

Capstone - Ecommerce

By



Group :

1. Sathish Devunuri
2. Vidhi Thakkar
3. Shalini Bulusu
4. Varsha Atmakuri





ElecKart is an e-commerce firm specialising in electronic products. Over the last one year, they had spent a significant amount of money in 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 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.



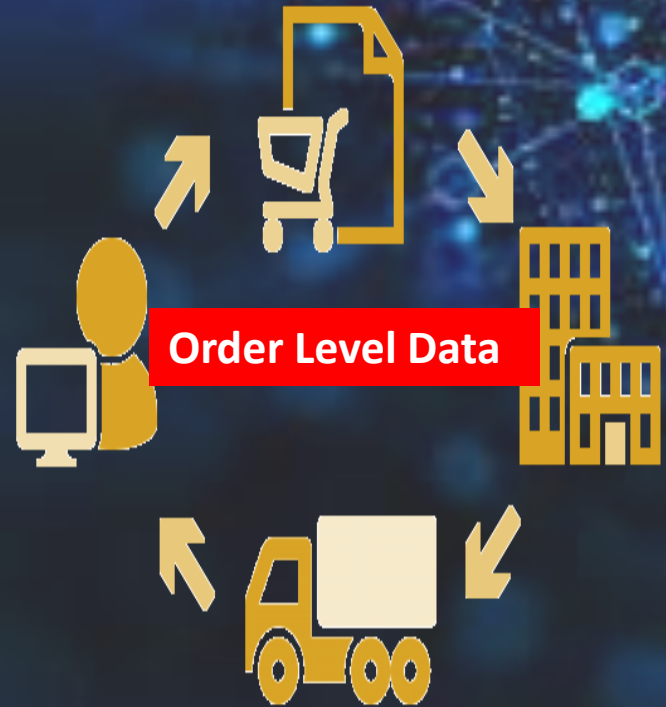
Develop a market mix model to observe the actual impact of different marketing variables over the last year.

Recommend the optimal budget allocation for different marketing levers for the next year.

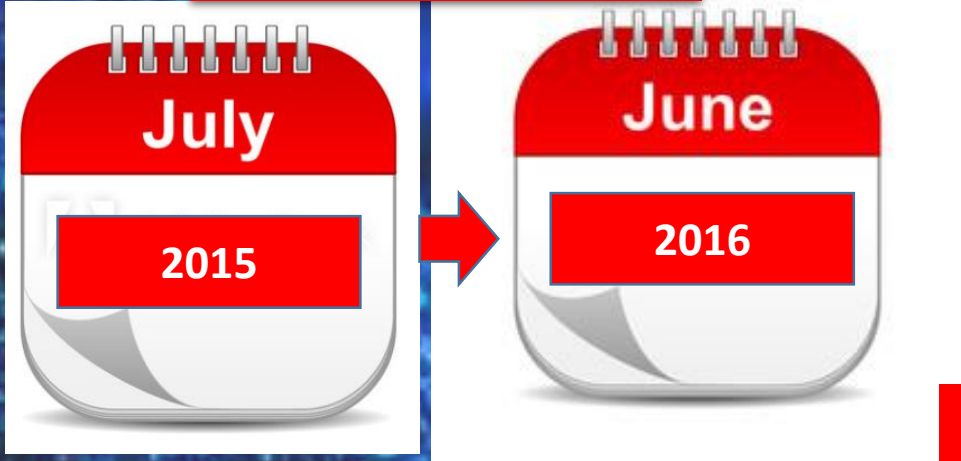




Source Data



Data Range



UpGrad

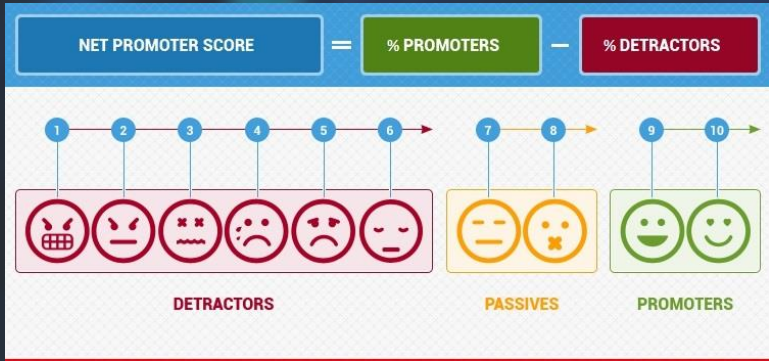
Days when there was any special sale

Monthly spends on various advertising channels



Display	Camera	Hardware	Battery	Phone features
5.5"	12 MP	4GB RAM	3600mAh	Released 2015, March
Optical image stabilization, Autoclick (Phase detection)				

Product Details



Monthly NPS score – this may work as a proxy to ‘voice of customer’



- Order Level Data
- Monthly Spend
- Special Sale
- Monthly NPS

July-2015 to June-2016

Weekly Σ

Data Prep

- NA Checks
- Missing Values
- Outlier Treatment
- Scaling the Data
- Data Types Conversion
- MRP > 0
- Camera Access. Gaming Access. Home Audio
- GMV 0 to 1

Engineered KPIs

- Selling Price
- Discount
- Cash on Delivery (COD)
- Pre Paid Delivery
- Product Cluster
- Ad Stock
- Lag/ inflation
- NPS - Score
- Moving Average
- Holidays in week
- GMV_Zero

EDA

- Data Grouping for further Analysis
- Create Data Insights
- Pricing Analysis
- Discount Analysis
- Gross Merchandise Value Analysis
- Marketing spend Across Channels
- Product Category and Sub Category Analysis







Subset of Data for Modeling

- Camera Accessory Data
- Gaming Accessory Data
- Home Audio Data

Modeling

- Basic Linear Model
- Multiplicative Model
- Koyck Model
- Distributed Lag Model
- Multiplicative and distributed model
- Significant KPI
- Model Coefficients
- Model Accuracy
- Model Optimization
- Plot the Model results



 <p>Selling Price</p>	<p>GMV(Gross Merchandise Value)/ Units</p>	<p>Selling Price / List Price: Price in which the product was sold and this will be helpful to compare the MRP and further analysis on the sale or offer days</p>
 <p>Discount</p>	<p>$(MRP - \text{Selling Price}) / MRP * 100$</p>	<p>Discount / Promotional Offer : With respect to the MRP on what % of Promotion provided to the customer on a given product</p>
 <p>Cash on Delivery (COD)</p>	<p>COD = 1</p>	<p>If COD =1 the transaction opted as Cash on Delivery</p>
 <p>Pre Paid Delivery</p>	<p>Prepaid=1</p>	<p>If Prepaid =1 the transaction paid online using E-payment</p>
 <p>Product Cluster</p>	<p>Cluster the product analytic vertical based on Selling Price, Units and MRP using Kmeans algorithm</p>	<p>Cluster divides the products into three categories based on MRP and number product units sold ; mass market, medium market and premium product are considered as 3 clusters</p>
 <p>Ad Stock</p>	<p><code>stats::filter(x=Media_Investment[i], filter=adstock_rate, method="recursive");</code></p>	<p>Calculation of adstock_rate for all Media modes in data set using stats::filter function and filter=adstock_rate, method="recursive");</p>



Lag/ inflation

For Discount and Selling Price, we add variables of inflation.
(How much Discount has varied within the past 3 weeks)
For Holidays, we see if upcoming 2 weeks have holidays in it.

Using Slide function

Discount at 3 levels (-1 ,-2,-3) , Selling Price at 3 leaves (-1,-2,-3) and IsHolidayWeek at 2 levels (1,2) –then derive the inflation using combinations of this levels



NPS - Score

NPS Score at Weekly

NPS Score provided at Month and we have considered the same number if the week dates fall in same month ; if the dates split across we have considered based on number of days by month



Holidays in week

Number of Sale Days / Holidays in a particular sale week

Based on the Sale Days / holidays data set we have counted the number of days (count) fall in a sale week of 1: 53



GMV_Zero

GMV = 0

Product sold free ; A product with a >0 MRP but the GMV =0 considered as free product during the sale / holiday period



