Chapter 2: Building the New Blue

Best Buy x
Data Rangers

Executive Summary

1. Get forecasts a week ahead with precision - within 3 units of actual units sold
2. Approach tried include a feature based and a univariate model
3. Final model based on combining apporaches

Bringing the data to life

Sparsity

23% SKUs have 33% or more 0 sale days historically

Scalability

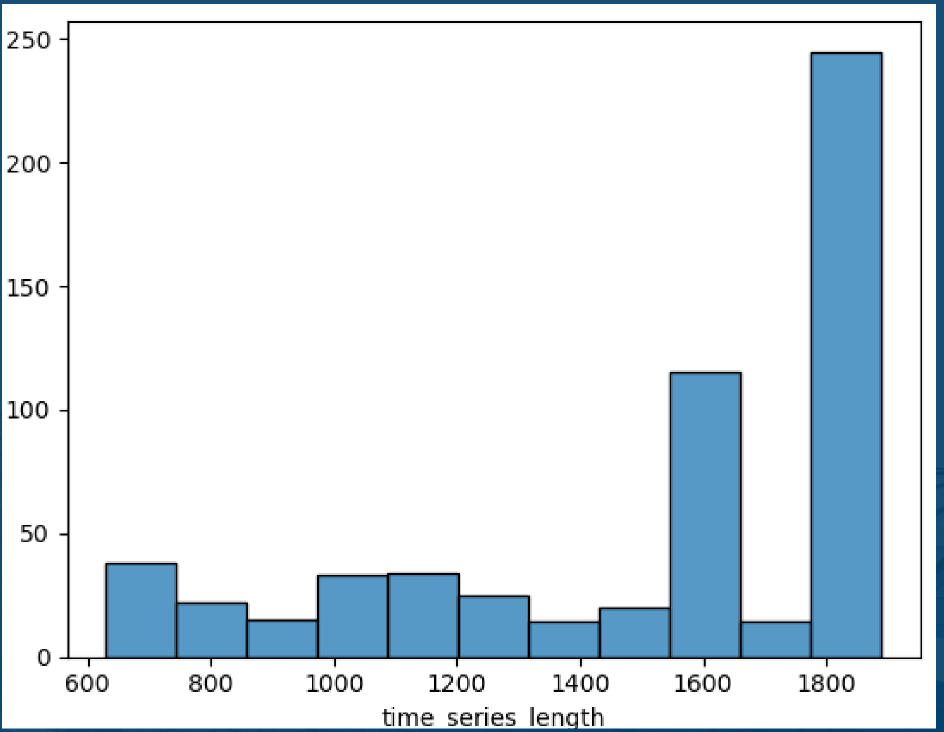
575 SKUs with at least 630 data points each

