**Problem Statement**

Play Video

**Please note that there are TWO submissions for the capstone project (Mid Submission and Final Submission), with different deadlines. Please check the details for the two submissions in the subsequent pages. There will be Viva that will be conducted for the Mid-Submission.**

**Background  - Business Understanding**

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.

Imagine that you are a part of the marketing team working on budget optimisation. You need to develop a market mix model to observe the actual impact of different marketing variables over the last year. Using your understanding of the model, you have to recommend the optimal budget allocation for different marketing levers for the next year.

You can download the dataset from this[link.](https://drive.google.com/drive/folders/1F5O1thcC43zAPxsWIw1WQMh58Qi5I5CB?usp=sharing)

**Data Understanding**

You have to use the data from July 2015 to June 2016. The data consists of the following types of information:

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

Apart from this, the following information is also available:

* 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

You can find additional product details from this [link](https://drive.google.com/drive/folders/1F5O1thcC43zAPxsWIw1WQMh58Qi5I5CB?usp=sharing).

**Data Preparation**

You have to create market mix models for three product subcategories  - camera accessory, home audio and gaming accessory. Also, the models have to be built at a weekly level for each of the sub-categories.

**Feature Engineering**

As the E-commerce company is situated in the Ontario region. Here, we will include its climate data to analyse if it has any effect on the revenue. Create as many features as possible with the available data. Macro-economic data like climatic information on a daily basis can be extracted from the [link](https://drive.google.com/drive/folders/1La-ABrzvMJ_BtV_-wVAbkimerVz9_xuk?usp=sharing).

Create separate columns if Pay date (if 1st or 15th of the month) and holidays by creating a flag as a 0 or 1. For eg. if Holiday is there, the value will be 1.

**Exploratory Data Analysis**

Perform Univariates, Bivariates, Correlation, Cross-tabs and visualizations on the pre-analytical dataset. Obtain insights and shortlist variables for modelling. Create the KPIs as required for modelling and write appropriate comments for choosing those KPIs.

**Modelling**

Build basic, logarithmic and multiplicative models or any other model that are possible with the available dataset. Analyse the impact of various attributes on the target variable through appropriate metrics.

**Presentation of Results**

Choose the best ones for each of the three product sub-categories and explain them to the CMO,  CFO etc.

Create a powerpoint presentation explaining the best models. The audience should be able to intuitively understand the variables and their importance in predicting future sales. Point out any surprising or unexpected trends you notice.