ElecKart Market Mix Modeling

- **Business & Data Understanding**
- Exploratory Data Analysis
- Model Building
- Recommendations

Business & Data Understanding – ElecKart Market Mix Modeling

Objective:

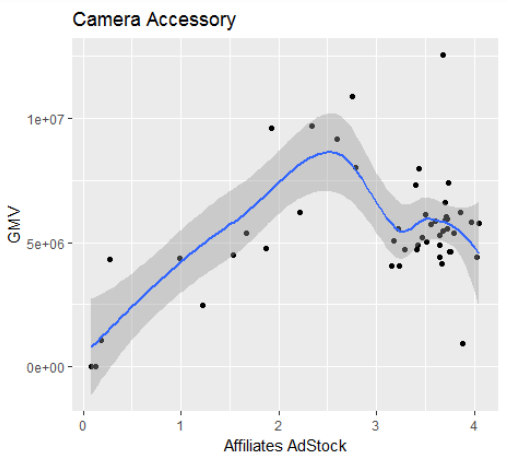
- To identify the most impactful channels for marketing spends
- To identify other factors having positive impact on sales of 3 product categories – Camera Accessory, Game Accessory & Home Audio

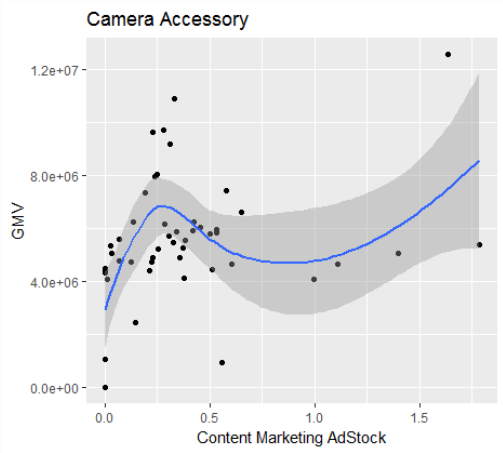
Data Understanding:

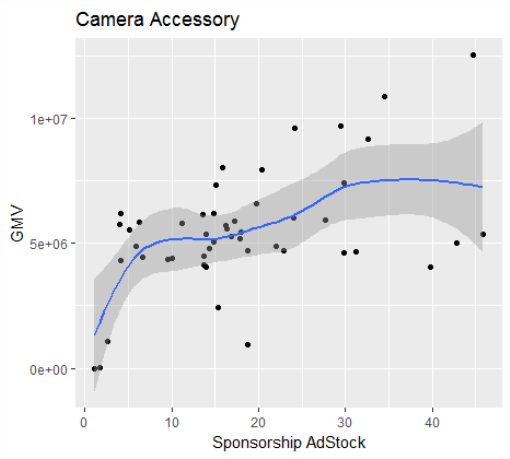
- Data available for a period of 1 year from July 2015 to June 2016
- The available data is at an order level for all the purchases made in the said time period
- The monthly advertising spends on the different marketing channels is also available
- The NPS (or brand perception) data is also available together with the holiday list for the year

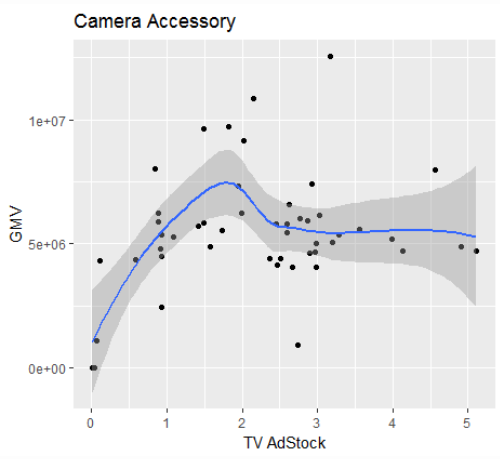
- Business & Data Understanding
- **Exploratory Data Analysis**
- Model Building
- Recommendations

