Capstone Project

ElecKart Market Mix Modeling

Group Name:

1. Ishita Aggarwal
2. Sundaresan V
3. Dhruti Contractor
4. Dhanilan M S

Eleckart Market Mix Model

ElecKartMarket Mix Modeling Objective:

- To identify the most useful revenue streaming channels for marketing spends.
- To identify other factors having positive impact on sales of three product categories –Camera Accessory, Gaming Accessory & Home Audio Accessory.

Data Understanding:

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

Approach

1.Data Preparation

1. Missing Values and Outlier Treatment
2. Variable Transformation
3. Unified date format

2.Data Exploration

1. Variable Identification
2. Univariate and Bi-Variate Analysis

3.Feature Engineering

1. Feature Extraction and Selection
2. Variable Interaction
3. Feature Creation

4.Model Building

1. Basic linear Model
2. Multiplicative Model
3. Koyck Model
4. Distributed Lag Model
5. Multiplicative + Distributed Lag Model

5.Model Evaluation and Recommendation

Source Code Details

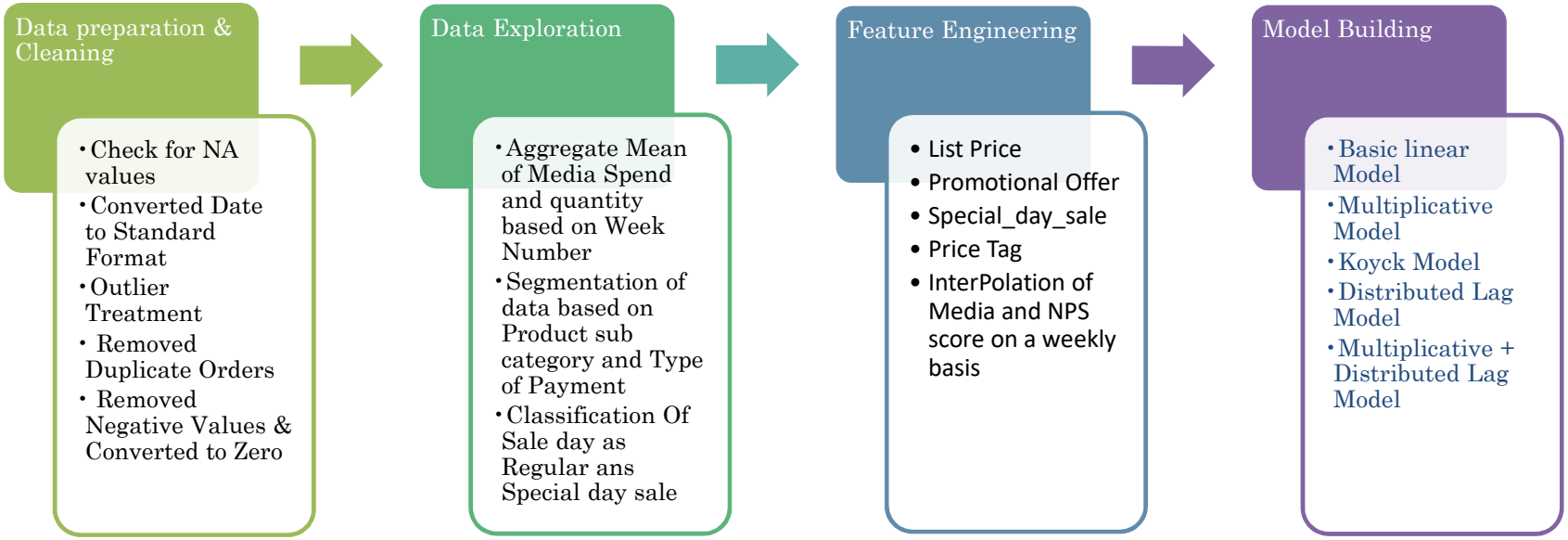
We have segregated our code into three parts.

1. Data Preparation and Model building is available under the filename “Ecommerce_Capstone_Model.R”
2. Data Preparation and EDA is available under “EDA_Capstone.R”

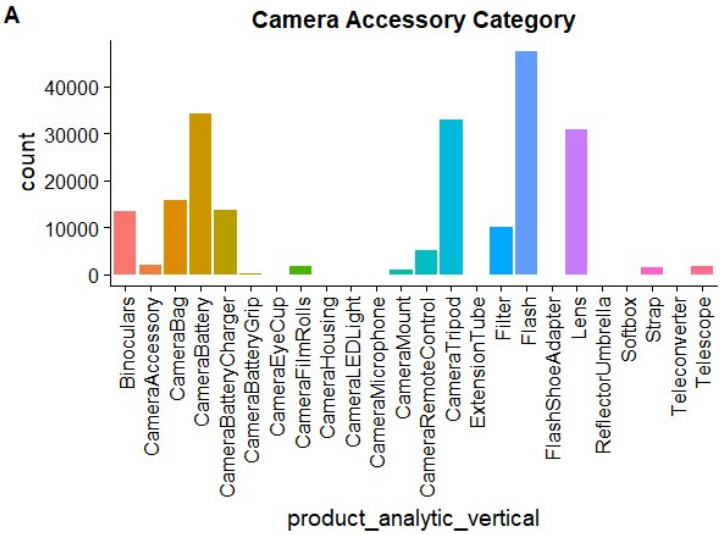
Note:

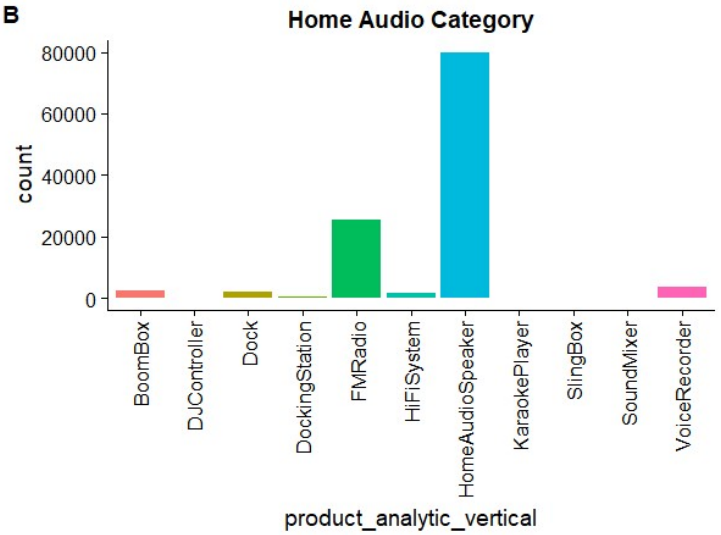
Data preparation is same under both files. Due to naming conventions used inside, we have segregated the files.

