
# unORG

***Supply Chain + Data Science***

**Team Hyperion**

Krish Sharma | Rajan Verma | Rajarshi Verma | Rujhaan Taneja

# B2B grocery supply chains by order prediction, SKU demand forecast and planned inventory model for UNORG

*B2B grocery delivery platform*

**Manufacturers**

*order_id*
*order_date*

unORG

*poc_id*
*warehouse_id*

**Local vendors/ customers**

*customer_id*

**4100+** Unique customers

**90+** POCs

**600+** SKUs

## Deliverables

### 1 Daily Order Prediction

*c_id: 1837*

| *D1* | *D2* | *D3* | *D14* |
|---|---|---|---|
| **0.28** | **0.14** | **0.80** | **0.21** |

### 2 SKU Forecast

*c_id: 1837* | *SKU_id* | *order_date*

*14-day aggregated quantity forecast*: **40**

### 3 Inventory Planning

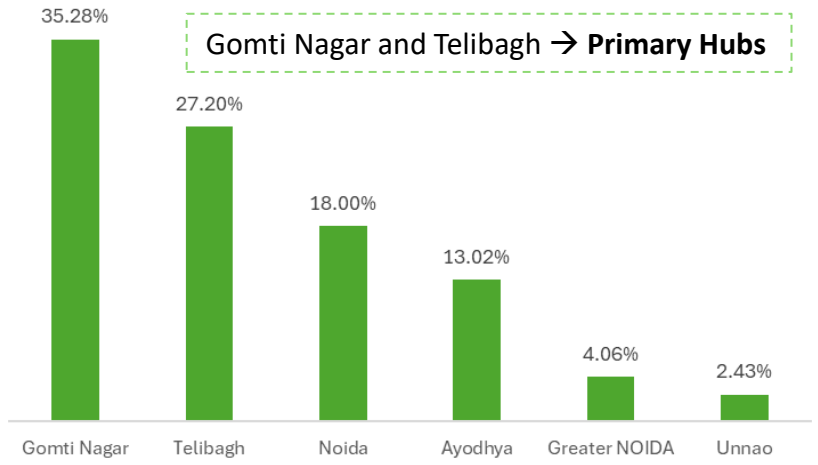
*Periodic review date**

*warehouse_id* | *SKU_id* | *Current_inventory* | *Target inventory*

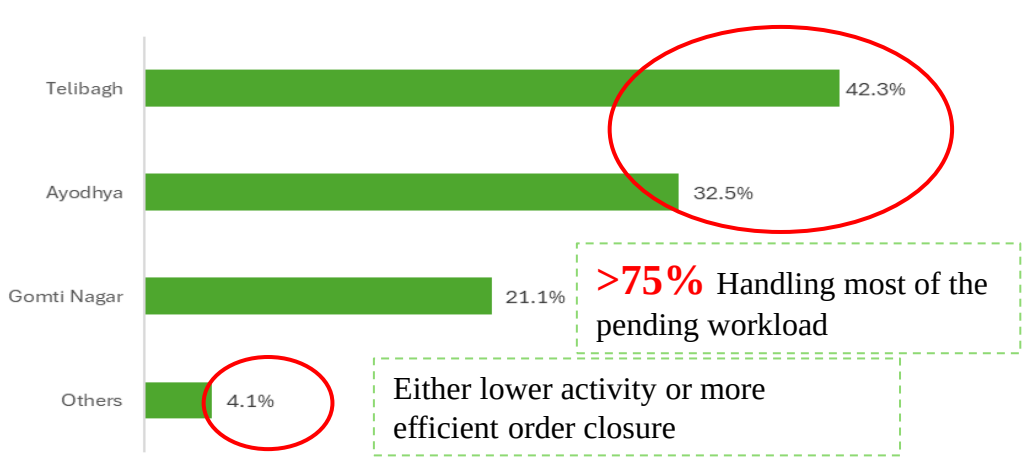
*Quantity*: **25**

# Gomti Nagar and Telibagh are the primary hubs for UnORG, strong weekday activity, and a notable post–New Year volume spike, highlighted operational efficiency and fulfillment trends across locations