Signal vs Noise

56 SKUs have more than 20% unusal sale days

Imputing

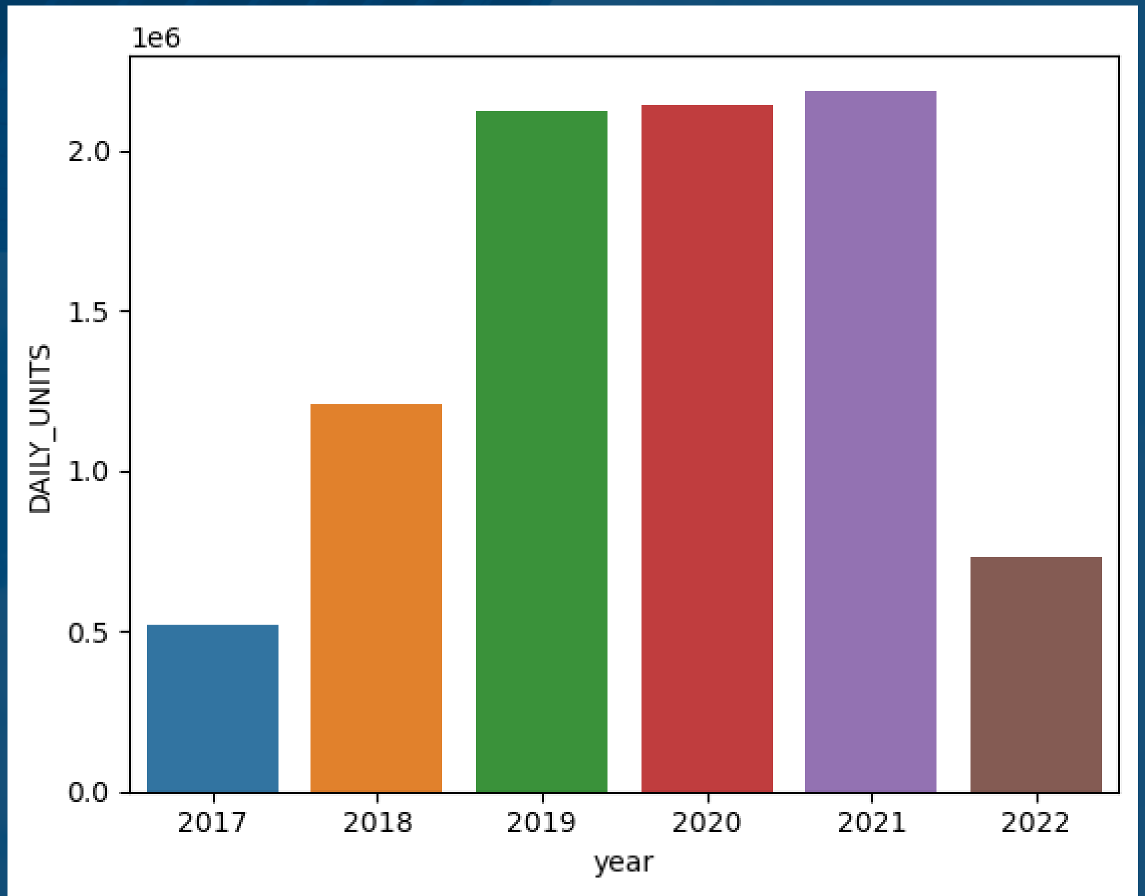
Only **50%** SKUs have complete data.
Adding **0** daily units in case of interim missing data for respective SKUs

