Marketing Mix Modeling for Mahou San Miguel and Pernalonga

Harshit Arora, Patrick He, Halle Steinberg, Sean Yu

Team 1

MKT 680 Marketing Analytics

Emory University

April 30th, 2020

Executive Summary

After careful analysis and consideration of the promotional partnership between Pernalonga and Mahou San Miguel, we have created various models that deconstruct the effects of various marketing vehicles on Pernalonga's sales quantity for three Mahou San Miguel products. The three products analyzed include:

- Product 138936951 (Single Can)
- Product 138936952 (6-pack)
- Product 138936953 (Case)

Each product was modeled separately due to differences in promotional vehicles used historically. In conclusion, we generally recommend different marketing channels to be used across the three products. We most strongly suggest the use of Email promotions for 6-packs and Cases, TV advertisements for Single Cans, and discounts and Radio advertisements for all three products. We strongly recommend discontinuing TV advertisements for both 6-packs and Cases and Flyers for Single Cans. Our report will dive into our analysis and insights and provide context to our recommendations and findings.

Background

Pernalonga is a leading supermarket chain with over 400 stores in Lunitunia that offers a vast variety of over 10 thousand products among more than 400 categories. Its business is currently heavily reliant on promotions, both in-store and through third-party suppliers, with nearly 30% of sales being driven by promotions. We have done ample research thus far with our client to understand how its customers behave and react to promotions, so we feel we have sufficient business understanding to push our analysis further. We know that there are many marketing channels through which Pernalonga can promote its products in order to maximize key performance metrics such as revenue and profit, and our task is to uncover how each of these channels plays a part. As marketers, we need to uncover which channels, or marketing "vehicles", are most responsible for revenue and profit and encourage Pernalonga to promote its products through the most efficient and successful channels possible.

Our goal with this report is to understand the contribution of various advertisement vehicles to Pernalonga's success and provide Pernalonga with information about how its customers are affected by each type. To understand this, we will focus on analyzing the effects of various marketing vehicles on sales quantity for Mahou San Miguel. Mahou San Miguel is a Spanish brewer that currently sells three of its products in Pernalonga's stores. The two businesses regularly partner together to promote the beers through more static platforms like weekly flyers and in-store promotions, but we know that Mahou San Miguel also has experience promoting through more dynamic vehicles such as email, web, TV, and radio. Mahou San Miguel is interested in figuring out which of these vehicles will help it to increase sales in Pernalonga's stores. By continuing to explore Pernalonga's transactional data, we were able to solve this problem and come up with a marketing mix model for Mahou San Miguel using explanatory regression modeling techniques to identify the best marketing vehicles and promotional opportunities for the coming year, ultimately helping to drive sales and boost Pernalonga's revenue.

Business Context and Data Understanding

Mahou San Miguel is currently the market leader in Spain, with around 36% of the country's beer market share, just ahead of Heineken and Damm. It currently produces more than 70% of all Spanish beer consumed worldwide and has a presence in over 70 countries around the world, offering its customers products across more than 50 different brands[1]. Mahou San Miguel holds a large share of the market in Spain, but it has unfortunately not seen the same success in Lunitunia, and specifically, in Pernalonga's stores[2]. Upon initial research, we discovered that Mahou San Miguel currently sells only three of its products in Pernalonga's stores. Each of these three products vary in how they are packaged and sold - one product is sold as a Single Can, one is sold as a 6-pack, and one is sold as a 24-pack or Case. When we look at all historical transactions we have in our data from the beginning of 2016 through the end of 2017, we see a range of prices and quantities sold across the three products, with the Case selling at the highest average unit price but the lowest price per can. In terms of volume, Cases are sold much less often than Single Cans or 6-packs, providing only marginally more total sales than Single Cans.

We can see that 6-packs account for the majority (56%) of Mahou San Miguel's total sales in Pernalonga's stores. Interestingly, Single Cans are sold the most frequently, but result in the least amount of sales. They are also discounted at the lowest rate on average, historically being sold with a discount of only about 4% of the average unit price, compared to about 5% and 8% for 6-packs and Cases, respectively. When we look at how Mahou San Miguel products compare to others in the beer category, these three products make up approximately 5% of all of Pernalonga's beer total sales.

Product ID	Product Type	Mean Unit Price	Mean Price Per Can	Total Quantity Sold	Total Sales	Mean Discount Amount
138936951	Single can	$0.75	$0.75	22,834	$17,028.66	$0.03
138936952	6-pack	$4.29	$0.72	11,120	$47,637.90	$0.21
138936953	Case	$15.64	$0.65	1,276	$19,957.74	$1.22

Figure 1: Descriptive Statistics for Mahou San Miguel Products in 2016 and 2017

We also have access to data regarding important national holidays, seasonality indicators, and various marketing vehicles that have been used to promote Mahou San Miguel products in the past. This data is provided at a weekly level, allowing us to easily implement these factors as variables in our models where we are trying to predict weekly sales quantity.

For both the beer category and Mahou San Miguel products, there's a clear effect from seasonality on the volume of beer sold at Pernalonga's stores. As we see in the graph below, beer sales appear to reach a peak in June and July, which are the warmest months in the year, and drop to the bottom at the end of the year in the winter months. Based on this knowledge, Pernalonga should consider promoting Mahou San Miguel products during the summer months, as this is when natural demand for beer is already the highest.

