Best Buy Project Week:
Sales Forecast for Slow-selling SKUs

Team: Random Sampler

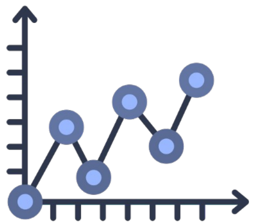
Ashish Dhiman, Anshit Verma, Yibei Hu

Executive Summury

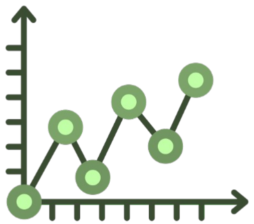
Objective	7-day sales forecast for 539 slow-moving items based on their 4 year performance and other exogenous factors.

Modelling

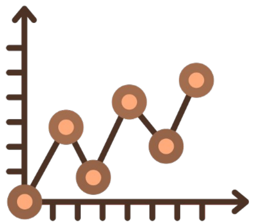
Ensemble Approach



Holt-Winters/ Prophet



STL/MSTL



HMM

Result

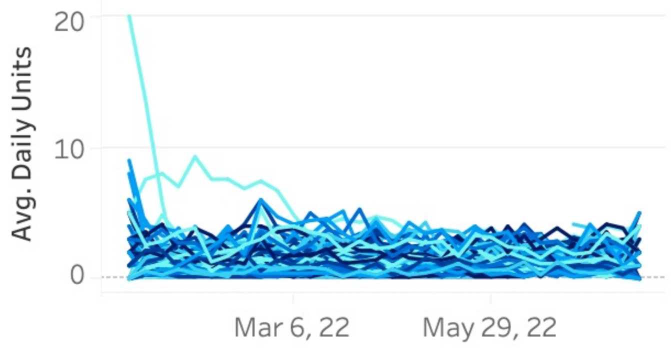
RMSE: 2.21

Run Time: ~ 4H

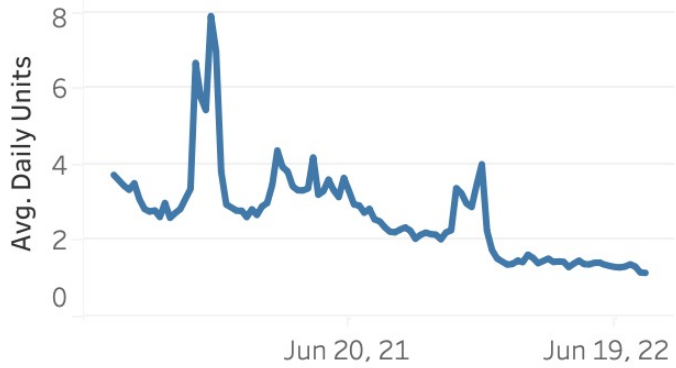
Exploratary Data Analysis: Patterns & Classifications

- **Bucket 1:** low var & low mean
- **Bucket 2:** mid var & low mean
- **Bucket 3:** high var & high mean

