Using Time-Series Analysis for Sales and Demand Forecasting

Stakeholder Report

AUGUST 2025

PERFORMED BY
LAUREN BRIXEY

1. Problem Statement & Objectives

Independent publishers risk costly over/under-stocking without reliable demand forecasts, particularly for seasonal titles. The objective of this project was to develop a time series model capable of delivering demand insights and informing an actionable stocking strategy. Using classical methods (SARIMA) as the benchmark, machine learning (XGBoost), deep-learning (LSTM) and hybrid (sequential/ parallel) models were evaluated using MAE/MAPE metrics and forecast prediction intervals to quantify uncertainty.

2. Data Preparation

2.1. Data Overview

The Nielsen BookScan dataset provided was first filtered to ISBNs with sales observations extending beyond 2024-07-01 (61 ISBNs) and from this set, *The Alchemist* (**TA**) and *The Very Hungry Caterpillar* (**TVHC**) were selected for further analysis. This project uses the sales data for TA (**Figure 1**) and TVHC (**Figure 2**) from ~2012-2024, targeting sales volume. The weekly series were used as the primary modelling input (628 points/title), while monthly series were also generated through aggregating by calendar month as a secondary input for comparison. A few zero-sales weeks/months were noted for later transformations. Both titles show strong annual seasonality with holiday peaks and a 2020 dip consistent with COVID-19 disruption. Evaluation used a holdout test set of the last 32 weeks (or 8 months) per title, which also defined the forecast horizon.

