Project 4: Marketing Models

MKT-680: Marketing Analytics

Team: Cheryl Cao, Kathy Ding, Roch Jia, Lessly Rocha, and Ye Wang

April 2023

Table of Contents

Introduction

ACSE Supermarket, a prominent retailer with over 40 stores in Locouria, offers a diverse product range spanning more than 100 categories. Their rewards program provides customers with access to enticing weekly sales and promotions. A considerable portion of ACSE's sales relies on these promotions, and the company actively partners with suppliers to fund them. One such partnership is with Mahou San Miguel, a Spanish brewer that entered the Locouria market in 2018. They sell two San Miguel beer products in ACSE's stores and actively promote these products through weekly flyers, in-store displays, web media, and traditional media channels. However, the effectiveness of these promotional efforts and the marketing partnership with ACSE needs to be verified.

With this business background, we have been entrusted by our client, Mahou San Miguel, to identify promotion and marketing activities that drive significant incremental sales for continuation into 2023. Leveraging our extensive experience in marketing mix modeling and data science, we aim to provide valuable insights into the marketing vehicles that generate substantial results.

Problem Definition

To address our client's concerns, we will analyze the transaction, product, and promotional data alongside supplementary data of our clients provided by the ACSE database. This comprehensive dataset will enable us to develop a robust model, taking into account factors such as seasonality, national holidays, and advertising half-life.

Our report will include model specification and diagnostics, decomposition of weekly sales volume for each San Miguel product, and recommendations on which promotional

activities should be continued in 2023. Ultimately, our objective is to optimize Mahou San Miguel's marketing strategy and maximize the effectiveness of their promotional activities in partnership with ACSE.

Data Understanding

Our dataset consists of 5 tables: transactions, products, seasonality, promo_ad and event_holiday. The transactions table contains 2-year transaction data from 2019 to 2020, with only two beer products included: San Miguel Especial and San Miguel Especial 6 PK. The two products are identical except the specification: single can and 6-can package respectively, and they have different pricing strategies. The Products table contains the detailed information for the 2 products. Seasonality table contains the quantified index to measure the effect of seasonality by week. Promo_ad table contains the detailed information of 6 different ads promotion vehicles, including starting week, amount of promotion and cost. Event_holiday table recorded the important events/holidays that may significantly influence the sales of the above products. We will combine all those tables except products to build up a dataframe for Marketing Mix Modelling.

Data Preparation

In order to tackle this business question, we aim to use marketing mix models to evaluate the contribution and ROI of each marketing vehicle and thus make recommendations to Mahou San Miguel for their next year's marketing strategy.

We began with data preparation to collect the information for our marketing mix model. In the transaction data, we aggregated the sales quantity and sales amount into the week level.

Since sales quantity and sales amount are measurements that our client wants to increase, we thought it is appropriate to aggregate them on a weekly basis in order to provide enough detail for the marketing performance. In addition, we also include the seasonality index into our model to capture the seasonal factors that would help the model be more accurate.

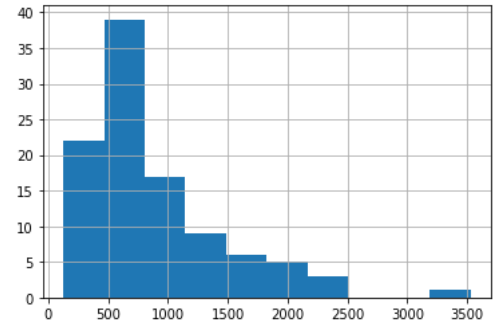
After this, we took a closer look at six marketing vehicles: TV, Radio, Google Paid Search, Flyer, Facebook Banner, and Display. For TV and Radio, we were given a half-life. According to this information, we calculated the decay parameter as well as the reach for these two particular vehicles. For instance, since the half-life for TV is 6 weeks we could use the formula for the decay parameter $\alpha = 1 - (0.5)^{1/h}$ in which h is equal to 6 and then get the value of α. With the decay parameter, we then continued calculating the adstock GRP using the formula $AdStock_t = \alpha GRP_t + (1 - \alpha)\, AdStock_{t-1}$ to calculate the total GRP. The final step in this calculation process was to use the GRPs that we got from the previous step and plug the number into the formula that we were given $Reach = 0.95 * (1 - e^{-0.40*GRP})$. We did the same process of calculations for the Radio. The remaining four vehicles had impressions and thus we decided to leave them as is. Lastly, because we aimed to build a regression model in the following step, we did one-hot encoding for the only categorical variable `event` in our dataset.