Moving Average

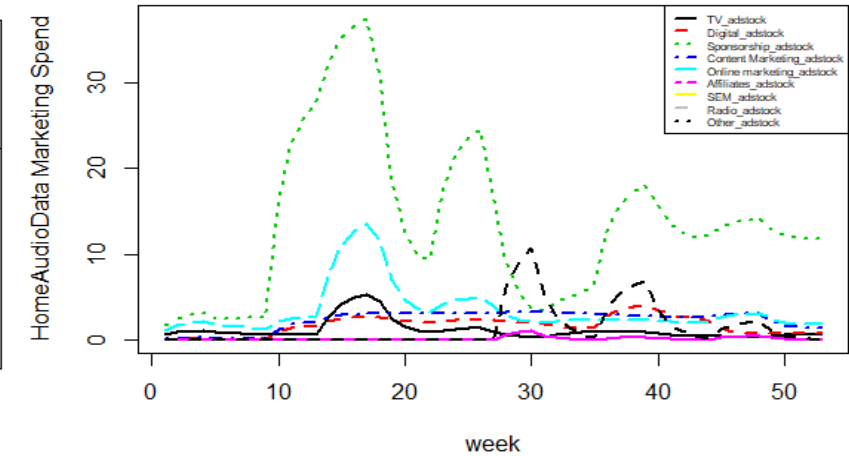
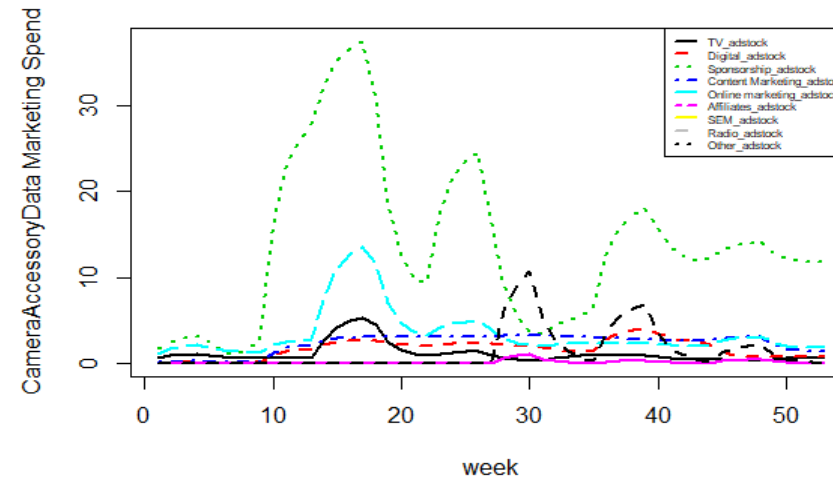
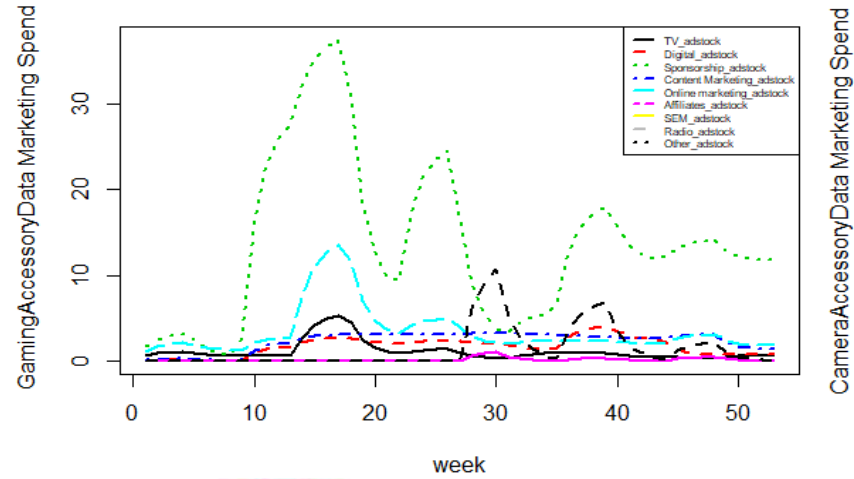
Moving Average for Selling Price and Discount

Moving average using movavg function for both Selling Price and Discount to see if the there is any variation in the average for last 3 months for a given product / SKU

Gaming Accessory Data

Camera Accessory Data

Home Audio Accessory Data



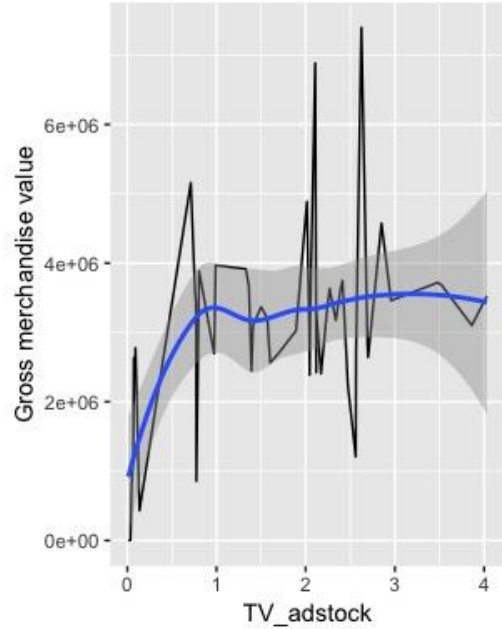
A large portion of marketing spend goes towards sponsorships followed by online
*Same trend was observed for all Categories identified



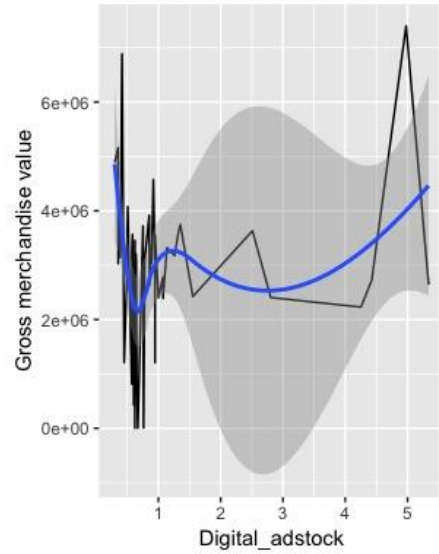
During Week 10 to 20 : Observed a significant Marketing spend

During Week 30: Observed a increase in “other type” of Marketing Spend while all other types of marketing channels observed as dropped down

