QUICKBITE EXPRESS: A DATA-DRIVEN RECOVERY

MADE BY : PIYUSH KUMAR JHA

PROBLEM STATEMENT

- QuickBite Express is a Bengaluru-based food-tech startup (founded in 2020) that connects customers with nearby restaurants and cloud kitchens. In June 2025, QuickBite faced a major crisis. A viral social media incident involving food safety violations at partner restaurants, combined with a week-long delivery outage during the monsoon season, triggered massive customer backlash. Competitors capitalized with aggressive campaigns, worsening the situation.
- Go through the metadata and analyse the datasets thoroughly. This is the most fundamental step.
- Begin the analysis by referring to the provided questions and datasets. You can use any tool of the choice (Python, SQL, Power BI, Tableau, Excel) to analyse and answer these questions
- Suggest some recovery strategies

DATASETS GIVEN

CLICK HERE FOR DATASETS

- dim_customer
- dim_delivery_partner_
- dim_menu_item
- dim_restaurant
- fact_delivery_performance
- fact_order_items
- fact_orders
- fact_ratings

DATA CLEANING

- **For Data Cleaning refer to the link below** 👇
- **Click Here**

EXECUTIVE SUMMARY

The Core Problem: 71% Revenue Collapse Driven by Operational Failure

- **The Collapse:** A massive, uniform 70.9% drop in revenue and orders across all major cities.
- **The Insight:** Average Order Value (AOV) remained stable. This proves the problem was not pricing, but a mass customer disengagement.
- **The Root Cause:** A complete operational breakdown. The average delivery delay more than doubled to 18.6 minutes (from 7.4).
- **The Failure:** The late delivery rate skyrocketed from 18% to 66.36% (two-thirds of all orders), even while handling only 30% of the original volume.

Immediate Consequences

- **Cancellations Doubled:** The fulfillment failure rate jumped from 6.06% to 11.93%.
- **Ratings Collapsed:** Average ratings fell from a stable 4.49 to a low of 2.58.
- **Top Complaints:** The 18.6-minute delay was the direct cause of the top complaints: "cold food," "late," and "bad quality.“

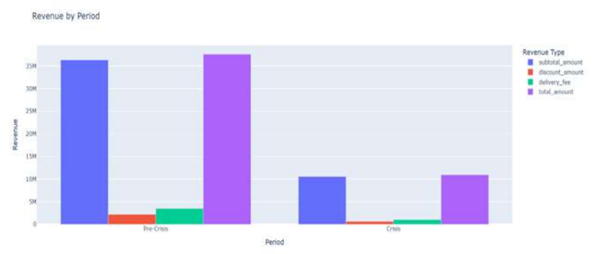
The Long-Term Damage: Catastrophic Loyalty Loss

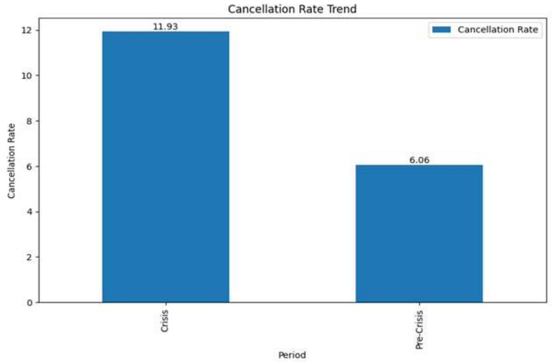
- **"Promoters" Wiped Out:** We saw a 93% churn rate (26 out of 28) of our best, high-frequency "promoter" customers.
- **Loyal Users Lost:** In total, the company lost 84.5% of all high-frequency users.

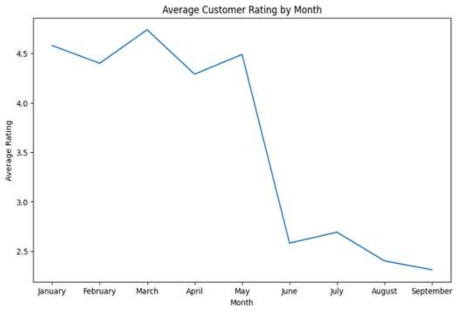
The Chain Reaction

- **Operational Failure:** Delivery delays doubled to 18.6 minutes.
- **Poor Experience:** Delays directly caused "cold food" complaints.
- **Reputation Collapse:** Ratings plummeted from 4.5 to 2.6.
- **Loyalty Wipeout:** 93% of our most valuable customers churned.

Executive Conclusion: The crisis exposed a fragile operational model. The failure to deliver orders on time was the single spark that started a fire, leading to a near-total wipeout of brand loyalty.





