#### **Retail & Marketing Analytics, Individual Assignment**

#### Seongmin Lee (CID: 01247436)

# Introduction

The primary goal of this analysis is to advise a fragrances corporation the optimal budget allocation by measuring the price elasticity, clout, vulnerability, and marginal effect and advertising elasticity based on the panel data-driven analysis. There are several challenges to solve before modelling, such as endogeneity, missing data, and data transformation.

With the brief explanatory analysis, this report is divided into three parts: preprocessing, the measuring price elasticity of demand, and calculating marketing communications. After breaking through the problems in the dataset, we will build three models; one is related to the price elasticity, clout, and vulnerability. Another is for the marginal effect and advertising elasticity. Moreover, building an adstock model will help to understand the investment effectiveness.

# Data Preprocessing

When it comes to the missing data, there are several options that we can handle in filling. Considering that the focal brands do not have any missing values, we could ignore the other brands’ columns when we build a model. However, if we just remove those columns including NA, we might lose the information from the competitor brands to understand the total perfume market. Even thought imputation could contaminate the regression output, we filled the NA by using the R library called “imputeTS”, which is highly used to impute the missing values in the time series dataset. If there is a sales department or others related to the cologne market, we could ask them which of the imputation algorithm results seems more like the real market. Under the circumstances that we have no one to ask, among several algorithms in the “imputeTS” library, the “Kalman” method has been applied to the dataset.

Secondly, we need to add dummy variables to control other factors that the raw dataset does not show us by understanding consumer purchase journey. When it comes to the fragrance market, there is an inevitable tendency that the demand sharply increases during the holidays. This is because people usually go for a trip during the holidays, which leads them to the airport and visits the duty-free shops. Also, perfumes are one of the typical purchase products as a gift. Therefore, we have to add the seasonality or holiday dummy variables based on the investment, price, and sell-out quantity as below.

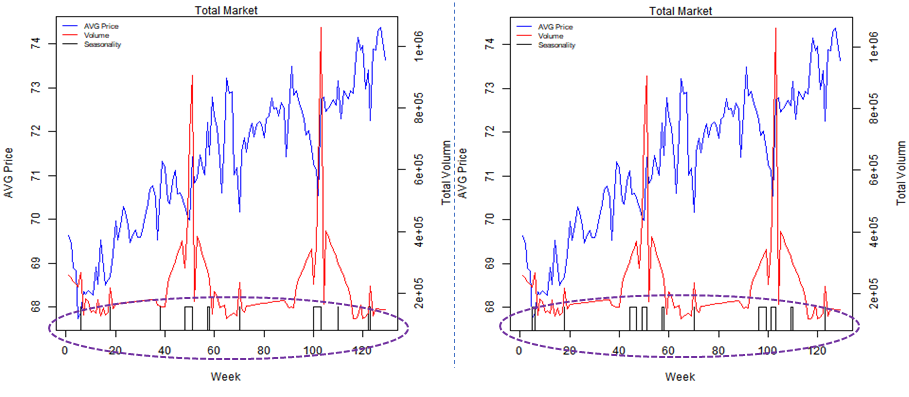


Figure 1. Total Volume and AVG. Price with Seasonality Dummy (Left: function, Right: manual)

The line plot on the left side demonstrates the weekly trend of price on average and total volume for two years and six months. In particular, the sale quantity coloured as a red line periodically soars during the Chrismas season. On top of that, the price on average plummeted during the same period, which implies that there was a promotional price event going on in the market. Those sudden changes in price and volume give a reason for building the seasonality dummy variable. In addition to the Christmas promotion, there must be other seasonal periods that we cannot explicitly recognise. So, at first, the seasonality dummy is one if the weekly investment value is larger than the one on average, and the weekly volume is greater than the one on average at the same time. Otherwise, the dummy has zero. As a result, the seasonality dummy has been fitted with the market demand and price as the black line of the graph on the left. To make week and date synchronised in each year, the value of the dummy variable has been manually filled, which results in the plot on the right. (Source: Table5 in the appendix)