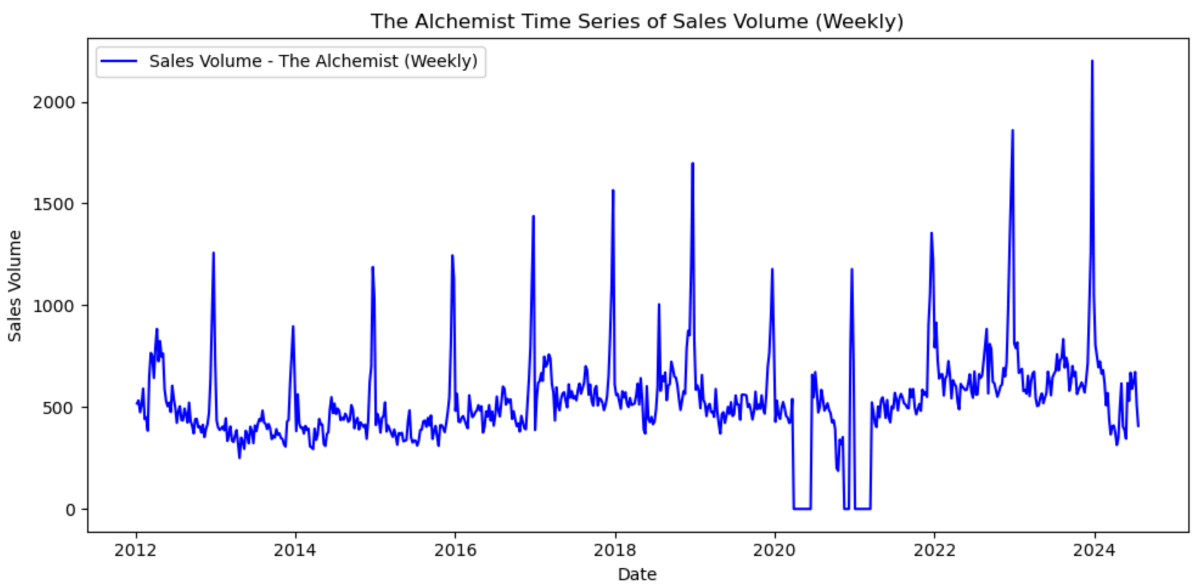

Figure 1*: Time Series of TA Weekly Sales Volume (2012-2024).*

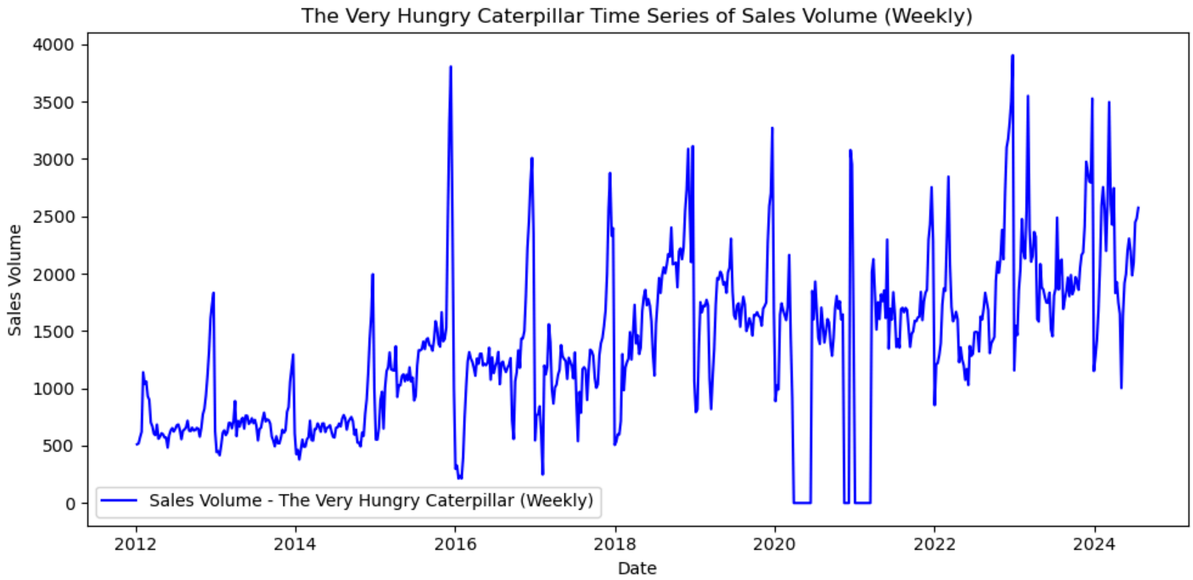

Figure 2*: Time Series of TVHC Weekly Sales Volume (2012-2024).*

2.2. Preprocessing

Dates were converted to a datetime index, regularised to a weekly frequency, and monthly totals derived by resampling. Seasonal-trend decomposition showed increasing seasonal amplitude for TVHC (weekly) and both monthly series, indicating multiplicative seasonality. In this instance, a Box-Cox transformation with a +1 shift (to handle zero months) was applied before modelling. Stationarity checks (ACF, PACF/ACF) found TA and TVHC approximately stationary weekly (though one differencing improved residual diagnostics) and both series required one difference at the monthly frequency.

3. Modelling

Three model families (SARIMA, XGBoost and LSTM) and two hybrid models (sequential and parallel) were evaluated. SARIMA models were fitted on Box-Cox variance stabilised series in the cases of multiplicative seasonality, with auto-arima ranges informed by ACF/PACF. Following training, residual diagnostics were performed (Ljung-box, ACF/PACF and distribution) to confirm white-noise error and the models were verified for convergence. The XGBoost pipeline applied a small positivity shift (for multiplicative seasonality), then deseasonalise, linear detrend and finally a recursive reduced XGBRegressor. Hyperparameters were tuned via grid search with expanding-window cross validation. The LSTM used sliding window sequences with architecture tuned by KerasTuner and early stopping. Hybrid models included a sequential model (LSTM on SARIMA residuals) and a parallel combination (weighted SARIMA + LSTM), with weights chosen on validation MAPE/MAE.

4. Results

4.1. The Alchemist (Weekly Sales)

The hybrid sequential model performed the best (MAPE: 134.29, MAPE: 19.8%, **Table 1**) improving on the SARIMA baseline (**Figure 3**). The hybrid parallel model (LSTM weighting=15%) performed closely to SARIMA while XGBoost and standalone LSTM performed relatively worse.

Table 1: *Comparison of TA Models performance metrics trained on Weekly Sales Volume data.*

Model	**MAE**	**MAPE (%)**
Hybrid Sequential	134.29	19.8
Hybrid Parallel	137.93	21.5
SARIMA Model	141.64	21.8
XGBoost Model	157.43	22.7
LSTM Model	185.16	28.5