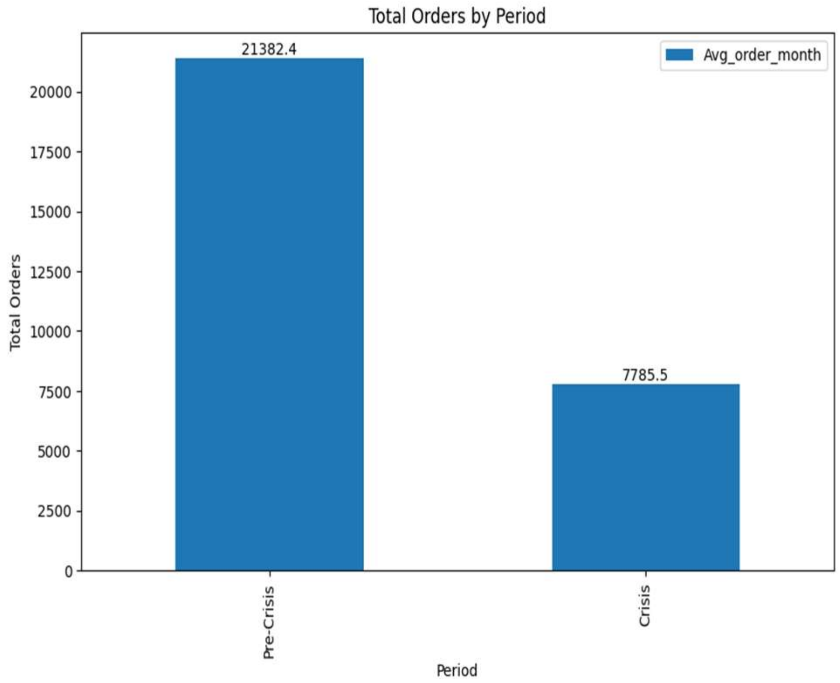


EXPLORATORY DATA ANALYSIS(EDA)

- MONTHLY ORDERS
- TOP 5 CITY GROUPS EXPERIENCED THE HIGHEST PERCENTAGE DECLINE IN ORDERS.
- RESTAURANT ANALYSIS
- CANCELLATION ANALYSIS
- DELIVERY SLA
- RATINGS FLUCTUATION
- SENTIMENT INSIGHTS
- REVENUE IMPACT
- LOYALTY IMPACT
- CUSTOMER LIFETIME DECLINE

EDA -1 COMPARE TOTAL ORDERS ACROSS PRE-CRISIS (JAN–MAY 2025) VS CRISIS (JUN–SEP 2025)

	month	total_orders
0	January	22076
1	February	21272
2	March	22133
3	April	20210
4	May	21221
5	June	8219
6	July	7768
7	August	7485
8	September	7670



EDA-1 INSIGHTS

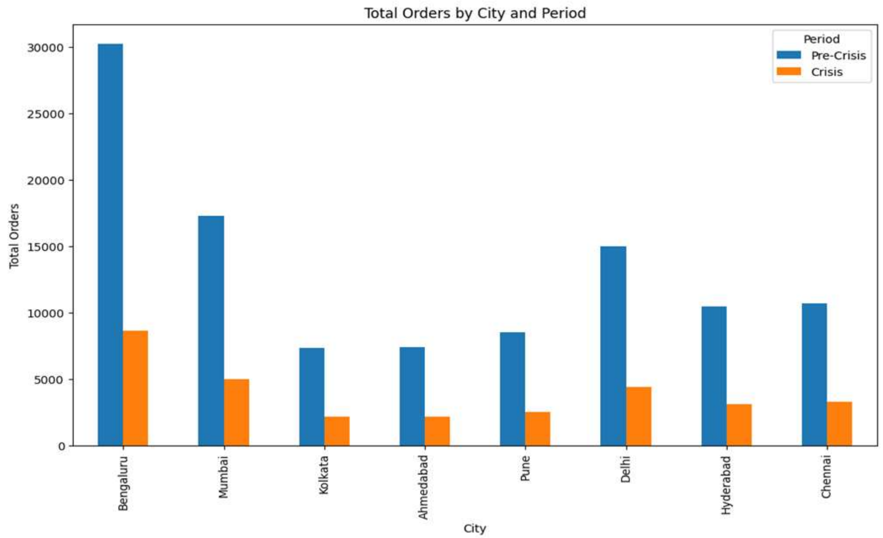
- **Pre-Crisis Average (Jan–May):** Orders were highly stable, averaging **21,382** per month.
- **The Drop (June):** Orders fell abruptly from 21,221 in May to **8,219** in June.
- **Magnitude of Decline:** The June drop alone represented a **61.2% loss** of monthly volume.
- **Crisis Average (June–Sept):** The volume stabilized at a much lower baseline, averaging **7,786** orders per month.