## Order Distribution by Warehouse




## Distribution of OPEN Orders by Warehouse

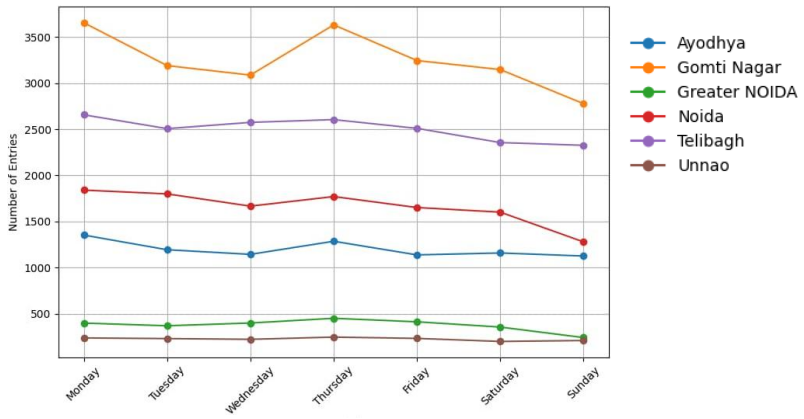



## General Observations

- Most open orders are from the Most Recent Date (10th April), sharp drop off for previous days (<1% before 8th April) → **System actively processing orders**
- High-value orders show exceptional fulfillment (98.9%), while general orders also perform well (95.76%), indicating possible **prioritization of high-value orders**

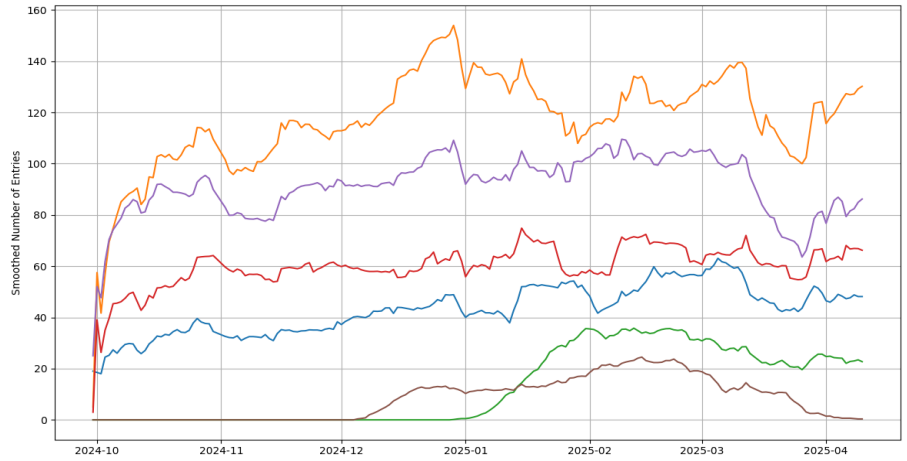
## Number of Entries per day for each Warehouse

- Gomti Nagar → **Busiest** warehouse throughout the week
- Monday, Thursdays → **Peak** Days
- Sundays → **Lowest**

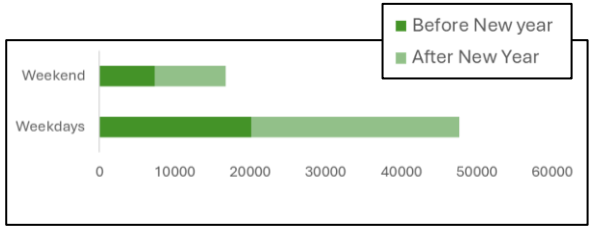



## 14-Day Rolling Average of Daily Entries per Warehouse

- Most warehouses show stable or growing trends, with no drastic drops except for **Unnao**, which tapers off sharply in early 2025