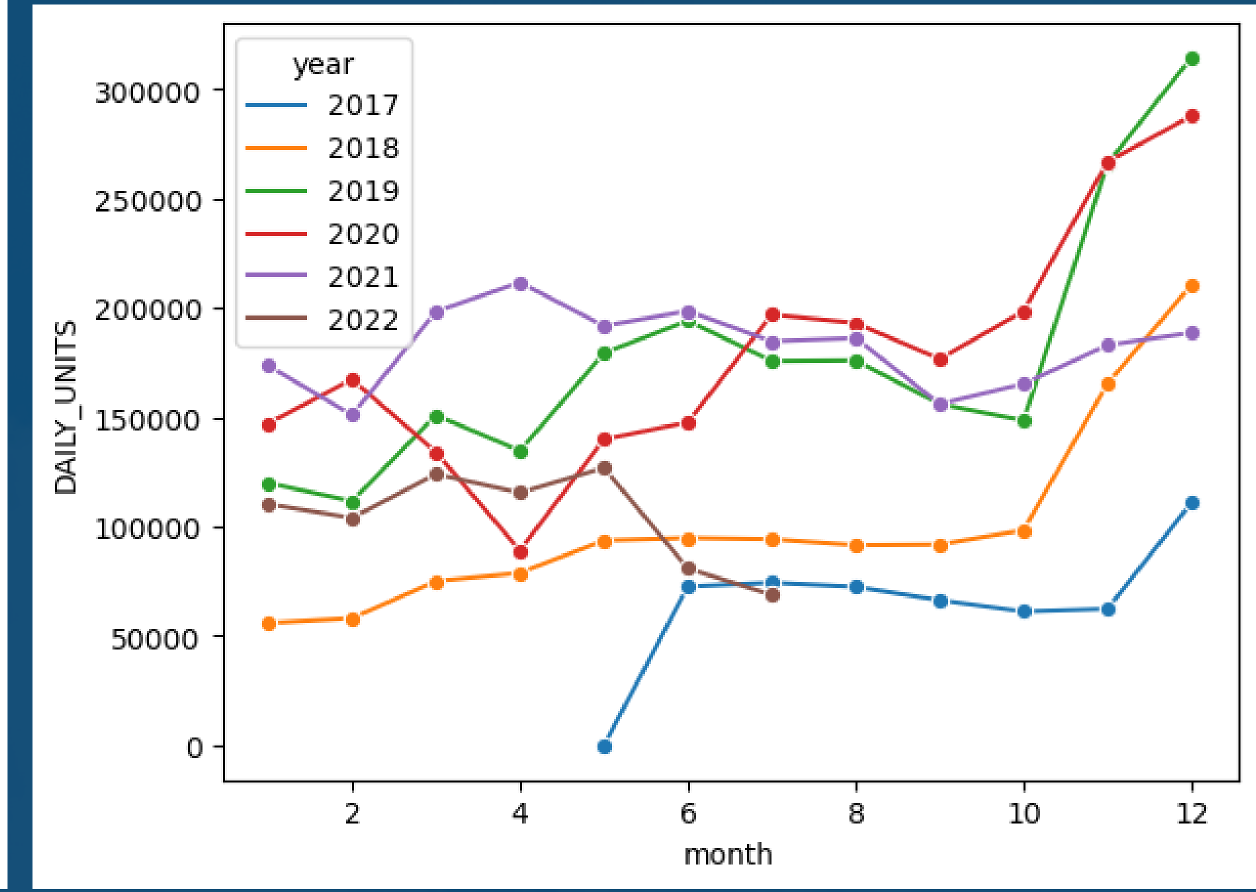


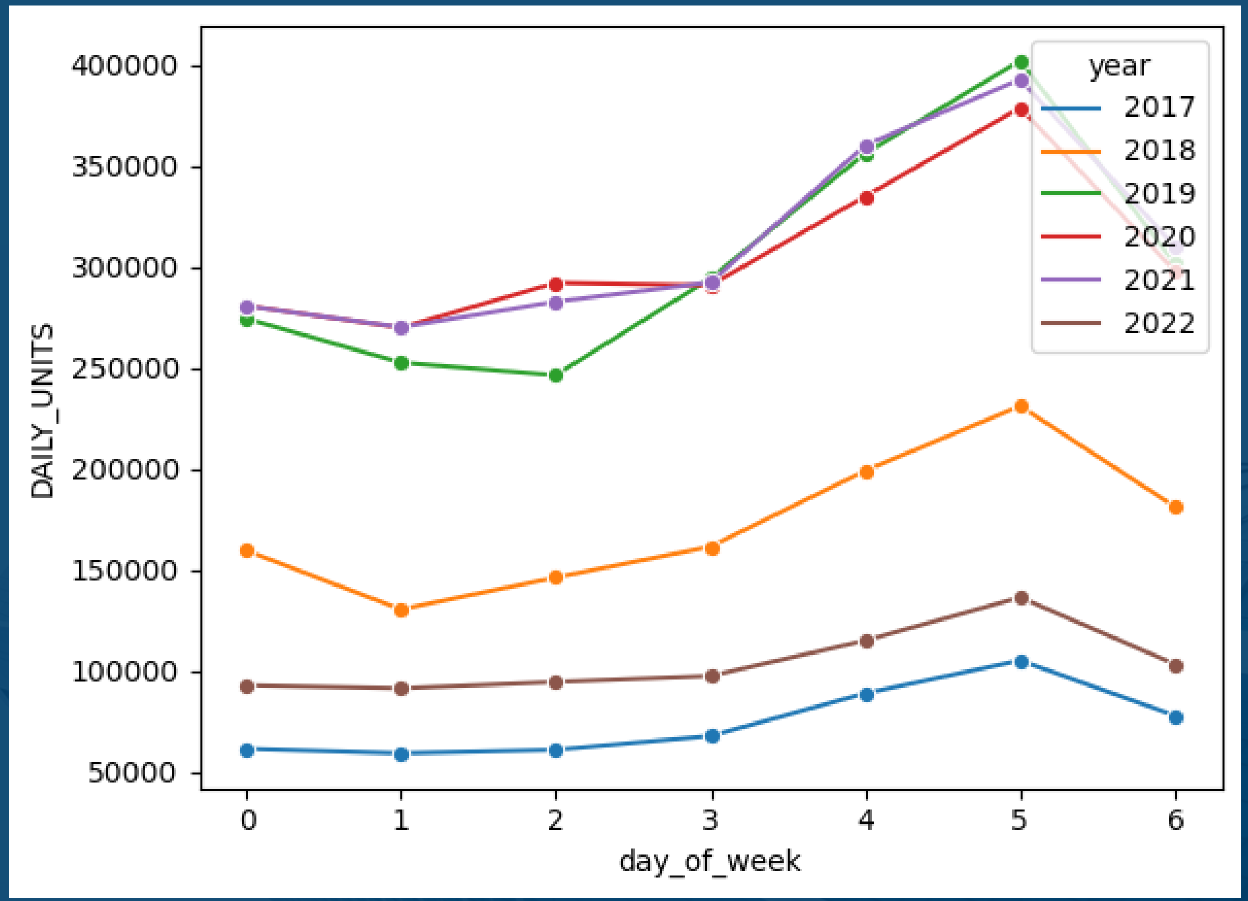
Handling Negative Units

Negative Daily Units - 753 cases where returns are greater than sales were replaced with 0s

Visualizing the data







Overall units by year indicate valuable trend information

Best Buy volumes are consistently higher in November & December

Day of Week seasonality is present with a consistent trend

Modelling the data

Model 1: Base-line (Average) Forecast

RMSE : 50.182

To set a baseline , we forcasted the daily units with the **historical average** of each SKU

Other Metrics:

- MAE: 8.5
- MSE: 2518.3
- R squared = 8%

Please note that in the following models, we have **rounded** the predictions for practical purpose (Decimal Inventory does not make sense)

Modelling the data

Model 2: Croston Model(TSB)*

Training Time - 0.8 seconds

RMSE : 4.717

Croston Model is a smoothing model(similar to Holt & Winters) specialised for forecasting demand of products with intermittent sales.

$$\text{if } d_t > 0 \begin{cases} a_{t+1} = \alpha d_t + (1-\alpha)a_t \\ p_{t+1} = \beta + (1-\beta)p_t \\ f_{t+1} = p_{t+1}a_{t+1} \end{cases}$$

$$\text{if } d_t = 0 \begin{cases} a_{t+1} = a_t \\ p_{t+1} = (1-\beta)p_t \\ f_{t+1} = p_{t+1}a_{t+1} \end{cases}$$

Underlying equations of the model.

Other Metrics:

- MAE: 1.846
- MSE: 32.85
- R squared = 93%

Fine -tuned parameters:

1. *alpha*: Increasing alpha gives more importance to the recent observation
   a. Increasing *alpha* gave us **17.5%** improvement

*Link: Croston, J. D. “Forecasting and Stock Control for Intermittent Demands.” Operational Research Quarterly (1970-1977), vol. 23, no. 3, 1972, pp. 289–303. JSTOR, https://doi.org/10.2307/3007885. Accessed 20 Jan. 2023.