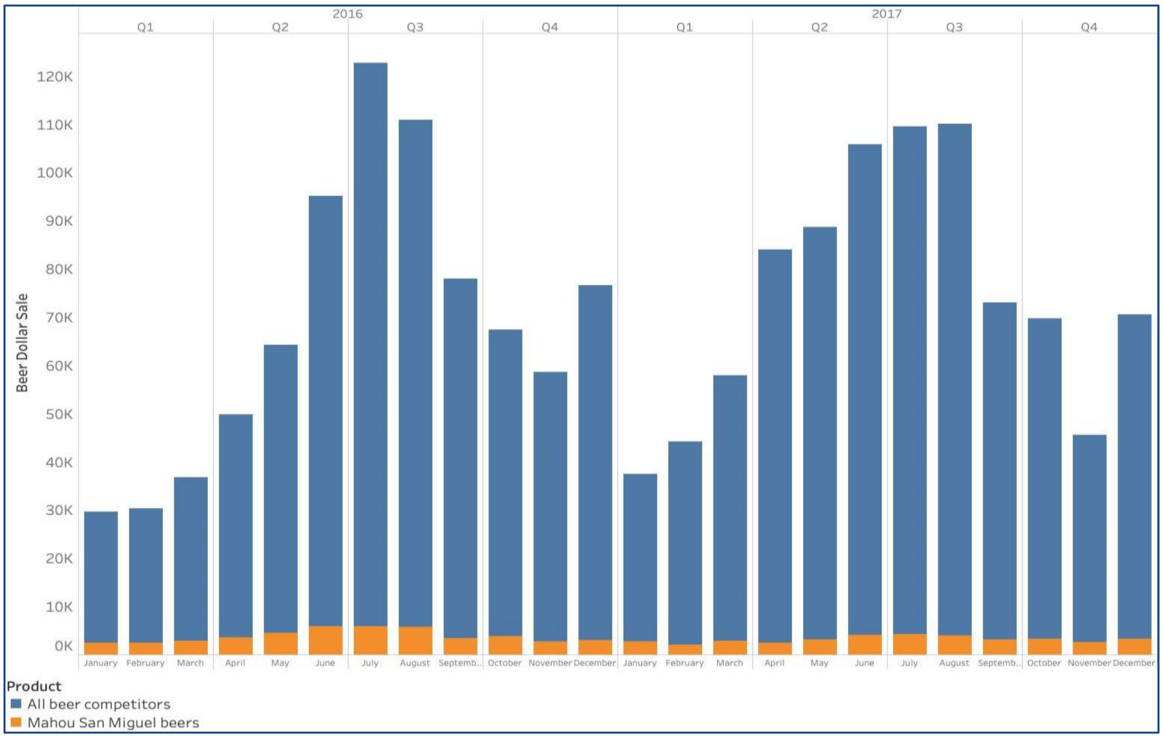

Figure 2: Beer Sales ($) by Month for Pernalonga's Beer Products

The marketing vehicles data has information for seven different types of promotional channels, but we see that only six of them have historically been used to promote all three of Mahou San Miguel's products: TV, Radio, Paid Search, Web Display, Flyers, and Emails. Store Displays have only been used to promote Mahou San Miguel's 6-pack and Case products and never the Single Can. The success and value of each of these marketing vehicles are quantified differently, with Paid Search and Web Displays measured in impressions, and TV and Radio measured in units of Gross Rating Points, or GRP, which is defined as the total number of impressions of an advertisement with respect to a target audience. GRP is the metric most commonly used to measure advertisements and the data used to compute it can be collected in a number of ways.

Our goal in this report is to predict weekly sales quantity given various levels of use for different marketing vehicles that Mahou San Miguel employs. Through regression modeling approaches, we will decompose Pernalonga's sales quantity for each product into the amount that can be attributed to each of these vehicles. In the next section of our report, we will detail how we prepared the data for modeling, including getting all of our data in a weekly format and converting certain variables into necessary measures to be used in a marketing mix model.

[1]*Source:* https://www.just-drinks.com/analysis/mahou-san-miguel-spains-beer-market-leader-just-the-facts_id118782.aspx
[2]*Source:* https://www.mahou-sanmiguel.com/en-gb/brands#gama

Data Preparation

To begin our analysis, we first merged the two datasets to create a file that contains all product and transaction information. This gave us a dataset containing nearly 30 million observations across 421 stores and 429 product categories. In total, our data consists of 10,767 unique products. We know from our previous analysis that there are some anomalies in the data that must be accounted for. One of the biggest errors in the data comes from the fact that there are only 753 unique transaction ID's assigned to the data. To account for the errors in these ID's, we again used a combination of customer ID, store ID, and transaction date to correctly define the transaction ID's, which resulted in 2.83 million unique transactions with a varying number of products purchased across all consumers. We also removed transactions in the last week of 2015, as in this week there was only one day that saw any transactions, which skews any results when we aggregate at the weekly level.

Before building our regression models and exploring the decomposition of sales quantity, there are general steps to take in marketing mix modeling and some variable considerations and impacts of marketing vehicles need to be properly addressed. Among all seven marketing vehicles in our data, both TV and Radio are measured in Gross Rating Points, or GRPs. As a general principle, for each promotional media channel, the measurement for each one should be converted from GRPs to Reach. The first step for us was to convert each TV and Radio GRP value to AdStock GRPs, which could then be converted to Reach. Given that we know TV advertisements have an 8-week half-life and Radio advertisements have a 4-week half-life, we were able to compute the alpha value (α), or decay parameter, for both TV and Radio using the following formulas.

$$\alpha_{TV} = 1 - (0.5)^{\frac{1}{8}}$$
$$\alpha_{Radio} = 1 - (0.5)^{\frac{1}{4}}$$

Next, we computed the AdStock GRPs for the period impacted by the TV and Radio marketing vehicles. AdStock GRPs help us account for the natural short-term retention of media effects on consumers over time, and they are calculated using the decay parameter and GRP, as shown below.

$$AdStock_t = \alpha GRP_t + (1-\alpha) AdStock_{t-1}$$

Finally, we converted AdStock GRPs to Reach at a weekly level. We know that Reach offers a better measurement of the actual effect of TV and Radio advertisements, so we mapped the AdStock GRPs into a representation of 2+ Reach. In our case, this represents the percentage of the target audience who has seen the advertisement at least twice. When we use Reach instead of GRP, we can model the percent of target audience that is expected to retain the message over time, rather than solely using raw GRPs or impressions. The following equations were used to calculate Reach for TV and Radio.

$$Reach_{TV} = 0.95(1 - e^{-0.020(AdStock\ GRPs)})$$

$$Reach_{Radio} = 0.90(1 - e^{-0.025(AdStock\ GRPs)})$$

It's also important to note that some marketing channels were only used in a certain period, so for the weeks that the channels were not used, we considered the effect of the marketing vehicle to be zero. For example, TV and Radio advertisements were both first distributed on June 5th, 2016, therefore, the sales quantity in the weeks before June 5th, 2016, are not considered to be impacted by TV and Radio advertisements. Additionally, not every marketing vehicle is applied to all three products - Flyers and Store Displays were only used for one or two Mahou San Miguel products at a given time. In order to uncover the effects of the various marketing factors on different products, we decided to build models separately for each of the three Mahou San Miguel products.

We know that TV and Radio GRPs needed to be appropriately converted for use in our model, so we made sure to appropriately transform them from GRPs to Reach. On the other hand, the other five marketing vehicles - Paid Search, Web Display, Email, Store Display, and Flyer - did not need to be converted or transformed in any way and could be used as originally measured in our data. These marketing vehicles were already measured by week, so we were able to simply use the measurements for these vehicles without any manipulations. For Email, it was represented as the number of circulations of the email that was sent in that week. For Paid Search and Web Display, the number of impressions in the corresponding week was used. For Flyer and Store Display, these were kept as binary measures, with a 1 indicating that the promotional vehicle was used in the week, and a 0 indicating it was not.

One thing that we noted for both Email and Web Display is that the values in the data for these two are very limited. For Email, there is only one unique value for circulations, and for Web Display, there are four unique values for impressions. We considered whether we should treat these two variables as factors, but ultimately, we decided to treat them both as integers because both circulations and impressions could realistically be any whole number. We are making this assumption in order to still capture the nuances and intricacies in these variables as best as possible.

Since our target variable is the weekly sales quantity of each product, we created two product-related variables to be used as independent variables in our models.

- ***Shelf Price*** - We calculated the weighted average shelf price for each product for each corresponding week. The weighted average shelf price is calculated by multiplying the average shelf price at each store by a weight, which in our case, was the total sales. Surprisingly, we found the weighted average price in each week is the same as simply the average shelf price. That said, in a particular week, the price of a product won't change across stores.
- ***Discount Amount*** - We used the average discount rate for each product for each corresponding week. The average discount for each product in each week is calculated by adding up the total discount amount then dividing it by the total sales volume in that week.

