Forecasting Category Sales using Data Mining and Predictive Analytics

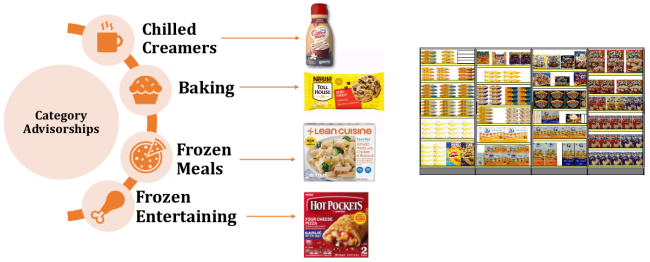
Amie Beckwith, Suelly Samudio Ortega, Joseph Goodman

University of Arkansas College of Engineering Industrial Engineering

Nestlé

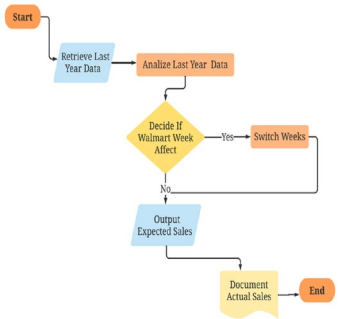
Nestlé USA's Category Advisorships

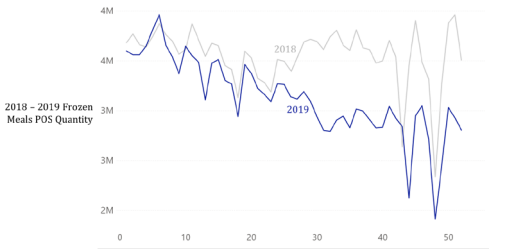
Nestlé USA is a segment of Nestlé, the world's largest food and beverage company. Nestlé USA provides a variety of food and beverage products such as frozen pizzas, chocolate chips, and coffee creamers to retailers across the United States. Nestlé USA's office in Rogers, Arkansas, focuses on its Walmart and Sam's Club operations. Nestlé USA has four category advisorships with Walmart: Baking, Chilled Creamers, Frozen Entertaining, and Frozen Meals. This means that Nestlé USA has been chosen by Walmart to offer professional advice on product assortment for these categories, including products produced by Nestlé USA's competitors.

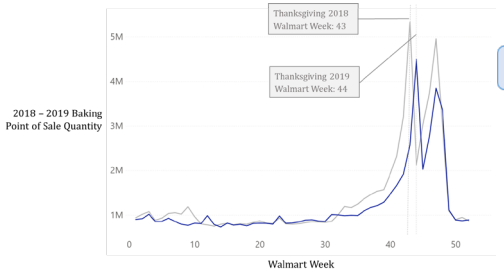


Nestlé USA's Current Forecasting System

Currently, Nestlé USA simply uses last year's sales as an estimate for this year's sales. This system has proven to be inefficient as it is entirely backwards looking. Holiday shifts from year to year and category growth or decline year over year are large contributors to error in the current system.

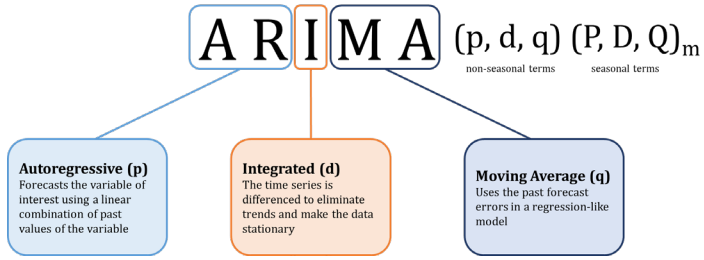






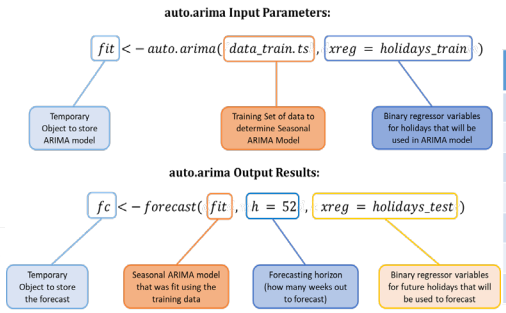
Seasonal ARIMA (Auto Regressive Integrated Moving Average)

To improve Nestlé USA's forecasting system, we first researched potential forecasting techniques. Given the highly seasonal nature of our data, we knew that we needed an approach that took seasonality into account. We narrowed down our focus to the three best options: seasonal ARIMA models, ARIMA models with regressors from Fourier series, and Prophet. Our analysis showed that seasonal ARIMA was the best option. These models incorporate binary holiday variables and detect growth and decline over time. Our analysis shows that the models offer a clear improvement to MAPE from the current system.



Dynamic Forecasting Model

Using the knowledge, we gained from our research on the seasonal ARIMA model, we were able to implement the ARIMA modeling process in a dynamic way so that it can be applied to any level of data input into the model.



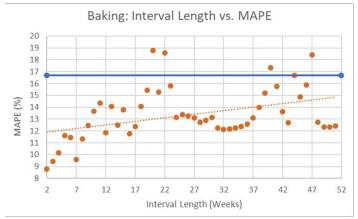
Holiday casual variables

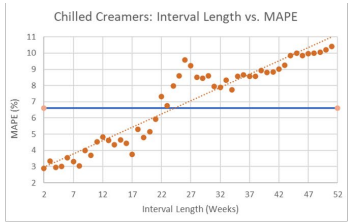
Walmart Week	Thanksgiving – Two Weeks	Thanksgiving – One Week	Thanksgiving
201941	0	0	0
201942	1	0	0
201943	0	1	0
201944	0	0	1
201945	0	0	0
201946	0	0	0
201947	0	0	0

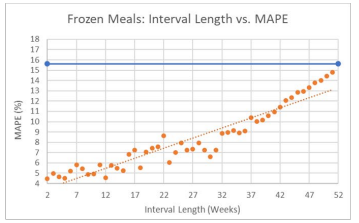
The model is self adjusting based on the data and parameters provided to it, this is done by use of both the "auto.arima" and "forecast" functions in R.. This function takes time series data and automatically fits an ARIMA model (including seasonal models).

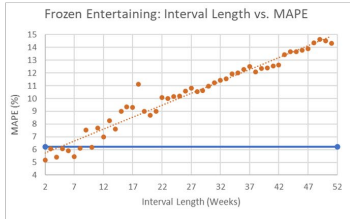
Analyzing the Forecasting Horizon

To measure forecast accuracy at different horizons, we forecasted the entire year of 2019 using different horizon lengths. The graphs show average MAPE values for each forecasting horizon length as well as a solid horizontal line to represent the current system's MAPE for 2019. We concluded that the new system should not forecast past 10 weeks in advance.









R Shiny Forecasting Interface

To add ease to the forecasting process, we also created a user interface for our forecasting models using R Shiny. This interface allows the user to make custom selections and press a button to output the desired forecast.