Price Optimization for Revenue Maximization at Scale

Nikhil Gupta, Massimiliano Moro, Kay Ayala, Dr. Bivin Sadler

Master of Science in Data Science, Southern Methodist University,  
Dallas, TX 75275 USA

{guptan, mmoro, kayala, bsadler}@mail.smu.edu

**Abstract.** Price optimization is a tool companies use to adjust the price of their products to best achieve their business objectives (such as maximizing revenue or profits). When price is high demand drops and when price is low demand increases. The business must find the right price for their products in order to maximize revenue. In reality, this concept becomes much more complex when we take into account factors such as cross cannibalization and the halo effects. This paper presents a novel approach on price optimization by using data analysis techniques and machine learning models to estimate the future demand of products after taking cross effects such as cannibalization, and the halo effect into consideration. The paper study data collected from 2013 to 2019 for Premium Whiskey’s and Economy Vodka’s from a large distributor of alcohol beverages. For each product and class of customer, this paper shows the expected demand and the ideal price strategy to maximize the revenue for the business.

# Introduction

Any distributor or retail business must periodically refine its price strategy in order to maximize business objectives such as profit and/or revenue. A small change in price may alter the demand with positive or negative consequences to the bottom-line profit of the business. Moreover, the price change in one product can have an impact on not only that product, but also on other products (A negative effect on other products is called cannibalization and a positive effect on other products is called the halo effect). These effects are hard to predict since there are many cross product relationships when hundreds or thousands of products are involved.

Recent innovations in machine learning provide vast computing power and flexibility in efficiently addressing predictive and optimization problems. Neural Network and Multivariate time series models can address a complex set of interdependent and sequential variables to unveil their patterns and provide a more accurate prediction of demand. Optimization methods can be used to navigate different scenarios and find the ideal price strategy. This paper applies these techniques to the complex interaction between demand and price for alcoholic beverage industry. While the techniques applied in this paper are on a specific industry, the concepts are generic enough to be applied to any industry.

This paper provides an answer to one of the most important question of the distributor/retail industry: what is the ideal pricing strategy for products across the board in order to drive up demand and maximize revenue?

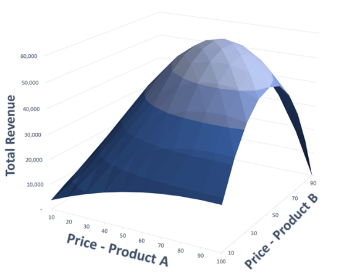
# Literature Review

## Price Elasticity

The concept of price elasticity has been studied in economics for several years using traditional methods [1]. Price elasticity is a metric quantifying the degree to which the decrease in the price of a product increases its demand. When price is first lowered, the increase in demand is enough to offset the decrease in price and this leads to a higher revenue. However, after a certain point, the increase in demand is not enough to offset the decrease in price and the revenue starts to fall. This phenomenon is indicative of the presence of an ideal price point to maximize revenue [2].



However, price elasticity tends to be complicated and the decrease in price of one product can not only increase the demand for that product but a decrease the demand for similar products as well (if customers started switching to the product with the discounted price) [3].



Moreover, this behavior is not only product dependent, but also customer dependent. In the “Product Cannibalization and the role of prices” paper, the authors talk about the idea of “core customers” (who are still likely to purchase products when price is increased) vs. fringe customers (who are likely to stop buying if price increases too much) [4]. The author also talks about the concept of asymmetric price elasticities, i.e. product A may be able to influence the demand for product B more so than product B can influence the price of product A. This can lead to more complexity in determining cross effects.

