ElecKart
Market Mixed Modelling

A CAPSTONE PROJECT SUBMITTED BY:

ABHINAV CHOUDHARY

OCT 11TH 2021

AGENDA

- Business Objective
- Action Plan
- Data Preparation
- Feature Engineering – KPIs Derived
- Exploratory Data Analysis - Insights
- Model Building & Evaluation
- Recommendations & Key Takeaways

Business Objective

To build a **Market Mix Model** for **ElecKart** (an e-commerce firm based out of Ontario, Canada) for 3 product sub-categories - **Camera Accessory, Gaming Accessory and Home Audio**

To observe the actual impact of various marketing variables over one year (July 2015 to June 2016) and recommend the optimal budget allocation for different marketing levers for the next year.

Business Objective - Action Plan

Performance Driver Analysis

KPIs impacting the top-line performance

Impact Analysis on Marketing ROI

Quantitative impact of each channel on revenue

Optimizing Marketing Spends

Recommendation to allocate the marketing budget to gain the highest outcome

4Ps of Marketing

We have the following datasets for analysis:

- Main Consumer dataset with order details at a daily level.

- Media Investment dataset with amount invested in each advertising medium for the past year.

- Sale Calendar dataset showing dates from past year when there was a promotional offer.

- NPS dataset showing net promotion score and company stock value for last year.

- Weather dataset having detailed weather reports from last year in the state of Ontario, Canada.

Price
- GMV
- Product MRP

Place
- Pin code
- Order Payment Type
- Week of the year seasonality
- Holiday/Events
- Sale calendar

Promotion
- Marketing Channel Investments
- Customer sentiment (NPS)
- Adstock

Product
- # of units sold
- Delivery days and SLAs
- Categories/subcategories
- Verticals Procurement SLA

Data Preparation

Incorrect Values in some columns
- Imputing NaN values in deliverybdays & deliverycdays with zero.
- Converting -ve values of some columns like deliverybdays & deliverycdays into absolute values

De-duplication of Data
- There were some duplicate entries which we decided to delete

Dropping Insignificant columns
- Removing columns with Single Unique Value (as it doesn't add any information to the analysis)
- Removing some of the 'Id' Columns which are insignificant to the analysis

Outlier Treatment
- Dropped the records where SLA is more than 30 days
- Performed outlier capping of GMV to 95 percentile

Limiting the data for the year
- As specified we deleted some rows outside the specified timeframe.
- We selected the data from July, 2015 to June, 2016

Aggregating the dataframes
- Extracting 3 separate data frames for 3 product subcategories camera accessory, home audio and gaming accessory
- Rolled up daily Order Data to Weekly Level by aggregating the numeric variables based on Week#

Preparing the data for modelling
- Scaling the dataframe & applied log transformation wherever required

Feature Engineering – KPIs Derived

Discount %
- Discount % = (product_mrp - selling_price)*100 / product_mrp

List Price
- List Price = GMV (Gross Merchandise Value) * Units

Payday Week
- If Payday falls within the week, then payday week = 1, else 0

Holiday Week
- If Holiday falls within the week, then payday week = 1,else 0

Product Premiumness
- If GMV value is greater than 75 percentile, then luxury, else mass-market

Week #
- Generate Week # column from the order date

AdStock Features
- Calculate Ad Stock values for all Advertising media - Affiliates, SEM, Radio, Content Marketing, Online Marketing

Lag Variables
- Lag variables(lag by 1, 2 & 3 days) for all KPIs can be taken for Distributive Lag Models

Exploratory Data Analysis – Insights

With respect to the count of orders, it has been observed that the Camera Accessory stands at the top followed by Gaming Accessory.