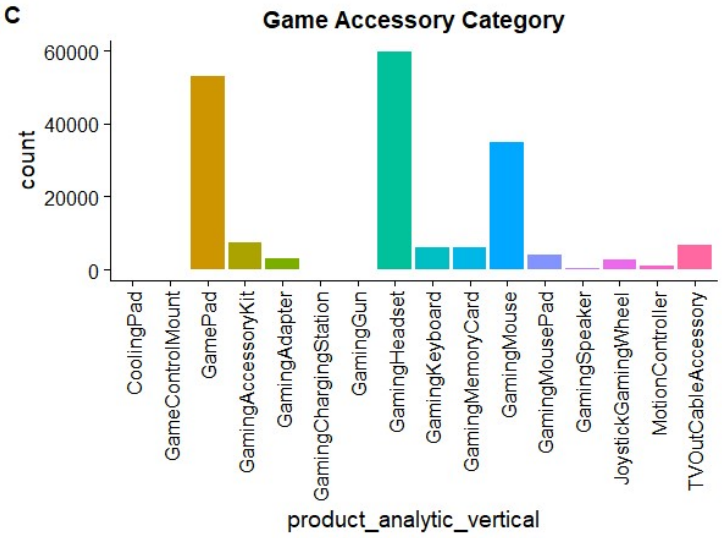
Problem Solving approach



EDA for Different Product Anlaytic Category

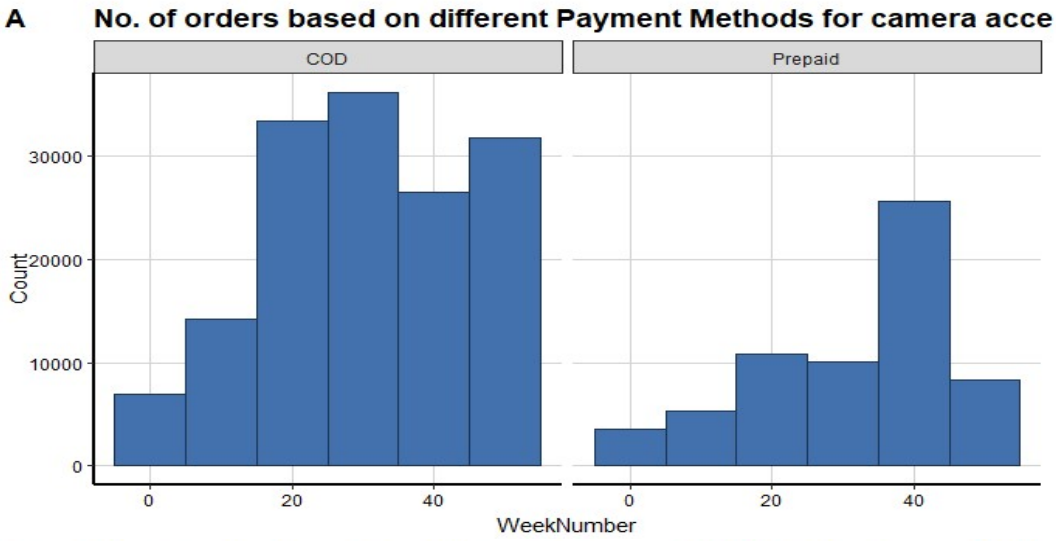


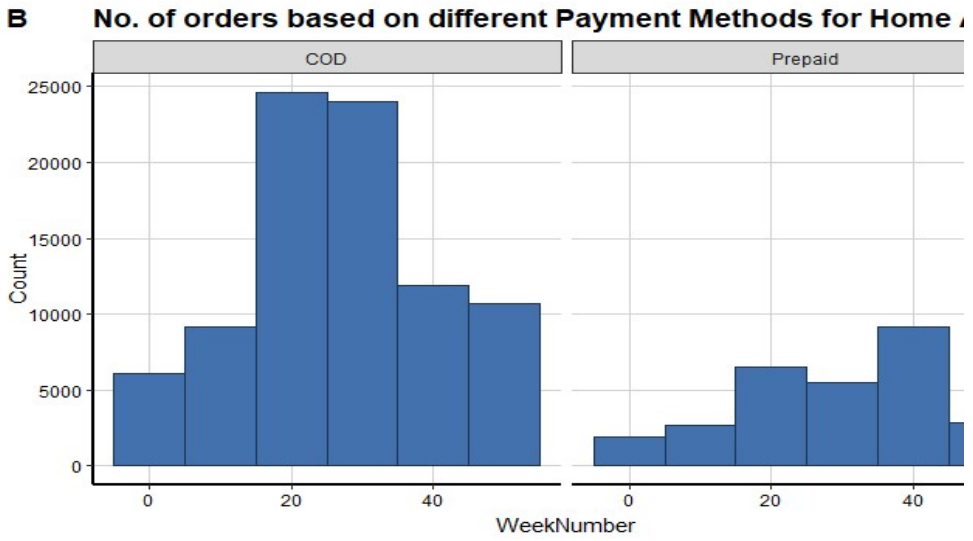


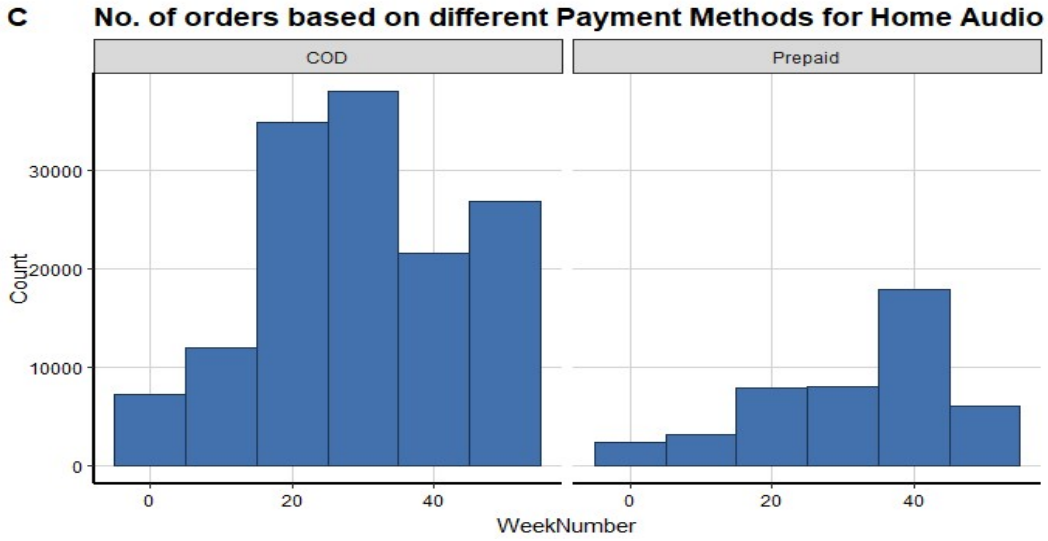


- Under Camera Accessories we found Camera Lens, Flash , Tripod are high selling categories , where as CamerLED Light, Extension tube and Camera Housing are low selling Categories.
- Under Home Audio we found Home Audi Speakers and FM Radio are high selling products,where as SingBox , Sound Mixer and DJ Controller does not have any demand.
- Under Gaming Accessories “Gaming Headset”, “GamePad” and GameMouse are high selling categories where as Cooling Pad , GameControl Mount and Gaming Charging Station got zero demand.

EDA on Payment methods for all three categories



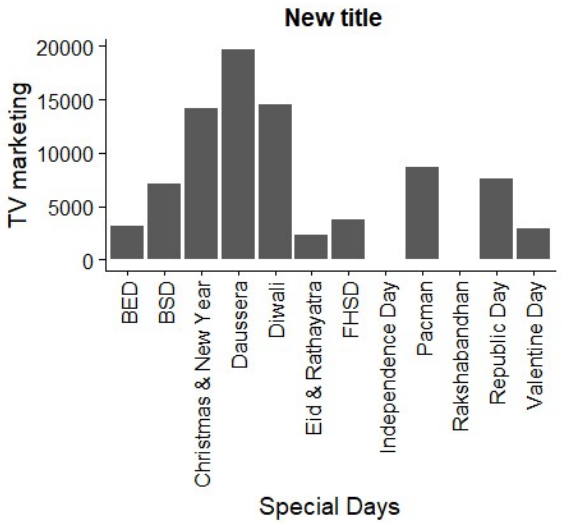


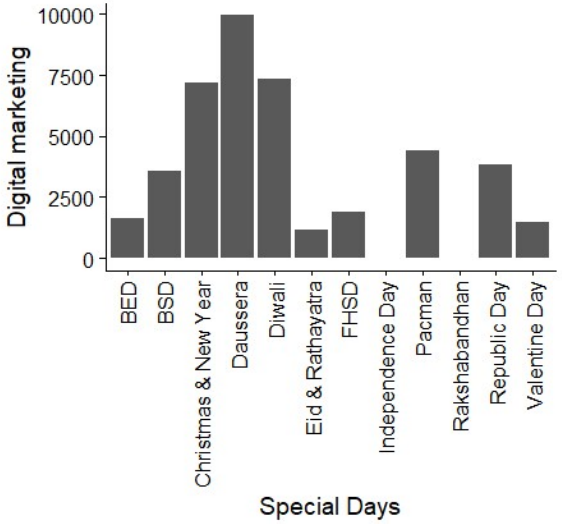


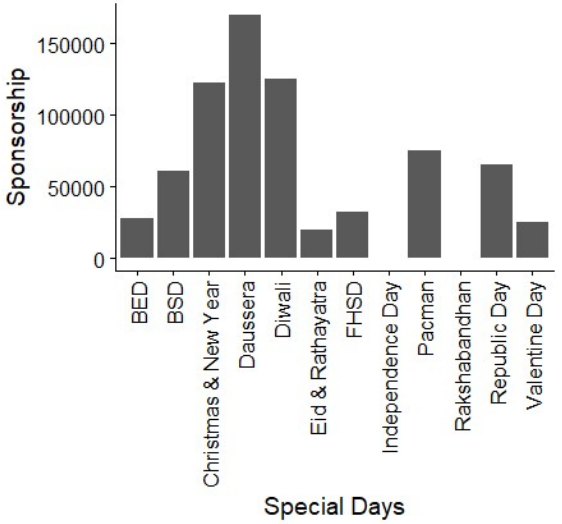
Between week 25 and 35 there is huge demand on COD Payment type across all three segments.

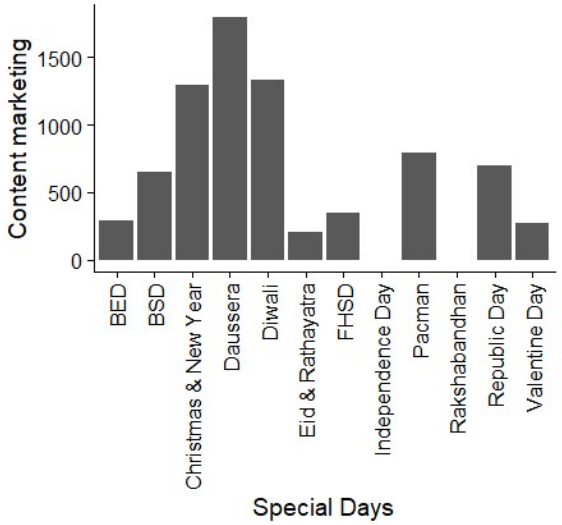
Between week 35 and 40 there is also relatively a small surge in prepaid (Online) Payment type

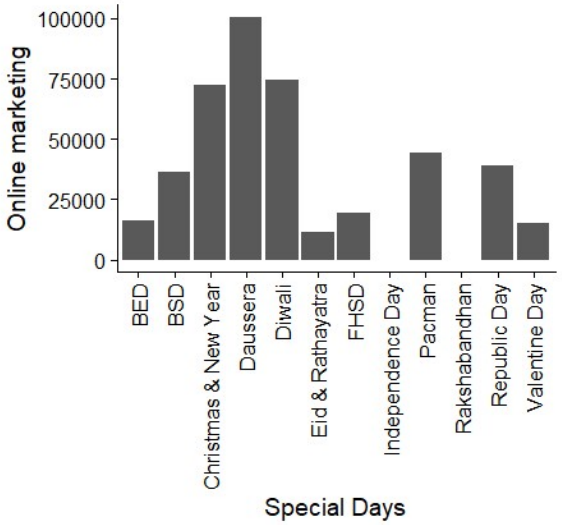
Media Marketing Spend on Special Days for Camera Accessory