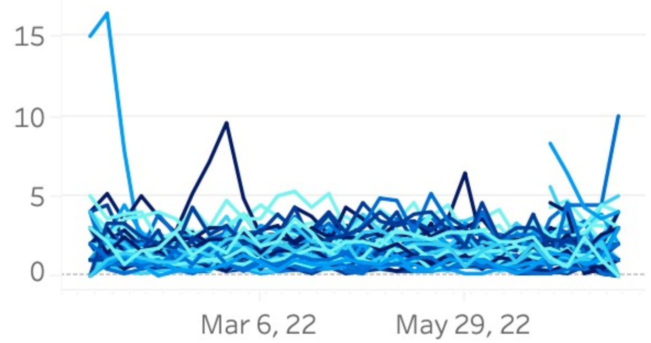
Bucket 1 Individuals



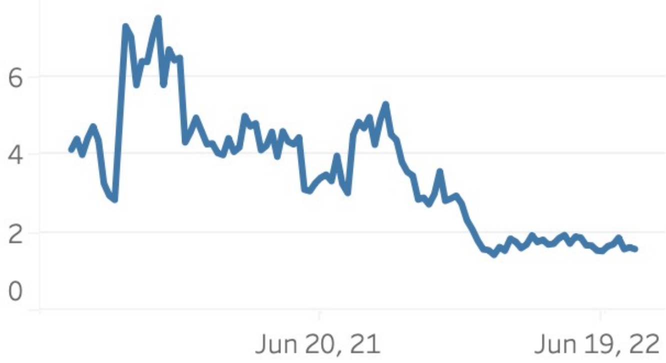
Bucket 1 Aggregate



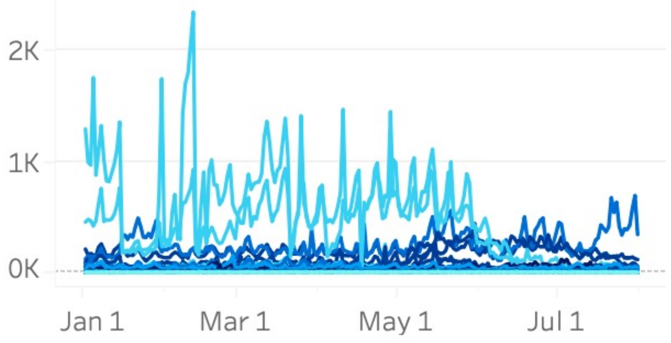
Bucket 2 Individuals



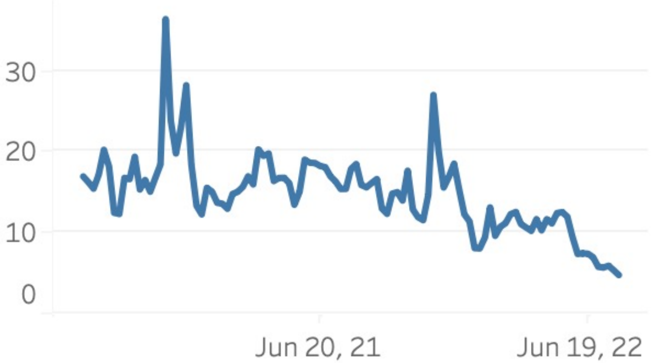
Bucket 2 Aggregate



Bucket 3 Individuals

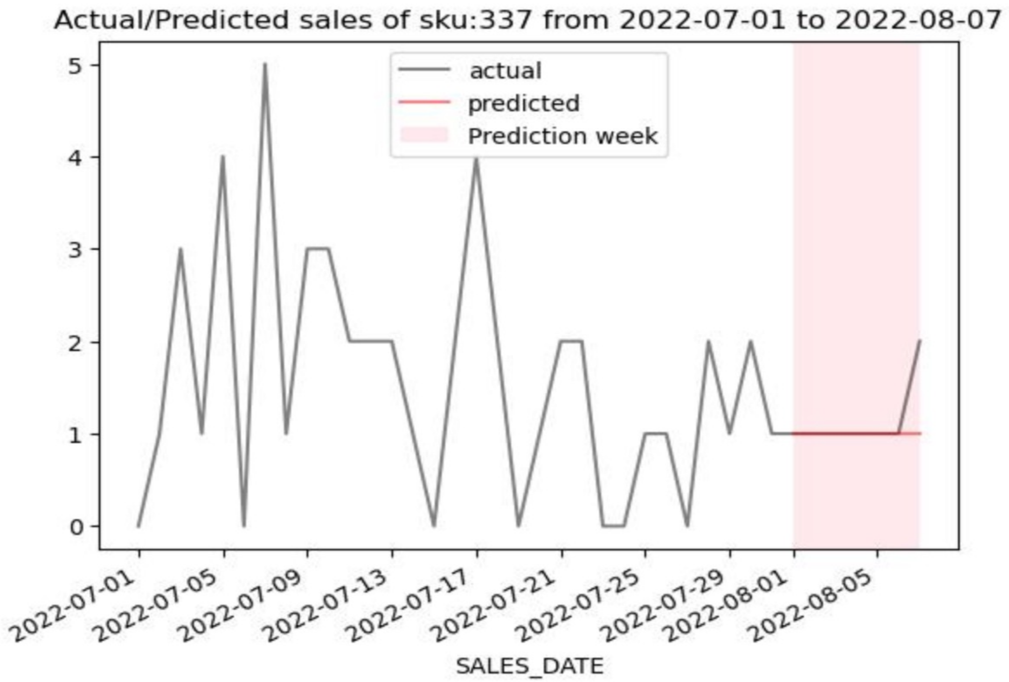


Bucket 3 Aggregate

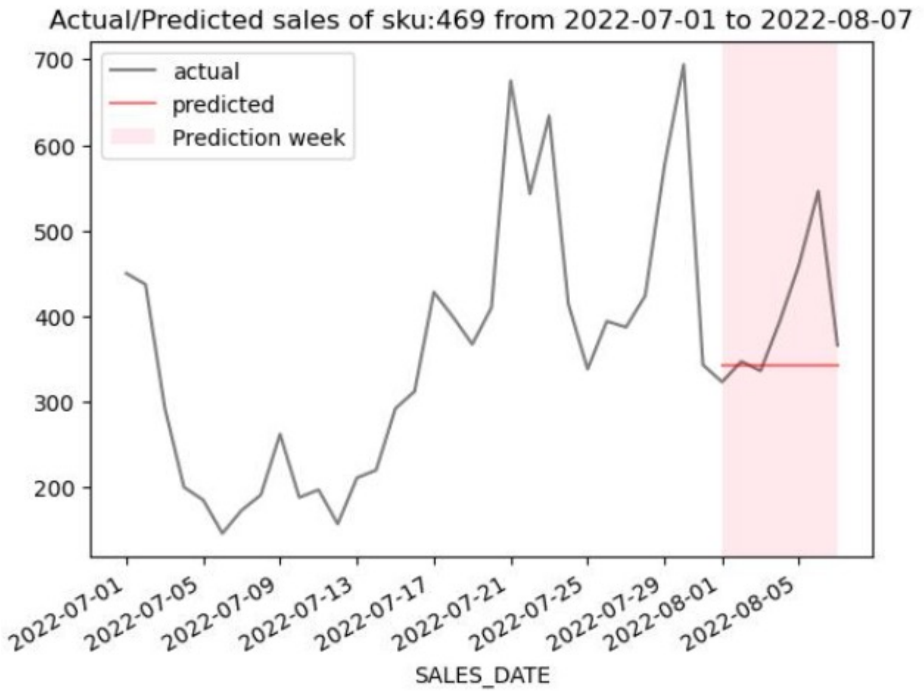


Benchmark: The Null Model

Null Good



Null Bad



Construction:

Naively predicting with last training day's value.