Whereas, in terms of revenue(GMV), Home Audio stands at the top followed by Camera Accessory
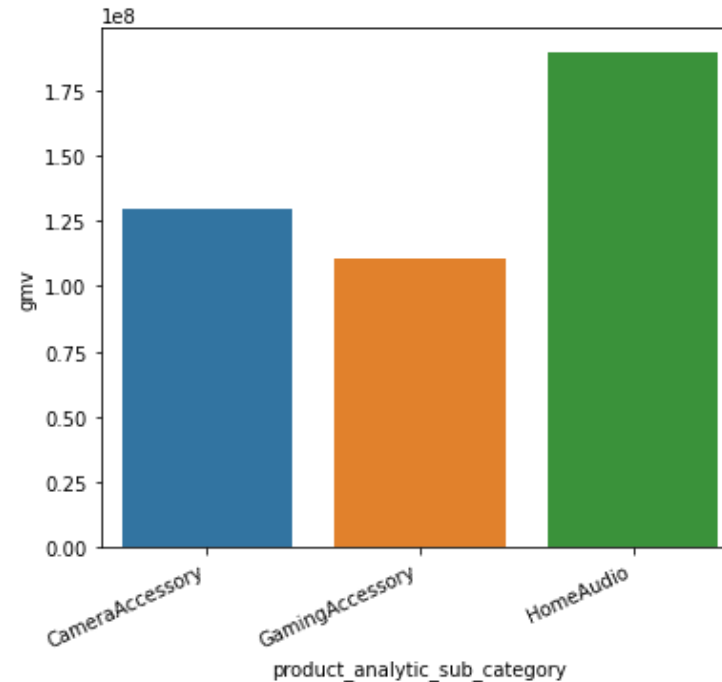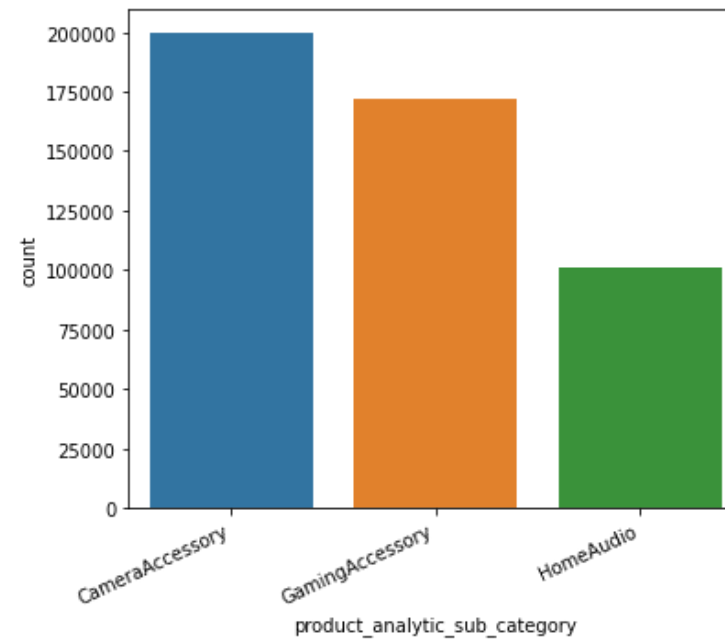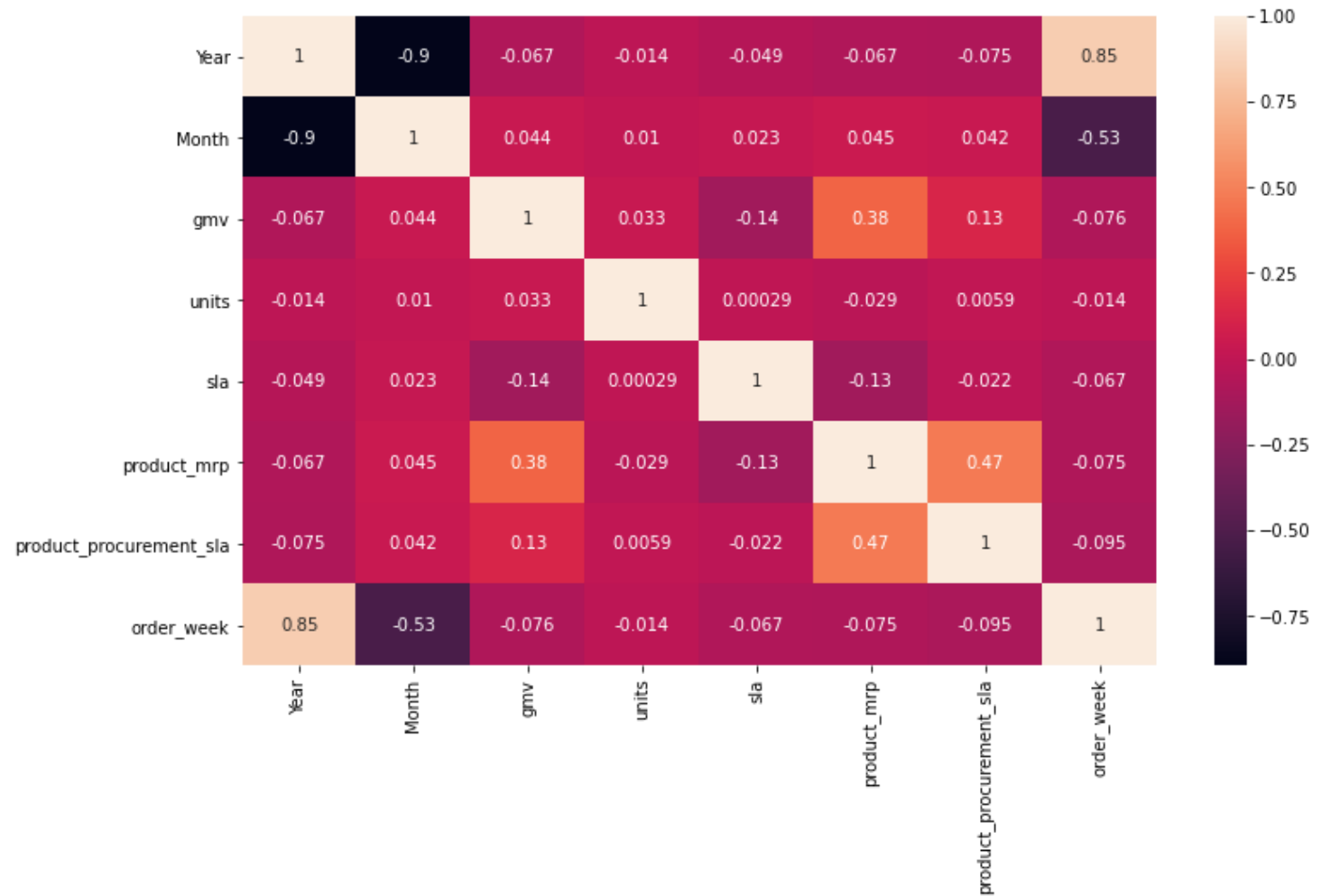
With the help of **Correlation Matrix**, we had observed that there is a high correlation between:
- Month & Year
- order week & Year

There is also a moderate correlation between:
- product mrp & product procurement sla
- GMV & product mrp

In terms of revenue, it was observed that the highest revenue was generated by the COD orders. And within COD orders, Home Audio followed by Camera Accessory had the higher revenue.