EDA -2 WHICH TOP 5 CITY GROUPS EXPERIENCED THE HIGHEST PERCENTAGE DECLINE IN ORDERS

- **Confirms Absolute Drop:** As we saw visually, **Bengaluru** had the largest *absolute* decline, losing 21,578 orders. This is followed by Mumbai (12,307) and Delhi (10,592), which were the three largest markets.
- **Confirms Relative Impact:** This table strongly supports the key insight that the crisis's impact was **remarkably uniform**. The "Decline %" for all top cities is tightly clustered between 69.32% and 71.40%.

Top 5 city with decline during the crisis period compared to the pre-crisis period are:

Period	city	Decline	Decline %
0	Bengaluru	21578	71.40
1	Mumbai	12307	71.30
2	Delhi	10592	70.62
3	Chennai	7394	69.32
4	Hyderabad	7389	70.50

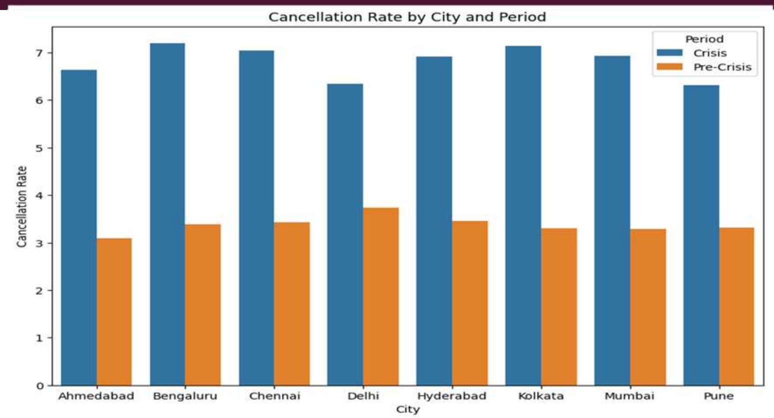


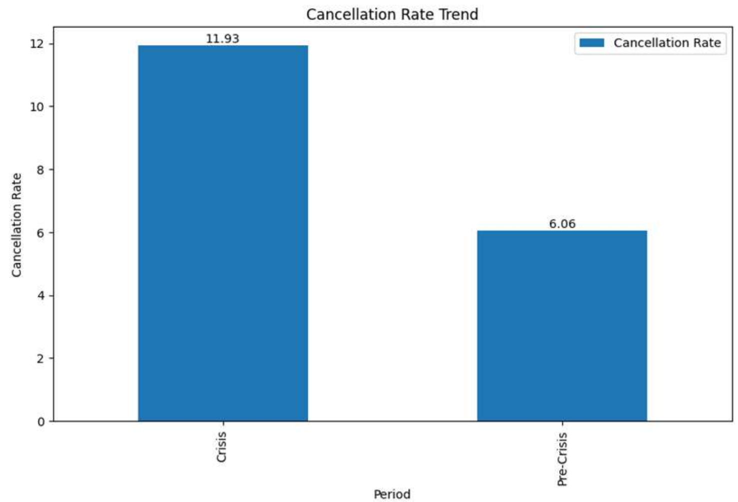
EDA-3 AMONG RESTAURANTS WITH AT LEAST 50 PRE-CRISIS ORDERS, WHICH TOP 10 HIGH-VOLUME RESTAURANTS EXPERIENCED THE LARGEST PERCENTAGE DECLINE IN ORDER COUNTS DURING THE CRISIS PERIOD?

Period	restaurant_name	Crisis	Pre-Crisis	Decline %
0	Fresh Tandoor Delight	4.0	54.0	92.59
1	Urban Kitchen Zone	8.0	63.0	87.30
2	Flavours of Tandoor Central	9.0	62.0	85.48
3	Classic Sweets Heaven	8.0	55.0	85.45
4	Grand Cafe Clouds	9.0	60.0	85.00
5	Punjabi Sweets Cafe	9.0	56.0	83.93
6	Hot & Crispy Mess Mahal	9.0	56.0	83.93
7	Thindi Mane Darshini Heaven	9.0	53.0	83.02
8	Punjabi Curry Delight	9.0	53.0	83.02
9	Spicy Express Bhojanalay	9.0	51.0	82.35

EDA-4 CANCELLATION ANALYSIS: WHAT IS THE CANCELLATION RATE TREND PRE-CRISIS VS CRISIS, AND WHICH CITIES ARE MOST AFFECTED?

- This indicates that for every 100 orders placed during the crisis, 12 failed. This isn't just lost revenue; it's an actively bad customer experience that damages brand trust and loyalty.
- This means that during the crisis, orders were being cancelled at almost **twice the rate** as before. This suggests significant operational issues, suchas restaurants being unable to fulfill orders, delivery partners being unavailable, or customers changing their minds due to uncertainty.
- While this was a system-wide failure, the operational breakdown was most severe in **Kolkata** and **Ahmedabad**, where cancellation rates exploded by over **114%**. This suggests the teams in those cities were the least prepared to handle the crisis, while **Delhi** proved to be the most resilient.