Finally, we also included seasonality for each week and a holiday index in our analysis to properly reflect the base of sales quantity which would not be affected by the addition of any marketing vehicles. For holidays, we used the data we were given as is, which indicated a week when there was a holiday and which holiday it was. We wanted to be able to see the impact that a specific holiday might have on sales quantity, so we chose to treat each holiday as a factor in the model so that we could understand any unique differences between weeks that have specific holidays and those that do not. We also merged the seasonality index for products in the beer category with each corresponding week in our data to properly reflect the seasonal sales pattern that we see in the base sales quantity.

Modeling Approach

The next step in our marketing mix modeling process was to develop models that would allow us to quantify the effect of each marketing vehicle on sales quantity. To achieve this, we relied upon logit-based models for each of the three Mahou San Miguel products. Logit models have the advantage of using a bounded dependent variable, and they allow for capturing complex implicit interaction which can occur due to multiple and similar marketing vehicles, such as TV and Radio. The base form of the model is:

$$\log\left(\frac{y_t}{1-y_t}\right) = \sum_i \beta_i g_i(x_{it}) + \epsilon_t$$

We used a log-transformed version of weekly sales quantity, the dependent variable in our models, by bounding the sales quantity with the theoretical maximum weekly sales quantity. It is possible that Pernalonga's stores may have a higher theoretical maximum than what has been seen in the two-year span of the data. Thus, we assumed this to be a value 20% higher than the historical maximum for each product.

In terms of the independent variables used in our models, we included all marketing campaigns: TV, Radio, Paid Search, Web Display, Email, Store Display, and Flyers. As described earlier, variables for the marketing vehicles included TV and Radio Reach, whether or not Flyers and Store Displays were used (binary variables), E-mail circulations, and Web Display and Paid Search impressions. Although we developed fairly identical models for all three products, it is important to note that product 138936951 (Single Can) did not have any Store Displays in the entirety of the data. Thus, this variable was excluded from this product's model. Moreover, we included factors such as seasonality index and holidays in all models, as these are factors that can heavily affect sales and are important to our baseline. Finally, weighted shelf price and average weekly discount percentage were included to capture the product statistics in each week.

Once we ran our initial models using all the mentioned independent variables, we noticed that three of the holiday factors were perfectly correlated with the Email marketing vehicle variable: "XMAS", "PrXMAS", and "NEWYEAR". This makes sense, as Email promotions only occurred in the same weeks as Christmas and New Year, causing multicollinearity between the variables. To ensure that we can include Email in the model, we verified which of these three holidays actually affect sales quantity. This was verified by comparing the sales quantity in the holiday weeks against the immediately preceding and succeeding weeks, and we found that sales tend to see a surge in the week of "PrXMAS" and "NEWYEAR". We also built a secondary model without Email to see if these two variables were statistically significant in predicting weekly sales quantity. Thus, our final models excluded the "XMAS" holiday as a variable, but included "PrXMAS", "NEWYEAR", and Email.

Model Evaluation and Results

Based on the F-test for Linear Regression, all three of our product-specific models were statistically significant with p-values < 0.05, indicating that the models do not have a lack of fit issue. We observed some intuitive yet interesting insights: for all products, variables such as discount, some holidays, and seasonality index seemed to have a positive impact on sales quantity, while variables such as shelf price and some holidays (for the Single Can) had a negative impact. That is, sales tend to see a dip on some specific holidays, as well as when shelf prices tend to increase. On the other hand, sales quantity saw a surge when higher discounts were provided. The table below summarizes the variables that seem to have a significant impact on sales quantity ($p < 0.05$) and whether that impact is positive or negative.

Product ID	Variables with Positive Impact on Sales Quantity	Variables with Negative Impact on Sales Quantity
138936951 (Single Can)	Discount (percentage), Seasonality Index	Shelf Price, Labor Day, New Year, Republic
138936952 (6-pack)	Discount (percentage), Seasonality Index, Restoration, Assumption	Shelf Price
138936953 (Case)	Discount (percentage), Seasonality Index	Shelf Price

Figure 3: Impact of Model Variables on Sales Quantity by Product

More importantly, we wanted to ensure that our models were able to perform reasonably well, as this was critical to achieving an accurate decomposition of sales quantity into the various marketing vehicles used in Pernalonga's stores. To verify this, we used several metrics to measure the model performance.

Product ID	RMSE	MAE	MAPE	Multiple R^2
138936951 (Single Can)	31.42	20.80	0.14	0.49
138936952 (6-pack)	16.28	12.36	0.13	0.75
138936953 (Case)	4.00	2.87	0.40	0.83

Figure 4: Model Performance Metrics by Product

It is interesting to note that while Product 138936953 (Case) has the smallest RMSE, it has the highest MAPE of all three product models. This can be explained by the fact that the Case sold at much lesser quantities across fewer transactions, as opposed to the Single Can or 6-pack. Given the very different nature of the three products in terms of how their quantities are measured, MAPE seems to be a good measure of model performance, along with R^2, which is a measure of the variance explained in quantities sold over time by the models. Despite the differences, all three models performed reasonably well at measuring the sales quantity variation over time. This can be best visualized using the below graphs that show the actual weekly sales quantity versus the predicted weekly sales quantity for each model.