## Weekdays show higher order volumes compared to weekends




After New Year, there's a visible spike noted, particularly Thursday

# LGBM-predicted demand quantities for each customer-SKU-date combination in the 14-day forecast window (11th - 24th April)

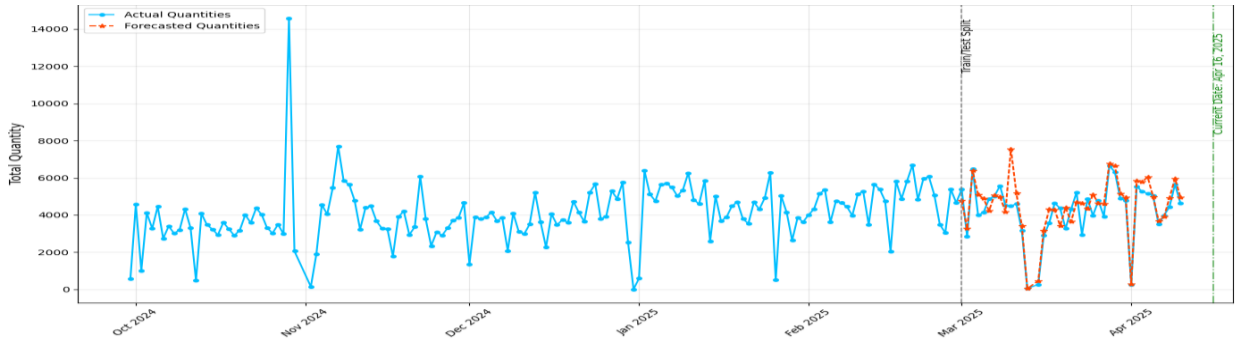
Orders.csv

| order_id | customer_id | order_date |
|---|---|---|
| ..... | poc_id | warehouse_id |

Items.csv

| order_id | quantity |
|---|---|
| ..... | order_item_id |

**Merged per customer per SKU basis**




**Daily Order Quantities**

30th Sept | 1st March (*Reference_date*) | 10th April | 24th April

TRAIN | VALIDATION | TEST

**Engineered 10+ features to capture Time trends, Calendar features, and discounts**

- *lag_1_day* — SKUs purchased in the prev. order
- *rolling_mean_7d* — Avg. qty across 7 most recent orders
- *discount_percentage* — Discount amount/ price per unit
- *is_holiday*, *day_of_week*, .....

```
LGBMRegressor
LGBMRegressor(learning_rate=0.05)
```

```
array([4.10740469e-04, 4.10740469e-04, 4.10740469e-04,
  ....  4.10740469e-04, 4.10740469e-04, 1.76423423e+00])
```

```
Test MAE: 0.007874720625264845
Test MSE: 0.6046870313062066
Test R2: 0.7080882290755617
```

| | customer_id | order_item_id | order_date | quantity |
|---|---|---|---|---|
| 0 | 5980 | 202967 | 2025-04-11 | 0.0 |
| 1 | 5980 | 202967 | 2025-04-12 | 0.0 |
| 2 | 5980 | 202967 | 2025-04-13 | 0.0 |
| 3 | 5980 | 202967 | 2025-04-14 | 0.0 |
| 4 | 5980 | 202967 | 2025-04-15 | 0.0 |
| ... | ... | ... | ... | ... |
| 1611297 | 7056 | 262487 | 2025-04-20 | 1.0 |
| 1611298 | 7056 | 262487 | 2025-04-21 | 1.0 |
| 1611299 | 7056 | 262487 | 2025-04-22 | 1.0 |
| 1611300 | 7056 | 262487 | 2025-04-23 | 1.0 |
| 1611301 | 7056 | 262487 | 2025-04-24 | 2.0 |

# Gomti Nagar and Telibagh are the primary hubs for UnORG, strong weekday activity, and a notable post–New Year volume spike, highlighted operational efficiency and fulfillment trends across locations

