Insights-driven eCommerce Forecasting

Causal Forecasting

Perspective on Forecasting for eCommerce Demand

Demand planning and forecasting are key business processes to determine the most accurate unconstrained customer demand for the business. Traditional forecasting typically uses time-series statistical modeling to determine upcoming forecasts using historical sales or shipments, while incorporating promotions, seasonality, and new product lifecycle data into the model as offsets or adjustors. The advantage of time-series forecasting is its simplicity and highly interpretable results. Time-series forecasting is the first step for companies to effectively plan for the demand and the appropriate supply required, deliver good customer service, and avoid excess inventory or obsolescence.

Across the world, consumers are shifting channels of purchase to online and realize the convenience of delivery and accessibility to purchase. eCommerce growth has been exponential since the COVID-19 pandemic, with the first half of 2020 seeing an increase equivalent to that of the previous ten years . To further emphasize this growth, total online spending in May 2020 was up 77% year-over-year, hitting $82.5 billion in the United States . As the eCommerce business grows and becomes an increasingly relevant sales channel over and above the traditional brick and mortar business, effective forecasting for eCommerce sales becomes a priority. Unlike the traditional retail distribution channel with lengthy and stable historic data and forecastable trends, the eCommerce business adds additional complexity. Volatile sales, customer hoarding, explosive growth, change in product management and change in product mix are the few but prominent challenges when forecasting for eCommerce. The above issues are compounded by less historical data for building the eCommerce forecast, and greater sensitivity of eCommerce sales to competitor pricing and promotions. Retailers are responding to this issue by aligning more resources to improve business planning for eCommerce and thus improve eCommerce margins. Ultimately, the traditional forecasting method is inadequate in its ability to accurately predict eCommerce demand. This has caused both inventory excess and shortages, and higher supply chain costs than required. These issues could have been alleviated with better forecasting mechanisms and use of eCommerce data that is readily available.

Improving demand planning and forecasting can provide a company an important competitive advantage in the marketplace and contribute to its commercial success.

Understanding New Data Elements to Augment Forecast

As consumer purchase behavior and methods of purchasing change rapidly in the ever-changing retail space, companies must tap into existing and novel data sources to better forecast demand and enable efficient omni-fulfillment.

Often companies leverage traditional inputs (Figure 1), such as Sales Orders, Sales Stock Outs, Actual Demand History, Product Category and Product Attributes, to forecast for upcoming demand. Organizations are overly dependent on the demand history, leaving the effect of other significant demand drivers such as price, promotions, holidays and special events to manual adjustments. Companies tend to be most reactive to the latest customer trends and to oversteer the forecast based on the most salient issues, such as a change in price, while not having any ability to statistically adjust the forecast for multiple simultaneous factors. Many supply chain planners find themselves spending time manually adjusting forecasts, a process that is cumbersome, error-prone and potentially filled with bias, instead of systematically integrating it mathematically. That time lost to manual analytics has an additional opportunity cost in that the time and attention could have been focused on other opportunities to better manage the supply chain.

Figure 1: Types of Data Elements

Companies with more mature planning processes in place such as Sales and Operations Planning (S&OP) tend to have collaborative planning. Merchandising, Sales & Marketing, Supply Chain and Finance share the latest knowledge on promotions, price changes, marketing activities and trends to better capture all foreseeable demand. A collaborative planning method can bring forth new data elements and forecast considerations. Integrating these insights in a structured, cohesive, and consistent manner often requires new processes and capabilities.

Fewer players in the market make use of the novel data elements. Companies with eCommerce as a channel of distribution often capture the eCommerce Point-of-Sale (POS) and can use POS data to adjust the demand plan. However, additional causal factors, such as website traffic, ad spend or price and promotions, short-term weather forecasts, social commentary or competitor pricing are data types that many companies struggle to have access to, or to know how to use strategically in forecasting the demand. With the highly competitive eCommerce environment and 24/7 accessibility for a customer to compare prices and complete purchases, companies must improve their leverage of novel data elements to better predict customer demand.

Competitor data, such as aggregate eCommerce demand, social listening, and competitor/marketing presence, can be vital factors impacting the demand of a product. Companies leverage media scrapping tools to monitor the latest trends and highlight the source of information. Social listening activities provide brand insights in a much timelier manner. With ever-growing social media platforms and rapid changes to consumer trends, companies need real-time direct feedback from their customers to understand true demand.

