Improve the share of seamless orders

Problem Statement:

Logistics management team wants to understand the causes of poor order experience, and would like data-driven insights to guide and improve the share of seamless orders.

An order is considered non-seamless if it has one or more of these "issues":

- Cancellations – which can be done by customer, rider, vendor, or app failures
- Delays – deliveries being delayed too much compared to originally promised time
- Customer Service Contacts – customers reaching out for help resolving issues

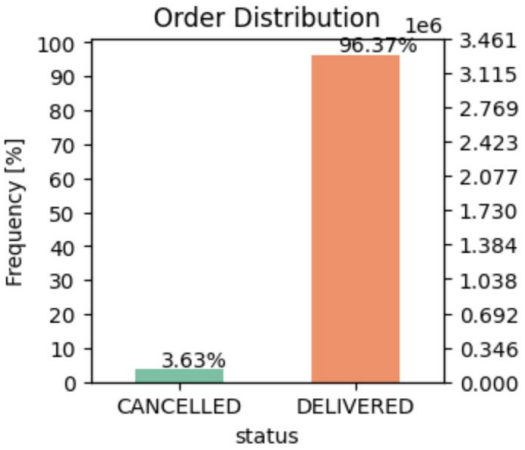
Data Exploration and Preprocessing:

- **Data Set:** Orders, Cancellations, Contacts, Deliveries

Basic Statistical Measures:
Order data: March 2021

- **Total Orders:** 3460862
- **Country Code:**
  - MY: 1706415
  - TH: 1754447
- **Order Distribution:**
  - Delivered: 3335142
  - Cancelled: 125720
- **Customer Support Contact:**
  - Served Chats: 94.7%
  - Missed Chats: 5.3%

*orders - contact: 1.65%
*cancelled - contact: 9.3%

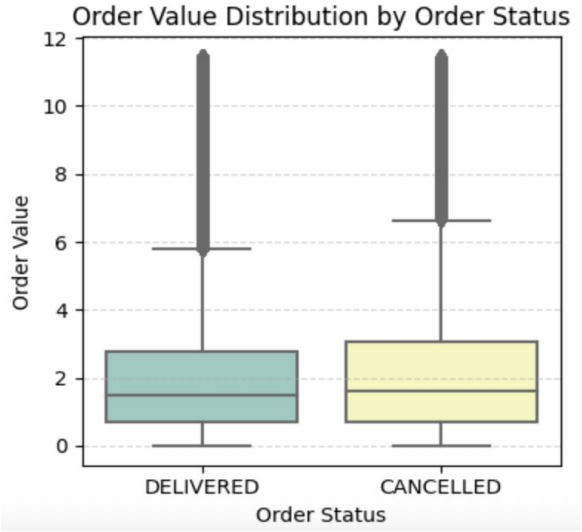


* ~3.6% of all orders were cancelled

* Restaurants drive 96% of all orders, while Local Shops and Groceries together contribute just 4%

* Mean order value for **cancelled orders is higher** than for delivered ones.

*** order value - ranging from 0 to 12



Seamless Vs Non-seamless orders

Non-Seamless Orders: Order issues - Cancellations, Delays, Customer Support Contacts
Classification order: CANCELLATION > DELAY > CONTACT (in that priority)

Assuming Delay Order: More than 10 minutes than promised delivery time
(Delay 10-15 minutes**; Noticeable delay,** but more than 15 minutes**; Long Delay**)

order_type	count	%
seamless	2970919	85.84%
non_seamless	489943	14.16%

delay_bucket	order_count	%
10-15 min late	141532	42.02
15-30 min late	141678	42.07
>30 min late	53583	15.91

issue_type	count	% of all orders	% of non_seamless
COMPLETED	2970919	85.84	0
DELAY	335850	9.7	68.55
CANCELLATION	125720	3.63	25.66
CONTACT	28373	0.82	5.79