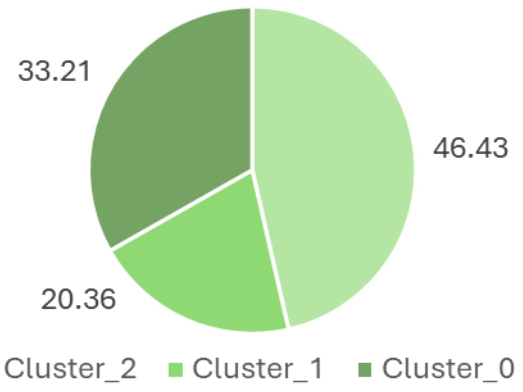
## Features engineering

- *average_order_gap* — Average time between orders
- *average_order_value* — (Total order value) / (No. of orders)
- *total_order_value*
- ...
- *Order Consistency Score* — *1/ (SD of order Gaps)*

**Clustered into 3 customer categories**

- Cluster_1 → total_orders <=4
- Cluster_0 → 4 < total_orders <= 25
- Cluster_2 → *average_order_gap >= 14*
- *else if not satisfied*

**% customers**




**Feature engineering**

### VALIDATION

**Cluster_1: Perfect**
- Accuracy: 51.50
- Precision: 0
- Recall: 0
- F1 Score: 0

**Cluster_2: Perfect**
- Accuracy: 14.03
- Precision: 83.53
- Recall: 4.25
- F1 Score:8.09

### TEST

**Cluster_1: Perfect**
- Accuracy: 51.93
- Precision0
- Recall: 0.0
- F1 Score: 0.0

**Cluster_1**
- Accuracy: 95.86
- Precision 47.95
- Recall: 50.0
- F1 Score: 48.95

**Cluster_2: Perfect**
- Accuracy: 14.24
- Precision: 86.08
- Recall: 4.35
- F1 Score: 8.27

**Cluster_2:**
- Accuracy: 82.13
- Precision: 84.048
- Recall: 51.91
- F1 Score: 48.85

**Cluster_1: Perfect**
- Accuracy: 51.93
- Precision: 7.14
- Recall: 0.01
- F1 Score: 0.02

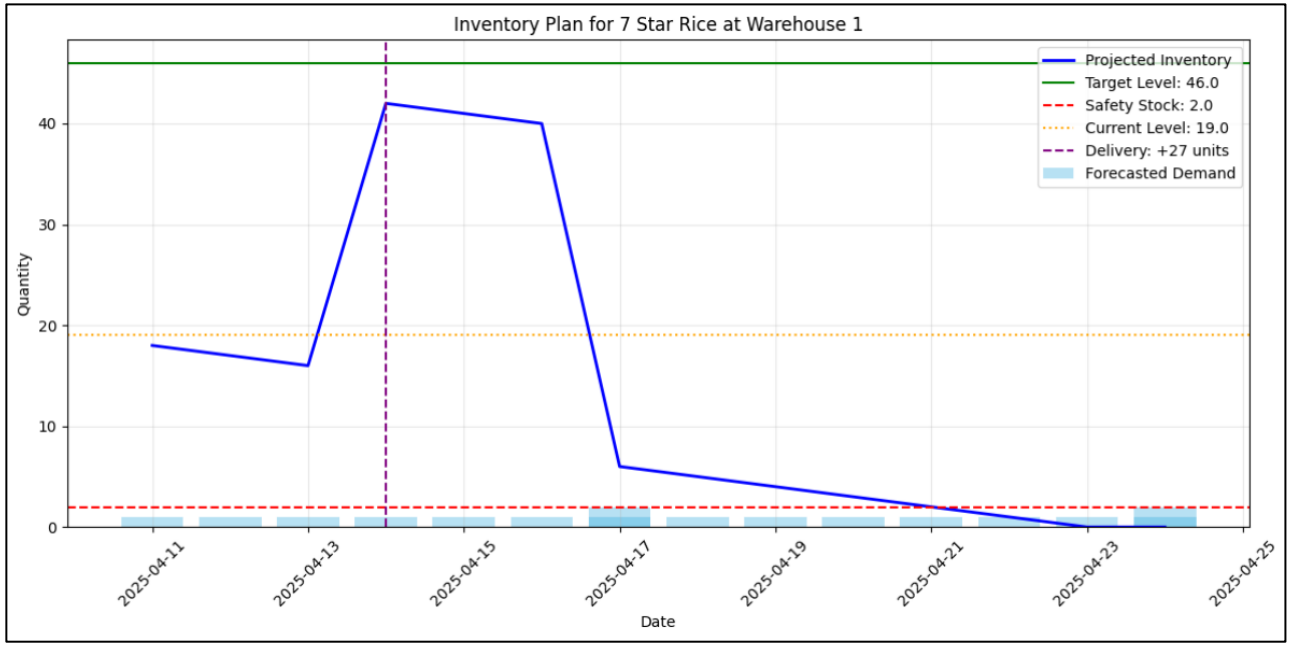
**Cluster_2: Perfect**
- Accuracy: 14.4
- Precision: 0.867
- Recall: 4.34
- F1 Score 0.826