# Explanatory Analysis

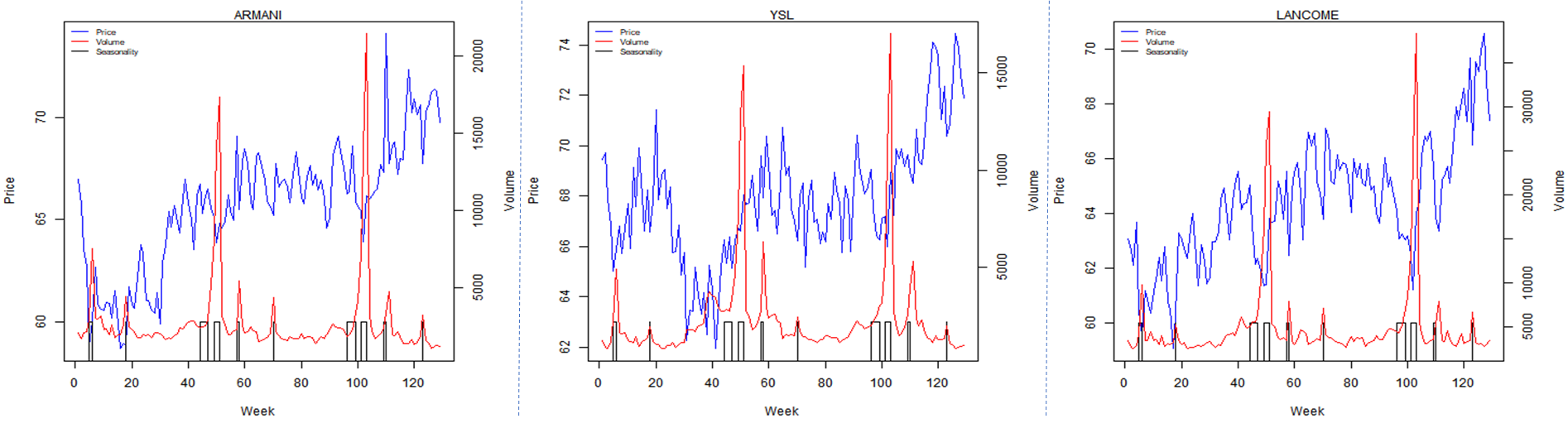
 [ARMANI] [YSL] [LANCOME]

Figure 2. Market Price and Volume for the focal brands

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Segmentation** | **Brand** | **AVG. Price** | **MAX Price** | **MIN Price** | **SD Price** | **AVG. Volume** |
| **$60 ~ $70** | **ARMANI** | **65.70** | **74.06** | **58.72** | **3.15** | **2,817** |
| **YSL** | **67.88** | **74.43** | **61.95** | **2.49** | **2,355** |
| **LANCOME** | **64.36** | **70.55** | **59.06** | **2.26** | **4,987** |
| DIOR P | 66.30 | 70.67 | 62.37 | 2.04 | 6,980 |
| **$70 ~ $80** | DIOR J | 74.67 | 80.89 | 69.98 | 2.21 | 6,971 |
| CHANEL NO5 | 80.85 | 86.41 | 73.85 | 2.70 | 5,512 |
| CHANEL COCO | 78.11 | 81.79 | 72.67 | 2.04 | 4,248 |
| NARCISO | 72.70 | 76.20 | 69.44 | 1.47 | 7260 |

As we saw the total market condition in the previous part, there was a sharp demand increase in a particular period with a price promotion. In this section, we have to take a look at the sales business for the three focal brands. The three charts above are conveying the same form of information on the movement of price and volume during the same period.

Above all, there are two points in common in the figure2. Firstly, the sell-out quantity for the focal brands during the second Christmas is higher than the one on the first Christmas, which might indicate that the investment for the latter Christmas was more effective than the one for the previous one. Moreover, the price is overall drawing the increasing curve for the two and half years, which could be the upselling brand marketing in a way that the retail prices for the perfume products of the three brands are reaching the seventy dollars groups such as Dior J and Chanel.

For ARMANI, the momentum of the volume change is highly similar to the total market trend in that each peak spot of the volume coloured as red is well matched up with the seasonality dummy variable coloured as black. However, the price promotion is not distinctively found during the holiday period. In particular, during the Christmas season, the price bargain is not as significant as the one on Valentine’s day in the beginning.