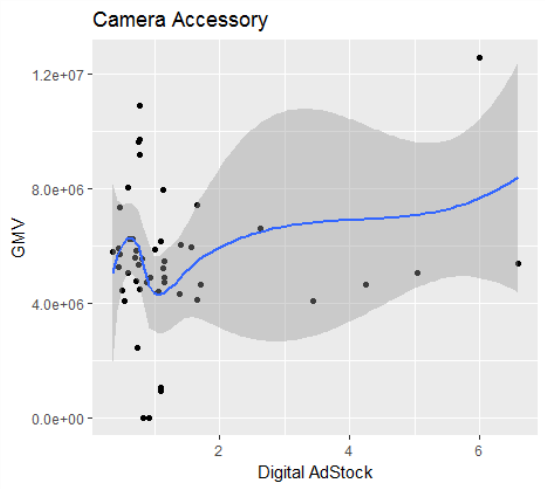
EDA – Figuring out Important Variables – Camera Accessory

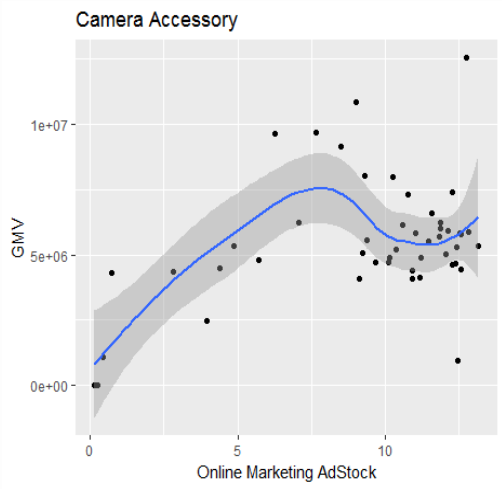


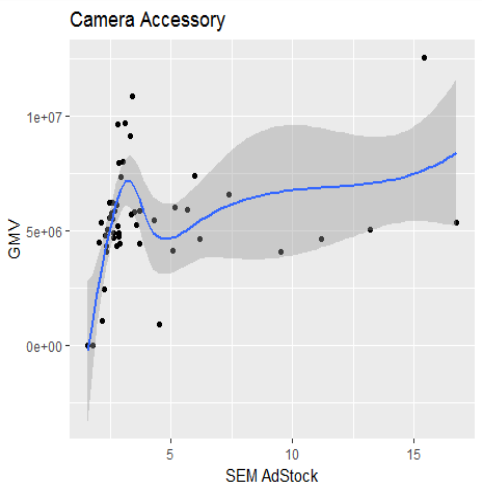






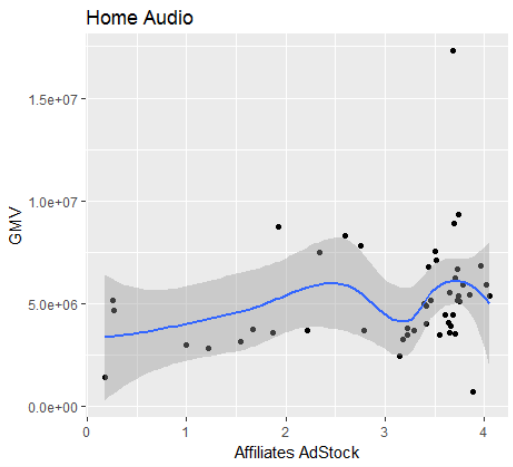