| Layer (type) | Output Shape | Param # |
|---|---|---|
| lstm (LSTM) | (None, 30, 128) | 74,240 |
| dropout (Dropout) | (None, 30, 128) | 0 |
| lstm_1 (LSTM) | (None, 64) | 49,408 |
| dropout_1 (Dropout) | (None, 64) | 0 |
| dense (Dense) | (None, 32) | 2,080 |
| dropout_2 (Dropout) | (None, 32) | 0 |
| dense_1 (Dense) | (None, 14) | 462 |

**LSTM Model**

# PS3 – Inventory Model

**Implemented a Periodic Review (P-System) for each customer-product-warehouse combination**




**7 days**
**Review Interval**

**3 days**
**Lead Time**

```
Running periodic review on 2025-04-11

Orders to place on 2025-04-11:
     warehouse_id                    item_name  current_inventory  \
0               1                  7 Star Rice               19.0
1               1  ASHOK BIRYANI MASALA (50 GM)                2.0
2               1  ASHOK KALI MIRCH POWDER (100g)              4.0
3               1    ASHOK SONTH MASALA (100gm)                2.0
4               1               Adani Mota Besan              30.0
..            ...                          ...               ...
986             6            Tata Salt Pkt(1Kg)              32.0
987             6                 Tibaar Rice                6.0
988             6              Tuta Basmati Rice                1.0
989             6                Urad Sabut Dal               14.0
990             6          Vibhor Soya Pouch(1L)              117.0

     target_inventory  order_quantity
0                46.0              27
1                 3.0               1
2                 7.0               3
3                 4.0               2
4               105.0              75
..                ...             ...
986              53.0              21
987              15.0               9
988               5.0               4
989              21.0               7
990             173.0              56

[991 rows x 5 columns]

Next review scheduled for: 2025-04-18
```

# Thankyou

# EDA

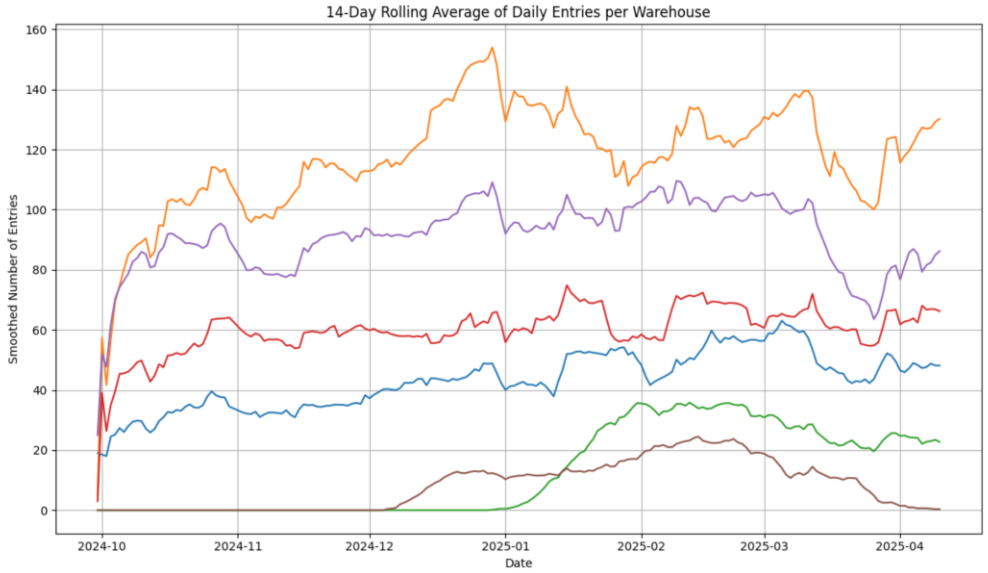
Engineered 20+ features across the dataset ← *order.csv & item_order.csv*