In the next observation it was found that the highest count of orders was generated from the Camera Accessory followed by Gaming Accessory for both the payment types.
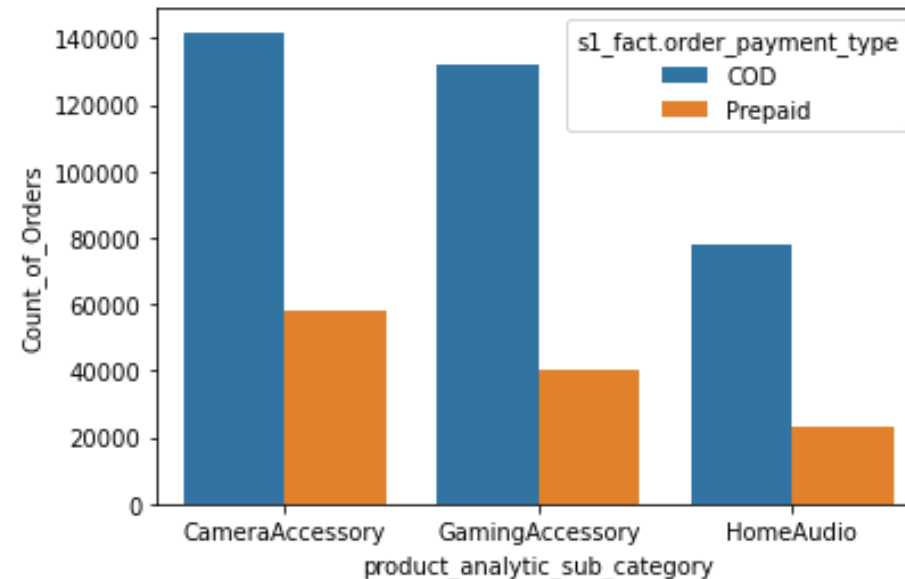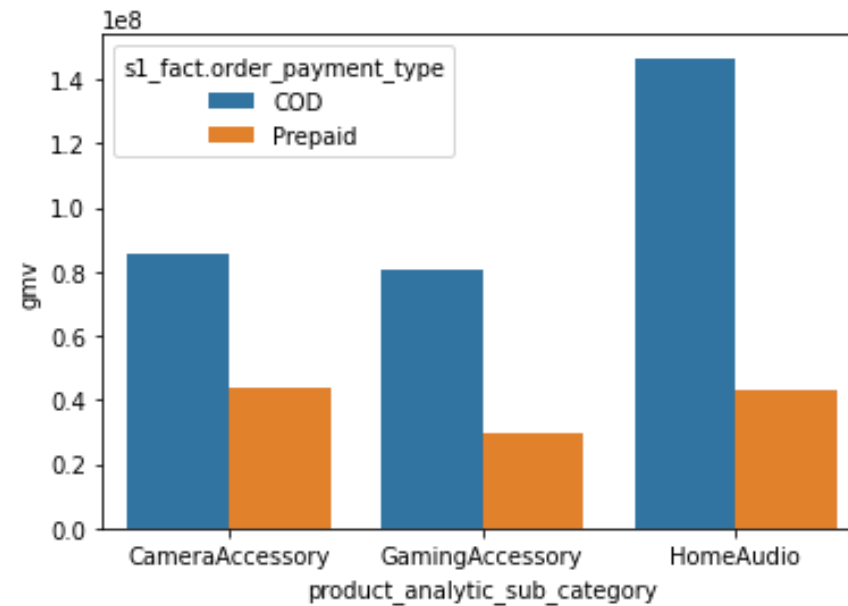
By analyzing the weekly spend over each channel, it was observed that between the weeks 39 to 45 for most of the channels, maximum amount of investments were done & that was done mainly through the Sponsorship Media.
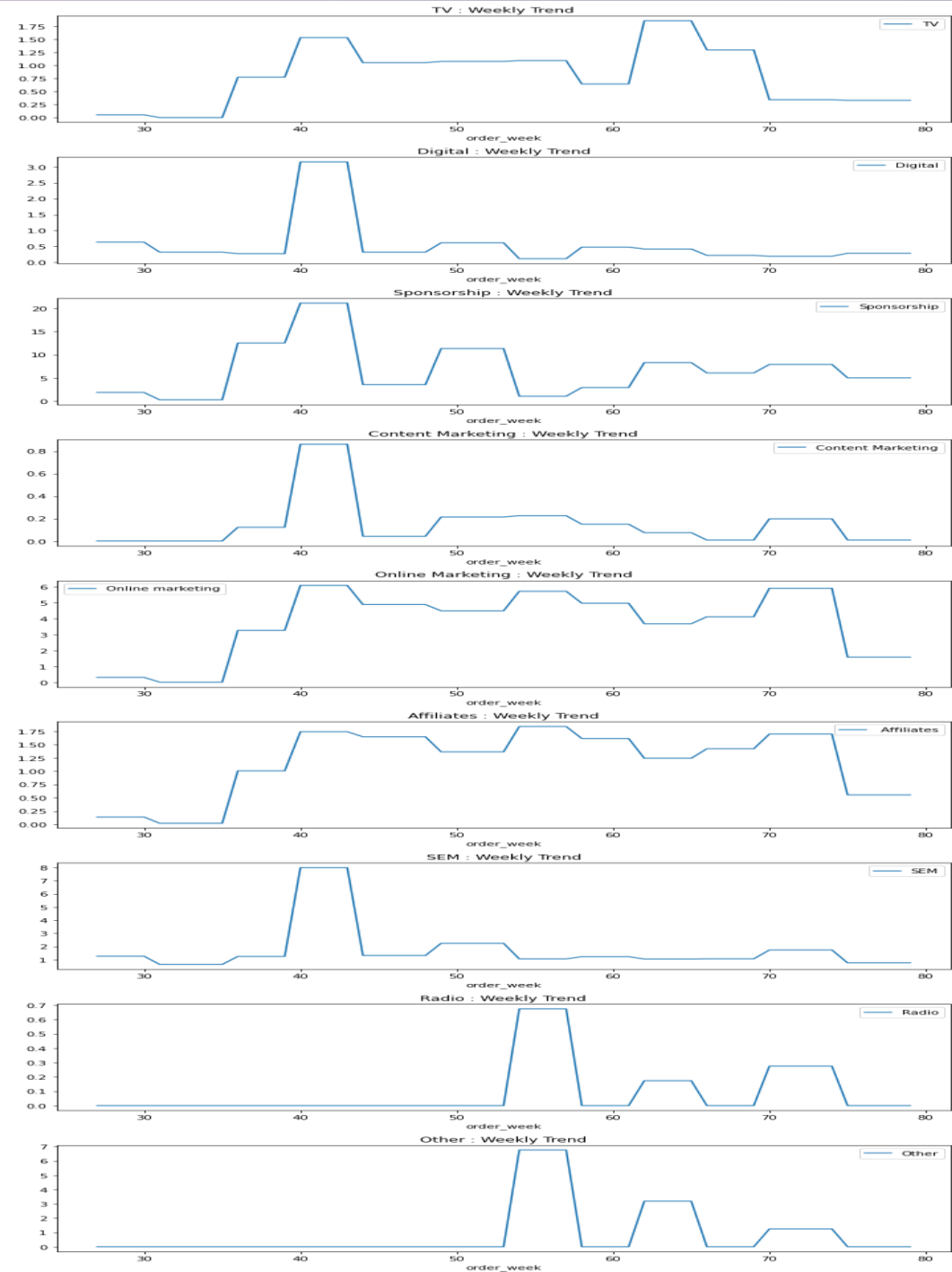
Out of the derived KPIs, with the help of Correlation Matrix it was observed that there is a high correlation between:

- adStock_Digital & adStock_SEM

- adStock_Content Marketing & adStock_SEM

- Month & adstock_radio; Month & adstock_others

- order_week & month; year & month

- adStock_Affiliates & adStock_Online marketing

- adStock_Other & adStock_Radio
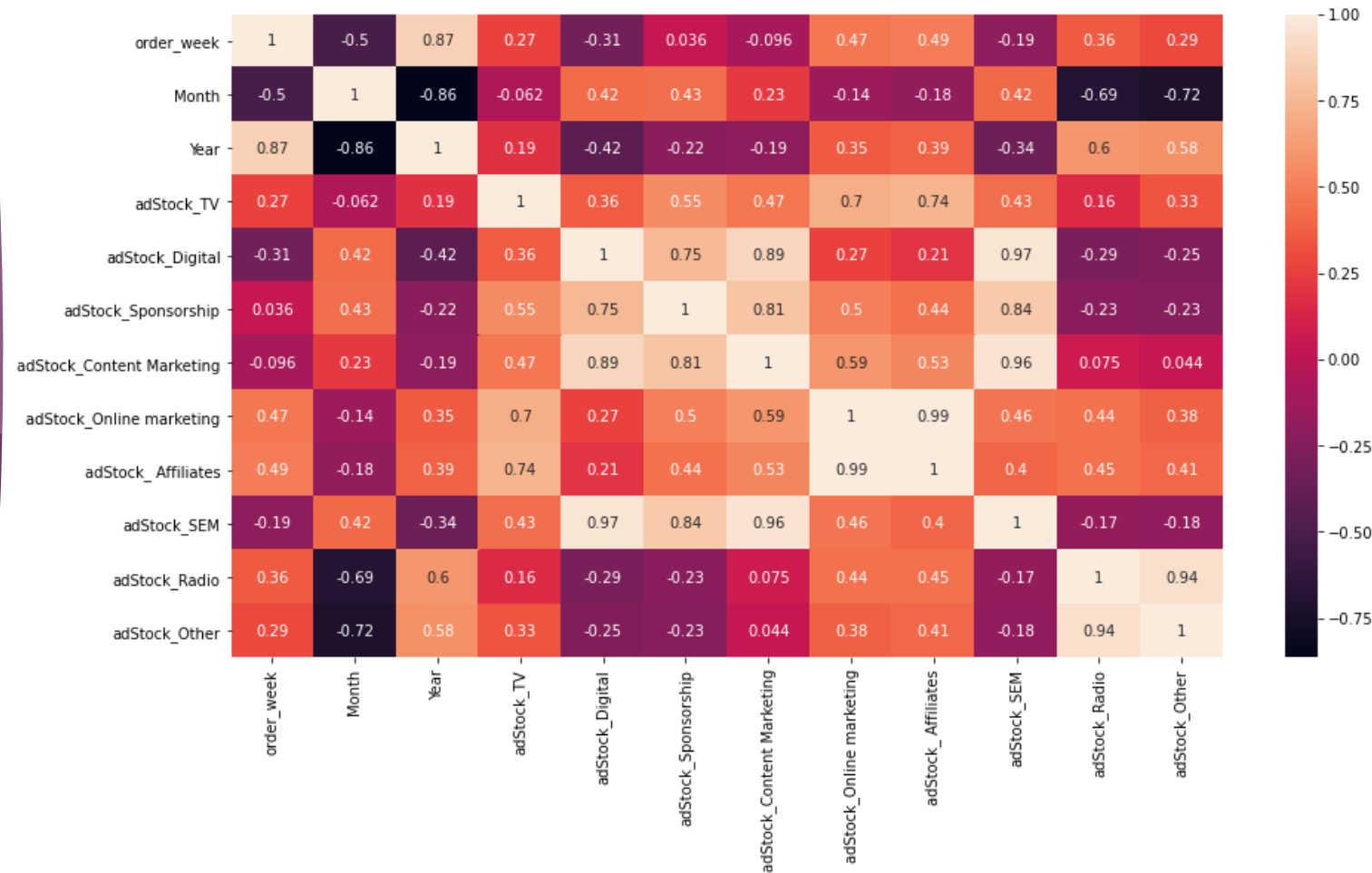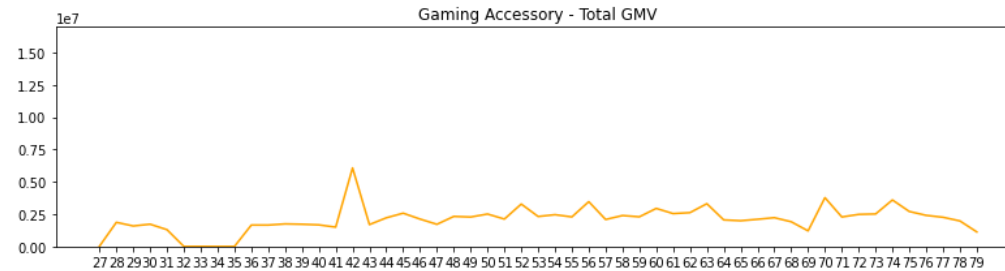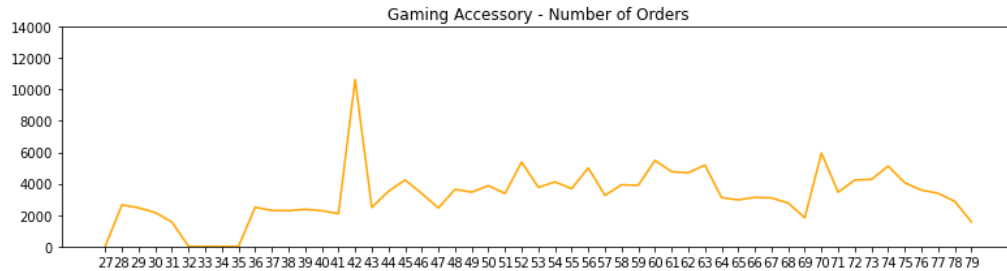