By incorporating additional relevant data elements with the right methods, it can augment visibility, improve overall forecast accuracy, and convert to online sales.

Proof of Concept: Causal Forecasting for eCommerce

Through a proof-of-concept experiment, Deloitte sought to prove that by using modern multivariate algorithms and the often-underutilized eCommerce data, organizations can improve their forecast accuracy on eCommerce channel demand.

Participating organization: A leading North American mattress and sleep accessory retailer.

Challenge: Having launched its eCommerce channel only months before the COVID-19 global pandemic, the company – like many others – faced a growth of online demand that exceeded expectations. With online shopping expected to continue playing a significant role in the post-pandemic era, Deloitte sought to uncover an alternative approach that can better predict the volatile eCommerce demand with limited referenceable demand history.

Goal: Deloitte partnered with Logility on a proof-of-concept exercise with the objectives of:

- Infusing new demand signals and SKU / category insights to adapt to eCommerce channel
- Evaluating and uncovering data features that impact forecast accuracy
- Quantifying how causal models that are infused with the novel data elements can improve eCommerce forecasts

Scope: This proof of concept used Google Analytics data that is wholly owned and supplied by the organization's eCommerce channel, including:

- Products – 278 SKUs across 22 product categories
- Reference - weekly eCommerce sales in Canada
- Timeframe – The model learned from a 64-week historical planning horizon (Feb. 2020 to May 2021) to identify trends and measure the relationship between different measures and the demand signal. Then, actual demand and predicted demand were compared across two six-week testing periods to evaluate the model's performance (Jun. 2021 to Aug. 2021) as shown in Figure 2.
- Technique – An ensemble of causal forecast models, including machine learning models (e.g., random forest) and statistical models with exogenous inputs (e.g., ARIMAX).

Partnership: This proof of concept was conducted in partnership with Logility, where we collaborated to:

- Identify success criteria for the proof of concept
- Define and extract pertinent client data
- Build demand forecasting model using Logility based on data provided
- Iterate model parameters and fine tune results
- Derive insights and root cause hypotheses

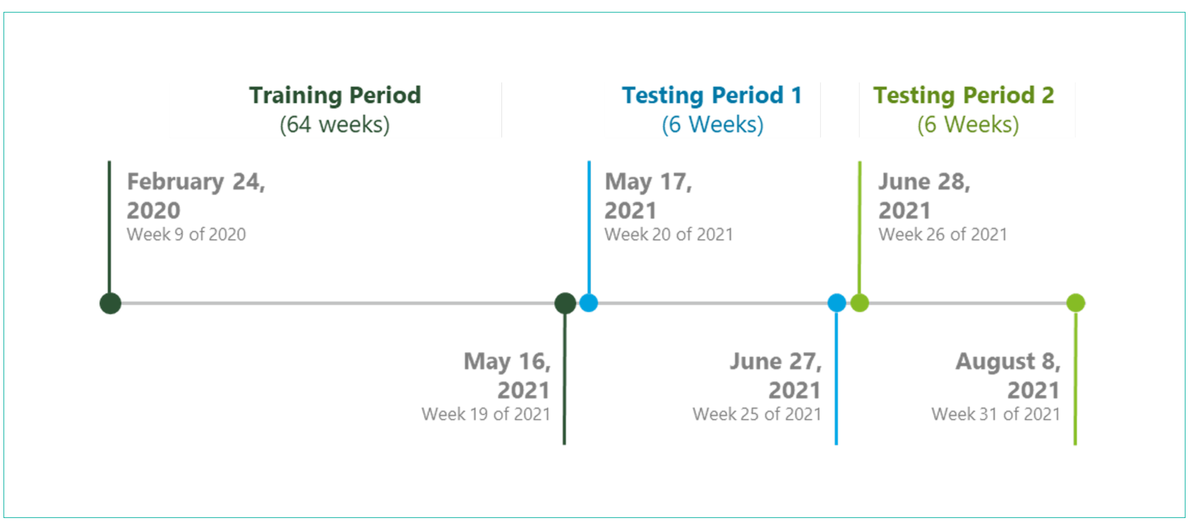

Figure 2: Project Timeframe

Approach: Logility's causal forecasting solution enables the integration of demand history and novel data elements, while providing a platform to explore and build demand forecast models from a library of traditional time-series models and causal models that can infuse causal inputs. This solution enables users to investigate multiple endogenous and exogenous factors influencing demand and how they can improve forecast accuracy, using a four-step approach (Figure 3).

