Market Mix Modelling
For
ElecKart

Prepared by
Praneeth, Sriharsha, Raju & Sravan

Background

ElecKart is an e-commerce firm based out of Ontario, Canada specialising in electronic products. Over the last one year, they had spent a significant amount of money on marketing. Occasionally, they had also offered big-ticket promotions (similar to the Big Billion Day). They are about to create a marketing budget for the next year, which includes spending on commercials, online campaigns, and pricing & promotion strategies. The CFO feels that the money spent over the last 12 months on marketing was not sufficiently impactful, and, that they can either cut on the budget or reallocate it optimally across marketing levers to improve the revenue response.

Problem Statement

- How to optimise spending on commercials, online campaigns and pricing & special promotions?

Analytics Solution

To develop a market mix model to observe the actual impact of different marketing variables over the last year and built model with potential market driver to generate maximum impact in sales so the ElecKart can optimise the marketing budget in future.

Data Description

Order level data

- FSN ID: The unique identification of each SKU
- Order Date: Date on which the order was placed
- Order ID: The unique identification number of each order
- Order item ID: Suppose you order 2 different products under the same order, it generates 2 different order Item IDs under the same order ID; orders are tracked by the Order Item ID.
- GMV: Gross Merchandise Value or Revenue
- Units: Number of units of the specific product sold
- Order payment type: How the order was paid – prepaid or cash on delivery
- SLA: Number of days it typically takes to deliver the product
- Cust id: Unique identification of a customer
- Product MRP: Maximum retail price of the product
- Product procurement SLA: Time typically taken to procure the product

Data Description cont….

Other data

- Monthly spend on various advertising channels
- Days when there was any special sale
- Monthly NPS score – this may work as a proxy to ‘voice of the customer’
- Stock Index of the company on a monthly basis

As per the provided instructions, data provided is from July 2015 to June 2016.

Solution Approach

1. Understanding Client Requirement
   - Understanding client business
   - Problem statement

2. Data Understanding
   - Describe data
   - Data quality check
   - Data cleaning

3. Data Preparation
   - Select relevant data
   - Integrate data
   - Format data

4. Model Building
   - Linear model (for mid submission)

Data Cleaning

- Data type of columns looked at and modified as per the need.
- Presence of missing values were looked at and were ignored.
- Dropped rows of column ‘gmv’ or column ‘product_mrp’ value as ‘0’.
- Relevant columns were picked up ignoring the outliers.
- After the basic cleaning process of sales, media, NPS score data, their missing values assessment was done.
- The entire data was looked at from weekly perspective.

Data Preparation

- Identified ‘gmv’ column with null values and replaced with ‘0’.
- Created new column ‘discount’ by subtracting ‘gmv’ from ‘product_mrp’.
- Created new column ‘discount_percentage’ using existing columns ‘product_mrp’ and ‘discount’.
- Filtered out dates which are not in July 2015 and June 2016.
- Converted order id and order item id to integer.
- Dropped columns 'customer ID' and 'PIN' as these columns are not used in analysis.
- Converted monthly data in to weekly.

Data Preparation cont ….

- Created new column ‘special_day_flag’ based on provided ‘Special Sale Calendar’.
- Created new column ‘pay_flag’ based on Payment Dates of every month.
- Converted ‘Marketing Spends’ to rupees which is provided in Crs.
- Prepared final data frame by merging ‘Media Investment’, ‘Monthly NPS Score’, ‘Product List’ and ‘ConsumerElectronics’.

