GENERAL CHAMPIONSHIP TECH

unORG

Supply Chain Data Science

MID PREP

Audited by:

Society of Business

PROBLEM DESCRIPTION

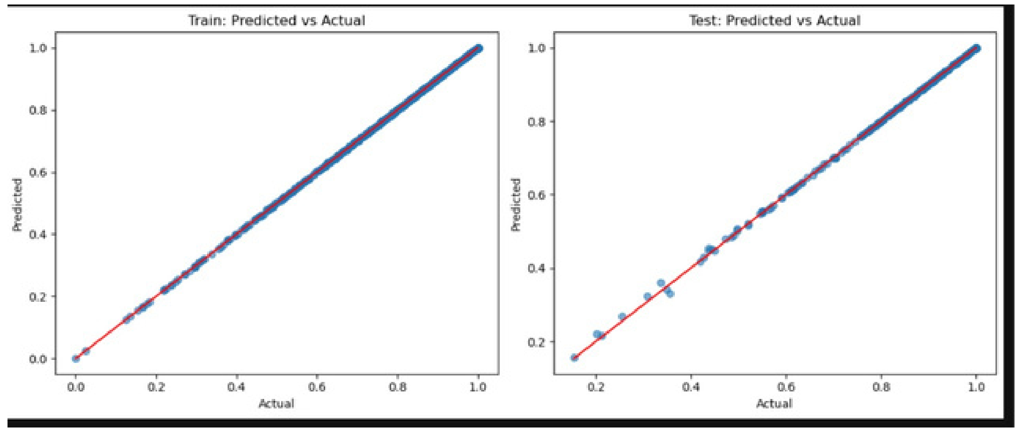
UNORG is looking for a predictive solution to forecast daily customer orders, including SKUs and quantities, over the next 14 days and generate an optimal inventory plan. The aim is to minimize stockouts, reduce waste, and ensure high service levels.

CONTENTS

Customer Order Prediction Model

```
MAE  : 0.0005
MSE  : 0.0000
RMSE : 0.0017
R²   : 0.9999
MAPE : 0.12%
```

```
Train MAE: 0.0002, R²: 1.0000
Test  MAE: 0.0005, R²: 0.9999
```