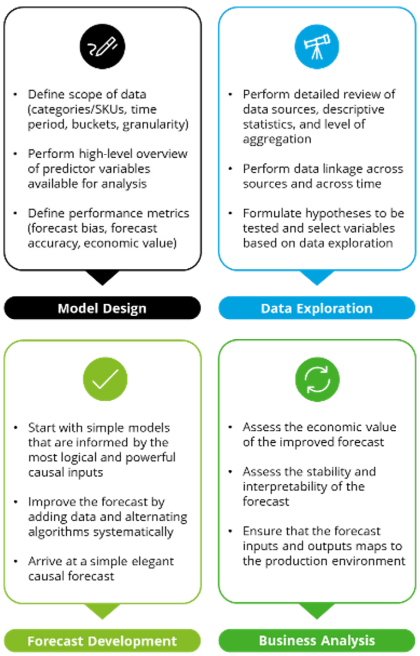

Figure 3: Causal Forecast Method

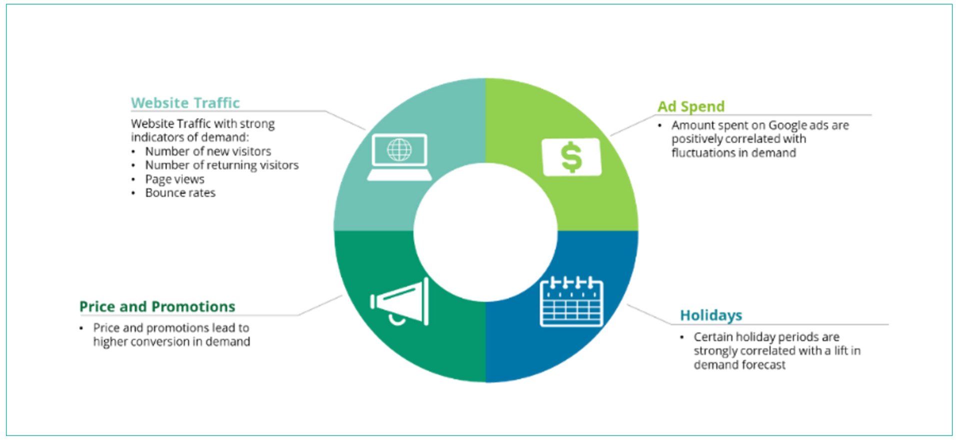

Figure 4: Initial Causal Factor Hypotheses

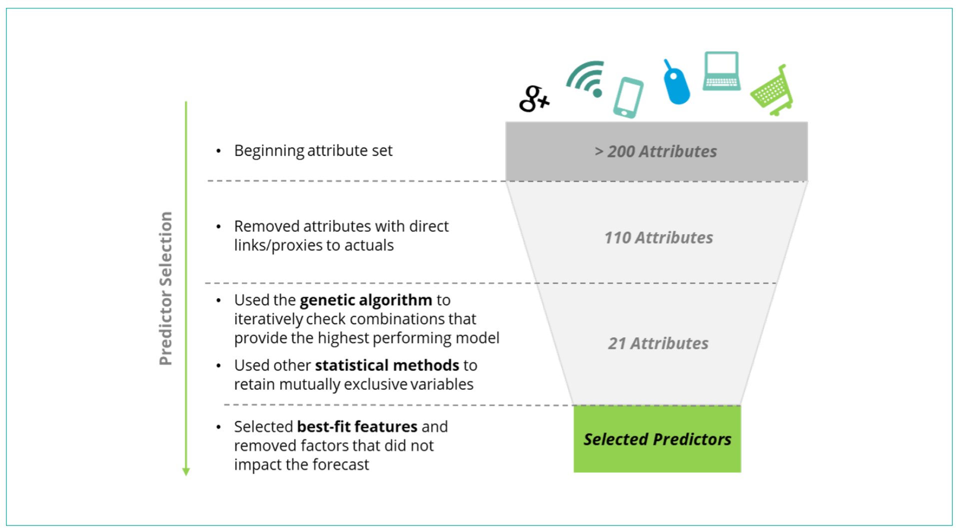

Figure 5: Predictor Selection Process

Once the scope was defined (Model Design), four hypotheses were defined as potential eCommerce demand conversion drivers (Data Exploration): Website Traffic, Price and Promotions, Ad Spend, and Holidays (Figure 4). Among over 1,000 available data measures on the client's Google Analytics platform, 200 were identified as key contenders and were retained for hypothesis testing.

