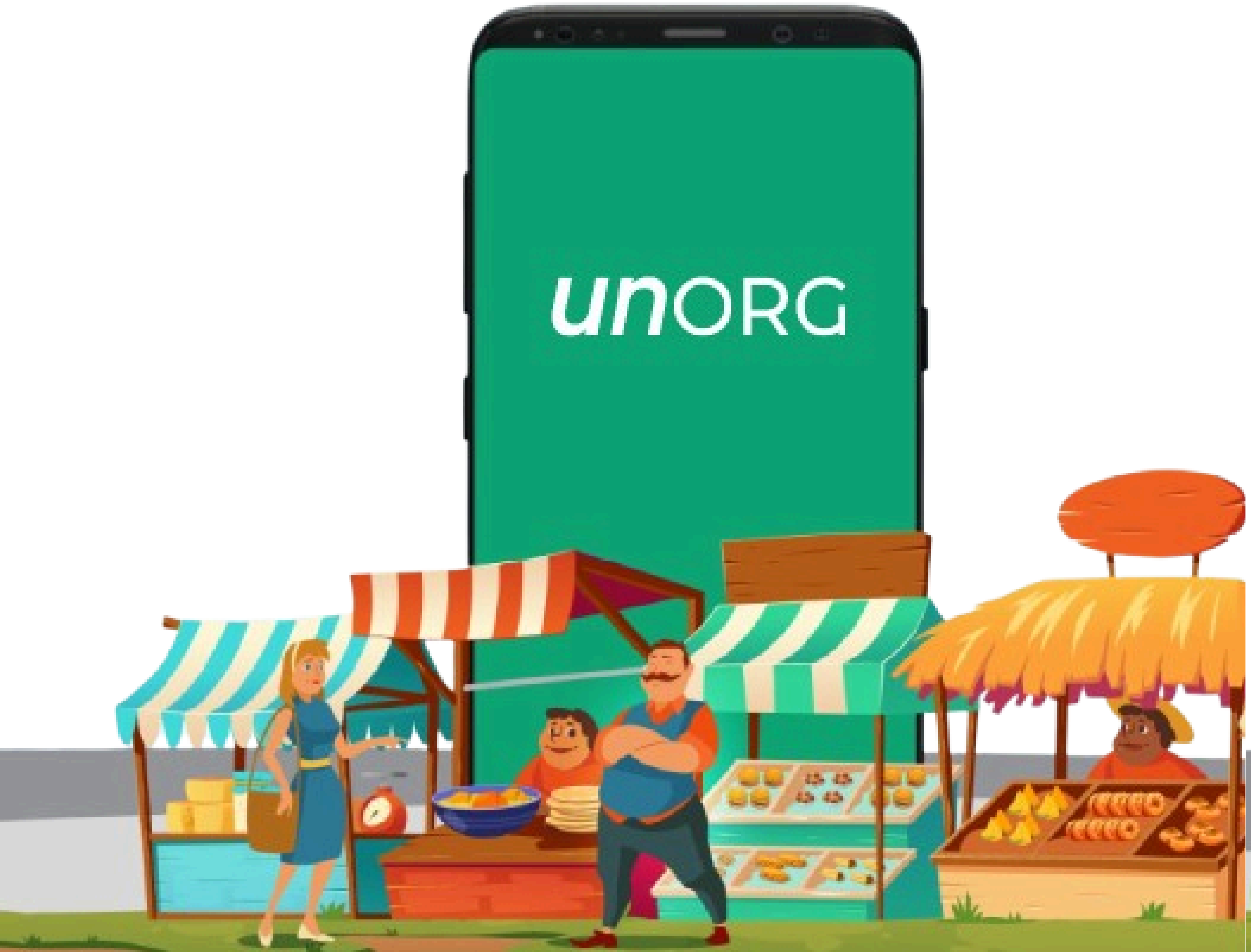
unORG

-Turning data into decisions — powering UNORG's growth through predictive intelligence.



Contingent - Bitmask

- Aman Kumar Kailaneya (23115010)
- Vanisha Choudhary(23117143)
- Hemant Bhakar(23113066)
- Parijat (23124020)

UnORG

UnOrg

UNORG is a fast-scaling B2B grocery delivery platform serving both traditional manufacturers (e.g., Petha, Daalmoth, Revdi makers) and modern F&B businesses (restaurants, cafés, hotels, and general stores).The platform connects these clients to essential goods with a strong emphasis on data-driven logistics, just-in-time delivery, and demand-supply alignment.