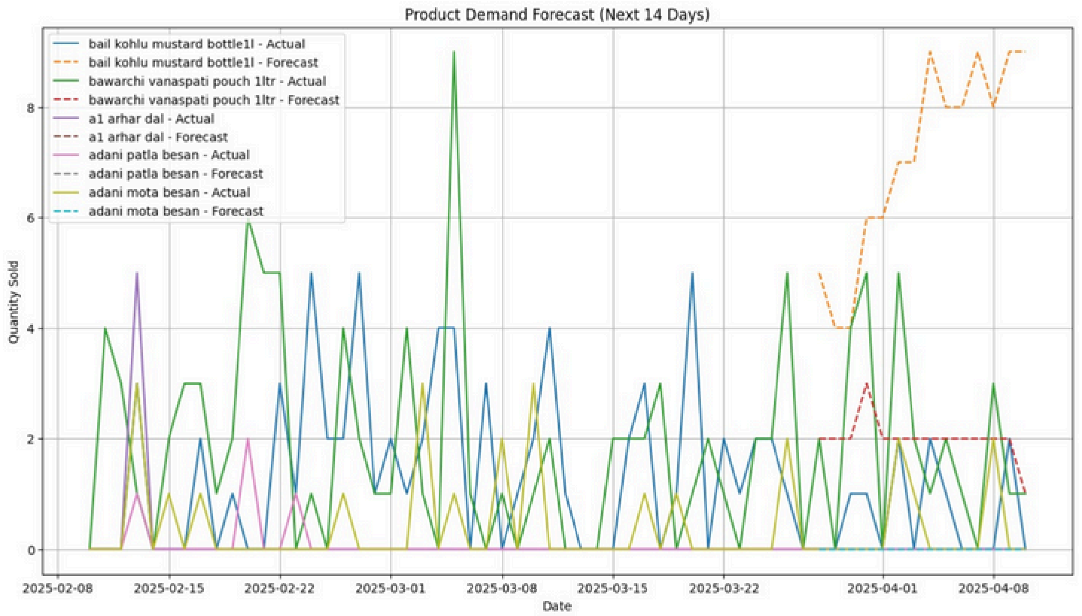


SKU-Level Demand Forecasting

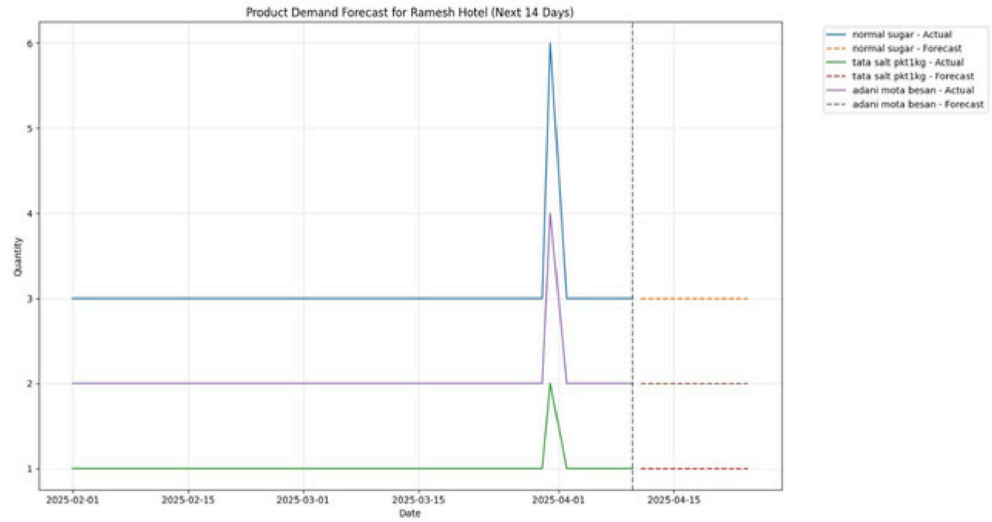
```
[ ] import pandas as pd
    from sklearn.metrics import r2_score

    # Calculate R-squared
    r2 = r2_score(forecast_results, X_test)
    print(f"R-squared (R²) Accuracy: {r2:.4f}")

R-squared (R²) Accuracy: 0.7529
```



Inventory Planning Engine




```
Forecast for top customer: Ramesh Hotel

Total quantities expected in next 14 days:
adani mota besan                 28.0
bawarchi vanaspati pouch 1ltr    14.0
chhola                           56.0
ganesh maida                    126.0
kezar aata                       84.0
kezar maida                      42.0
kn gold maida                   112.0
nature fresh soya tin15l         14.0
normal sugar                     42.0
prakash aata                    112.0
rice moti gold steam             14.0
riceman jeera rice               42.0
shreshtha aata                  140.0
swastik maida                   140.0
tata salt pkt1kg                 14.0
independence steam rice          42.0
steam rice                       42.0
adani patla besan                 0.0
besan                             0.0
dalda vanaspati pouch             0.0
saccha heera chhola               0.0
soyum soya tin15l v               0.0
dtype: float64
```

Data Handling Techniques

- Missing value imputation via zero-fill and calendar reindexing ensured complete time-series continuity.
- Outlier detection combined Z-score thresholding and IQR bounds at SKU-level for anomaly removal.
- Group-by aggregation on SKU-date and customer dimensions ensured normalized, deduplicated records.
- Regex-based text normalization and label encoding/factorization prepped categorical data for modeling.
- Time index alignment created consistent date grids for Prophet and Croston model inputs.

Feature Engineering

- **Recency:** days_since_last_order, days_since_first_order captured product/customer freshness.
- **Frequency:** Total orders, order-rate per SKU enabled behavior-based segmentation.
- **Monetary (extended):** Quantity-based aggregates as value proxies supported SKU prioritization.
- **Temporal decomposition:** Extracted day_of_week, week, month, year to model trend and seasonality.
- **Demand segmentation:** Labeled SKUs as smooth, intermittent, or lumpy using statistical thresholds.
- Multi-output feature sets supported simultaneous forecasting across SKU-warehouse pairs.