Through further data analysis, feature engineering, model robustness and accuracy assessments (Figure 5), eight out of the 200 data attributes demonstrated the strongest performance for predicting units sold for 278 SKUs across 22 product categories (Figure 6).

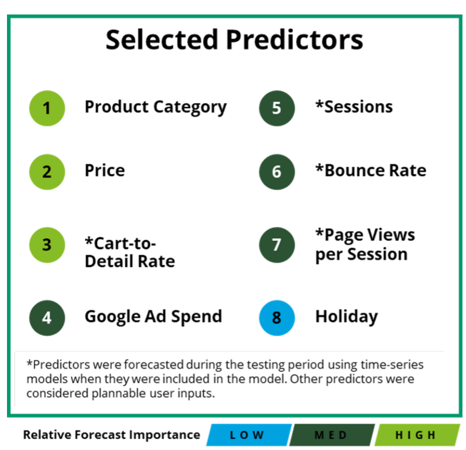

Figure 6: Selected Predictors and Forecast Importance

A restricted machine learning model was used to forecast the independently plannable measures, while ARIMA was used to statically forecast the forecastable measures ("*" in Figure 6). These measures were integrated into initial causal forecast models developed using ARIMAX . The causal forecasts were evaluated against traditional time-series forecasts based on variance from actuals as well as aggregated weighted Mean Average Percent Error (MAPE) at the Product (SKU) level (see Figure 7 for more on the key differences between traditional and causal forecast models).

Statistical Forecast
(Tournament Ensemble)

The tournament algorithm compares traditional statistical forecasting models (ARIMA, HoltWinters, Croston's, Narve, Exponential Smoothing), and then forecasts using the best fitting model. It only requires historical data as input.

Advantages: Useful for benchmarking accuracy before introducing causal inputs, simple and highly interpretable, adaptive

Disadvantages: Assumes stationarity, unable to respond to fluctuations and non-linear relationships

Causal Forecast Models

Machine Learning Models are the most flexible model and can integrate the largest volume of causal inputs. Inputs can be categorical or continuous.

Advantages: Automated tuning and trend/pattern identification, continuous improvement with more data, can learn non-linear relationships, robust to outliers

Disadvantages: Requires large amount of data, interpretation of results, susceptibility to bias, limited capacity to extrapolate

ARIMAX uses a combination of time series with regression causal inputs that must be continuous.

Advantages: Straightforward to understand and explain, can extrapolate in scenario testing

Disadvantages: Requires more expert knowledge and tuning skills, assumes linear relationships with predictors, need to manually include the right interaction terms/polynomials (requires inslght), unable to incorporate categorical variables

Figure 7: Statistical vs. Causal Forecast Models

The forecast validation and performance/accuracy monitoring cycles were conducted using different configurable dashboards supplied by Logility, including:

- Causal Forecast Review Dashboard (Figure 8) for a quick accuracy check during the validation period . This dashboard includes a Model Contribution window, outlining the variables used and their relative importance to the forecast model output.