Insights

Top Customers (by number of orders):
- Bikaner Sweets (Harola) — 231 orders
- Dev Chole Bhature — 183 orders
- Ramesh Hotel — 175 orders

Top 10 Ordered Items:
- Ruchi Gold Palm Pouch (1L) – 6,945 orders
- Normal Sugar – 6,233 orders
- Prakash Maida – 4,952 orders

Orders with Negative Profit: 80,856
- ~70.25% of total orders
- Primarily from Gomti Nagar and Telibagh warehouses

High-discount, Loss-making Orders: 34,367
- Average Discount: ₹1,929.50
- Signs of over-discounting driving losses

Zero Invoiced Quantity : 3,576 items (~3.11%)
- Ramesh Hotel, Khushi Nainital Momos, and 93 Bataliyan CRPF among most affected

Statistics

Total Orders
115,093

Unique Customers
4,161

Unique SKUs
607

Average Items per Order
1.79

Warehouses
6

WWW/UNORG.IN

Problem Overview – Why Predicting B2B Grocery Demand is Challenging Yet Critical

B2B ≠ B2C

- **B2C**: Regular, small basket sizes, consistent user habits
- **B2B**: Irregular, bulk purchases, unpredictable cycles
- **UNORG's Clients**: From traditional sweet makers (e.g., Petha, Daalmoth) to modern restaurants, dhabas, and cafés—each behaves very differently

Unlike B2C platforms where behavior can be predicted based on regular app usage, in B2B, orders depend on external factors

Order Patterns of B2B

- B2B orders skip dates, vary across weekdays, and may cluster around specific occasions
- There are long gaps of inactivity, especially for seasonal sellers
- Order volumes can swing from 0 to 200+ units in a day

SKU Diversity and Client Preferences

UNORG's catalog spans hundreds of SKUs across grains, spices, oils, packaged foods but

- Each client has unique preferences based on their business type
- Restaurants might regularly order high volumes of cooking oil and spices
- General stores prefer diverse, smaller quantities of fast-moving consumer goods
- Preferences may shift seasonally or due to changing customer tastes

Inventory Misalignment

- Overstocking causes spoilage, unnecessary storage costs, and cash flow issues.
- Understocking results in lost orders, service delays, and reduced client satisfaction
- Higher logistics cost, due to reactive fulfillment .

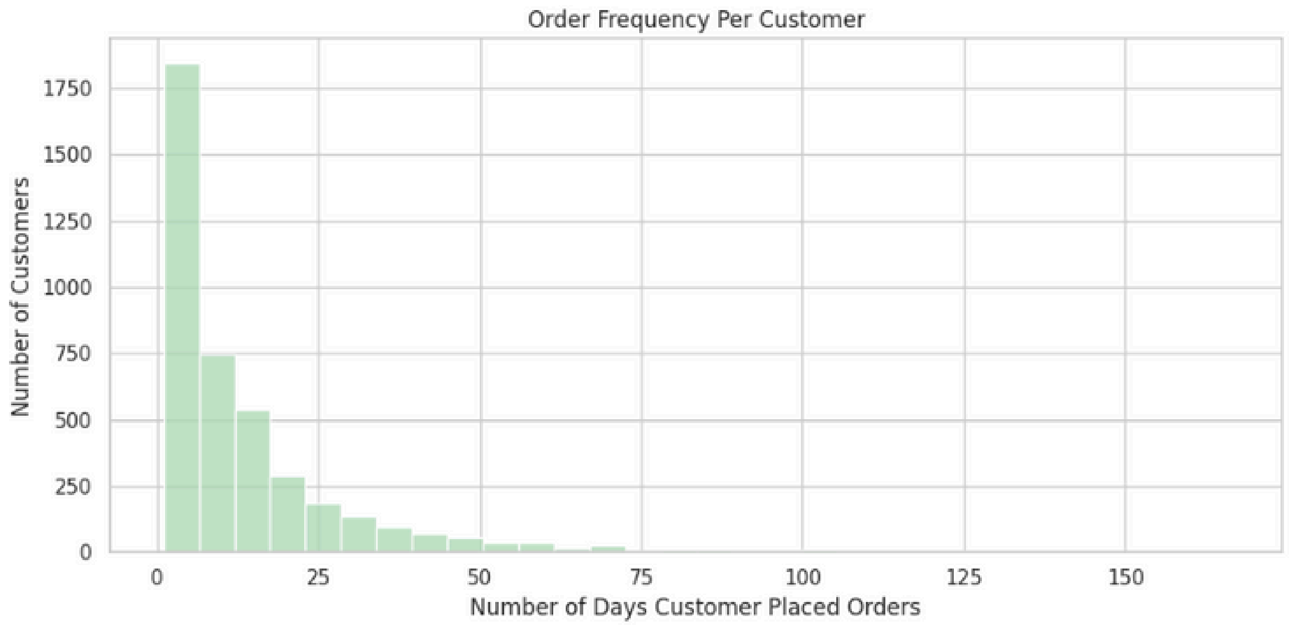
Insights

- We're dealing with segmented, non-standard behavior—requiring customized modeling per client segment.
- Traditional time-series models that expect regular intervals struggle with this level of sparsity and volatility.
- Forecasting needs to be granular, at the SKU-client level, rather than generic.
- Misjudging demand has a direct and measurable impact on UNORG's KPIs and profitability.