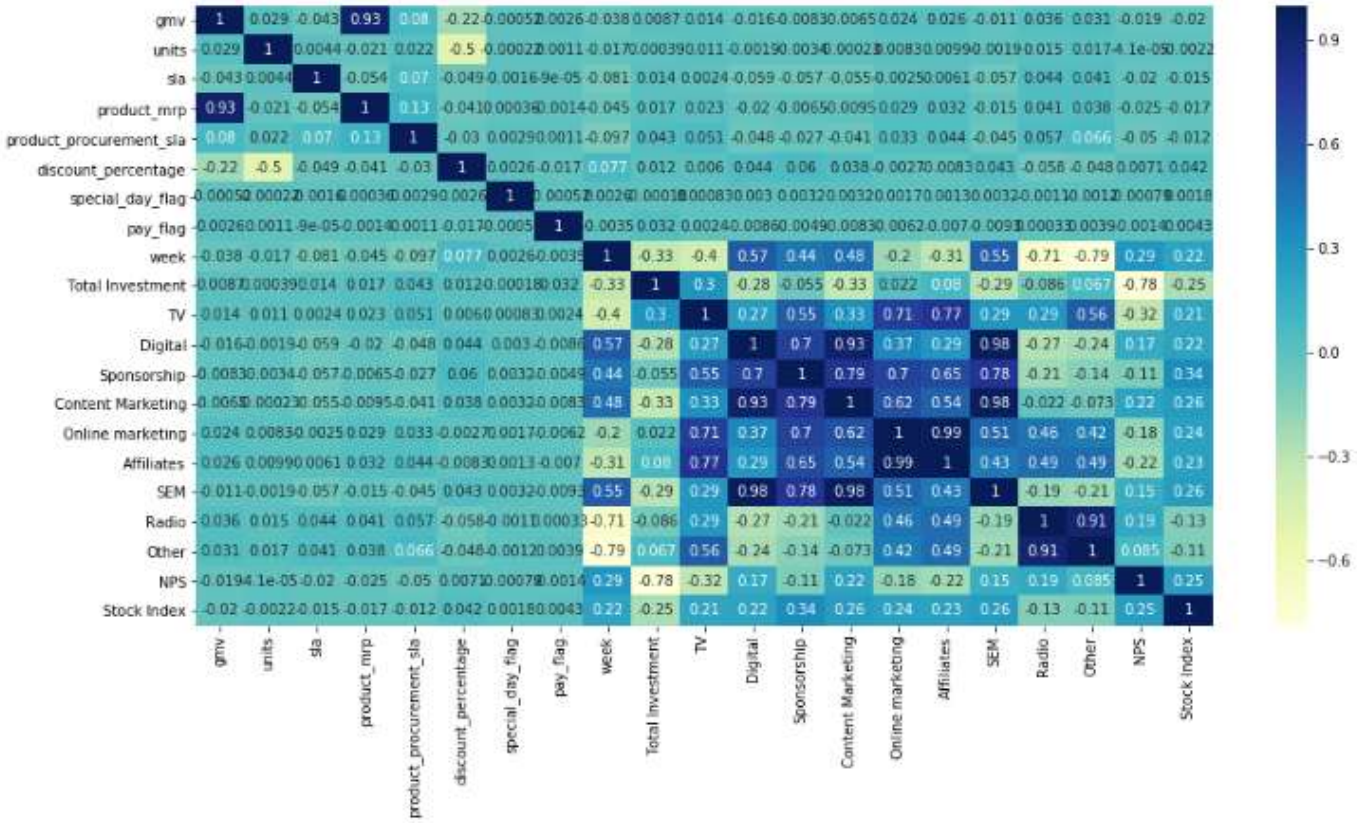
Unit Scaling

- GMV, Product_MRP are in terms of INR, while marketing spend is in INR Cr. For better model explanation converted the marketing spend to INR.

KPIs

- week
- gmv
- sla
- units
- product_mrp
- product_procurement_sla
- discount_percentage
- special_day_flag
- pay_flag
- TV
- Digital
- Sponsorship
- Content Marketing
- Online marketing
- Affiliates
- SEM
- Radio
- Others
- NPS
- Stock Index

Correlation of different market levers



Linear Model Metrics

- In the first iteration of Linear model the Adjusted R^2 is 48% and RSE came out to be 1058000 which implies that the Adjusted R^2 is moderate but error is too high.

Future Roadmap

- As the Linear model yielded the Adjusted R^2 to be 48% with high RSE, other models like Multiplicative, Koyck, Distributive and Distributive multiplicative needs to be carried out to Evaluate the best model with high Adjusted R^2 and low RSE

Thank you..