EDA-5 DELIVERY SLA: MEASURE AVERAGE DELIVERY TIME ACROSS PHASES. DID SLA COMPLIANCE WORSEN SIGNIFICANTLY IN THE CRISIS PERIOD?

- As per Industry standards delivery time threshold is **10–15-minute deviation** from the Estimated Time of Arrival (ETA) is typically seen as being within the customer's "zone of tolerance“, **so we took 10 min as the delivery threshold**.
- **Pre-Crisis:** The average time delay was **7.39 minutes**. This is a manageable operational variance.
- **Crisis:** The average time delay exploded to **18.60 minutes**.
- **More orders were late:** 66% of all deliveries missed the 10-min SLA window.
- **They were *much* later:** Those orders weren't just a little late; they were arriving, on average, **18.6 minutes** behind schedule.

'The Avg Time Delay'

Period	Abs_error
Crisis	18.60
Pre-Crisis	7.39

On Time Delivery %

```
threshold=10
count_ontime_pre_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Pre-Crisis'].Abs_error<=threshold).sum()
count_ontime_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Crisis'].Abs_error<=threshold).sum()
total_pre_crisis_orders=len(df_merge_delivery[df_merge_delivery['Period']=='Pre-Crisis'])
total_crisis_orders=len(df_merge_delivery[df_merge_delivery['Period']=='Crisis'])
on_time_precrisis_percentage=round((count_ontime_pre_crisis/total_pre_crisis_orders)*100,2)
on_time_crisis_percentage=round((count_ontime_crisis/total_crisis_orders)*100,2)
print('On Time Delivery % during pre-crisis within threshold (10 min) is',on_time_precrisis_percentage,'%')
print('On Time Delivery % during crisis within threshold (10 min) is',on_time_crisis_percentage,'%')
```

```
On Time Delivery % during pre-crisis within threshold (10 min) is 73.72 %
On Time Delivery % during crisis within threshold (10 min) is 33.64 %
```

Late Delivery %

```
count_delay_pre_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Pre-Crisis'].Delivery_diff<-threshold).sum()
count_delay_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Crisis'].Delivery_diff<-threshold).sum()
delay_precrisis_percentage=round((count_delay_pre_crisis/total_pre_crisis_orders)*100,2)
delay_crisis_percentage=round((count_delay_crisis/total_crisis_orders)*100,2)
print('Late Delivery % during pre-crisis after threshold (10 min) is',delay_precrisis_percentage,'%')
print('Late Delivery % during crisis after threshold (10 min) is',delay_crisis_percentage,'%')
```

```
Late Delivery % during pre-crisis after threshold (10 min) is 17.93 %
Late Delivery % during crisis after threshold (10 min) is 66.36 %
```

Early Delivery Rate%(reached super early above threshold)

```
count_early_pre_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Pre-Crisis'].Delivery_diff>threshold).sum()
count_early_crisis=(df_merge_delivery[df_merge_delivery['Period']=='Crisis'].Delivery_diff>threshold).sum()
early_precrisis_percentage=round((count_early_pre_crisis/total_pre_crisis_orders)*100,2)
early_crisis_percentage=round((count_early_crisis/total_crisis_orders)*100,2)
print('Early Delivery % during pre-cris is',early_precrisis_percentage,'%')
print('Early Delivery % during crisis is',early_crisis_percentage,'%')
```

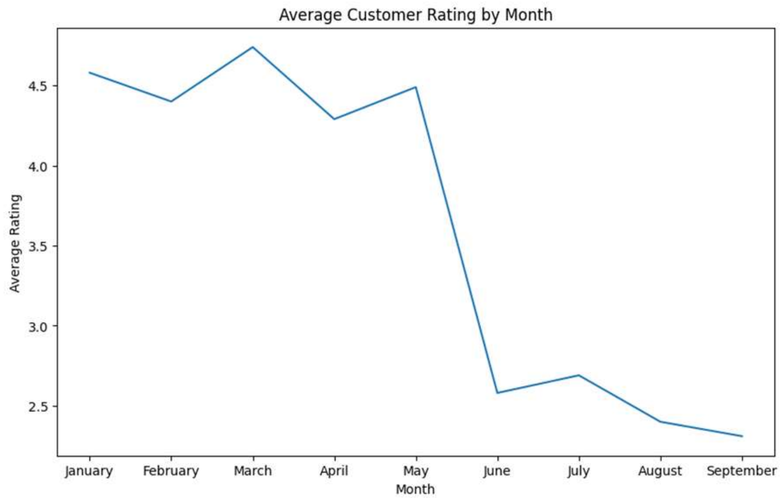
```
Early Delivery % during pre-cris is 8.36 %
Early Delivery % during crisis is 0.0 %
```

EDA -6 RATINGS FLUCTUATION: TRACK AVERAGE CUSTOMER RATING MONTH-BY-MONTH. WHICH MONTHS SAW THE SHARPEST DROP?

