Session – Summary

Market Mix Modeling

Market mix modeling involves the use of multiple regression techniques to help predict the optimal mix of marketing variables. Regression is based on a number of inputs (or independent variables) and how these relate to an outcome (or dependent variable) such as sales or profits. Once the model is built and validated, the input variables (advertising, promotion, etc.) can be manipulated to determine the net effect on a company's sales or profits.

What is market mix modelling?

There are various reasons for which accountability is needed in marketing:

1. To manage brands as a transformed marketplace
2. To be able to deliver better value through enhanced media and channel experience
3. To best allocate human and other resources for marketing activities
4. To be able to leverage cutting edge research tools and new sources of information to stay ahead of competition

Market mix modelling brings accountability to the spending and decision making in marketing. It uses various types of statistical models to model the relationship between the different categories of spending and their impact on the sales and revenue.

MMM helps the CMO solve four top questions:

- Performance driver analysis: Which KPIs drive the top line performance? Which among these drivers could be controlled by internal influence (e.g. pricing, promotion, advertising, loyalty offers, etc) and which are external to our control (e.g. competition, new disruptive technologies, industry trends, macro-economic and demographic policies, etc)?
- Impact analysis on MROI: What is the quantitative impact of each commercial levers on the outcome parameters – revenue, traffic, customer's perception or loyalty to the brand or company? This tells us how much a 1-unit increase in each marketing lever impacts the outcome parameters. For example, if you increase the frequency of TV commercial by 1%, what % incremental traffic would it bring?
- Trade-off between marketing levers: What is the trade-off among marketing levers? Since these impacts are not necessarily additive, what is the compound influence of multiple levers at a time? There could be levers which improve traffic, but, not necessarily revenue. What is the impact on revenue growth when advertising and promotions are done in isolation as compared to when done together?
- Optimising marketing spends: How can you best allocate the marketing budget to gain the highest outcome? How to allocate the budget between commercials and promotion? Once you know the commercial budget, how should you spend it among various media solutions such as TV, radio, print media or digital?

Impact of advertising on sales

There are seven major effects of advertising on sales or revenues:

1. **Current effect** of advertisement is the effect in sales or any other outcome parameters at the same time as when it is aired or exposed.
2. **Carry-over effect** of advertisement is the effect in sales or any other outcome parameters that follows even after the day of the advert exposure.
3. **Shape effect** of advertisement is the effect in sales with the increasing intensity of the advertising
4. **Competitive effect** of advertisement is the effect in sales or any other outcome parameters when your competitors also advertise similar products or services
5. **Dynamic effect** of advertisement is the effect in sales or any other outcome parameters that changes with time
6. **Content effect** of advertisement is the effect in sales or any other outcome parameters due to a change in the advertising content
7. **Media effect** of advertisement is the effect in sales or any other outcome parameters, due to change in media channels, e.g. TV, radio, newspapers, magazines, digital, social network, etc.

There are 3 types of dynamic effects as well:

1. **Wear-in effect:** When, in response to an advertisement, sale or traffic increases from one week to the next, even though advertising intensity remain same for all the weeks
2. **Wear-out effect**: When in response to an advertisement, sale or traffic declines from one week to the next, even though the advertising intensity remain same.
3. **Hysteresis effect** is the effect on Sales or Traffic, that persists even after the TVC been stopped airing or withdrawing the print advertising.

Impact of pricing/promotion on sales

There are 5 major pricing/promotion strategies that are used:

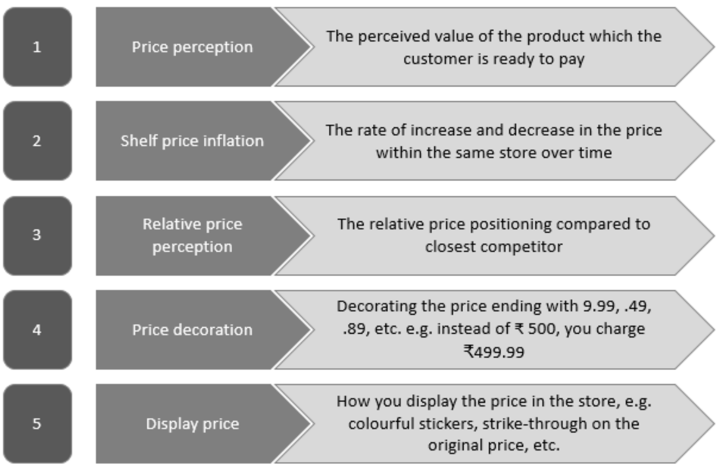

Figure: Pricing strategies

There are 10 common types of promotional pricing:

1. **Free shipping and free return** (e-commerce companies): The supply chain cost and the reverse supply chain costs are borne by the company.
2. **Flash sale:** These are sales of a very short duration. For example, Xiaomi smartphones are always sold through flash sales. Van Heusen holds a premier sale for a limited time period for their loyal customers when fresh stocks arrive.
3. **Branded freebies:** For example, towards the end of winter, Nescafé or Bru offers a coffee mug along with a pack of coffee with minimal extra cost. This is to ease out the price hike, and set a higher price for the next season.
4. **Buy more and save more:** Promotion such as 'Buy 2 get 10% off', 'Buy 3 get 15% off' and 'Buy 5 get 30% off' happen to clear stock.
5. **Loyalty points:** For example, you get points for each ₹100 spent and then you can redeem them after you have accumulated a certain number of points. This can be seen as 1% discount throughout. It creates a loyal customer base and increases repeat purchases.
6. **Organising an event or competition among customers**. For example, home furniture company IKEA asked their customers to click snaps of their IKEA furniture and post them on social media. The winner got a gift voucher. This increased the company's visibility with zero cost of advertising.
7. **Gift coupon:** Once you buy products of a certain amount, you receive a gift coupon worth 10% of the purchase which you can redeem during your next purchase. This once again ensures repeat purchase.
8. **Promised price match:** Big Bazaar declared that if someone can get the same product at a cheaper price anywhere, then it will give them double the amount in return. It was a bold step to create price-sensitive loyal customers.
9. **Holiday discounts:** For example, offering discounts on a certain holiday time, when customers tend to buy more. This creates a feel-good factor and creates a traffic push. Flipkart’s 'The Big Billion Day' is one such promotion.
10. **Discounted products:** The example from the previous page is the best example of discounted products. This happens mainly to clear old stocks or even clear slow moving/non-moving stocks.

How does product assortment impact revenues?

Any retail business can increase their sales and revenue by following these 3 strategies:

1. Sale more to the same customer: When a customer buys Pasta, you can’t sale more pasta to him. You need to have right pasta seasoning in your stock to cross sale
2. Increase the frequency of their visit: To make a customer visit more often, you must have some convenience and premium products which he will come for apart from regular visit. E.g. having a Liquor store along with super market improve weekend visit of customers.
3. Acquire new customers: To acquire new customer, you need to know what attracts them. If you have to target a different segment of customer, then keep products which would appeal them is important.

Thus, right selection of products is of vital importance. However, what is of even more importance is the granular tagging of the products:

- When you calculate the KPIs that represent pricing & promotion, you can't do that at each SKU level, nor can you do that at an overall store or broad category level
- At the SKU level, it will be too noisy, and strategic planning doesn't really happen at that granularity
- You have to get the right kind of tagging for each SKU to capture their thematic values