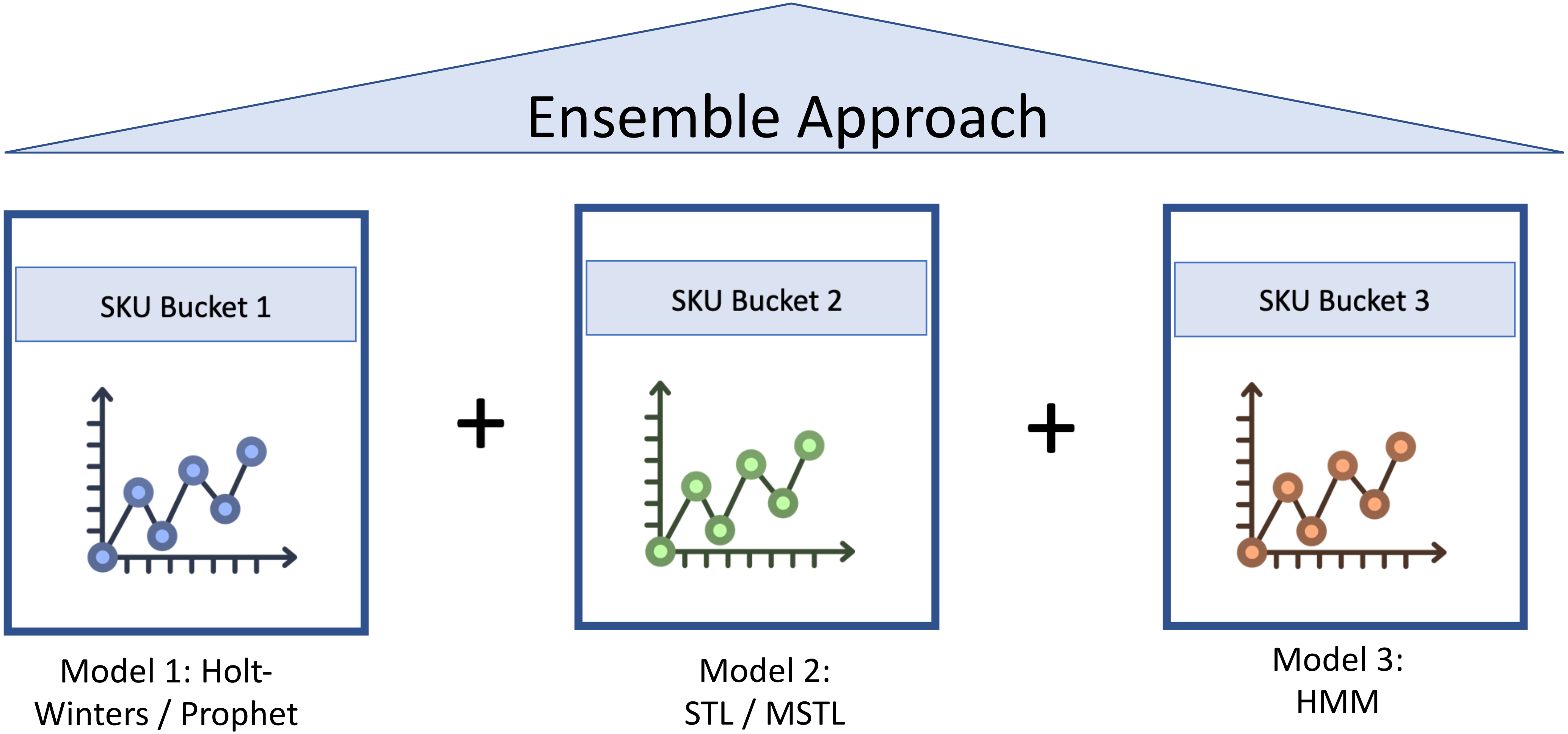
RMSE:

5.29

Performance:

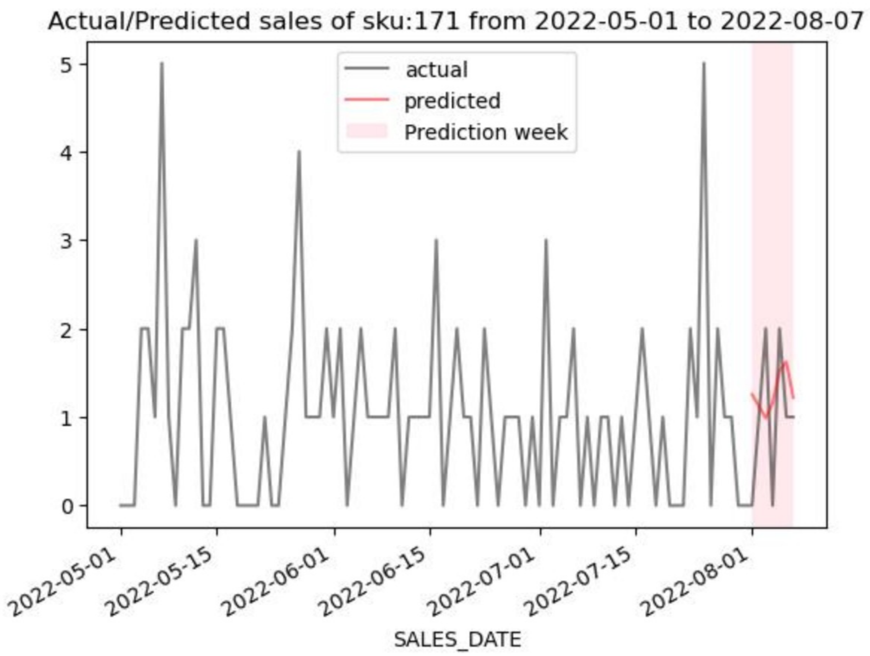
Good in low volatility SKUs

Modeling Approach: Overview of the Ensemble Approach

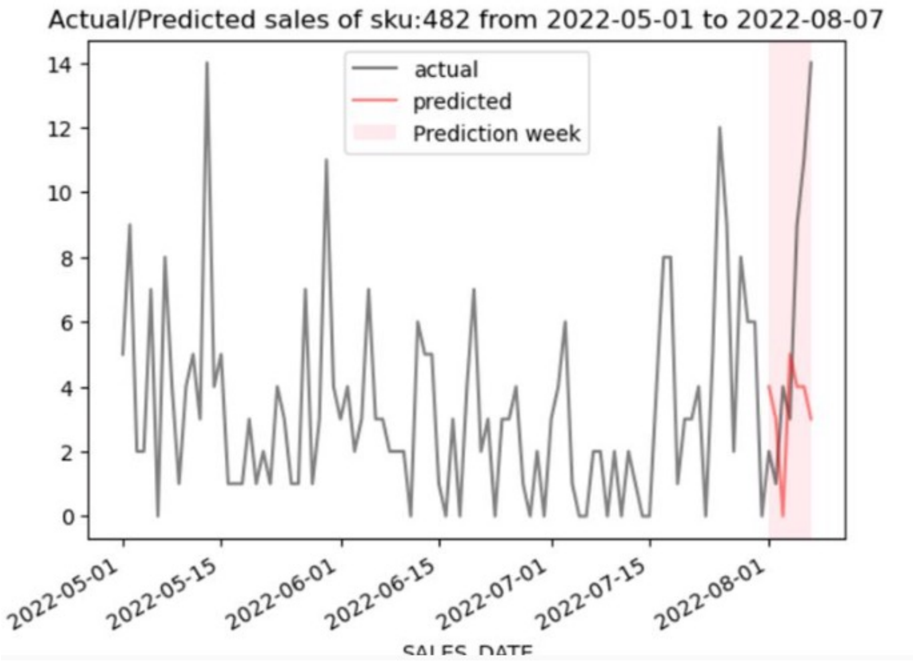


Model 1: Basic Models (Holt-Winters & Prophet)

Holt-Winter Good



Holt-Winters Bad



Bucket Size:

311 SKUs

RMSE:

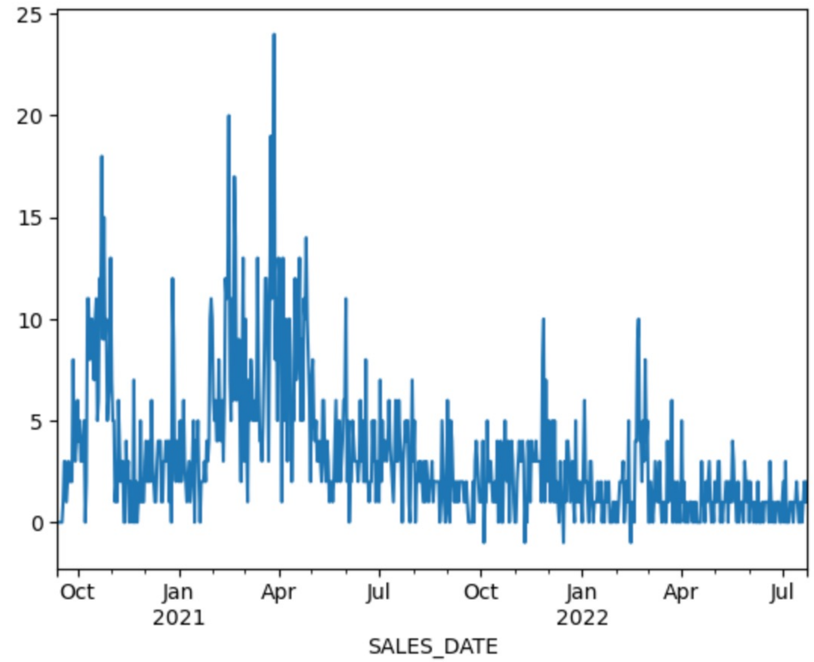
4.83

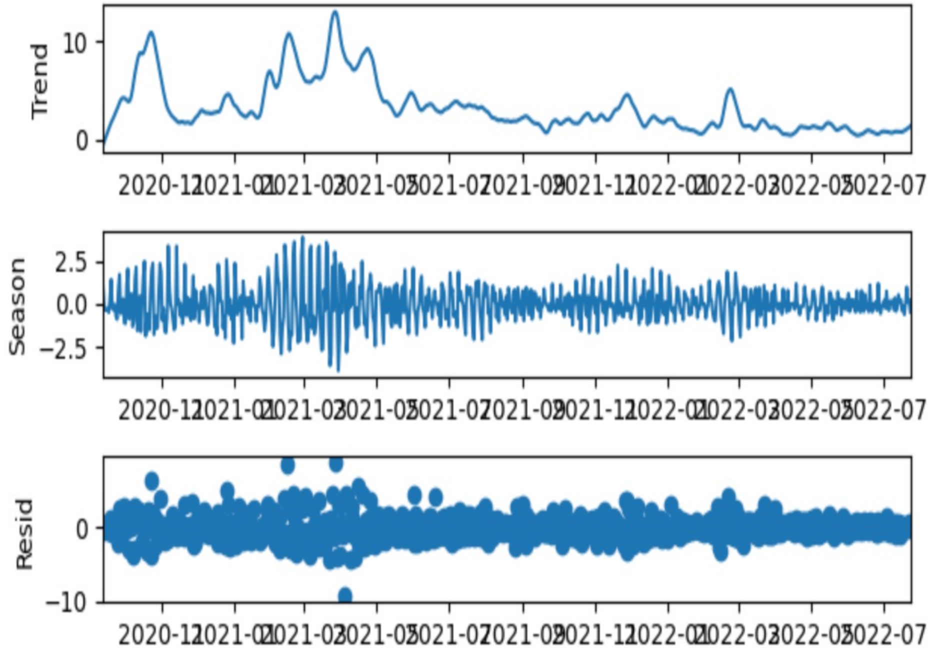
9.6% improvement over null model

Runtime:

~2.7s / SKUs

Model 2: ML Models over STL/MSTL decomposition (1/2)





Auto Regression

+

Xgboost

+

Ridge Regress.