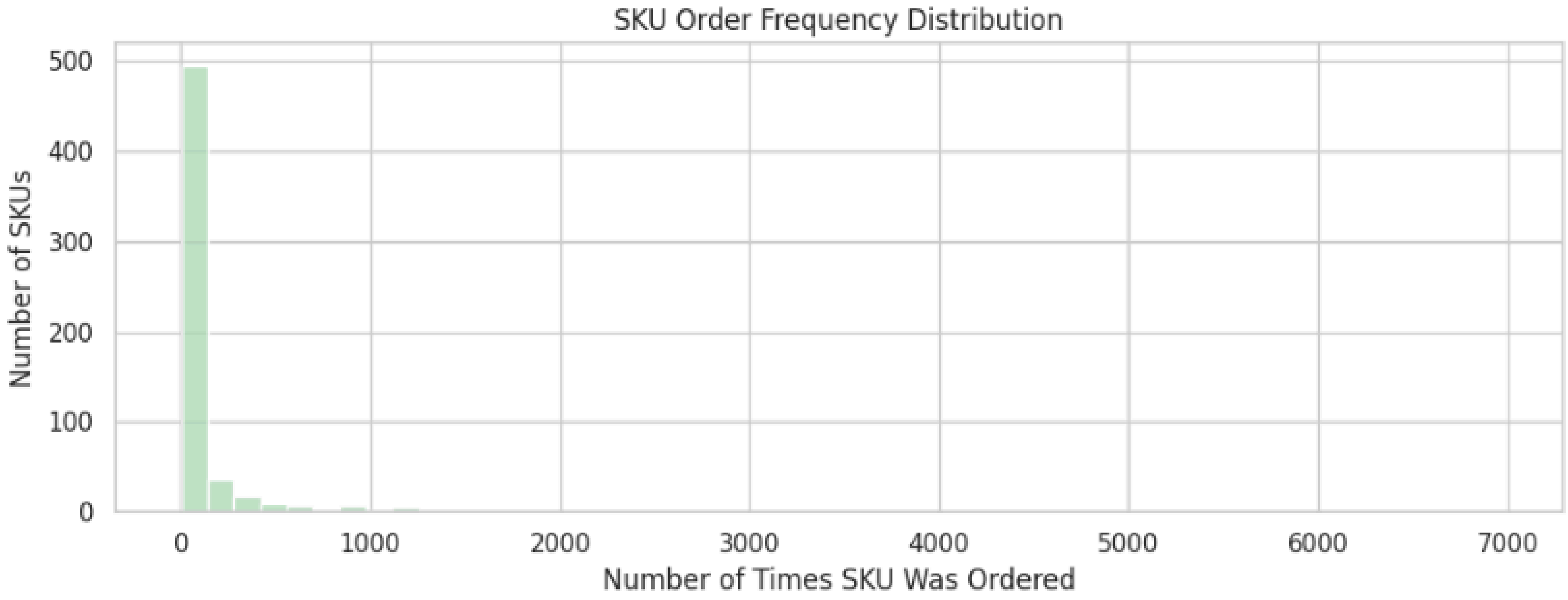
Opportunity

- Identify patterns in client behavior, even in sparse data
- Forecast demand at the individual SKU level per client
- Generate aggregate inventory requirements, helping with optimized stocking, vendor negotiations, and delivery planning