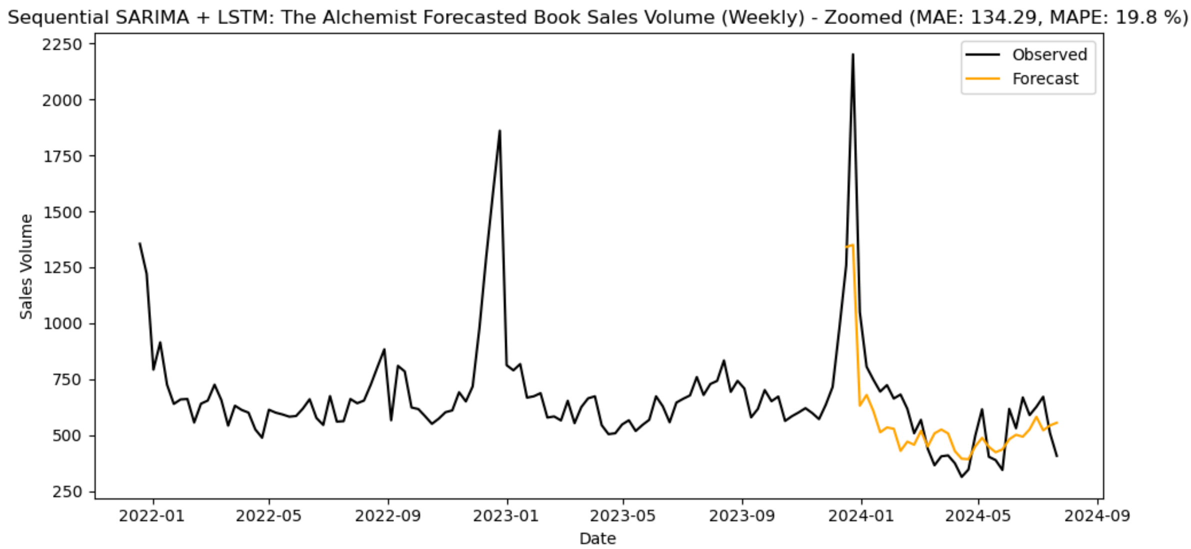

***Figure 3**: TA Weekly Book Sales Volume forecasted (orange) using a Hybrid Sequential SARIMA + LSTM Model.*

4.2. The Very Hungry Caterpillar (Weekly Sales)

The hybrid parallel model (LSTM weighting=23%) performed the best (MAE: 504.46, MAPE: 22.0%, **Table 2**), improving on the SARIMA baseline (**Figure 4**). It also edged the hybrid sequential variant, whilst XGBoost and standalone LSTM performed relatively worse.

Table 2: *Comparison of TVHC Models performance metrics trained on Weekly Sales Volume data.*

Model	MAE	MAPE (%)
Hybrid Parallel	504.46	22.0
SARIMA Model	523.72	23.3
Hybrid Sequential	524.11	23.3
XGBoost Model	562.55	24.4
LSTM Model	653.21	28.4

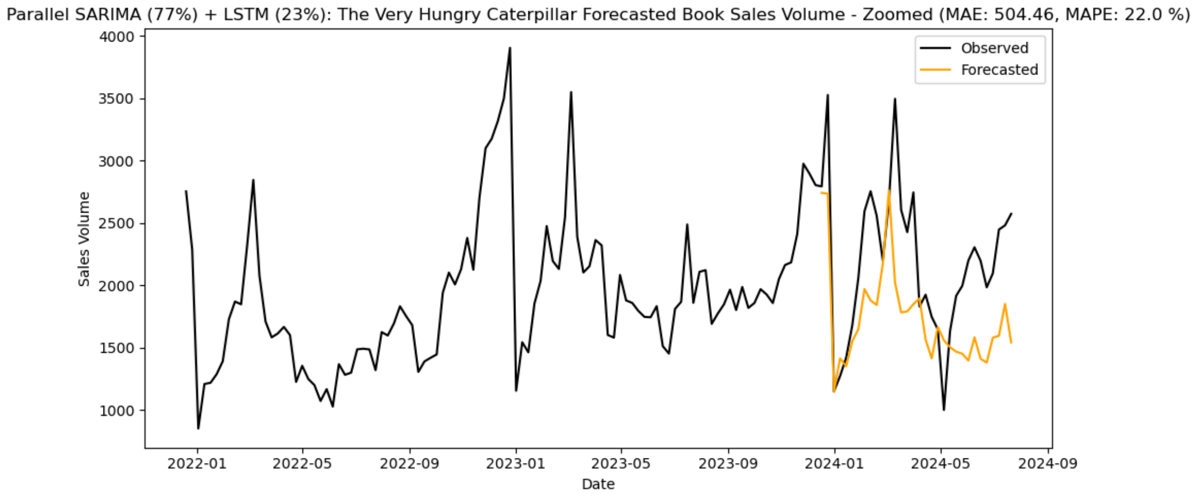

Figure 4*: TVHC Weekly Book Sales Volume forecasted (orange) using a Hybrid Parallel SARIMA + LSTM Model.*