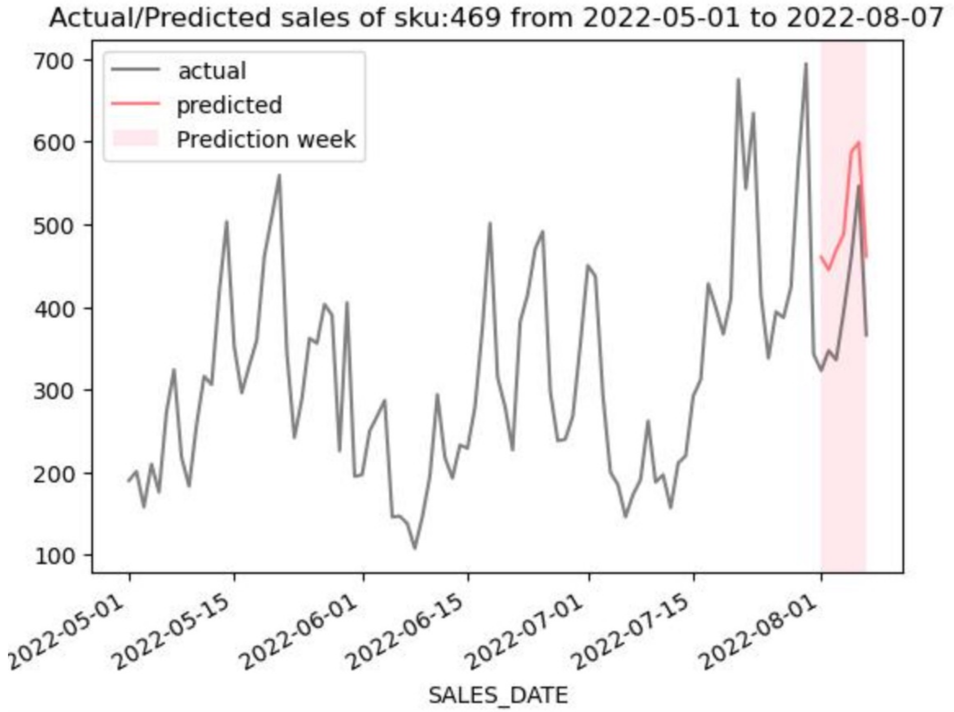
Prediction

Model 2 expands on the philosophy of Prophet by applying STL decompostion to SKUs and using models to indviduall predict those. Holiday variables created from prophet also feed in as exogeneous variables.

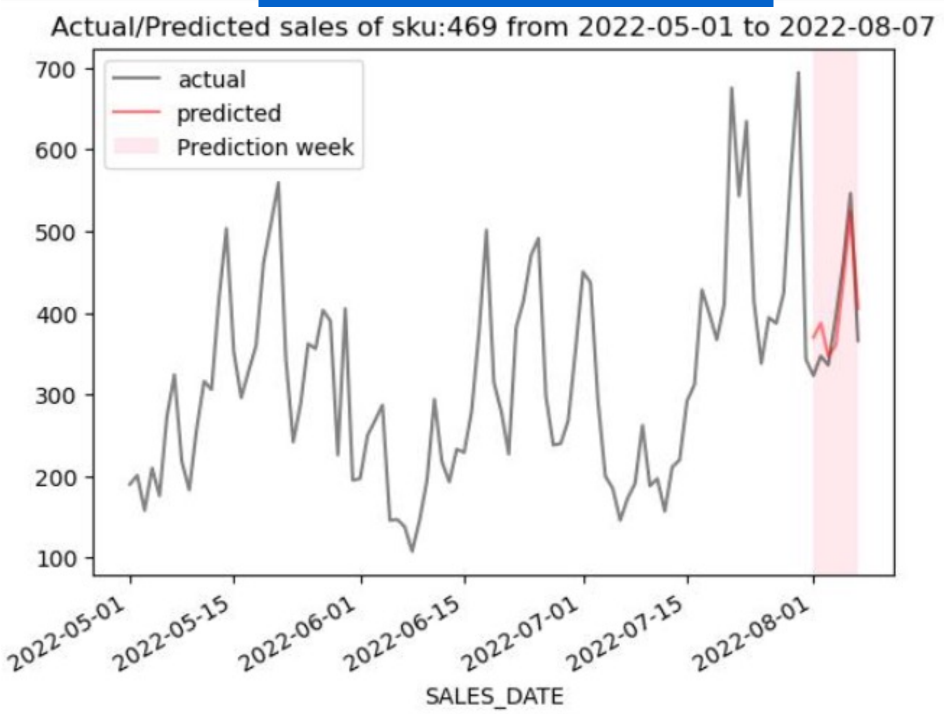
Model 2: ML Models over STL/MSTL decomposition (2/2)

BEST BUY

Prophet



STL



Bucket Size:

80 SKUs

RMSE:

3.23

39% improvement over null model

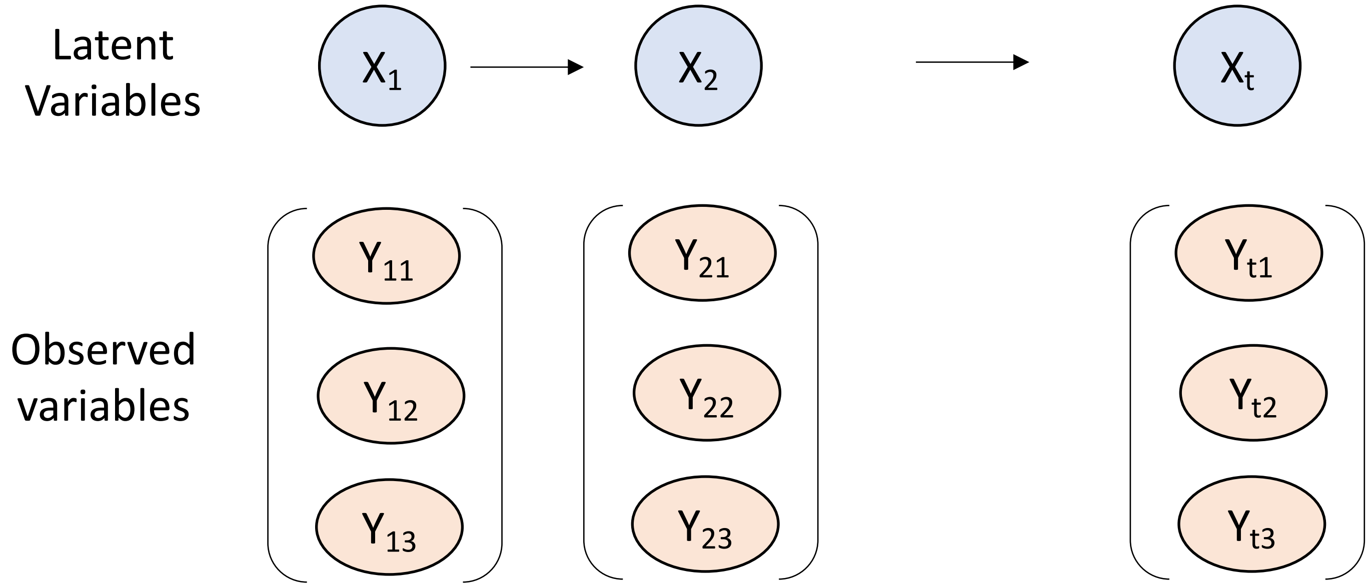
Runtime:

~70s / SKUs

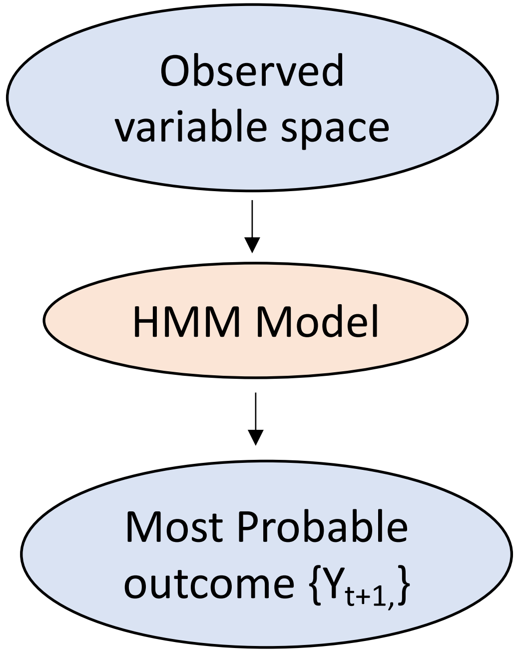
* Includes Prophet, Training + Scoring

Model 3: Probabilistic Models | HMM (1/2)

Hidden Markov Model



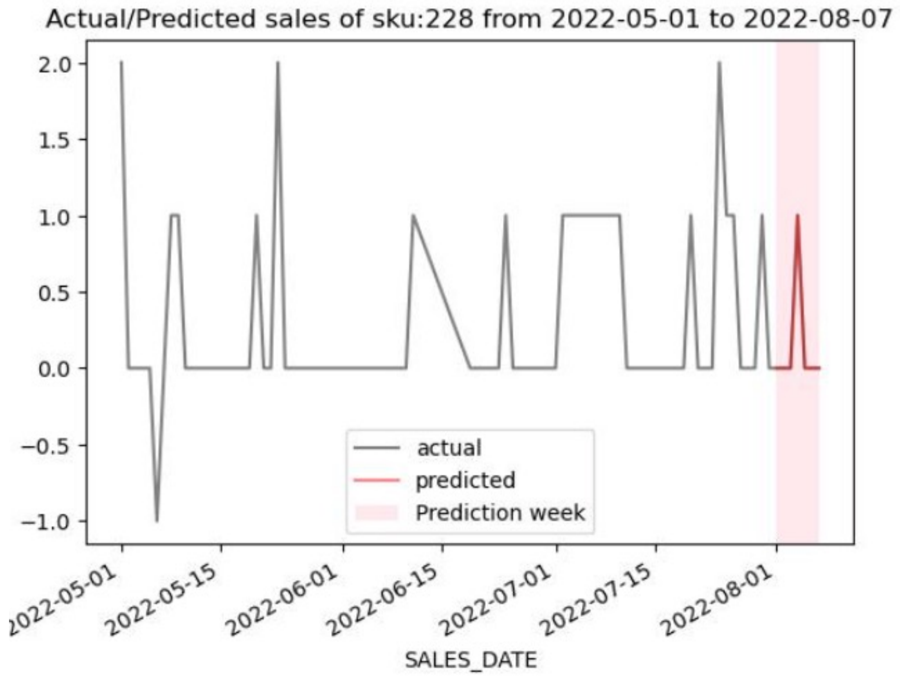
Forecast:



