Homework 2

Information, Impacts and Insights

E-FRI

Asha Mairh

Jeet Patel

Kavish Hukmani

Keshore Suryanarayanan

Neon Zhang

10-23-2021

Contents

1 Executive Summary

Marketing Mix Modelling has had a massive increase in adoption over the past two decades. The offering of improving sales by optimally reallocating the advertising spend is something that nobody can pass on. This report determines the effectiveness of advertising activities on the sales of a product for a cosmetics firm.

The firm uses a variety of marketing channels, both online and offline to target various types of customers. These can be narrowed down to three focal channels - catalogs for win back customers, newsletters, and portal ads. To capture the diminishing returns of advertising spend, a square root transformation is applied and to account for carryover effect of advertising, a lag in sales is incorporated.

We use a focal model that determines Sales using lagged Sales and respective advertising expenditures on Catalogs for Winback Customers, Portals and Newsletters. The model is statistically significant and results in positive linear relationships between the expenditures and sales, but fails to explain majority of the variation in Sales.

It is concluded that display advertising and newsletters are better in generating sales. We recommend to the business to account for advertising synergy and seasonality in future models to better determine sales and thereby better allocate advertising budget across media.

2 Introduction

The ability to track every action a person makes online and using it to draw inferences about sales, combined with existing traditional advertising techniques, has led to the formulation of "Advertising Analytics 2.0" [Nichols, 2013]. It has a heavy focus on the interdependency of the various advertising activities and understands how multiple activities can collectively contribute towards making a sale, trying to use this information to improve the advertising strategy. It does so in three steps - Attribution, Optimization and Allocation.

This report explores the effectiveness of a cosmetic firm's advertising activities on sales of a product it launched about 4 years ago. It also analyzes potential reasons behind the result and provides recommendations that help better determine advertising effectiveness and thereby the advertising budget allocation for both this product and similar products in the future. [Albers et al., 2010]

3 Problem Formulation

The cosmetics firm has employed both offline and online media channels. Catalogs and Mailings comprise the offline media, with expenditure on catalogs bifurcated by the loyalty of the recipient customers. The online media channels employed are banner, search, social media, newsletter, retargeting and portals. Banners, newsletters and retargeting require the customers to have entered the website, whereas search advertising, social media advertising and portals could be triggered by searches either on the product type or on the brand performed by internet users. [Dinner et al., 2014]

There is an implicit hierarchical structure in the dataset. *ADV_Total* is broken down by *ADV_Offline* and *ADV_Online*. *ADV_Offline* is further broken down by *Catalog_ExistCust*, *Catalog_Winback*, *Catalog_NewCust* and *Mailings*. *ADV_Online* is further broken down by *Banner*, *Search*, *SocialMedia*, *Newsletter*, *Retargeting* and *Portals*.

Advertising expenditure is known to have a carryover effect on sales which is represented in the model by including the lagged version of sales. [Breuer and Brettel, 2012] Additionally, advertising expenditure has diminishing returns which is represented by the square root operation on advertising variables. Using this framework, the focal model is devised and the intuition behind selection of variables is explained in 'Model Development'. The appendix explores enhancements to the focal model.

4 Data Description

The data considered is the spend across the offline and online advertising channels, and unit sales over 42 months. The bifurcation of the channels was discussed in the previous section. We have excluded the variables SocialMedia and Banner from the given data due to lack of sufficient data. The firm never indulged in social media and the spend on banner has been infrequent to be judged on effectiveness. Additionally, we observe that Search, Portals, and Retargeting are highly correlated. To remedy the same, we dropped Search (as Portals is a closer touchpoint to Sales) and Retargeting (as the firm invested in retargeting only later in the dataset). Apart from this, we have kept all other variables under consideration. We believe that sending out catalogs to winback customers is powerful in reinforcing brand awareness, [Griffin and Lowenstein, 2002] and that display advertising and newsletters explain sales as they are potential last touch channels prior to a sale. Hence, we have limited our analysis to Catalog_Winback, Portals and

Newsletter. On average, the firm has spent $20.93 on Newsletters, $5.25 on Portals, and $211.01 on catalogs for winback customers periodically.

