Machine Learning to Accurately Forecast Demand

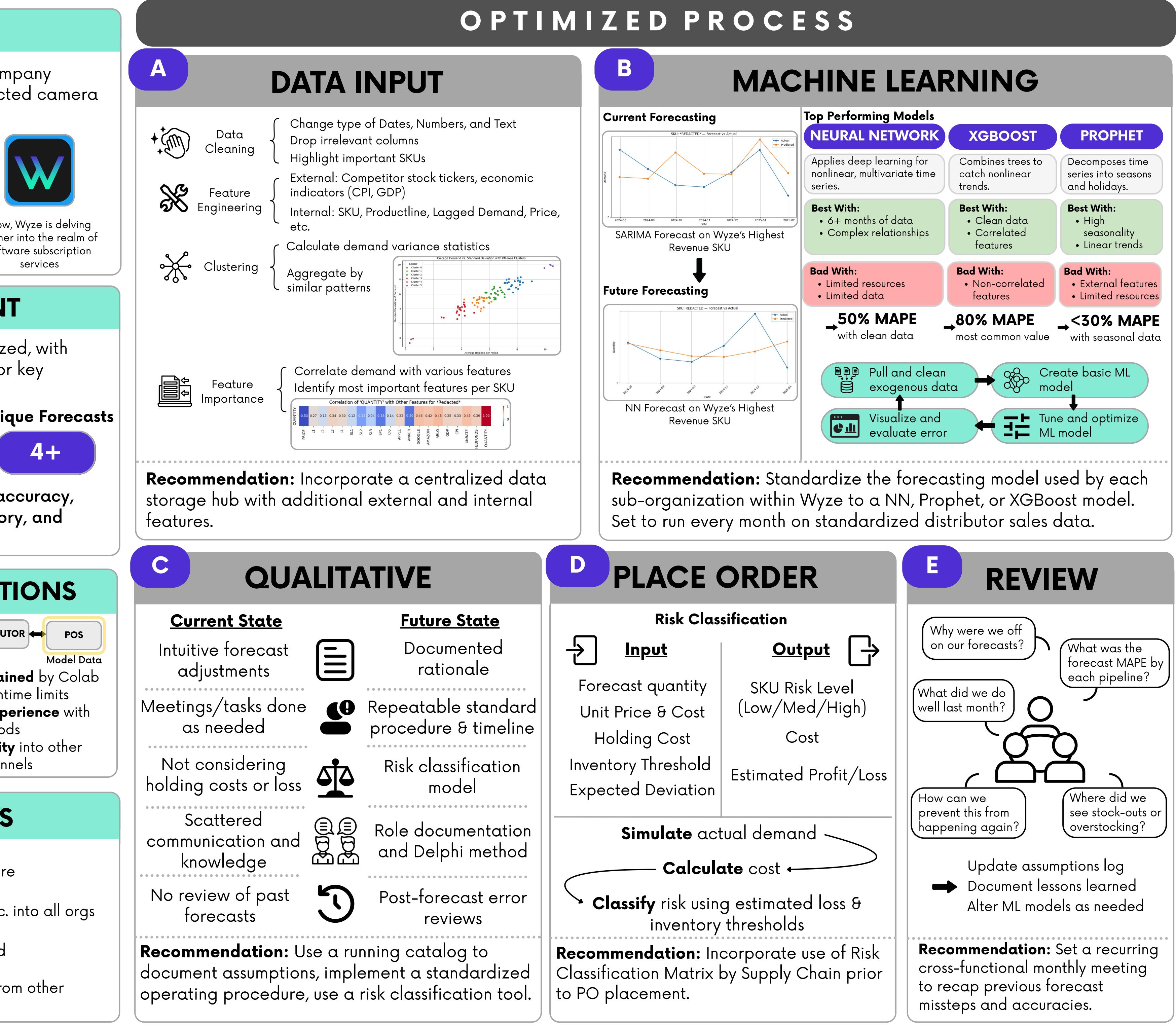**WYZE** | **W**

Thanks to our Project Sponsors: Yuepeng Li, Kevin Wang, Sam Sun, Luke Li, Patricia Cecilia Buchanan

LET'S GET WYZE

Wyze Labs is a smart home technology company founded on manufacturing internet connected camera devices.