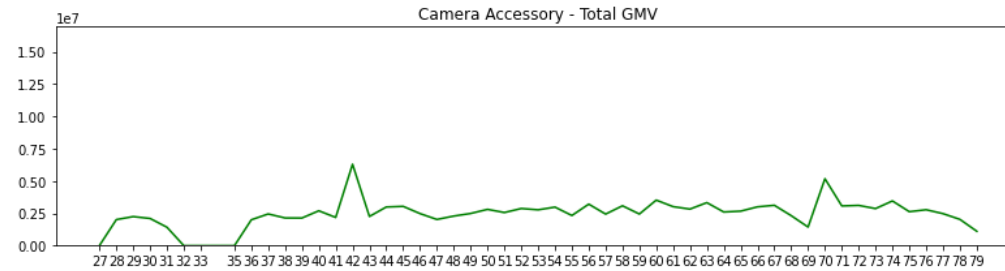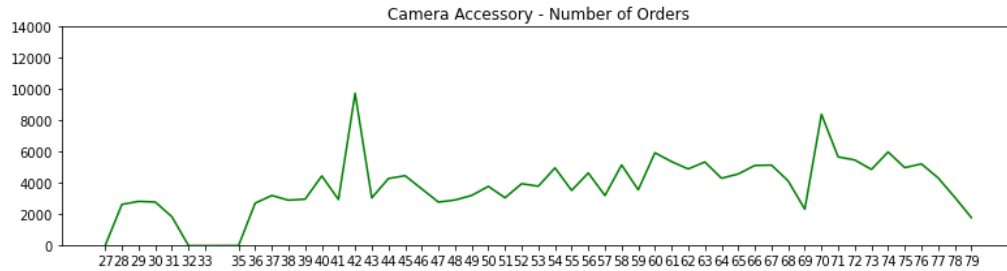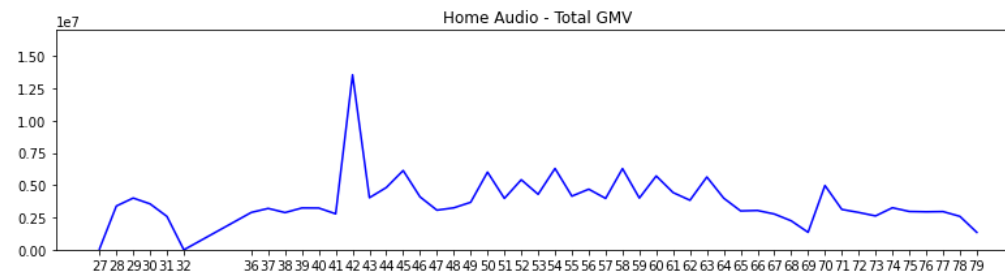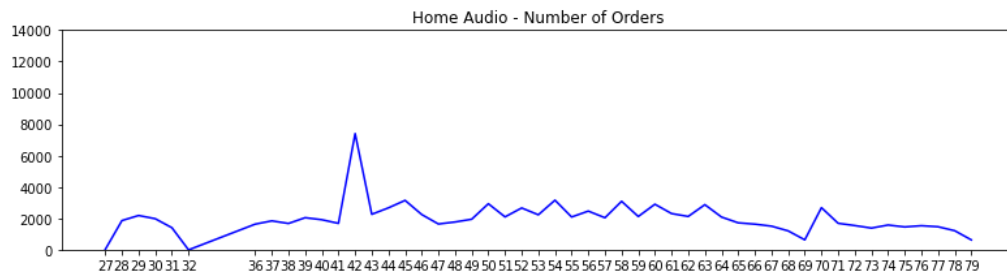
From the weekly trends, below were the major highlights:
- Highest number of orders as well as the highest revenue was generated in week-42.
- The maximum Revenue was generated for Home Audio followed by Camera Accessory.