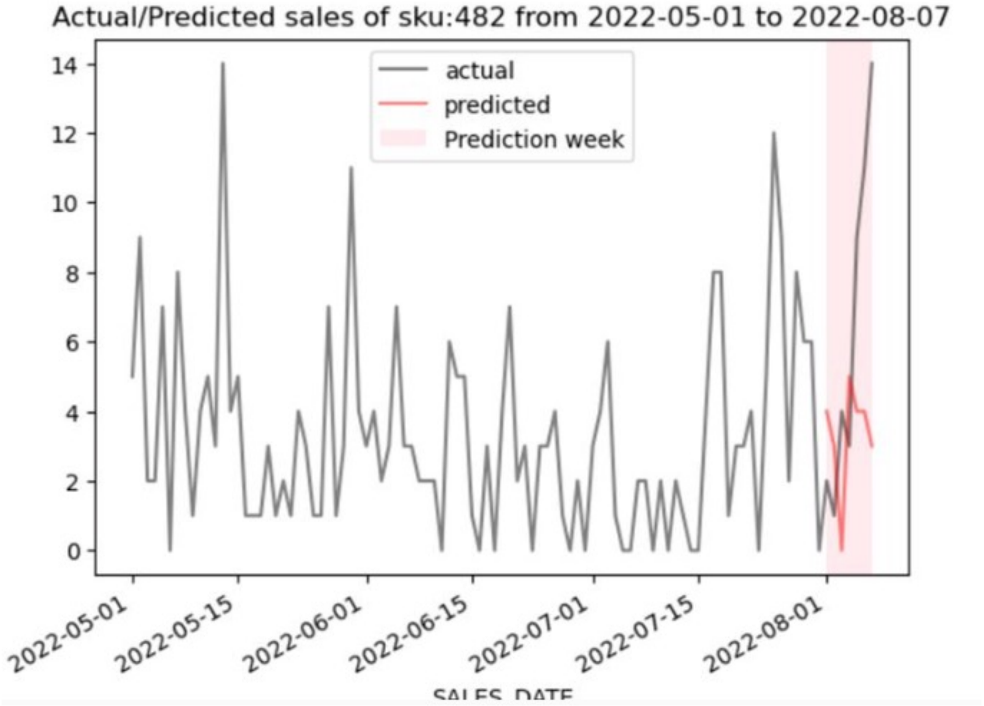
Specific SKUs are modelled as Hidden Markov Chain, with observed variables constrcuted out of lagged daily units. Forecasting is implemented as a search through probable observation variables to find the most probable observation vector.

Model 3: Probabilistic Models | HMM (2/2)

HMM Good



HMM Bad



Bucket Size:

109 SKUs

RMSE:

2.21

58.23% improvement over null model

Runtime:

~80s / SKUs

Training + Scoring

Aggregated Result

RMSE

2.21

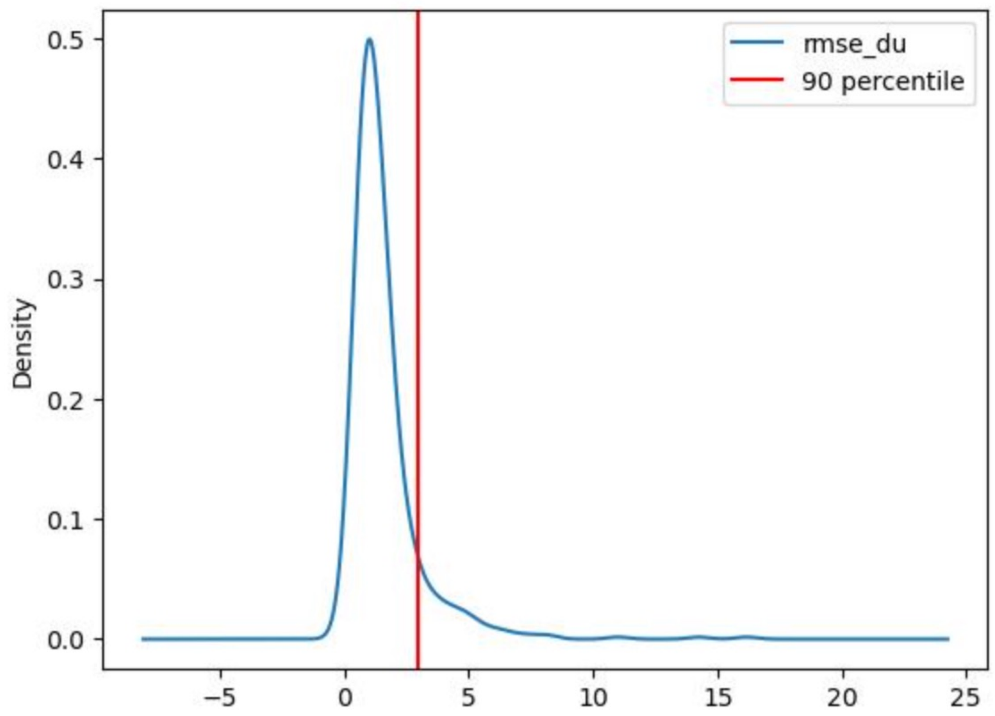
58.23% improvement over null model.
9.6% -Prophet / Holt-Winters
29.4% - STL/MSTL
28.8% - HMM

Run Time

~4H

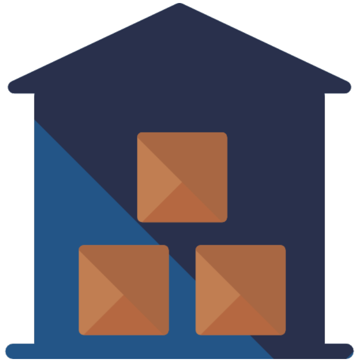
Training + Scoring

RMSE Dist. for single SKUs



The 90% percentile of single SKU RSME is 2.93. The RMSE is right skewed, which distorted the prediction

Business Impact and Future Outlook



Inventory

Better prediction would reduce the cost for inventary storage and purchase. Improve **inventary turnover** while lower **storage cost**.

