CASE WHEN co.total_orders > 1 THEN d.customer_unique_id END) AS repeat_customers,
       COUNT(DISTINCT CASE WHEN co.total_orders > 1 THEN d.customer_unique_id END)
       / COUNT(DISTINCT d.customer_unique_id) AS repeat_rate
FROM delivery_status d
JOIN customer_orders co ON d.customer_unique_id = co.customer_unique_id
GROUP BY d.delivery_status;
```

Synthesis — Connecting Both Reports

Taken together, the two Olist analyses tell a coherent story about the business's core challenge and where to focus improvement efforts.

Report 1 Finding	Report 2 Connection
97% one-time buyers	Delivery delays hurt satisfaction but don't fully explain retention — structural marketplace factors are likely dominant
Top 10% drive 38% of revenue	Protecting high-value customers from late delivery experiences is critical — a 4.29→2.57 score drop risks losing your most valuable segment
\$159.83 avg order value	12-day avg delivery is misaligned with mid-range pricing expectations; faster delivery could justify and support higher AOV
Health, Gifts, Lifestyle lead revenue	Gift and lifestyle purchases are inherently time-sensitive — late delivery is especially damaging in these categories

Strategic Recommendations

Reduce Late Deliveries — Target the 8.1%

With review scores dropping from 4.29 to 2.57 for late orders, even a 50% reduction in the late rate (from 8.1% to ~4%) would meaningfully improve overall platform satisfaction. Priority actions: tighter seller SLA enforcement, more accurate delivery date estimation, and proactive customer communication when delays occur.

Address the 12-Day Delivery Baseline

An average of 12 days between order and delivery is a structural disadvantage in modern e-commerce. Investing in logistics partnerships, regional fulfillment centers, or seller incentives for faster processing could compress this window significantly — improving customer experience even without reducing the late rate.

Build Loyalty Mechanisms — Retention Is the Real Problem

Since delivery performance does not fully explain the 97% one-time buyer rate, retention strategies need to address the underlying marketplace behavior. Loyalty programs, personalized re-engagement emails, and subscription or bundle offers targeting repeat-purchase intent are higher-leverage than operational fixes alone.

Protect High-Value Customers From Late Experiences

The top 10% of customers drive 38% of revenue. Implementing priority fulfillment or proactive service recovery (automatic discount or credit for late delivery) for this segment would protect the revenue concentration identified in Report 1 while improving satisfaction scores where it matters most.

Conclusion

The order fulfillment funnel is operationally sound — over 98% conversion at every stage and a 97% overall delivery rate. But operational efficiency is not the same as customer experience quality. A 12-day average delivery window and an 8.1% late rate produce a measurable and significant satisfaction penalty, with review scores nearly halved for affected customers.

At the same time, this analysis clarifies an important limit: delivery problems alone cannot account for a 97% one-time buyer rate. The retention challenge is larger than fulfillment. It reflects the nature of marketplace purchasing — customers who arrive for a specific product and have no built-in reason to return. Solving retention requires both operational improvement and deliberate loyalty strategy.

Together, these two reports map the full picture: a business with strong revenue fundamentals, an efficient operational backbone, a significant customer satisfaction vulnerability in delivery timing, and a structural retention gap that represents the largest untapped growth opportunity.