Variable	**Description**
Months	Time in Months
Sales (units)	Sales of items in units in the month
Catalogs_Winback	Amount spent on Shopping Catalogs sent to Customers (who have not bought for at least 6 months) in the month
Newsletter	Amount spent on Newsletter ads in the month
Portals	Amount spent on ad portal advertising in the month

Table 1: Variable descriptions

5 Model Development

The focal model represents the carryover effect of advertising by regressing Sales in the previous time period with that in the given time period. It focuses on the individual effects of the media and does not incorporate advertising synergy. Offline media is more intricate in its reach. Catalogs and mailings are likely to reach customers of many demographics irrespective of their internet usage. [Lesscher et al., 2021] So we believe that they are a more powerful means to generate brand awareness and product curiosity among customers, as they cast a wider net. In the data, we see that Catalog_Winback and Catalog_NewCust primarily focus on brand awareness, but we only include Catalog_Winback because the customer composition of 'New Customers' is unclear.

Online media, on the other hand, drive customers to the consideration stage of the purchase funnel. [Bayer et al., 2020] Higher investment in Search Advertising demands higher investment in Display Advertising to push customers across the funnel, explaining the high correlation between Search and Portals. Newsletter is also powerful in the consideration stage because it advertises the product specifically to the customers, and any click-through leads to the customers landing on the purchase page. Thus, we have employed the following focal model:

$$Sales_t = \lambda Sales_{t-1} + \beta_0 + \beta_1\sqrt{Catalog\_Winback} + \beta_2\sqrt{Portals} + \beta_3\sqrt{Newsletter} + \epsilon$$

6 Results

As a result, the focal equation is expressed as follows:

$$Sales_t = 0.202Sales_{t-1} + 1837.096 + 19.713\sqrt{Catalog\_Winback} + 727.554\sqrt{Portals} + 63.032\sqrt{Newsletter} + \epsilon$$

R-squared	Adj. R-squared	F-statistic	Prob (F-statistic)	AIC
0.271	0.19	3.346	0.0199	660.1

Table 2: Regression output of focal model

According to the table showing above, R^2 is 0.271, P-value is 0.0199, which means that under 95% confidence, Catalog_Winback, Newsletter, Portals, and Sales in the last month can explain 27.1% change of sales in the last 42 months with statistical significance.

Variable	coef	std err	t	P-value
const	1837.096	972.624	1.889	0.067*
Catalogs_Winback	19.713	16.345	1.206	0.236
Newsletter	63.032	127.759	0.493	0.625
Portals	727.554	302.440	2.406	0.021**
Sales_Lag	0.202	0.163	1.234	0.225

Table 3: Regression summary of focal model

Catalog_Winback, Newsletter, and Portals in the last month all have positive linear relationships with Sales, while only Portals has a statistically significant positive and the strongest impact on the sales at 95% confidence level. Specifically, (1) increasing one unit of advertisement spending on Catalog_Winback can lift 19.713 units sales; (2) improving the advertising investment on Newsletter by one unit would enlarge the Sales by 63.032; (3) 727.544 units of Sales can be achieved when the company put money into the Portal channel for one unit; (4) the Sales of the past month will promote the sales of the new month by 0.202 units, though the effect is not significant. Similarly, the advertisement elasticity also proves the effectiveness of investing money on the Portals, with 10% increase spending on the Portals, the sales would increase by 1.8%.

Variable	Elasticity
Catalogs_Winback	0.005098
Newsletter	0.016301
Portals	0.188156

Table 4: Variable elasticities of focal modal

Consequently, Portals advertisement is the channel with the highest return on advertising investment. In the future, the company should allocate more budget on display advertising.

7 Recommendations and Managerial Implications

Online media are more effective in generating sales than offline media, however, when used without synergy, they fail to explain the variation in sales significantly. We see that accounting for carryover effect and diminishing returns boosts the determination of sales, and thereby judging advertising media effectiveness.