Distribution

More accurate in time needs prediction would lead to better supply chain decisions that increase **batch delivery** and lower **deliver cost**.

Delivery

More precise prediction would assist customer to make in-time purchase, which would improve long-term **customer loyalty**.

RS

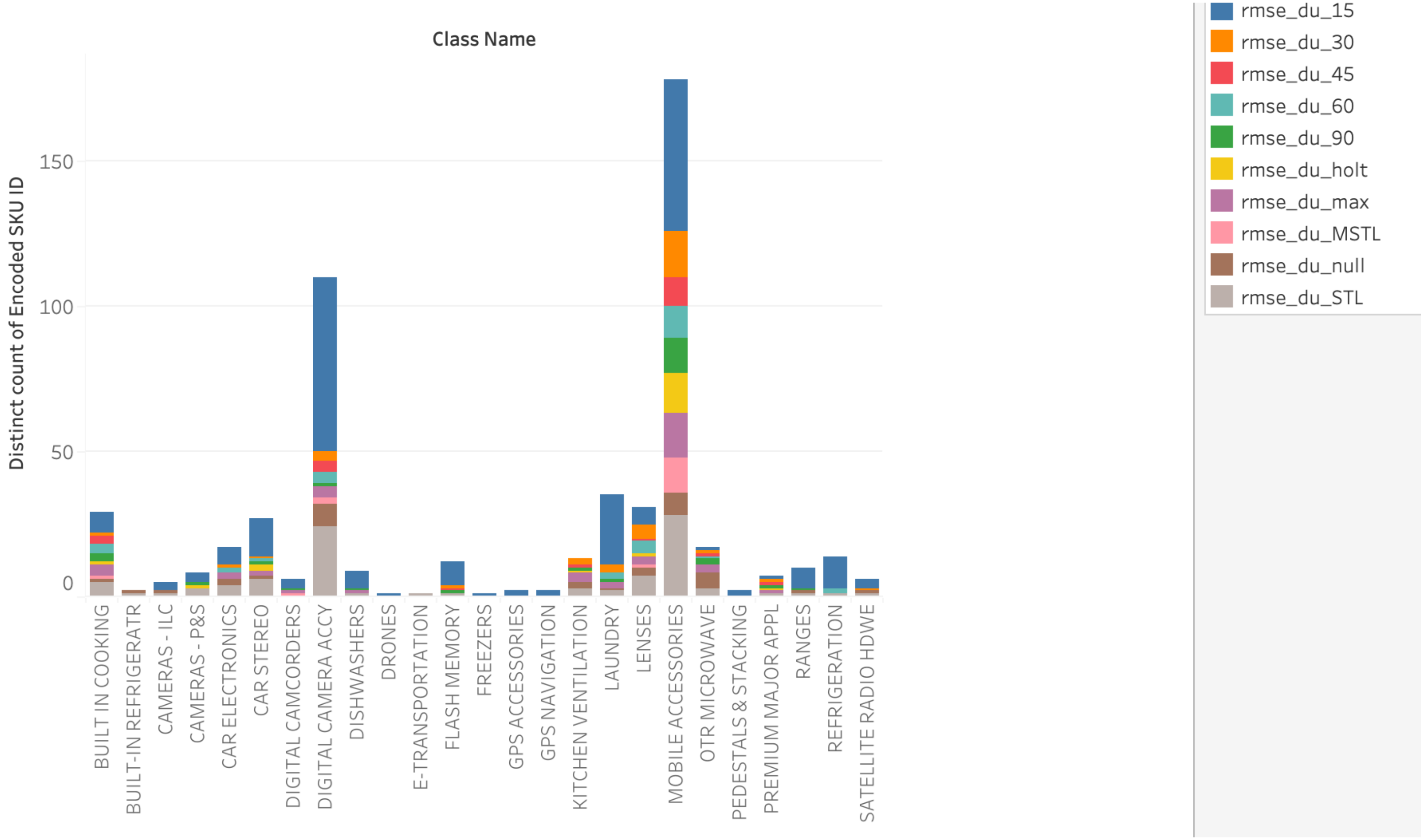
Thank You!

Ashish Dhiman: https://www.linkedin.com/in/ashish1610dhiman/
Anshit Verma: https://www.linkedin.com/in/anshitverma/
Yibei Hu: https://www.linkedin.com/in/olivia-hu9908/

Appendix 1: Run Time Calculation

rmse_du_null	205	2							
rmse_du_prophet	68				null	205	2	410	
rmse_du_STL	65				stl/mstl/prophet	150	70	10500	
rmse_du_15	39				hmm	39	75	2925	
rmse_du_60	30							13835	3.84305556
rmse_du_90	29								
rmse_du_max	29								
rmse_du_30	27								
rmse_du_45	23								
rmse_du_MSTL	15								
rmse_du_holt	9								

Appendix 2: Class/ model mapping



Appendix 3: Each Model's effect in average RMSE reduction(Train data)