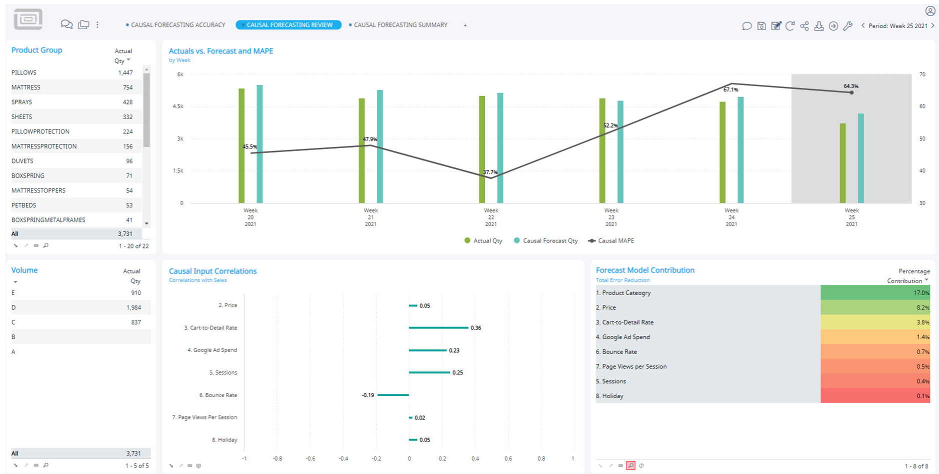

Figure 8: Logility Dashboards - Causal Forecasting Review

- Causal Forecasting Accuracy Dashboard (Figure 9) for a deeper dive into forecast accuracy by week and over selected time periods. This dashboard allows for a MAPE comparison between Base and Causal models displaying the Forecast Value Added (FVA).

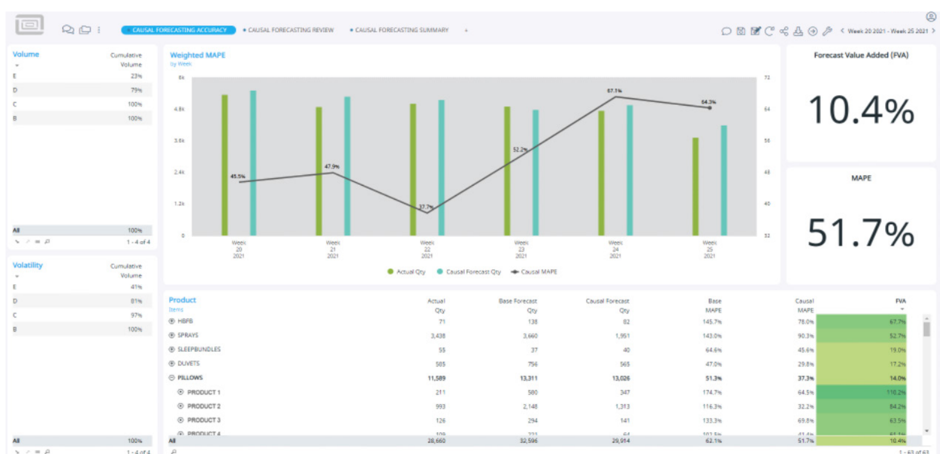

Figure 9: Logility Dashboards - Causal Forecasting Accuracy

- Causal Forecasting Summary Dashboard (Figure 10) to provide a high-level summary of projected demand by product group and item. This dashboard includes multiple slices and drill down capabilities to leverage all historical data at the most granular level.