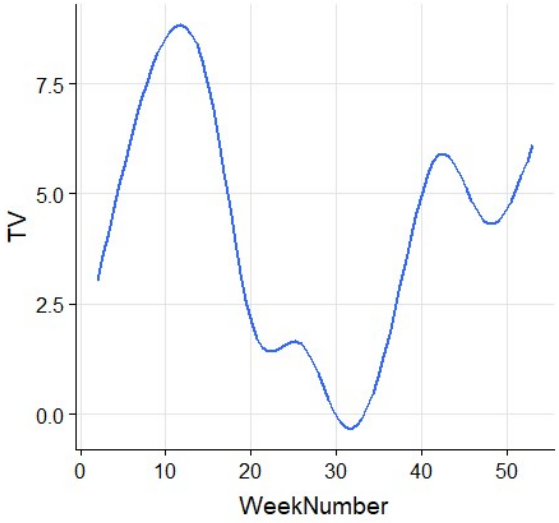


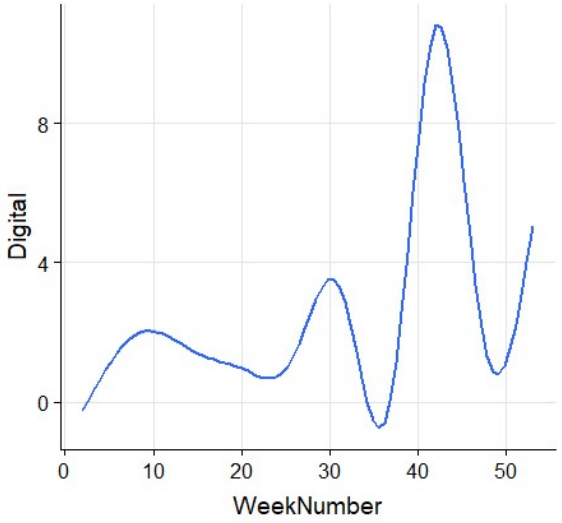


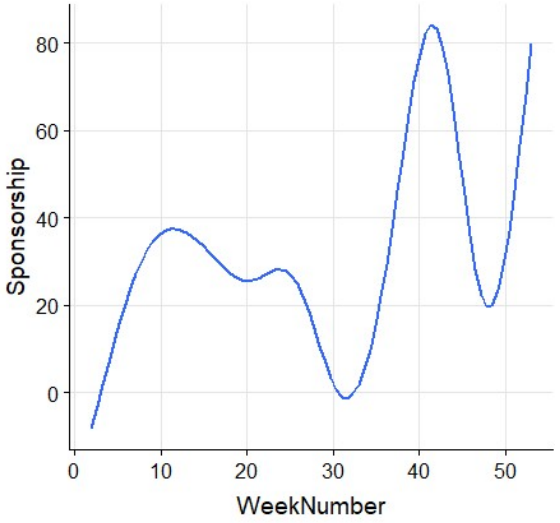


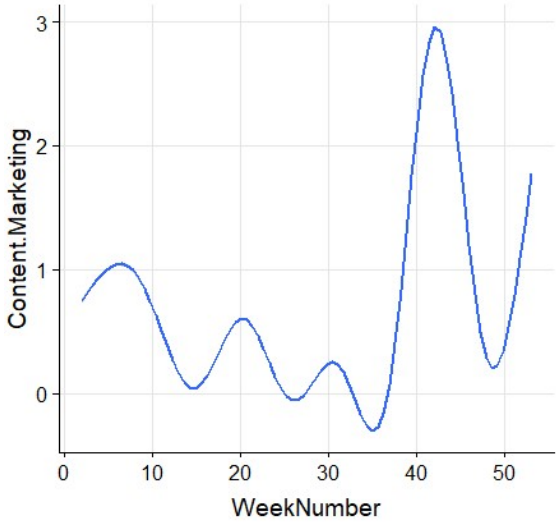
- On Independence Day and Rakshabandan Day there is no spend on media advertisement. They should try leveraging their media spend on these days too.

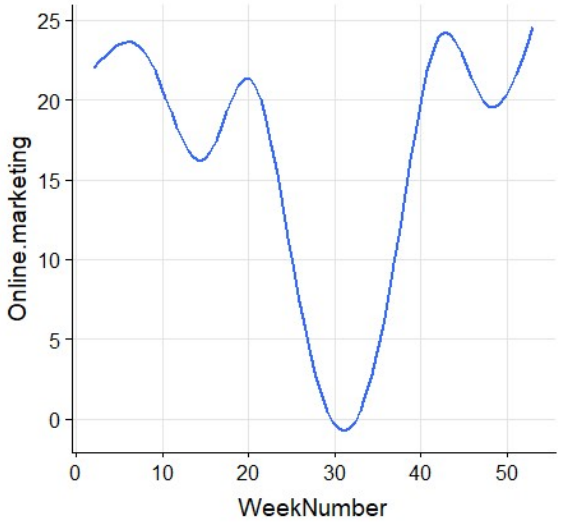
Media Marketing Spend on Regular Days for Camera Accessory



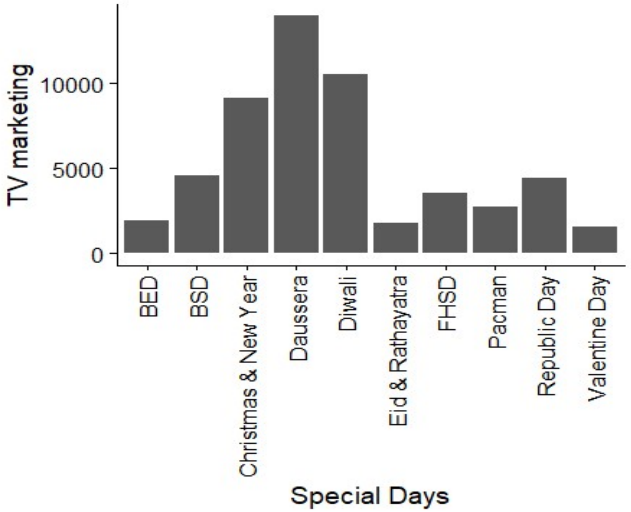


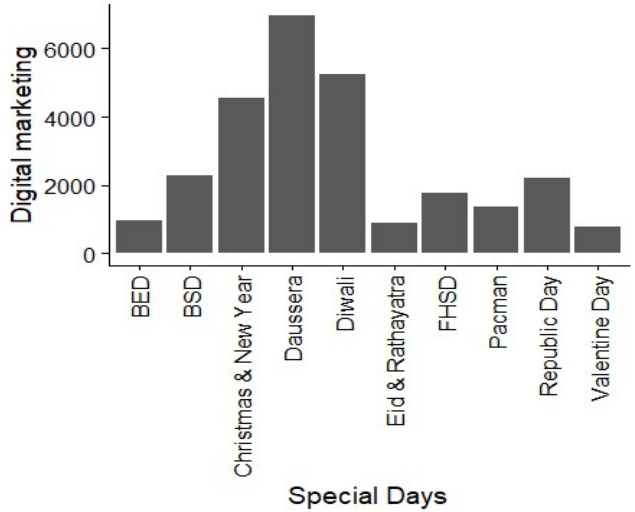


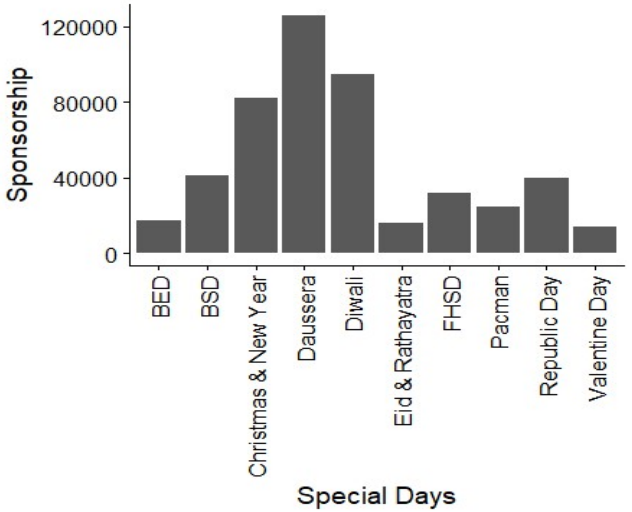


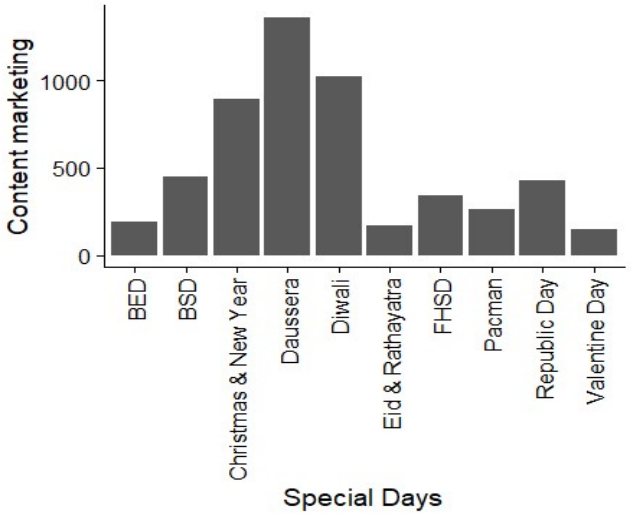


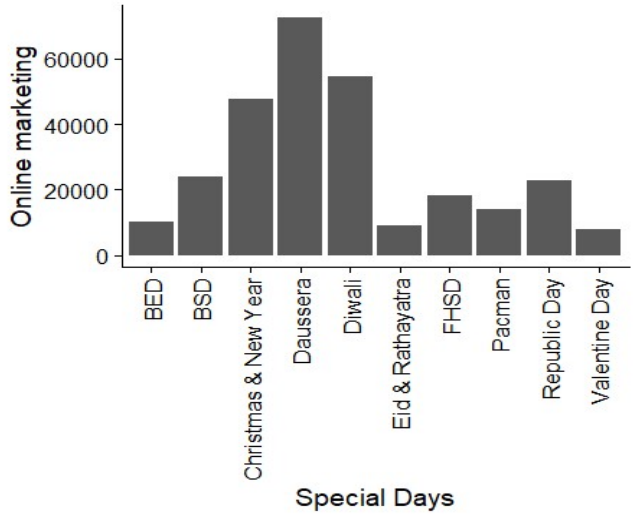
Media Marketing Spend on Special Day for Home Audio accessory