Modeling

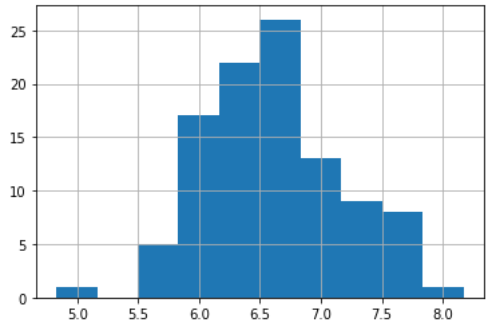
As part of the modeling process, we started by separating the two San Miguel beer products offered at ACSE, which are the single beer and the six-pack. It made more sense to treat the two different products separately to analyze the impact of the marketing vehicles on each. We also made the decision to use sales quantity instead of sales amount as our dependent variable, as

we think the price of the product varies over time and cannot be taken as an appropriate metric for measuring sales.

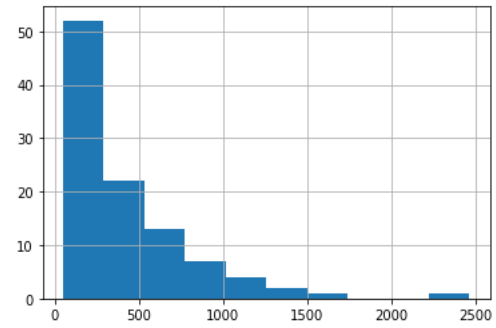
After this, we applied log transformations to the sales quantity of both products since the distribution for each was skewed.



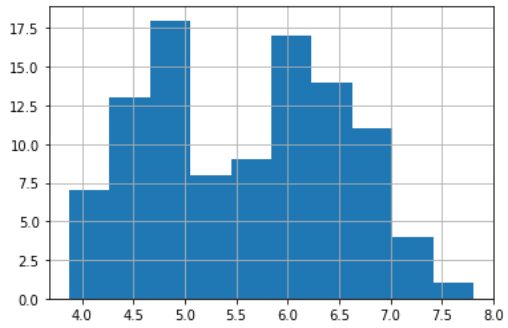
Sales_qty for single beer before log transformation



Sales_qty for single beer after log transformation



Sales_qty for 6-pack beer before log transformation



Sales_qty for 6-pack beer after log transformation

We also created dummy variables for each event holiday type (e.g. Christmas, Easter, etc.). In summary, the aforementioned action enabled us to prepare the data in a linear and machine-learning way so that the OLS Regression (which has a linear assumption) could be trained on our data. We feed each individual product's features, which are shown in the footnote.[1] The adjusted R-square for the San Miguel single beer was .887 and the adjusted R-squared for

[1] Features feed to the model: seasonality index, event holiday Christmas, event holiday Easter, event holiday Halloween, event holiday Labor Day, event holiday Memorial Day, event holiday National Day, event holiday New Year, event holiday Pre Easter, event holiday Pre Super Bowl, event holiday Super Bowl, promo display, promo facebook, promo flyer, promo Google, promo radio reach, and promo TV reach.

the San Miguel 6-pack was .901. Both of these numbers let us know that the marketing vehicles and the uncontrollable factors (like seasonality) account explain that approximately 88.7% of the variations in the San Miguel single beer and approximately 90.1% of the variation in the San Miguel six-pack beer can be explained by the predictor variables used in the regression model.

OLS Regression Results

Dep. Variable:	y	R-squared:	0.906
Model:	OLS	Adj. R-squared:	0.887
Method:	Least Squares	F-statistic:	47.64
Date:	Fri, 21 Apr 2023	Prob (F-statistic):	7.76e-36
Time:	00:07:44	Log-Likelihood:	33.051
No. Observations:	102	AIC:	-30.10
Df Residuals:	84	BIC:	17.15
Df Model:	17		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	5.1846	0.143	36.331	0.000	4.901	5.468
seas_index	0.8796	0.155	5.685	0.000	0.572	1.187
event_holiday_Christmas	0.4022	0.148	2.726	0.008	0.109	0.696
event_holiday_Easter	-0.0004	0.140	-0.003	0.998	-0.278	0.277
event_holiday_Halloween	0.0972	0.143	0.681	0.498	-0.187	0.381
event_holiday_Labor Day	-0.0487	0.198	-0.246	0.806	-0.442	0.345
event_holiday_Memorial Day	0.0805	0.142	0.567	0.572	-0.202	0.363
event_holiday_National Day	0.2686	0.155	1.728	0.088	-0.040	0.578
event_holiday_New Year	-0.8119	0.150	-5.420	0.000	-1.110	-0.514
event_holiday_Pre Easter	-0.0253	0.142	-0.178	0.859	-0.309	0.258
event_holiday_Pre Super Bowl	0.0701	0.201	0.349	0.728	-0.329	0.469
event_holiday_Super Bowl	-0.0162	0.142	-0.114	0.910	-0.299	0.266
promo_display	0.1730	0.294	0.587	0.558	-0.413	0.759
promo_fb	-1.468e-06	7.93e-07	-1.852	0.068	-3.04e-06	1.08e-07
promo_flyer	0.6661	0.053	12.578	0.000	0.561	0.771
promo_google	2.514e-05	6.82e-06	3.689	0.000	1.16e-05	3.87e-05
promo_radio_reach	0.3477	0.071	4.910	0.000	0.207	0.489
promo_tv_reach	0.4003	0.090	4.454	0.000	0.222	0.579