- The single sharpest drop in average customer ratings occurred between **May and June**.
- **May Rating:** 4.49
- **June Rating:** 2.58
- This represents a massive **drop of 1.91 points** in a single month.
- **Key Insight:** As the line graph visually confirms, customer ratings were stable and high (all above 4.2) from January to May. The drop from May to June is a "cliff-edge" fall, directly correlating with the "Crisis" period
- The data shows that customers were severely impacted by the service breakdown and responded with extremely low ratings, wiping out all the goodwill built up in the pre-crisis period.

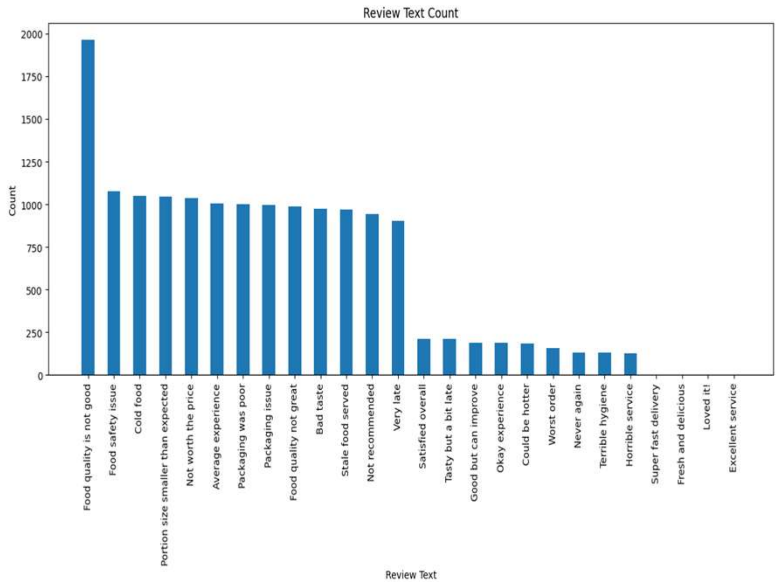
The Average Rating MoM is

month	rating
January	4.58
February	4.40
March	4.74
April	4.29
May	4.49
June	2.58
July	2.69
August	2.40
September	2.31



EDA-7 SENTIMENT INSIGHTS: DURING THE CRISIS PERIOD, IDENTIFY THE MOST FREQUENTLY OCCURRING NEGATIVE KEYWORDS IN CUSTOMER REVIEW TEXTS.

- In this I have used **Power BI Word Cloud visual** for identify the most frequently occurring negative keywords.
- Based on the word cloud, the most frequently occurring negative feedback is clustered around **food quality** and **delivery issues**.
- On Food Quality & Condition: Cold, Bad, Stale, Horrible, Worst, poor
- On Delivery & Experience: late, Terrible experience, Never again
- On Portioning & Value: smaller, Portion
- This sentiment analysis directly correlates with the other findings: customers were receiving their orders **late** (which explains "Cold" food) and were having a **"Terrible experience"** due to the high cancellation and delay rates.

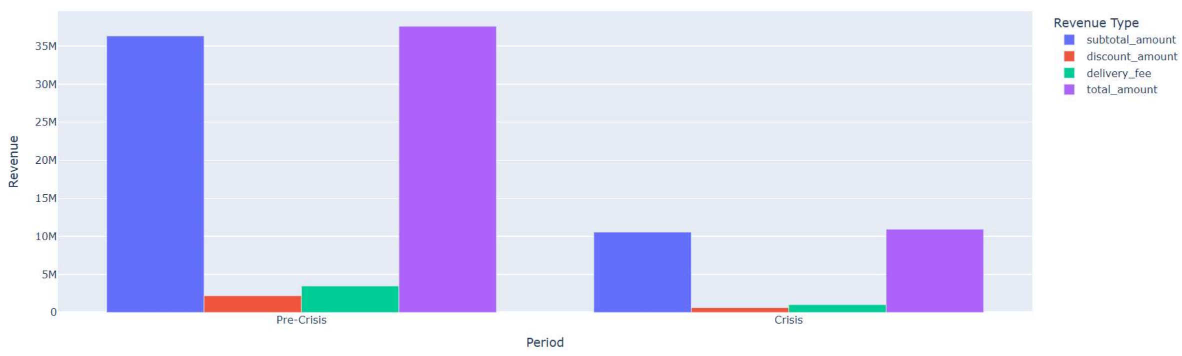


EDA-8 REVENUE IMPACT: ESTIMATE REVENUE LOSS FROM PRE-CRISIS VS CRISIS (BASED ON SUBTOTAL, DISCOUNT, AND DELIVERY FEE).