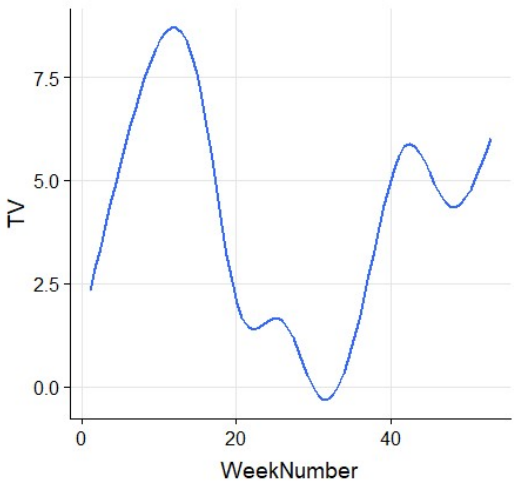


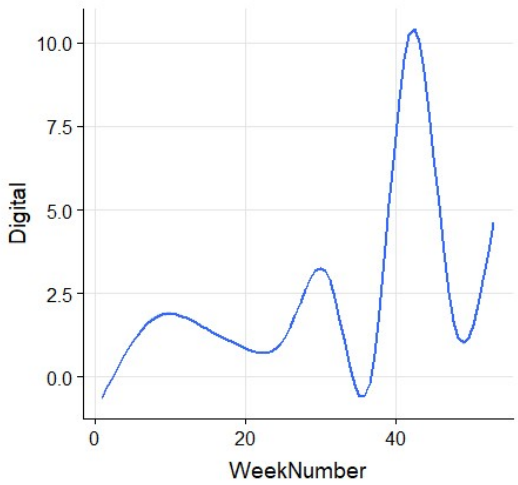


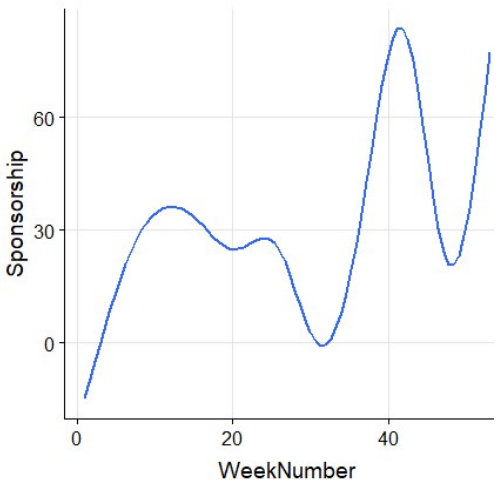


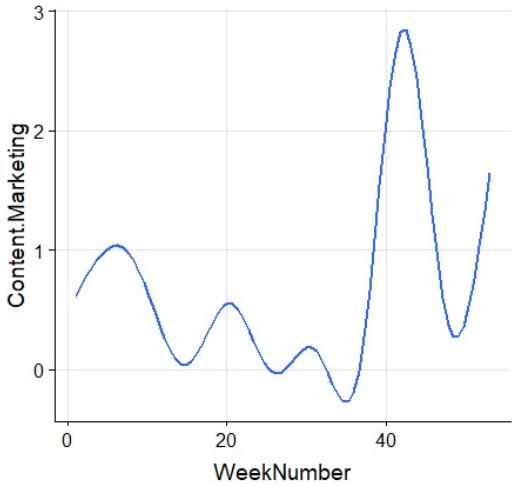


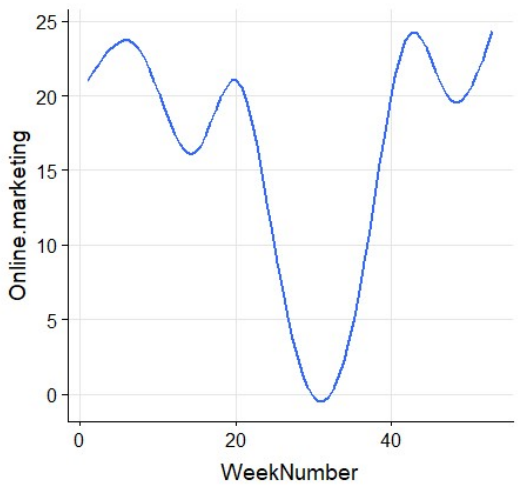
Media Marketing Spend on Regular Days for Home Audio Accessory



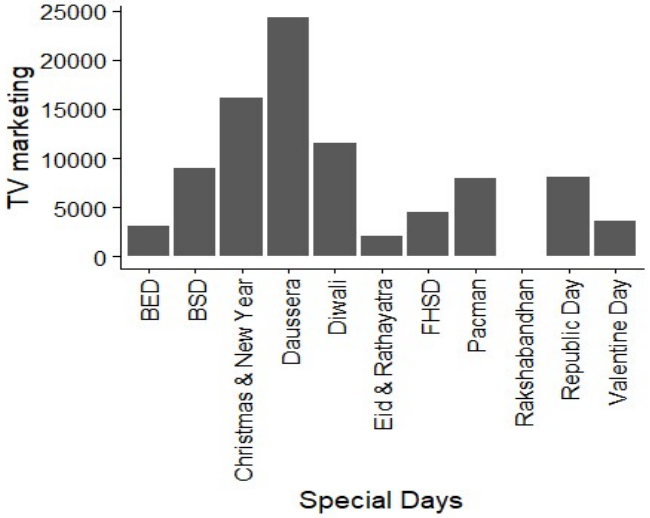


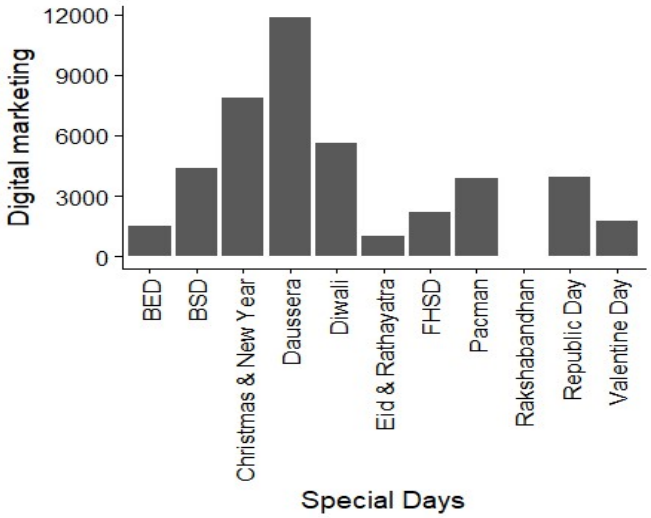


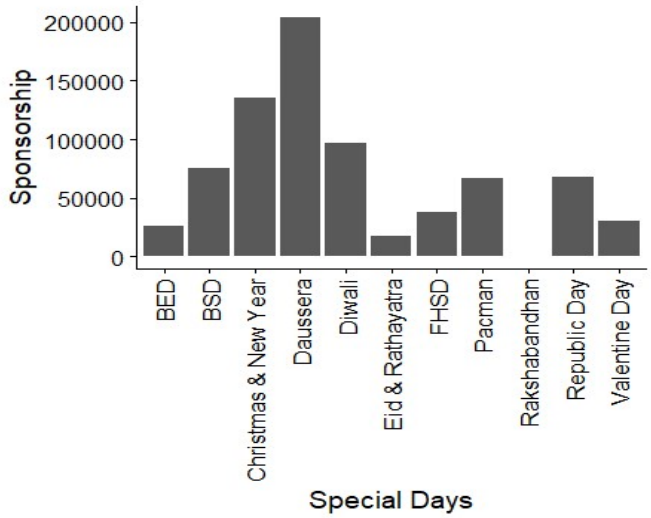


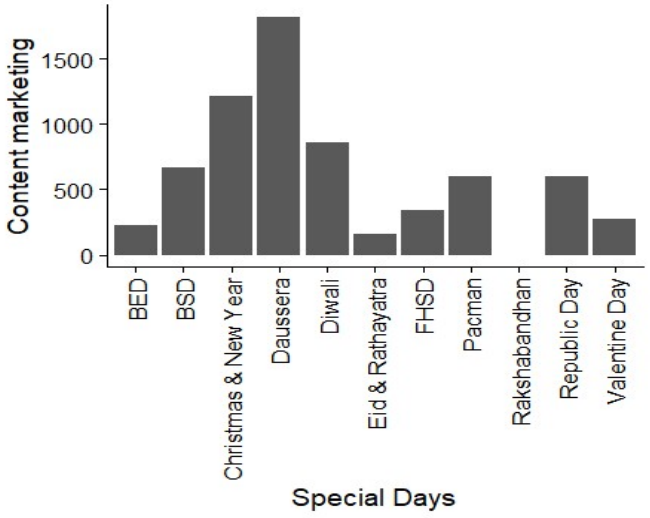


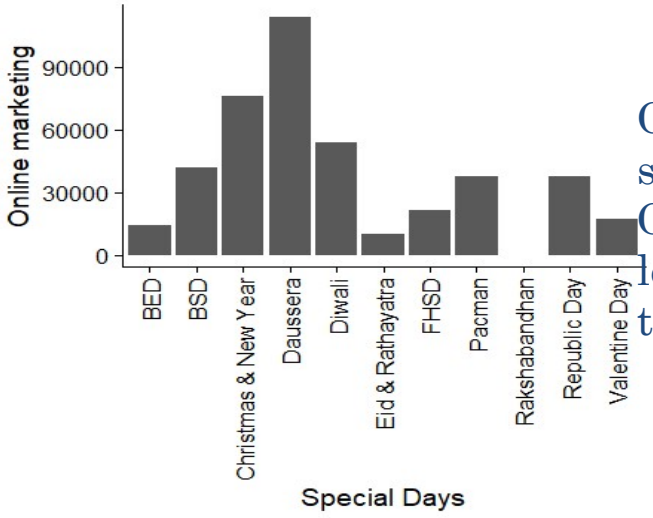
Media Marketing Spend on Specials Days for Gaming Accessory