Founded in 2017, Wyze was originally created on smart cameras

Starting 2019, Wyze began diversifying into the general smart home area

Now, Wyze is delving further into the realm of software subscription services

PROBLEM STATEMENT

Forecasting is inaccurate and unstandardized, with Wyze frequently experiencing stock-outs for key revenue generating products.
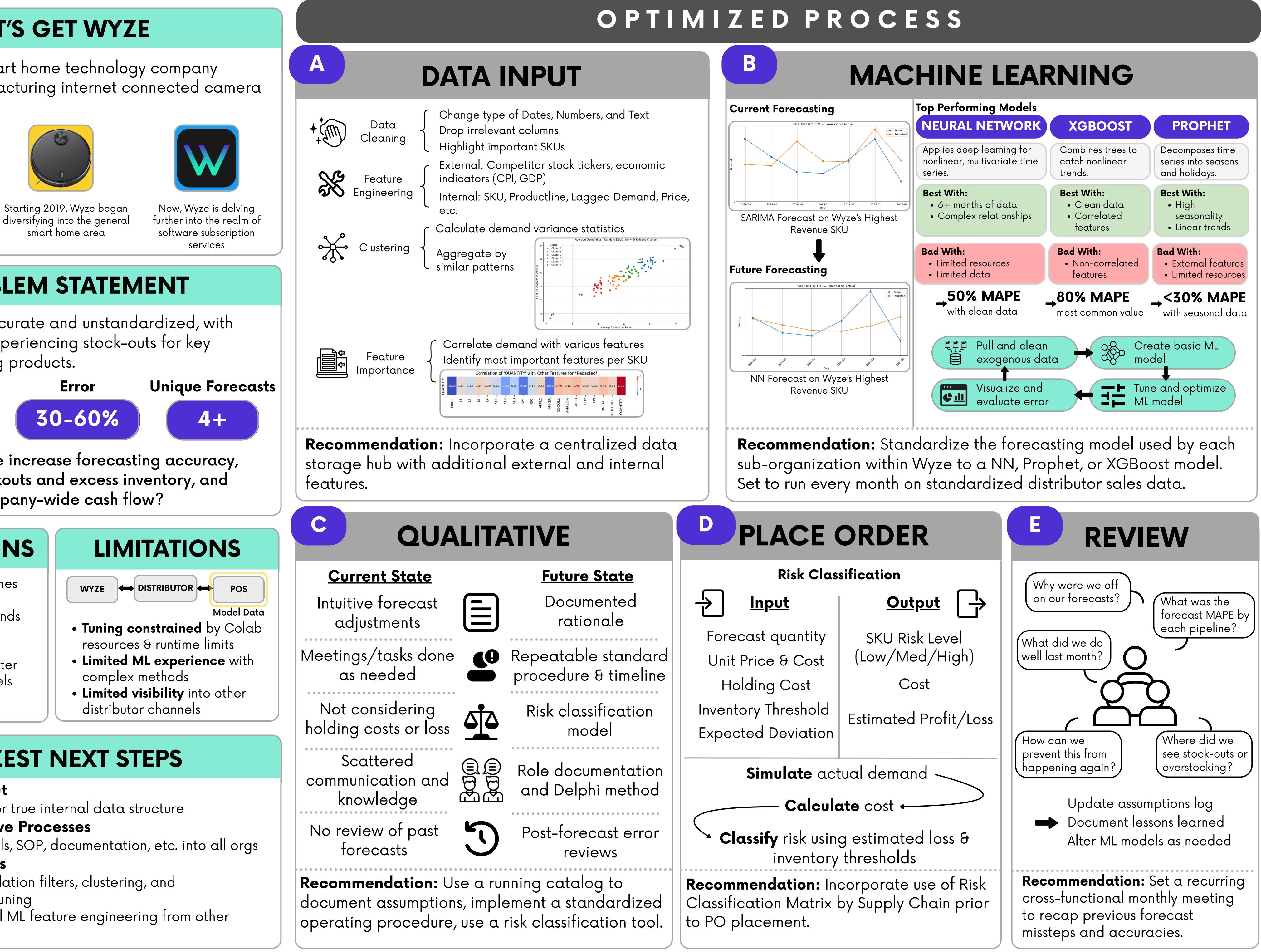
Annual Stockouts	Error	Unique Forecasts
150,000+	30-60%	4+

➡ How can Wyze increase forecasting accuracy, minimize stockouts and excess inventory, and maximize company-wide cash flow?