Omnibus:	15.981	Durbin-Watson:	1.155
Prob(Omnibus):	0.000	Jarque-Bera (JB):	60.420
Skew:	-0.232	Prob(JB):	7.59e-14
Kurtosis:	6.742	Cond. No.	5.36e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.36e+05. This might indicate that there are
strong multicollinearity or other numerical problems.

The above table shows the result for the single beer regression. According to the table, we can tell that some of the holidays, such as Christmas and National Day, have a statistically significant effect on the sales quantity. For example, the coefficient for event_holiday_Christmas is 0.4022, which means that during the Christmas holiday, the sales quantity of the San Miguel single beer can be expected to increase by e^(0.4022) = 1.495 multiplicatively. As for different vehicles, we found from the summary table that flyers, google, radio, and tv all have a statistically significant effect on the sales quantity. Among these vehicles, flyers have the biggest effect on sales quantity: an increase of 1 unit in the flyer can be expected to increase the sales quantity by 1.946 multiplicatively.

Similar to the single beer product, we got the summary table for the six-pack beer as well. Through the result table, we observed that some of the holiday events that are unlikely to have a statistically significant effect on the sales quantity of single beer do have an effect on the six-pack product. For example, pre-Easter and pre-Super Bowl are expected to increase the sales quantity multiplicatively by 1.685 and 1.936 respectively. These observations make sense as people are likely to buy six-packs to get prepared for watching the Super Bowl with a couple of friends or family members. Looking at the vehicle for the six-pack product, we found that except for Facebook, all other five vehicles have a statistically significant effect on the sales quantity.

DueTos

After this, we moved on to calculate the Duetos of each marketing vehicle. That is, what portion of the sales can be attributed to each marketing vehicle. In order to do this, we calculated the sales quantity for the San Miguel single beer and for the San Miguel 6-pack without each of the marketing vehicles. That is, we calculated the “base” of what quantity of each product would

be sold if each of the marketing vehicles was not present. We did this by setting, for example, a flyer as 0 to calculate the sales quantity that would happen if flyers were not to be used as a marketing vehicle in the weeks that it was used—what we call base sales. Then, we subtracted the base sales quantity prediction that we calculated by setting the flyer to 0 for a given week from the number of our predicted sales for that same week when the flyer was actually set to 1 (meaning that flyers were used as a promotion vehicle). This way we see the additional sales quantity that happened due to the flyer being used as a marketing vehicle and we can attribute those sales to the flyer marketing vehicle. We proceeded to do a similar process with the remaining five marketing vehicles.

After this step, we came up with a way of estimating the profit for the product: we calculated the average price of two products respectively, and then multiply the average price by a profit margin to get the final estimated profit for each product. We also would like to note here that we use a profit margin of 0.5 for the single beer and 0.4 for the six-pack beer. A couple of reasons led us to make such an estimation: first, San Miguel is one of the leading companies in the industry, and thus has the advantageous bargaining power to keep the profit margin in the range between 0.3 to 0.5; second, the average unit price we have calculated for every single beer is lower for the six-pack beer, and thus we decided to use 0.4 for the six-pack beer and 0.5 for the single beer to reflect the price differences between two products we've noticed in the data. With unit profit, we can easily calculate the DueTos profits for each vehicle and then compare them with the actual costs. As for ROI, we calculate it by using the DueTo profit divided by Cost and minus 1. The resulting DueTo profits and ROI for the San Miguel single beer, the six-pack, and two products in total are summarized below:

Marketing Vehicle	**DueTo Profits**	**Cost**	**ROI**

Display	$679.2	$1,500	-54.7%
Flyer	$45,364.6	$32,875	37.9%
Google Paid Search	$18,444.3	$37,464.5	-50.8%
Radio	$29,769.2	$24,600	21.0%
TV	$39,168.6	$69,333.33	-43.5%
Facebook Banner	-$4,362.4	$28,445.25	-115.3%