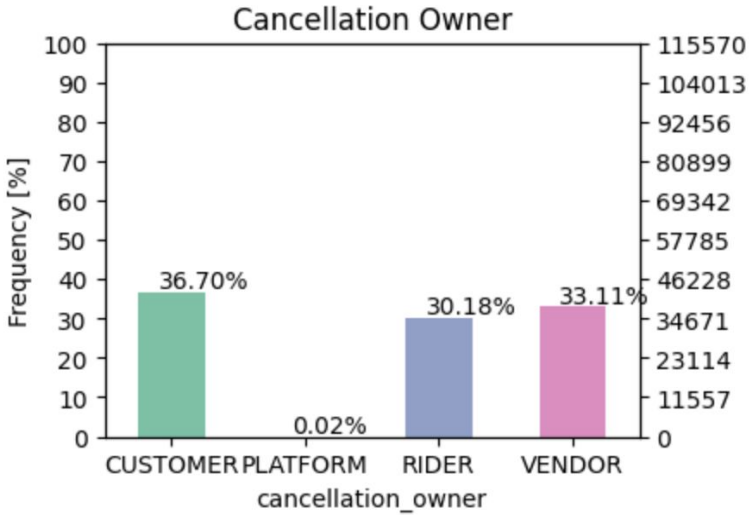
vertical	Groceries	Local shops	Restaurants
delay	7817	5808	323166
not late	57703	69867	2988519
delay %	**11.93%**	7.67%	9.76%
delay: >30 min late	**3.53%**	1.46%	1.51%

Delay is the top reason for poor order experiences — accounting for **~69% of all issues**;

Out of all delays, **~16% of delays are longer** — impacting **~11% of non-seamless orders** (30+ mins over promised)

Cancellation Share:

- Customer cancellations lead, but vendor-driven and rider cancellations are the top other contributors — **~60**% of all cancellations.
- Local Shops account for only **2%** of order volume, but **11%** of cancellations — with a **17.75%** cancellation rate, the highest among all verticals.
- Order being late / waiting time is the leading reason for customer contact for cancelled orders.



vertical	DELIVERED	CANCELLED	cancelled %
Groceries	63954	2003	3.04%
Local shops	63303	13657	17.75%
Restaurants	3207885	110042	3.32%

Top customer contact reasons
Cancelled Orders

contact_reason	count
Request: order will take longer than expected	1689
Request: order is late, does not want to wait	1246
Complain about late order	1227
Wrong order	1151
Delivery time/date	1075
Food items	851
Order marked as delivered but didn't receive	840
Request: changed mind	777
Contact without CR	676
Address	397

Recommendations:

- Address Long Delivery Delays; Reducing long delays (>30 min) (cancellations due to delays & groceries)
  How? → Flag orders predicted to be “at risk” and reroute early.
- Reduce vendor cancellations by flagging late preps (to mention local shops delivery cancellations)
  How? → Auto-assign delayed orders to backup vendors after timeout.
  → Restrict availability for vendors with frequent failures.
- Enable Customer-Side Interventions -> Customer Cancellations
  How? → Introduce reorder options featuring faster alternatives to stay rather than cancel.

Metric	Definition	Owner	Notes
Seamless Order Rate	Orders without any issues /Total orders	Vendor, Rider, Customer	Measures smooth order flow
Order Completion Rate	Orders delivered on or before promised time / Total orders	Vendor, Rider	Tracks timely deliveries
Cancellation Rate	Cancelled orders / Total orders	Vendor, Rider, Customer	Overall cancellations
Long Delivery Delays	Orders with long delay / Total orders	Vendor, Rider	Identifies delay issues
Reorder Rate	% of customers who reorder using faster alternatives after initial cancellation attempt	Customer	Measures effectiveness of intervention

Impact Estimation

Recommendation	Affected Orders	% of non-seamless orders	**Reduction %	Recovered Orders	Seamless Rate Lift %
Long Delays (>30 min)	53583	10.94%	30%	16075	+0.46%
Vendor Cancellations	38264	7.81%	30%	11479	+0.33%
Customer Cancellations	42342	8.64%	10%	4234	+0.12%

Recommendation	Impact	Notes
Long Delays (>30 min)	High	Platform-wide, predictive routing
Vendor Cancellations	Medium	Ops-led, Local Shops focus
Customer Reorder Option	Low	UX/product feature, preventive

*** Estimated impacts assume 10–30% reduction in issue volumes for directional analysis

*** Many cancellations are driven by long delays / long waiting times, though categorized separately.

Limitations & Assumptions

Assumptions:

- Each order was labeled with a single primary issue: CANCELLATION > DELAY > CONTACT (in that priority)
- Multiple contacts per order were aggregated and counted once
- Delay order: More than 5 minutes than promised delivery time (actual_delivery_duration_secs > promised_delivery_time)
- Impact estimates are based on assumed reduction rates (10–30%) for each issue type. Actual improvements may vary depending on implementation and operational constraints.

Limitations:

- Didn't deep dive in deliveries - timing and day of the week and cancellations data
- Needed to work more on customer contact reason - find more insights and severity analysis for contacts
- Issue types are mutually exclusive — some orders may have had multiple simultaneous issues
- Order value did not take into consideration for cancellations and neither for delays.
- Next: A/B test interventions