- The highest count of orders were generated for Gaming Accessory
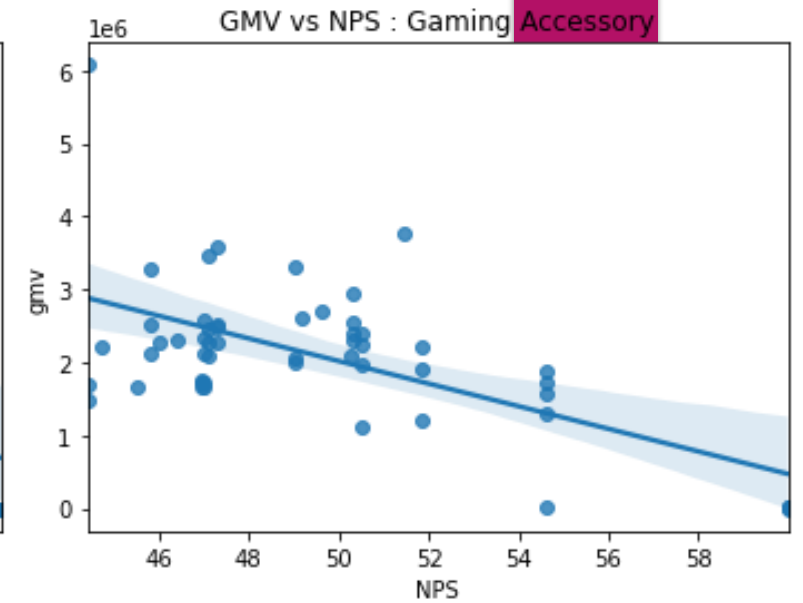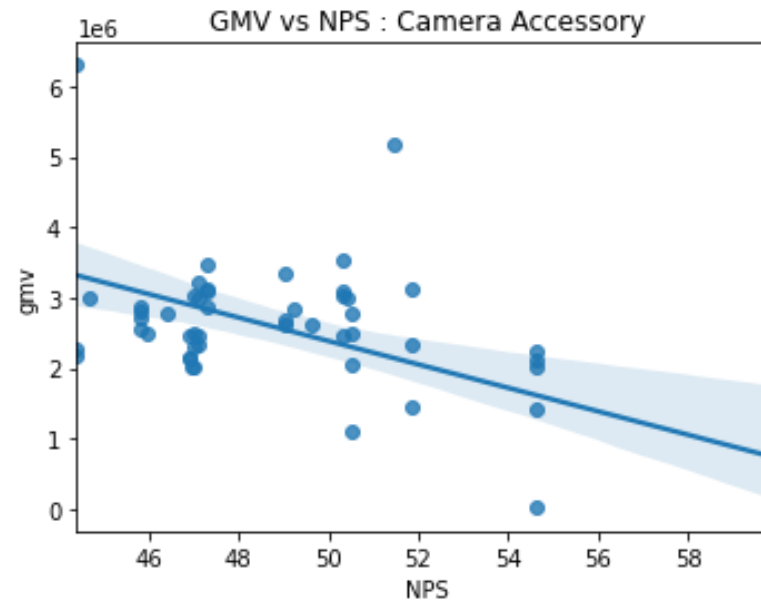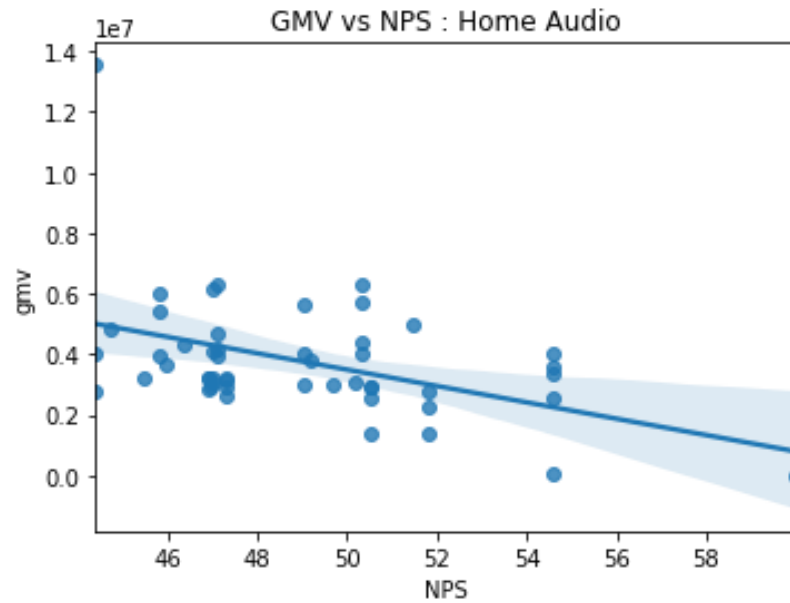
Finally, with the help of above regplot, we had observed that there is a negative correlation between the NPS & total sales(GMV) across all the product sub-categories