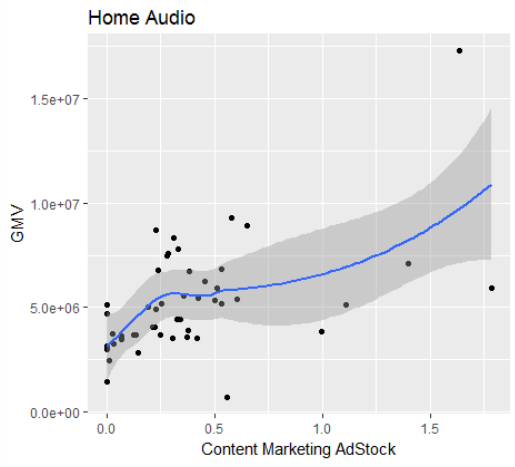


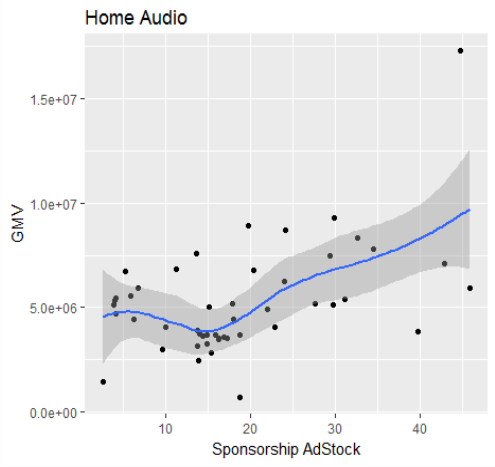


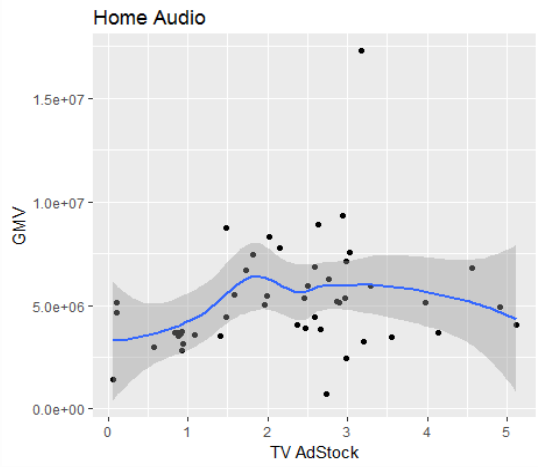
UpGrad

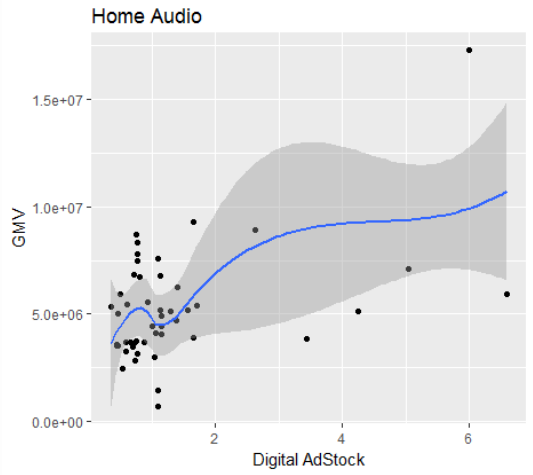
EDA – Figuring out Important Variables – Home Audio