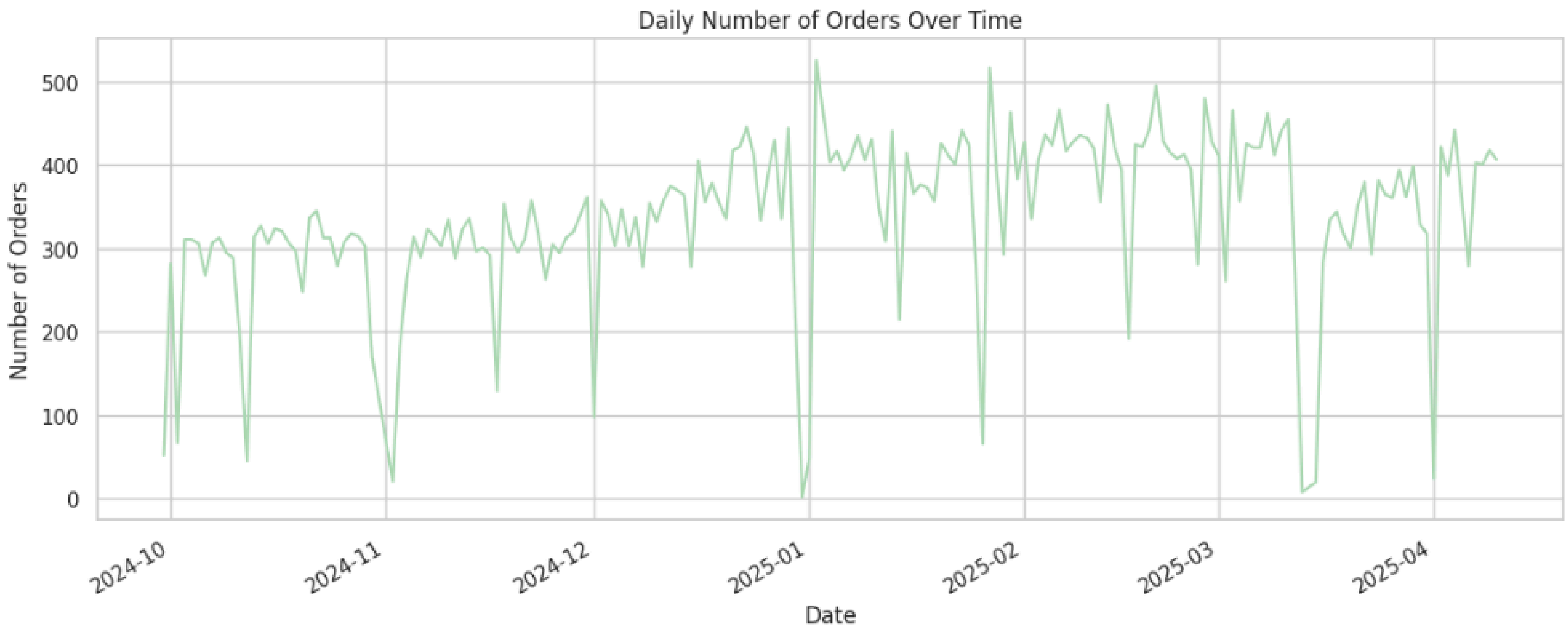
Understanding UNORG’s Clients & Data Landscape



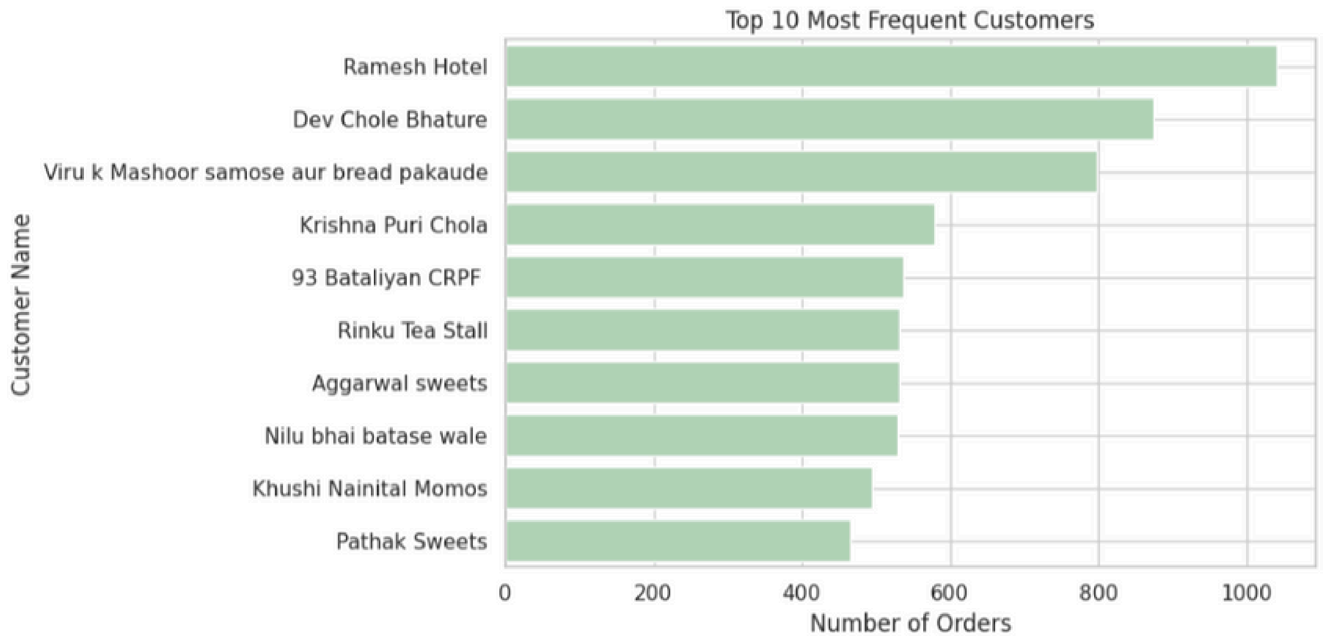
Most customers place orders on very few days, indicating that ordering is irregular and bulk-driven — typical of B2B setups where restocking happens in bursts rather than consistently.