- The revenue loss was catastrophic, dropping by **26,680,812.94**, which represents a **70.92%** decline.
- The Most Critical Insight: AOV Remained Stable
  - **Order Volume Decline:** ~70-71%
  - **Revenue Decline:** 70.92%
- The revenue collapse was **not** because customers were spending less per order. The problem was not pricing or value. The revenue loss was a direct, proportional symptom of two core problems:
  - **Massive Drop in Demand:** A ~70% drop in order volume.
  - **Operational Failure:** A doubling of the cancellation rate (to 11.93%), meaning that even when you *did* get an order, you failed to fulfill it 1 out of 8 times, wiping out that potential revenue.

Period	subtotal_amount	discount_amount	delivery_fee	total_amount
Pre-Crisis	36338591.91	2190304.60	3472677.74	37620964.25
Crisis	10555201.30	627678.17	1012628.18	10940151.31

Revenue by Period




```
revenue_loss=round(df_revenue['total_amount'][0]-df_revenue['total_amount'][1],2)
print('Revenue loss from pre-crisis vs crisis is',revenue_loss,'which is',round((revenue_loss/df_revenue['total_amount'][0])*100,2),'%')
```

```
Revenue loss from pre-crisis vs crisis is 26680812.94 which is 70.92 %
/tmp/ipython-input-1838731011.py:1: FutureWarning:
```

EDA-9 LOYALTY IMPACT: AMONG CUSTOMERS WHO PLACED FIVE OR MORE ORDERS BEFORE THE CRISIS, DETERMINE HOW MANY STOPPED ORDERING DURING THE CRISIS, AND OUT OF THOSE HOW MANY HAD AN AVERAGE RATING ABOVE 4.5?

- **Mass Churn of Loyal Users (84.5%):**
  - There was a core group of **58** loyal customers who had each placed 5 or more orders before the crisis.
  - Of those 58, **49 stopped ordering** during the crisis.
  - This is an **84.5% churn rate** among most active, high-frequency users.
- **Near-Total Loss of "Promoters" (93%):**
  - Within that loyal group, there were **28** "promoters"— best customers who both ordered frequently *and* gave high ratings (> 4.5).
  - Of those 28, **26 stopped ordering**.
  - This is a **93% churn rate** among biggest fans and advocates.

```
df_grouped_cust=df_orders.groupby(['Period','customer_id'])['order_id'].count()
df_grouped_cust=df_grouped_cust.rename('total_orders')
df_grouped_cust=df_grouped_cust.reset_index()
df_grouped_cust_precrisis=df_grouped_cust[df_grouped_cust['Period']=='Pre-Crisis']
df_grouped_cust_precrisis=df_grouped_cust_precrisis[df_grouped_cust_precrisis['total_orders']>=5]
cnt_precrisis=df_grouped_cust_precrisis['customer_id'].count()
print('Users placed 5 or more than 5 orders before the crisis are',cnt_precrisis)
```

```
Users placed 5 or more than 5 orders before the crisis are 58
```

```
still_ordering=df_grouped_cust_crisis['customer_id'].count()
print('Users stopped ordering after crisis are',cnt_precrisis-still_ordering)
```

```
Users stopped ordering after crisis are 49
```

```
df_stopped_rating=pd.merge(stopped_order,df_ratings,on='customer_id')
df_stopped_rating=df_stopped_rating.groupby('customer_id')['rating'].mean().round(2)
count_high_rating_stopped = df_stopped_rating[df_stopped_rating > 4.5].count()
print('Number of users who stopped ordering and had an average rating above 4.5:', count_high_rating_stopped)
```

```
Number of users who stopped ordering and had an average rating above 4.5: 26
```

```
high_rating=pd.merge(df_grouped_cust_precrisis,df_ratings,on='customer_id')
a=high_rating.groupby('customer_id')['rating'].mean().round(2)
count_high_rating = a[a > 4.5].count()
print('Number of users who had an average rating above 4.5 with 5 or more than 5 orders:', count_high_rating)
```

```
Number of users who had an average rating above 4.5 with 5 or more than 5 orders: 28
```

EDA-10 CUSTOMER LIFETIME DECLINE: WHICH HIGH-VALUE CUSTOMERS (TOP 5% BY TOTAL SPEND BEFORE THE CRISIS) SHOWED THE LARGEST DROP IN ORDER FREQUENCY AND RATINGS DURING THE CRISIS? WHAT COMMON PATTERNS?