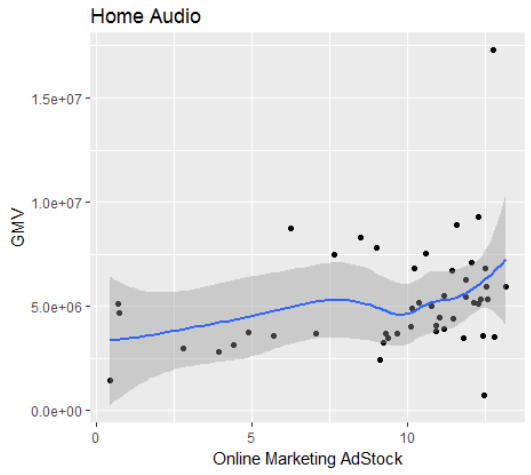


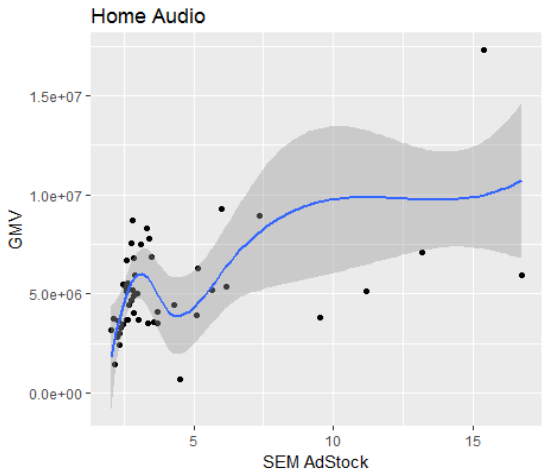






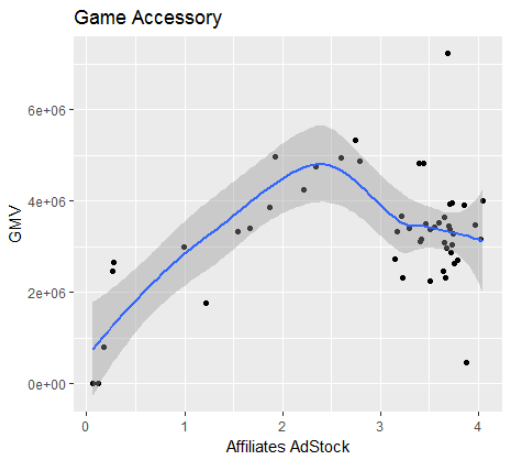


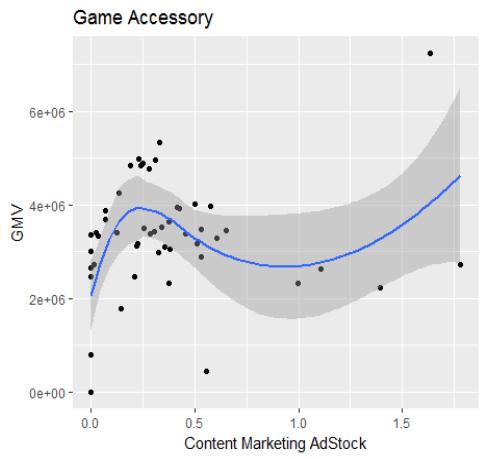


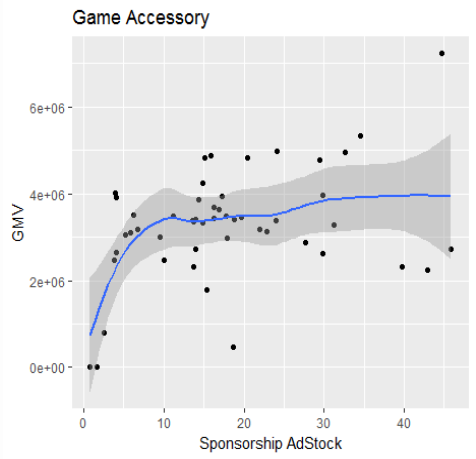


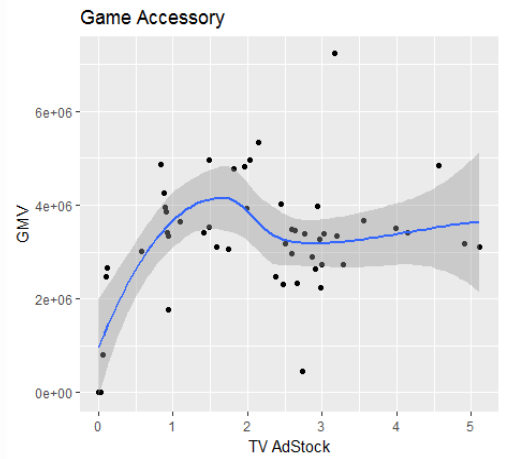
UpGrad

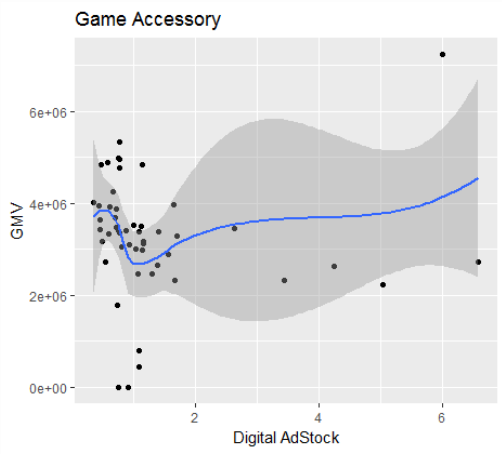
EDA – Figuring out Important Variables – Game Accessories