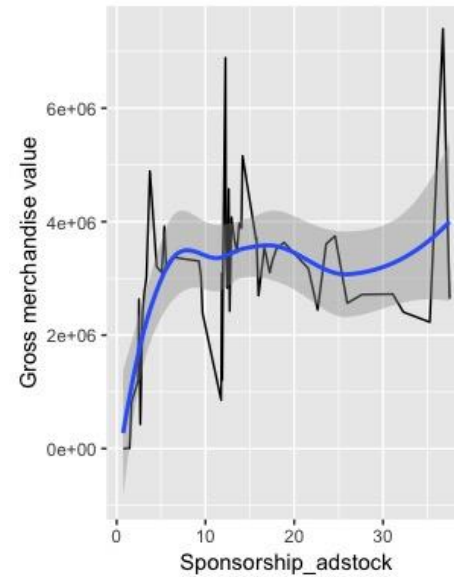
TV Adstock and GMV



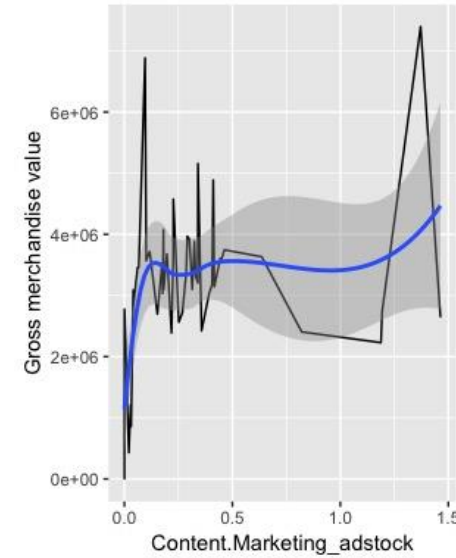
Digital_adstock and GMV



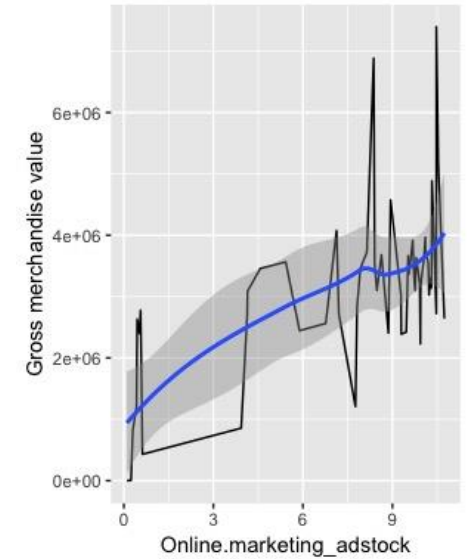
Sponsorship_adstock and GMV



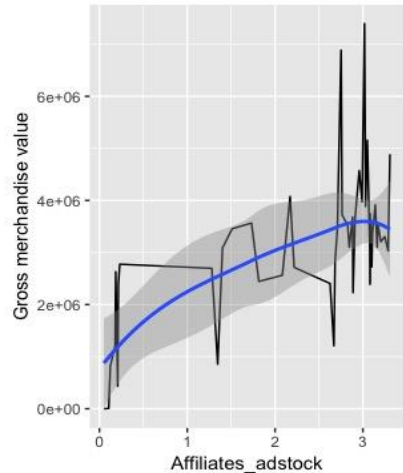
Content.Marketing_adstock and GMV



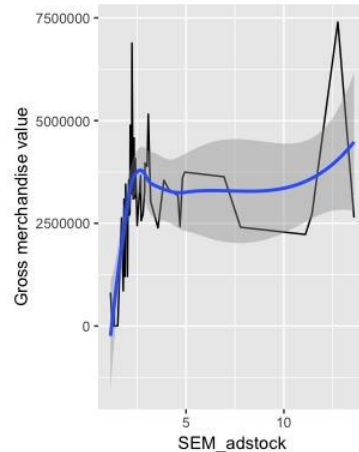
Online.marketing_adstock and GMV



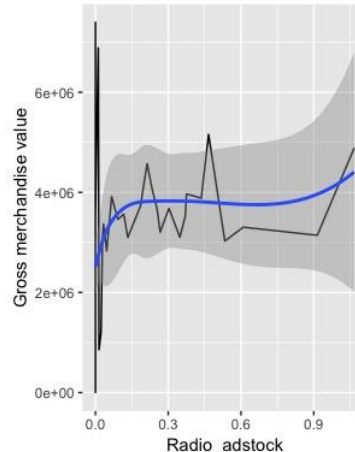
Affiliates_adstock and GMV



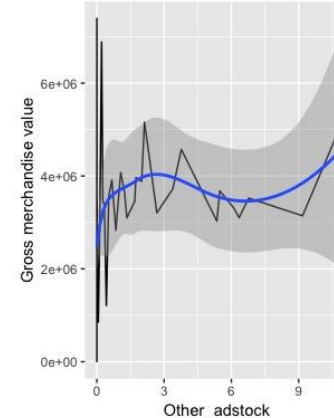
SEM_adstock and GMV



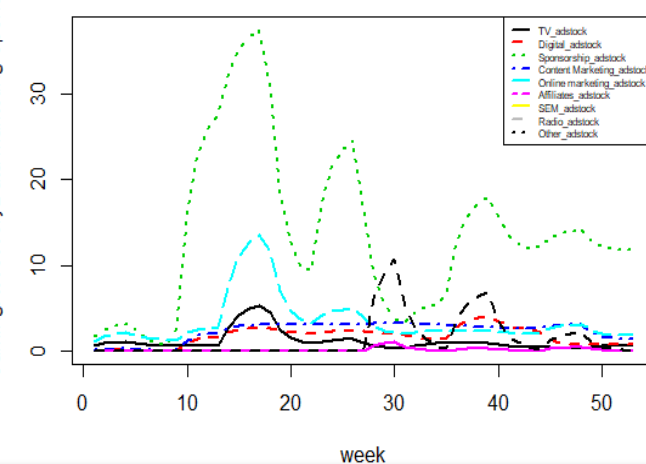
Radio_adstock and GMV



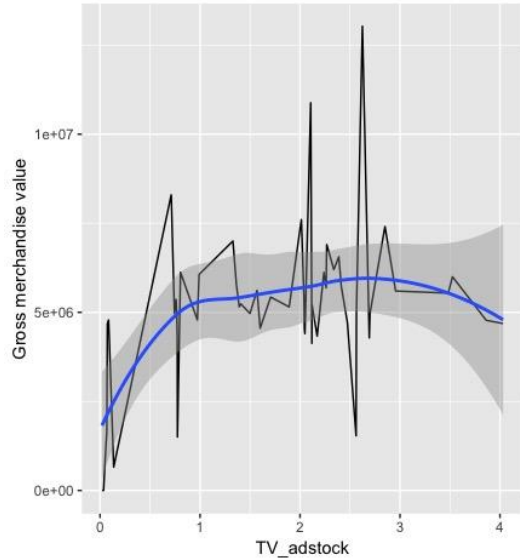
Other_adstock and GMV



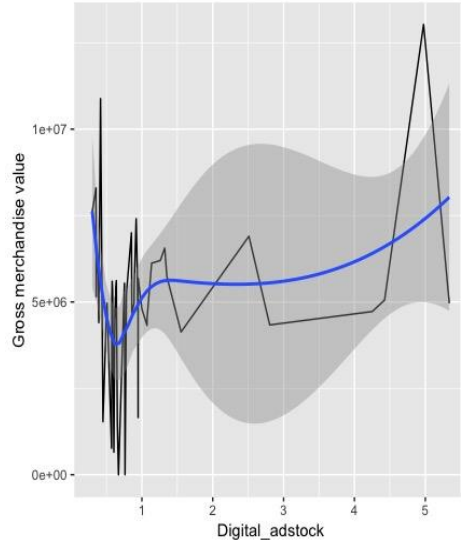
GamingAccessoryData Marketing Spend



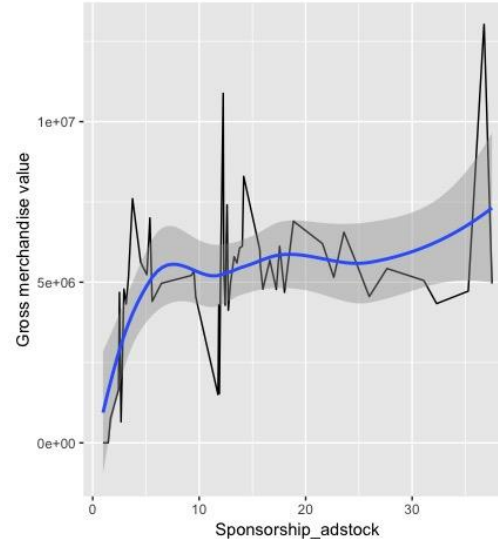
TV Adstock and GMV



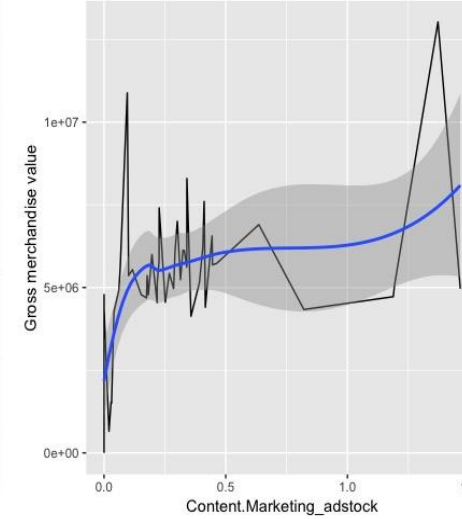
Digital_adstock and GMV



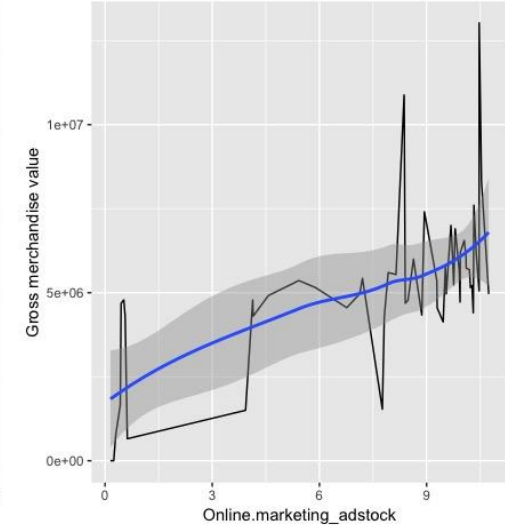
Sponsorship_adstock and GMV



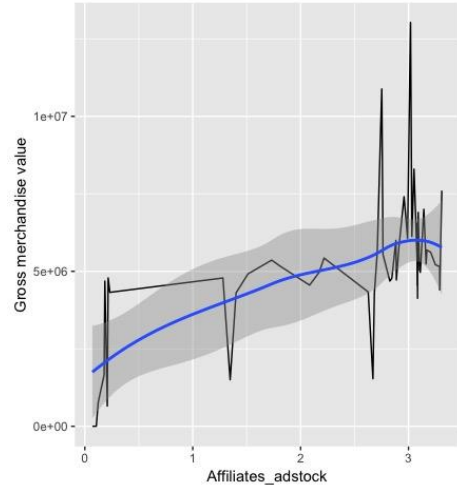
Content.Marketing_adstock and GMV



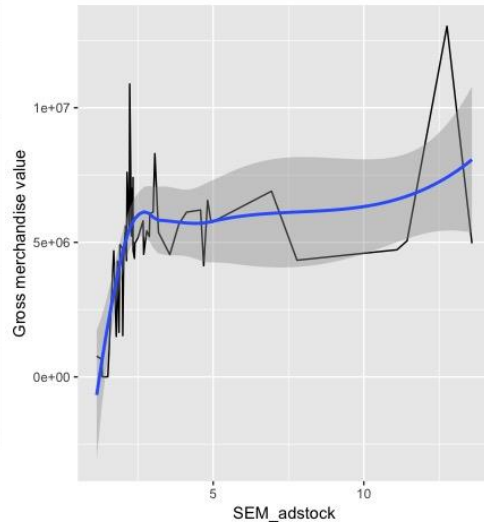
Online.marketing_adstock and GMV



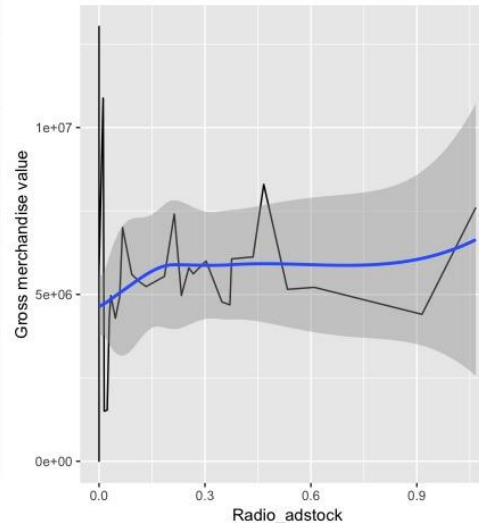
Affiliates_adstock and GMV



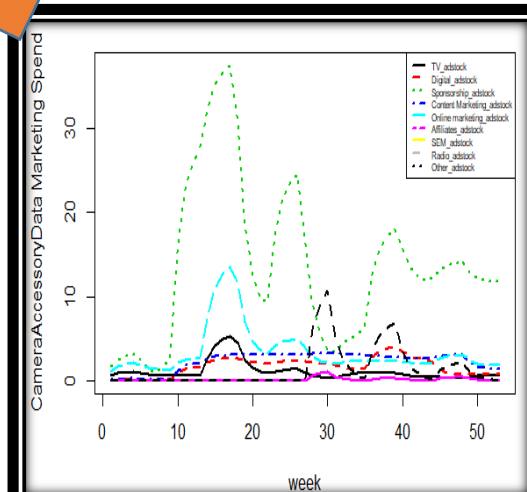
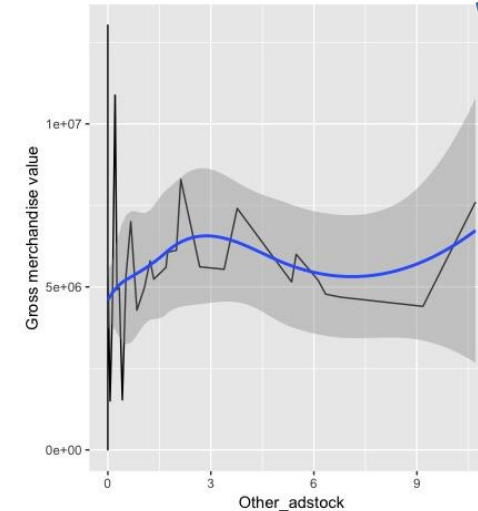
SEM_adstock and GMV