The fit and explanation of variation of the existing model present an opportunity for improvement. We explored the robustness of logarithmic transformations over square root to represent diminishing returns, and obtained better results. Furthermore, we investigated the presence of synergy by incorporating offline-online and online-online synergy elements in the model and obtained drastically better results in terms of both the validity of the overall model and explanation of variation. This shows that synergy does exist in advertising and needs to be incorporated for both description and prediction of sales. Seasonality is another factor that needs to be considered especially in a time series data, which has not been explored in the paper but presents opportunities for future research.

8 Conclusions

For the given model specifications, we have devised an intuitive model that fares relatively better in explaining Sales with the given data. The explanatory variables were chosen based on an assumption that offline media primarily drive awareness, on collinearity checks as well as data availability.

The model is limited in its ability to determine variation in Sales, implying the existence of other variables that drive it. The additional variables could be the offline-online and online-online advertising synergy, which form the basis of our extended analysis. It also fails to have statistical significance in one of its coefficients, but was still used because of its overall validity. It does not account for seasonality, which could comprise any future analysis.

References

[Albers et al., 2010] Albers, S., Mantrala, M. K., and Sridhar, S. (2010). Personal selling elasticities: A meta-analysis. *Journal of Marketing Research*, 47(5):840–853.

[Bayer et al., 2020] Bayer, E., Srinivasan, S., Riedl, E. J., and Skiera, B. (2020). The impact of online display advertising and paid search advertising relative to offline advertising on firm performance and firm value. *International Journal of Research in Marketing*, 37(4):789–804.

[Breuer and Brettel, 2012] Breuer, R. and Brettel, M. (2012). Short- and long-term effects of online advertising: Differences between new and existing customers. 26(3):155–166.

[Dinner et al., 2014] Dinner, I. M., Van, H. J. H., and Neslin, S. A. (2014). Driving online and offline sales: The cross-channel effects of traditional, online display, and paid search advertising. *Journal of Marketing Research*, 51(5):527–545.

[Griffin and Lowenstein, 2002] Griffin, J. and Lowenstein, M. W. (2002). *Customer winback: How to recapture lost customers–And keep them loyal*. John Wiley & Sons.

[Lesscher et al., 2021] Lesscher, L., Lobschat, L., and Verhoef, P. C. (2021). Do offline and online go hand in hand? cross-channel and synergy effects of direct mailing and display advertising. *International Journal of Research in Marketing*, 38(3):678–697.

[Nichols, 2013] Nichols, W. (2013). Advertising analytics 2.0. *Harvard business review*, 91(3):60–68.

A Data Description

Variable	Description
Months	Time in Months
Sales (units)	Sales of items in units in the month
ADV_Total	Total Advertising Spend in the month, comprises ADV_Offline and ADV_Online
ADV_Offline	Total Offline Advertising Spend, comprises Catalogs_ExistCust, Catalogs_Winback, Catalogs_NewCust in the month
Catalogs_ExistCust	Amount spent on Shopping Catalogs sent to existing Customers in the month
Catalogs_Winback	Amount spent on Shopping Catalogs sent to Customers (who have not bought for at least 6 months) in the month
Catalogs_NewCust	Amount spent on Shopping Catalogs sent to New Customers in the month
Mailings	Amount spent on Mailings (excluding Catalogs) sent to Customers. Mailing include flyers, postcards and letters in the month
ADV_online	Total Online Advertising Spend, comprises Banner, Search, SocialMedia, Newsletter, Retargeting and Portals in the month
Banner	Amount spent on Banner ads in the month
Search	Amount spent on Search ads in the month
SocialMedia	Amount spent on Social Media ads in the month
Newsletter	Amount spent on Newsletter ads in the month
Retargeting	Amount spent on Retargeting ads in the month
Portals	Amount spent on ad portal advertising in the month

Table 5: Variable descriptions

A correlation matrix is created and represented in the form of a heatmap in Figure 1, to study the correlations among the advertising expenditures and study multicollinearity. We observe high correlations especially among the online advertising media, especially between Search advertising and Display advertising. This is intuitive because as a company spends more to get customers in the awareness stage of the purchase funnel, they need to proportionately spend on display advertising to transition them into the consideration stage.