Model Building & Evaluation

Camera Accessory Model

Model	Important Features	Accuracy (on test dataset)
Linear / Additive	adStock_Online marketing, Binoculars, CameraBattery, Filter, Strap	96%
Multiplicative	discount, total_premiumn_products, CameraTripod	99%
Distributed Lag	Binoculars, CameraMount, CameraTripod, Lens, Strap, gmv_lag1, product_procurement_sla_lag1	88%

Gaming Accessory Model

| Model	Important Features	Accuracy (on test dataset)
Linear / Additive	adStock_TV, GamePad, GamingHeadset, GamingMemoryCard, GamingMouse, JoystickGamingWheel, MotionController	98%
Multiplicative	GamingHeadset, JoystickGamingWheel	94%
Distributed Lag	GamePad, GamingAccessoryKit, GamingHeadset, GamingKeyboard, GamingMemoryCard, GamingMouse, product_procurement_sla_lag1, NPS_lag3	99%

Home Audio
Model

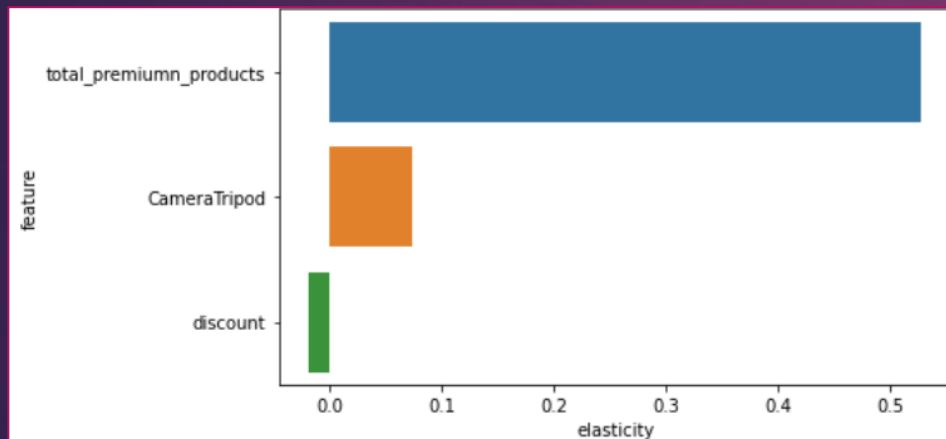| Model	Important Features	Accuracy (on test dataset)
Linear / Additive	FMRadio', 'HomeAudioSpeaker', 'VoiceRecorder	99%
Multiplicative	sla, HomeAudioSpeaker, SoundMixer	99%
Distributed Lag	FMRadio, HomeAudioSpeaker, adStock_Online marketing_lag3	99%

Recommendations & Key Takeaways
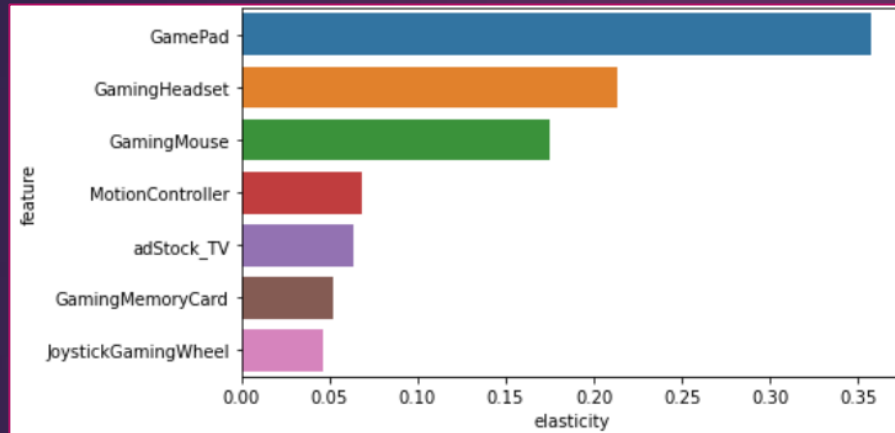
Camera Accessory – Recommendations
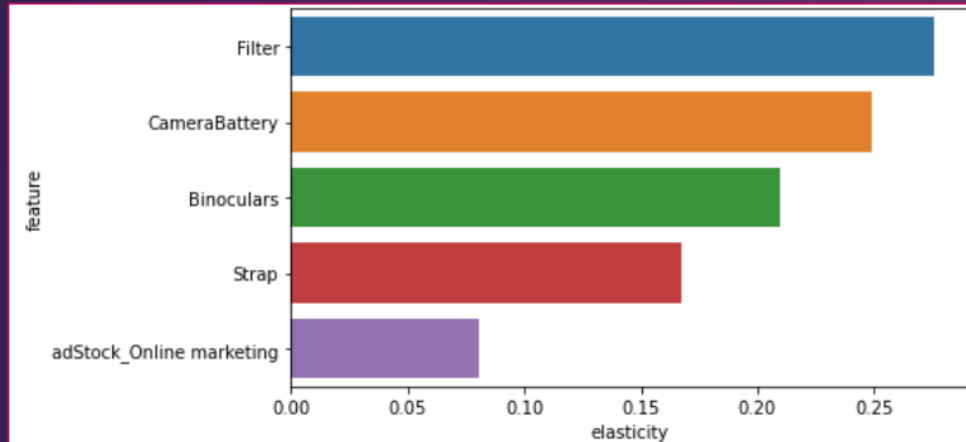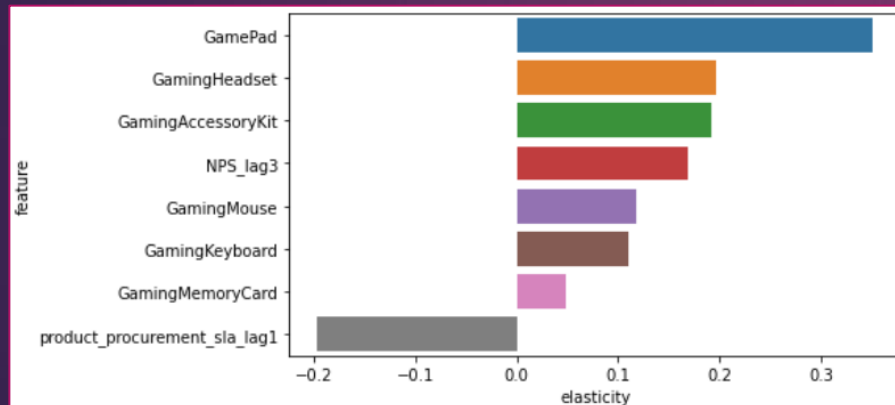