ASSUMPTIONS

- Combined approaches decrease error
- Amazon demand trends are similar to other distributors
- ML tracks trends better than statistical models

LIMITATIONS

WYZE → DISTRIBUTOR → POS

Model Data

- **Tuning constrained** by Colab resources & runtime limits
- **Limited ML experience** with complex methods
- **Limited visibility** into other distributor channels

WYZEST NEXT STEPS

1. **Adjust Data Input**
   - Adjust ML inputs for true internal data structure
2. **Adopt Qualitative Processes**
   - Incorporate models, SOP, documentation, etc. into all orgs
3. **Optimize Models**
   - Incorporate correlation filters, clustering, and hyperparameter tuning
   - Explore additional ML feature engineering from other datasets

OPTIMIZED PROCESS

A — DATA INPUT

Data Cleaning
- Change type of Dates, Numbers, and Text
- Drop irrelevant columns
- Highlight important SKUs

Feature Engineering
- External: Competitor stock tickers, economic indicators (CPI, GDP)
- Internal: SKU, Productline, Lagged Demand, Price, etc.

Clustering
- Calculate demand variance statistics
- Aggregate by similar patterns

Feature Importance
- Correlate demand with various features
- Identify most important features per SKU

Recommendation: Incorporate a centralized data storage hub with additional external and internal features.

B — MACHINE LEARNING

Current Forecasting

SARIMA Forecast on Wyze's Highest Revenue SKU

Future Forecasting

NN Forecast on Wyze's Highest Revenue SKU

Top Performing Models

NEURAL NETWORK	XGBOOST	PROPHET
Applies deep learning for nonlinear, multivariate time series.	Combines trees to catch nonlinear trends.	Decomposes time series into seasons and holidays.
Best With: 6+ months of data, Complex relationships	**Best With:** Clean data, Correlated features	**Best With:** High seasonality, Linear trends
Bad With: Limited resources, Limited data	**Bad With:** Non-correlated features	**Bad With:** External features, Limited resources
➡ **50% MAPE** with clean data	➡ **80% MAPE** most common value	➡ **<30% MAPE** with seasonal data

- Pull and clean exogenous data → Create basic ML model
- Tune and optimize ML model ← Create basic ML model
- Visualize and evaluate error ← Tune and optimize ML model

Recommendation: Standardize the forecasting model used by each sub-organization within Wyze to a NN, Prophet, or XGBoost model. Set to run every month on standardized distributor sales data.

C — QUALITATIVE

Current State		Future State
Intuitive forecast adjustments		Documented rationale
Meetings/tasks done as needed		Repeatable standard procedure & timeline
Not considering holding costs or loss		Risk classification model
Scattered communication and knowledge		Role documentation and Delphi method
No review of past forecasts		Post-forecast error reviews

Recommendation: Use a running catalog to document assumptions, implement a standardized operating procedure, use a risk classification tool.

D — PLACE ORDER

Risk Classification

Input	Output
Forecast quantity	SKU Risk Level (Low/Med/High)
Unit Price & Cost	Cost
Holding Cost	Estimated Profit/Loss
Inventory Threshold	
Expected Deviation	

Simulate actual demand
Calculate cost
Classify risk using estimated loss & inventory thresholds

Recommendation: Incorporate use of Risk Classification Matrix by Supply Chain prior to PO placement.

E — REVIEW

- Why were we off on our forecasts?
- What was the forecast MAPE by each pipeline?
- What did we do well last month?
- How can we prevent this from happening again?
- Where did we see stock-outs or overstocking?

➡ Update assumptions log
Document lessons learned
Alter ML models as needed

Recommendation: Set a recurring cross-functional monthly meeting to recap previous forecast missteps and accuracies.