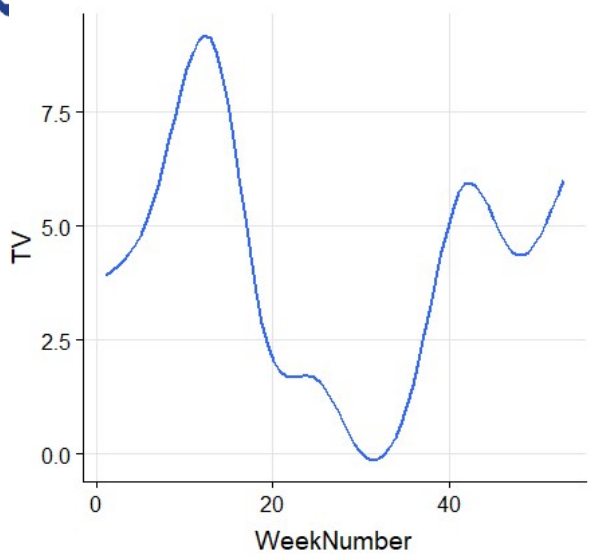


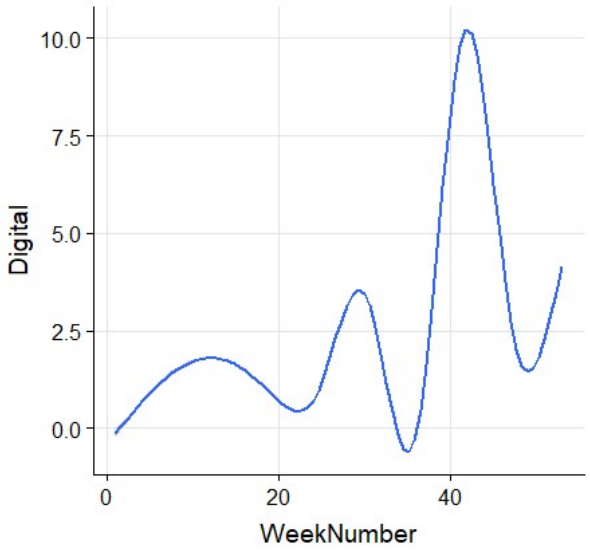


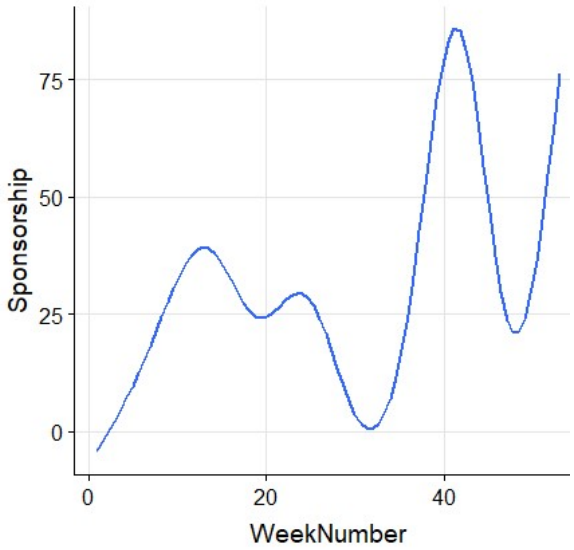


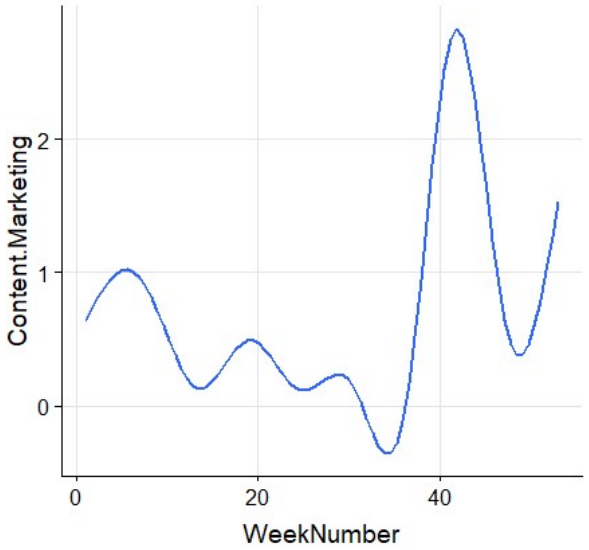
On Rakshabandan Day there is no spend on media advertisement for Gaming accessory. They should try leveraging their media spend on these days too.

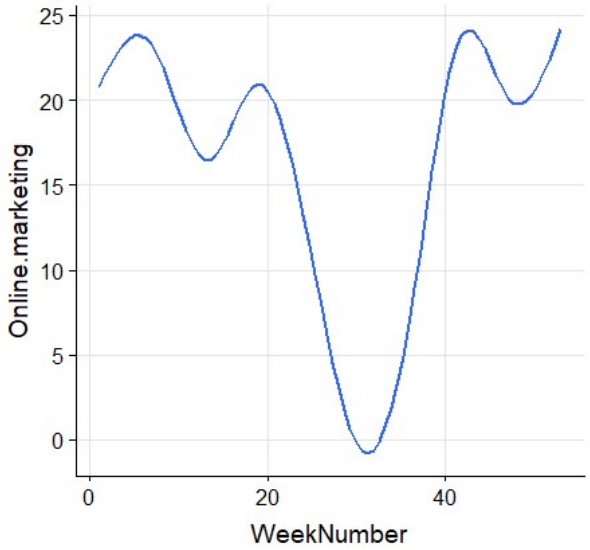
Media Marketing Spend on Regular Days for Gaming Accessory