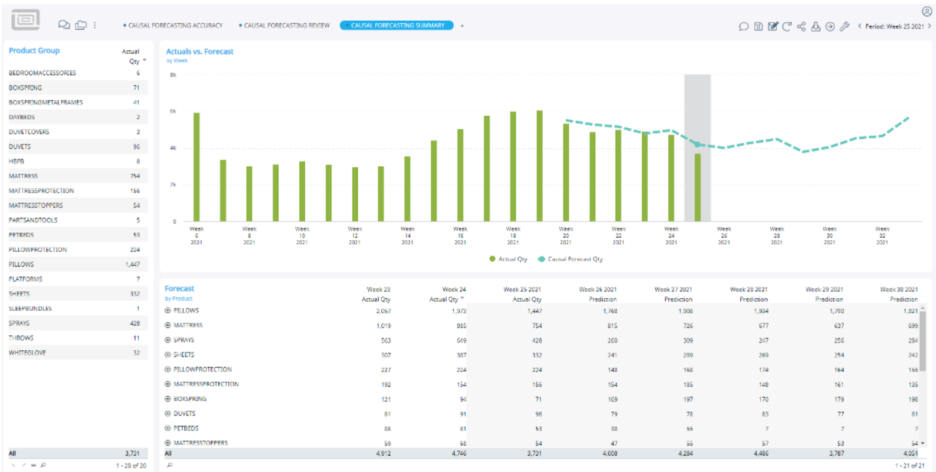

Figure 10: Logility Dashboards - Causal Forecasting Summary

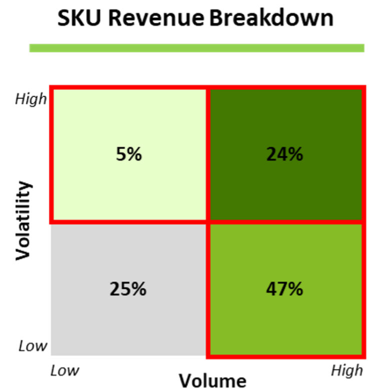

Figure 12: SKU Revenue Breakdown

Forecast diagnosis and tuning were performed based on segmentation by volume, variability and by the statistical algorithm that best fits the data.

Results:

We evaluated the value-add of novel data elements and advanced algorithms by comparing the weekly forecast accuracy of the best-performing causal model against the best performing time-series forecast. The models were compared over two independent six-week testing periods.

The best performing causal model across 278 SKUs consistently exhibited improved forecast accuracy based on a reduction of the weighted MAPE, with an overall improvement in forecast accuracy increase of 10% for all products (see Figure 11).

The most significant impact was observed on high-volume and/or high-volatility SKUs, which make up 75% of total revenue in the sample eCommerce sales data set (see Figure 12). Both categories have individually shown an improvement of 13% over statistical forecasting with regards to forecast accuracy (Figure 11), with the combination of both high-volatility and high-volume SKUs having the most significant forecast value add benefits at 16.1%. This analysis highlights the importance of having the ability to efficiently classify item groupings and drill down to categories impacting forecast accuracy through the right tools.

The proof of concept has shown promising results on the short-term eCommerce demand forecast (first six weeks). Beyond that horizon (i.e., for the second six-week testing period), there is a sharp degradation in forecast accuracy regardless of the model used given the volatile nature of eCommerce.

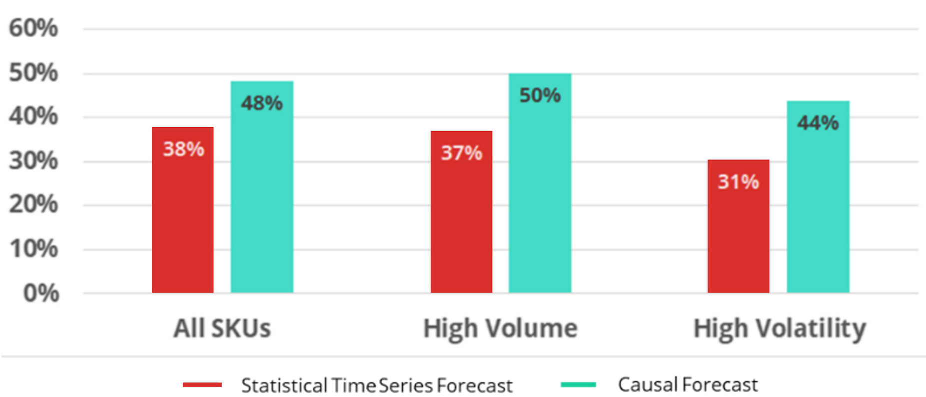

Figure 11: Forecast Accuracy Based on MAPE
(Statistical Time Series Forecast vs. Best Performing Causal Model)

Conclusion

As a result of the surge in eCommerce business and the channel shift, consumer business companies are forced to:

- Increase responsiveness to the changing product volume and mix
- Anticipate changes to demand patterns proactively
- Fulfill the omnichannel demand with eCommerce taking a larger share of the overall business and driving more volatility

Predicting demand patterns is a challenge magnified by the digital channel, driven by how consumers interact with eCommerce and unmeasured actions of competitors. The eCommerce channel comes with its own unique levers that strongly correlate to sales performance, can be measured with high precision, and serve as valuable inputs to demand forecasting. Adequately managing these data points, especially within a company's own organization, and leveraging them for forecast management will foster more effective and more profitable operations.

The proof of concept unlocked the potential of channel-specific indicators by introducing them into enhanced forecasting methods, such as a causal model. Consumer businesses with eCommerce channels should explore alternative approaches, leveraging data and advanced forecast models to improve their eCommerce demand forecast, and ultimately their bottom line.