Hand crafted features

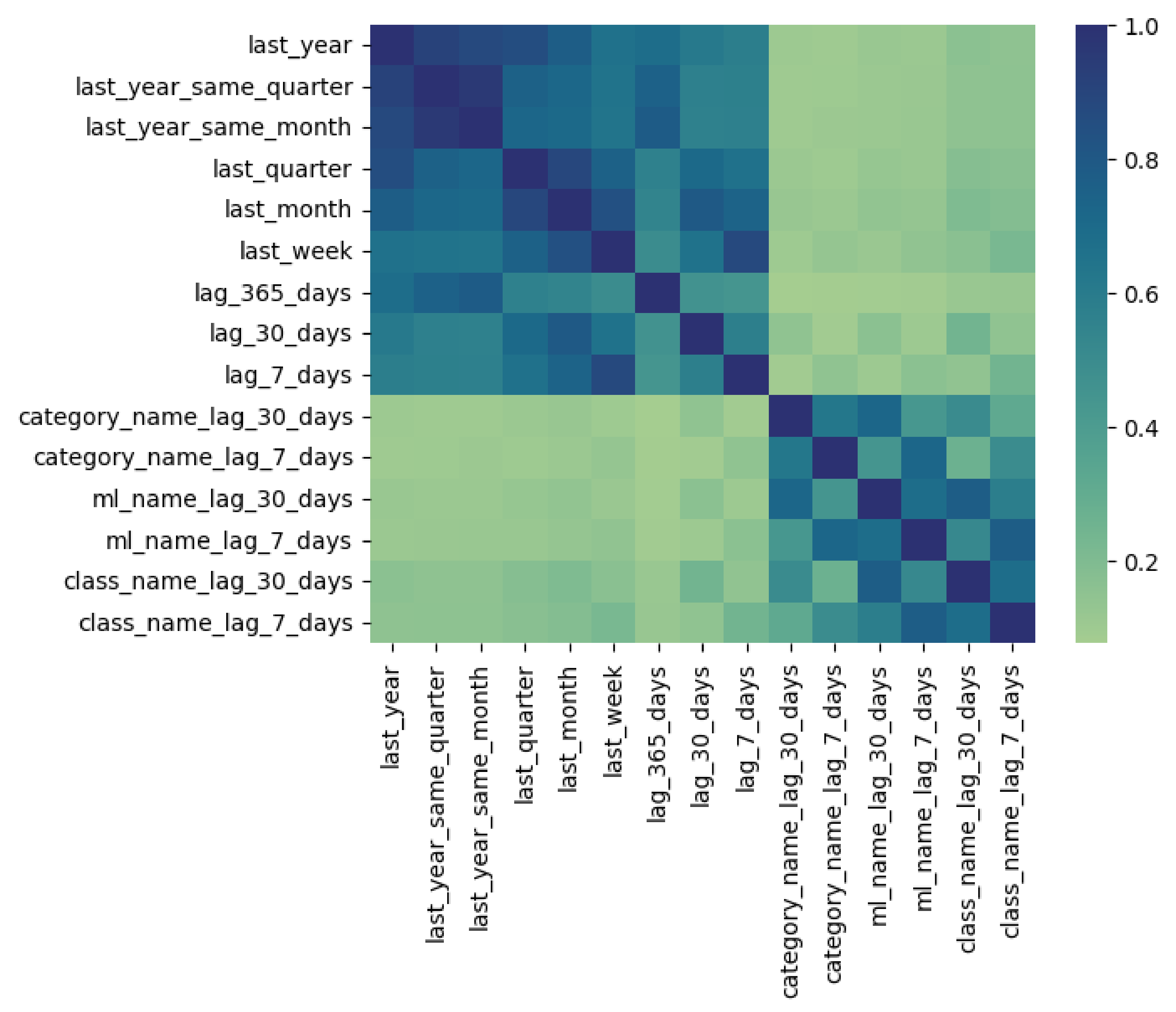
Local Indicators

- Date based features
    - Month, Quarter, Year, Day of the Week etc.
- Historical Daily units of SKU
    - Last year same quarter, Last year same month
    - Lag 30 day, Lag 7 day
- Inventory based
    - One hot encoded
    - Lag 1 day inventory

Global Indicators

- Historical Daily units of Category/ ML_name etc.
    - Category Units 7,30 day lags
    - ML Name units 7,30 day lags
- External indicators
    - Global Supply Chain Pressure Index
    - Consumer Sentiment Pressure Index

Hand crafted features



- Time based features have decent correlation with actual Daily Units
- However, category level features seem to have a weak correlation
- This alludes some indication to why hierarchical time series were not performing as well