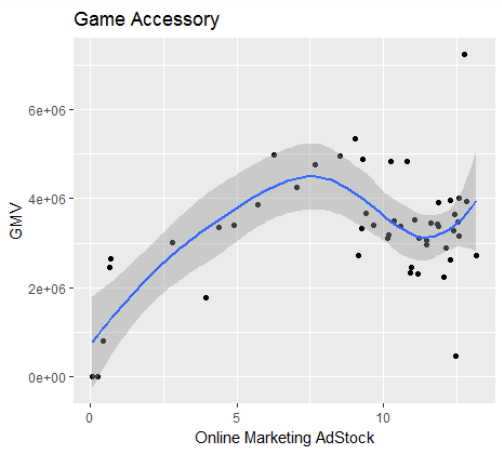


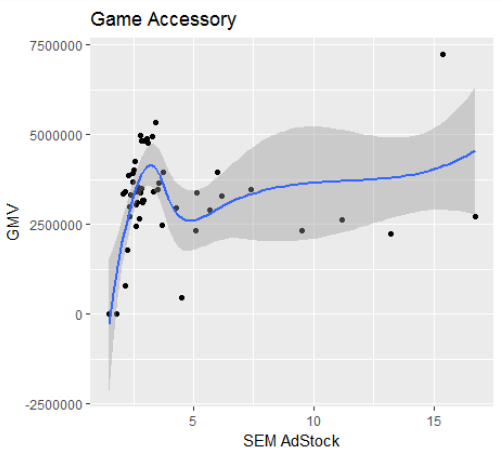






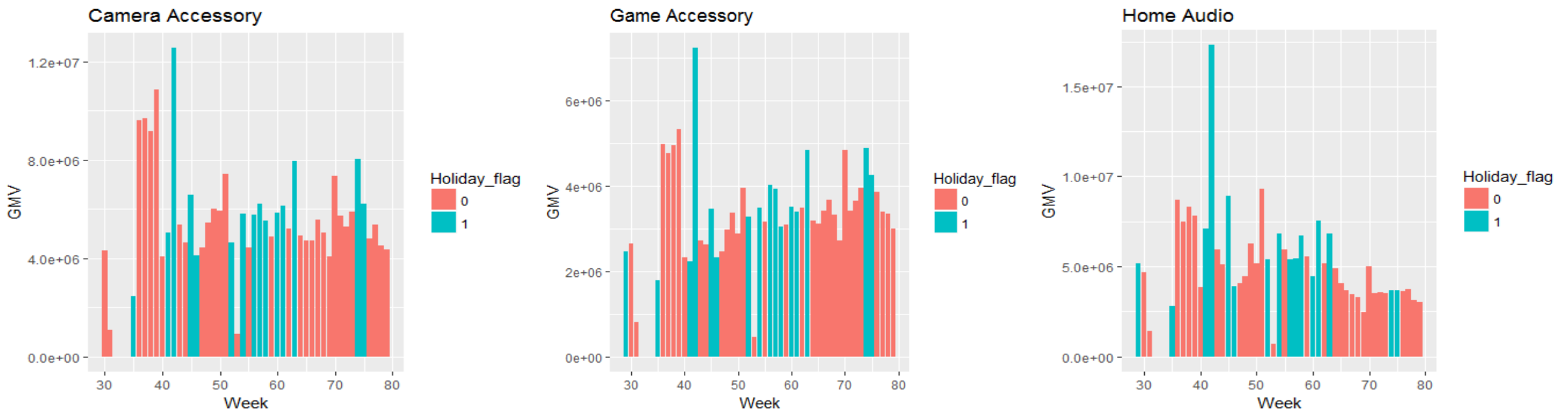






UpGrad

EDA – Figuring out Important Variables – Other Important Variables (Holidays)



The Blue bars represent the Sales value in weeks with Holiday/Festival falling within them. As we can see, these weeks have higher sales figure than their neighbouring weeks.