Further supporting the benefits of tapping into underutilized internal data

Our proof of concept, as well as recent academic research and studies from MIT and Hindawi , highlight promising results when using key indicators to predict demand patterns and eCommerce transaction trends:

- Subscriptions
- Censored demand
- Website analytics
- Express delivery

Enhance forecasting methods

Alternative forecasting approaches such as causal modelling and machine learning technologies should be leveraged to drive material improvement in near-term forecast accuracy.

Overall, consumer business companies with eCommerce channels are recommended to explore alternative approaches to improve eCommerce demand forecast leveraging new data objects and advanced forecasting models.

How can we help?

Deloitte is the world's leading professional services firm. Deloitte's supply chain consulting practice has delivered supply planning strategies, process improvements, solution implementations and integrations. Logility is a leading supply chain solution provider that has helped many companies across a variety of industries to improve their integrated business planning, through leveraging AI, machine learning and automation.

Deloitte is a leader in the supply chain planning process and is focused on providing end-to-end thinking, building capability for its clients, and identifying opportunities. Our professionals can leverage the company's existing data assets to perform a series of analyses, assess current capabilities and provide solutions to improve on eCommerce forecasting. Leveraging 'Leading Indicators' unique to the eCommerce channel combined with enhanced forecasting methods may lead to material improvements of forecast accuracy that can be leveraged to increase sales while optimizing distribution and inventory levels.

Logility provides alternative forecasting approaches, such as machine learning technologies, to drive improvements in near-term forecast accuracy. With an increased layer of complexities to eCommerce, all relevant demand signals are key to driving success in forecast accuracy.

In the proof-of-concept analysis, Deloitte and Logility together have shown that forecast accuracy can be improved by leveraging additional social analytics and demand signals in the causal forecast models. Beyond the ability to build and run various sophisticated demand forecast models, Logility also provides a user-centric interface to display KPIs in its solution.

Together we can help you explore the potential in improving your omnichannel demand forecast.

Get in Touch

Contact us with any questions related to the proof-of-concept causal forecasting exercise and how the Deloitte Supply Chain team is helping organizations to identify capabilities in today's ever-changing consumer business landscape.

Endnotes

1. Gartner. Supply Chain Brief: COVID-19 Impact on eCommerce and Sustainable Packaging (2020)
2. Koetsier. COVID-19 Accelerated E-Commerce Growth '4 To 6 Years' (2020)
3. ARIMA = autoregressive integrated moving average
4. ARIMAX = ARIMA with exogenous variables
5. The Pearson Correlation Coefficient reviews all variables analyzed and measured on a scale between -1 to +1, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation.
6. MIT Academic Paper. Improving Online Demand Forecast using Novel Features in Website Data: A Proof-of-Concept at Zara (2018)
7. Research Article, Forecast of E-Commerce Transactions Trend Using Integration of Enhanced Whale Optimization Algorithm and Support Vector Machine (2021)

Laurie Wang
Partner & Canadian Supply Chain Planning Practice Leader
lauwang@deloitte.ca

Sarah Bazmi
Deloitte Supply Chain Senior Manager
sbazmi@deloitte.ca

Simon Booth
Deloitte Supply Chain Manager
simbooth@deloitte.ca

Hang Nguyen
Deloitte Supply Chain Manager
hnguyen@deloitte.ca

Partner Contact:
Accelerating the sustainable digital supply chain, Logility helps companies seize new opportunities, sense and respond to changing market dynamics and more profitably manage their complex global businesses.

The Logility® Digital Supply Chain Platform leverages an innovative blend of artificial intelligence [AI] and advanced analytics to automate planning, accelerate cycle times, increase precision, improve operating performance, break down business silos and deliver greater visibility. Logility is a wholly owned subsidiary of American Software, Inc. [NASDAQ: AMSWA].

To learn how Logility can help you make smarter decisions faster, visit www.logility.com.

Worldwide Headquarters 800.762.5207 / United Kingdom +44 (0) 121 629 7866

Other Contributors
Cindy Yang, Tatiana Aoun, Bonnie Tsang, Vinder Sodhi, Alison Weyland, Duong Thai, Don Lange