A small set of SKUs dominates order frequency, while the majority are rarely ordered. This shows a long-tail distribution in product demand.



There is noticeable seasonality, daily fluctuations, and occasional demand surges (e.g., before festivals or restock cycles).



These customers represent consistent high-volume buyers, likely businesses with regular demand (like hotels, CRPF mess halls, etc.).

A robust solution must handle irregularity, diversity, and data sparsity—making feature design and modeling strategy critical.

Customer Order Prediction Model

Objective

Predict the daily probability of each customer placing an order for the next 14 days, enabling better demand planning and inventory management.

Data Overview

- **Orders Dataset:** customer_id, order_id, order_date
  Used to derive order frequency & recency patterns
- **Items Dataset:** order_id, item_name, quantity
  Used to understand product preferences & ordering habits
- **Time Range:** Last 6 months of customer activity

Feature Engineering

- **Recency & Frequency Metrics:** days_since_last_orderorders_last_7_days, orders_last_14_days , total_orders, avg_days_between_orders
- **Customer Preferences:** common_item_encoded ,distinct_items_ordered
- **Temporal Features:** day_of_week, month, is_weekend

Modeling Strategy

Problem Type: Binary Classification
(Will a customer order in next 14 days? Yes → 1, No → 0)

Models Evaluated:

- Logistic Regression (Baseline)
- Random Forest
- K-Nearest Neighbors
- XGBoost
- **LightGBM** – Best performer based on Precision, Recall, and F1-score

Classification Report Highlights

Metric	Class 0(No Order)	Class 1(Order)
Precision	0.97	0.60
Recall	1.00	0.15
F1- Score	0.98	0.24

Business Impact

- Improve customer retention through early intervention
- Optimize inventory stocking based on predicted demand
- Enhance marketing ROI with focused targeting
- Increase repeat purchases and overall revenue
- Enable data-driven automation for retargeting strategies

Future Improvements

- Include monetary value features like average order value or total spend to better capture customer intent
- Incorporate customer demographics and location data for more personalized predictions
- Explore sequence models (e.g., RNNs/LSTMs) to capture time-dependent purchasing patterns more effectively

WWW/UNORG.IN

SKU-Level Forecasting Model

Objective

Predict the specific SKUs each customer is likely to order and estimate the expected quantity — enabling targeted replenishment and personalized marketing.

Data Overview

- **Orders Dataset:** customer_id, order_id, order_date
- **Order Items Dataset:** order_id, item_name, quantity
- **Merged View enables:** Mapping customer-SKU purchase relationships and Understanding frequency and quantity behavior
- **Time Range:** Last 6 months

Feature Engineering

- **Customer-SKU Interaction Metrics:** order_frequency, last_order_date, days_since_last_order
- **Quantity Metrics:** total_quantity, avg_quantity, quantity_std
- **Behavioral Indicators:** Variability in quantity (suggests bulk vs. frequent small purchases)

Modeling Strategy

Problem Type: Multi-output Regression
- SKU Selection: Classification (Will customer order this SKU?)
- Quantity Prediction: Regression (How much will they order?)

Models Evaluated:
- Random Forest
- XGBoost
- **LightGBM** (Best for both classification & quantity prediction)

Classification Report Highlights

Metric	Class 0(No Order)	Class 1(Order)
Precision	0.94	0.40
Recall	0.85	0.64
F1- Score	0.89	0.49

Business Impact

- Improve inventory planning by forecasting product-level demand
- Enable personalized SKU-level recommendations
- Reduce stockouts for frequently purchased SKUs
- Drive basket size growth through cross-sell predictions

Sample output

customer_id	item_name	selection_probability	will_order	predicted_quantity
4	Ruchi Gold Palm Pouch(1L)	0.707095	1	4
4	Prakash Maida	0.547655	1	2
4	Bawarchi Vanaspati Pouch 1ltr	0.337773	0	0
4	Raag Gold Palm Pouch	0.337449	0	0
4	Normal Sugar	0.287165	0	0

Future Improvements

- Integrate seasonality trends and promotion data for quantity spikes
- Explore sequence learning (e.g., Transformers) for next-product prediction

Inventory Planning: Demand Aggregation & Stock Optimization

Objective

Generate an optimal inventory plan by forecasting aggregate demand across all customers for the next 14 days, minimizing stockouts, reducing overstocking, and enhancing supply chain efficiency.