For both titles, the XGBoost/LSTM models captured level and trend but tended to miss event-driven/seasonal spikes, whereas SARIMA handled seasonality better. Combining the two in hybrid models recovered seasonal structure and delivered the strongest accuracy overall.

4.3. The Alchemist (Monthly Sales)

To ensure a like-for-like comparison, weekly forecasts from the XGBoost and SARIMA models were aggregated to calendar months and evaluated against monthly actuals, alongside models trained directly on monthly data.

Direct monthly modelling underperformed the weekly models in all cases in terms of accuracy (MAE/MAPE) (**Table 3, Table 4, Figure 5, Figure 6**). These results support using weekly models as the primary forecaster and then aggregating to monthly for reporting if needed, rather than fitting directly on the monthly frequency.

Table 3: *Comparison of TA SARIMA performance metrics trained on Weekly vs Monthly Sales Volume data.*

Model	MAE	MAPE (%)
SARIMA Weekly (Aggregated)	468.70	17.1
SARIMA Monthly (Direct)	897.47	44.0

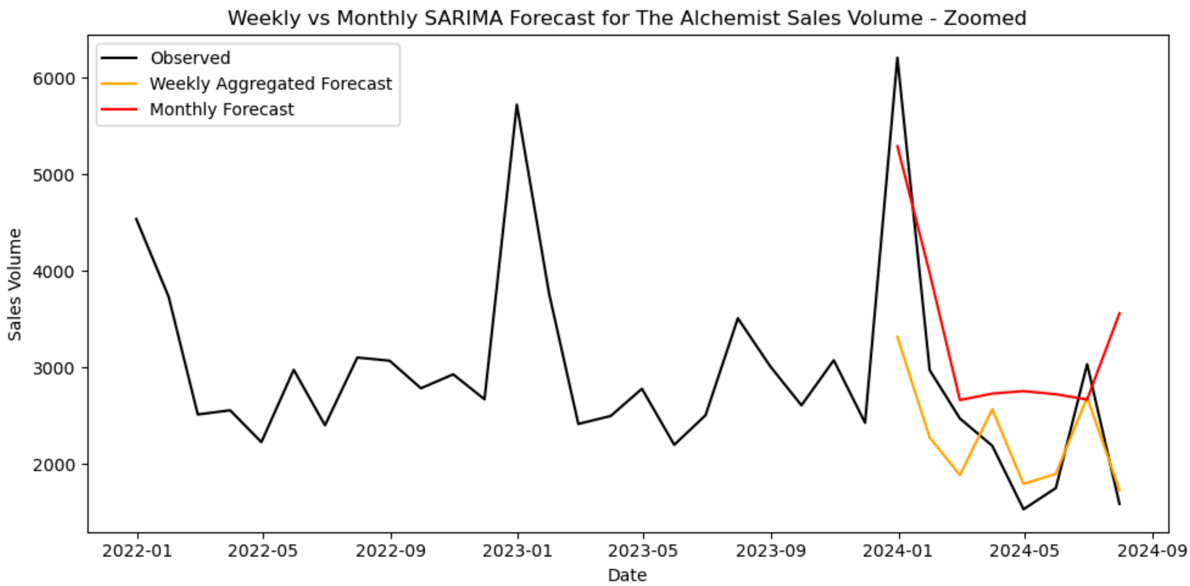

Figure 5*: TA Monthly Book Sales Volume forecasted using a SARIMA Model trained on Weekly (orange) vs Monthly (red) Sales Volume data.*

Table 4: *Comparison of TA XGBoost performance metrics trained on Weekly vs Monthly Sales Volume data.*

Model	MAE	MAPE (%)
XGBoost Weekly (Aggregated)	551.38	19.2
XGBoost Monthly (Direct)	571.57	26.6