- *Total_order* :The total number of orders placed by the customer across the observed time period.
- *Average Order Gap Days* The average no. of days between two consecutive orders placed by the customer.
- ⋮
- *Total_order* The total number of orders placed by the customer indicates their overall purchase activity.
- *Order Consistency Score* measures how regularly and predictably a customer places orders over time.

**Major problems faced:**

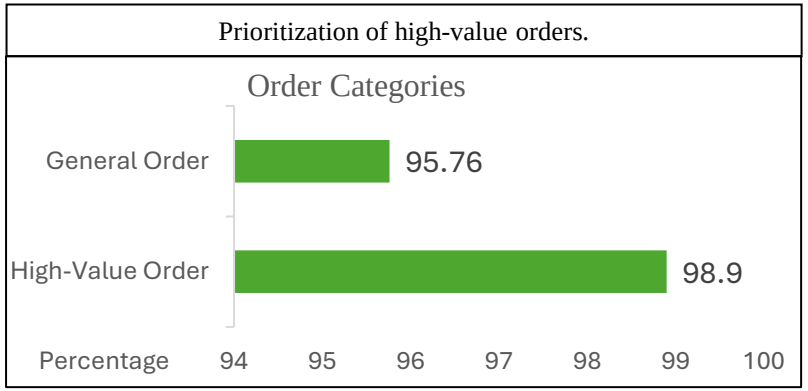
1. Same customers with multiple IDs:
   89 customers appeared with multiple IDs.
   Generate a **primary ID** by replacing each customer ID with the one having the most orders.

2. Same POC with multiple IDs:
   2 POCs appeared with multiple IDs.
   Generate a primary POC ID by replacing each POC ID with the one having the most orders.