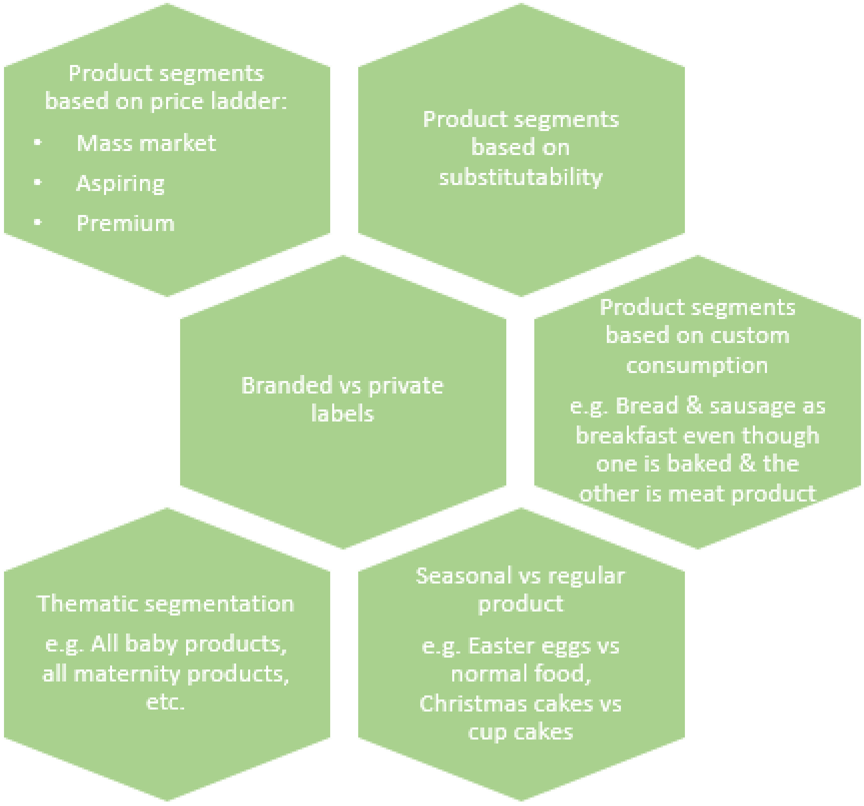

Figure 2: Granular tagging of the products

Modelling the impacts of KPIs

The basic linear model can capture the Current effect. This model can be estimated using Multivariate linear regression method.

$$Y = \alpha + \beta1At + \beta2Pt + \beta3Dt + \beta4Qt + \beta5Tt + \epsilon$$

Here, the different variables that are present are:

Y	Where Y is the dependent variable e.g. revenue, traffic, etc.	A	KPIs related to advertising
α	α is the constant term, which means, in the absence of other explanatory variables, α amount of sale happens, i.e. it's the organic sale without the influence of any marketing levers.	P	KPIs related to pricing
β_i	β_i are the coefficients to the explanatory variables which are most relevant	D	KPIs related to promotion or discounts
t	The subscript *t* indicates various time periods, measured as week or day	Q	KPIs related to product assortment or quality
ϵ	ϵ is the error term. This is the un-explained part of the model. This is generally a random noise and does not correlate with time or any other KPI. This is called "white noise" and follows a normal distribution.	T	KPIs related to industry trend, seasonality, etc.

Figure 3: Simple linear model

However, as you can see, this model assumes an additive relationship between the different KPIs. In such a case we use multiplicative model.

The name itself suggests that in this model, independent factors are multiplied together to get the marketing mix. This model includes the interaction of the independent variable through multiplication.

$$Y = e^{\alpha} * At^{\beta 1} * P_t^{\beta 2} * D_t^{\beta 3} * Q_t^{\beta 4} * T_t^{\beta 5} * \epsilon \qquad \text{Eq 2}$$

This model looks complex, but if you use Logarithmic transformation, it becomes a linear model again. The only difference is that the explanatory variables and the dependent variable (Y) are introduced to the model in their log form. In such a model, we do not really explain the revenue or traffic directly, but only their growth.

$$\ln Y = \alpha + \beta 1 \ln(At) + \beta 2 \ln(Pt) + \beta 3 \ln(Dt) + \beta 4 \ln(Qt) + \beta 5 \ln(Tt) + \epsilon' \qquad \text{Eq 3}$$

Once the equation is converted to linear form, then, a multivariate linear regression can be used to estimate the α and βi (i=1,2,...,5). ϵ' is the $\ln(\epsilon)$ the White Noise and assumed to be independently and identically normally distributed (or 'IID normal' in short).

Figure 4: Multiplicative model

The multiplicative model has three major **benefits**:

1. This model implies that there exists an interaction effect of the explanatory KPIs. For example, if your TV advertising has a 2x revenue impact, and your newsprint has a 1.2x revenue impact, it's not necessary that TVC + newsprint would have a 3.2X revenue impact. When done together, it may impact 2.5x vs 5x. This depends on the interaction effect which can be modelled through the multiplicative method.
2. The multiplicative model also implies that, based on the coefficients, the model can estimate a variety of shapes. It's not necessarily stuck into a linear format. Therefore, the model is more flexible to estimate the relationship between explanatory and dependent variables, as compared to Eq1.
3. Since the model does not estimate the actual revenue but its first derivative, which is the growth of revenue, the coefficients also reflect the 'elasticity' i.e. the rate of change of revenue with a change in advertising spending.