In addition, traditional techniques for modeling cross effects have relied on basic methods like OLS, but this may not be sufficient to capture the non-linear nature of the problem space. [5]

## Models used in literature

Previous studies have looked at modeling demand for the alcoholic distribution market and both concluded that the demand needs to be modeled at a per customer and a per product basis because of the different buying patterns, product preferences and tolerance to price changes. Moreover, they also found that there is “no one sized fit all” model. The type of model that worked for one customer-product combination did not work well for another. While both these papers focused on a handful of customers (less than 5) and products (less than 5), they indicated that the ability to scale this would need an automated machine learning system (AutoML) to develop these models at scale. [6] [7]

## AutoML for Time Series

Due to recent advances in computing, and data science, as well as the democratization of machine learning, several AutoML systems have been introduced in the marketplace. One such framework is auto-sklearn [8] which is built on the popular scikit-learn framework in python. Demand models are inherently time series based as the demand is a function of time (in addition to other exogenous variables). However, the scikit-learn framework tends to focus more on non-time series models. Although it has support for time series based sliding window resampling methods as well as imputing and feature engineering methods, the models themselves are not leveraging any correlation between the time series data as a traditional autocorrelation based model (such as Vector Auto Regressive model) would. In addition, the classical models such as ARIMA are missing from the learn framework. [9]

Similarly, a recent entry into the world of AutoML is PyCaret, but it also does not have support for Time Series models [10]. Other systems such as the h2o.ai Driverless AI do support time series forecasting, but these are paid products costing thousands of dollars [11].

Hence, in order to truly scale this across several customers and products, we need to develop an AutoML framework ourselves that can support a variety of machine learning models and methods such as Seasonal ARIMA, Vector Auto Regressive models, XGBoost, and stacking. Fortunately, libraries such as statsmodels do have extensive support for Time Series Analysis [12]. In addition, scikit-learn provides the option to add custom models in the framework[13]. Based on our research, we believe that using these available libraries and frameworks, a free and scalable time series AutoML framework can be developed to perform demand forecasting and revenue optimization at scale.

## Optimization

The idea of optimization is not new; many books have been written on thetopic [14], [15], [16]. The challenging however is as the design space becomes more non-linear and high dimensional, optimizers can struggle to reach the global minima [17]. More recently, Bayesian optimization has emerged as a promising technique that allows for a more efficient search of the sample space leading to more optimal results. [18]

# Methods

We plan to explore the following methods for various sections of the analysis

* **Identifying Cross Effects**: Graph Theory and Association Analysis to identify most related products
* **Modeling Demand**: Seasonal ARIMA with exogenous variables, Vector Autoregressive Models, Neural Networks such as LSTMs, Prophet Library from Facebook, others.
* **AutoML**: scikit-learn, statsmodels
* **Optimization**: Bayesian Optimization

# Results

* 1. Cross-cannibalization and Halo Effects
     1. These items [\_\_\_\_\_] had the highest influence on each other
  2. Demand Forecasting
     1. ASE
     2. Model Fit Metrics
     3. Confidence
  3. Price Elasticity
     1. Metrics, Charts to show the Price Elasticity
     2. Variable Importance?
  4. Optimal Revenue Estimation
     1. Supply Quantity
     2. Associated Demand
     3. Price points
     4. Anonymized projected revenue per product and overall

# Discussion

* + Now that we have projected demand, estimated optimal price, and calculated projected revenue.
  + Our projection is best during [\_\_\_\_] time period due to our forecast horizon
  + We recommend using this methodology to calculate price points [\_\_\_] time in advance (days, weeks, or months). Based on the forecast horizon and the variability of global politics and other phenomena
  + While we see [\_\_\_\_] interaction patterns in our cross-cannibalization and halo effects analysis, ultimately the true interaction pattern is best captured with experimentation.
  + Since this analysis is not an experiment but rather an exploration of data we cannot make any population or causal claims. However, we were able to elucidate patterns in the data.

# Conclusion

* We believe this methodology can increase profit for products other than just alcohol
  + The economic size of the retail sector is large making this kind of analysis have broad implications
* Future work
  + New products
  + Navigating the pandemic ‘cold start’ problem
  + Including spatial analysis
  + Including a broader feature set outside of just the company’s data to include
    - Economic indicators
    - Social network data
    - Marketing effects
    - Weather and other global effects
  + Building in a control system style experiment for capturing cannabalization and halo effects.
  + Consulting literature on price point psychology

# References

1. OpenStax. (2016, May 18). Principles of Economics. Retrieved from BC Open Textbooks: https://opentextbc.ca/principlesofeconomics/

