ANALYTICS & ADVERTISING: QUANTIFYING MARKETING EFFECTS ON SALES

GROUP 6

10/30/2019
BAX 401 Intro to Business Analytics

Zunfeng Huang
Mitesh Jain
John Elmer Loretizo
Mayank Mani
Shalini Mishra
Prakriti Rastogi

Table of Contents

Executive Summary

Advertising has always been about connecting consumers to products. It's centered on the idea of making them aware and enticing them to eventually buy various products. This report quantifies the effects of advertising mediums in the sales of a cosmetic firm. The mediums are generally divided into two major groups – online and offline platforms with the model considering 3.5 years-worth of monthly data. Initial analysis revealed that only 4 advertising mediums generated quality observations that can be included in the model. From these, the following results have been discovered:

- *Limited available data.* Out of the 9 possible variables to be considered in the model, only 4 variables showed enough observations to be considered in the model.
- *Nonlinear effect of advertising on sales.* Models generated from selected variables failed to explain at least 50% of the variation in sales.
- *Two significant advertising mediums.* Both online search advertising as well as catalogs sent to existing customers exhibited the highest effect in terms of change in sales.

In addition, diminishing effects of these advertising mediums were also assessed. A final recommended model was developed with the following key points:

- *Reallocation of budget.* Budget for search ads resulted in negative effects to the sales and should be reduced. In addition, budget increase for catalogs sent to existing customers must be considered.
- *Ensure data quality.* Additional measures to ensure clean and accurate data must be made in order to build a more robust model.
- *Include carryover effects of advertising media.* This measure will help gauge how the effect of advertising media from previous period affect current sales.

Introduction

Marketing has always been about connecting to customers at the right time in the right place. Marketing mix refers to a set of tactics/strategies that a company uses for a successful product offering. When properly understood and applied, it's a key factor driving a product's success. A cosmetic firm launched a product 4 years ago and has been utilizing multiple media for advertising. For Offline Advertising, catalogs and mailings were used. Mailings included flyers, postcards, and letters. For effective targeting, different catalogs were sent to different customer segments. These segments include existing customers, win-back customers; those who had not made any purchase in the previous 6 months and potential new customers. Their online marketing activities included Banner ads, Search ads, Social Media ads and Newsletter ads .The firm wants to measure the efficacy of advertising activities on product sales.

Problem Formulation

Multi-channel marketing allows companies to expand their product's reach. We are provided with 42 months of sales data for the product in terms of units sold. The dataset contains the following expenses information (in $):

1. Monthly spend on mailings and catalogs targeted for new customers, existing customers and win-back customers separately. Also, spends on digital channels like banners, search, social media, portals and newsletter
2. Expenditure in advertising strategies like retargeting which allows one to focus on clients which already has shown some degree of interest in their product

We are trying to understand the impact of advertising expenses on product sales to gauge the effectiveness of a channel. Our model in focus will have a structure like the skeletal equation explained : $Y_t \sim f(X_{1t}, X_{2t}, \ldots, X_{nt}, Z_{1t}, Z_{2t}, \ldots, Z_{nt}, Y_{t-1})$, for n media types

We have considered the following three data components:

1. Response Data is Sales in month t denoted by Y_t
2. Advertising spends for each channel in month t are denoted by X_{it}, where i denotes the i^{th} media type
3. Metrics to account for lag effect and diminishing returns in advertising

   Lag effect of Advertising: Lag of sales for each month t represented by Y_{t-1}

   Diminishing returns: Every dollar spent on advertising does not generate an equal return.

   Diminishing returns in month t for i^{th} media type is represented by Z_{it}

$$Z_{it} = \sqrt{X_{it}} \text{ or } Z_{it} = \ln(X_{it})$$

Data Description

We have to carefully select variables from our dataset to avoid a 'Garbage In Garbage Out' model. We removed certain variables from our dataset based on the following criteria:

- Variables like Catalogs_WinBack, Catalogs_NewCust, Mailings, Banner, Social Media & Retargeting can't be relied on for modeling as they have >40% data missing
- Overall spend metrics like ADV_Total, Adv_offline and Adv_online were removed as they are a linear combination of other variables included in the data. Incorporating them into the model would result in multicollinearity.

Thus, we are working with remaining 4 variables: Catalog_ExistCust, Search, Portals, Newsletter and Lag_Sales: this variable considers the effect of sales with a lag of 1 month.

Descriptive statistics of the dependent as well as the independent variables:

	Sales	*CatExistCust*	*Search*	*Newsletter*	*Portals*	*Sales_lag*
Min	3355	0.0	38.17	7.057	2.544	3355
1st Qrt	4406	328.7	45.38	16.691	3.393	4169

Median	4690	598.0	66.11	19.779	4.707	4631
Mean	4809	567.6	69.83	20.734	5.246	4716
3^{rd} Qrt	5195	625.6	88.19	25.139	6.867	5165
Max.	6976	1298.7	134.87	53.609	9.303	6467

Model Development

Features selected and definition:

1. Sales with 1-month lag – t-1
2. Diminishing effect of amount spent on Catalog for Existing Customers – Catalogs_ExistCust_dim, calculated by taking the square root of data within the Catalog for Existing Customer column
3. Diminishing effect of amount spent on Newsletter – Newsletter_dim, calculated similarly as above
4. Diminishing effect of amount spent on Search – Search_dim, calculated similarly as above
5. Diminishing effect of amount spent on Portals – Portals_dim, calculated similarly as above
6. Intercept
7. Error term

$$Y_t = \lambda Y_{t-1} + \beta_0 + \beta_1\left(\sqrt{\text{Portals}}\right) + \beta_2\left(\sqrt{\text{Newsletter}}\right) + \beta_3\left(\sqrt{\text{Catalog for existing csutomers}}\right) + \beta_4\left(\sqrt{\text{Search}}\right) + \varepsilon$$

We believe that advertising should have positive effects in either boosting or retaining the sales, at least normally, will not hamper the sales.

Model Effectiveness: By using R’s linear regression model simulation, we receive the general effectiveness (Adjusted R-squared) of the model, which is 0.20. The score is substantially low and suggests that further improvements are required.

Summary of Focal Model:

Coefficients:	Estimate	t value	Pr(>|t|)
(Intercept)	3108	3.052	0.00432
lag_sales	0.003037	0.018	0.98609
Portals_dim	1167	1.981	0.05554
Newsletter_dim	119.5	0.889	0.37983
Catalogs_ExistCust_dim	-27.85	-1.654	0.10699
Search_dim	-99.88	-0.607	0.54758

Variables Significance: According to above summary table, only intercept is statistically significant.

Results

A more robust model was built to consider effects of advertising medium on sales. In total, there are 5 variables in the final model. These are the (1) lag of sales, the carryover effect of previous month's sales, (2) search advertising medium, (3) shopping catalog for existing customers. The last two variables are derived from variables (2) and (3), which were transformed by taking the square root of their values. The transformation was done because advertising exhibits a non-constant and decreasing marginal return on sales[1].

The relationship can be defined by the equation below:

$$Y_t = \lambda Y_{t-1} + \beta_0 + \beta_1(\text{Catalog for existing customers}) + \beta_2(\text{Search}) + \beta_3\left(\sqrt{\text{Catalog for existing csutomers}}\right) + \beta_4\left(\sqrt{\text{Search}}\right) + \varepsilon$$

Summarized in the table below are estimates of the coefficient of each variable as well as their elasticity. A positive value for coefficient estimates and elasticity[2] denote a positive change for sales and vice versa. In addition, one can see a very high intercept which meant that a number of other variables/factors that might explain sales are not taken into account by the given model.

[1] The law of diminishing returns states that the relationship between marginal input and marginal output will be more than proportional at the initial stage, but after a certain point, with all other inputs held constant, the marginal product of each unit of input will drop as the amount of the input increases (Heibroner and Thurow 1981; Picard 1989; Samuelson and Nordhaus 1989).

[2]

The negative value for sales shows an indication of overbudget for that medium that negatively affects sales. However, one can also notice that catalogs sent to existing customers provide a positive effect to sales.

Advertising Medium:	Coefficient Estimate	Elasticity
(Intercept)	1522.0062	
lag_sales	-0.2124	
Search	-58.1452	-0.009972182808
Catalogs_ExistCust	4.7061	0.000807118894
Search_dim	1223.6891	0.2098685946
Catalogs_ExistCust_dim	-193.4651	-0.03318019964

Recommendation

- *Advertising Budget Reallocation:* Increase in budget on catalogs for existing customers can yield more positive effect on the product sales
- *Enriching the dataset*: R-squared ~ 50% of our model suggests that our existing data wasn't enough to explain the effect of advertising on sales properly. Enriching the current dataset will enable us to get a more robust model. Our given dataset can be enriched by adding important marketing metrics like price, promotion, product distribution etc. which also impacts sales of a product.
- Funnel effects of an advertising channel must be considered to avoid misguided attribution of sales to a specific ad channel.

Conclusion

Implementing infrastructure that captures advertising data from different platforms is only the first step towards developing Advertising Analytics 2.0. Knowing what data to capture and how to use it to the firm’s advantage is a completely different feat at which many have failed. Sales of a company doesn’t depend only on their advertising channels, they also hinge upon the company’s goodwill in the market and their price points, none of which was captured in this case.

References

Chan, David and Perry, Michael(2017) , Challenges And Opportunities In Media Mix Modeling, pp. 6 -18. Retrieved from https://ai.google/research/pubs/pub45998

Shephard R.W., Färe R. (1974) The Law of Diminishing Returns. In: Eichhorn W., Henn R., Opitz O., Shephard R.W. (eds) Production Theory. Lecture Notes in Economics and Mathematical Systems (Mathematical Economics), vol 99. Springer, Berlin, Heidelberg Retrieved from https://doi.org/10.1007/978-3-642-80864-7_17

Valos, M. J Database Mark Cust Strategy Manag (2008) 15: 239. Retrieved from https://doi.org/10.1057/dbm.2008.22

Gert Assmus, John U. Farley, Donald R. Lehmann First Published February 1, 1984 Research Article, How Advertising Affects Sales: Meta-Analysis of Econometric Results, pp. 65-74. Retrieved from https://doi.org/10.1177/002224378402100107