When it comes to YSL, around the 40th week, there was a price drop during the off-season, which could be regarded as an offensive strategy which is for occupying the volume from the other brands. Moreover, it is distinctive that the volume increase momentum of YSL during the peak-season is powerful than the one in the other brands, which might point out the higher price elasticity in YSL.

LANCOME has the usual strategy of the price promotion for Christmas in a way that the price of the product is sharply down, which is significantly lower than the other two brands. On top of that, regarding the size of the volume, LANCOME has the most substantial sales quantity compared to other ones, which means that LANCOME has the solid strategy based on the firm market position.

# Price Elasticity

The underlying goal of measuring the impact of prices is to understand how sensitive consumers are to a price change. In other words, the sales quantity could increase if we use a promotional price by investing. Furthermore, we can help the decision maker to optimise the price and promotion allocation at the end. Additionally, we can assess the competitive effects, which indicates how much a change of the brand price affects the other brands, which can be shown by vulnerability and clout.

For those focal brands, we need to set up regression models with different functional forms, such as linear, semi-log, and log-log. Among the three models, we need to pick a model with the best performance having the lowest root-mean-square error, which is the fitted value – true value due to the different dependent variable in each model. If the dependent variable is a form of the natural logarithm, then the root-mean-square error can be calculated as the natural logarithm to the power of the fitted value – true value. When we add the lag volume, the adjusted R-squared is higher, which improves the model performance. Therefore, the basic forms of the three equations are as following:

1. Linear

: Volume\_x ~ lag\_Volume\_x + Price\_x + Price\_y + Price\_z +

Invest\_x + Invest\_y + Invest\_z + seasonality

1. Semi-Log

: log\_Volume\_x ~ lag\_Volume\_x + Price\_x + Price\_y + Price\_z +

Invest\_x + Invest\_y + Invest\_z + seasonality

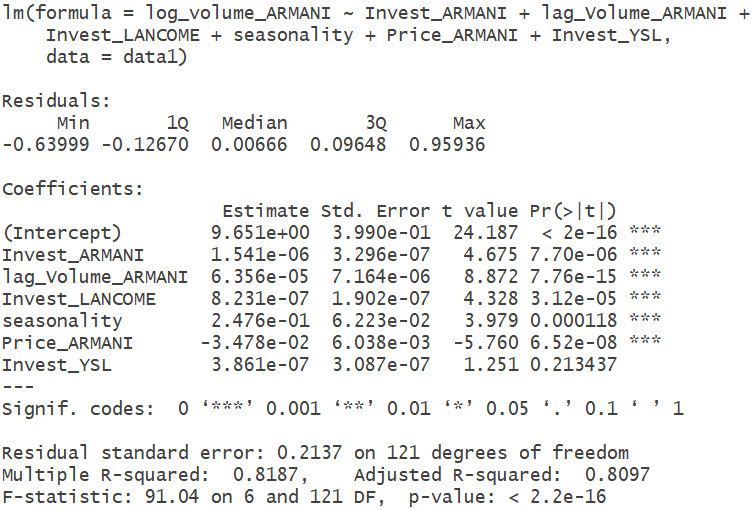
1. Log-Log

: log\_Volume\_x ~ lag\_Volume\_x + log\_Price\_x+ log\_Price\_y + log\_Price\_z +

Invest\_x + Invest\_y + Invest\_z + seasonality

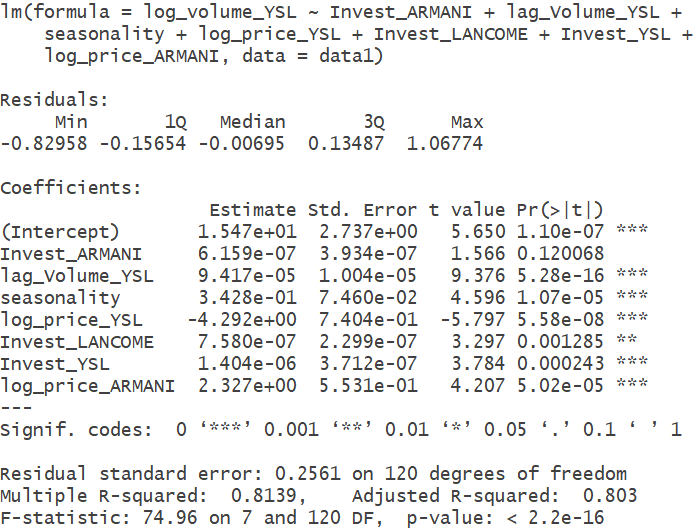
where x, y, and z are the three brands.