UpGrad

- Data Understanding & Data Preparation
- Exploratory Data Analysis
- **Model Building**
- Recommendations

Camera Accessories – Building the 5 Models

- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 10 fold cross validation again on the training data
- Even though, the Multiplicative model has a decent Adj. R sq figure, but it has just 2 important variables. Also, these variables may not be too important from business point of view.
- The combination of multiplicative and distributed lag model is ruled out because of high SSE value.
- So, we select the simple linear model & Koyck model on the account of higher Adj. R Sq values

Model	Variables	Adj. R Square	SSE (10 - CV)
Simple Linear model	SLA + Digital + Sponsorship + inc_PO_MA1	0.3539	0.861
Multiplicative model	per_order + NPS	0.834	0.686
Koyck model	Digital + Sponsorship + Online_marketing	0.3503	0.904
Distributed Lag model	Sponsorship + promotional_offer.3	0.2604	0.797
Multiplicative + Distributed Lag model	promotional_offer + per_order + gmv_1	0.7581	3.069

Game Accessories – Building the 5 Models

- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 10 fold cross validation again on the training data
- The multiplicative model and the combination of the multiplicative & distributed lag model are ruled out because of high SSE value
- We are also looking for models that include more business actionable variables
- So, we select the Koyck model and the Distributed on the account of higher Adj. R Sq values

Model	Variables	Adj. R Square	Cross-Validation
Simple Linear model	NPS + SEM + inc_PO_MA2	0.3606	0.821
Multiplicative model	per_order + Affiliates	0.5844	1.575
Koyck model	NPS + inc_LP_MA3 + inc_PO_MA1 + gmv.1	0.4061	0.857
Distributed Lag model	holiday_freq + inc_LP_MA1 + inc_PO_MA1 + promotional_offer.2 + NPS.3 + gmv.1	0.5335	0.722
Multiplicative + Distributed Lag model	Digital + Sponsorship + promotional.offer.2 + NPS_2 + gmv_1	0.6379	1.942

Home Audio– Building the 5 Models

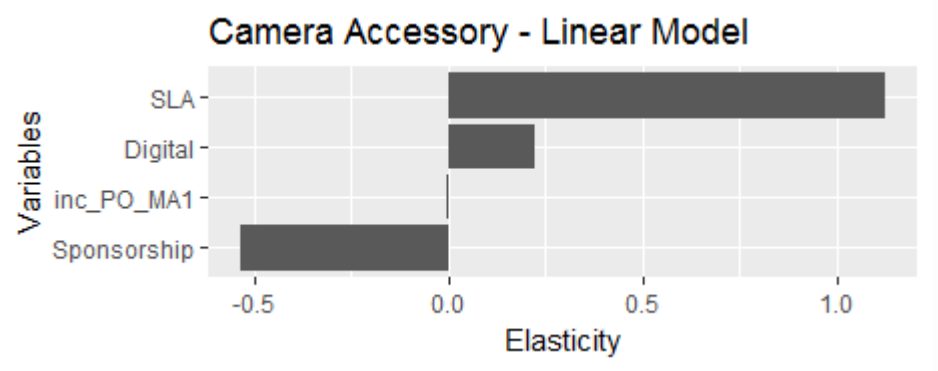
- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 10 fold cross validation again on the training data
- The multiplicative model and the combination of the multiplicative & distributed lag model are ruled out because of the very low Adj. R Sq values
- Out of the remaining 3, the decision was made on the basis off the elasticity of KPIs and their business sense
- So, we select the Koyck model and the Distributed on the account of higher Adj. R Sq values