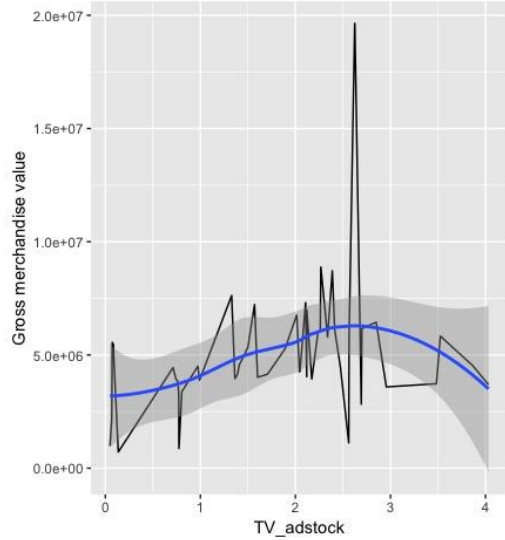
Radio_adstock and GMV



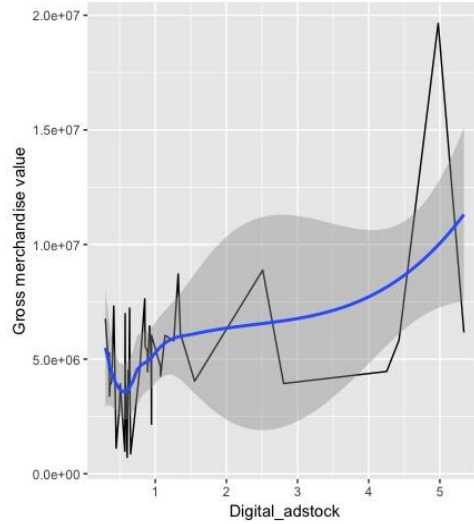
Other_adstock and GMV



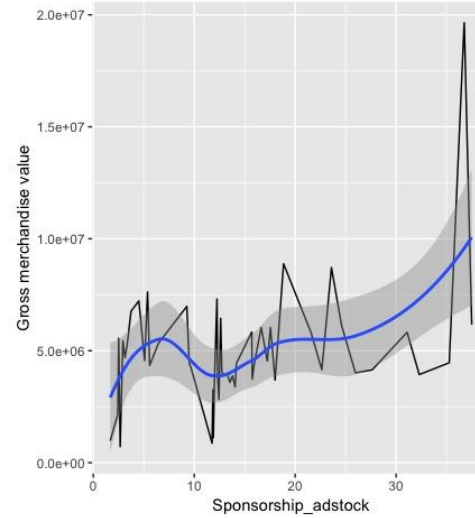
TV Adstock and GMV



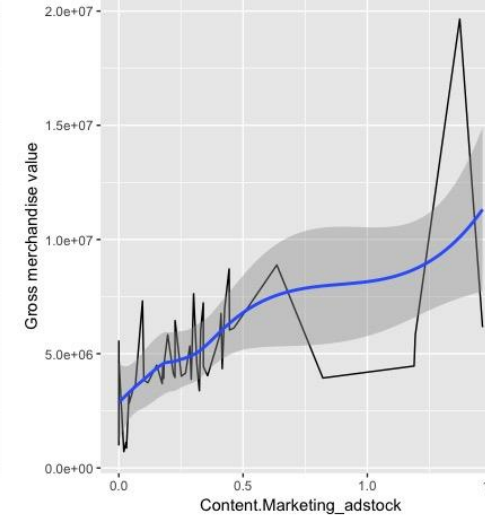
Digital_adstock and GMV



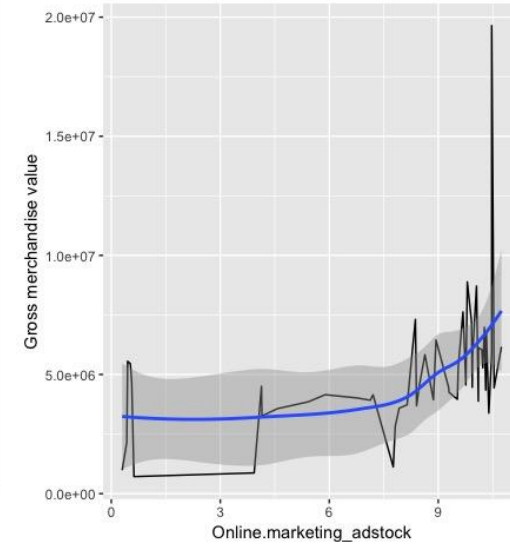
Sponsorship_adstock and GMV



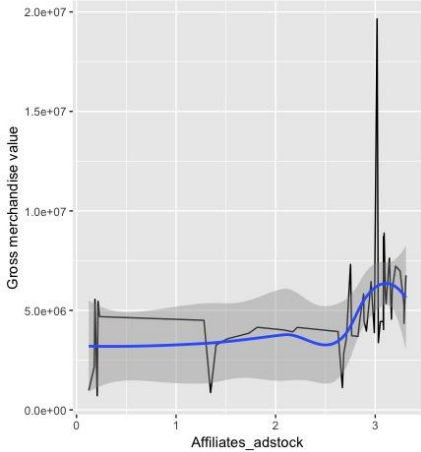
Content.Marketing_adstock and GMV



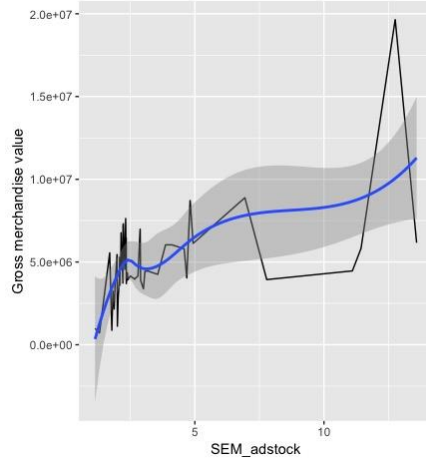
Online.marketing_adstock and GMV



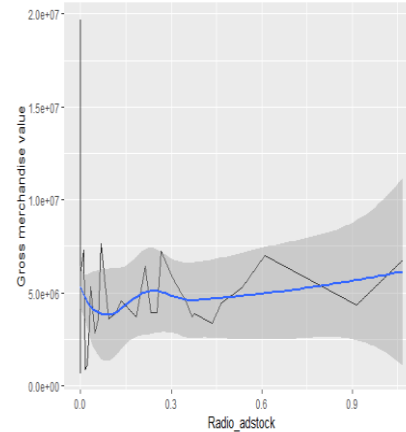
Affiliates_adstock and GMV



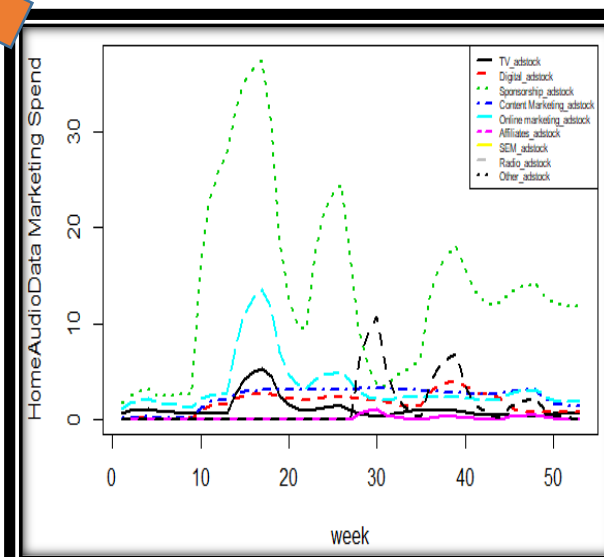
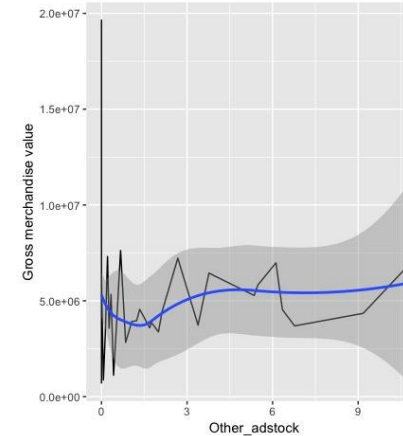
SEM_adstock and GMV



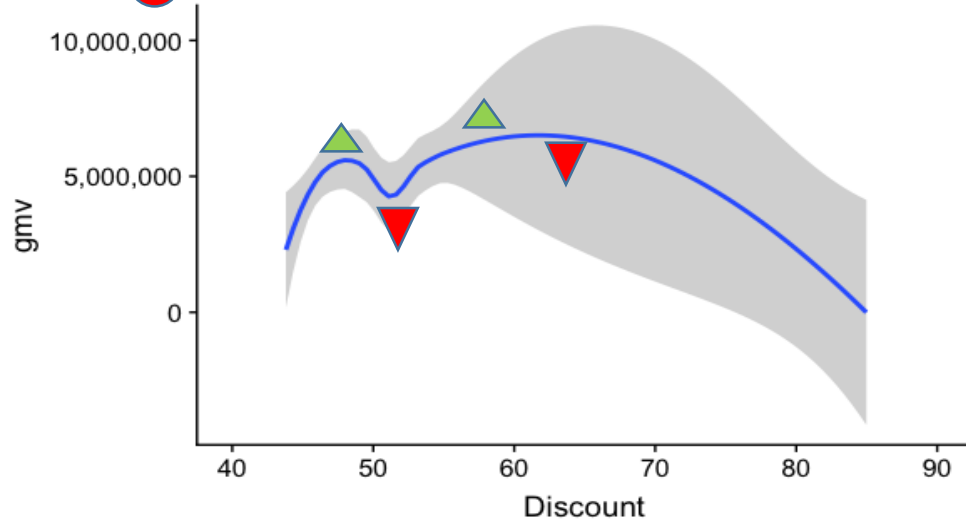
Radio_adstock and GMV