- **Pattern 1: Location (City)** The churn of HVCs was concentrated in the largest, most important markets.
  - **Bengaluru:** 2,726 HVCs were lost
  - **Mumbai:** 1,755 HVCs were lost
  - **Delhi:** 1,738 HVCs were lost This shows the operational failure was worst in the very cities that were the engine of the business.
- **Pattern 2: Cuisine Preference** The data shows a clear preference among both churned HVCs and the retained-but-unhappy HVCs.
  - **Top Cuisines for Churned HVCs:** Biryani (1,327), Chinese (970), North Indian (603).
  - **Top Cuisines for Unhappy Retained HVCs:** Biryani (91), Chinese (65), North Indian (43).
  - **Insight:** The pattern is identical. The problem was **not** the cuisine type. The problem was a complete failure to **deliver these popular cuisines** acceptably.
- **Pattern 3: The Service Failure (The "Why")** This is the root cause. The "delivery delays" and other issues are why they left. The most common complaints from these HVCs are:
  - **Food quality is not good** (41)
  - **Food safety issue** (24)
  - **Cold food** (17)
  - **Very late** (15)

The count of High LTV churned User City Wise

city	count
Bengaluru	2726
Mumbai	1755
Delhi	1738
Chennai	1158
Hyderabad	1124
Kolkata	1060
Ahmedabad	940
Pune	901

The Top cuisine Prefernce of High LTV churned User

Top_Cuisine_Preference	count
Biryani	1327
Chinese	970
North Indian	603
Desserts	408
Fast Food	369
Pizza	232
South Indian	215
Healthy	164

'Most common review during crisis by high LTV retained customers'

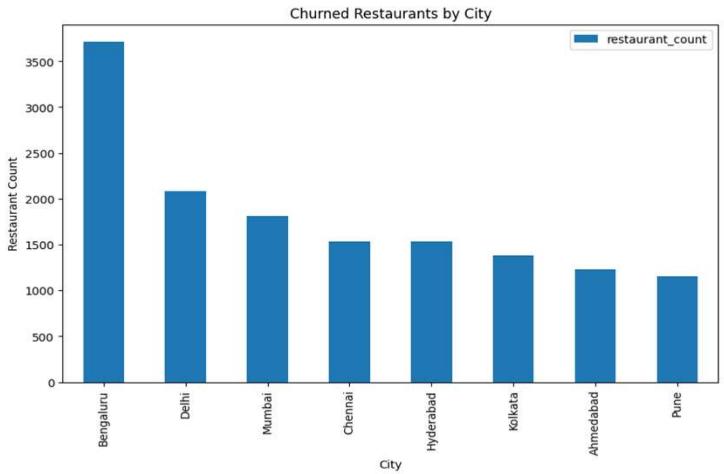
review_text	count
Food quality is not good	41
Food safety issue	24
Not worth the price	22
Food quality not great	22
Average experience	21
Packaging was poor	21
Not recommended	21
Packaging issue	21
Bad taste	20
Portion size smaller than expected	19
Cold food	17
Very late	15
Stale food served	15
Good but can improve	9
Worst order	5
Okay experience	4
Could be hotter	3
Satisfied overall	3
Tasty but a bit late	3
Terrible hygiene	3
Never again	2

Top ordered Cuisine by retained High LTV customers where review was extremely low

Top_Cuisine_Preference	count
Biryani	91
Chinese	65
North Indian	43
Desserts	31
Fast Food	24
South Indian	16
Healthy	13
Pizza	12

EDA-11 RESTAURANT CHURN ANALYSIS

- Widespread Churn in Key Markets (especially Tier 1 Cities)
  - **Bengaluru:** 3,717 restaurants churned.
  - **Delhi:** 2,085 restaurants churned.
  - **Mumbai:** 1,815 restaurants churned.
- Loss of High-Volume Partners
  - The "**High Order Volume Restaurant Churn**" table proves this isn't just small, insignificant partners leaving.
    - You lost **147 high-volume restaurants** (with >15 orders each), including key accounts like "Punjabi Curry Delight" (34 orders) and "Royal House Palace" (33 orders).
- Consolidated Insight
  - The operational failure (66% late orders, 12% cancellations) created a toxic cycle
    - Customers left, so order volume collapsed by 70%.
    - The few remaining orders resulted in "cold food" and "bad quality" reviews.
    - his data proves the final step: **Restaurants left the platform** because it brought them **no volume** and actively **damaged their brand reputation.**



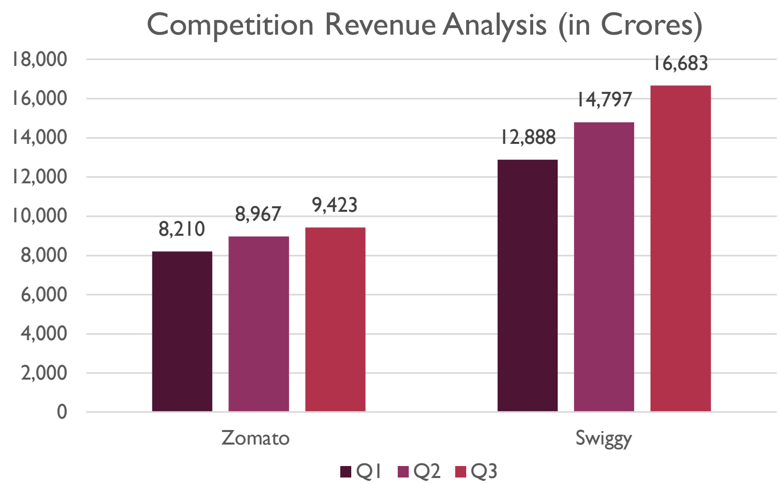
'High Order Volume Restaurant Churn'