For the ARMANI, the equation formed in the semi-log has the lowest root-mean-square error. The estimated equation is :



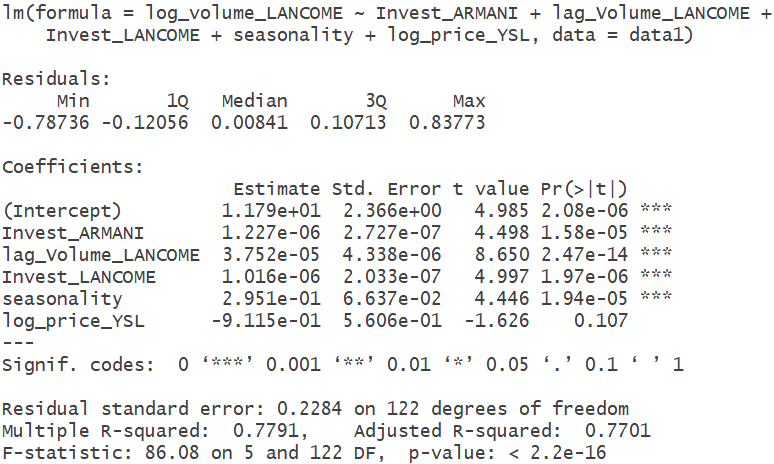
where n = 128, adjusted R-squared = 0.8097. As we can see the result, the price of ARMANI is statistically significant, which indicates that the increase of the unit price will decrease three percent of the sales quantity in ARMANI, holding other factors in the equation fixed. Moreover, the price changes by YSL and LANCOME does not affect the sales volume of ARMANI.

Secondly, when it comes to the YSL, the log-log form has the best goodness of fit among the three statistical models. The estimated equation is :



Where n = 128, adjusted R-squared = 0.8030. The both of log prices for YSL and ARMANI are statistically significant, which demonstrates that 1% of the price increase of YSL will drop 4.3% of sell-out quantity. On top of that, the 1% of the price increase of ARMANI will decrease the sales volume of YSL by 2.3%. And the change of LANCOME price has no effect on the YSL sales volume while controlling other factors.

Regarding the regression model for LANCOME, the log-log form performed the best. The estimated equation is :



where n = 128, adjusted R-squared = 0.7701. Compared to the previous models, none of the price variables is statistically significant, which indicates that the price change of LANCOME does not have any effect on brands volumes, vice versa. However, the sale volume of LANCOME might be affected by the other brands like Chanel, which needs adding the price variables for the other brands.

Based on the regression models above, we can come up with the price elasticity matrices, which helps to understand the demand movements by price changes.

Table 1. Price Elasticity Matrices

|  |  |  |  |
| --- | --- | --- | --- |
|  | ARMANI | YSL | LANCOME |
| ARMANI | -2.29 | 0 | 0 |
| YSL | 2.33 | -4.29 | 0 |
| LANCOME | 0 | 0 | 0 |

The figures in diagonal are called “Own-Price Elasticity”, which is an indicator of the sell-out quantity change to the price promotion in one brand. Because perfumes are a product group that the higher the price goes up, the lower the sales volume goes down, the sign of the “Own Price Elasticity” makes sense to be negative. For instance, the own price sensitivity for ARMANI is -2.29, which indicates that 1% increase of the price in ARMANI will decrease 2.29% of the ARMANI sales volumes. Regarding YSL, the sell-out volume will drop by 4.29% if the price of a YSL product increase by 1%. And the change of the price in LANCOME has nothing to do with the sales ramp-up. Furthermore, the absolute values of the ARMANI and YSL own price elasticities indicate that the sell-out volume is elastic. In other words, a small number of the price drop can result in the more substantial increase of the sales volumes, which consequently achieve higher revenue.