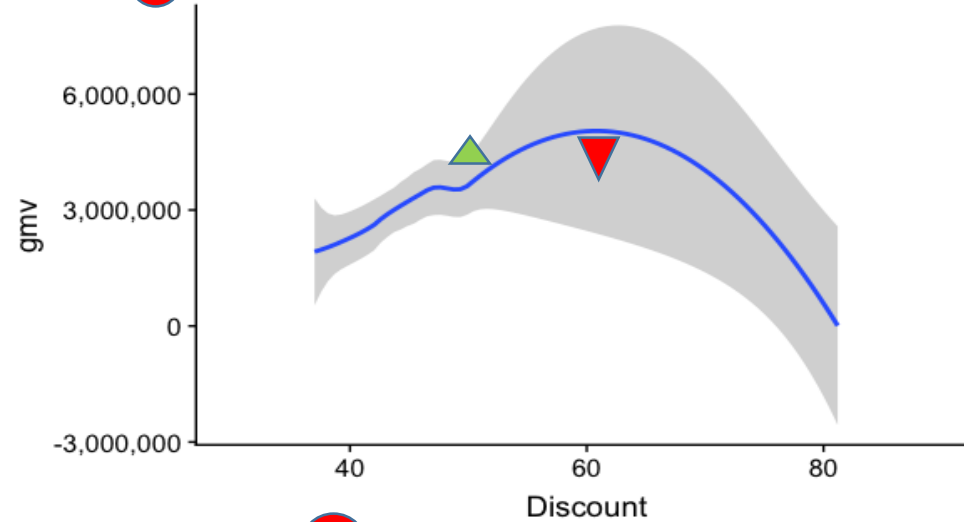
Other_adstock and GMV



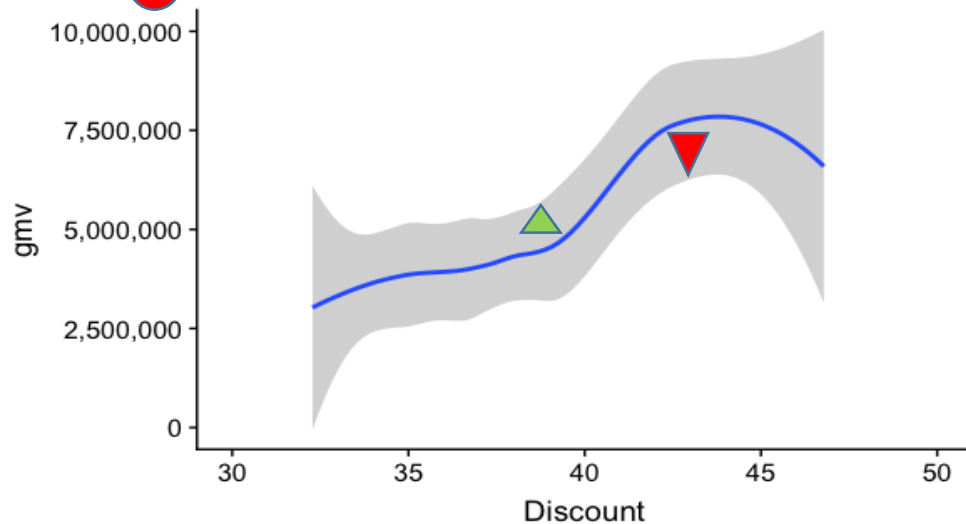
1 Discount vs Revenue/ GMV CameraAccessoryData



2 Discount vs Revenue/ GMV GamingAccessoryData



3 Discount vs Revenue/ GMV HomeAudioData



1

2

Camera Accessory :

1. Discount ranges from 43 % to 85 %
2. GMV increased initially with the increase of discount
3. GMV dropped down at 50% range and a slow raise in GVM up to 60%
4. The GMV started decreasing with increase in Discount after 60%

Gaming Accessory :

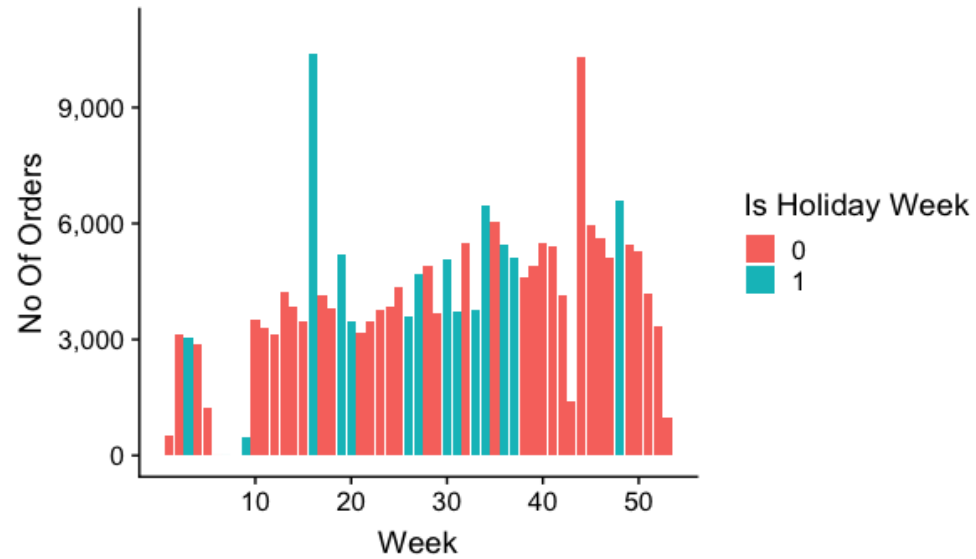
1. Discount ranges from 37 to 81 %
2. GMV increased with the growth in Discount up to 60%
3. The GMV started decreasing with increase in Discount after 60%