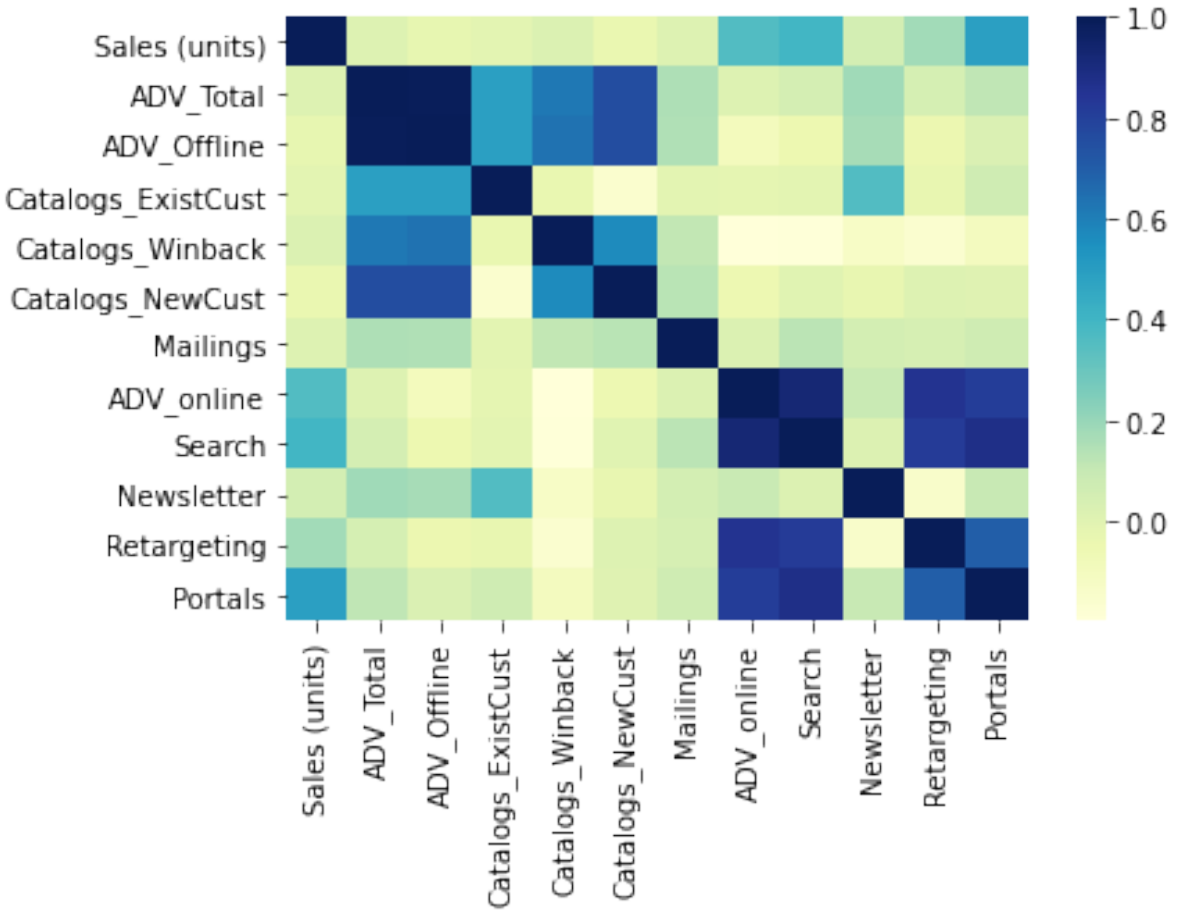

Figure 1: Variable correlations

B Model Experiments

B.1 Experimenting without intercept

We interpret the intercept as the steady state condition in which we turn off the advertisement across all channels for a long time and then observe the unit sales. We experimented running the focal model without intercept and observed that the overall model becomes extremely statistically significant with 0.977 $adjusted - R^2$. However, within the context of this problem, even under steady state conditions, we will still observe some unit sales because advertising is not the sole driver of sales. Hence, it is likely that removing the intercept would end up in gross over fitting of the model.