The elements in the off-diagonal are “Cross-Price Elasticity”, which is a factor to explain how much the price change of the brand in a row affects the volume of the brand in a column. As mentioned before, the sign of the elements should be positive from the general sales point of view. For example, 2.33 shows that 1% of the price increase concerning YSL would increase 2.33% of the volume increase in ARMANI.

In view of competitive effect, we can also calculate the clout and vulnerability as the table 3 shows below. The clout is measured by the column sum of the price elasticity matrices, which points out the total price effects of each brand on the competitors. In this case, ARMANI and LANCOME do not have much impact on the market. However, -4.29 indicates that the price change in YSL will have a more significant impact on the market compared to the other brands. On the other hand, when it comes to vulnerability, the focal brands are not easily affected by the price change of each other as the negative values explain. Therefore, we can conclude that from the competitive effect point of view, YSL has the most enormous impact on the market.

Table 2. Clout and Vulnerability

|  |  |  |
| --- | --- | --- |
| Brand | Clout | Vulnerability |
| ARMANI | 0.04 | -2.29 |
| YSL | -4.29 | -1.97 |
| LANCOME | 0 | 0 |

# Advertising elasticity

To calculate the communication impact for the three focal brands, we need to build the six different regression models: three models without carryover effects and three models with carryover effects. Among the six models, we need to select the highest adjusted R-squared one, which refers to the best goodness of fit. Before the regression analysis, we made sure that the time series dataset is stationary by ADF, PP, and KPSS tests. The functional forms are as follows:

* **Without Carryover**

1. Linear Response

: Volume\_x ~ Invest\_x + Invest\_x\_lag + seasonality

1. Concave Response

: Volume\_x ~ Invest\_x\_**log** + Invest\_x\_lag + seasonality

1. Concave-Quadratic Response

: Volume\_x ~ Invest\_x + **Invest\_x\_sqr** + Invest\_x\_lag + seasonality

* **With Carryover**

1. Linear Response

: Volume\_x ~ **Volume\_x\_lag** + Invest\_x + Invest\_x\_lag + seasonality

1. Concave Response

: Volume\_x ~ **Volume\_x\_lag** + Invest\_x\_**log** + Invest\_x\_lag + seasonality

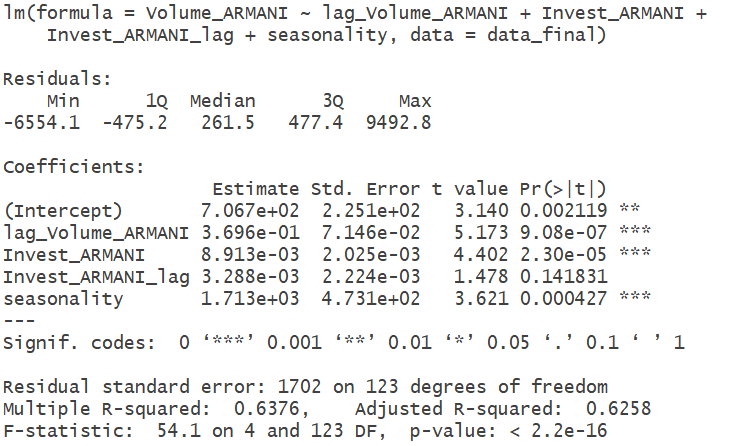
1. Concave-Quadratic Response

: Volume\_x ~ **Volume\_x\_lag** + Invest\_x + **Invest\_x\_sqr** + Invest\_x\_lag + seasonality