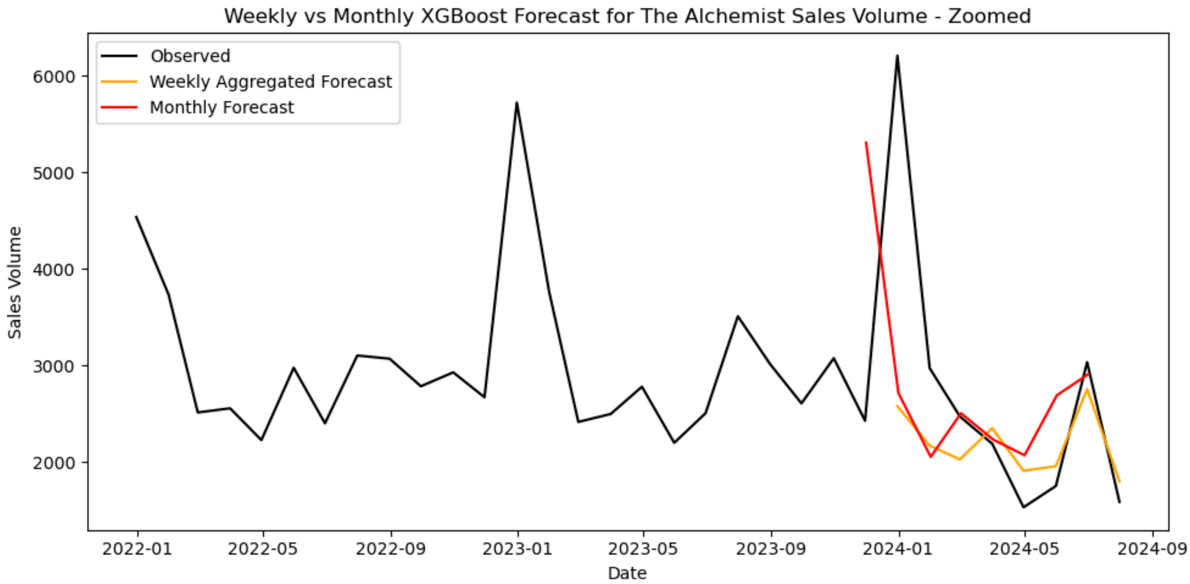

Figure 6*: TA Monthly Book Sales Volume forecasted using an XGBoost Model trained on Weekly (orange) vs Monthly (red) Sales Volume data.*

4.4. The Very Hungry Caterpillar (Monthly Sales)

Similarly to observations derived from TA, direct monthly modelling underperformed the weekly models in all cases in terms of accuracy (MAE/MAPE) (**Table 5, Table 6, Figure 7, Figure 8**).

Table 5: *Comparison of TVHC SARIMA performance metrics trained on Weekly vs Monthly Sales Volume data.*

Model	MAE)	MAPE (%)
SARIMA Weekly (Aggregated)	1719.86	18.3
SARIMA Monthly (Direct)	2179.89	25.7