Modelling the data

Model 3: Feature Based XGBoost

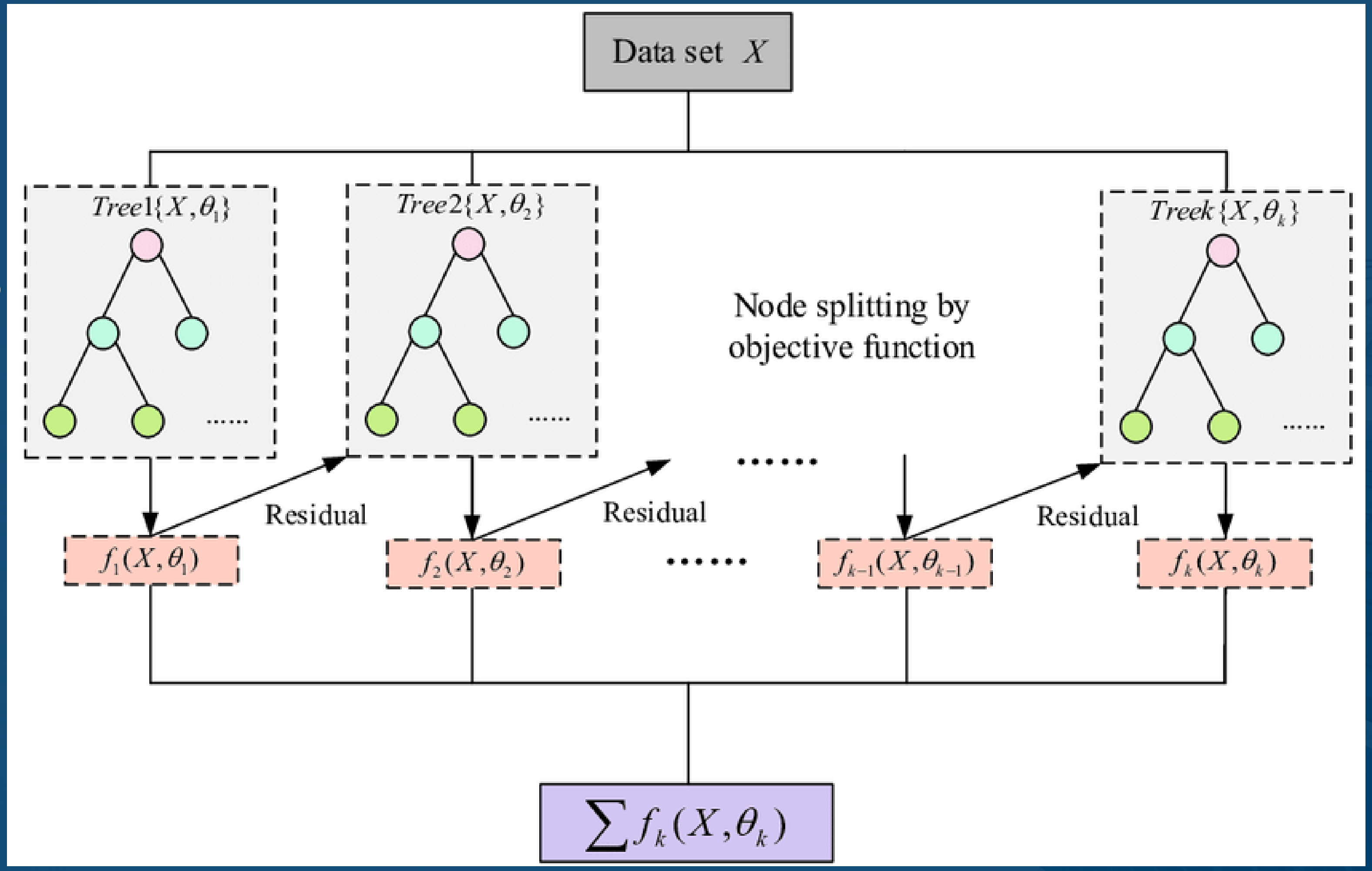
Training Time - 33.7 seconds

RMSE : 4.59

As we know that Tree-based models work effectively on tubular data and a boosting ensemble model would make sure predictions have good accuracy without considerable overfitting.

Other Metrics:

- MAE: 1.834
- MSE: 21.13
- R squared = 93.4%



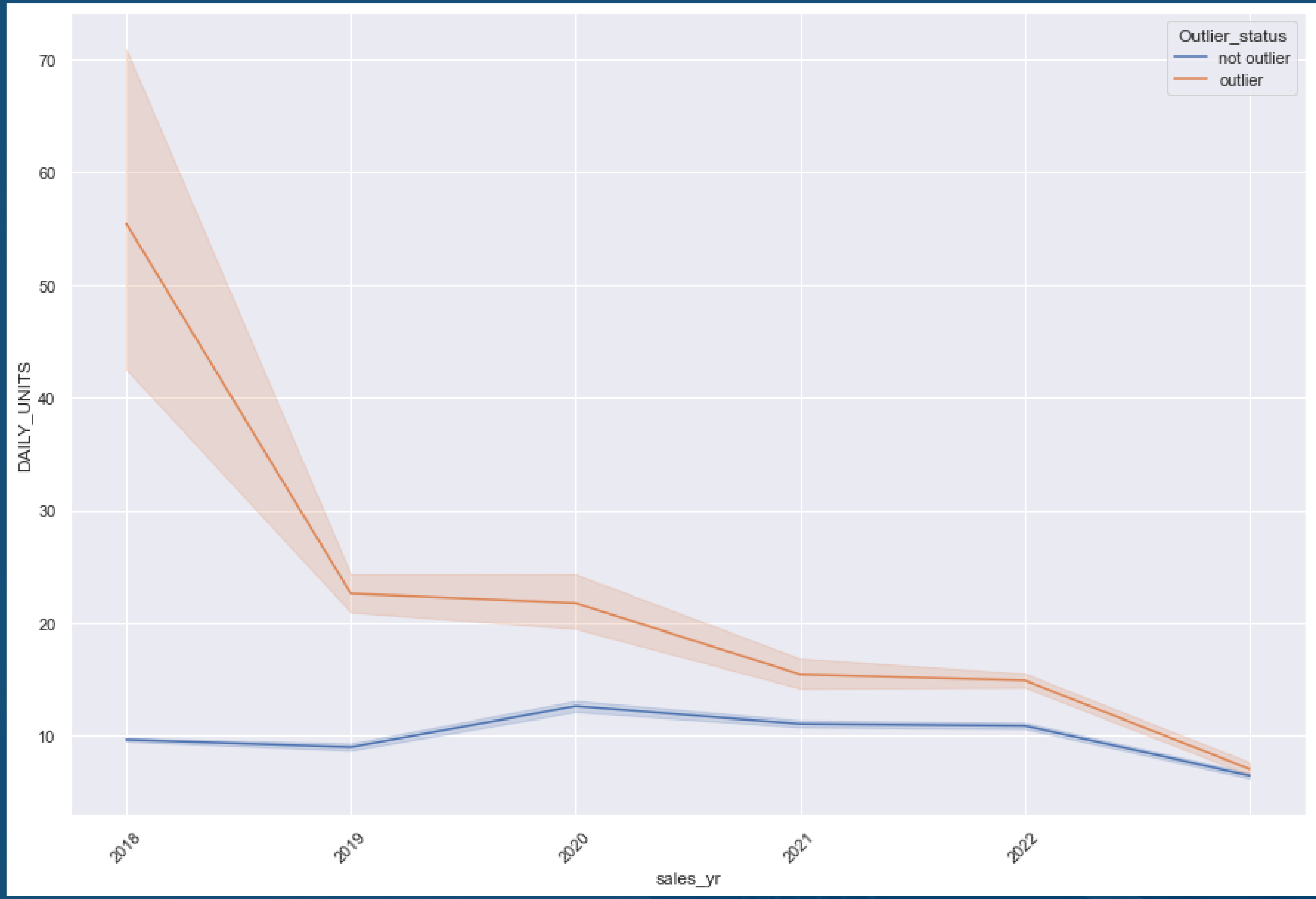
Model Architecture

Evaluation deep dive

- We realised the behaviour of all SKUS were not the same

- There are few SKU's **(26)** which are very slow moving and have large variability as compared to the other **(549)** SKUs and hence there is a high chance that our **XG Boost** model might overfit

- **Croston Model** works better if the demand is intermittent (have a lot of days with 0 demand)



The graph shows the variance in DAILY_UNITS of SKUs

Solution: Using a combination of the 2 models

Modelling the data

Final Model : Ensemble Model

RMSE : 3.49

Based on the best performing model for each SKU, we employ the appropriate strategy to predict for SKUs

Other Metrics:

- MAE: 1.419
- MSE: 12.19
- R squared = 93.4%

Training Time - 33.7 seconds

26 SKU's

XG Boost Model

549 SKU's

Croston Model

23.9% Increase in RMSE wrt XG Boost

Forecasting in Real Life

We have tried to make sure the forecasting is as close to real life and not just optimise for the RMSE score

- Note on available data
    - We realize that data like GSCPI and Consumer Sentiment are normally posted with a lag of a month, hence we have used the lag of these features for practicality.

- We also tested a **Hierarchial Time Series and Prophet** model which gave us great RMSE results but took **~2 hours** to compute.

We also analysed other metrics like MAE, R Squared to make sure we don't just optimise RMSE.