Technical Innovation

Modular time-series engine using Prophet & Croston tailored to demand variability.

Dynamic model selection driven by SKU-level demand segmentation.

End-to-end ML pipelines with XGBoost, Random Forest, and hyperparameter tuning for robust prediction.

Custom outlier detection ensuring clean, reliable input data.

Multi-output forecasting for SKU x warehouse combinations.

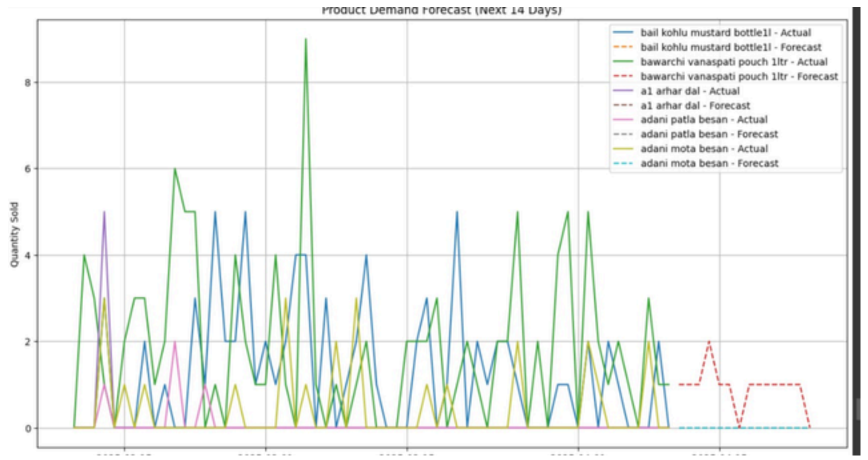
Clustering integration for product/customer segmentation

Inventory Planning Logic

Criteria	How It's Addressed?
Effectiveness	Consolidates granular demand to SKU-level plan
Scalability	Works for any number of SKUs / warehouses
Simplicity	Transparent logic, easy to automate
Adaptability	Buffer %, lead time, and window are tunable

```
Top 10 Products Forecast (Next 14 Days):
masodha double filter sugar    770.0
sapna gold maida               572.0
madhur gold maida              272.0
ruchi gold palm pouch1l        194.0
hum tum rice                   165.0
tata salt pkt1kg               146.0
normal sugar                   108.0
ruchi gold palm                101.0
fortune soya pouch1l            78.0
sapna gold aata                 62.0
dtype: float64

Forecast Summary:
Total products forecasted: 124
Forecast period: 2025-04-11 to 2025-04-24
Total units forecasted across all products: 2668
```



How it was planned?

1. **Demand Aggregation**
   Consolidate predicted daily demand per customer across all SKUs and days
   Output: 14-day total demand per SKU at each warehouse

2. **Safety Stock Addition**
   Apply a 10% buffer to accommodate demand variability and uncertain lead times. Safety stock tuned per SKU type

3. **Stocking Quantity Estimation**
   Final Stocking Qty = Aggregated Demand + Safety Stock
   Ensures buffer for service SLAs and lead-time lags

Cost to Implement

Model Development & Training: Time and resources spent on data cleaning, feature engineering, model building (e.g., time series + ML).

Infrastructure: Basic cloud resources or in-house servers to run daily forecasts.

Integration: Minor dev effort to integrate predictions into inventory and order planning systems.

Maintenance: Low—once deployed, model retraining can be scheduled periodically

ROI

Reduced Waste: By accurately forecasting quantity per SKU per customer,Save costs on overstock and wastage

Reduced Stockouts: Timely order predictions

Faster Decision-Making: Automated forecasts reduce manual planning effort and error

Save costs on overstock and wastage

Capture missed revenue from unfulfilled orders.

THANK YOU

www.reallygreatsite.com