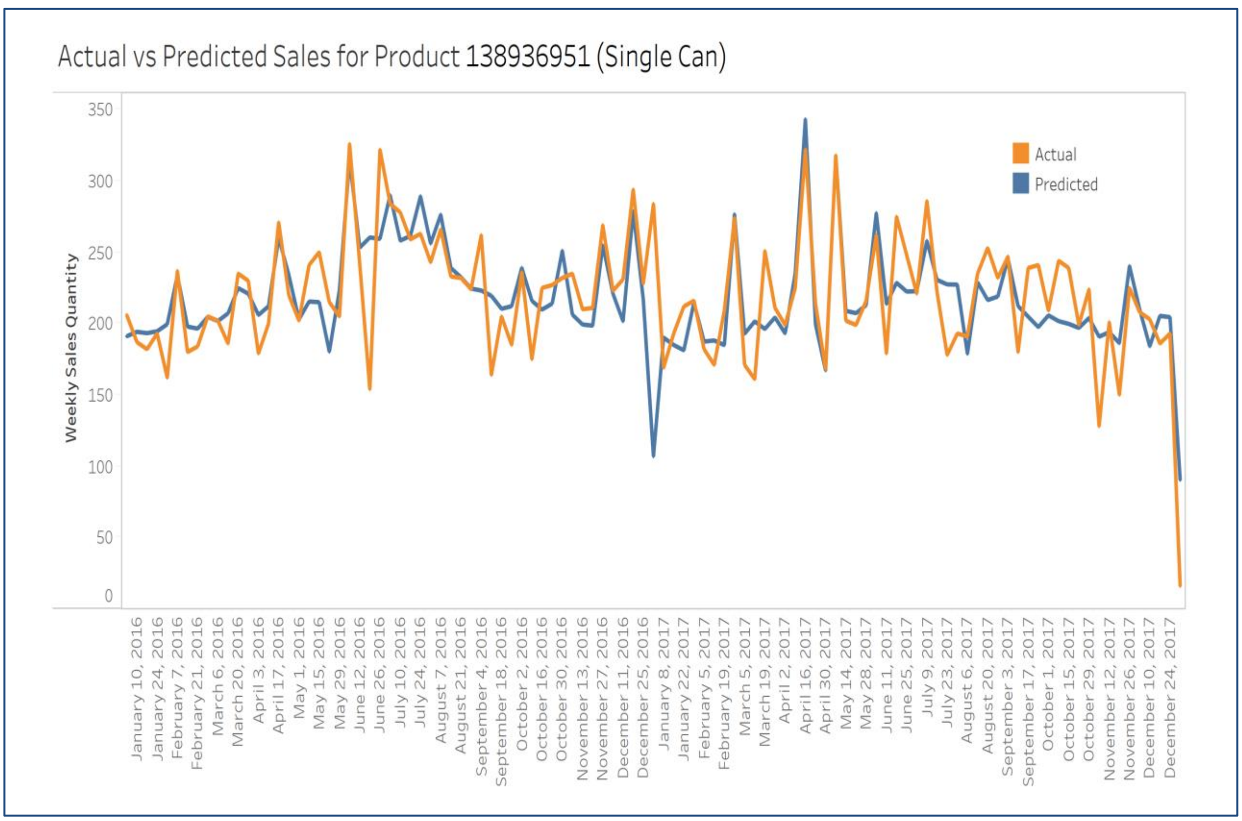

Figure 5: Actual vs. Predicted Sales for Single Cans

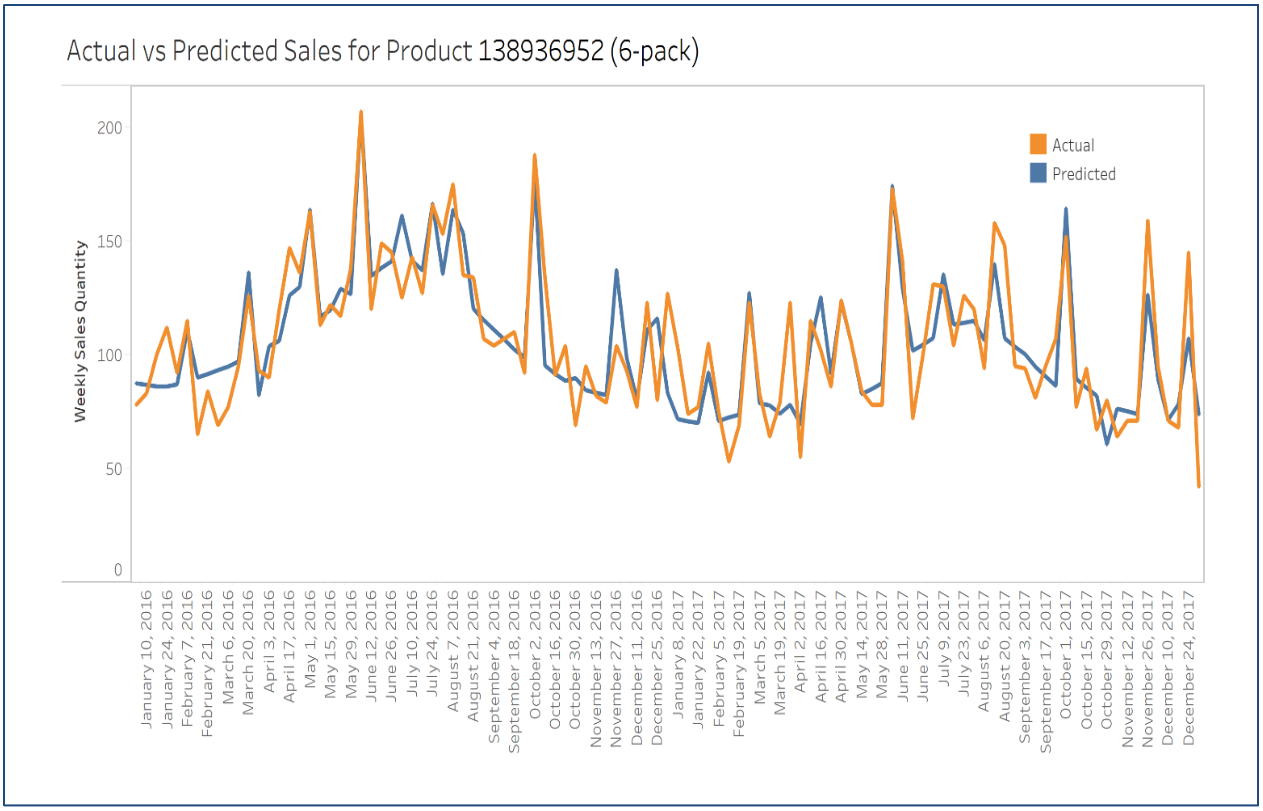

Figure 6: Actual vs. Predicted Sales for 6-packs

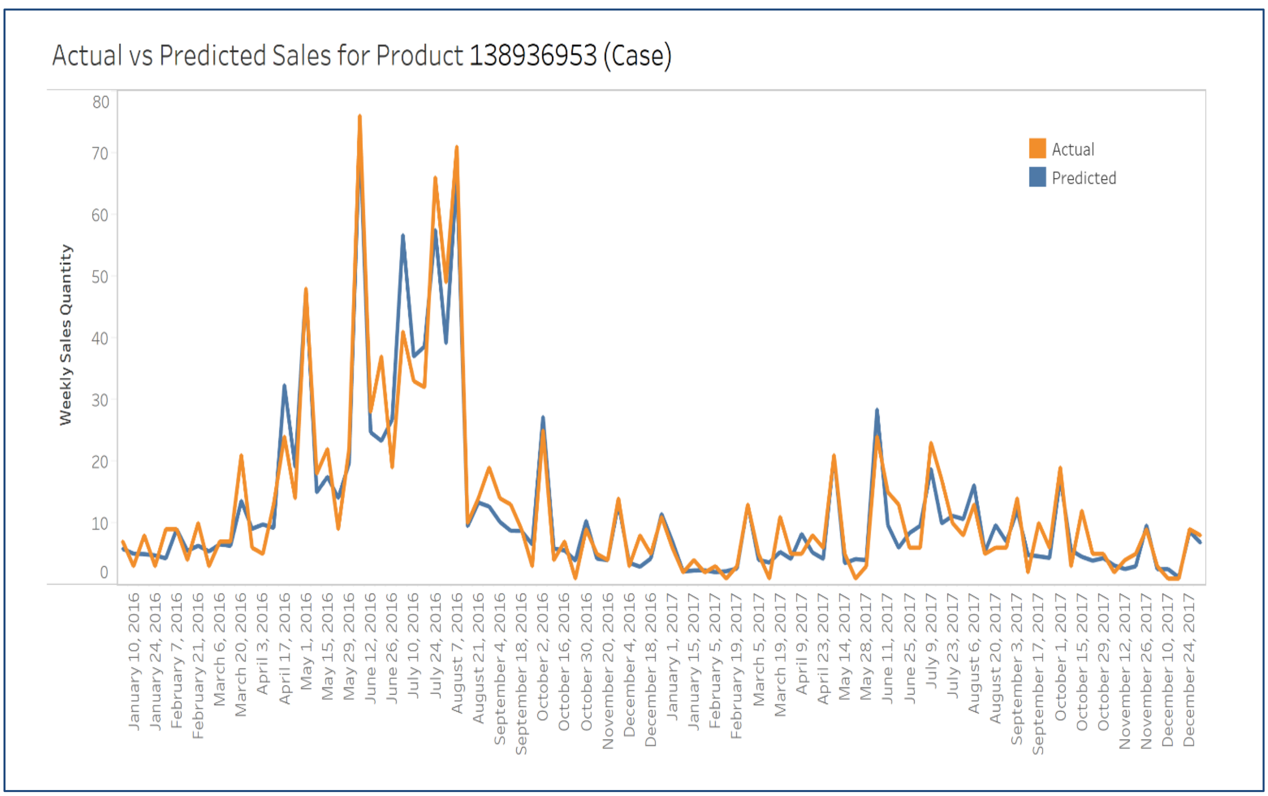

Figure 7: Actual vs. Predicted Sales for Cases

In general, all three models did reasonably well at measuring the variation in sales quantity. However, given the time series nature of the data, it was important to see if there was any presence of autocorrelation, which can create a bias in the standard errors of the model coefficients. To measure this, we used the Durbin-Watson Statistic, for which the result is shown below.