Methodology Overview

Input
Predicted Order Quantities from output from customer-level demand prediction (Challenge 1 & 2).

Date-Aware Expansion
Predictions are cross-joined with a 14-day future window.

Demand Aggregation
Sum daily predicted quantities per SKU. Compute total 14-day demand for each product.

Safety Stock Buffer
Add 20% buffer to cover variability in demand. Fixed buffer of 10 units ensures minimum availability.

Inventory Calculation
Compare current stock with projected demand + safety stock. Compute Recommended Order Quantity

Planning Logic

Recommended Order = max(buffer stock , (14_day_demand + safety_stock))

Sample output

	item_name	14_day_demand	current_stock	safety_stock	recommended_order
0	7 Star Rice	1106	0	231	1337
1	AJMAL BASMATI RICE	322	0	74	396
2	ASHOK BIRYANI MASALA (50 GM)	280	0	66	346
3	Aashirwad Chakki Atta (10kg)	5446	0	1099	6545
4	Aashirwad Chakki Atta (5kg)	5474	0	1105	6579
5	Adani Mota Besan	994	0	209	1203
6	Adani Patla Besan	560	0	122	682
7	Agni Tata Tea	1456	0	301	1757
8	Agro Chana Besan Mahin	14	0	13	27
9	Agro Chana Besan Mota	56	0	21	77

Supply Chain Impact

- **Demand-Driven Restocking:** Helps avoid manual forecasting errors and over-reliance on past data.
- **Reduced Stockouts:** Ensures product availability, improving customer satisfaction and retention.
- **Inventory Cost Optimization:** Prevents unnecessary overstocking and warehousing costs.
- **Improved Supply Chain Responsiveness:** Enables agile planning and procurement with predictive signals.
- **Data-Informed Procurement:** Supports supplier negotiations and reorder triggers with accurate projections.