3

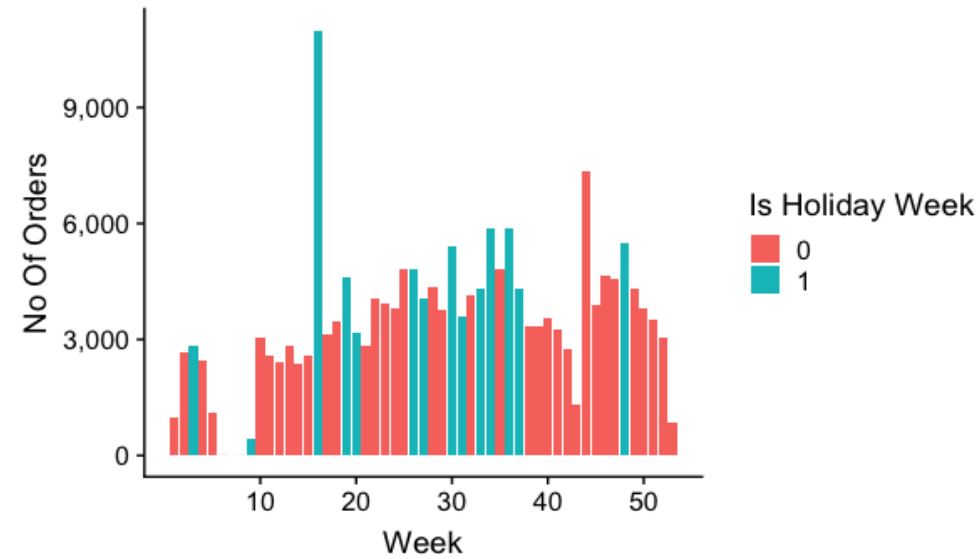
Home Audio Accessory :

1. Discount ranges from 32 to 46%
2. Slow growth in GMV up to 43% and a drop between the 43 to 46 % range
3. Compare to other categories the Sum of Revenue / GMV range was reported low for Home Audio

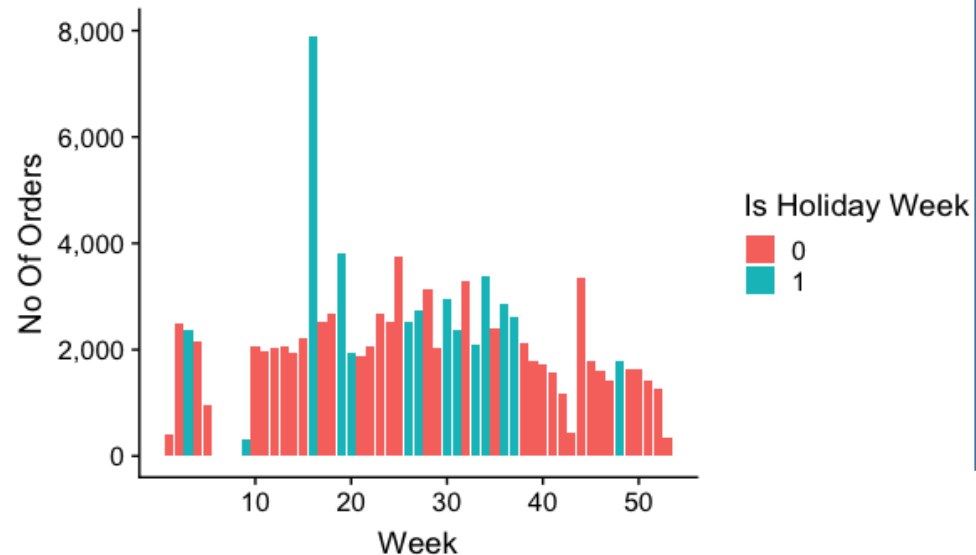
No_of_orders vs Is Holiday Week CameraAccessoryData



No_of_orders vs Is Holiday Week GamingAccessoryData



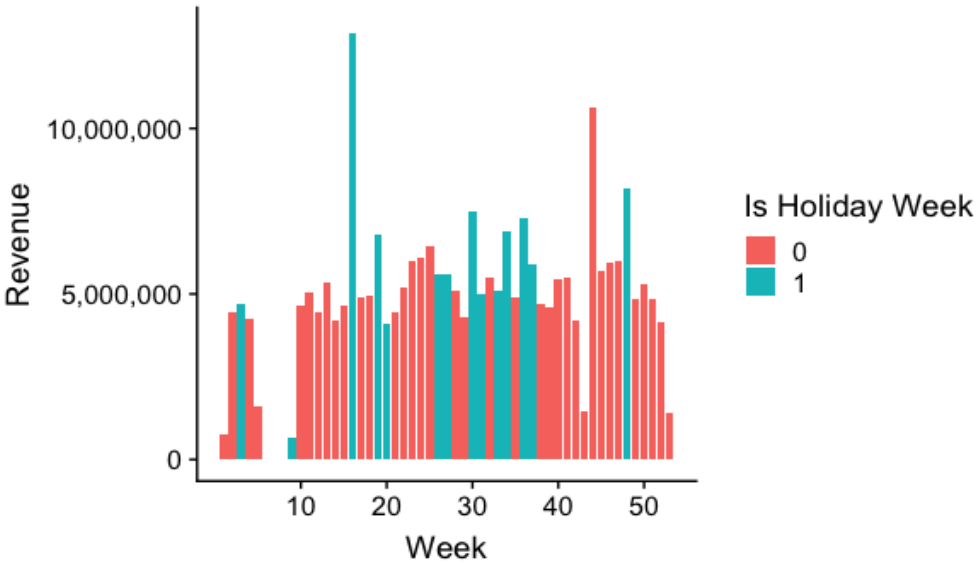
No_of_orders vs Is Holiday Week HomeAudioData



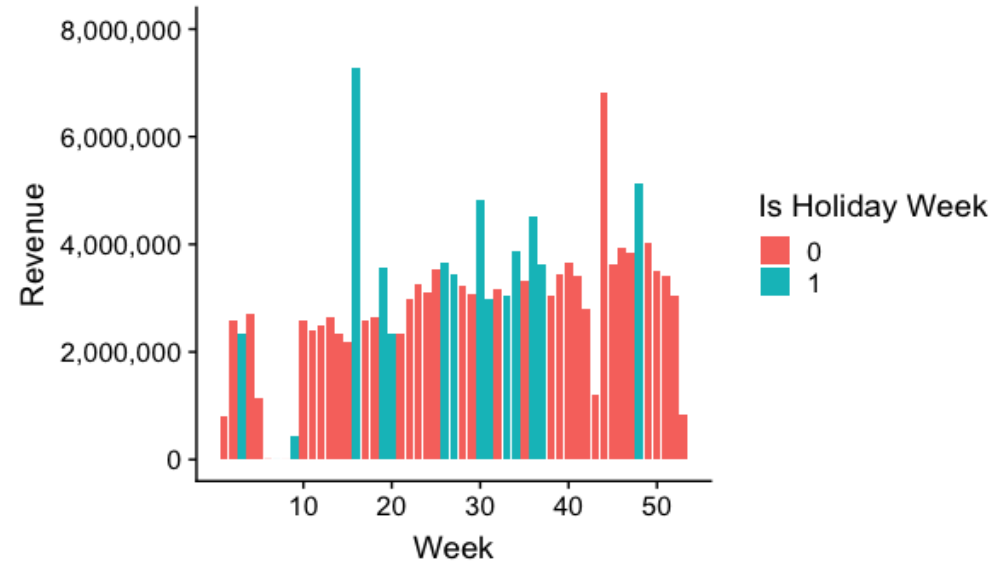
Observations :

1. Week 16 reported highest number of orders for all 3 segments of product category
2. Week preceded or followed by a holiday shows a significant increase in the sales
3. Immediate to a holiday week recorded a low number of transactions / orders
4. Camera Accessory : reported a low transaction during the week 6 and 7
5. Gaming Accessory : reported a low transaction during the week 6 , 7 and 8
6. Home Audio : No transactions / orders reported during the weeks 6 , 7 and 8