Product ID	Value of *d* (DW Statistic Test)
138936951 (Single Can)	1.86
138936952 (6-pack)	2.17
138936953 (Case)	1.92

Figure 8: Durbin-Watson Statistic by Product

A value close to 2 indicates that there is strong indication of no autocorrelation, which roughly resembles the values our models have. This is verified by the p-value, which is > 0.05 for all three models, indicating that the null hypothesis (H_0 = There is no first-order autocorrelation) cannot be rejected. Thus, the estimated standard errors should be close to the true standard errors.

Finally, it is important to ensure that there are no multicollinearity issues between the predictor variables in our models. Thus, we used the Variance Inflation Factor, or VIF, to quantify the measure of multicollinearity. Generally, for any variable, if VIF > 5, the variable is considered to have a high level of collinearity with other variables, indicating that at least 80% of the variance of the variable can be explained by the other variables. For all our products, however, the VIF values were lower than 5, indicating that the model did not contain multicollinearity issues.

Variable	138936951 (Single Can)	138936952 (6-pack Can)	138936953 (Case)
Weekly Discount	3.13	1.35	1.38
Weekly Shelf Price	1.39	3.06	4.75
Flyer	2.35	1.92	4.53
Web Display	1.16	2.04	2.81
Paid Search	1.17	1.17	1.20
Email	2.18	1.18	1.21
Store Display	-	1.31	4.18
TV	3.30	3.27	3.37
Radio	2.88	2.82	2.95

Seasonality Index	1.74	1.78	1.80
Holiday *(maximum among all holidays)*	1.16	1.14	1.25

Figure 9: Variable VIF values by Product

DueTo Analysis and Results

The final step of marketing mix modeling includes the computation of DueTos in order to decompose sales quantity into the sales quantity Pernalonga would have seen regardless of any additional marketing activity and the sales quantity due to each of Mahou San Miguel's various marketing initiatives. Ultimately, the goal of a marketing mix model is to decompose sales quantity into its base and other components that can be attributed to each promotional vehicle. These components are called DueTos, and they should be calculated in order to determine which promotional channel or vehicle is contributing most to Mahou San Miguel's sales quantity in Pernalonga's stores.

To begin, we needed to first set a base for sales quantity. The base portion of a marketing goal, sales quantity in our case, is that portion that is not affected by the marketing vehicles of interest in the measurement. In our case, we define the base for each product as the base value plus any contribution from seasonality and holidays. Next, we needed to figure out the base value of each of our marketing vehicles. There are some heuristics we can follow, such as for media and promotions, the base is set equal to 0 in general. In our case, this includes any holidays or Email promotions. For price, the base can typically be set as either the historical average price of the product or the first price the product was sold at in our time frame, and in our case, we have chosen to set it equal to the first shelf price we see in the history of the corresponding product.

Once we established the base values for the marketing vehicles for each of our three products, we could then begin calculating each DueTo. Since we used a logit model, which is a form of multiplicative models, we can calculate the DueTo as the difference between the predicted value of the dependent variable when a marketing vehicle is added and the predicted value of the dependent variable when the marketing vehicle is kept at its base. For each product, we calculated the predicted sales quantity of each product in each week using the ***actual variable values***, and then we calculated the predicted sales quantity of the product where each marketing vehicle is set to its ***corresponding base value***. The difference between these values represents the corresponding DueTo for that marketing vehicle prior to any re-scaling. We followed this process for each of the three Mahou San Miguel products, ultimately re-scaling the data to match the original scale and units of our sales quantity variable, which is simply dollars.

Inherently in logit models, there is a transformation bias in each of the predicted values. In order to appropriately debias our model results, we needed to scale each DueTo so that the sum of all DueTos was equal to the corresponding product's actual weekly sales quantity. To do this, we first calculated the sum of ***all*** DueTos for each product in each week based on each week's predicted sales quantity. To get each re-scaled individual DueTo, we then divided the corresponding DueTo for each marketing vehicle by the sum of all DueTos in that week and multiplied it by the actual weekly sales quantity.

The weekly sales quantity of each Mahou San Miguel product was ultimately decomposed into the following components: Base (which includes Seasonality and Holidays), Shelf Price, Discount, Paid Search, Web Display, Email, Store Display, Flyer, TV, and Radio. Below are the results and visual representations of each product's DueTo decomposition after scaling them back to the original value of our sales quantity variable and appropriately debiasing the results.