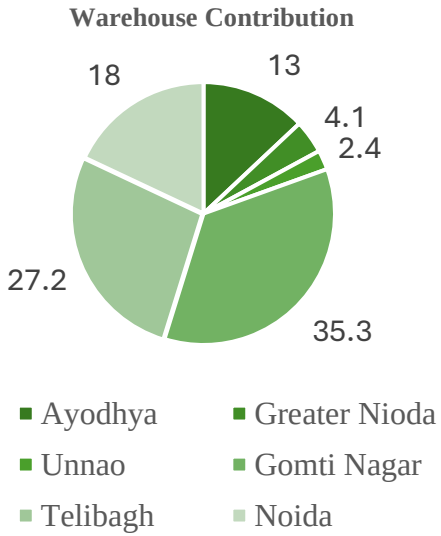



**Greater Noida** and **Unnao** warehouses were launched later, showing activity starting well after others.

**Warehouses near religious locations experienced noticeable entry spikes after the New Year.**

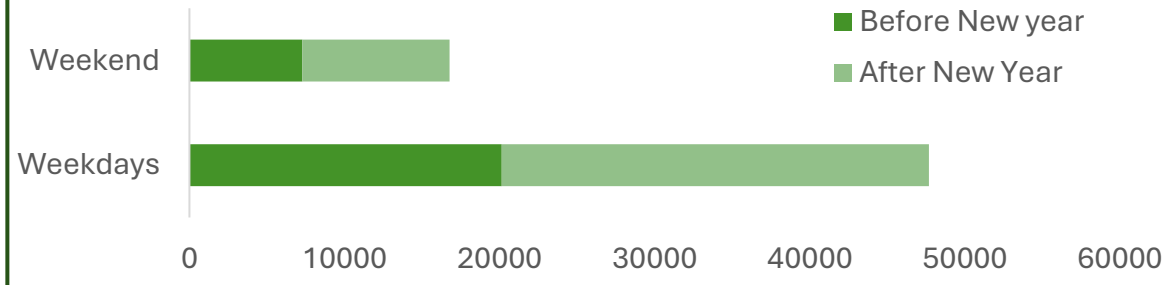



High-value orders show exceptional fulfillment (98.9%), while general orders also perform well (95.76%), indicating possible **prioritization of high-value orders**.




## Order activity is significantly higher after New Year across all days

Weekdays show higher order volumes compared to weekends.