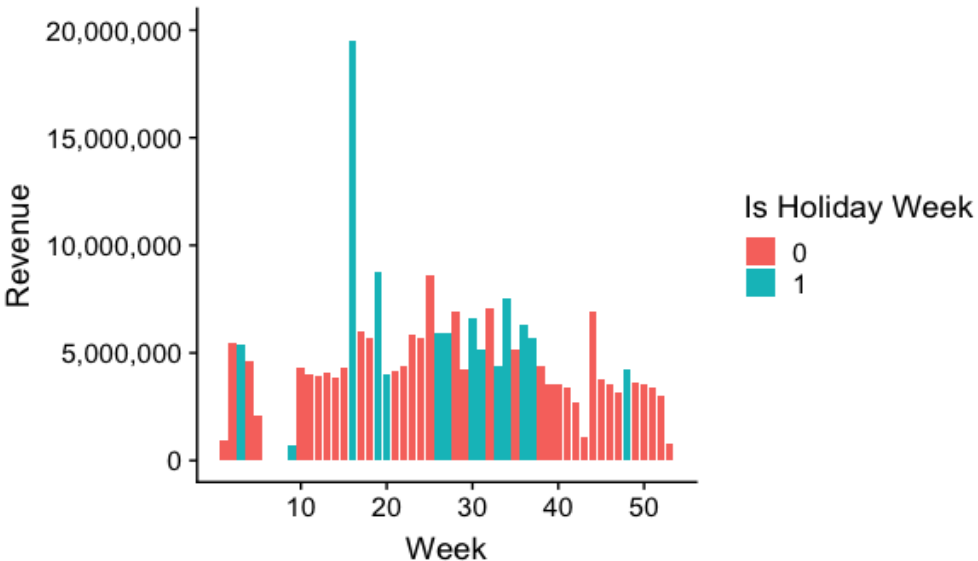
Revenue/GMV vs Is Holiday Week CameraAccessoryData



Revenue/GMV vs Is Holiday Week GamingAccessoryData



Revenue/GMV vs Is Holiday Week HomeAudioData



Observations :

1. Week 16 reported highest Revenue / GMV for all 3 segments of product category
2. Week preceded or followed by a holiday shows a significant increase in the sales
3. Immediate to a holiday week recorded a low Revenue / GMV
4. Camera Accessory : reported a low Revenue during the week 6 and 7
5. Gaming Accessory : reported a low Revenue during the week 6 , 7 and 8
6. Home Audio : Zero Revenue reported during the weeks 6 , 7 and 8

Model	Significant Variables	Final Adjusted R-squared	10 fold cross validation(Mean Square Error)	Elasticity - Analysis	# of Model iteration s used	Final Conclusion
Basic Linear Model	<ol style="list-style-type: none"> Online marketing ad stock IsHolidayWeek2 SLA 	0.479	0.572		13	Positive impact on gmv if we spend 2 weeks before the holiday and invest in Online Marketing ; SLA will negatively impact if we miss any and organization should coordinate this with the delivery partners
Multiplicative Model	<ol style="list-style-type: none"> Discount -1.9199 Negative SLA 1.1011 Positive deliverycdays -0.0554 Negative Online.marketing_adstock 0.1022 Positive 	0.756	1.84		17	Positive impact on gmv if we spend in online marketing , SLA ; However there is a negative impact if we spend more on discounts and impact with deliverycdays
Koyck model	<ol style="list-style-type: none"> SLA -2.28e-01 Negative IsHolidayWeek -1.08e+15 Negative deliverybdays -9.85e-02 Negative Online.marketing_adstock 5.67e-01 Positive 	0.501	0.567		17	Positive impact on gmv if we spend in online marketing and negative impact if we miss SLA and spend on holiday week and deliverybdays and organization should coordinate this with the delivery partners
Distributed Lag Model	<ol style="list-style-type: none"> SLA 1.02e+00 Positive procurement_sla -6.71e-01 Negative IsHolidayWeek -7.91e+14 Negative deliverybdays 5.58e-02 Positive Online.marketing_adstock 1.20e-01 Positive 	0.506	0.588		25	Positive impact on gmv if we spend in online marketing and maintain the SLA and deliverybdays ; Negative impact if we spend on Holiday week , not maintaining the procurement SLA with vendors
Multiplicative and distributed model	<ol style="list-style-type: none"> Selling Price 1.1138 Positive SLA 0.5154 Positive Online.marketing_adstock 0.1417 Positive IsHolidayWeek1 0.0417 Positive 	0.825	0.839		21	Positive impact on gmv if we spend in Online Marketing and Selling Price should maintain with coordination of with SLA and maintaining the spending one week before the holiday



Camera Accessory Data



UpGrad

© Can Stock Photo

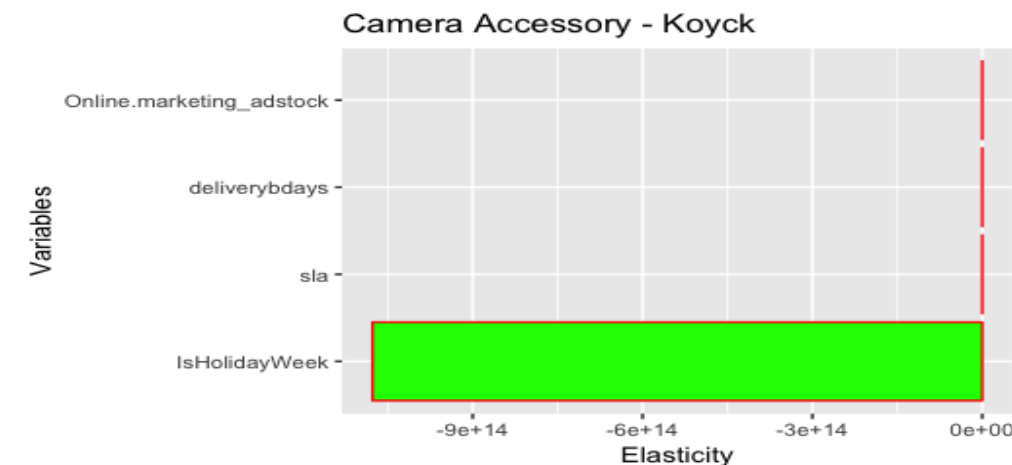
Koyck model	<div><div>1. SLA</div><div>-2.28e-01 Negative</div></div> <div><div>2. IsHolidayWeek</div><div>-1.08e+15 Negative</div></div> <div><div>3. deliverybdays</div><div>-9.85e-02 Negative</div></div> <div><div>4. Online.marketing_adstock</div><div>5.67e-01 Positive</div></div>	0.501	0.567		17	Positive impact on gmv if we spend in online marketing and negative impact if we miss SLA and spend on holiday week and deliverybdays and organization should coordinate this with the delivery partners
-------------	---	-------	-------	--	----	--

Though Multiplicative and Distributed Model have positive R-Square value we have observed a high 10 fold cross validation(Mean Square Error) ; Next best Model has selected as Final Model : Koyck Model with a low error rate and good R-Square value.

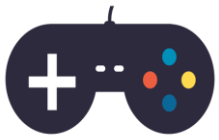
Model recommends to invest more in Online Marketing for a positive influence on Revenue / GMV ;

While we spend on Online Marketing, Model suggesting to maintain the delivery SLA and Deliverybdays with the delivery partner to maintain the customer satisfaction which in-directly influence the revenue / gmv

Model suggests not to spend during the Holiday week as it may influence negatively



Model	Significant Variables	Final Adjusted R-squared	10 fold cross validation(Mean Square Error)	Elasticity - Analysis	# of Model iterations used	Final Conclusion
Basic Linear Model	<ol style="list-style-type: none"> procurement_sla -1.26e+00 Negative deliverybdays 1.82e-01 Positive Online.marketing_adstock 5.58e-01 Positive HolidayWeek2 9.48e+14 Positive 	0.546	0.486		10	Positive impact on gmv if we spend 2 weeks before the holiday and invest in Online Marketing ; procurement SLA and deliverybdays will negatively impact if we miss any and organization should coordinate this with the delivery partners
Multiplicative Model	<ol style="list-style-type: none"> procurement_sla 0.2064 Positive Online.marketing_adstock 0.1136 Positive SP_Lag_3 0.0252 Positive 	0.773	0.992		9	Positive impact on gmv if we spend in online marketing , Selling Price of Lag 3 and organization should work with Vendor on the procurement SLA as it is creating a positive impact
Koyck model	<ol style="list-style-type: none"> procurement_sla -1.25e+00 Negative IsHolidayWeek -1.46e+15 Negative deliverybdays 2.59e-01 Positive NPS_Score -8.19e-01 Negative 	0.535	0.542		14	Negative impact on gmv if organization not maintains the procurement SLA and NPS score from the customers ; also have a negative impact if they spend on week of Holiday ; Positive impact on deliverybdays
Distributed Lag Model	<ol style="list-style-type: none"> procurement_sla -2.02 Negative deliverybdays -0.12 Negative Affiliates_adstock 0.72 Positive 	0.522	0.578		27	Positive impact on gmv if we spend in Affiliates and ; Negative impact if vendor not maintaining the procurement SLA and deliverybdays.
Multiplicative and distributed model	<ol style="list-style-type: none"> Selling Price 0.8319 Positive procurement_sla 0.1698 Positive Affiliates_adstock 0.0524 Positive 	0.823	1.01		23	Positive impact on gmv if we spend in Affiliates and Selling Price should maintain along with Vendor for procurement SLA's