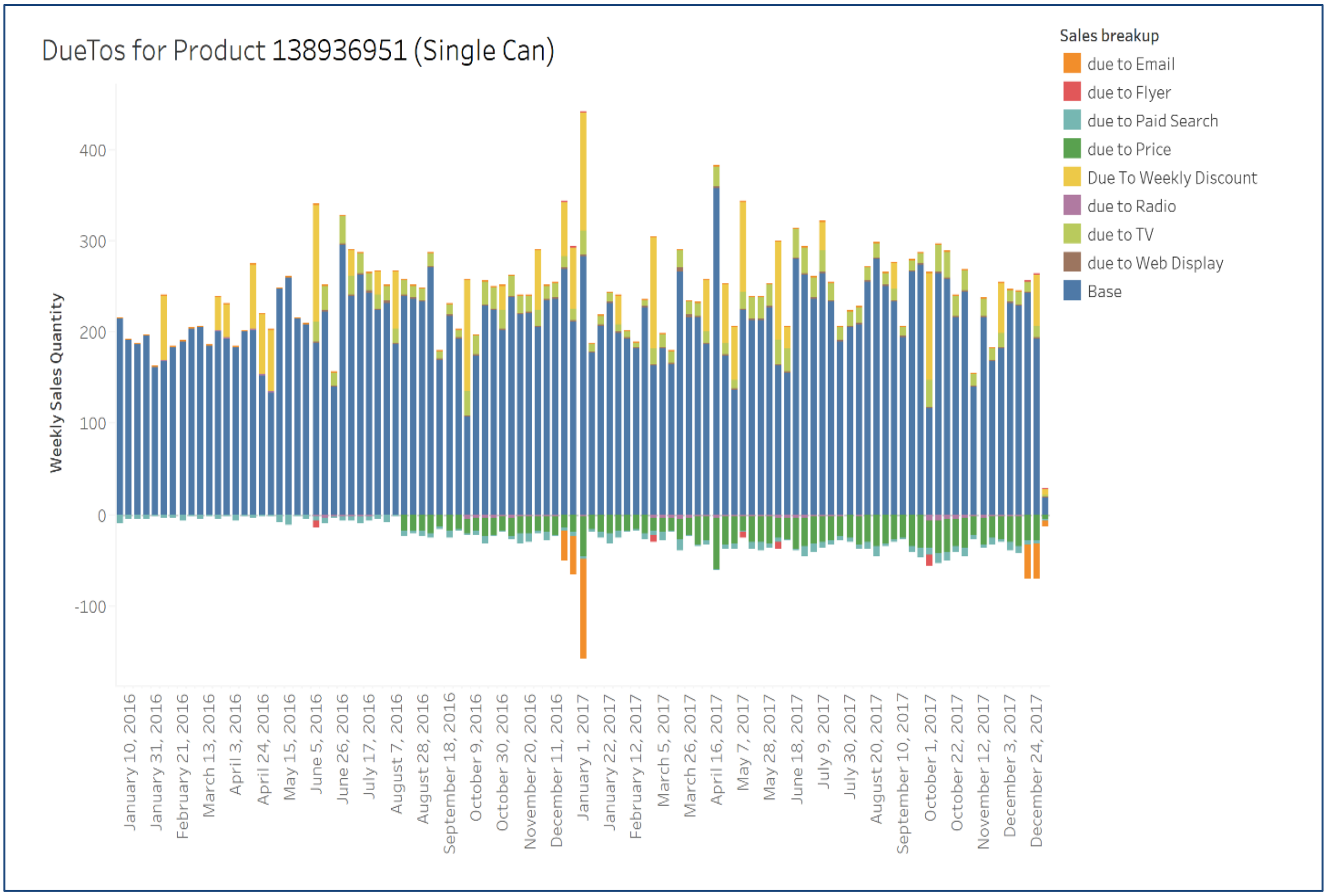

Figure 10: DueTo Decomposition for Single Cans

We can see from the DueTo decomposition of the Single Can product that the additional positive contributions of marketing vehicles to sales quantity can most heavily be attributed to discounts and TV, which both appear all throughout the year. In most weeks, discounts appear to add the biggest positive contribution to sales quantity. We also see positive contributions from Web Displays throughout most weeks.

Interestingly, we see what appears to be a negative effect from Emails in several of the holiday weeks and from price in most weeks throughout the history of our data. This indicates that Emails do not help in selling Mahou San Miguel Single Cans in Pernalonga's stores. When we compare this to other traditional marketing methods such as TV, we see that TV appears to have a positive effect on sales quantity for the Single Cans. This is something that we should be wary of as it is unlikely that a promotional event had a negative effect on sales. What is more likely is that sales quantity decreased at the same time that a promotional event was occurring, but it isn't necessarily accurate to conclude that the promotional event caused it. In future models and further analysis, Mahou San Miguel and Pernalonga should look to uncover additional underlying reasons for the decrease in sales quantity during these time periods.

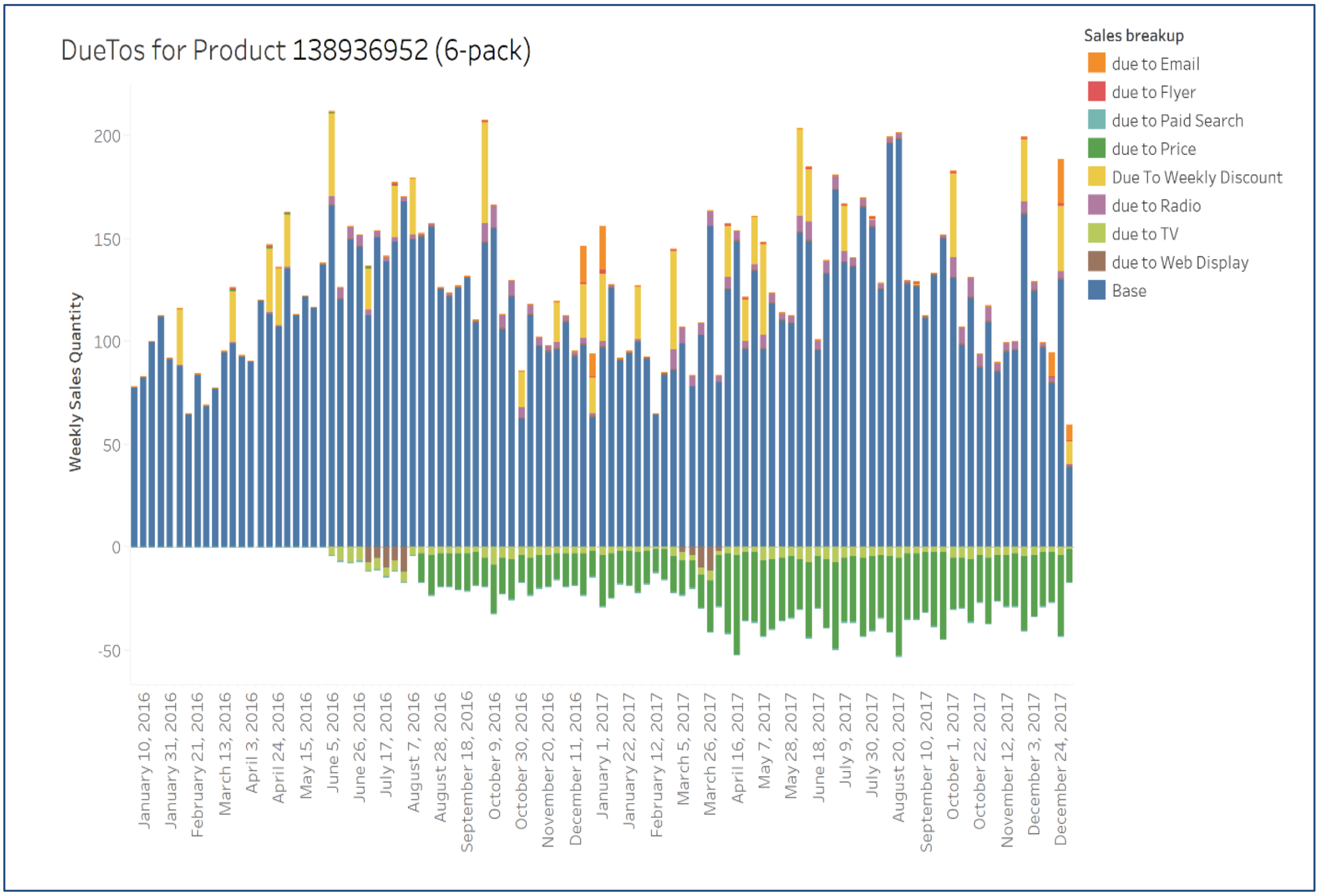

Figure 11: DueTo Decomposition for 6-packs

For Mahou San Miguel 6-packs, we see a different story than we did for Single Cans. We see that Email has high positive effects on sales quantity around the winter holidays, while the biggest negative effects appear to come from price, with TV advertisements negatively affecting sales quantity in several weeks as well. Web Displays also seem to have a negative effect on sales quantity. We also see positive effects from discounts and Radio advertisements at large proportions. We see the largest additions on top of the base come around the winter months, but surprisingly there are much fewer total sales in those months. The 6-packs appear to be much more popular during the warmer months, when people may be buying them to share with friends when there is good weather.