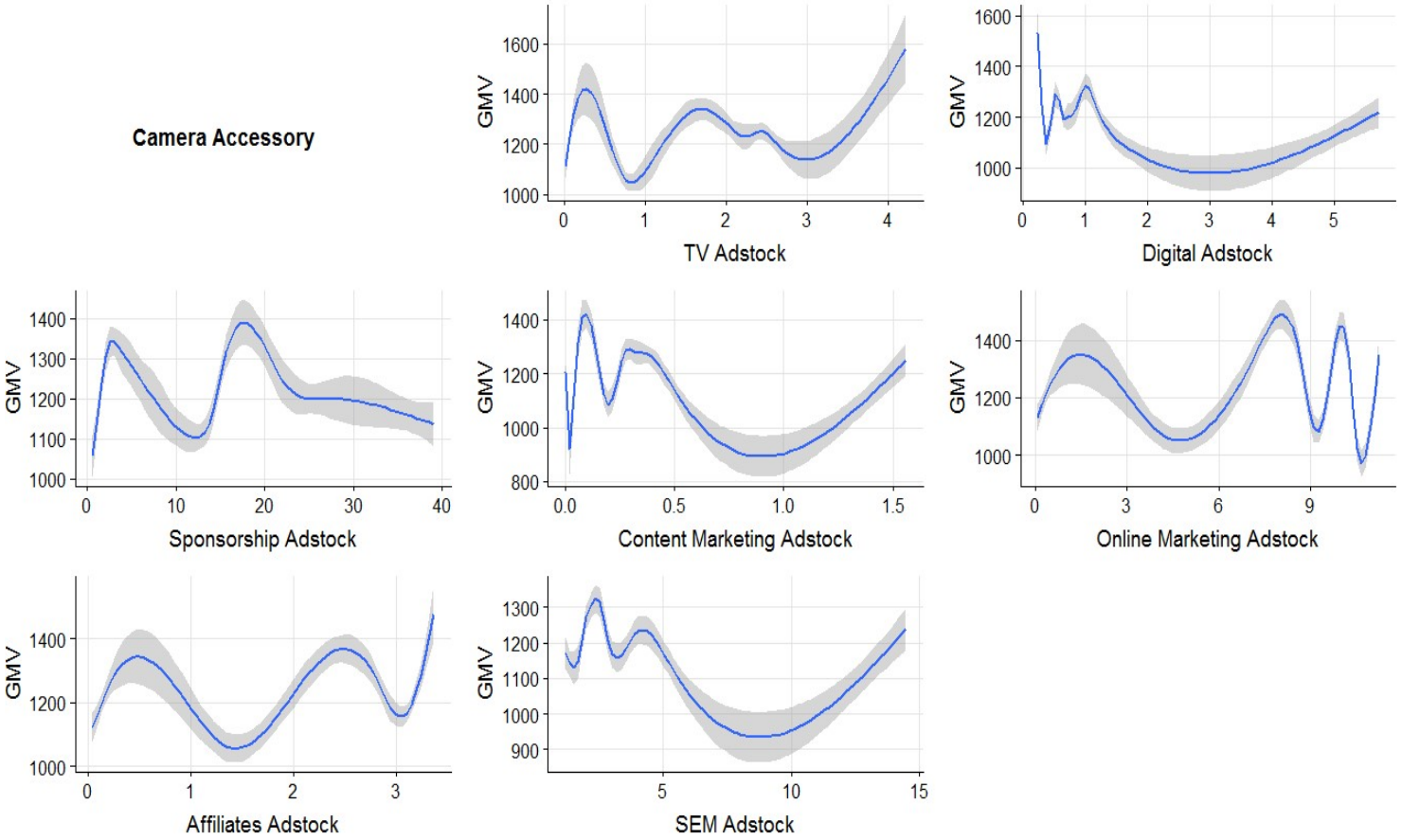




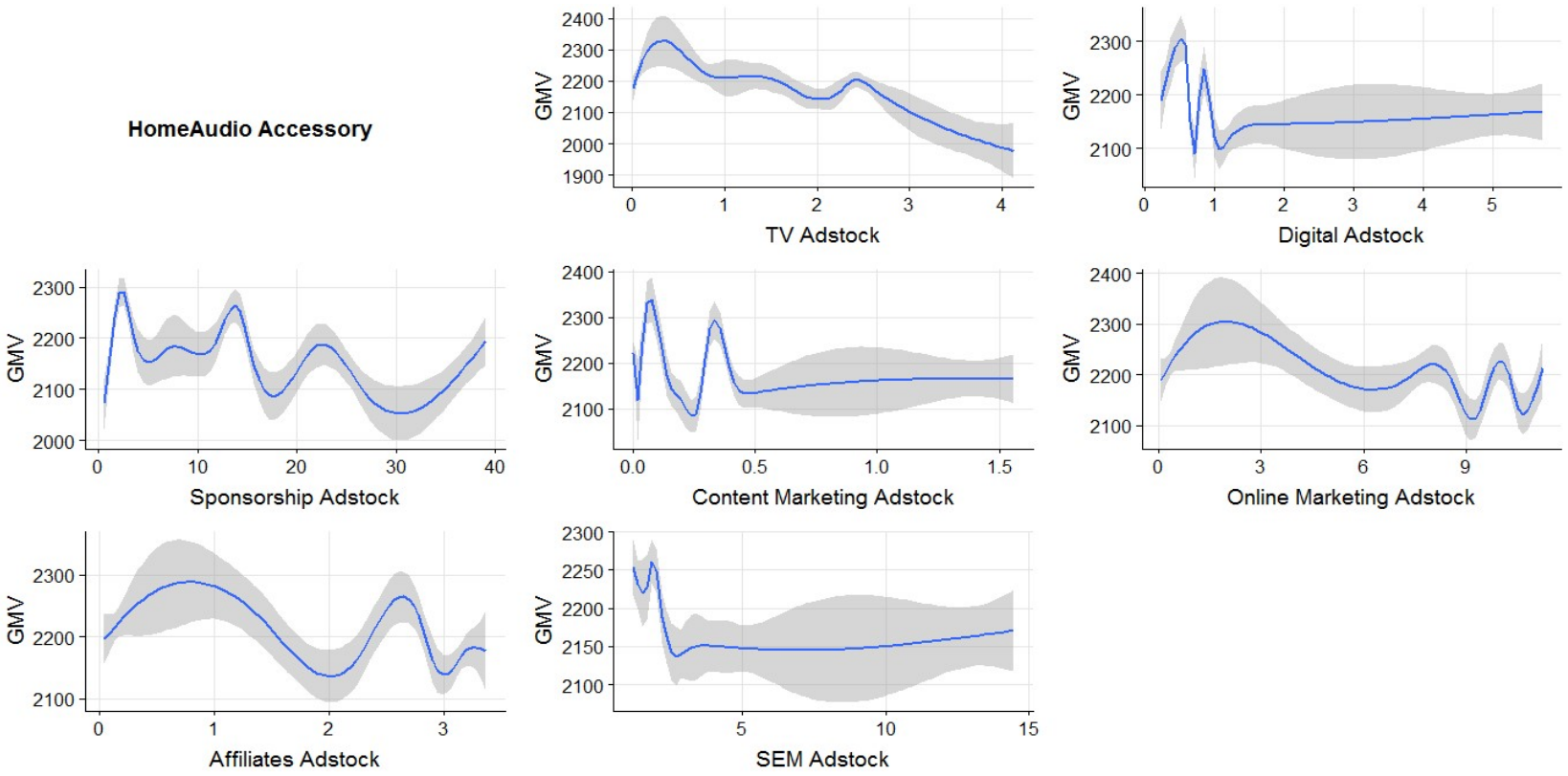




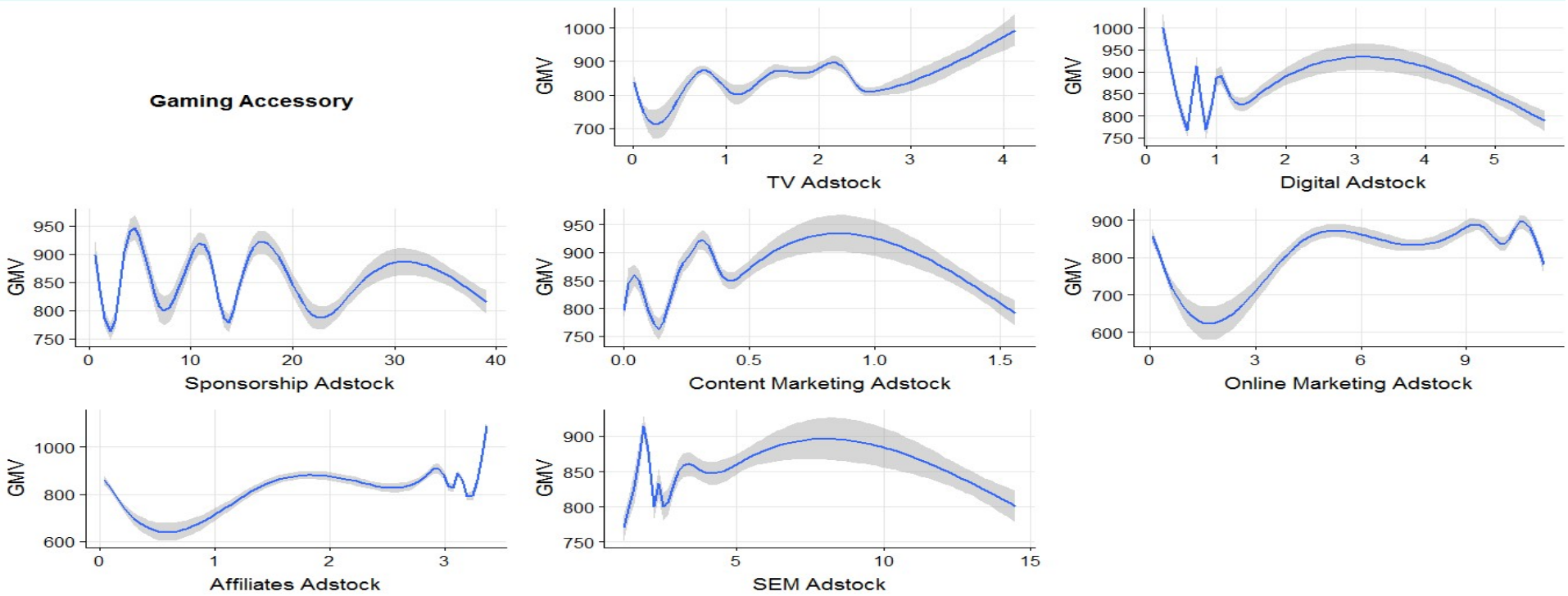
Impact of Adstock on Sales for Camera Accessory



Impact of Adstock on Sales for HomeAudio Accessories



Impact of Adstock on Sales for Gaming Accessory



Model Building

Following models are built to facilitate marketing team to make strategic decisions on their media spends:-

- Simple linear Model- Build the Basic Linear Model with all the KPI
- Multiplicative Model- Build the multiplicative model using the log of the individual KPIs
- Koyck Model- Build the Koyck model using the lag of the dependent variable
- Distributed Lag Model-Build the distributed lag model using the past lags of both the dependent and the independent variables