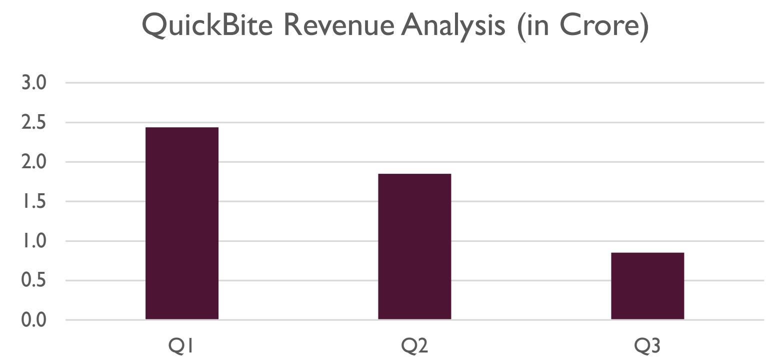
restaurant_name	order_id
Punjabi Curry Delight	34
Royal House Palace	33
Fresh Pizza Bhojanalay	29
Delhi Cafe Clouds	29
Hot & Crispy Thali Bhojanalay	29
...	...
Fresh Mess Point	16
Tandoori Delights Stop	16
Hot & Crispy Kitchen Hub	16
Thindi Mane Sweets Central	16
Grand Kitchen Heaven	16

147 rows × 1 columns

SUMMARY

- The revenue drop(**~70-71%**) and order drop(**70.92%**) were identical. This proves the **Average Order Value (AOV) was stable**. The problem wasn't pricing; it was a massive **disengagement** of the entire customer base.
- **SLA (Delivery) Failure**: The average delay more than doubled, from 7.4 minutes to **18.6 minutes** which means The late delivery rate skyrocketed from 18% to **66.36%**. (Two-thirds of all orders were late).
- The overall cancellation rate nearly doubled from **6.06% to 11.93% , Kolkata** and **Ahmedabad** saw the worst relative breakdown, with cancellation rates exploding by over **114%**.
- **Customer Experience:** Average ratings were stable and high (4.2-4.7) until May (4.49). In June, they fell off a cliff to **2.58**. The top complaints were **"Food quality issue," "Cold," "late," "Bad,"** and **"Terrible experience"**. The data proves that the **18.6-minute average delay** directly caused the "cold food" and "bad quality" complaints, which in turn caused the ratings to collapse.
- The company lost **84.5%** (49 out of 58) of its high-frequency customers (those with 5+ pre-crisis orders). The damage was even worse among the *best* customers. Of the 28 "promoters" (high-frequency users who also gave >4.5 ratings), **26 left**—a **93% churn rate** results in the **increase in CAC .**





RECOMMENDATIONS

- **Phase 1: Triage**

  Goal: Stop the churn. Fix the core product.

  - **Pause Ads:** Stop all new customer acquisition ads. You are paying to create new detractors.
  - **Fix Operations:** Re-route the ad budget to fix the **66% late rate**. Hire more partners and pay bonuses for on-time delivery.
  - **Launch Audits:** Immediately roll out a visible **"Safety Certified" badge** for restaurants to address "food safety" complaints.

- **Phase 2: Stabilize**

  Goal: Use your data to win back the customers you lost.

  - **Win Back Promoters (The 26):** Send a personal apology and a **100% free meal credit** to the 26 promoters you lost. A discount is not enough
  - **Retain Recoverable Customers (The 634):** Proactively send a service credit (e.g., "100 back") to the 634 high-value customers who are unhappy but still orders.
  - **Re-Engage Restaurants:** Call your top restaurants ("Fresh Tandoor Delight"). Help fix "cold food" issues and offer a temporary commission cut to rebuild the partnership.

- **Phase 3: Growth**

  Goal: Only after you have fixed the product, you can start marketing.

  - **Restart "Smart" Acquisition:** Once your average rating from retained customers is back above 4.0, you can slowly restart new customer ads. Your Customer Acquisition Cost (CAC) will naturally fall as your positive reviews will start supporting your marketing instead of fighting it.
  - **Diversify or Die.** Your data shows you are a food-only business in a 2025 market where your competitors (Swiggy/Zomato) are diversified logistics giants. Your core operation is too fragile. You *must* leverage your (now fixed) delivery network to enter **quick commerce (groceries)** and **pharmacy.** This diversifies your revenue and makes you less fragile to a "food-only" crisis.

THANK YOU