The multiplicative model also has three major **limitations**:

1. It fails to estimate the competitive, content, media and dynamic effects. Therefore, more advanced models are required, which we will learn about later.
2. Even though the multiplicative model has the flexibility to estimate various shapes, it fails to estimate the S-shaped curve, which is one of the most prevalent shapes in practice.
3. Since the multiplicative model is modelled at a first derivative level, it implies elasticity is constant, which is not always necessarily true.

In Exponential Attraction model, the market share of target brand is a function of its share of marketing effect in the total marketing effort done by the brands of similar category or brand value. This is called Kotler's fundamental theorem of marketing. This model, as a matter of it's design, captures the completion response in advertising also.

$$M_i = S_i / \Sigma_j S_j = A_i / \Sigma_j A_j \quad \text{Eq 4}$$

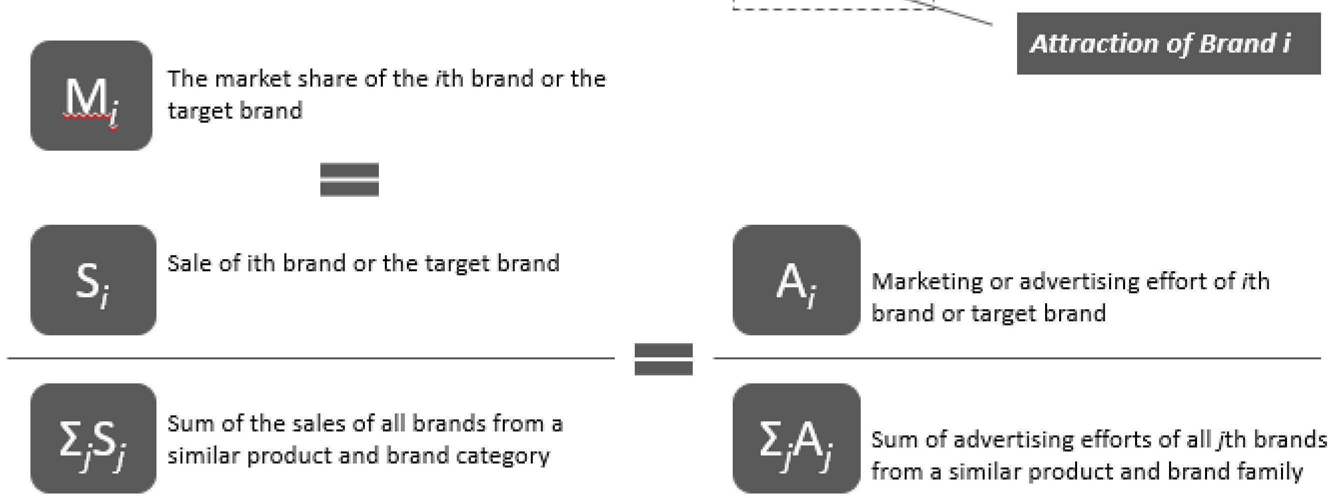

Figure 5: Attraction model

Now the attraction model is good with the linear responses but fails to capture the non-linear responsiveness. In such cases, we use exponential attraction model or multinomial logit model.

Till now, you have seen the models that capture the current effect of advertising. However, if you want to capture the carry-over effect, you would want to model the current sales or revenue figures on the past figures of the advertising spends and other KPIs. The models that can help you model these impacts are the Koyck and the distributed lag models.

The Koyck model is an extension of Basic linear model that we learnt at Eq1 with time lagged version of the dependent variable is also considered as an independent variable in estimating Sale, Revenue or Traffic. This can be thought as a combination of Autoregressive (AR) model & multivariate linear model

$$Y = \alpha + \mu Y_{t-1} + \beta 1At + \beta 2Pt + \beta 3Dt + \beta 4Qt + \beta 5Tt + \epsilon$$

Since we considered A_t as the Advertising effects, β1 is the current effect of advertising, while, carry over effect of advertising is $\beta1 * \mu/(1-\mu)$

- ⇒ Higher the value of μ, longer is the carry over effect of the advertising.
- ⇒ The smaller the value of μ, shorter is the carry over effect of the advertising, which means, sale depend more on current effect.
- ⇒ The total effect of advertising = Current effect + Carry over effect =

$$\beta1 + \beta1 * \mu/(1-\mu) = \beta1 /(1-\mu)$$

This model seems very simple, but mathematically, this model has many implications, especially in terms of the shape of carry over effect.