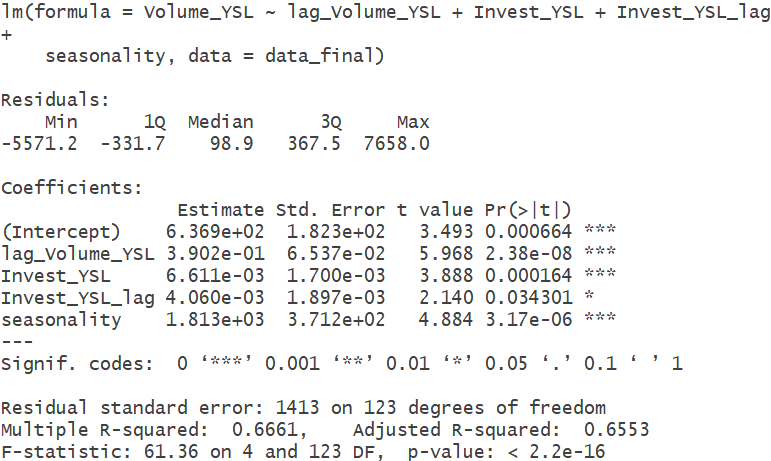
where x, y, and z are the three brands.

Based on the formulas above, we found out that the linear response model with carryover has the best performance for the three brands. The estimated equation for each brand is as follows:

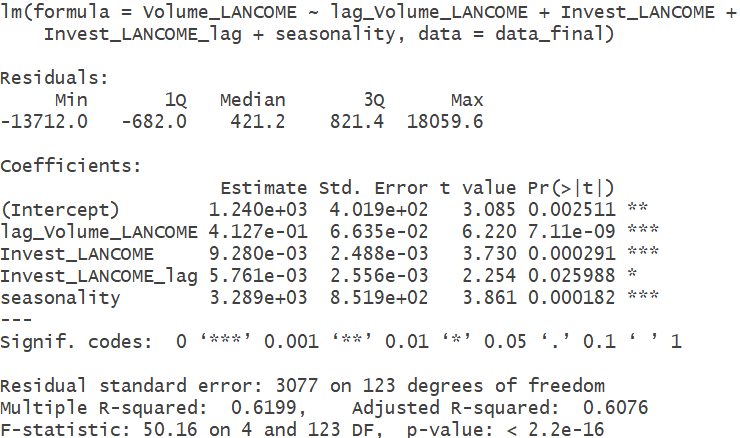
1. **ARMANI (Linear Response with Carryover)**



1. **YSL (Linear Response with Carryover)**



1. **LANCOME(Linear Response with Carryover)**



By using the previous regression models based on carryover effect, we can summarise the marginal effect and advertising elasticities as the table3.

Table 3. Marginal Effect and Advertising elasticities

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Brand | Marginal effect | | Ad. Elasticity | |
| Short Run | Long Run | Short Run | Long Run |
| ARMANI | 0.0089 | 0.0141 | 0.1971 | 0.3127 |
| YSL | 0.0066 | 0.0108 | 0.1252 | 0.2054 |
| LANCOME | 0.0093 | 0.0158 | 0.1363 | 0.2321 |

The advertising elasticities in the short run indicate that a 1% more investment in an advertisement by ARMANI, YSL, and LANCOME will impact sell-out quantity in the short-term in 0.20%, 0.13%, and 0.14%, and in the long-term 0.31%, 0.21%, and 0.23% respectively. It indicates that the advertising spends by ARMANI has the highest efficientness compared to the other brands.

In addition to the regression analysis, we could also use the adstock method to measure the level of carryover effect. Based on the linear response model with carryover, we can add the adstock variable. By changing the coefficient (alpha) of the adstock from 0.01 to 0.99 by 0.001, we run the regression model. Among the 981 models, we can select the right model with the highest adjusted R square. As a result, the alpha for each brand investment is as below:

Table 4. Adstock Model Result

|  |  |  |  |
| --- | --- | --- | --- |
|  | ARMANI | YSL | LANCOME |
| Alpha | 0.182 | 0.168 | 0.197 |

The alpha value indicates that how long consumers memorise any events formed by each brand. As the table4 shows, the three focal brands do not have much difference in view of the memory effect. In other words, after the focal brands executed a budget, only 17 ~ 20% of the memory would linger on the consumers.Therefore, from the communication impact point of view, the proportion of the investment should be more focused on ARMANI to increase the sale quantity.

# Conclusion and budget optimisation

In this report, we analysed the price sensitivity, competitive effects(clout and vulnerability), and communication impact for the focal three brands. By doing so, in view of the price change, we could understand that YSL has the most powerful to the other brands. Furthermore, LANCOME does not take or give influence to ARMANI and YSL. When it comes to the memory effect, all of the three brands have the similar level of the impression from consumers if an event promotion is executed. However, regarding the spending efficiency, ARMANI can increase the sell-out quantity more than the other two brands.