Elasticity with Additive Model

Elasticity with Multiplicative Model

- ► Out of the various ML models that we have built, the Additive/Linear model & the Multiplicative model fits best for the Camera Accessory.

- ► Looking at the Elasticity of the key Features from both the best fit models, below are the recommendations to the business:

  - ► ElecKart should promote products like - Filter, Camera Battery, Binoculars, Strap & Camera Tripod as it has a very positive & incremental impact on the overall Revenue(GMV) growth.

  - ► More marketing efforts should be spent on the promotion of Premium products.

  - ► For marketing & products promotion, it's a good idea to invest more in Online Marketing channels as it has shown a positive impact in boosting the overall revenue for the company.

  - ► A slight increase in product discounts could also help to increase the sales marginally.
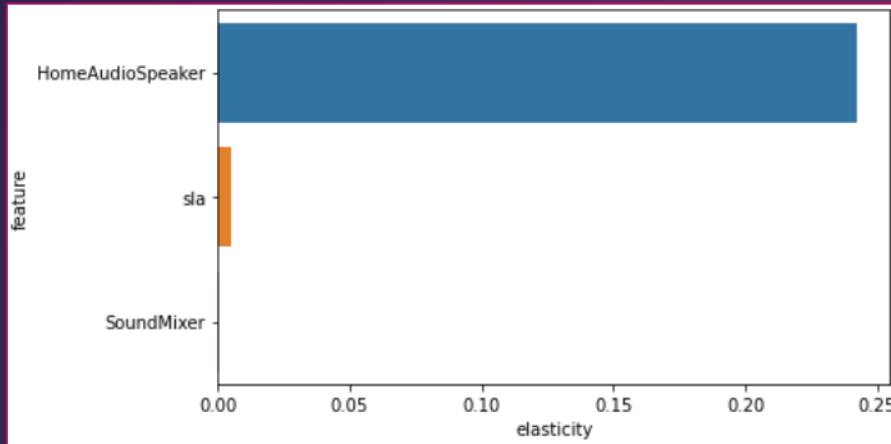
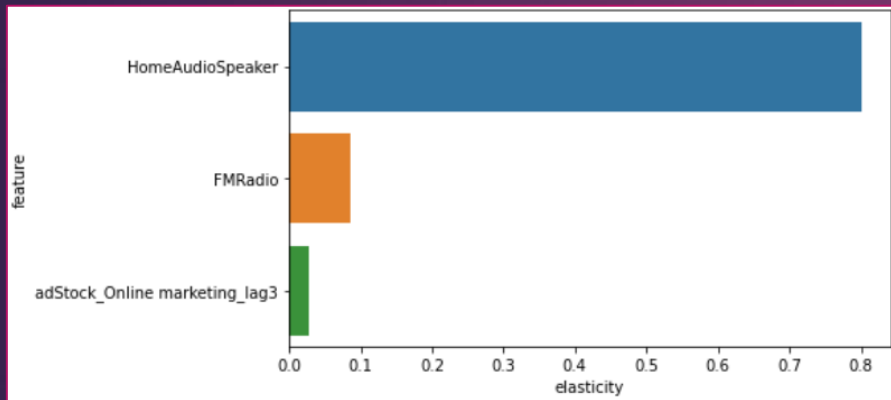# Gaming Accessory – Recommendations

Elasticity with Additive Model

Elasticity with Distributed Lag Model

▶ Out of the various ML models that we have built, the Additive/Linear model & the Distributed Lag model fits best for the Gaming Accessory.

▶ Looking at the Elasticity of the key Features from both the best fit models, below are the recommendations to the business:

　▶ ElecKart should promote products like – GamePad, Gaming Headset, Gaming Mouse, Gaming Memory Cards & Accessory Kits as they fetch the highest revenue(GMV) for the company.

　▶ Marketing & Advertising spends on TV should be slightly increased as it has a positive impact on the revenue.

　▶ An increase in NPS scores for the last 3 weeks has also contributed positively to the revenue growth of the business.

　▶ Product Procurement SLA should be reduced for the last 1 week.

Home Audio – Recommendations

Elasticity with Multiplicative Model

Elasticity with Distributed Lag Model

► Out of the various ML models that we have built, the Multiplicative model & the Distributed Lag model fits best for the Home Audio.

► Looking at the Elasticity of the key Features from both the best fit models, below are the recommendations to the business:

  ► ElecKart should promote products like – Home Audio Speakers & FM Radio as they have shown a significant positive impact on revenue(GMV) growth (especially through the sale of Home Audio Speakers) of the company.

  ► Marketing & Advertising spends on Online Marketing channels for the last 3 weeks should be slightly increased as it can contribute to the revenue growth marginally.

  ► Even if in case the SLA increases slightly, it hasn't negatively affected the revenue growth for the business.

Thank You ☺