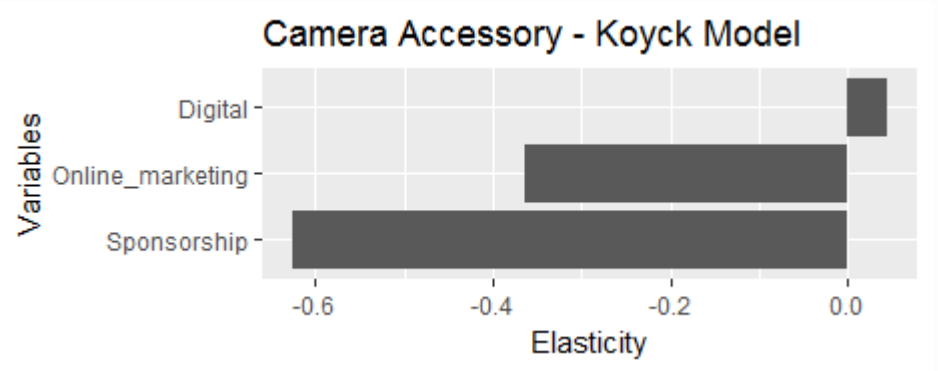
Model	Variables	Adj. R Square	Cross-Validation
Simple Linear model	holiday_freq + Sponsorship	0.3064	0.897
Multiplicative model	promotional_offer + Digital	0.2593	0.198
Koyck model	Digital + inc_PO_MA3	0.3118	0.890
Distributed Lag model	promotional_offer + Sponsorship	0.4291	0.672
Multiplicative + Distributed Lag model	promotional_offer + NPS_1	0.2492	0.181

- Business & Data Understanding
- Exploratory Data Analysis
- Model Building
- **Recommendations**

Camera Accessories – Recommendations Based on Elasticity of KPIs

- The adjoining figures represent the elasticity of the different variables w.r.t the overall sales figure.
- Positive elasticity means that increasing the value of the KPI would lead to increase in the sales figure
- Therefore, the objective would be to allocate more spending on the KPIs that have positive elasticity w.r.t the sales.
- Thus, going with the Basic Linear Model, ElecKart should focus more on the Digital channel and decrease its spending on Sponsorships.
- Similar conclusions can be drawn from the Koyck model – Increasing the spending on digital and decreasing on Online Marketing and Sponsorship.