Therefore, ARMANI should pay more attention to the advertising spend than the price promotion. Regarding YSL, there should be offensive or defensive price strategy for the seasonality due to the higher price sensitivity. LANCOME can focus on the communication and price equally in that the brand does not have clout or vulnerability and has the high advertising elasticity.

Regarding the budget allocation for the rest of the year 2017, we added the monthly and the weekly dummy in the previous regression models. In the three models having the month dummy, the December was statistically significant, which indicates that the demand for perfumes indeed rises in comparison to January. To be specific, when we re-build the three models with the weekly dummy, the result is as the table6 in the appendix.

Concerning ARMANI, we recommend focussing on the weeks from 48th to 51st. Especially, increasing the communication impact and executing price promotion together will help to increase the sell-out quantity, compared to other weeks. When it comes to YSL, due to the higher price elasticity, the price promotion will be more critical than the advertising. Specifically, the strong price promotion will be effective on the 50th week in this year. Moreover, from 48th to 51st, the advertising spending has also a positive effect on the demand increase. Regarding LANCOME, in the week 48th, the price bargain will lead to an increase in the sell-out quantity. Furthermore, from the 49th to 51st, the price promotion and communication spending should be executed at the same time.

# Appendix

Table 5. Seasonality(Holiday) Dummies

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Week number** | **Total Volume** | **Holiday** |
| 2015-02-01 | 5 | 220,892 | Valentine's Day |
| 2015-02-08 | 6 | 268,636 |
| 2015-05-03 | 18 | 219,013 | Mother's Day |
| 2015-11-01 | 44 | 302,700 | Thanksgiving Day |
| 2015-11-08 | 45 | 325,771 |
| 2015-11-15 | 46 | 348,842 |
| 2015-11-22 | 47 | 371,913 |
| 2015-12-06 | 49 | 436,928 | Christmas |
| 2015-12-13 | 50 | 719,021 |
| 2015-12-20 | 51 | 905,847 |
| 2016-01-31 | 5 | 274,732 | Valentine's Day |
| 2016-02-07 | 6 | 246,199 |
| 2016-05-01 | 18 | 237,433 | Mother's Day |
| 2016-10-30 | 44 | 284,288 | Thanksgiving Day |
| 2016-11-06 | 45 | 304,713 |
| 2016-11-13 | 46 | 325,137 |
| 2016-11-20 | 47 | 345,562 |
| 2016-12-04 | 49 | 353,151 | Christmas |
| 2016-12-11 | 50 | 598,366 |
| 2016-12-18 | 51 | 1,058,889 |
| 2017-01-29 | 5 | 319,746 | Valentine's Day |
| 2017-02-05 | 6 | 298,665 |
| 2017-05-07 | 19 | 224,855 | Mother's Day |

Table 6. Regression model added "Week dummy variable"

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Brand** | **Regression**  **Model** | **Week (Month)** | **Coefficient** | **Spending** |
| ARMANI | Semi-Log | 48 (November) | 0.689 | Price Promotion |
| 49 (December) | 0.894 |
| 50 (December) | 1.493 |
| 51 (December) | 2.058 |
| Linear Response with Carryover | 48 (November) | 2,401 | Communication |
| 49 (December) | 3,590 |
| 50 (December) | 9,587 |
| 51 (December) | 14,430 |
| YSL | Log-Log | 50 (December) | 0.758 | Price Promotion |
| Linear Response with Carryover | 48 (November) | 1,655 | Communication |
| 49 (December) | 2,162 |
| 50 (December) | 6,204 |
| 51 (December) | 8,435 |
| LANCOME | Log-Log | 48 (November) | 0.709 | Price Promotion |
| 49 (December) | 0.616 |
| 50 (December) | 1.116 |
| 51 (December) | 1.337 |
| Linear Response with Carryover | 49 (December) | 4,331 | Communication |
| 50 (December) | 16,150 |
| 51 (December) | 25,310 |