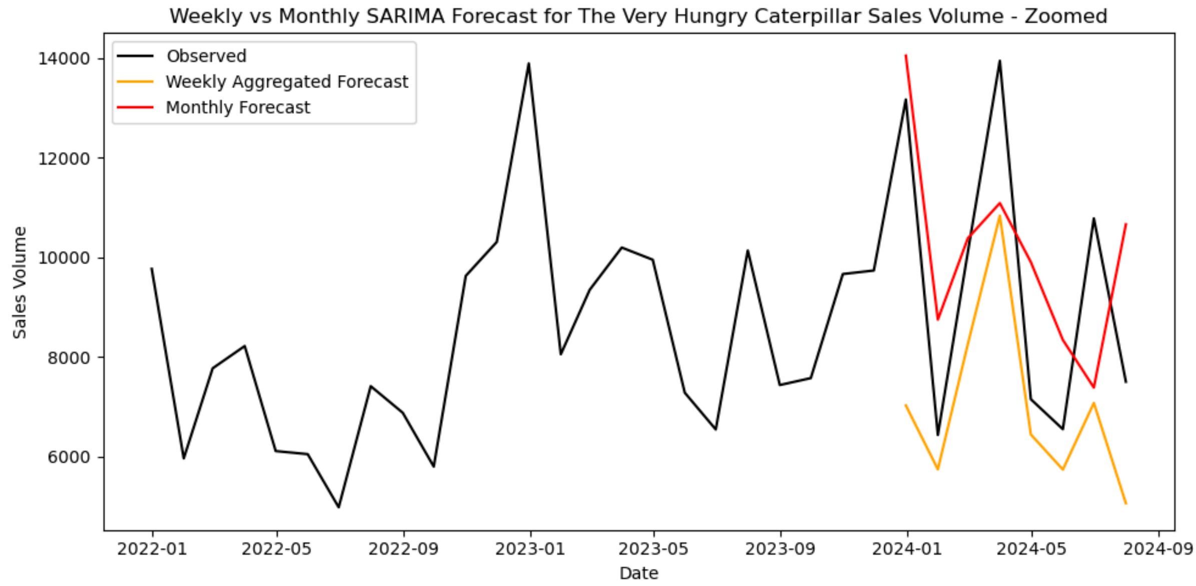

***Figure 7**: TVHC Monthly Book Sales Volume forecasted using an SARIMA Model trained on Weekly (orange) vs Monthly (red) Sales Volume data.*

***Table 6:** Comparison of TVHC XGBoost performance metrics trained on Weekly vs Monthly Sales Volume data.*

Model	MAE	MAPE (%)
XGBoost Weekly (Aggregated)	2147.92	23.5
XGBoost Monthly (Direct)	2460.50	26.2

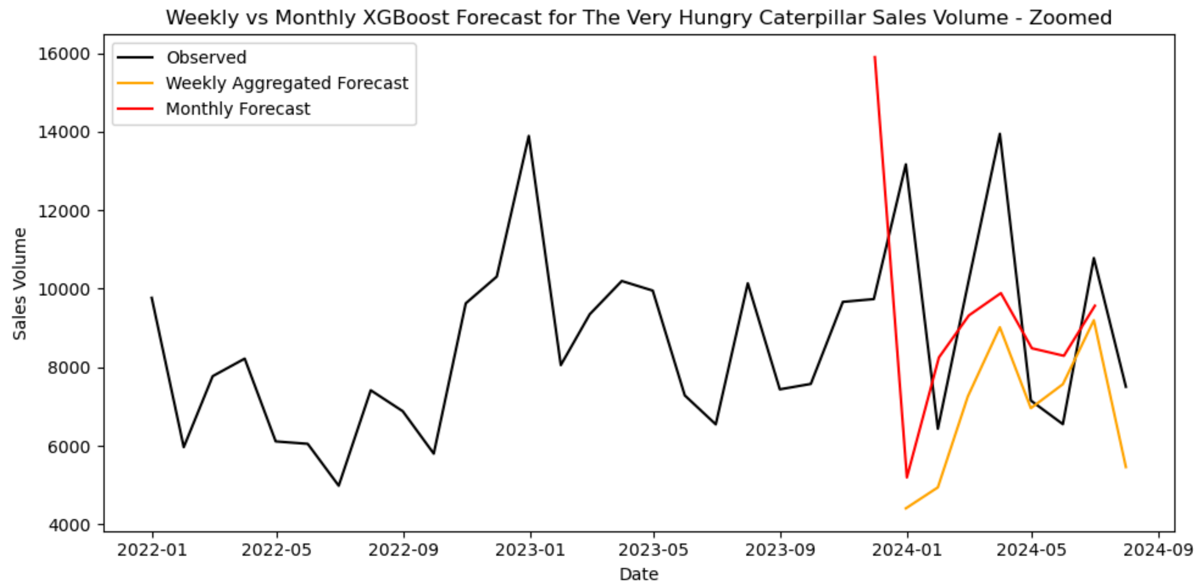

***Figure 8**: TVHC Monthly Book Sales Volume forecasted using an XGBoost Model trained on Weekly (orange) vs Monthly (red) Sales Volume data.*

8. Conclusions and Recommendations

Across both titles, weekly-trained hybrid models delivered the strongest accuracy: for TA the sequential SARIMA + LSTM gave the best results (MAPE=19.8%) while for TVHC, the parallel SARIMA + LSTM performed best (MAPE=22.0%). Direct monthly models underperformed the weekly trained models while standalone ML/DL models underperformed the SARIMA benchmark, capturing level/trend but missing seasonal or event-driven spikes. Future improvements include adding calendar derived lag features and exogenous regressors (e.g. holidays, promotions).