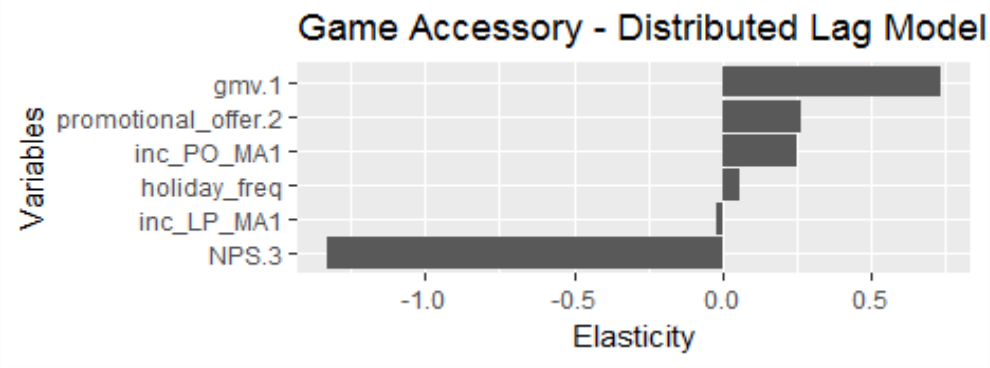


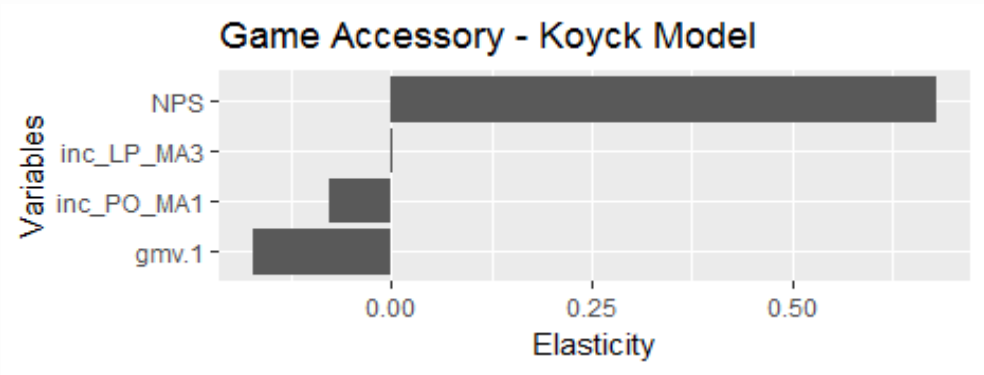


UpGrad

Game Accessories – Recommendations Based on Elasticity of KPIs

- Here, the Distributed Lag model suggests that the total sales figure of the previous week, promotional discount offered in the past, increase in the discount percentage w.r.t the previous weeks and finally presence of holidays in the week have the most impact on the sales figure.
- Thus, ElecKart should channelize more of the resources in offering discounts rather than focusing on marketing channels.
- Not many actionable insights can be drawn using the Koyck model.

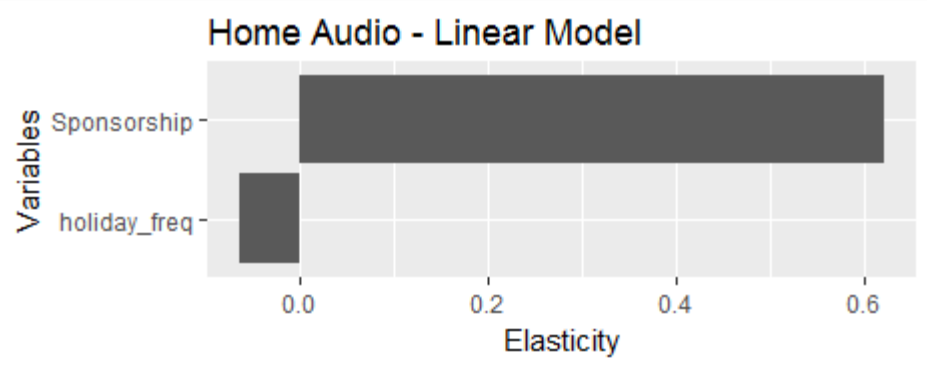


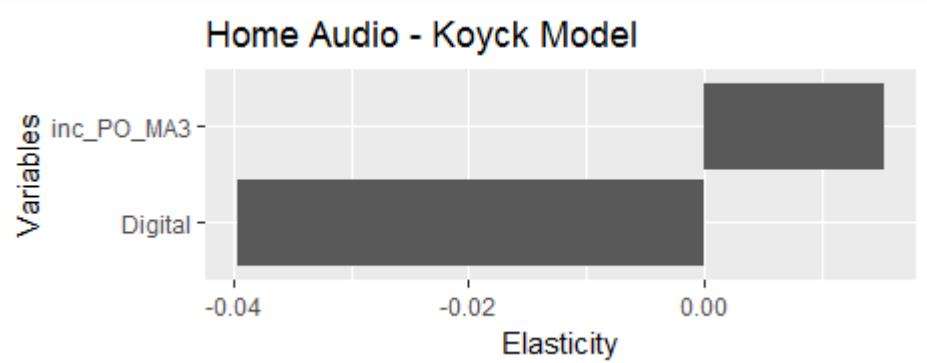


UpGrad

Home Audio – Recommendations Based on Elasticity of KPIs

- For Home Audio segment, the linear model suggests that more resources be allocated in Sponsorships.
- Similarly, the Koyck model suggests that the resources should not be allocated through the Digital channel. Moreover, offering promotional discounts can be beneficial.





UpGrad