DueTos and ROI for San Miguel single beer

Marketing Vehicle	**DueTo Profits**	**Cost**	**ROI**
Display	$29,811.9	$63,000	-52.7%
Flyer	$34,291.6	$31,375	9.3%
Google Paid Search	$72,436.9	$37,464.5	93.3%
Radio	$39,232.8	$24,600	59.5%
TV	$64,842.3	$69,333.33	19.6%
Facebook Banner	-$2,976.7	$28,445.25	-110.5%

DuseTos and ROI for San Miguel six-pack beer

Marketing Vehicle	**DueTo Profits**	**Cost**	**ROI**
Display	$30,491.1	$64,500.0	-52.7%
Flyer	$79,656.1	$64,250.0	23.9%
Google Paid Search	$90,881.2	$74,929.0	21.3%
Radio	$69,002.0	$49,200.0	40.2%
TV	$104,010.9	$138,666.67	-24.9%
Facebook Banner	-$7,339.1	$56,890.5	-112.9%

DuseTos and ROI for San Miguel products in total

According to the result, we found that for the single beer, flyer and radio have relatively high ROI, while for the sic-pack beer, flyer, radio, as well as google paid search all have relatively high ROI. If we take both products into account, the result is similar to what we've noticed for the six-pack beer—flyer, radio, and google paid search seems to be effective vehicles with relatively high ROI.

In order to provide our client with more actionable insights, we also tried to calculate the ROIs across different years. In this case, we have done the calculations for the years 2019 and 2020. The summary table were shown as follows:

Marketing Vehicle	**DueTo Profits**	**Cost**	**ROI**
Display	$15,944.5	$34,500.0	-53.8%
Flyer	$40,251.8	$31,250.0	28.8%
Google Paid Search	$39,881.8	$37,964.0	5.1%
Radio	$33,955.5	$24,000.0	41.5%
TV	$46,696.4	$104,000.0	-55.1%
Facebook Banner	-$5,110.1	$42,062.75	-112.1%

DuseTos and ROI for San Miguel products in total in 2019

Marketing Vehicle	**DueTo Profits**	**Cost**	**ROI**
Display	$14068.9	$30000.0	-53.1%
Flyer	$38658.3	$33000.0	17.1%
Google Paid Search	$50563.7	$36965.0	36.8%
Radio	$34529.3	$25200.0	37.0%
TV	$56786.1	$34666.67	63.8%

Facebook Banner	-$2068.4	$14827.75	-113.9%

DuseTos and ROI for San Miguel products in total in 2020

Comparing the ROI for all six vehicles for years 2019 and 2020 respectively, we observed that TV in the year 2020 showed a high ROI, while in the year 2019, it returns a negative ROI. The ROI for Google paid search also increased greatly in 2020, compared to its ROI in 2019. In contrast, the ROI for flyers decreased in 2020. These observations make sense as the pandemic hits in 2020. During the quarantine, it is hard for flyers to reach people, and people might spend more time watching tv and searching online, making ROIs of google paid search and tv go up.

Deployment

Based on our analysis, we provide the following recommendations to Mahou San Miguel to optimize their marketing strategy in partnership with ACSE:

1) **Continue focusing on marketing vehicles with high ROI:** Flyers, radio, and Google Paid Search have demonstrated relatively high ROIs for both San Miguel single beer and six-pack beer products. These vehicles should be the primary focus of Mahou San Miguel's promotional activities moving forward.
2) **Reevaluate the effectiveness of low-performing marketing vehicles:** Display, TV, and Facebook Banner have shown negative or low ROIs. Mahou San Miguel should reevaluate the strategy and performance of these

vehicles, considering factors like targeting, creative, and media placements. For instance, TV showed a high ROI in 2020, but not in 2019, so it is important to identify the reasons behind this change and determine if it can be a consistently effective marketing vehicle.

3) **Consider the impact of time-series factors:** We observed significantly different evaluation results for 2019 and 2020 respectively. Apart from budget allocation changes in different vehicles, there may be two main reasons. On the one hand, some promotion vehicles (TV, radio) have long-term influence and can influence the performance for the next year. On the other hand, the pandemic has significantly impacted the effectiveness of certain marketing vehicles, as seen with the decrease in ROI for flyers and the increase in ROI for TV and Google Paid Search in 2020. Mahou San Miguel should monitor the ongoing impact of those factors and adapt their marketing strategy accordingly.
4) **Tailor marketing vehicles to specific products:** Certain marketing vehicles may be more effective for specific products, such as the six-pack beer being more influenced by pre-Super Bowl and pre-Easter events. Mahou San Miguel should tailor their marketing strategy for each product to maximize the effectiveness of their promotional activities.

By following these recommendations, Mahou San Miguel can optimize their marketing strategy and maximize the effectiveness of their promotional activities in partnership with ACSE, ultimately driving significant incremental sales and improving their return on investment.