There are 2 major drawbacks of Koyck model. They are:

- Koyck model can capture the smooth decay immediately after the current effect jump, however, does not estimate any humped effects
- Estimating the carry-over effect of one explanatory variable, say, advertising is not very clear, since there are other independent variables too, which also may have some carry over effect. For example, immediately after heavy sale, there comes a deep a sale for next couple of weeks, which is called the effect of pantry loading. This can not be separated from carry over effect of advertising through Koyck model.

This therefore, gives birth of distributed lag model. In Distributed lag model not only dependent variable is entered in their lagged version, but also the independent variables.

$$\begin{gathered} Y = \alpha + \mu Y_{t-1} + \mu Y_{t-2} + \mu Y_{t-3} + \ldots \\ + \beta_1 A_t + \beta_1 A_{t-1} + \beta_1 A_{t-2} + \ldots \\ + \beta 2Pt + \beta_2 P_{t-1} + \beta_2 P_{t-2} + \ldots \\ + \beta 3Dt + \beta_3 D_{t-1} + \beta_3 D_{t-2} + \\ + \beta 4Qt + \beta_4 Q_{t-1} + \beta_4 Q_{t-2} + \ldots \\ + \beta 5Tt + \beta 5Tt\text{-}1 + \beta 5Tt\text{-}2 + \ldots \\ + \epsilon \end{gathered}$$

So far we have discussed multiple types of marketing mix models, however, none of them captures the last 4 advertising effects, i.e. Content, Media, Wear-in or Wear-out effects.

These can only be modeled by using Dummy KPIs or by using **Hierarchical modeling** technique

$$Y = \alpha + \beta 1At + \Lambda AtA'_t + \beta 2Pt + \beta 3Dt + \beta 4Qt + \beta 5Tt + \epsilon$$

Where, A'_t is a dummy variable.

A'_t =0 when the Advert1 is used at time t

$A'_t = 1$ when the Advert2 is used at time t

In this case, the main coefficient of advertising is β1 captures the effect of advertisement 1, while the interaction coefficient ʎ captures the effectiveness of advertisement 2.

Since in such model, the effectiveness is calculated in stages, are called as Hierarchical model.

Pricing KPIs

Suppose you are considering the Exponential attraction model or distributed lag model.

You have the following 3 data points for each SKU at a weekly level

1. MRP
2. List price
3. Discounted price

Calculate **Shelf price inflation** using the list price or base price. This will show week over week how the base price has changed.

$$\text{Shelf price inflation}_{\text{ith week}} = \text{List price}_{\text{ith week}} / \text{List price}_{\text{(i-1)th week}}$$

Calculate **% Discount offered** using the (a) list price and discounted price (b) MRP and discounted price. This is to check human behavior, whether they do the mental math against the MRP or the List price it was earlier. Either way, you will get the effect. Since these 2 KPIs will be highly correlated, ultimately, model will choose only one which will be of highest significance, and best explanatory factor.

$$\text{\% discount offered}_{\text{ith week}} = \text{Discounted price}_{\text{ith week}} / \text{List price}_{\text{ith week}} \text{ ------}\textit{(wrt List price)}$$

$$\text{\% discount offered}_{\text{ith week}} = \text{Discounted price}_{\text{ith week}} / \text{MRP}_{\text{ith week}} \text{ ------}\textit{(wrt MRP)}$$

- Reference price is a great way to measure the price perception customers has regarding that store or brand.
- This is generally done through Survey. At the starting of this concept, question like “Prices are high compared to what I feel” / “I feel Better prices are available elsewhere” and marking the level agreeing to these statements on a scale of 1-10.
- However, it was noticed, that most customers used to answer to please the researcher, without telling the actual fact. Thus the processes got changed.
- It was observed that there exists a strong correlation between historical prices vs the price perception among customers. Thus, instead if surveying the customers, today we use moving average of n previous weeks’ price for the same SKU as the reference price of the product.
- So, this is similar to Shelf price inflation, with a difference, that instead of comparing with previous day or previous week’s price, we are comparing with an average price of previous 2/3/4/6 weeks.
- Thus Shelf price inflation is a specific form of reference price.

Presenting results through visualisation

Once you have done the market mix modelling, next important step is to present your findings through visualisations. There are 3 common ways to represent your findings:

- Sales curve showing the inflection points at the time of advertisement campaigns
- Elasticity of each explanatory KPIs
- Breaking down the change in sales/revenues figure and explaining each component of it