Gaming Accessory Data



UpGrad

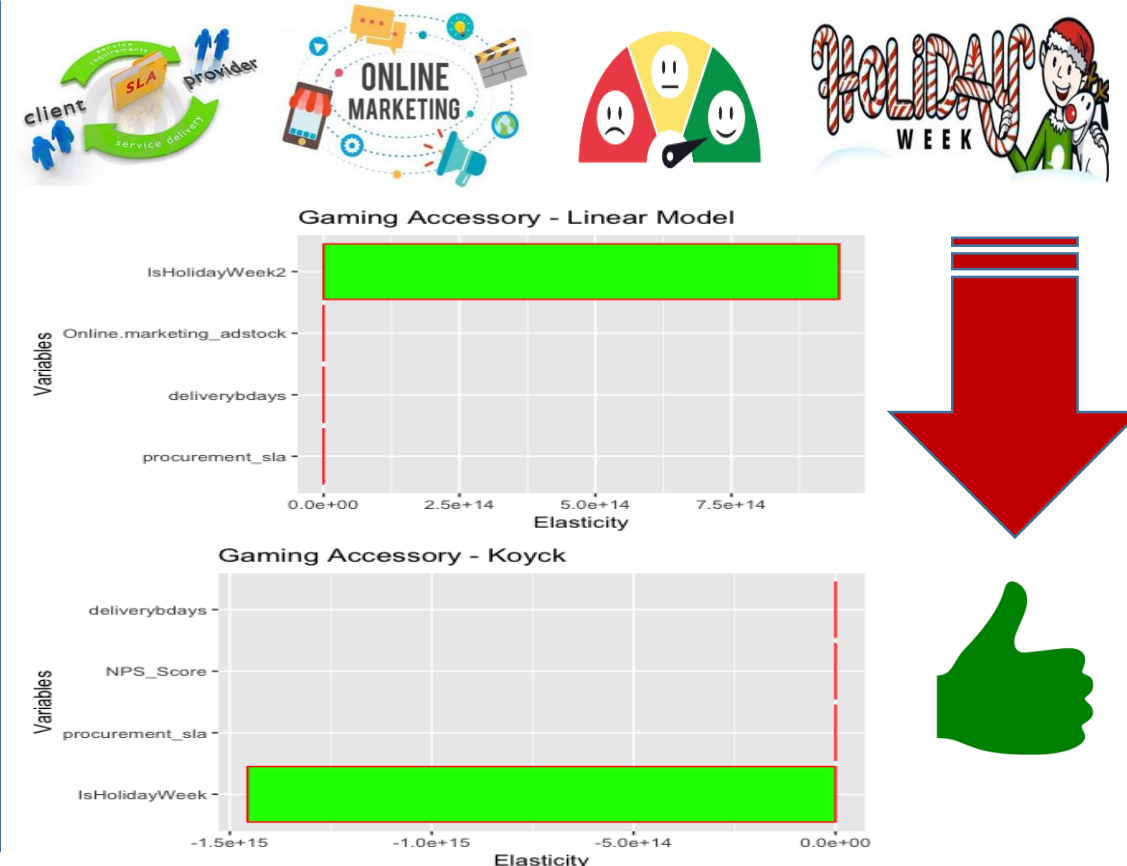
Basic Linear Model	<ol style="list-style-type: none">1. procurement_sla -1.26e+00 Negative2. deliverybdays 1.82e-01 Positive3. Online.marketing_adstock 5.58e-01 Positive4. HolidayWeek2 9.48e+14 Positive	0.546	0.486		10	Positive impact on gmv if we spend 2 weeks before the holiday and invest in Online Marketing ; procurement SLA and deliverybdays will negatively impact if we miss any and organization should coordinate this with the delivery partners
Koyck model	<ol style="list-style-type: none">1. procurement_sla -1.25e+00 Negative2. IsHolidayWeek -1.46e+15 Negative3. deliverybdays 2.59e-01 Positive4. NPS_Score -8.19e-01 Negative	0.535	0.542		14	Negative impact on gmv if organization not maintains the procurement SLA and NPS score from the customers ; also have a negative impact if they spend on week of Holiday ; Positive impact on deliverybdays

Final Model : Combination of Basic Linear and Koyck model

Since both models scored a good rating with respective to the R-Square and Error value

Model recommends to spend on Online Marketing , Investment two weeks before the holiday week shows a positive results , Deliverybdays will also impact positively on Revenue / GMV

However, Model recommends to maintain the NPS Score, Maintain the procurement SLA's as they may influence negatively on the Revenue / GMV



Model	Significant Variables	Final Adjusted R-squared	10 fold cross validation(Mean Square Error)	Elasticity - Analysis	# of Model iteration s used	Final Conclusion
Basic Linear Model	1. Discount 10.495 Positive 2. SEM_adstock 0.864 Positive	0.467	0.754	 <p>Home audio Accessory - Linear Model</p>	16	Positive impact on gmv on spending at SME and Discounts
Multiplicative Model	1. Discount 0.68426 Positive 2. SLA 0.25630 Positive 3. Other_adstock -0.00784 Negative 4. NPS_Score -1.16877 Negative	0.51	0.199	 <p>Home Audio - Multiplicative Model</p>	13	Positive impact on gmv if we spend in Discount sale and organization should work with Vendor on the procurement SLA as it is creating a positive impact ; Negative impact if we spend on other add stocks and NPS should keep high
Koyck model	1. Discount 11.6 Positive 2. Digital_adstock -0.7 Negative	0.489	0.744	 <p>Home Audio - Koyck Model</p>	17	Positive impact on gmv if we spend in Discount sale and organization should stop investing in Digital add stock.
Distributed Lag Model	1. SLA -0.0906 Negative 2. Digital_adstock -0.1533 Negative 3. Discount_Lag_3 -0.0945 Negative 4. 'Discount-3' -2.4238 Negative	0.502	0.74	 <p>Home Audio -Distributive Lag Model</p>	25	Negative impact on gmv if we spend in Digital add stock and Discount with Lag of 3 weeks ; Organization should work with vendor to maintain the SLA
Multiplicative and distributed model	1. Discount 0.6563 Positive 2. SLA 0.2232 Positive 3. Affiliates_adstock 0.0219 Positive	0.4932	0.228	 <p>Home Audio- Distributed + Multiplicative</p>	30	Positive impact on gmv if we spend in Affiliates and Discount should maintain along with Vendor for procurement SLA's



Home Audio Data



UpGrad

Multiplicative Model	<div><div>1. Discount0.68426 Positive</div><div>2. SLA0.25630 Positive</div><div>3. Other_adstock-0.00784 Negative</div><div>4. NPS_Score-1.16877 Negative</div></div>	0.51	0.199	<p>Home Audio - Multiplicative Model</p> <p>Variables: Discount, sla, Other_adstock, NPS_Score</p> <p>Elasticity: -1.2, -0.8, -0.4, 0.0, 0.4</p>	13	Positive impact on gmv if we spend in Discount sale and organization should work with Vendor on the procurement SLA as it is creating a positive impact ; Negative impact if we spend on other add stocks and NPS should keep high
Multiplicative and distributed model	<div><div>1. Discount0.6563 Positive</div><div>2. SLA0.2232 Positive</div><div>3. Affiliates_adstock0.0219 Positive</div></div>	0.4932	0.228	<p>Home Audio- Distributed + Multiplicative</p> <p>Variables: Discount, sla, Affiliates_adstock</p> <p>Elasticity: 0.0, 0.2, 0.4, 0.6</p>	30	Positive impact on gmv if we spend in Affiliates and Discount should maintain along with Vendor for procurement SLA's

Final Model : Combination of Multiplicative Model and Multiplicative and Distributed Model

Since both models scored a good rating with respective to the R-Square and Error value

Model recommends to spend on Affiliates Marketing , Discounts , and maintain the delivery SLA with the help of trusted delivery partners to maintain the positive growth in revenue /GMV.

However, Model recommends to maintain the NPS Score, Reduce the spending in Other type of marketing models as they may influence negatively on the Revenue / GMV

Home Audio - Multiplicative Model

Variables: Discount, sla, Other_adstock, NPS_Score

Elasticity: -1.2, -0.8, -0.4, 0.0, 0.4

Home Audio- Distributed + Multiplicative

Variables: Discount, sla, Affiliates_adstock

Elasticity: 0.0, 0.2, 0.4, 0.6

Thank You!