2. Bromley, R. (2006). Elasticity. Retrieved from pcecon.com class notes: https://www.pcecon.com/notes/elasticity.html

3. McColl, R., Macgilchrist, R., & Rafiq, S. (2020). Estimating cannibalizing effects of sales promotions: The impact of price cuts and store type. Journal of Retailing and Consumer Services. Retrieved from https://doi.org/10.1016/j.jretconser.2019.101982

4. Meredith, L., & Maki, D. (2001). Product cannibalization and the role of prices. Applied Economics, 1785-1793.

5. Pickett, J. C., & Reilly, D. P. (2020, May). Testing Marketing Hypothesis. Retrieved from Automatic Forecasting Systems: https://www.autobox.com/pdfs/cannibalize.pdf

6. Jiang, L., Rollins, K. M., Ludlow, M., & Sadler, B. (2020). Demand Forecasting for Alcoholic Beverage Distribution. SMJ Data Science Review. Retrieved from https://scholar.smu.edu/cgi/viewcontent.cgi?article=1131&context=datasciencereview

7. Arora, T., Chandna, R., Conant, S., Sadler, B., & Slater, R. (2020). Demand Forecasting In Wholesale Alcohol Distribution: An Ensemble Approach. SMU Data Science Review. Retrieved from https://scholar.smu.edu/cgi/viewcontent.cgi?article=1137&context=datasciencereview

8. Machine Learning Professorship Freiburg. (2019). auto-sklearn. Retrieved from auto-sklearn: https://automl.github.io/auto-sklearn/master/index.html#

9. Rosenthal, E. (2018, MArch 22). Time Series for scikit-learn People (Part II): Autoregressive Forecasting Pipelines. Retrieved from Ethan Rosenthal : <https://www.ethanrosenthal.com/2018/03/22/time-series-for-scikit-learn-people-part2/>

10. Moez, A. (2020, 05). PyCaret. Retrieved from PyCaret: https://pycaret.org/

11. H2O.ai. (2018, June 12). H2O.ai. Retrieved from Time is Money! Automate Your Time-Series Forecasts with Driverless AI: https://www.h2o.ai/blog/time-money-automate-time-series-forecasts-driverless-ai/

12. Perktold, J., Seabold, S., & Taylor, J. (2020, Feb 21). Time Series Analysis. Retrieved from statsmodels v.0.11.1: https://www.statsmodels.org/stable/tsa.html

13. scikit-learn developers. (2019). Developing scikit-learn estimators. Retrieved from Scikit Learn: https://scikit-learn.org/stable/developers/develop.html

14. Griva, I., Nash, S. G., & Sofer, A. (2009). Linear and Nonlinear Optimization. Philadelphia: siam. Retrieved from https://pdfs.semanticscholar.org/5f8d/2a70f62baf4f4971467ae03fad2db61dfe48.pdf

15. Floudas, C. A. (1995). Nonlinear and Mixed-Integer Optimization. New York: Oxford University Press. Retrieved from https://pdfs.semanticscholar.org/6687/8b5770d65e781a0b9f8c95867f11336008c8.pdf

16. Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., & Mahajan, A. (2013). Mixed-integer nonlinear optimization. Acta Numerica, 1-131. Retrieved from https://search-proquest-com.proxy.libraries.smu.edu/docview/1322231410?accountid=6667

17. Dauphin, Y. N., Pascanu, R., Gulcehre, C., Cho, K., Ganguli, S., & Bengio, Y. (2014). Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. Cornell University Library. Retrieved from https://search-proquest-com.proxy.libraries.smu.edu/docview/2084532558?accountid=6667&rfr\_id=info%3Axri%2Fsid%3Aprimo

18. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & De Freitas, N. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. IEEE, 148-175. Retrieved from https://ieeexplore-ieee-org.proxy.libraries.smu.edu/stamp/stamp.jsp?tp=&arnumber=7352306&tag=1

19. Lomax, W., Hammond, K., East, R., & Clemente, M. (1997). The measurement of Cannibalization. Journal of Product & Brand Management, 27-39. Retrieved from http://mktg.uni-svishtov.bg/ivm/resources/prod\_canibalism.pdf

20. Ito, S., & Fujimaki, R. (2016). Optimization Beyond Prediction: Prescriptive Price Optimization. Cornell University. Retrieved from https://arxiv.org/pdf/1605.05422.pdf