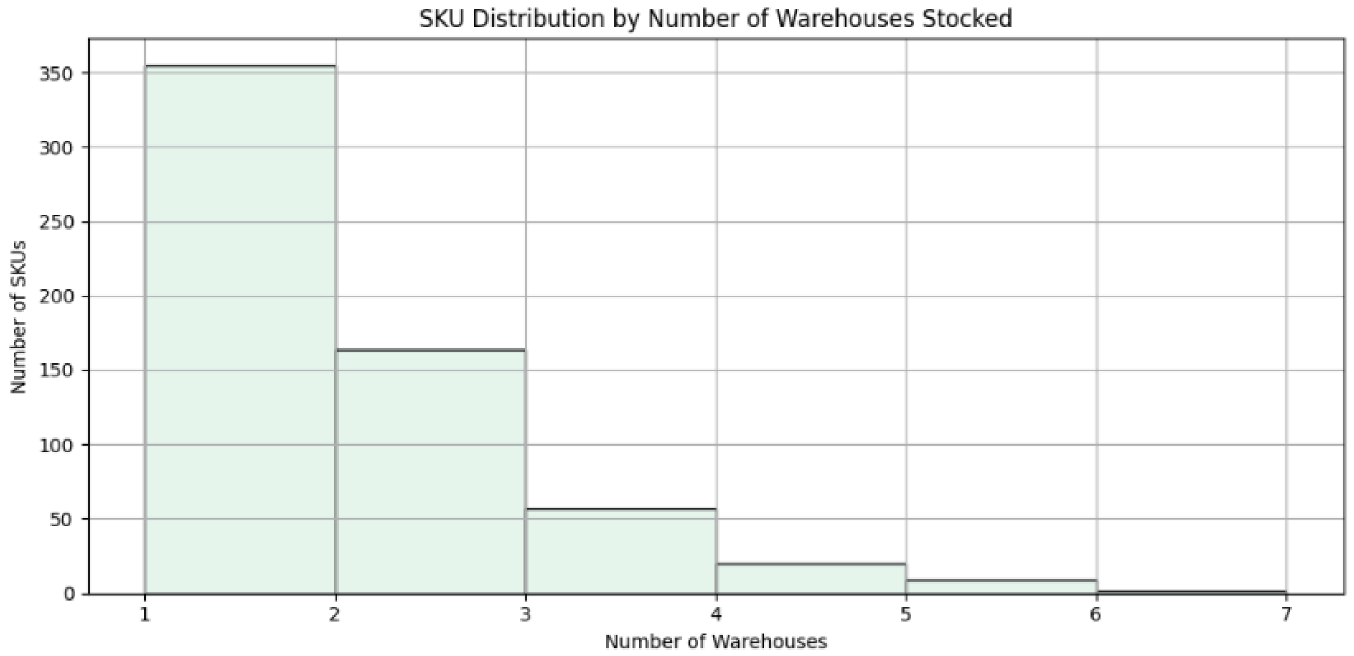
WWW.UNORG.IN

A Hidden Bottleneck : Warehouse-Wise SKU Distribution

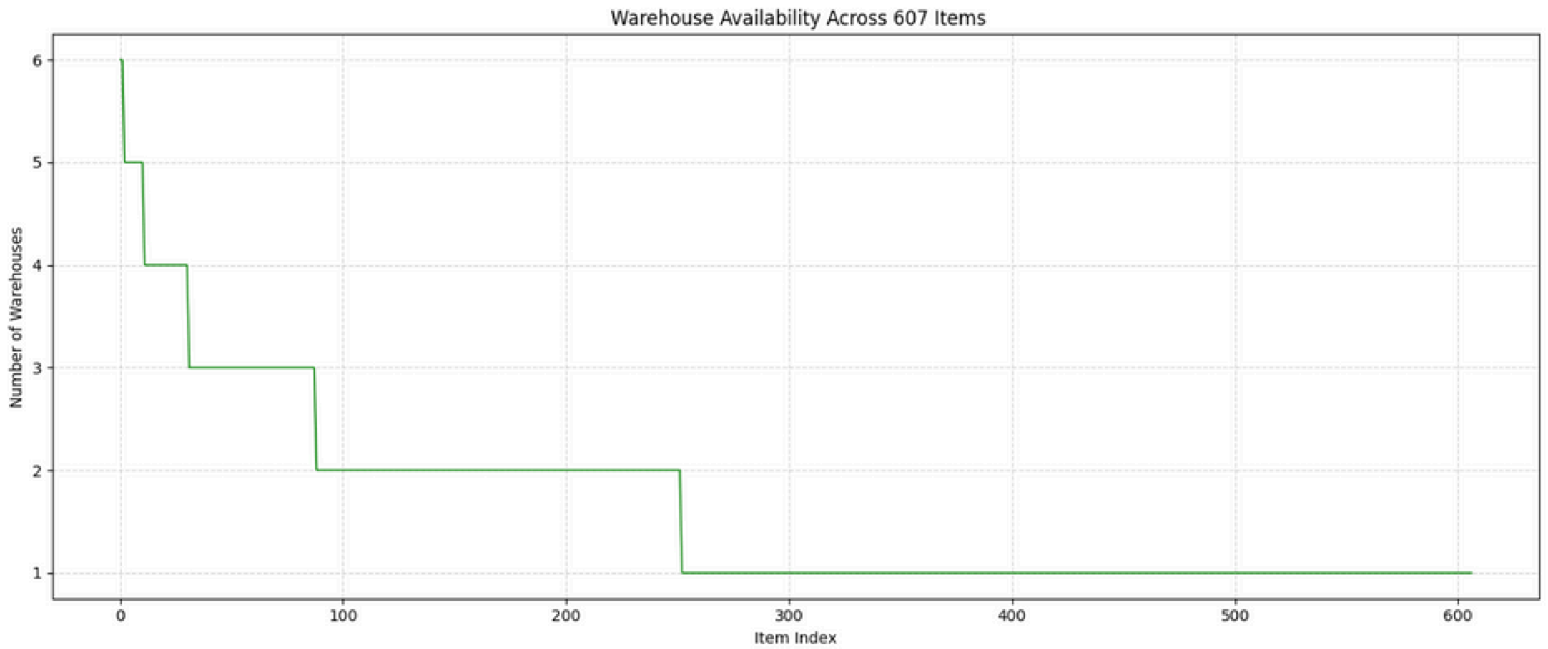
Business Context

In historical order data, each product (SKU) is fulfilled from a specific warehouse. However, not all SKUs are stocked in every warehouse — a fact clearly visualized in the graphs.

Insights



Over 350 SKUs are stocked in only one warehouse, and less than 10 SKUs are stocked in all 6 warehouses.



A line plot of 607 individual SKUs shows sharp drops — suggesting most SKUs are not widely available across warehouses.

Business Implications

Without accounting for this constraint:

- Product recommendations become ineffective if they suggest SKUs unavailable at the customer's assigned warehouse.
- Demand forecasting models overestimate SKU needs in warehouses that never stock them.
- Order fulfillment teams face confusion, delays, and increased manual intervention.