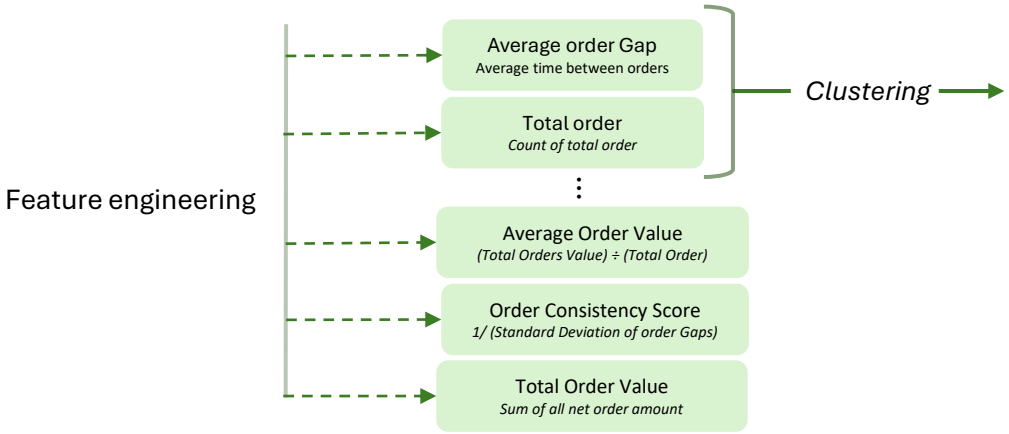

Overall most active day: Monday & Thursday

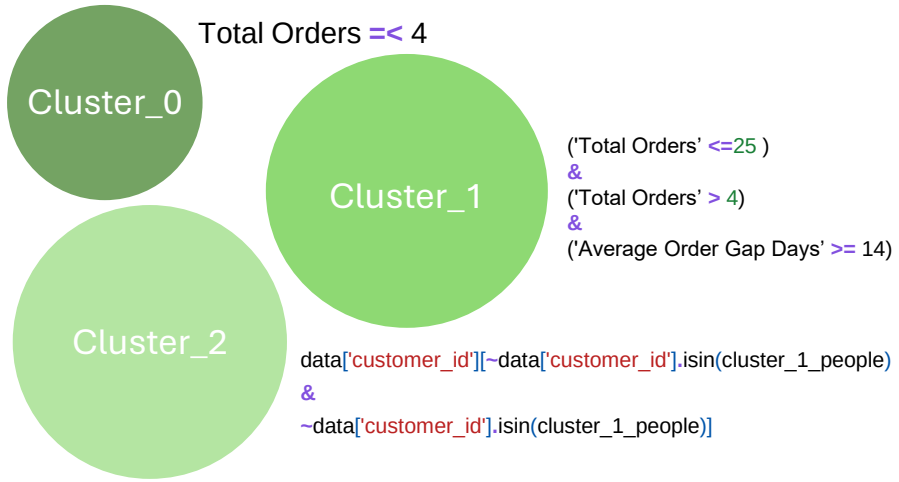
Overall Least active day: Sunday

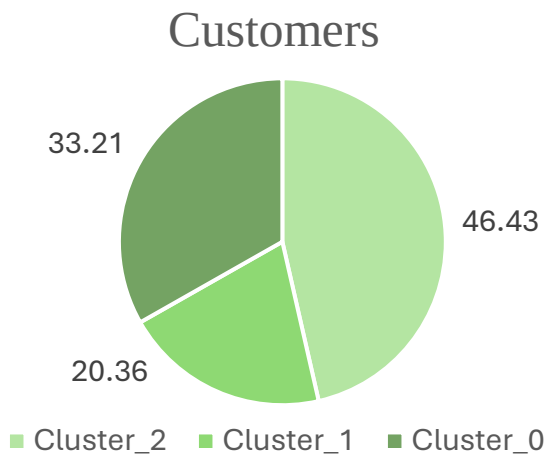
**After New Year**, there's a visible spike noted, particularly **Thursday**

# PS1 – LSTM based Daily Order Prediction Model

## *Order.csv dataset*




## *LSTM model Implementation*

cluster_x — Feature engineering — *Preprocessing for LSTM* →

| customer_id | 0 | 1 | 2 | ... | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|
| 1 | 0.29416 | 0.30853 | 0.31522 | ... | 0.31571 | 0.31441 | 0.31520 |
| 4 | 0.00000 | 0.00000 | 0.00000 | ... | 0.00000 | 0.00000 | 0.00000 |
| 5 | 0.22865 | 0.23321 | 0.23301 | ... | 0.24403 | 0.24347 | 0.25080 |
| 9 | 0.00000 | 0.00000 | 0.00000 | ... | 0.00000 | 0.00000 | 0.00000 |
| 12 | 0.00000 | 0.00000 | 0.00000 | ... | 0.00000 | 0.00000 | 0.00000 |
| 13 | 0.26654 | 0.30609 | 0.30896 | ... | 0.32041 | 0.31685 | 0.32203 |
| 14 | 0.23439 | 0.23842 | 0.23755 | ... | 0.24574 | 0.24513 | 0.25224 |
| 15 | 0.00000 | 0.00000 | 0.00000 | ... | 0.00000 | 0.00000 | 0.00000 |
| 18 | 0.20650 | 0.21019 | 0.21010 | .... | 0.22965 | 0.22955 | 0.23857 |

### TRAIN

| Layer (type) | Output Shape | Param # |
|---|---|---|
| lstm (LSTM) | (None, 30, 128) | 74,240 |
| dropout (Dropout) | (None, 30, 128) | 0 |
| lstm_1 (LSTM) | (None, 64) | 49,408 |
| dropout_1 (Dropout) | (None, 64) | 0 |
| dense (Dense) | (None, 32) | 2,080 |
| dropout_2 (Dropout) | (None, 32) | 0 |
| dense_1 (Dense) | (None, 14) | 462 |

**Cluster_1**
- Accuracy: 0.5192
- Precision: 0
- Recall: 0
- F1 Score: 0

**Cluster_2**
- Accuracy: 14.4
- Precision: 0.867
- Recall: 4.34
- F1 Score 0.826

### VALIDATION

**Cluster_1**
- Accuracy: 51.50
- Precision: 0
- Recall: 0
- F1 Score: 0

**Cluster_2**
- Accuracy: 14.03
- Precision: 83.53
- Recall: 4.25
- F1 Score:8.09

### TEST

**Cluster_1**
- Accuracy: 0.5185
- Precision 0
- Recall: 0
- F1 Score: 0

**Cluster_2: Perfect**
- Accuracy: 82.13
- Precision: 84.048
- Recall: 51.91
- F1 Score: 48.85

**Cluster_2**
- Accuracy: 14.24
- Precision: 86.08
- Recall: 4.35
- F1 Score: 8.27