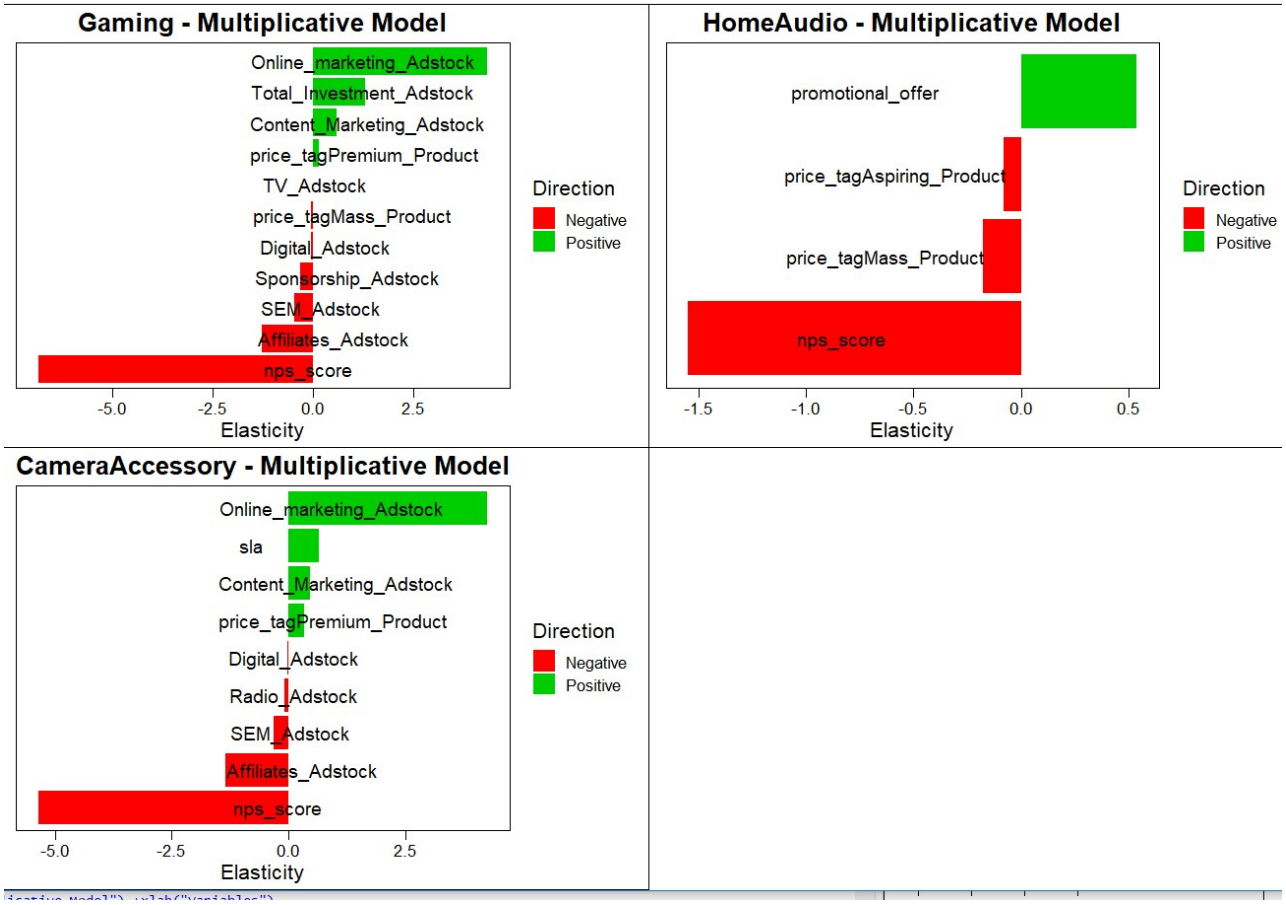
Linear Model for Camera Linear Model, HomeAudio and Gaming

```
Camera Linear Model: lm(formula = gmv ~ deliverycdays + promotional_offer + sla +
    product_procurement_sla + is_special_week + nps_score + Digital_Adstock +
    Online_marketing_Adstock + SEM_Adstock + price_tagMass_Product,
    data = lm_data_camera_1)
Home Audio Linear Model: lm(formula = gmv ~ promotional_offer + is_special_week + SEM_Adstock +
    price_tagMass_Product, data = lm_data_homeaudio_1)
Gaming  Linear Model: lm(formula = gmv ~ deliverycdays + is_special_week + nps_score +
    price_tagAspiring_Product + price_tagMass_Product, data = lm_data_gaming_1)

================================================================================================================
                                Camera Linear Model        Home Audio Linear Model       Gaming  Linear Model
----------------------------------------------------------------------------------------------------------------
  (Intercept)                2463062.446***  (74826.480)  1672187.654*** (107898.289)  1000513.195*** (36560.100)
  deliverycdays               434283.506*** (106245.832)                                196835.166*** (38480.544)
  promotional_offer           723986.950*** (171902.170)   624530.220*** (133586.790)
  sla                         339068.797***  (89563.259)
  product_procurement_sla     540118.282*** (115679.338)
  is_special_week             353133.424***  (85073.142)   393887.897*** (110223.508)   169777.805*** (37968.267)
  nps_score                  -676874.220*** (193158.301)                               -309988.913*** (37392.653)
  Digital_Adstock            1743190.818**   (607824.496)
  Online_marketing_Adstock    434007.380**   (157946.971)
  SEM_Adstock                -2010264.978**  (678740.723)   447306.371*** (109392.493)
  price_tagMass_Product       -987582.214*** (172475.844)  1590227.128*** (132748.239)   938510.436*** (43366.777)
  price_tagAspiring_Product                                                             416714.718*** (43173.790)
----------------------------------------------------------------------------------------------------------------
  R-squared                          0.589                        0.738                        0.782
  sigma                         755710.401                   1294779.472                  452223.464
  F                                 13.022                       97.688                      105.460
  adj. R-squared                     0.543                        0.730                        0.775
================================================================================================================
```

Multiplicative Model for Camera Linear Model, HomeAudio and Gaming

```
Calls:
Camera Multiplicative Model: lm(formula = gmv ~ sla + nps_score + Digital_Adstock + Content_Marketing_Adstock +
    Online_marketing_Adstock + Affiliates_Adstock + SEM_Adstock +
    Radio_Adstock + price_tagPremium_Product, data = data_CameraAccessory_Multi_1)
Home Audio Multiplicative Model: lm(formula = gmv ~ promotional_offer + price_tagMass_Product +
    nps_score + price_tagAspiring_Product, data = data_HomeAudio_Multi_2)
Gaming  Multiplicative Model: lm(formula = gmv ~ nps_score + Total_Investment_Adstock + TV_Adstock +
    Digital_Adstock + Sponsorship_Adstock + Content_Marketing_Adstock +
    Online_marketing_Adstock + Affiliates_Adstock + SEM_Adstock +
    price_tagMass_Product + price_tagPremium_Product, data = data_Gaming_Multi_2)

=============================================================================================================
                          Camera Multiplicative Model  Home Audio Multiplicative Model  Gaming  Multiplicative Model
-------------------------------------------------------------------------------------------------------------
  (Intercept)                    34.988*   (14.685)           29.383*** (4.930)            47.081*** (13.546)
  sla                             5.125*** (0.557)
  nps_score                     -19.701*** (3.694)           -5.163*** (1.139)           -22.441***  (3.436)
  Digital_Adstock                 3.303*** (0.408)                                         3.677***  (0.422)
  Content_Marketing_Adstock      -2.929*** (0.276)                                        -3.129***  (0.267)
  Online_marketing_Adstock       34.865*** (3.643)                                        32.525***  (3.761)
  Affiliates_Adstock            -30.525*** (3.459)                                       -27.803***  (3.827)
  SEM_Adstock                    -4.205*** (0.573)                                        -5.593***  (0.686)
  Radio_Adstock                   0.603*** (0.124)
  price_tagPremium_Product       -2.077*** (0.343)                                        -1.106***  (0.069)
  promotional_offer                                  1.992*** (0.423)
  price_tagMass_Product                              1.526*** (0.093)                     0.363***  (0.067)
  price_tagAspiring_Product                          0.677*** (0.079)
  Total_Investment_Adstock                                                                4.996***  (0.908)
  TV_Adstock                                                                             -1.360***  (0.400)
  Sponsorship_Adstock                                                                    -1.743***  (0.300)
-------------------------------------------------------------------------------------------------------------
  R-squared                       0.826                    0.825                        0.860
  sigma                           0.760                    0.817                        0.787
  F                              48.579                  163.564                       78.494
  adj. R-squared                  0.809                    0.820                        0.849
=============================================================================================================
>
```

Elasticity of the different variables w.r.t the overall sales for Multiplicative Model



- Thus going with the Multiplicative Model,

I. For gaming Accessory ElecKart should focus more on the Online channel ,Content Marketing and decrease its spending on Sponsorships w.r.t the sales.

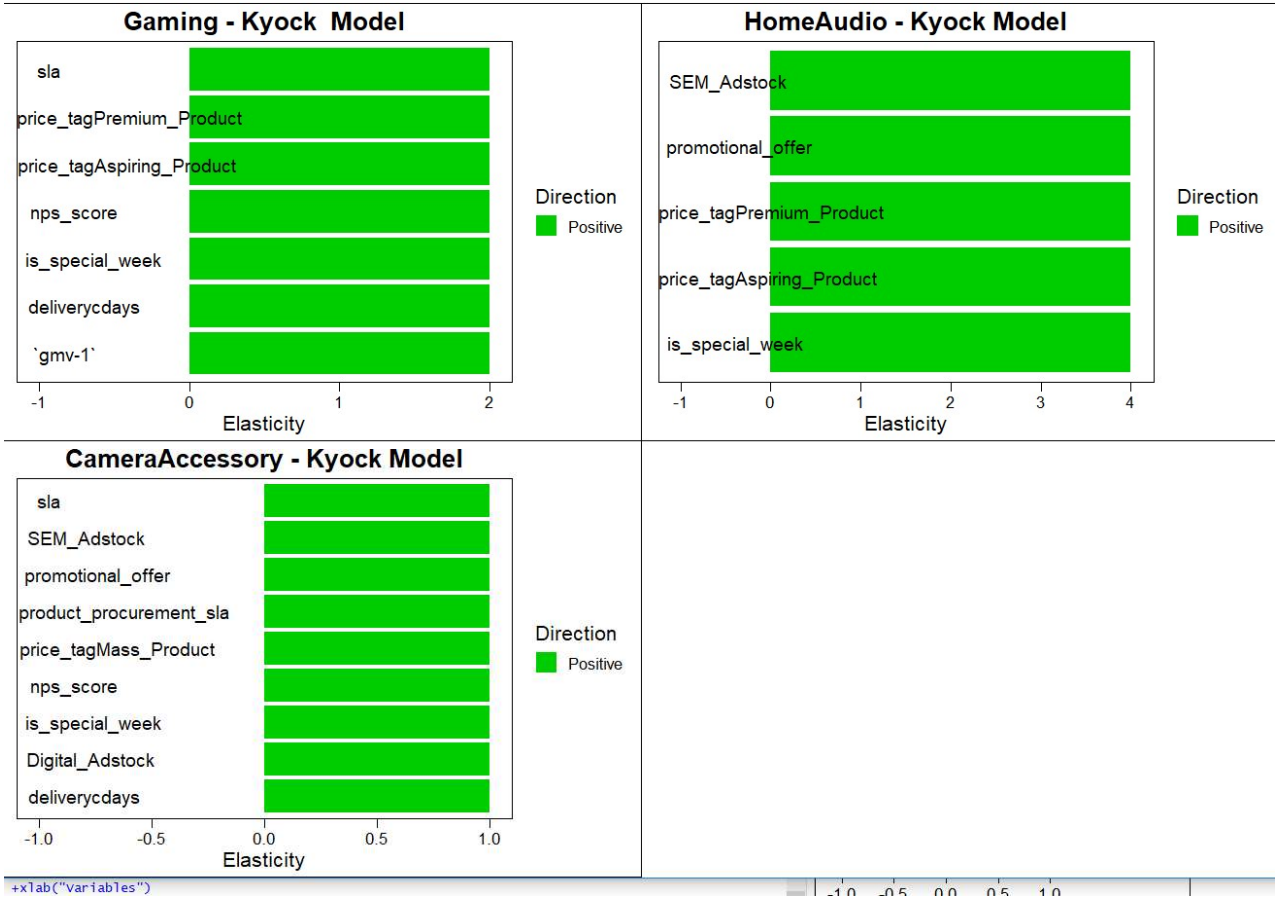
II. For Home Accessory ElecKart should focus more on the Promotion