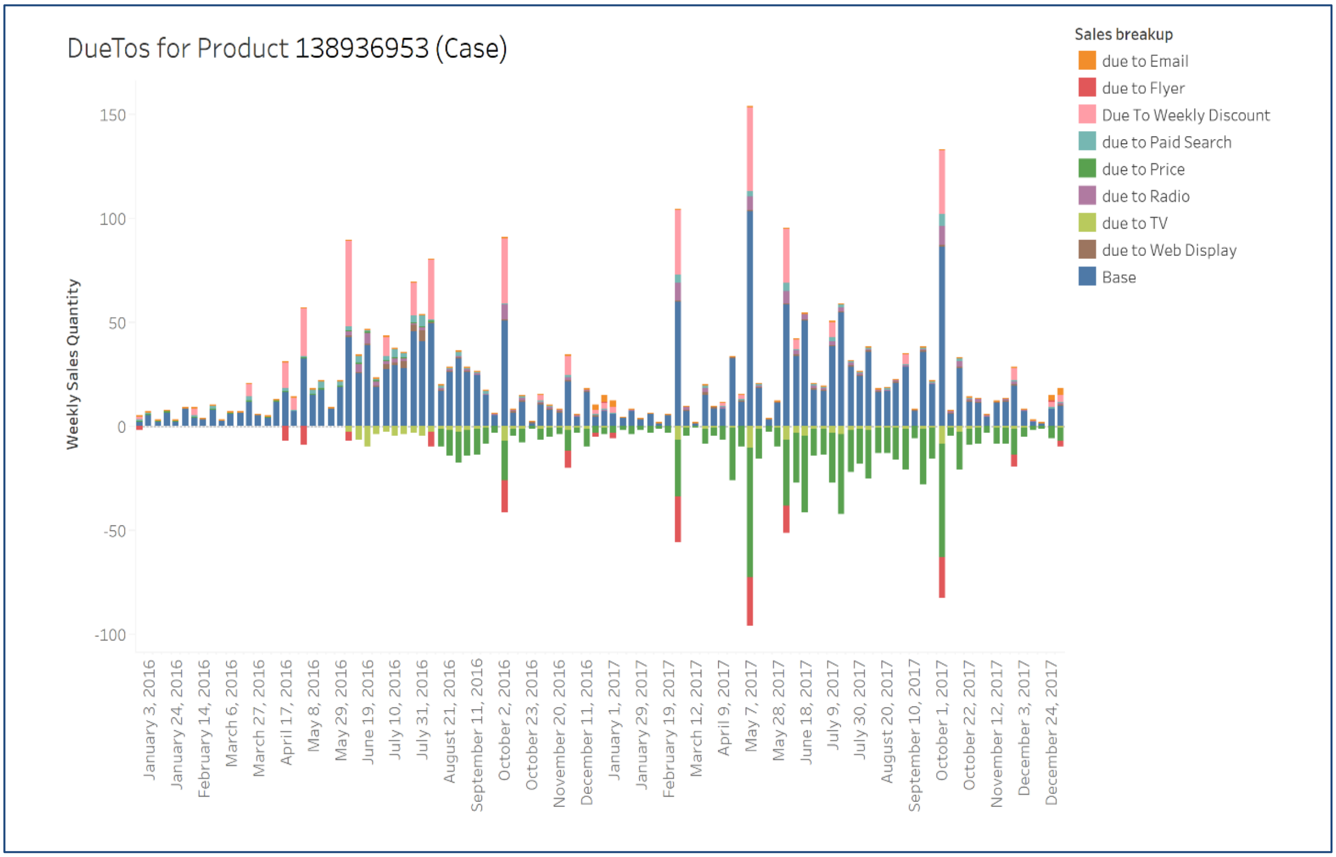

Figure 12: DueTo Decomposition for Cases

Lastly, we analyzed the DueTo decomposition for Mahou San Miguel's Case product. We see much lower contributions to sales quantity across all marketing channels, and we seem to see a more significant negative effect from Flyers. This product appears to have more of a seasonal sales pattern, with high points in the warmer months and lower points near the winter holidays. Like the other two products, we see what appears to be negative effects from price and large positive effects from discounts. Again, Radio appears to be positively affecting sales quantity, and TV appears to be negative.

Conclusion and Next Steps

Based on everything we've learned from the sales quantity decomposition process, we can conclude that there are indeed some Mahou San Miguel promotional campaigns that are very effective in driving sales. That said, we found that some promotional techniques do not have any reasonable impact on driving incremental sales. Thus, Mahou San Miguel should focus on the marketing vehicles that seem to drive the highest incremental sales for each product, as suggested by our DueTos analysis.

For both 6-packs and Cases, Emails have been very effective at driving incremental sales during winter, and it may be possible to consider using this media for a wider time period as well.

Similarly, discounts can have a great impact on incremental sales throughout the year, and we see this having positive effects for all three products. To improve reach, Radio media promotions should also be considered, as they have a positive effect for all three products, but a cost-benefit analysis may be helpful as a next-step in determining if they are profitable before making a decision to continue using this media. We strongly recommend discontinuing TV for both 6-packs and Cases, as they do not seem to help drive sales for these two products. For Single Cans, however, TV appears to be positively contributing to sales. Further, since Flyers did not seem to help drive sales, most prominently for Product 138936953 (Case), they should not be used as a marketing vehicle.

Finally, we recognize that promotional effectiveness can be somewhat dynamic. Thus, any implemented recommendations should be thoroughly tested to see if they are making the impact they are expected to make. One possibly approach could be A/B testing across similar stores, which would allow for Pernalonga and Mahou San Miguel to test the success of different marketing vehicles, while controlling for other factors such as price and discount. This would allow Mahou San Miguel to keep pace with using the marketing vehicles that drive the most incremental sales in the future.

Appendix

General Model Equation:

log(sales/theoretical maximum sales) ~ weekly_shelf_price + weekly_avg_disc + Flyer + Email + Web_Display + Paid_Search + Store_Display + TV + Radio + seas_index + holiday