Experiment	**AIC**	**P Value of Model**	**Adjusted R Squared**
No Intercept	663.9669	<2.2e-16	0.9772

Table 6: Focal Model without intercept

B.2 Experimenting with different Functional Form

It is widely accepted that advertising suffers from the law of diminishing returns. To capture this effect, we used the square root function for the focal model. We explore another functional form using the natural log of the advertisements and check if the observations are robust to the choice of the functional form. We observe that using natural log improves the explain-ability of the model.

Experiment	AIC	P Value of Model	Adjusted R Squared
Log	661.3732	0.01505	0.2041

Table 7: Focal Model with log functional form

B.3 Experimenting with online-online synergy

Newsletter - Search: We posit that sending newsletters to customers could prompt search behavior impacting sales.

Search - Portal: Searches drive display advertising i.e. advertising through portals. They may not directly impact sales strongly, but when synergized with portals, they can potentially influence customers both in awareness and consideration stages.

Portal - Retargeting: Now, searches and portals could still have customer churn wherein customers could enter the website but end up not making a purchase. This is where synergy in searches and portals with retargeting could come into play and result in Sales.

Experiment	AIC	P Value of Model	Adjusted R Squared
Newsletter-Search	665.0446	0.05625	0.164
Search-Portal	665.1827	0.05876	0.1612
Portal-Retargeting	664.1052	0.04164	0.183

Table 8: Focal Model with online-online synergy

B.4 Experimenting with offline-online synergy

Catalogs and Mailings individually have negative or negligible impacts on sales, but that could be because they only generate awareness in customers, who enter the consideration and purchase stages through online channels. Therefore synergy between offline and online media can drive sales. We examine the synergy of offline channels with Newsletter and Portals. We observe that Portals perform marginally better than Newsletter and hence consider Portals for further analysis.

We further break down the analysis of the offline channels into Catalogs and Mailings and examine their synergy with portals. Upon inspection, the synergy with mailing and portal fares worse than the aggregate offline spend, hence we proceed with bifurcating the catalog spend and running the synergy check on win-back customers and combination of new and existing customers. We finally conclude that the combination of catalogs sent to new and existing customers, despite being marginally inferior in AIC to the aggregate offline channel, best explains the synergy between offline and online channels.

Experiment	**AIC**	**P Value of Model**	**Adjusted R Squared**
Total Offline - Newsletter	646.6087	8.06E-05	0.4668
Total Offline - Portal	645.5683	5.43E-05	0.4801
Mailing - Portal	664.9018	0.05375	0.1669
Catalog_Winback - Portal	662.1213	0.021	0.206
Catalog_NewExist - Portal	646.5368	7.84E-05	0.4677

Table 9: Focal Model with offline-online synergy

B.5 Experimenting with offline-online & online-online synergy

Having established the offline-online synergy, we further experiment it in a combination with online-online synergy. For this, we take the best models from the individual experiments i.e. Portal-Retargeting and Catalog_NewExist-Portal. We observe that this combination performs worse than the offline-online model of synergy.

Experiment	**AIC**	**P Value of Model**	**Adjusted R Squared**
NewExist - Portal & Portal-Retargeting	649.2264	0.0003654	0.4522

Table 10: Focal Model with offline-online & online-online synergy

B.6 Experimenting with removing the lag variable

We finally conclude our analysis by examining the effect of removal of lag variables of sales and observe that in the best case scenarios with all the above analysis, removing the lag variable, consistently performs worse than the one with it included.

Experiment	**AIC**	**P Value of Model**	**Adjusted R Squared**
No Intercept	685.9801	<2.2e-16	0.9728
Log	678.1072	0.007203	0.2112
Portal-Retargeting	679.8672	0.01749	0.2106
NewExist - Portal	662.2541	1.96E-05	0.481
NewExist - Portal & Portal-Retargeting	665.055	0.0001149	0.4659

Table 11: Focal Model without lag sales