III. For Camera Accessory ElecKart should focus more on the Online channel ,Content Marketing and decrease its spending on Radio channel, Affiliates w.r.t the sales.

Koyck Model

```
Calls:
Camera Kyock Model: lm(formula = gmv ~ deliverycdays + promotional_offer + sla +
    product_procurement_sla + is_special_week + nps_score + Digital_Adstock +
    SEM_Adstock + price_tagMass_Product, data = Kyock_data_camera_1)
Home Audio Kyock Model: lm(formula = gmv ~ promotional_offer + is_special_week + SEM_Adstock +
    price_tagAspiring_Product + price_tagPremium_Product, data = Kyock_data_homeaudio_1)
Gaming  Kyock  Model: lm(formula = gmv ~ deliverycdays + sla + is_special_week + nps_score +
    `gmv-1` + price_tagPremium_Product + price_tagAspiring_Product,
    data = Kyock_data_gaming_1)

================================================================================================================
                                Camera Kyock Model          Home Audio Kyock Model        Gaming  Kyock  Model
----------------------------------------------------------------------------------------------------------------
  (Intercept)                2463062.446***  (77444.514)   1672187.654*** (108023.684)    637772.913*** (95616.346)
  deliverycdays               430231.080*** (109952.575)                                  166287.015*** (39550.584)
  promotional_offer           644570.746*** (175384.060)    668719.538*** (144115.282)
  sla                         326218.357***  (92570.450)                                  112336.126**  (40840.011)
  product_procurement_sla     431131.709*** (112469.178)
  is_special_week             412691.589***  (85143.995)    393231.337*** (110354.488)    123041.258**  (38382.067)
  nps_score                  -926812.636*** (176363.097)                                 -259928.092*** (37137.489)
  Digital_Adstock            1210336.085*   (596215.640)
  SEM_Adstock               -1433652.573*   (668071.885)    449409.555*** (109549.428)
  price_tagMass_Product      -887870.873*** (174514.838)
  price_tagAspiring_Product                               -1611324.188*** (136792.398)   -436134.016*** (57811.298)
  price_tagPremium_Product                                -1465775.252*** (163635.254)  -1168779.486*** (65932.432)
  `gmv-1`                                                                                      0.363***     (0.089)
----------------------------------------------------------------------------------------------------------------
  R-squared                       0.555                       0.739                        0.814
  sigma                      782151.253                  1296284.209                  420996.469
  F                              12.724                      78.105                       90.435
  adj. R-squared                  0.511                       0.729                        0.805
================================================================================================================
> |
```

Elasticity for Koyck Model



- Positive elasticity means that increasing the value of the KPI would lead to increase in the sales figure
- Thus going with the Kyock Model

I. For gaming Accessory ElecKart should focus more on Special Weeks

II. For Home Accessory ElecKart should focus more on the Promotion, Special Weeks

III. For Camera Accessory ElecKart should focus more on the Promotion, Special Weeks,Digital Marketing,Product Procurement SLA

Camera Accessories

- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 5 fold cross validation again on the training data
- Multiplicative model has a decent Adj. R sq figure, but in other three models error is more due to high SSE value and less add R.square value

- So, we select the Multiplicative linear model on the account of higher Adj. R Sq values

Model	Variables	Adj. R Square	Cross Validation
Linear Model	deliverycdays+promotional_offer+sla+product_procurement_sla+is_special_week+nps_score+Digital_adstock+Online_marketing_adstock+SEM_adstock +Price_tagMassProduct+Price_tagAspiringProduct	0.589	2938108
Multiplicative Model	sla+nps_score+digital_adstock+Content_Marketing_adstock+Online_marketing_adstock+Affiliates_adstock + SEM_adstock + Radio_adstock+Price_tagPremiumProduct	0.826	0.29
Koyck Model	deliverycdays+promotional_offer+sla+product_procurement_sla+is_special_week+nps_score+Digital_adstock+ SEM_adstock +Price_tagMassProduct	0.55	741644
Distributed Lag Model	**lm(formula = gmv ~ deliverycdays + prepaid_percentage + promotional_offer + sla + product_procurement_sla + deliverybdays + nps_score + `deliverycdays-3` + `promotional_offer-3` + `sla-3` + `product_procurement_sla-2` + `product_procurement_sla-3` + `deliverybdays-2` + `gmv-3`, data = data_dlag_camera)**	0.60	52891

Home Audio Accessories

- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 5 fold cross validation again on the training data
- Multiplicative model has a decent Adj. R square figure, but in other three models error is more due to high SSE value and less add R.square value

- So, we select the Multiplicative linear model on the account of higher Adj. R Square values

Model	Variables	Adj. R Square	Cross Validation
Linear Model	promotional_offer+is_special_week+SEM_adstock +Price_tagMassProduct	0.738	4581898
Multiplicative Model	nps_score+promotional_offer +Price_tagMassProduct+Price_tagAspiringProduct	0.825	0.72
Koyck Model	promotional_offer+is_special_week+SEM_adstock +Price_tagMassProduct+Price_tagAspiringProduct	0.739	900556
Distributed Lag Model	lm(formula = gmv ~ deliverycdays + promotional_offer + product_procurement_sla + deliverybdays + nps_score + Digital_Adstock + Online_marketing_Adstock + SEM_Adstock + Radio_Adstock + `deliverycdays-1` + `deliverycdays-3` + `prepaid_percentage-1` + `promotional_offer-2` + `sla-1` + `sla-3` + `product_procurement_sla-1` + `deliverybdays-1` + `deliverybdays-2` + `deliverybdays-3` + `gmv-1` + `gmv-2` + `gmv-3`, data = data_dlag_home_audio)	0.74	780012

Gaming Accessories

- Adjusted R Square figures are based on the performance of the model on the training data.
- The SSE figures are based on the 5 fold cross validation again on the training data
- Multiplicative model has a decent Adj. R sq figure, but in other three models error is more due to high SSE value and less add R.square value

- So, we select the Multiplicative linear model on the account of higher Adj. R Square values

Model	Variables	Adj. R Square	Cross Validation 5 fold
Linear Model	deliverycdays+is_special_week+nps_score+Price_tagMassProduct+Price_tagAspiringProduct	0.782	2507662
Multiplicative Model	nps_score+digital_adstock+Content_Marketing_adstock+Online_marketing_adstock+Affiliates_adstock + SEM_adstock+Price_tagPremiumProduct +Price_tagAspiringProduct+TV_adstock +Total Investment_adstock+ TV_adstock	0.860	0.4
Koyck Model	deliverycdays+sla+is_special_week+nps_score+Price_tagMassProduct+Price_tagAspiringProduct+gmv-1	0.814	455814
Distributed Model	lm(gmv ~ prepaid_percentage + sla + deliverybdays + nps_score + is_special_week + TV_Adstock + Digital_Adstock + Sponsorship_Adstock + `prepaid_percentage-2` + `prepaid_percentage-3` + `promotional_offer-1` + `promotional_offer-2` + `promotional_offer-3` + `sla-1` + `sla-2` + `sla-3` + `product_procurement_sla-1` + `product_procurement_sla-2` + `deliverybdays-3` + `is_special_week-2` + `is_special_week-3` + `nps_score-3` + `gmv-1` + `gmv-3`,data_dlag_GamingAccessory)	0.5924	1888542

Then what should we do to improve Sale 3X?

Multiplicative model is performing well in all three categories . So we choose multiplicative model elasticity for improving their sales.

Thus going with the Multiplicative Model,

I. For gaming Accessory ElecKart should focus more on the Online channel ,Content Marketing and decrease its spending on Sponsorships w.r.t the sales.

II. For Home Accessory ElecKart should focus more on the Promotion.

III. For Camera Accessory ElecKart should focus more on the Online channel ,Content Marketing and decrease its spending on Radio channel, Affiliates w.r.t the sales.

Thank You