138936951 (Single Can) Model Output:

```
Call:
lm(formula = formula, data = model_51)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.95232 -0.14068 -0.00246  0.15159  1.95232 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)   
(Intercept)          1.779e+00  1.066e+00   1.669  0.09904 . 
weekly_avg_disc      5.047e+00  1.844e+00   2.737  0.00765 **
weekly_shelf_price  -3.065e+00  1.454e+00  -2.107  0.03826 * 
Flyer               -1.164e-01  4.774e-01  -0.244  0.80804   
Web_Display          8.303e-08  5.592e-07   0.148  0.88234   
Paid_Search         -2.114e-06  3.563e-06  -0.593  0.55471   
Email               -2.167e-06  2.035e-06  -1.065  0.29020   
TV                   6.149e-01  9.053e-01   0.679  0.49899   
Radio               -9.075e-02  6.766e-01  -0.134  0.89364   
seas_index           7.180e-05  3.565e-05   2.014  0.04739 * 
holidayALLSAINTS     1.126e-01  3.321e-01   0.339  0.73545   
holidayASSUMPTION    8.156e-02  3.329e-01   0.245  0.80710   
holidayCARNIVAL     -3.587e-01  5.111e-01  -0.702  0.48485   
holidayCORPUS       -4.641e-01  3.419e-01  -1.358  0.17848   
holidayEASTER       -2.526e-01  3.412e-01  -0.740  0.46138   
holidayIMMACULATE    2.744e-01  3.324e-01   0.825  0.41160   
holidayLABOR        -8.913e-01  4.169e-01  -2.138  0.03561 * 
holidayLIBERTY      -5.583e-01  4.170e-01  -1.339  0.18453   
holidayNEWYEAR      -1.201e+00  4.437e-01  -2.706  0.00834 **
holidayPOPEVISIT     1.336e-01  7.351e-01   0.182  0.85621   
holidayPORTUGAL     -5.182e-01  6.649e-01  -0.779  0.43806   
holidayPrASSUMPTION -5.024e-01  3.533e-01  -1.422  0.15896   
holidayPrEASTER     -1.565e-01  3.752e-01  -0.417  0.67765   
holidayPrLIBERTY    -2.191e-01  3.577e-01  -0.613  0.54191   
holidayPrXMAS        6.483e-01  4.582e-01   1.415  0.16107   
holidayREPUBLIC     -1.181e+00  5.833e-01  -2.025  0.04621 * 
holidayRESTORATION  -1.144e-01  4.018e-01  -0.285  0.77653   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4429 on 79 degrees of freedom
Multiple R-squared:  0.4921,	Adjusted R-squared:  0.325 
F-statistic: 2.944 on 26 and 79 DF,  p-value: 0.0001233
```

138936952 (6-pack) Model Output:

```
Call:
lm(formula = formula, data = model_52)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.73045 -0.17004 -0.01219  0.18173  0.76165 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)          2.259e+00  7.726e-01   2.924  0.00452 ** 
weekly_shelf_price  -9.324e-01  1.851e-01  -5.038 2.97e-06 ***
weekly_avg_disc      3.519e+00  1.479e+00   2.379  0.01982 *  
Flyer                2.149e-02  1.529e-01   0.141  0.88857    
Email                1.312e-06  1.422e-06   0.923  0.35889    
Web_Display         -4.317e-07  4.241e-07  -1.018  0.31184    
Paid_Search         -1.840e-07  2.679e-06  -0.069  0.94541    
Store_Display       -8.051e-03  1.999e-01  -0.040  0.96797    
TV                  -2.276e-01  6.724e-01  -0.338  0.73596    
Radio                2.683e-01  4.957e-01   0.541  0.58985    
seas_index           1.582e-04  2.726e-05   5.803 1.33e-07 ***
holidayALLSAINTS    -3.630e-01  2.537e-01  -1.431  0.15643    
holidayASSUMPTION    4.818e-01  2.496e-01   1.930  0.05723 .  
holidayCARNIVAL     -5.841e-02  3.419e-01  -0.171  0.86477    
holidayCORPUS        9.835e-02  2.609e-01   0.377  0.70718    
holidayEASTER       -3.299e-01  2.507e-01  -1.316  0.19196    
holidayIMMACULATE    3.178e-01  2.488e-01   1.277  0.20535    
holidayLABOR         3.637e-01  2.898e-01   1.255  0.21312    
holidayLIBERTY      -1.416e-01  2.903e-01  -0.488  0.62711    
holidayNEWYEAR      -5.063e-01  3.397e-01  -1.490  0.14021    
holidayPOPEVISIT    -3.661e-01  4.252e-01  -0.861  0.39189    
holidayPORTUGAL      5.805e-01  3.851e-01   1.507  0.13580    
holidayPrASSUMPTION -9.875e-02  2.819e-01  -0.350  0.72708    
holidayPrEASTER      1.466e-01  2.868e-01   0.511  0.61075    
holidayPrLIBERTY    -1.801e-01  2.607e-01  -0.691  0.49175    
holidayPrXMAS       -1.180e-01  3.536e-01  -0.334  0.73950    
holidayREPUBLIC      4.780e-01  3.800e-01   1.258  0.21214    
holidayRESTORATION   4.943e-01  2.801e-01   1.764  0.08156 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3314 on 78 degrees of freedom
Multiple R-squared:  0.7518,	Adjusted R-squared:  0.6659 
F-statistic:  8.75 on 27 and 78 DF,  p-value: 2.517e-14
```

138936953 (Case) Model Output:

```
Call:
lm(formula = formula, data = model_53)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.45934 -0.31545  0.05925  0.36161  1.06924 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)          6.350e+00  1.587e+00   4.001 0.000148 ***
weekly_shelf_price  -7.581e-01  1.029e-01  -7.366 2.03e-10 ***
weekly_avg_disc      1.368e+01  4.410e+00   3.102 0.002717 ** 
Flyer               -6.329e-01  7.579e-01  -0.835 0.406359    
Email               -7.376e-07  3.438e-06  -0.215 0.830701    
Web_Display          5.384e-07  7.653e-07   0.703 0.483959    
Paid_Search          6.384e-06  4.931e-06   1.295 0.199437    
Store_Display       -9.401e-01  1.023e+00  -0.919 0.361046    
TV                  -6.347e-01  1.217e+00  -0.521 0.603583    
Radio                5.977e-01  9.153e-01   0.653 0.515814    
seas_index           3.719e-04  4.872e-05   7.634 6.35e-11 ***
holidayALLSAINTS     2.159e-01  4.508e-01   0.479 0.633424    
holidayASSUMPTION   -7.012e-01  4.368e-01  -1.605 0.112703    
holidayCARNIVAL     -1.187e+00  8.471e-01  -1.402 0.165157    
holidayCORPUS       -1.125e-01  4.593e-01  -0.245 0.807161    
holidayEASTER        2.231e-01  4.418e-01   0.505 0.615051    
holidayIMMACULATE   -9.697e-02  4.385e-01  -0.221 0.825600    
holidayLABOR         8.201e-01  7.783e-01   1.054 0.295486    
holidayLIBERTY      -1.107e+00  7.028e-01  -1.576 0.119343    
holidayNEWYEAR       3.164e-01  9.518e-01   0.332 0.740539    
holidayPOPEVISIT    -1.897e-01  1.141e+00  -0.166 0.868370    
holidayPORTUGAL     -4.604e-01  1.095e+00  -0.421 0.675289    
holidayPrASSUMPTION  5.966e-01  5.028e-01   1.187 0.239152    
holidayPrEASTER     -2.583e-01  5.959e-01  -0.434 0.665910    
holidayPrLIBERTY     2.429e-01  5.062e-01   0.480 0.632695    
holidayPrXMAS       -5.368e-01  8.398e-01  -0.639 0.524658    
holidayREPUBLIC     -9.244e-01  1.067e+00  -0.866 0.389214    
holidayRESTORATION   4.198e-01  6.588e-01   0.637 0.525878    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5804 on 74 degrees of freedom
Multiple R-squared:  0.8279,	Adjusted R-squared:  0.7651 
F-statistic: 13.18 on 27 and 74 DF,  p-value: < 2.2e-16
```