Opportunities for Improvement

Incorporating warehouse-level availability enables us to:

- Filter product recommendations based on what's actually in stock at each warehouse.
- Optimize inventory placement to ensure high-demand SKUs are distributed more effectively.
- Reduce delivery time and improve fulfillment rate, especially for repeat or high-value customers.

Business Impact

Feasibility

Compatibility

- Solution is based on historical order and SKU-level data already collected by UNORG.
- Works with current data pipelines, requiring only moderate engineering to operationalize.
- Flexible enough to integrate with existing warehouse and order management systems.

Modular Architecture

- The system is built in layers (Customer Prediction → SKU Forecasting → Inventory Aggregation), enabling staged rollout.
- Lightweight models (e.g., LightGBM) ensure fast inference and low compute cost — ideal for real-time or batch modes.

Anticipated Business Impact on KPIs

Fulfillment Rate	Increase	Accurate SKU-level forecasting ensures warehouses are pre-stocked based on expected demand.
Waste Reduction	Decrease	Demand-aligned inventory reduces excess stocking and spoilage, especially for perishables
Service SLA Compliance	Increase	Enhanced planning reduces delivery delays by improving SKU availability at the right warehouse.
Profitability	Decrease	Reduces losses from over-discounting and reactive procurement.

Scalability Potential

Business Scaling

Easily generalizes across cities, warehouses, and product categories with retraining.

Technical Scaling

Compatible with cloud-based infrastructure (AWS/GCP) for real-time expansion.

Decision-Making Scaling

Empowers procurement, marketing, and ops teams with predictive insights.

Phase-Wise Implementation Plan

Phase	Description	Timeline	Potential Limitations	Next Steps
Phase 1: Data Readiness & Model Finalization	Clean and structure order + SKU data Finalize model selection (LightGBM/XGBoost) Prepare features for daily order prediction	Weeks 1–3	Incomplete or inconsistent historical data Low ordering frequency for new clients	Set up automated data pipelines- Define model retraining frequency
Phase 2: Controlled Pilot Deployment	Deploy model for select clients & 1–2 warehouses Run in parallel with manual planning Collect feedback from operations team	Weeks 4–6	Prediction-action gap Model bias toward high-frequency buyers	Host training sessions for ops- Document use cases where model wins
Phase 3: Integration with Inventory & Procurement	Use 14-day demand forecast for replenishment Build stock recommendation engine Incorporate warehouse constraints	Weeks 7–9	Limited SKU visibility in certain warehouses Warehouse-level SKU mapping inconsistencies	Align inventory teams Refine buffer logic for understocked zones
Phase 4: Full-Scale Rollout & Monitoring	Expand to all customers and warehouses Enable auto-alerts for low-stock, missed orders Launch dashboards for business visibility	Weeks 10–13	Model performance may degrade over time Operational inertia in larger teams	Implement model monitoring Regular feedback loop with stakeholders
Phase 5: Continuous Improvement & Scaling	Add seasonal, promo, and vendor data Tune model based on real-time trends Explore deep learning for long-term prediction	Ongoing	Requires ongoing data updates Computational cost may increase with scale	Plan quarterly model reviews Benchmark vs. manual forecasts regularly

Cost-Benefit Analysis

Cost of Implementation

Cost Component	Estimated Effort/Cost
Data Engineering	2–3 weeks of setup time to clean and pipeline order & inventory data
Model Development & Tuning	~1 month of ML engineering (modeling, testing, evaluation)
System Integration	Integration with inventory planning tools and warehouse systems (moderate backend effort)
Operational Rollout & Training	Minimal cost with internal team sessions and phased rollout

Tangible Benefits

Benefit Area	Estimated Impact
Reduced Inventory Waste	~15–25% decrease in overstocked items, lowering spoilage and capital lock-up
Higher Fulfillment Rate	8–12% improvement in order fulfillment via preemptive stocking
Lower Stockouts	~30% fewer missed orders due to real-time demand awareness
Profitability Boost	Recovery of losses from 70%+ negative profit orders through targeted procurement & discounting optimization

ROI Outlook

- **Breakeven Timeline**: Within **2–3 months** post-deployment due to immediate improvements in waste reduction and service levels.
- **Long-Term Value:** Solution compounds over time as accuracy improves with more data.
- **Scalability:** Minimal marginal cost to extend solution to new regions, warehouses, or customer segments.

Conclusion

The solution delivers a high ROI with low technical debt. Its ability to reduce waste, improve planning, and enhance fulfillment directly improves UNORG's bottom line and customer experience — with scalability built in.

Thank-You